Electronic Colloquium on Computational Complexity, Report No. 154 (2017)

The Relation between Polynomial Calculus, Sherali-Adams, and Sum-of-Squares Proofs

Christoph Berkholz
Humboldt-Universität zu Berlin
berkholz@informatik.hu-berlin.de

October 12, 2017

Abstract

We relate different approaches for proving the unsatisfiability of a system of real polynomial equations over Boolean variables. On the one hand, there are the static proof systems Sherali-Adams and sum-of-squares (a.k.a. Lasserre), which are based on linear and semi-definite programming relaxations. On the other hand, we consider polynomial calculus, which is a dynamic algebraic proof system that models Gröbner basis computations.

Our first result is that sum-of-squares simulates polynomial calculus: any polynomial calculus refutation of degree d can be transformed into a sum-of-squares refutation of degree $2d$ and only polynomial increase in size. In contrast, our second result shows that this is not the case for Sherali-Adams: there are systems of polynomial equations that have polynomial calculus refutations of degree 3 and polynomial size, but require Sherali-Adams refutations of degree $\Omega(\sqrt{n}/\log n)$ and exponential size.

1 Introduction

The area of *proof complexity* was founded in [6] and studies the complexity of proofs for co-NP complete problems. Traditionally, one considers proof systems for proving the unsatisfiability of (or *refuting*) a propositional formula in conjunctive normal form. If one faces a proof system, there are two important questions to ask:

1. Does the system always produce proofs of polynomial size?
2. How strong is the system compared to other proof systems?

If the answer to the first question is *yes*, in which case the system is called *p-bounded*, then NP = co-NP. Therefore, it is conjectured that no proof system is p-bounded and this has been proven for a number of weak proof systems. For the second question, one considers the notion of *polynomial simulation*: A proof system P polynomial simulates a proof system Q if for every Q-proof of size S there is a P-proof of size $poly(S)$.

Nowadays, a large part of proof complexity focuses on weak proof systems, for which the first question has already been answered negatively. One reason for this is that they often model algorithms for solving hard problems and understanding the complexity of proofs might shed light on the complexity of algorithmic approaches that implicitly or explicitly search for proofs in the underlying proof system. The (semi-)algebraic proof systems we consider in this paper also fall into this category and are used to prove the unsatisfiability of a system $\mathcal{F}$ of real polynomial equations $f_i = 0$ over n Boolean variables $x_j \in \{0, 1\}$.[1] On the one hand, we consider *polynomial calculus*, which is a dynamic algebraic proof system that allows to derive new

[1] Note that this subsumes the problem of refuting 3-CNF formulas, because a clause $x \vee \overline{y} \vee z$ can be encoded as polynomial equation $(1 - x)y(1 - z) = 0$.

ISSN 1433-8092

polynomial equations that follow from $\mathcal{F}$ line-by-line. This proof system was introduced in [5] to model Gröbner basis computations and proofs of *degree d* (where the degree of all polynomials in the derivation is bounded by d) can be found in time $n^{O(d)}$ by a bounded-degree variant of the Gröbner basis algorithm.

On the other hand, we consider the semi-algebraic proof system *Sherali-Adams* and the stronger *sum-of-squares* proof system. They are based on the linear and semi-definite programming hierarchies of Sherali-Adams [19] and Lasserre [13] and can be used to prove the unsatisfiability of a system of polynomial equations and inequalities. Proofs of degree d can be found algorithmically by solving a linear program (for Sherali-Adams) or a semi-definite program (for sum-of-squares) of size $n^{O(d)}$. Contrary to polynomial calculus, both systems are static in the sense that they provide the whole proof at once.

In order to compare these semi-algebraic proof systems with polynomial calculus, we first remark that it is known that both systems cannot be simulated by polynomial calculus. A simple example is the linear equation $\sum_{i=1}^{n} x_i = n + 1$, which has a refutation of linear size and degree 2 in Sherali-Adams and sum-of-squares, but requires polynomial calculus refutations of degree $\Omega(n)$ and size $2^{\Omega(n)}$ [11]. Our first theorem states that sum-of-squares is strictly stronger than polynomial calculus.

Theorem 1.1. *Let $\mathcal{F}$ be a system of polynomial equations over the reals. If $\mathcal{F}$ has a polynomial calculus refutation of degree d and size S, then it has a sum-of-squares refutation of degree $2d$ and size poly(S).*

For the author of this paper, this theorem was highly unexpected. In fact, there has been some evidence that the converse might be true. First, in the *non-Boolean* setting there are systems of equations that are easier to refute for polynomial calculus than for sum-of-squares [10] (see Section 2.4 for a discussion). Second, even for systems of polynomial equations over Boolean variables, separations of polynomial calculus from its static version *Nullstellensatz* were known [4].

Since sum-of-squares extends Nullstellensatz, it follows that the semi-definite lifts in the sum-of-squares/Lasserre hierarchy are necessary for "flattening" a dynamic polynomial calculus proof into a static one, although polynomial calculus is a purely algebraic system without semi-definite components. Our second theorem concerns the question whether the weaker Sherali-Adams linear programming hierarchy is already able to simulate polynomial calculus. Here we have a negative answer (that we would have expected for sum-of-squares as well).

Theorem 1.2. *There is a system $\mathcal{F}$ of polynomial equations over $\mathbb{R}[x_1, \ldots, x_n]$ such that:*

1. *$\mathcal{F}$ has a polynomial calculus refutation of degree 3 and size $O(n^2)$.*
2. *Every Sherali-Adams refutation of $\mathcal{F}$ has degree $\Omega(\sqrt{n}/\log n)$ and size $2^{\Omega(\sqrt{n}/\log n)}$.*

The lower bound is based on a modified version of the pebbling contradictions. The original pebbling contradictions have already been used to separate Nullstellensatz degree from polynomial calculus degree [4], but it turns out that they are easy to refute in Sherali-Adams. To obtain contradictions that are hard for Sherali-Adams (and still easy for polynomial calculus), we apply a substitution trick twice: first to show that the resulting contradiction requires high degree in Sherali-Adams and second to obtain a size lower bound from a degree lower bound. We believe that both techniques are also helpful for future lower bound arguments for static proof systems.

Acknowledgements Part of this work was done during the Oberwolfach workshop 1733 *Proof and complexity and Beyond* and the author acknowledges helpful discussions with Albert Atserias, Edward Hirsch, Massimo Lauria, Joanna Ochremiak, and Iddo Tzameret on Theorem 1.1. The author thanks Edward Hirsch for providing helpful comments on an earlier version the paper. We acknowledge the financial support by the German Research Foundation DFG under grant SCHW 837/5-1.

2 Proof Systems

For this section we fix a system of real polynomial equations $\mathcal{F} = \{f_1 = 0, \ldots, f_m = 0\}$ and a system of polynomial inequalities $\mathcal{H} = \{h_1 \geqslant 0, \ldots, h_s \geqslant 0\}$ over variables $x_1, \ldots, x_n$. As it is common in propositional proof complexity, we focus on the special case of polynomial equations (and inequalities) over *Boolean* variables and consider the task of proving that a system of polynomial equations (and/or inequalities) has no 0/1-solution. To enforce Boolean variables, the axioms $x_j^2 = x_j$ are always included in the proof systems. In Section 2.4 we briefly discuss non-Boolean variants.

Algebraic proof systems are used for proving the unsatisfiability of a system of multivariate polynomial equations over some field $\mathbb{F}$. As we focus on real polynomials we set $\mathbb{F} = \mathbb{R}$, unless mentioned otherwise. Semi-algebraic proof systems are used to prove the unsatisfiability of a system of polynomial equations and/or polynomial inequalities (in this setting the polynomials are always real).

2.1 Algebraic Proof Systems: Nullstellensatz and Polynomial Calculus

Nullstellensatz [2] is a static algebraic proof system that is based on Hilbert's Nullstellensatz. A *Nullstellensatz proof* of $f = 0$ from $\mathcal{F}$ is a sequence of polynomials $(g_1, \ldots, g_m; q_1, \ldots, q_n)$ such that

$$\sum_{i=1}^{m} g_i f_i + \sum_{j=1}^{n} q_j(x_j^2 - x_j) = f. \tag{1}$$

Note that the proof is sound in the sense every 0/1-assignment that satisfies $\mathcal{F}$ also satisfies $f = 0$. The *degree* of the Nullstellensatz proof is $\max\big(\{\deg(g_i f_i) \ : \ i \in [m]\} \cup \{\deg(h_j) + 2 \ : \ j \in [n]\}\big)$. The *size* of the derivation is the sum of the sizes of the binary encoding of the polynomials f, $g_i f_i$, $q_j(x_j^2 - x_j)$, each represented as a sum of monomials. A *Nullstellensatz refutation* of $\mathcal{F}$ is a proof of $-1 = 0$ from $\mathcal{F}$, in which case $\mathcal{F}$ is *unsatisfiable* (i. e., has no 0/1-solution). The Nullstellensatz system is also complete: If $\mathcal{F}$ is an unsatisfiable system of multi-linear polynomials, then it has a refutation of degree at most n.

Nullstellensatz is a static (or one-shot) proof system, as it provides the whole proof at once. The dynamic version of Nullstellensatz is *polynomial calculus* (PC) [5]. It consists of the following derivation rules for polynomial equations $(f_i = 0) \in \mathcal{F}$, polynomials f, g, variables x_j, and numbers $a, b \in \mathbb{R}$:

$$\frac{}{f_i = 0}, \quad \frac{}{x_j^2 - x_j = 0}, \quad \frac{f = 0}{x_j f = 0}, \quad \frac{g = 0 \qquad f = 0}{ag + bf = 0}. \tag{2}$$

A *polynomial calculus derivation* of $f = 0$ from $\mathcal{F}$ is a sequence $(r_1 = 0, \ldots, r_L = 0)$ of polynomial equations that are iteratively derived using the rules (2) and lead to $f = r_L = 0$. The *degree* of a derivation is the maximum degree of the polynomials in the derivation and the *size* is the sum of the sizes of the binary encoding of the polynomials in the derivation. A *polynomial calculus refutation* is a derivation of $-1 = 0$. It is straightforward to check that polynomial calculus simulates Nullstellensatz: If $\mathcal{F}$ has a Nullstellensatz refutation of degree d and size N, then it has a polynomial calculus refutation of degree d and size polynomial in N.

In both systems proofs of bounded degree d can be found in time $n^{O(d)}$: for Nullstellensatz the coefficients of the polynomials can be computed by solving a system of linear equations of size $n^{O(d)}$, and for polynomial calculus this can be done by using a bounded degree variant of the Gröbner basis algorithm [5].

2.2 Semi-algebraic proof systems: Sherali-Adams, Sum-of-Squares, Positivstellensatz

Sherali-Adams is a static proof system that models the Sherali-Adams lift-and-project hierarchy of linear programming relaxations [19]. It can also be viewed as an extension of the Nullstellensatz system. A *Sherali-Adams proof* of $f \geqslant 0$ from $(\mathcal{F}, \mathcal{H})$ is a sequence of polynomials $(g_1, \dots, g_m; q_1, \dots, q_n; p_0, \dots, p_s)$ such that

$$\sum_{i=1}^{m} g_i f_i + \sum_{j=1}^{n} q_j(x_j^2 - x_j) + p_0 + \sum_{\ell=1}^{s} p_\ell h_\ell = f, \tag{3}$$

and where every p_ℓ has the form $p_\ell = \sum_{A,B} a^\ell_{A,B} \prod_{j\in A} x_j \prod_{j\in B}(1 - x_j)$ with non-negative coefficients $a^\ell_{A,B}$.[2] Note that the polynomials $p_\ell \colon \mathbb{R}^n \to \mathbb{R}$ are positive in $[0,1]^n$ and hence the proof is sound in the sense every 0/1-assignment that satisfies $\mathcal{F}$ and $\mathcal{H}$ also satisfies $f \geqslant 0$. The *degree* (sometime called *rank*) of a Sherali-Adams proof is the maximum degree of the polynomials $g_i f_i$, $q_j(x_j^2 - x_j)$, p_0, $p_\ell h_\ell$ and the *size* is the sum of the sizes of their encoding. A Sherali-Adams refutation of $(\mathcal{F}, \mathcal{H})$ is a proof of $-1 \geqslant 0$ from $(\mathcal{F}, \mathcal{H})$. Note that every Nullstellensatz refutation of $\mathcal{F}$ is a Sherali-Adams refutation of $(\mathcal{F}, \emptyset)$ by choosing $p_0 = 0$.

Sum-of-squares (SOS) is a semi-algebraic proof system that extends Nullstellensatz and Sherali-Adams. It models the Lasserre hierarchy of semi-definite programming relaxations [13], for which reason it is sometimes called *Lasserre*, and also builds on Putinar's Positivstellensatz [18]. The difference to Sherali-Adams is that the positive polynomials p_ℓ are now sums of squares. Formally, a *sum-of-squares proof* of $f \geqslant 0$ from $(\mathcal{F}, \mathcal{H})$ is a sequence of polynomials $(g_1, \dots, g_m; q_1, \dots, q_n; p_0, \dots, p_s)$ such that

$$\sum_{i=1}^{m} g_i f_i + \sum_{j=1}^{n} q_j(x_j^2 - x_j) + p_0 + \sum_{\ell=1}^{s} p_\ell h_\ell = f, \tag{4}$$

and where every p_ℓ has the form $p_\ell = \sum_{c=1}^{t_\ell} (p_{\ell,c})^2$ (and is encoded as such) for arbitrary polynomials $p_{\ell,c}$ (in standard monomial form). Again, the *degree* of a proof is the maximum degree of the polynomials $g_i f_i$, $q_j(x_j^2 - x_j)$, p_0, $p_\ell h_\ell$, the *size* is the sum of the sizes of their encoding. A *sum-of-squares refutation* is a proof of $-1 \geqslant 0$. It is not hard to see that the positive polynomials $p = \sum_{A,B} a_{A,B} \prod_{j\in A} x_j \prod_{j\in B}(1 - x_j)$ in the Sherali-Adams proof system have a sum-of-squares proof (from $\mathcal{F} = \mathcal{H} = \emptyset$) of degree $|A| + |B| + 1$ and size *poly*(p). It immediately follows that sum-of-squares simulates Sherali-Adams.

Lemma 2.1. *If $(\mathcal{F}, \mathcal{H})$ has a Sherali-Adams refutation of degree d and size N, then it has a sum-of-squares refutation of degree $d + 1$ and size poly(N).*

Another semi-algebraic system that is related to sum-of-squares is *Positivstellensatz*. It builds on Stengle's Positivstellensatz (independently proven by Krivine [12] and Stengle [20]), which has also been used to define a hierarchy of relaxations, see [17]. Our definition of the Positivstellensatz proof system follows the one introduced in [10], a different way of formalising Stengle's Positivstellensatz as a proof system (without focusing on complexity) was presented in [14]. We remark that Stengle's Positivstellensatz and the Positivstellensatz proof system as defined in [10] do not necessarily include the Boolean axioms $x_j^2 - x_j$ and also work for polynomials over non-Boolean variables. To be precise, we will call the system that is named "Positivstellensatz" in [10] "non-Boolean Positivstellensatz" in this paper (see Section 2.4). To define the proof system, we consider for the system of polynomial inequalities $\mathcal{H} = \{h_1 \geqslant 0, \dots, h_s \geqslant 0\}$ the system $\widehat{\mathcal{H}} = \{\prod_{\ell\in I} h_\ell \geqslant 0 \ : \ I \subseteq [s]\}$, which extends $\mathcal{H}$ by taking products of polynomial inequalities.

[2] We assume that the p_ℓ are explicitely provided in this form, whereas g_i and q_j are arbitrary polynomials encoded in the standard way as a sum of monomials.

Clearly, $(\mathcal{F},\mathcal{H})$ is satisfiable if and only if $(\mathcal{F},\widehat{\mathcal{H}})$ is satisfiable. A *Positivstellensatz proof* of $f \geqslant 0$ from $(\mathcal{F},\mathcal{H})$ is a sum-of-squares proof of $f \geqslant 0$ from $(\mathcal{F},\widehat{\mathcal{H}})$. Note that on systems of polynomial equations (where $\mathcal{H}=\emptyset$) sum-of-squares and Positivstellensatz are the same.

One way of combining polynomial calculus with semi-algebraic proof systems is as follows. Note that a Sherali-Adams, sum-of-squares, or Positivstellensatz proof of $f \geqslant 0$ can be decomposed to

$$g + p_0 + \sum_{\ell} p_\ell h_\ell = f, \tag{5}$$

$$\sum_{i=1}^{m} g_i f_i + \sum_{j=1}^{n} q_j(x_j^2 - x_j) = g, \tag{6}$$

where (6) is a Nullstellensatz proof of $g=0$. By replacing this Nullstellensatz proof of $g=0$ with a polynomial calculus proof of $g=0$, we obtain dynamic versions of the static semi-algebraic proof systems. The dynamic version of Positivstellensatz is called *Positivstellensatz calculus* and was also introduced in [10]. However, the proof of Theorem 1.1 (in particular Lemma 3.1) implies that Positivstellensatz and Positivstellensatz calculus can simulate each other.

Corollary 2.2. *If $(\mathcal{F},\mathcal{H})$ has a Positivstellensatz calculus refutation of degree d and size S, then it has a Positivstellensatz refutation of degree $2d$ and size poly(S).*

Proof. By definition, a Positivstellensatz calculus refutation of $(\mathcal{F},\mathcal{H})$ is a polynomial calculus derivation of $-1-p_0-\sum_\ell p_\ell h_\ell$ from $\mathcal{F}$, where $h_\ell \in \widehat{\mathcal{H}}$. By Lemma 3.1, there is a degree-$2d$, size *poly*(S) sum-of-squares proof of non-negativity of

$$-\big(-1-p_0-\textstyle\sum_\ell p_\ell h_\ell\big)^2 = -1-2p_0-p_0^2-(2+2p_0)\big(\sum_\ell p_\ell h_\ell\big)-\big(\sum_\ell \sum_{\ell'} p_\ell p_{\ell'} h_{\ell'} h_\ell\big), \tag{7}$$

from $(\mathcal{F},\widehat{\mathcal{H}})$, which in turn is a Positivstellensatz refutation of $(\mathcal{F},\mathcal{H})$. □

For completeness, we mention that there are also dynamic semi-algebraic proof systems that are based on the Lovász-Schrijver lift-and-project method [15] and where one can infer polynomial *inequalities* line-by-line (see [9] for an overview). These systems are, however, much stronger and somewhat different from the proof systems considered in this paper.

2.3 Twin variables

In all the proof systems mentioned above, it might be useful to introduce *twin variables*: for every variable x_j one has available the formal variable $x_j^{\neg}$ that expresses its "negation" $1-x_j$. To ensure that they are complementary, the additional polynomial equality $x_j + x_j^{\neg} = 1$ is always present in $\mathcal{F}$. Except for Sherali-Adams this does not change the definition of the proof systems, as it only affects the input encoding. For Sherali-Adams with twin variables, it is additionally assumed that every p_ℓ has now the form $p_\ell = \sum_{A,B} a^\ell_{A,B} \prod_{j\in A} x_j \prod_{j\in B} x_j^{\neg}$ [7].

Note that inclusion of twin variables does not affect the degree of a refutation, but it might affect the size, as for example the polynomial $\prod_{j\in[n]}(1-x_j)$, which has size $2^{\Theta(n)}$, can be more succinctly expressed as $\prod_{j\in[n]} x_j^{\neg}$, which is of size $\Theta(n)$. We are, however, not aware of any formal separation of (semi-)algebraic proof systems with and without twin variables with respect to proof size.

Twin variables are particularly useful when encoding CNF formulas into polynomial equations. It is known that polynomial calculus with twin variables, which is called *polynomial calculus resolution* (PCR) [1], can polynomially simulate the resolution calculus [5, 1]. The same is true for Sherali-Adams [7] and hence sum-of-squares, but not for Nullstellensatz[3].

Remark 2.3. Theorem 1.1 and Theorem 1.2 remain true in the presence of twin variables.

[3] This essentially follows from the degree lower bounds in [4] and Lemma 4.8.

2.4 The non-Boolean case

It is also conceivable to consider (semi-)algebraic proof systems over non-Boolean variables. In this case the additional Boolean axioms $x_j^2 - x_j = 0$ are omitted in the definitions (formally, we require that $q_j = 0$ in the above definitions). Note that there is no meaningful non-Boolean variant of the Sherali-Adams proof system, as its correctness (specifically, the non-negativity of the polynomials p_ℓ) crucially depends on the fact that all variables are between 0 and 1. However, non-Boolean variants of Nullstellensatz, polynomial calculus, sum-of-squares, and Positivstellensatz are still sound proof systems. It follows from Stengle's Positivstellensatz [20], that Positivstellensatz is also refutational complete in this setting. For sum-of-squares this does only hold if we put additional requirements on $\mathcal{F} \cup \mathcal{H}$ (being *Archimedian* [18]). Non-Boolean Nullstellensatz and polynomial calculus are only complete over algebraically closed fields (such as the complex numbers).

We remark that in these systems it is no longer the case that every unsatisfiable multi-linear system of equations over n variables has a refutation of degree n: for example, the so-called telescopic system $\mathcal{F}_n^{\text{ts}} := \{yx_1 = 1,\ x_1^2 = x_2,\ x_2^2 = x_3,\ \ldots,\ x_{n-1}^2 = x_n,\ x_n = 0\}$ requires exponential refutation degree in Nullstellensatz [3] and sum-of-squares [10]. Moreover, the same example shows that the simulation of polynomial calculus by sum-of-squares (Theorem 1.1) does not hold in the non-Boolean case:

Theorem 2.4 ([10])**.** *Let $\mathcal{F}_n^{ts}$ be the telescopic system as defined above.*

1. *$\mathcal{F}_n^{ts}$ has a non-Boolean Nullstellensatz (hence sum-of-squares) refutation of degree $2^{O(n)}$.*
2. *$\mathcal{F}_n^{ts}$ has a non-Boolean polynomial calculus refutation of degree $O(n)$.*
3. *Every non-Boolean sum-of-squares refutation of $\mathcal{F}_n^{ts}$ has degree $2^{\Omega(n)}$.*

3 Sum-of-Squares Simulates Polynomial Calculus

This section is dedicated to the proof of Theorem 1.1. Let us fix an unsatisfiable system of polynomial equations $\mathcal{F} = \{f_1 = 0, \ldots, f_m = 0\}$. Let $(r_1 = 0, \ldots, r_L = 0)$ be a polynomial calculus derivation of $r_L = 0$ from $\mathcal{F}$ of degree d and size S. Let $\mathfrak{a}$ be the minimal integer such that every coefficient c in the proof satisfies $\mathfrak{a}^{-1} \leqslant 4c^2 \leqslant \mathfrak{a}$. Hence, the largest encoding size of coefficient is $\Theta(\log \mathfrak{a})$. Theorem 1.1 follows immediately from the following inductive lemma.

Lemma 3.1. *There are polynomials $q_1, \ldots, q_L$ and $p_1, \ldots, p_L$ of size at most poly(S) such that for every $\widehat{L} \leqslant L$ there are nonnegative coefficients $\mathfrak{a}^{-\widehat{L}} \leqslant a_i, b_\ell, c_\ell \leqslant \mathfrak{a}^{\widehat{L}}$, such that*

$$\sum_{i=1}^{m}(-a_i f_i)f_i + \sum_{\ell=1}^{\widehat{L}} b_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{\widehat{L}} c_\ell p_\ell^2 = -(r_{\widehat{L}})^2 \tag{8}$$

is a sum-of-squares proof of $-(r_{\widehat{L}})^2 \geqslant 0$ of degree $2d$.

Proof. First note that (8) is indeed a sum-of-squares proof of the form (4) since

$$\sum_{\ell=1}^{\widehat{L}} b_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) = \sum_{j=1}^{n} \Big(\sum_{\ell:\ j_\ell = j} b_\ell q_\ell \Big)\,(x_j^2 - x_j) \tag{9}$$

and $c_\ell p_\ell^2 = (\sqrt{c_\ell} p_\ell)^2$ (as we require $c_\ell \geqslant 0$). Although we shall first provide the polynomials q_ℓ and p_ℓ, we just assume that we have already done so and postpone their definition for ease of exposition. The proof is now by induction on $\widehat{L}$ and we do a case analysis on the four types of derivation rules (2). First suppose that $r_{\widehat{L}} = f_i$ is an axiom from $\mathcal{F}$. Then we can easily derive

$-(r_{\widehat{L}})^2$ in sum-of-squares by defining $p_{\widehat{L}} = q_{\widehat{L}} := 0$, setting a_i to 1 and all other coefficients to 0. The case of a Boolean axiom $r_{\widehat{L}} = x_j^2 - x_j$ is also simple. We define $q_{\widehat{L}} := -(x_j^2 - x_j)$ as well as $p_{\widehat{L}} := 0$, set $b_{\widehat{L}}$ to 1 and all other coefficients to 0 in order to derive $-(r_{\widehat{L}})^2$.

Now suppose that $r_{\widehat{L}} = x_{j'} r_{L'}$ is obtained by multiplying a previously derived polynomial $r_{L'}$ (for some $L' < L$) by a variable $x_{j'}$. By induction assumption we have a sum-of-squares proof of $-(r_{L'})^2 \geqslant 0$ of degree $2d$:

$$\sum_{i=1}^{m}(-a_i f_i)f_i + \sum_{\ell=1}^{L'} b_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{L'} c_\ell p_\ell^2 = -(r_{L'})^2. \tag{10}$$

Now we want to turn this proof into a proof of $-(x_{j'} r_{L'})^2 \geqslant 0$. Of course, we could do this by just multiplying everything by $x_{j'}^2$. However, this would increase the degree of the refutation to $2d+2$! Instead, we use the sum of squares polynomials in order to simulate the multiplication rule in polynomial calculus without increasing the degree. We define $p_{\widehat{L}} := r_{L'} - x_{j'} r_{L'}$ as well as $q_{\widehat{L}} := -2(r_{L'})^2$ and observe that

$$\begin{aligned}(p_{\widehat{L}})^2 + q_{\widehat{L}} \cdot (x_{j'}^2 - x_{j'}) &= (r_{L'})^2 - 2x_{j'}(r_{L'})^2 + x_{j'}^2(r_{L'})^2 - 2x_{j'}^2(r_{L'})^2 + 2x_{j'}(r_{L'})^2 &(11)\\ &= (r_{L'})^2 - (x_{j'} r_{L'})^2. &(12)\end{aligned}$$

By adding them to (10) we derive $-(x_{j'} r_{L'})^2 \geqslant 0$ without increasing the degree. Formally, we define $j_{\widehat{L}} := j'$, set $b_{\widehat{L}} = c_{\widehat{L}} = 1$ and obtain

$$\sum_{i=1}^{m}(-a_i f_i)f_i + \sum_{\ell=1}^{\widehat{L}} b_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{\widehat{L}} c_\ell p_\ell^2 = -(x_{j'} r_{L'})^2 = -(r_{\widehat{L}})^2. \tag{13}$$

The remaining case is derivation of $r_{\widehat{L}} = a \cdot r_{L'} + b \cdot r_{L''}$ for $a, b \in \mathbb{R}$ as a linear combination of two previously derived polynomials $r_{L'}$ and $r_{L''}$. By induction assumption we have

$$\sum_{i=1}^{m}(-a'_i f_i)f_i + \sum_{\ell=1}^{L'} b'_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{L'} c'_\ell p_\ell^2 = -(r_{L'})^2 \qquad \text{and} \tag{14}$$

$$\sum_{i=1}^{m}(-a''_i f_i)f_i + \sum_{\ell=1}^{L''} b''_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{L''} c''_\ell p_\ell^2 = -(r_{L''})^2. \tag{15}$$

Our goal is to devise a sum-of-squares proof of $-(r_{\widehat{L}})^2 = -a^2(r_{L'})^2 - 2ab \cdot r_{L'} r_{L''} - b^2(r_{L''})^2$. For this we define $p_{\widehat{L}} := a \cdot r_{L'} - b \cdot r_{L''}$ and $q_{\widehat{L}} := 0$. To derive $-(r_{\widehat{L}})^2$, we multiply the sum-of-squares proof (14) by $2a^2$, multiply (15) by $2b^2$, and then add both proofs together with $(p_{\widehat{L}})^2$. More precisely, we set $a_i = 2a^2 a'_i + 2b^2 a''_i$ for all $i \in [m]$; $b_\ell = 2a^2 b'_\ell + 2b^2 b''_\ell$, $c_\ell = 2a^2 c'_\ell + 2b^2 c''_\ell$ for all $\ell \leqslant \max(L', L'')$; $c_{\widehat{L}} = 1$ and set the remaining coefficients to 0. Then we obtain

$$\begin{aligned}\sum_{i=1}^{m}(-a_i f_i)f_i + \sum_{\ell=1}^{\widehat{L}} b_\ell q_\ell (x_{j_\ell}^2 - x_{j_\ell}) + \sum_{\ell=1}^{\widehat{L}} c_\ell p_\ell^2 &= -2a^2(r_{L'})^2 - 2b^2(r_{L''})^2 + (p_{\widehat{L}})^2 &(16)\\ &= -a^2(r_{L'})^2 - b^2(r_{L''})^2 - 2ab \cdot r_{L'} r_{L''} &(17)\\ &= -(r_{\widehat{L}})^2 &(18)\end{aligned}$$

By the definition of $\mathfrak{a}$, the factors $2a^2$ and $2b^2$ are bounded by $2\mathfrak{a}^{-1}$ and $\frac{1}{2}\mathfrak{a}$ from below and above. Since by induction assumption we have $\mathfrak{a}^{-\widehat{L}+1} \leqslant a'_i, b'_\ell, c'_\ell, a''_i, b''_\ell, c''_\ell \leqslant \mathfrak{a}^{\widehat{L}-1}$, it follows that $\mathfrak{a}^{-\widehat{L}} \leqslant a_i, b_\ell, c_\ell \leqslant \mathfrak{a}^{\widehat{L}}$. This concludes the proof of Lemma 3.1. □

Proof of Theorem 1.1. The theorem follows immediately from Lemma 3.1, since every degree-d polynomial calculus derivation of $-1 = 0$ can be transformed into a degree-$2d$ sum-of-squares proof of non-negativity of $-(-1)^2 = -1$. By the requirements in the Lemma the size of the sum-of-squares proof is $poly(S)$. □

4 Sherali-Adams does not Simulate Polynomial Calculus

The system of polynomial equations that separates Sherali-Adams from polynomial calculus (Theorem 1.2) is a variant of the pebbling contradictions, which are unsatisfiable propositional formulas that are based on the black pebble game. These formulas and their variants have found several applications in propositional proof complexity. For an in-depth treatment of the history and some of the applications of pebbling in proof complexity we refer the reader to the survey [16].

Let us fix some notation. In a directed graph $\mathcal{G} = (V, E)$ we let $N^-(v) = \{u \ : \ (u, v) \in E\}$ be the set of incoming and $N^+(v) = \{w \ : \ (v, w) \in E\}$ be the set of outgoing neighbours of a vertex $v \in V$. The vertex sets $S = \{v \ : \ N^-(v) = \emptyset\}$ and $T = \{v \ : \ N^+(v) = \emptyset\}$ are called the *sources* and the *sinks* of G. A *circuit* is a directed acyclic graph $\mathcal{G}$ with a unique sink t and where every non-source vertex $v \in V \setminus S$ has two incoming neighbours.

The *(black) pebble game* is a one-player game played on a circuit $\mathcal{G} = (V, E)$. The player has available a pool of P pebbles and the game proceeds by placing and removing pebbles on the vertices of $\mathcal{G}$. In each round the player can do one of the following moves:

1. place a pebble on a sink vertex $s \in S$,
2. place a pebble on $w \in V \setminus S$ if there are pebbles on both vertices in $N^-(w)$, or
3. remove an arbitrary pebble.

The player wins the game when he places a pebble on the sink node t. It is obvious, that the player can always win the game with $|V|$ pebbles and the *(black) pebbling price* $\mathrm{Peb}(\mathcal{G}) \leqslant |V|$ is the minimal number P such that the player wins the black pebble game on $\mathcal{G}$ with P pebbles. For our lower bounds we will consider circuits $\mathcal{G}$ with high pebbling price.

Theorem 4.1 ([8])**.** *For every large enough n there is a circuit $\mathcal{G}$ with n vertices and* $\mathrm{Peb}(\mathcal{G}) = \Omega(n/\log n)$.

The *pebbling contradiction* $\mathcal{F}_\mathcal{G}$ for a circuit $\mathcal{G} = (V, E)$ is the system of polynomial equations over Boolean variables $\{x_v \ : \ v \in V\}$ that contains the following equations:

$$x_s = 1, \qquad \text{for all } s \in S, \tag{19}$$

$$x_u x_v = x_u x_v x_w, \qquad \text{for all } w \in V \setminus S \text{ and } N^-(w) = \{u, v\}, \text{ and} \tag{20}$$

$$x_t = 0, \qquad \text{for the sink } t. \tag{21}$$

It is easy to see that this system is unsatisfiable. Moreover, we remark that $\mathcal{F}_\mathcal{G}$ is the standard encoding of the CNF pebbling contradiction, which contains clauses x_s, $\overline{x}_u \vee \overline{x}_v \vee x_w$, and $\overline{x}_t$. As this CNF can be easily refuted in resolution using unit propagation, it follows that this system is easy to refute in any proof system that simulates resolution, such as polynomial calculus, Sherali-Adams, and sum-of-squares. For later reference, the next lemma formulates this claim for polynomial calculus.

Lemma 4.2. *$\mathcal{F}_\mathcal{G}$ has a polynomial calculus refutation of degree 3 and size $O(n)$ for any n-vertex circuit $\mathcal{G}$.*

Proof. For a vertex $v \in V$ let $\mathrm{dist}(v)$ be the smallest distance to a source vertex in S. By induction on $\mathrm{dist}(v)$ we derive the equation $x_v = 1$. If $\mathrm{dist}(v) = 0$, then this equation is an axiom. For the induction step let w be a vertex with incoming neighbours $N^-(w) = \{u, v\}$ and assume that (a) $x_u = 1$ as well as (b) $x_v = 1$ was already derived. We also have the axiom (c) $x_u x_v = x_u x_v x_w$ available. Multiplying (a) with x_v gives (d) $x_u x_v = x_v$ and by a linear combination with (b) we get (e) $x_u x_v = 1$. Multiplying (e) with x_w results in (f) $x_u x_v x_w = x_w$. Now we can derive $x_w = 1$ by a linear combination of (c), (e), and (f). The lemma follows since from $x_t = 1$ and the axiom $x_t = 0$ we can derive $-1 = 0$. □

In [4] it was shown that every Nullstellensatz refutation of $\mathcal{F}_\mathcal{G}$ requires degree $\mathrm{Peb}(\mathcal{G})$ and hence this system separates Nullstellensatz degree from polynomial calculus degree. However, it is not hard to construct a Nullstellensatz refutation of $\mathcal{F}_\mathcal{G}$ that has size $poly(n)$. Therefore, this example does *not* separate both systems with respect to proof size. Moreover, as mentioned before, this system is also easy for Sherali-Adams (with respect to size *and* degree). To prove our separation theorem between Sherali-Adams and polynomial calculus, we modify the formula a bit in order to make it hard for Sherali-Adams, while at the same time it remains easy for polynomial calculus. We do this by substituting for every variable x_v the sum of fresh variables according to the following definition.

Definition 4.3. Let $\mathcal{F}$ be a set of polynomial equations over variables $x_1, \ldots, x_n$ and $k \geqslant 1$. The system $\mathcal{F}[+_k]$ is obtained from $\mathcal{F}$ be replacing every variable x_i in every $f \in \mathcal{F}$ by the sum $x_{i,1} + \cdots + x_{i,k}$ of k new variables and including the additional polynomial equations $x_{i,\ell}x_{i,\ell'} = 0$ for all $i \in [n]$ and $1 \leqslant \ell < \ell' \leqslant k$.

The following lemma shows that after substitution the system remains easy to refute in polynomial calculus.

Lemma 4.4. *Let $\mathcal{F}$ be a set of polynomial equations and suppose there is a polynomial calculus refutation of $\mathcal{F}$ of degree d and size S. Then $\mathcal{F}[+_k]$ has a polynomial calculus refutation of degree d and size $O(k^d S)$.*

Proof. We obtain the new proof by substituting all variables x_i by $x_{i,1} + \cdots + x_{i,k}$ and expand the polynomials to monomial form (this increases the size by a factor of k^d). It remains to check that the substituted equations form a polynomial calculus refutation of $\mathcal{F}[+_k]$. It is clear that a former derivation of an axiom $f \in \mathcal{F}$ is now a derivation of an substituted axiom from $\mathcal{F}[+_k]$. A derivation of a Boolean axiom $x_i^2 = x_i$ translates to $(\sum_{\ell\in[k]} x_{i,\ell})^2 = \sum_{\ell\in[k]} x_{i,\ell}$, which can be derived using the Boolean axioms $x_{i,\ell}^2 = x_{i,\ell}$ and the additional equations $x_{i,\ell}x_{i,\ell'} = 0$ (see Definition 4.3). The substituted variant of a linear combination of two previously derived polynomials f, g is just the linear combination of the substituted versions of f and g. Multiplication by a variable x_j to a polynomial in the original proof translates to multiplying by $\sum_{\ell\in[k]} x_{j,\ell}$, which can be simulated by k separate multiplications of $x_{j,1}, \ldots, x_{j,k}$ and subsequent addition steps. □

To obtain a system of equations that is hard for Sherali-Adams and easy for polynomial calculus we apply two substitution steps to the formula $\mathcal{F}_\mathcal{G}$ for circuits from Theorem 4.1. First we prove that every refutation of $\mathcal{F}_\mathcal{G}[+_n]$ in Sherali-Adams requires degree $d = \mathrm{Peb}(\mathcal{G})$. In the second step we show that a degree d lower bound for an arbitrary instance $\mathcal{F}$ translates to a $2^{\Omega(d)}$ size lower bound for $\mathcal{F}[+_2]$. Together we obtain that $\mathcal{F}_\mathcal{G}[+_n][+_2]$ requires high degree *and* size in Sherali-Adams. We will use a common approach for proving lower bounds in static proof systems and define a solution for the "dual" system.

Definition 4.5. A mapping $D\colon \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ is a *d-evaluation* if it satisfies the following conditions.

(D1) D is linear: $D(af + bg) = aD(f) + bD(g)$ for all $f, g \in \mathbb{R}[x_1, \ldots, x_n]$

(D2) D is multi-linear: $D(\prod_j x_j^{d_j}) = D(\prod_j x_j)$

(D3) $D(f \cdot f_i) = 0$ for every axiom $f_i \in \mathcal{F}$ and $f \in \mathbb{R}[x_1, \ldots, x_n]$ with $\deg(f) \leqslant d - \deg(f_i)$

(D4) $D\big(\prod_{j\in A} x_j \prod_{j\in B}(1 - x_j)\big) \geqslant 0$ for all $A, B \subseteq [n]$ with $|A \cup B| \leqslant d$.

It is not hard to verify that the existence of a d-evaluation implies that there is no Sherali-Adams refutation of degree d: suppose for contradiction that there is a Sherali-Adams refutation of degree d of the form

$$\sum_{i=1}^{m} g_i f_i + \sum_{j=1}^{n} q_j(x_j^2 - x_j) + p_0 = -1, \tag{22}$$

with $p_0 = \sum_{A,B} a^0_{A,B} \prod_{j\in A} x_j \prod_{j\in B}(1-x_j)$. Now we apply D to both sides of the equation. From (D3) it follows that $D(g_i f_i) = 0$, from (D2) we obtain $D(q_j(x_j^2 - x_j)) = 0$, and from (D4) it follows that $D(p_0) \geqslant 0$. By linearity (D1) the left hand side is evaluated to something non-negative, whereas on the right-hand side we have $D(-1) = -1$.

Due to the multi-linearity (D2) the lower bound technique actually proves something stronger. The *ml-degree* of a polynomial is the degree of its multi-linearisation, i. e., the maximum number of distinct variables in a monomial. We immediately get the following lemma.

Lemma 4.6. *If a system of multi-linear equations $\mathcal{F}$ has a d-evaluation D, then there is no Sherali-Adams refutation of $\mathcal{F}$ that has ml-degree $\leqslant d$.*

The next lemma is proven by constructing a d-evaluation.

Lemma 4.7. *Let $\mathcal{G}$ be a circuit with n vertices. Every Sherali-Adams refutation of $\mathcal{F}_{\mathcal{G}}[+_k]$ requires ml-degree at least* $\min(\mathrm{Peb}(G), k/2)$.

Proof. Let $d < \min(\mathrm{Peb}(G), k/2)$ and suppose for contradiction that there is a Sherali-Adams refutation of ml-degree d. By Lemma 4.6 it suffices to define an operator D that satisfies (D1)–(D4). We start by defining D on multi-linear terms. We call a multi-linear term *inconsistent*, if it contains two distinct variables $x_{v,\ell}$ and $x_{v,\ell'}$ for some $v \in V$. If $g = \prod_{(v,\ell)\in I} x_{v,\ell}$ is an inconsistent term, we define $D(g) := 0$. Otherwise, $g = \prod_{(v,\ell)\in I} x_{v,\ell} = \prod_{u\in U} x_{u,\ell_u}$ and the value of $D(g) := \widetilde{D}(U)$ will only depend on the set $U \subseteq V$. To define the mapping $\widetilde{D}\colon 2^V \to \mathbb{R}$, we say that $U \subseteq V$ is *reachable*, if the player has a strategy in the black pebble game with d pebbles to reach a position where exactly the vertices in U are pebbled. The mapping is now defined as follows.

$$\widetilde{D}(U) := \begin{cases} \left(\frac{1}{k}\right)^{|U|}, \text{ if } U \text{ is reachable,} \\ 0, \text{ otherwise.} \end{cases} \tag{23}$$

We extend the definition of D to all polynomials by (multi-)linearity. Note that this completes the definition of D and immediately satisfies (D1), (D2), as well as (D3) for the axioms $x_{i,\ell}x_{i,\ell'} = 0$ introduced by Definition 4.3. To verify (D4), we have to show that $D(p) \geqslant 0$ for every polynomial $p = \prod_{(v,\ell)\in I} x_{v,\ell} \prod_{(v,\ell)\in J}(1 - x_{v,\ell})$ of degree at most d. First note that if $I \cap J \neq \emptyset$, then $D(p) = 0$ since the mapping D satisfies (D2). Therefore, we may assume that p is multi-linear when multiplied out to monomial form. If $\prod_{(v,\ell)\in I} x_{v,\ell}$ is either inconsistent or it is consistent and defines a non-reachable position, then $D(p) = 0$ and we are done. Otherwise, $D(\prod_{(v,\ell)\in I} x_{v,\ell}) = k^{-|I|}$ and we get

$$\begin{aligned} D(p) &= \left(\frac{1}{k}\right)^{|I|} + \sum_{\emptyset\neq K\subseteq J} (-1)^{|K|} D\left(\prod_{(v,\ell)\in K\cup I} x_{v,\ell}\right) \\ &\geqslant \left(\frac{1}{k}\right)^{|I|} - \sum_{\emptyset\neq K\subseteq J} \left(\frac{1}{k}\right)^{|I|+|K|} \\ &= \left(\frac{1}{k}\right)^{|I|} \left(1 - \sum_{z=1}^{|J|} \binom{|J|}{z} \left(\frac{1}{k}\right)^z\right). \end{aligned}$$

Because we have have $|J| \leqslant d < k/2$ it follows that

$$\sum_{z=1}^{|J|} \binom{|J|}{z} \left(\frac{1}{k}\right)^z < \sum_{z=1}^{|J|} \left(\frac{k}{2}\right)^z \left(\frac{1}{k}\right)^z < \sum_{z=1}^{\infty} 2^{-z} = 1.$$

Hence, $D(p) > 0$ and property (D4) is proven. It remains to verify (D3) for all three types of substituted axioms. For every multi-linear term g we need to check:

$$D\Big(g \cdot \Big(\textstyle\sum_{\ell=1}^{k} x_{s,\ell}\Big)\Big) = D(g)\,, \tag{24}$$

$$D\Big(g \cdot \Big(\textstyle\sum_{\ell=1}^{k} x_{u,\ell}\Big)\Big(\textstyle\sum_{\ell=1}^{k} x_{v,\ell}\Big)\Big) = D\Big(g \cdot \Big(\textstyle\sum_{\ell=1}^{k} x_{u,\ell}\Big)\Big(\textstyle\sum_{\ell=1}^{k} x_{v,\ell}\Big)\Big(\textstyle\sum_{\ell=1}^{k} x_{w,\ell}\Big)\Big)\,, \tag{25}$$

$$D\Big(g \cdot \Big(\textstyle\sum_{\ell=1}^{k} x_{t,\ell}\Big)\Big) = 0, \tag{26}$$

where $s \in S$ is a source, $w \in V \setminus S$ with $N^-(w) = \{u, v\}$, and t is the sink. First suppose that g is either inconsistent or defines a position U that is not reachable. In both cases everything above evaluates to 0. Hence, let $g = \prod_{u \in U} x_{u,\ell_u}$ for a reachable vertex set U. In the case of (24), we have $|U| \leqslant d - 1$. If $s \in U$, then $D(x_{s,\ell_s} g) = D(g)$, since we satisfy (D2). Since the other summands $x_{s,\ell} g$ are inconsistent for $\ell \neq \ell_s$, they evaluate to 0 and the equality (24) holds. Now assume that $s \notin U$. We have that $U \cup \{s\}$ is reachable as well, since the player has at least one pebble remaining and can place it on the source s. It follows that $D\Big(g \cdot \Big(\sum_{\ell=1}^{k} x_{s,\ell}\Big)\Big) = k \cdot (\frac{1}{k})^{|U|+1} = (\frac{1}{k})^{|U|} = D(g)$.

Checking (25) for non-source vertices w with $N^-(w) = \{u, v\}$ is similar. Here we have $|U| \leqslant d - 3$ and by the rules of the game we know that $U \cup \{u, v\}$ is reachable if and only if $U \cup \{u, v, w\}$ is reachable. Hence, if $U \cup \{u, v\}$ is not reachable, both sides evaluate to 0. Otherwise, by a case analysis on the shape of $U \cap \{u, v, w\}$, one can easily verify that both sides evaluate to $(\frac{1}{k})^{|U|}$.

For the source vertex t, note that since $d < \mathrm{Peb}(G)$, no position that contains t is reachable. Hence, $D(g x_{t,\ell}) = 0$ for all $\ell \in [k]$ and the equality (26) holds. This concludes the proof of the lemma. □

Lemma 4.7 (together with Theorem 4.1 and Lemma 4.2) already provides a separation between degree in Sherali-Adams and polynomial calculus. To separate the proof size we need the following lifting lemma.

Lemma 4.8. *Let $\mathcal{F}$ be a system of multi-linear polynomial equations and let P be one of the proof systems Nullstellensatz, Sherali-Adams, or sum-of-squares. If every P-refutation of $\mathcal{F}$ has ml-degree at least d, then every P-refutation of $\mathcal{F}[+_2]$ has ml-degree at least d and size $\Omega(2^d)$.*

Proof. Let $\mathcal{F} = \{f_1 = 0, \ldots, f_m = 0\}$ over variables $\{x_1, \ldots, x_n\}$ and consider a P-refutation

$$\sum_{i=1}^{m} g_i f_i' + \sum_{j=1}^{n} \sum_{\ell=1}^{2} q_{j,\ell}(x_{j,\ell}^2 - x_{j,\ell}) + p_0 = -1 \tag{27}$$

of $\mathcal{F}[+_2] = \{f_1' = 0, \ldots, f_m' = 0\} \cup \{x_{1,1}x_{1,2} = 0, \ldots, x_{n,1}x_{n,2} = 0\}$. Suppose that this refutation has size 2^{d-1} and let $L \leqslant 2^{d-1}$ be the total number of large monomials of ml-degree $\geqslant d$ in the refutation (i. e. in the polynomials $g_i f_i$, $q_j(x_j^2 - x_j)$, and p_0). We consider the set Γ of all restrictions that set for every j exactly one of the variables $x_{j,1}$ and $x_{j,2}$ to 0 and leaves the other variable unset. It follows that for every $\gamma \in \Gamma$ the set of equations $\mathcal{F}_\gamma$ that results from $\mathcal{F}[+_2]$ by restricting the variables according to γ agrees with $\mathcal{F}$ (modulo renaming the variables). Moreover, by applying the restriction to (27) one obtains a P-refutation of $\mathcal{F}_\gamma$. It remains to argue that if L is too small, then choosing a restriction $\gamma \in \Gamma$ uniformly at random might end up with a refutation of $\mathcal{F}_\gamma$ of ml-degree $< d$, contradicting the assumption. This follows by a simple union bound argument. First note that the probability that a monomial of ml-degree $\geqslant d$ is not set to 0 by a restriction γ is $\leqslant (\frac{1}{2})^d$.[4] Furthermore, the probability that the restricted

[4]This claim does only hold for "ml-degree" and not for "degree" and that's why we consider this notion.

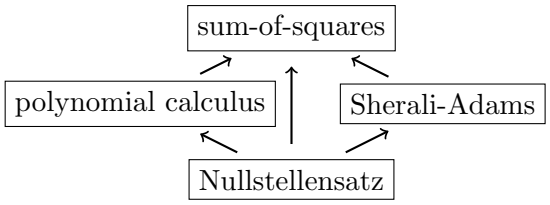

Figure 1: Relation between the proof systems. An arrow $\mathsf{P} \to \mathsf{Q}$ indicates that a proof in system P of degree d and size S can be converted into a proof in system Q of degree $O(d)$ and size $poly(S)$. Whenever there is no irreflexive arrow, it is known that the simulation does not hold.

refutation has ml-degree $< d$ is bounded by the probability that at least one large monomial does not vanish. Since we have

$$\mathbb{P}_{\gamma\in\Gamma}[\text{ex. monomial of ml-degree} \geqslant d \text{ that does not vanish}] \leqslant L \cdot \left(\tfrac{1}{2}\right)^d \leqslant \tfrac{1}{2}, \tag{28}$$

which is bounded away from 1, the lemma follows. □

By combining Lemma 4.7 and Lemma 4.8 we can now prove Theorem 1.2.

Proof of Theorem 1.2. Let $\mathcal{G}$ be a circuit from Theorem 4.1 on k vertices. By Lemma 4.7 we obtain that $\mathcal{F}_{\mathcal{G}}[+_k]$ requires Sherali-Adams refutations of ml-degree $\Omega(k/\log k)$. By Lemma 4.8 it follows that every Sherali-Adams refutation of $\mathcal{F}_{\mathcal{G}}[+_k][+_2]$ requires ml-degree (and hence degree) $\Omega(k/\log k)$ and size $2^{\Omega(k/\log k)}$. On the other hand, Lemma 4.2 combined with Lemma 4.4 shows that $\mathcal{F}_{\mathcal{G}}[+_k][+_2]$ has a polynomial calculus refutation of degree 3 and size $O(k^4)$. Since $\mathcal{F}_{\mathcal{G}}[+_k][+_2]$ has $n = 2k^2$ variables, the theorem follows. □

5 Conclusions

We compared the static semi-algebraic proof systems Sherali-Adams and sum-of-squares with polynomial calculus, a dynamic algebraic proof system. The main results show that sum-of-squares simulates polynomial calculus (Theorem 1.1), while Sherali-Adams is not able to do so (Theorem 1.2). The relations between the proof systems considered in this paper are described in Figure 1.

One open question concerns the separation between polynomial calculus and Sherali-Adams. Note that the pebbling contradiction $\mathcal{F}_{\mathcal{G}}$ that separates polynomial calculus degree from Nullstellensatz degree is a system of polynomial equations that encodes a CNF formula. This is no longer the case for the substituted formula $\mathcal{F}_{\mathcal{G}}[+_k][+_2]$ that separates polynomial calculus from Sherali-Adams, and encoding $\mathcal{F}_{\mathcal{G}}[+_k][+_2]$ as a CNF blows up its size exponentially. It would therefore be nice to know whether there is a separating CNF. Note that such a CNF would have to be hard for resolution as well, which is not the case for the substituted variants of the pebbling contradictions (that are in conjunctive normal form) considered in the literature (see [16]).

References

[1] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Space complexity in propositional calculus. *SIAM J. Comput.*, 31(4):1184–1211, 2002. URL: `https://doi.org/10.1137/S0097539700366735`, `doi:10.1137/S0097539700366735`.

[2] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower bounds on hilbert's nullstellensatz and propositional proofs. *Proceedings of the London Mathematical Society*, s3-73(1):1–26, 1996. URL: `http://dx.doi.org/10.1112/plms/s3-73.1.1`, `doi:10.1112/plms/s3-73.1.1`.

[3] W. Dale Brownawell. Bounds for the degrees in the nullstellensatz. *Annals of Mathematics*, 126(3):577–591, 1987. URL: http://www.jstor.org/stable/1971361.

[4] Josh Buresh-Oppenheim, Matthew Clegg, Russell Impagliazzo, and Toniann Pitassi. Homogenization and the polynomial calculus. *Computational Complexity*, 11(3-4):91–108, 2002. URL: https://doi.org/10.1007/s00037-002-0171-6, doi:10.1007/s00037-002-0171-6.

[5] M. Clegg, J. Edmonds, and R. Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th annual ACM symposium on Theory of computing*, pages 174–183, 1996.

[6] Stephen A. Cook and Robert Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, March 1979.

[7] Stefan S. Dantchev, Barnaby Martin, and Mark Nicholas Charles Rhodes. Tight rank lower bounds for the sherali-adams proof system. *Theor. Comput. Sci.*, 410(21-23):2054–2063, 2009. URL: https://doi.org/10.1016/j.tcs.2009.01.002, doi:10.1016/j.tcs.2009.01.002.

[8] John R. Gilbert and Robert Endre Tarjan. Variations of a pebble game on graphs. Technical Report STAN-CS-78-661, Stanford University, 1978. Available at http://infolab.stanford.edu/TR/CS-TR-78-661.html.

[9] Dima Grigoriev, Edward A. Hirsch, and Dmitrii V. Pasechnik. Complexity of semi-algebraic proofs. *Moscow Mathematical Journal*, 2(4):647–679, 2002. URL: http://www.ams.org/distribution/mmj/vol2-4-2002/abst2-4-2002.html.

[10] Dima Grigoriev and Nicolai Vorobjov. Complexity of null- and positivstellensatz proofs. *Annals of Pure and Applied Logic*, 113(1–3):153–160, 2001.

[11] Russell Impagliazzo, Pavel Pudlák, and Jirí Sgall. Lower bounds for the polynomial calculus and the gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999. URL: https://doi.org/10.1007/s000370050024, doi:10.1007/s000370050024.

[12] J. L. Krivine. Anneaux préordonnés. *Journal d'Analyse Mathématique*, 12(1):307–326, Dec 1964. URL: https://doi.org/10.1007/BF02807438, doi:10.1007/BF02807438.

[13] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. doi:10.1137/S1052623400366802.

[14] H. Lombardi, N. Mnev, and M.-F. Roy. The positivstellensatz and small deduction rules for systems of inequalities. *Math. Nachr.*, 181:245–259, 1996.

[15] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991. URL: https://doi.org/10.1137/0801013, doi:10.1137/0801013.

[16] Jakob Nordström. Pebble games, proof complexity, and time-space trade-offs. *Logical Methods in Computer Science*, 9(3), 2013. URL: https://doi.org/10.2168/LMCS-9(3:15)2013, doi:10.2168/LMCS-9(3:15)2013.

[17] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.

[18] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993. URL: http://www.jstor.org/stable/24897130.

[19] 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(3):411–430, 1990.

[20] Gilbert Stengle. A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207(2):87–97, Jun 1974. URL: `https://doi.org/10.1007/BF01362149`, `doi:10.1007/BF01362149`.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il