Exponential Lower Bound for Static Semi-Algebraic Proofs

Dima Grigoriev[1], Edward A. Hirsch[2⋆], and Dmitrii V. Pasechnik[3]

[1] IRMAR, Université de Rennes, Campus de Beaulieu, 35042 Rennes, cedex France.
Email: dima@maths.univ-rennes1.fr.
Web: http://www.maths.univ-rennes1.fr/~dima/
[2] Steklov Institute of Mathematics at St.Petersburg, 27 Fontanka, 191011 St.Petersburg, Russia.
Email: hirsch@pdmi.ras.ru. Web: http://logic.pdmi.ras.ru/~hirsch/
[3] Department of Technical Mathematics and Informatics, Faculty ITS, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.
Email: d.pasechnik@its.tudelft.nl.
Web: http://ssor.twi.tudelft.nl/~dima/

Abstract. Semi-algebraic proof systems were introduced in [1] as extensions of Lovász-Schrijver proof systems [2, 3]. These systems are very strong; in particular, they have short proofs of Tseitin's tautologies, the pigeonhole principle, the symmetric knapsack problem and the clique-coloring tautologies [1].
In this paper we study *static* versions of these systems. We prove an exponential lower bound on the length of proofs in one such system. The same bound for two tree-like (dynamic) systems follows. The proof is based on a lower bound on the "*Boolean degree*" of Positivstellensatz Calculus refutations of the symmetric knapsack problem.

1 Introduction

Algebraic proof systems. An observation that a propositional formula can be written as a system of polynomial equations has lead to considering *algebraic* proof systems, in particular, the Nullstellensatz (NS) and the Polynomial Calculus (PC) proof systems, see Subsection 2.2 below (we do not dwell much here on the history of this rich area, several nice historical overviews one could find in e.g., [4–9]).

For these proof systems several interesting complexity lower bounds on the degrees of the derived polynomials were obtained [6, 7, 9]. When the degree is close enough to linear (in fact, greater than the square root), these bounds imply exponential lower bounds on the proof complexity (more precisely, on the number of monomials in the derived polynomials) [7]. If polynomials are given by formulas rather than by sums of monomials as in NS or in PC, then the complexity could decrease significantly. Several gaps between these two kinds of proof systems are demonstrated in [10].

⋆ Partially supported by grant #1 of the 6th RAS contest-expertise of young scientists projects (1999) and grants from CRDF, RFBR and NATO.

Semi-algebraic proof systems. In [1], we have introduced several *semi-algebraic* proof systems. In these system, one deals with polynomial inequalities, and new inequalities can be derived by algebraic operations like the sum, the multiplication and the division. The simplest semi-algebraic systems are the so-called Lovász-Schrijver calculi (see [2, 3], cf. also [11] and Subsection 2.3 below), where the polynomials are restricted to quadratic ones. No exponential lower bounds are known so far even for these restricted systems (a number of lower bounds on the number of steps of Lovász-Schrijver *procedure* is known [12–15, 1], but they do not imply exponential lower bounds on the size of proofs [1]). Moreover, general semi-algebraic proof systems, (where one allows polynomials of arbitary degree, see [1] and Subsection 2.3 below), appear to be very strong. In [1], it is proved that such systems have short proofs of Tseitin's tautologies, the pigeonhole principle, clique-coloring tautologies and the symmetric knapsack problem. They also polynomially simulate the Cutting Planes proof system [16–19] with polynomially bounded coefficients. Another (and much stronger) kind of semi-algebraic proof system was introduced in [20] with no focus on the complexity.

Static systems and our results. Another proof system manipulating polynomial inequalities called the Positivstellensatz Calculus was introduced in [21]. Lower bounds on the degree in this system were established for the parity principle, for Tseitin's tautologies [22] and for the knapsack problem [23]. Lower bounds on the Positivstellensatz Calculus degree are possible because its "dynamic" part is restricted to an ideal and an element of a cone is obtained from an element of ideal by adding the sum of squares to it. On the contrary, the semi-algebraic proof systems introduced in [2, 3, 1] are completely "dynamic" proof systems. (The discussion on static and dynamic proof systems can be found in [21]. Briefly, the difference is that in the dynamic semi-algebraic proof systems a derivation constructs gradually an element of the cone generated by the input system of inequalities, while in the Positivstellensatz Calculus the sum of squares is given explicitly.) We consider a static version of Lovász-Schrijver calculi and prove an exponential lower bound on the size of refutation of the symmetric knapsack problem (Section 4); this bound also translates into the bound for the tree-like version of (dynamic) LS. The key ingredient of the proof is a linear lower bound on the "Boolean degree" of Positivstellensatz Calculus refutations (Section 3). Note that exponential lower bounds on the size of (static!) Positivstellensatz refutations are still unknown.

Organization of the paper. We start with the definitions of proof systems in general and the particular proof systems we use in our paper (Section 2). We then prove a lower bound on the "Boolean degree" of Positivstellensatz Calculus refutations of the symmetric knapsack problem (Section 3), and derive from it an exponential lower bound on the size of proofs in a static semi-algebraic proof system and in the tree-like versions of two dynamic semi-algebraic proof systems (Section 4). Finally, we formulate open questions (Section 5).

2 Definitions

2.1 Proof systems

A *proof system* [24] for a language L is a polynomial-time computable function mapping words (proof candidates) onto L (whose elements are considered as theorems).

A *propositional proof system* is a proof system for any fixed co-NP-complete language of Boolean tautologies (e.g., tautologies in DNF).

When we have two proof systems Π_1 and Π_2 for the same language L, we can compare them. We say that Π_1 *polynomially simulates* Π_2, if there is a function g mapping proof candidates of Π_2 to proof candidates of Π_1 so that for every proof candidate π for Π_2, one has $\Pi_1(g(\pi)) = \Pi_2(\pi)$ and $g(\pi)$ is at most polynomially longer than π.

Proof system Π_1 is *exponentially separated* from Π_2, if there is an infinite sequence of words $t_1, t_2, \ldots \in L$ such that the length of the shortest Π_1-proof of t_i is polynomial in the length of t_i, and the length of the shortest Π_2-proof of t_i is exponential.

Proof system Π_1 is *exponentially stronger* than Π_2, if Π_1 polynomially simulates Π_2 and is exponentially separated from it.

When we have two proof systems for different languages L_1 and L_2, we can also compare them if we fix a reduction between these languages. However, it can be the case that the result of the comparison is more due to the reduction than to the systems themselves. Therefore, if we have propositional proof systems for languages L_1 and L_2, and the intersection $L = L_1 \cap L_2$ of these languages is co-NP-complete, we will compare these systems as systems[1] for L.

2.2 Algebraic proof systems

There is a series of proof systems for languages consisting of unsolvable systems of polynomial equations. To transform such a proof system into a propositional proof system, one needs to translate Boolean tautologies into systems of polynomial equations.

To translate a formula F in k-DNF, we take its negation $\neg F$ in k-CNF and translate each clause of $\neg F$ into a polynomial equation. A clause containing variables $v_{j_1}, \ldots, v_{j_t}$ $(t \leq k)$ is translated into an equation

$$(1 - l_1) \cdot \ldots \cdot (1 - l_t) = 0, \tag{1}$$

where $l_i = v_{j_i}$ if variable v_{j_i} occurs positively in the clause, and $l_i = (1 - v_{j_i})$ if it occurs negatively. For each variable v_i, we also add the equation $v_i^2 - v_i = 0$ to this system.

[1] If one can decide in polynomial time for $x \in L_1$, whether $x \in L$, then any proof system for L_1 can be restricted to $L \subseteq L_1$ by mapping proofs of elements of $L_1 \setminus L$ into any fixed element of L. For example, this is the case for L_1 consisting of all tautologies in DNF and L consisting of all tautologies in k-DNF.

Remark 1. Observe that it does not make sense to consider this translation for formulas in general DNF (rather than k-DNF for constant k), because an exponential lower bound for any system using such encoding would be trivial (note that $(1-v_1)(1-v_2)\dots(1-v_n)$ denotes a polynomial with exponentially many monomials).

Note that F is a tautology if and only if the obtained system S of polynomial equations $f_1 = 0$, $f_2 = 0$, ..., $f_m = 0$ has no solutions. Therefore, to prove F it suffices to derive a contradiction from S.

Nullstellensatz (NS) [4]. A proof in this system is a collection of polynomials $g_1, \dots, g_m$ such that

$$\sum_i f_i g_i = 1.$$

Polynomial Calculus (PC) [8]. This system has two derivation rules:

$$\frac{p_1 = 0;\ p_2 = 0}{p_1 + p_2 = 0} \qquad \text{and} \qquad \frac{p = 0}{p \cdot q = 0}. \tag{2}$$

I.e., one can take a sum[2] of two already derived equations $p_1 = 0$ and $p_2 = 0$, or multiply an already derived equation $p = 0$ by an arbitrary polynomial q. The proof in this system is a derivation of $1 = 0$ from S using these rules.

Positivstellensatz [21]. A proof in this system consists of polynomials $g_1, \dots, g_m$ and $h_1, \dots, h_l$ such that

$$\sum_i f_i g_i = 1 + \sum_j h_j^2 \tag{3}$$

Positivstellensatz Calculus [21]. A proof in this system consists of polynomials $h_1, \dots, h_l$ and a derivation of $1 + \sum_j h_j^2 = 0$ from S using the rules (2).

2.3 Dynamic semi-algebraic proof systems

To define a propositional proof system manipulating with inequalities, we again translate each formula $\neg F$ in CNF into a system S of linear inequalities, such that F is a tautology if and only if S has no 0-1 solutions. Given a Boolean formula in CNF, we translate each its clause containing variables $v_{j_1}, \dots, v_{j_t}$ into the inequality

$$l_1 + \dots + l_t \geq 1, \tag{4}$$

where $l_i = v_{j_i}$ if the variable v_{j_i} occurs positively in the clause, and $l_i = 1 - v_{j_i}$ if v_{j_i} occurs negatively. We also add to S the inequalities

$$x \geq 0, \tag{5}$$

$$x \leq 1 \tag{6}$$

for every variable x.

[2] Usually, an arbitrary linear combination is allowed, but clearly it can be replaced by two multiplications and one addition.

Lovász-Schrijver calculus (LS) [2, 3] (cf. also [11]). In the weakest of Lovász-Schrijver proof systems, the contradiction must be obtained using the rule

$$\frac{f_1 \geq 0;\ \ldots;\ f_t \geq 0}{\sum_{i=1}^{t} \lambda_i f_i \geq 0} \quad (\text{where } \lambda_i \geq 0), \tag{7}$$

applied to linear or quadratic f_i's and the rules

$$\frac{f \geq 0}{fx \geq 0}; \qquad \frac{f \geq 0}{f(1-x) \geq 0} \qquad (\text{where } f \text{ is linear, } x \text{ is a variable}). \tag{8}$$

Also, the system S is extended by the axioms

$$x^2 - x \geq 0, \quad x - x^2 \geq 0 \tag{9}$$

for every variable x.

LS$_+$ [2, 3, 11]. This system has the same axioms and derivation rules as LS, and also has the axiom

$$l^2 \geq 0 \tag{10}$$

for every linear l.

Note that the Lovász-Schrijver systems described above deal either with linear or quadratic inequalities. In [1], several extensions of Lovász and Schrijver proof systems are introduced. The main idea is to allow a proof to contain monomials of degree up to d.

LS^d. This system is an extension of LS. The difference is that rule (8) is now restricted to f of degree at most $d-1$ rather than to linear inequalities. Rule (7) can be applied to any collection of inequalities of degree at most d.

Remark 2. Note that LS=LS^2.

2.4 Static semi-algebraic proof systems

Nullstellensatz is a "static" version of Polynomial Calculus; Positivstellensatz is a "static" version of Positivstellensatz Calculus. Similarly, we define "static" versions of the semi-algebraic proof systems defined in the previous subsection.

Static LS^n. A proof in this system is a a refutation of a system of inequalities $S = \{s_i \geq 0\}_{i=1}^{t}$, where each $s_i \geq 0$ is either an inequality given by the translation (4), an inequality of the form $x_j \geq 0$ or $1 - x_j \geq 0$, or an inequality of the form $x_j^2 - x_j \geq 0$. The refutation consists of positive real coefficients $\omega_{i,l}$ and multisets $U_{i,l}^+$ and $U_{i,l}^-$ defining the polynomials

$$u_{i,l} = \omega_{i,l} \cdot \prod_{k \in U_{i,l}^+} x_k \cdot \prod_{k \in U_{i,l}^-} (1 - x_k)$$

such that

$$\sum_{i=1}^{t} s_i \sum_{l} u_{i,l} = -1. \tag{11}$$

Static $\mathbf{LS}^n_+$. The difference from the previous system is that S is extended by inequalities $s_{t+1} \geq 0, \ldots, s_{t'} \geq 0$, where each polynomial s_j ($j \in [t+1..t']$) is a square of another polynomial s'_j. The requirement (11) transforms into

$$\sum_{i=1}^{t'} s_i \sum_l u_{i,l} = -1. \tag{12}$$

Static $\mathbf{LS}_+$. The same as static LS^n_+, but the polynomials s'_i can be only linear.

Remark 3. Note that static LS_+ includes static LS^n.

Remark 4. Note that these static systems are not propositional proof systems in the sense of Cook and Reckhow [24], but are something more general, since there is no clear way to verify (11) in deterministic polynomial time (cf. [25]). However, they can be easily augmented to match the definition of Cook and Reckhow, e.g., by including a proof of the equality (11) or (12) using axioms of a ring (cf. F-NS of [10]). Clearly, if we prove a lower bound for the original system, the lower bound will be valid for any augmented system as well.

Remark 5. The size of a refutation in these systems is the length of a reasonable bit representation of all polynomials $u_{i,l}$, s_i (for $i \in [1..t]$) and s'_j (for $j \in [t+1..t']$) and is thus at least the number of $u_{i,l}$'s.

Example 1. We now present a very simple static LS_+ proof of the propositional pigeonhole principle. The negation of this tautology is given by the following system of inequalities:

$$\sum_{\ell=1}^{m-1} x_{k\ell} \geq 1; \qquad 1 \leq k \leq m; \tag{13}$$

$$x_{k\ell} + x_{k'\ell} \leq 1; \qquad 1 \leq k < k' \leq m;\ 1 \leq \ell \leq m-1. \tag{14}$$

(That says that the k-th pigeon must get into a hole, while two pigeons k and k' cannot share the same hole ℓ.)

Here is the static LS_+ proof:

$$\begin{aligned}
&\sum_{k=1}^{m} \left(\sum_{\ell=1}^{m-1} x_{k\ell} - 1 \right) + \\
&\sum_{\ell=1}^{m-1} \left(\sum_{k=1}^{m} x_{k\ell} - 1 \right)^2 + \\
&\sum_{\ell=1}^{m-1} \sum_{k=1}^{m} \sum_{k \neq k'=1}^{m} (1 - x_{k\ell} - x_{k'\ell}) x_{k\ell} + \\
&\sum_{\ell=1}^{m-1} \sum_{k=1}^{m} (x_{k\ell}^2 - x_{k\ell})(m-1) \\
&= -1.
\end{aligned}$$

□

3 Linear lower bound on the "Boolean degree" of Positivstellensatz Calculus refutations of the knapsack

We use the following notation from [7, 23]. For a polynomial f, its *multilinearization* $\overline{f}$ is a polynomial obtained by the reduction of f modulo $(x - x^2)$ for every variable x, i.e., f is the unique multilinear polynomial equivalent to f modulo these ("Boolean") polynomials. When $f = \overline{f}$ we say that f is reduced.

For a monomial t one can define its *Boolean degree* $\mathrm{Bdeg}(t)$ as $\deg(\overline{t})$, in other words, the number of occurring variables; then one extends the concept of Bdeg to polynomials: $\mathrm{Bdeg}(f) = \max \mathrm{Bdeg}(t_i)$, where the maximum is taken over all non-zero monomials t_i occurring in f. Thereby, one can define Bdeg of a derivation in PC and subsequently in Positivstellensatz and Positivstellensatz Calculus as maximum Bdeg of *all* polynomials in the derivation (in Positivstellensatz and Positivstellensatz Calculus, this includes polynomials h_j^2, cf. definition in Subsection 2.2).

The following lemma extends the argument in the proof of [7, Theorem 5.1] from deg to Bdeg.

Lemma 1. *Let* $f(x_1, \ldots, x_n) = c_1x_1 + \ldots + c_nx_n - m$, *where* $c_1, \ldots, c_n \in \mathbb{R}\setminus\{0\}$. *Let* q *be deducible in PC from the knapsack problem* $f = 0$ *with* $\mathrm{Bdeg} \leq \lceil (n-1)/2 \rceil$. *Then one can represent*

$$q = \sum_{i=1}^{n} (x_i - x_i^2) g_i + fg, \tag{15}$$

where $\deg(fg) \leq \mathrm{Bdeg}(q)$.

Proof. Similarly to the proof of [7, Theorem 5.1], we conduct the induction along a (fixed) deduction in PC. Assume (15) and consider a polynomial qx_1 obtained from q by multiplying it by a variable x_1. W.l.o.g. one can suppose that g is reduced. Then $\overline{qx_1} = \overline{fgx_1}$; denote $h = \overline{gx_1}$. Let $d = \deg(h) - 1$. We need to verify that $d + 2 = \deg(fh) \leq \mathrm{Bdeg}(qx_1)$. Taking into account that

$$d + 1 = \deg(h) \leq \deg(g) + 1 = \deg(fg) \leq \mathrm{Bdeg}(q) \leq \mathrm{Bdeg}(qx_1),$$

the mere case to be brought to a contradiction is when $\mathrm{Bdeg}(qx_1) = \mathrm{Bdeg}(q) = \deg(g) + 1 = d + 1$.

We write $g = p + x_1p_1$ where all the terms of g not containing x_1 are gathered in p. Clearly, $\deg(p) \leq \deg(g) = d$. Moreover, $\deg(p) = d$ because if $\deg(p) < d$, we would have $d+1 = \deg(h) \leq \mathrm{Bdeg}(gx_1) \leq \max(\mathrm{Bdeg}(x_1p), \mathrm{Bdeg}(x_1^2p_1)) \leq d$.

On the other hand, $d = \mathrm{Bdeg}(q) - 1 \leq \lceil (n-1)/2 \rceil - 1$. Therefore, [7, Lemma 5.2] applied to the instance $c_2x_2 + \ldots + c_nx_n - 0$ of symmetric knapsack states that

$$\deg(\overline{(c_2x_2 + \ldots + c_nx_n)p}) = \deg(p) + 1 = d + 1$$

(one should add to the formulation of [7, Lemma 5.2] the condition that p is reduced).

Hence there exists a monomial $x^J = \prod_{j\in J} x_j$ occurring in p for a certain $J \subseteq \{2,\ldots,n\}$, $|J| = d$, and besides, there exists $i \in [2..n]$ such that the monomial $x_i x^J$, being of the degree $d+1$, occurs in the polynomial $\overline{(c_2x_2 + \ldots + c_nx_n)p}$, in particular $i \notin J$.

Because of that the monomial $T = x_i x^J x_1$ with $\deg(T) = d+2$ occurs in

$$p' = \overline{(c_2x_2 + \ldots + c_nx_n)px_1}.$$

Furthermore, T occurs in

$$\overline{fgx_1} = \overline{((c_2x_2 + \ldots + c_nx_n) + (c_1x_1 - m))(p + x_1p_1)x_1}$$

since after opening the parenthesis in the right-hand side of the latter expression we obtain only p' and two subexpressions

$$\overline{(c_1x_1 - m)(p + x_1p_1)x_1} = \overline{(c_1 - m)gx_1} \quad \text{and} \quad \overline{(c_2x_2 + \ldots + c_nx_n)x_1p_1x_1}$$

of Boolean degree at most $d+1$ (thereby, any monomial from these subexpressions cannot be equal to the *reduced* monomial T). Finally, due to the equality $\overline{qx_1} = \overline{fgx_1}$, we conclude that $\mathrm{Bdeg}(qx_1) \geq \deg(\overline{qx_1}) = \deg(\overline{fgx_1}) \geq d+2$; the achieved contradiction proves the induction hypothesis for the case of the rule of the multiplication by a variable (note that the second rule in (2) can be replaced by the multiplication by a variable with a multiplicative constant).

Now we proceed to the consideration of the rule of taking the sum of two polynomials q and r. By the induction hypothesis we have

$$r = \sum_{i=1}^{n}(x_i - x_i^2)u_i + fu,$$

where u is reduced and $\deg(fu) \leq \mathrm{Bdeg}(r)$. Then making use of (15) we get $\overline{r+q} = \overline{fv}$ where $v = \overline{g+u}$. The inequality

$$\begin{aligned}\deg(v) &\leq \max\{\deg(g),\deg(u)\} \leq \max\{\mathrm{Bdeg}(q),\mathrm{Bdeg}(r)\} - 1\\ &\leq \lceil (n-1)/2\rceil - 1 \leq \lceil n/2\rceil - 1\end{aligned}$$

enables us to apply [7, Lemma 5.2] to v, this implies that $\deg(\overline{fv}) = \deg(v)+1 = \deg(fv)$. Therefore, $\mathrm{Bdeg}(r+q) \geq \deg(\overline{r+q}) = \deg(\overline{fv}) = \deg(fv)$. □

The next corollary extends [7, Theorem 5.1].

Corollary 1. *Any PC refutation of the knapsack f has* Bdeg *greater than* $\lceil (n-1)/2\rceil$.

Now we can formulate the following theorem extending the theorem of [23] from deg to Bdeg. Denote by δ a stairs-form function which equals to 2 out of the interval $(0,n)$ and which equals to $2k+4$ on the intervals $(k,k+1)$ and $(n-k-1,n-k)$ for all integers $0 \leq k < n/2$.

Theorem 1. *Any Positivstellensatz Calculus refutation of the symmetric knapsack problem* $f = x_1 + \ldots + x_n - m$ *has* Bdeg *greater or equal to* $\min\{\delta(m), \lceil (n-1)/2 \rceil + 1\}$.

Proof. The proof of the theorem follows the proof of the theorem [23]. First, we apply Lemma 1 to the deduction in PC being an ingredient of the deduction in Positivstellensatz Calculus (see definitions in 2.2). This provides a refutation in Positivstellensatz Calculus of the form

$$-1 = \sum_{i=1}^{n} (x_i - x_i^2) g_i + fg + \sum_j h_j^2. \tag{16}$$

The rest of the proof follows literally the proof from [23] which consists in applying to (16) the homomorphism B introduced in [23]. It is worthwhile to mention that B is defined on the quotient algebra $\mathbb{R}[x_1, \ldots, x_n]/(x_1 - x_1^2, \ldots, x_n - x_n^2)$, thereby, the proof in [23] actually, estimates Bdeg rather than just deg. □

4 Exponential lower bound on the size of static LS_+ refutations of the symmetric knapsack

In this section we apply the results of Section 3 to obtain an exponential lower bound on the size of static LS_+ refutations of the symmetric knapsack. We follow the notation introduced in Subsection 2.4 and Section 3. The *Boolean degree of a static LS (LS_+) refutation* is the maximum Boolean degree of the polynomials $u_{i,l}$ in Subsection 2.4.

Let us fix for the time being a certain (threshold) d.

Lemma 2. *Denote by* M *the number of* $u_{i,l}$*'s occurring in (12) that have Boolean degrees at least* d*. Then there is a variable* x *and a value* $a \in \{0, 1\}$ *such that the result of substituting* $x = a$ *in (12) contains at most* $M(1 - d/(2n))$ *non-zero polynomials* $u_{i,l}|_{x=a}$ *of Boolean degrees at least* d*. (Note that by substituting in (12) a value* a *for* x *we obtain a valid static* LS_+ *refutation of the system* $S|_{x=a}$*).*

Proof. Since there are at least Md polynomials $u_{i,l}$ of Boolean degrees at least d containing either x or $1 - x$, there is a variable x such that either x or $1 - x$ occurs in at least $Md/(2n)$ of these polynomials. Therefore, after substituting the appropriate value for x, at least $Md/(2n)$ polynomials $u_{i,l}$ vanish from (12). □

For the symmetric knapsack problem

$$x_1 + x_2 + \ldots + x_n - m = 0 \tag{17}$$

we can rewrite its static LS_+ refutation in the following way. Denote

$$\begin{aligned} f_0 &= x_1 + \ldots + x_n - m, \\ f_i &= x_i - x_i^2 \quad (1 \le i \le n), \\ f_i &= (s_i')^2 \quad (n+1 \le i \le n') \end{aligned}$$

(m is not an integer). The refutation can be represented in the form

$$\sum_{i=0}^{t} f_i \sum_l g_{i,l} + \sum_{j=n+1}^{n'} f_j t_j + \sum_{j=n'+1}^{n''} t_j = -1, \tag{18}$$

where

$$g_{i,l} = \gamma_{i,l} \cdot \prod_{k \in G^+_{i,l}} x_k \cdot \prod_{k \in G^-_{i,l}} (1 - x_k),$$

$$t_j = \tau_j \cdot \prod_{k \in T^+_j} x_k \cdot \prod_{k \in T^-_j} (1 - x_k)$$

for appropriate multisets $G^-_{i,l}$, $G^+_{i,l}$, T^-_j and T^+_j, positive real τ_j and *arbitrary* real $\gamma_{i,l}$.

Lemma 3. *If $n/4 < m < 3n/4$, then the Boolean degree D of any static LS_+ refutation of the symmetric knapsack problem is at least $n/4$.*

Proof. Replacing in t_j each occurrence of x_i by $f_i + x_i^2$ and each occurrence of $1 - x_i$ by $f_i + (1 - x_i)^2$ and subsequently opening the parentheses in t_j, one can gather all the terms containing at least one of f_i and separately the products of squares of the form x_i^2, $(1 - x_i)^2$. As a result one gets a representation of the form

$$\sum_{i=0}^{n} f_i g_i + \sum_{j=1}^{n'''} h_j^2 = -1$$

for appropriate polynomials g_i, h_j of Boolean degrees $\mathrm{Bdeg}(g_i), \mathrm{Bdeg}(h_j^2) \le D$, thereby a Positivstellensatz (and Positivstellensatz Calculus) refutation of the symmetric knapsack of Boolean degree at most $D + 2$. Then Theorem 1 implies that $D \ge \lceil (n-1)/2 \rceil - 1 \ge n/4$. □

Theorem 2. *For $m = (2n + 1)/4$ the number of $g_{i,l}$'s and t_j's in (18) is* $\exp(\Omega(n))$.

Proof. Now we set $d = \lceil n/8 \rceil$ and apply Lemma 2 consecutively $\kappa = \lfloor n/4 \rfloor$ times. The result of all these substitutions in (18) we denote by (18′), it contains $n - \kappa$ variables; denote by $u'_{i,l}$ the polynomial we thus get from $u_{i,l}$. We denote by f'_0 the result of substitutions applied to f_0. Note that after all substitutions we obtain again an instance of the knapsack problem. Taking into account that the free term m' of f'_0 ranges in the interval $[m - \kappa, m]$ and since $(n - \kappa)/4 < m - \kappa < m < 3(n - \kappa)/4$, we are able to apply Lemma 3 to (18′). Thus, the degree of (18′) is at least $(n - \kappa)/4 > d$.

Denote by M_0 the number of $u_{i,l}$'s of the degrees at least d in (18). By Lemma 2 the refutation (18′) contains at most $M_0(1 - d/(2n))^\kappa \le M_0(1 - 1/16)^{n/4}$ non-zero polynomials $u'_{i,l}$ of degrees at least d. Since there is at least one polynomial $u'_{i,l}$ of such degree, we have $M_0(1 - 1/16)^{n/4} \ge 1$, i.e. $M_0 \ge (16/15)^{n/4}$, which proves the theorem. □

Corollary 2. *Any static LS_+ refutation of (17) for $m = (2n+1)/4$ must have size* $\exp(\Omega(n))$.

Corollary 3. *Any treelike LS_+ (or LS^n) refutation of (17) for $m = (2n+1)/4$ must have size* $\exp(\Omega(n))$.

Proof. The size of such treelike refutation (even the numer of instances of axioms f_i used in the refutation) is at least the number of polynomials $u_{i,l}$. □

Remark 6. The value $m = (2n+1)/4$ in Theorem 2 and its corollaries above can be changed to any non-integer value between $\lceil n/4 \rceil$ and $\lfloor 3n/4 \rfloor$ by tuning the constants in the proofs (and in the $\Omega(n)$ in the exponent).

5 Open questions

1. Prove an exponential lower bound for a static semi-algebraic *propositional* proof system. Note that we have only proved an exponential lower bound for static LS_+ as a proof system for the co-NP-complete language of *systems of 0-1 linear inequalities*, because the symmetric knapsack problem is not obtained as a translation of a Boolean formula in DNF.
2. Prove an exponential lower bound for a dynamic semi-algebraic proof system, e.g., for LS.
3. Can static LS be polynomially simulated by a certain version of the Cutting Planes proof system?

References

1. Grigoriev, D., Hirsch, E.A., Pasechnik, D.V.: Complexity of semi-algebraic proofs. In: Proceedings of the 19th International Symposium on Theoretical Aspects of Computer Science, STACS 2002. Volume 2285 of Lecture Notes in Computer Science., Springer (2002) 419–430
2. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0–1 optimization. SIAM Journal on Optimization **1** (1991) 166–190
3. Lovász, L.: Stable sets and polynomials. Discrete Mathematics **124** (1994) 137–153
4. Beame, P., Impagliazzo, R., Krajíček, J., Pitassi, T., Pudlák, P.: Lower bounds on Hilbert's Nullstellensatz and propositional proofs. Proc. London Math. Soc. **73** (1996) 1–26
5. Beame, P., Impagliazzo, R., Krajíček, J., Pudlák, P., Razborov, A.A., Sgall, J.: Proof complexity in algebraic systems and bounded depth Frege systems with modular counting. Computational Complexity **6** (1996/97) 256–298
6. Razborov, A.A.: Lower bounds for the polynomial calculus. Computational Complexity **7** (1998) 291–324
7. Impagliazzo, R., Pudlák, P., Sgall, J.: Lower bounds for the polynomial calculus. Computational Complexity **8** (1999) 127–144
8. Clegg, M., Edmonds, J., Impagliazzo, R.: Using the Groebner basis algorithm to find proofs of unsatisfiability. In: Proceedings of the 28th Annual ACM Symposium on Theory of Computing, STOC'96, ACM (1996) 174–183

9. Buss, S., Grigoriev, D., Impagliazzo, R., Pitassi, T.: Linear gaps between degrees for the polynomial calculus modulo distinct primes. Journal of Computer and System Sciences **62** (2001) 267–289
10. Grigoriev, D., Hirsch, E.A.: Algebraic proof systems over formulas. Technical Report 01-011, Electronic Colloquim on Computational Complexity (2001) `ftp://ftp.eccc.uni-trier.de/pub/eccc/reports/2001/TR01-011/index.html`.
11. Pudlák, P.: On the complexity of propositional calculus. In: Sets and Proofs: Invited papers from Logic Colloquium'97. Cambridge University Press (1999) 197–218
12. Stephen, T., Tunçel, L.: On a representation of the matching polytope via semidefinite liftings. Math. Oper. Res. **24** (1999) 1–7
13. Cook, W., Dash, S.: On the matrix-cut rank of polyhedra. Math. Oper. Res. **26** (2001) 19–30
14. Dash, S.: On the Matrix Cuts of Lovász and Schrijver and their use in Integer Programming. Technical report tr01-08, Rice University (2001) `http://www.caam.rice.edu/caam/trs/2001/TR01-08.ps`.
15. Goemans, M.X., Tunçel, L.: When does the positive semidefiniteness constraint help in lifting procedures. Mathematics of Operations Research (2001) to appear.
16. Gomory, R.E.: An algorithm for integer solutions of linear programs. In: Recent Advances in Mathematical Programming. McGraw-Hill (1963) 269–302
17. Chvátal, V.: Edmonds polytopes and a hierarchy of combinatorial problems. Discrete Math. **4** (1973) 305–337
18. Cook, W., Coullard, C.R., Turán, G.: On the complexity of cutting-plane proofs. Discrete Appl. Math. **18** (1987) 25–38
19. Chvátal, V., Cook, W., Hartmann, M.: On cutting-plane proofs in combinatorial optimization. Linear Algebra Appl. **114/115** (1989) 455–499
20. Lombardi, H., Mnev, N., Roy, M.F.: The Positivstellensatz and small deduction rules for systems of inequalities. Mathematische Nachrichten **181** (1996) 245–259
21. Grigoriev, D., Vorobjov, N.: Complexity of Null- and Positivstellensatz proofs. Annals of Pure and Applied Logic **113** (2001) 153–160
22. Grigoriev, D.: Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. Theoretical Computer Science **259** (2001) 613–622
23. Grigoriev, D.: Complexity of Positivstellensatz proofs for the knapsack. Computational Complexity **10** (2001) 139–154
24. Cook, S.A., Reckhow, A.R.: The relative efficiency of propositional proof systems. Journal of Symbolic Logic **44** (1979) 36–50
25. Pitassi, T.: Algebraic propositional proof systems. In Immerman, N., Kolaitis, P.G., eds.: Descriptive Complexity and Finite Models. Volume 31 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society (1997)