Tight Size-Degree Bounds for Sums-of-Squares Proofs*

Massimo Lauria
KTH Royal Institute of Technology

Jakob Nordström
KTH Royal Institute of Technology

April 7, 2015

Abstract

We exhibit families of 4-CNF formulas over n variables that have sums-of-squares (SOS) proofs of unsatisfiability of degree (a.k.a. rank) d but require SOS proofs of size $n^{\Omega(d)}$ for values of $d = d(n)$ from constant all the way up to n^{δ} for some universal constant δ. This shows that the $n^{\mathrm{O}(d)}$ running time obtained by using the Lasserre semidefinite programming relaxations to find degree-d SOS proofs is optimal up to constant factors in the exponent. We establish this result by combining NP-reductions expressible as low-degree SOS derivations with the idea of relativizing CNF formulas in [Krajíček '04] and [Dantchev and Riis '03], and then applying a restriction argument as in [Atserias, Müller, and Oliva '13] and [Atserias, Lauria, and Nordström '14]. This yields a generic method of amplifying SOS degree lower bounds to size lower bounds, and also generalizes the approach in [ALN14] to obtain size lower bounds for the proof systems resolution, polynomial calculus, and Sherali-Adams from lower bounds on width, degree, and rank, respectively.

1 Introduction

Let $f_1, \ldots, f_s \in \mathbb{R}[x_1, \ldots, x_n]$ be real, multivariate polynomials. Then the Positivstellensatz proven in [Kri64, Ste73] says (as a special case) that the the system of equations

$$f_1 = 0, \ldots, f_s = 0 \tag{1.1}$$

has no solution over $\mathbb{R}^n$ if and only if there exist polynomials $g_j, q_\ell \in \mathbb{R}[x_1, \ldots, x_n]$ such that

$$\sum_{j=1}^{s} g_j f_j = -1 - \sum_{\ell} q_\ell^2 \ . \tag{1.2}$$

That there can exist no solution given an expression of the form (1.2) is clear, but what is more interesting is that there always exists such an expression to certify unsatisfiability. We refer to (1.2) as a *Positivstellensatz proof* or *Sums-of-squares (SOS) proof* of unsatisfiability, or as an *SOS refutation*,[1] of (1.1). We remark that the Positivstellensatz also applies if we add inequalities $h_1 \geq 0, \ldots, h_t \geq 0$ to the system of equations and allow terms $-h_j \sum_\ell q_{j,\ell}^2$ on the right-hand side in (1.2).

The *degree*[2] of an SOS refutation is the maximal degree of any $g_j f_j$. The search for proofs of constant degree d is *automatizable* as shown in a sequence of works by Shor [Sho87], Nesterov [Nes00], Lasserre [Las01], and Parrilo [Par00]. What this means is that if there exists a degree-d SOS refutation for a system of polynomial equalities (and inequalities) over n variables, then such a refutation can be found in polynomial time $n^{\mathrm{O}(d)}$. Briefly, one can view (1.2) as linear system of equations in the

[*] This is the full-length version of the paper with the same title to appear in *Proceedings of the 30th Annual Computational Complexity Conference (CCC '15).*

[1] All proofs for systems of polynomial equations or for formulas in conjunctive normal form (CNF) in this paper will be proofs of unsatisfiability, and we will therefore use the two terms "proof" and "refutation" interchangeably.

[2] This is sometimes also referred to as the "rank," but we will stick to the term "degree" in this paper.

coefficients of g_j and $u = \sum_\ell q_\ell^2$ with the added constraint that u is a sum of squares, and such a system can be solved by semidefinite programming in $d/2$ rounds of the Lasserre SDP hierarchy.

In the last few years there has been renewed interest in sums-of-squares in the context of constraint satisfaction problems (CSPs) and hardness of approximation, as witnessed by, for instance, [BBH$^+$12, OZ13, Tul09]. These works have highlighted the importance of SOS degree upper bounds for CSP approximability, and this is currently a very active area of study.

Our focus in this paper is not on algorithmic questions, however, but more on sums-of-squares viewed as a proof system (also referred to in the literature as *Positivstellensatz* or *Lasserre*). This proof system was introduced by Grigoriev and Vorobjov [GV01] as an extension of the Nullstellensatz proof system studied by Beame et al. [BIK$^+$94], and Grigoriev established SOS degree lower bound for unsatisfiable $\mathbb{F}_2$-linear equations [Gri01b] (also referred to as the 3-XOR problem when each equation involves at most 3 variables) and for the knapsack problem [Gri01a].

Given the connections to semidefinite programming and the Lasserre SDP hierarchy, it is perhaps not surprising that most works on SOS lower bounds have focused on the degree measure. However, from a proof complexity point of view it is also natural to ask about the minimal *size* of SOS proofs, measured as the number of monomials when all polynomials in each term in (1.2) are expanded out as linear combinations of monomials. Such SOS size lower bounds were proven for knapsack in [GHP02] and $\mathbb{F}_2$-linear systems of equations in [KI06],[3] and tree-like size lower bounds for other formulas were also obtained in [PS12].

A wider interest in this area of research was awakened when Schoenebeck [Sch08] (essentially) rediscovered Grigoriev's result [Gri01b], which together with further work by Tulsiani [Tul09] led to integrality gaps for a number of constraint satisfaction problems. There have also been papers such as [BPS07] and [GP14] focusing on *semantic* versions of the proof system, with less attention to the actual syntactic derivation rules used. We refer the reader to, for instance, the introductory section of [OZ13] for more background on sums-of-squares and connections to hardness of approximation, and to the survey [BS14] for an in-depth discussion of SOS as an approximation algorithm and the intriguing connections to the so-called Unique Games Conjecture [Kho02].

1.1 Our Contribution

As discussed above, if a system of polynomial equalities and inqualities over n variables can be shown inconsistent by SOS in degree d, then by using semidefinite programming one can find an SOS refutation of the system in time $n^{\mathrm{O}(d)}$. It is natural to ask whether this is optimal, or whether there might exist "shortcuts" that could lead to SOS refutations more quickly.

We prove that there are no such shortcuts in general, but that the running time obtained by using the Lasserre semidefinite programming relaxations to find SOS proofs is optimal up to the constant in the exponent. We show this by constructing formulas on n variables (which can be translated to systems of polynomial equalities in a canonical way) that have SOS refutations of degree d but require refutations of size $n^{\Omega(d)}$. Our lower bound proof works for d from constant all the way up to n^δ for some constant δ.

Theorem 1.1 (informal). *Let $d = d(n) \leq n^\delta$ where $\delta > 0$ is a universal constant. Then there is a family of 4-CNF formulas $\{F_n\}_{n\in\mathbb{N}^+}$ with $\mathrm{O}(n^2)$ clauses over $\mathrm{O}(n)$ variables such that F_n is refutable in sums-of-squares in degree $\Theta(d)$ but any SOS refutation of F_n requires size $n^{\Omega(d)}$.*

This theorem extends an analogous result joint by the two authors with Atserias in [ALN14] for the proof systems resolution, polynomial calculus, and Sherali-Adams,[4] where upper bounds on refutation size in terms of width, degree, and rank, respectively, were shown to be tight up to the multiplicative constant in the exponent. Theorem 1.1 works for all of these proof systems, since the upper bound is in fact on resolution width (i.e., the size of a largest clause in a resolution refutation), not just SOS degree,

[3] It might be worth pointing out that definitions and terminology in this area have suffered from a certain lack of standardization, and so what [KI06] refers to as "static Lovász-Schrijver calculus" is closer to what we mean by SOS/Lasserre.

[4] The exact details of these proof systems are not important for this discussion, and so we choose not to elaborate further here, instead referring the interested reader to [ALN14].

and in this sense the theorem subsumes the results in [ALN14]. The concrete bound we obtain for the exponent inside the asymptotic notation in the $n^{\Omega(d)}$ size lower bound is very much worse, however, and therefore the gap between upper and lower bounds is very much larger than in [ALN14].

We want to emphasize that the size lower bound in Theorem 1.1 holds for SOS proofs of arbitrary degree. Thus, going to higher degree (i.e., higher levels of the Lasserre SDP hierarchy) does not help, since even arbitrarily large degree cannot yield shorter proofs. This is an interesting parallel to the paper [LRST14] exhibiting problems for which a (symmetric) SDP relaxation of arbitrary degree but bounded size n^d does not do much better than the systematic relaxation of degree d.

1.2 Techniques

We obtain the result in Theorem 1.1 as a special case of a more general method of amplifying lower bounds on width (in resolution), degree (in polynomial calculus) and rank/degree (in Sherali-Adams and Lasserre/SOS) to size lower bounds in the corresponding proof systems. This method is in some sense already implicit in [ALN14], which in turn relies heavily on an earlier paper by Atserias et al. [AMO13], but it turns out that extracting the essential ingredients and making them explicit is helpful for extending the results in [ALN14] to an analogue for sums-of-squares. We give a brief, informal description of the three main ingredients of the method below.

(i) Find a base CNF formulas hard with respect to width/degree/rank To start, we need to find a base problem, encoded as an unsatisfiable CNF formula, that is "moderately hard" for the proof system at hand. What this means is that we should be able to prove asymptotically tight bounds on width if we are dealing with resolution, on degree for polynomial calculus, and on degree/rank for Sherali-Adams and sums-of-squares. It then follows by a generic argument (as discussed briefly above for SOS) that a bound $\mathrm{O}(d)$ on width/degree/rank implies an upper bound $n^{\mathrm{O}(d)}$ on proof size.

In [AMO13, ALN14] the pigeonhole principle served as the base problem. This principle, which has been extensively studied in proof complexity, is encoded in CNF as *pigeonhole principle (PHP) formulas* saying that there is a one-to-one mapping of m pigeons into n pigeonholes for $m > n$. For sums-of-squares we cannot use PHP formulas, however, since they are not hard with respect to SOS degree. Instead we construct an SOS reduction in low degree from inconsistent systems of $\mathbb{F}_2$-linear equations to the clique problem, and then appeal to the result in [Gri01b, Sch08] briefly discussed above to obtain the following degree lower bound.

Theorem 1.2 (informal). *Given $k \in \mathbb{N}^+$, there is a graph G and a 3-CNF formula k-$\mathrm{Clique}(G)$ of size polynomial in k with the following properties:*

1. *The graph G does not contain a k-clique, but the formula k-$\mathrm{Clique}(G)$ claims that it does.*
2. *Resolution can refute k-$\mathrm{Clique}(G)$ in width k.*
3. *Any sums-of-squares refutation of k-$\mathrm{Clique}(G)$ requires degree $\Omega(k)$.*

(ii) Relativize the CNF formulas The second step is to take the formulas for which we have established width/degree/rank lower bounds and *relativize* them. Relativization is an idea that seems to have been considered for the first time in the context of proof complexity by Krajíček [Kra04] and that was further developed by Dantchev and Riis [DR03]. Very loosely, it can be described as follows.

Suppose that we have a CNF formula encoding (the negation of) a combinatorial principle saying that some set S has a property. For instance, the CNF formula could encode the pigeonhole principle discussed above, or could claim the existence of a totally ordered set of n elements where no element in the set is minimal with respect to the ordering (these latter CNF formulas are known as *ordering principle formulas*, *least number principle formulas*, or *graph tautologies* in the literature).

The formula at hand is then relativized by constructing another formula encoding that there is a (potentially much larger) set T containing a subset $S \subseteq T$ for which the same combinatorial principle holds.

For the ordering principle, we can encode that there exists a non-empty ordered subset $S \subseteq T$ of arbitrary size such that it is possible for all elements in S to find a smaller element inside S. This relativization step transforms the previously very easy ordering principle formulas into relativized versions that are exponentially hard for resolution [Dan06, DM14]. For the PHP formulas, we specify that we have a set of $M \gg m$ pigeons mapped into into $n < m$ holes such that there exists a subset of m pigeons that are mapped injectively.

In our setting, it will be important that the relativization does not make the formulas too hard. We do not want the hardness to blow up exponentially and instead would like the upper bound obtained in the first step above to scale nicely with the size of the relativization. For our general approach to work, we therefore need formulas talking about some domain being mapped to some range, where we can enlarge the domain while keeping the range fixed, and where in addition the mapping is symmetric in the sense that permuting the domain does not change the formula.

For this reason, relativizing the ordering principle formulas does not work for our purposes. Pigeonhole principle formulas have this structure, however, which is exactly why the proofs in [ALN14] go through. As already mentioned, PHP formulas will not work for sums-of-squares, but we can relativize the formulas in Theorem 1.2 by saying that there is a large subset of vertices such that there is a k-clique hiding inside such a subset.

(iii) Apply random restrictions to show proof size lower bounds In the final step, we use random restrictions to establish lower bounds on proof size for the relativized CNF formulas obtained in the second step. This part of the proof is relatively standard, except for a crucial twist in the restriction argument introduced in [AMO13].

Assume that there is a small refutation in sums-of-squares (or whatever proof system we are studying) of the relativized formula claiming the existence of a subset of size $m \ll M$ with the given combinatorial property. Now hit the formula (and the refutation) with a random restriction that in effect chooses a subset of size m, and hence gives us back the original, non-relativized formula. This restriction will be fairly aggressive in terms of the number of variables set to fixed truth values, and hence it will hold with high probability that the restricted refutation has no monomials of high degree (or, for resolution, no clauses of high width), since all such monomials will either have been killed by the restriction or at least have shrunk significantly. (We remark that making use of this shrinking in the analysis is the crucial extra feature added in [AMO13].) But this means that we have a refutation of the original formula in degree smaller than the lower bound established in the first step. Hence, no small refutation can exist, and the lower bound on proof size follows.

This concludes the overview of our method to amplify lower bounds on width/degree/rank to size. It is our hope that developing such a systematic approach for deriving this kind of lower bounds, and making explicit what conditions are needed for this approach to work, can also be useful in other contexts.

1.3 Organization of This Paper

The rest of this paper is organized as follows. We start in Section 2 by reviewing the definitions and notation used, and also stating some basic facts that we will need. In Section 3, we prove a degree lower bound for CNF formulas encoding a version of the clique problem. We then present in Section 4 a general method for obtaining SOS size lower bounds from degree lower bounds (or from width, degree, and rank, respectively, for proof systems such as resolution, polynomial calculus, and Sherali-Adams). We conclude with a brief discussion of some possible directions for future research in Section 5.

2 Preliminaries

For a positive integer n, we use the standard notation $[n] = \{1, 2, \dots, n\}$. All logarithms in this paper are to base 2. A CNF formula F is a conjunction of clauses, denoted $F = \bigwedge_j C_j$, where each clause C is a disjunction of literals, denoted $C = \bigvee_i a_i$. Each literal a is either a propositional variable x (a *positive* *literal*) or its negation $\overline{x}$ (a *negative literal*). We think of formulas and clauses as sets, so that there is no repetition and order does not matter. We consider polynomials on the same propositional variables, with the convention that, as an algebraic variable, x evaluates to 1 when it is true and to 0 when it is false. All polynomials in this paper are evaluated on 0/1-assignments, and live in the ring of real multilinear polynomials, which is the ring of real polynomials modulo the ideal generated by polynomials $x_i^2 - x_i$ for all variables x_i. In other words, all variables in all monomials have degree at most one, and monomial multiplication is defined by $\left(\prod_{i \in A} x_i\right) \cdot \left(\prod_{i \in B} x_i\right) = \prod_{i \in A \cup B} x_i$.

Since sums-of-squares derivations operate with polynomial equations and inequalities, in order to reason about CNF formulas we need to encode them in this language. For a clause $C = C^+ \vee C^-$, where we write C^+ and C^- to denote the subsets of positive and negative literals, respectively, we define

$$S(C) = \sum_{x \in C^+} x + \sum_{\overline{x} \in C^-} (1 - x) \tag{2.1}$$

and encode C as the inequality

$$S(C) \geq 1 \ . \tag{2.2}$$

Clearly, a clause C is satisfied by a 0/1-assignment if and only if the same assignment satisfies the inequality $S(C) \geq 1$. For a variable x and a bit $\beta \in \{0, 1\}$, we define

$$\delta_{x=\beta} = \begin{cases} 1 - x & \text{if } \beta = 0, \\ x & \text{if } \beta = 1; \end{cases} \tag{2.3}$$

and for a sequence of variables $\vec{x} = (x_{i_1}, \ldots x_{i_w})$ and a binary string $\beta = (\beta_1, \ldots \beta_w)$, we define the *indicator polynomial*

$$\delta_{\vec{x}=\beta} = \prod_{j=1}^{w} \delta_{x_{i_j}=\beta_j} \tag{2.4}$$

expanded out as a linear combination of monomials. That is, $\delta_{\vec{x}=\beta}$ is the polynomial that evaluates to 1 for 0/1-assignments satisfying the equalities $x_{i_j} = \beta_j$ for $j = 1, \ldots, w$ and to 0 for all other 0/1-assignments. We have the following useful fact.

Fact 2.1. *For every sequence of variables $\vec{x}$ the syntactic equality $\left(\sum_{\beta \in \{0,1\}^w} \delta_{\vec{x}=\beta}\right) = 1$ holds (after cancellation of terms).*

Let F be a CNF formula over some set of variables denoted as $\mathit{Vars}(F)$, and let ρ be a *partial assignment* on $\mathit{Vars}(F)$. We write $F\restriction_\rho$ to denote the formula F *restricted by* ρ, where all clauses $C \in F$ satisfied by ρ are removed and all literals falsified by ρ in other clauses are removed. For a polynomial p over variables $\mathit{Vars}(F)$ (written, as always, as a linear combination of distinct monomials), we let $p\restriction_\rho$ denote the polynomial obtained by substituting values for assigned variables and removing monomials that evaluate to 0. We extend this definition to sets of formulas or polynomials in the obvious way by taking unions.

Definition 2.2 (Sums-of-squares proof system). A *sums-of-squares derivation*, or *SOS derivation* for short, of the polynomial inequality $p \geq 0$ from the system of polynomial constraints

$$f_1 = 0, \ldots, f_s = 0, \ h_1 \geq 0, \ldots, h_t \geq 0 \tag{2.5}$$

is a sum

$$p = \sum_{j=1}^{s} g_j f_j + \sum_{j=1}^{t} u_j h_j + u_0 \ , \tag{2.6}$$

where $g_1, \ldots, g_s$ are arbitrary polynomials and each u_j is expressible as a sums of squares $\sum_\ell q_{j,\ell}^2$. A derivation of the equation $p = 0$ is a pair of derivations of $p \geq 0$ and $-p \geq 0$. A *sums-of-squares refutation* of (2.5) is a derivation of the inequality $-1 \geq 0$ from (2.5).

The *degree* of an SOS derivation is the maximum degree among all the polynomials $g_j f_j$, $u_j h_j$, and u_0 in (2.6). The *size* of an SOS derivation is the total number of monomials (counted with repetition) in all polynomials $g_j f_j$, $u_j h_j$, and u_0 (all expanded out as linear combinations of distinct monomials). The size and degree of refuting an unsatisfiable system of polynomial constraints are defined by taking the minimum over all SOS refutations of the system with respect to the corresponding measure.

Remark 2.3. Readers more familiar with the usual definition of Positivstellensatz/sums-of-squares in the literature might be a bit puzzled by the use of multilinearity in Definition 2.2, and might also wonder where the axioms $x_i^2 - x_i = 0$, $x_i \geq 0$, and $1 - x_i \geq 0$ for every variable x_i disappeared. It is important to note that we have these axioms in our multilinear setting as well, although they are not explicitly mentioned. Equations of the form $x_i^2 - x_i = 0$ are tautological due to multilinearity, and the inequalities $x_i \geq 0$ and $1 - x_i \geq 0$ are derivable by the squaring rule since in the multilinear setting we have $x_i = x_i^2$ and $1 - x_i = (1 - x_i)^2$.

Our choice of the multilinear setting is without any loss of generality and only serves to simplify the technical arguments slightly. It is easy to see that applying the multilinearization operator mapping x_i^ℓ to x_i for every $\ell \geq 1$ to any SOS derivation over real polynomials yields a legal SOS derivation over multilinear real polynomials in at most the same size and degree. Thus, working in the multilinear setting can only make our lower bounds stronger. As to the upper bounds in this paper, we prove them in the resolution proof system discussed below, and the simulation of resolution by sums-of-squares in Lemma 2.6 below works also in the standard setting without multilinearization.

Let us state some useful basic properties of multilinear polynomials for later reference (and also provide a proof just for completeness).

Proposition 2.4 (Unique multilinear representation). *Every function $f : \{0,1\}^n \to \mathbb{R}$ has a unique representation as a multilinear polynomial. In particular, if p is a multilinear polynomial such that $p(\alpha) \in \{0,1\}$ for all $\alpha \in \{0,1\}^n$, then for every positive integer ℓ the equality $p^\ell = p$ holds (where this is a syntactic equality of multlinear polynomials expanded out as linear combinations of distinct monomials).*

Proof. The set of functions from $\{0,1\}^n$ to $\mathbb{R}$ is a vector space of dimension 2^n. Any function $f(\vec{x})$ in this space can be represented as a linear combination $\sum_{\beta \in \{0,1\}^n} f(\beta) \cdot \delta_{\vec{x}=\beta}(\vec{x})$. Since each $\delta_{\vec{x}=\beta}$ is a multilinear polynomial the multilinear monomials on n variables are a set of 2^n generators of the vector space. By linear independence they also form a basis, and hence the representation of a function as a linear combination of multilinear monomials is unique. The second part of the proposition now follows immediately since p^ℓ and p compute the same function. □

The upper bounds in this paper are shown in the weaker proof system *resolution*, which is defined as follows. A *resolution derivation* of a clause D from a CNF formula F is a sequence of clauses $(D_1, D_2, \ldots, D_\tau)$ such that $D_\tau = D$ and for every clause D_i it holds that it is either a clause of F (an *axiom*), or is obtained by *weakening* from some $D_j \subseteq D_i$ for $j < i$, or can be inferred from two clauses D_ℓ, D_j, $\ell < j < i$, by the *resolution rule* that allows to derive the clause $A \vee B$ from two clauses $A \vee x$ and $B \vee \overline{x}$ (where we say that $A \vee x$ and $B \vee \overline{x}$ are *resolved on* x to yield the *resolvent* $A \vee B$). If in a resolution derivation $(D_1, D_2, \ldots, D_\tau)$ each clause D_j is only used once in a weakening or resolution step to derive some D_i for $i > j$, we say that the derivation is *tree-like* (such derivations may contain multiple copies of the same clause). A *resolution refutation* of F, or *resolution proof* for F, is a derivation of the empty clause (the clause containing no literals) from F.

The *width* of a clause is the number of literals in it, and the width of a CNF formula or resolution derivation is the maximal width of any clause in the formula or derivation. The *size* of a resolution derivation is the total number of clauses in it (counted with repetitions). The size and width of refuting an unsatisfiable CNF formula F is defined by taking the minimum over all resolution refutations of F with respect to the corresponding measure.

The following standard fact is easy to establish by forward induction over resolution derivations. We omit the proof.

Fact 2.5. *Consider a partial assignment ρ which assigns ℓ variables. Let A be the unique clause of width ℓ such that A evaluates to false under ρ. If resolution can derive C in width w and size S from $F{\restriction}_\rho$, then resolution can derive $A \lor C$ in width at most $w + \ell$ and size at most $S + 1$ from F.*

Let us also state for the record the formal claim that SOS is more powerful than resolution in term of degree (and for constant degree also in terms of size). The next lemma is essentially Lemma 4.6 in [ALN14], except that there the lemma is stated for the Sherali-Adams proof system. Since SOS simulates Sherali-Adams efficiently with respect to both size and degree, however, the same bounds apply also for SOS. Referring to the discussion in Remark 2.3, it should also be pointed out that the lemma in [ALN14] is proven in the more common non-multilinear setting with explicit axioms $x_i^2 - x_i = 0$, $x_i \geq 0$, and $1 - x_i \geq 0$ for all variables x_i.

Lemma 2.6 (SOS simulation of resolution). *If a CNF formula $F = \bigwedge_{j=1}^{t} C_j$ has a resolution refutation of size S and width w, then the constraints $\{S(C_j) \geq 1\}_{j=1}^{t}$ as defined in (2.1) and (2.2) have an SOS refutation of size $\mathrm{O}\bigl(w2^w S\bigr)$ and degree at most $w + 1$.*

The next lemma will be useful as a subroutine when we prove upper bounds in resolution.

Lemma 2.7. *Let k and $m_1, m_2, \ldots m_k$ be positive numbers. Then the CNF formula consisting of the clauses*

$$y_{i,0} \qquad i \in [k], \tag{2.7a}$$

$$\overline{y}_{i,j-1} \lor x_{i,j} \lor y_{i,j} \qquad i \in [k],\ j \in [m_i], \tag{2.7b}$$

$$\overline{y}_{i,m_i} \qquad i \in [k], \tag{2.7c}$$

$$\overline{x}_{1,j_1} \lor \overline{x}_{2,j_2} \cdots \lor \overline{x}_{k,j_k} \qquad (j_1, \ldots, j_k) \in [m_1] \times \cdots \times [m_k], \tag{2.7d}$$

has a resolution refutation of width $k + 1$ and size $\mathrm{O}\bigl(\prod_{i=1}^{k} m_i\bigr)$.

Proof. We prove the lemma by backwards induction over k. Consider any clause A of the form

$$A = \overline{x}_{1,j_1} \lor \overline{x}_{2,j_2} \cdots \lor \overline{x}_{(i-1),j_{(i-1)}} \tag{2.8}$$

for $1 \leq i \leq k$ (and note that for $i = 1$ this is the empty clause). We will show how to derive A in width $i + 1$ given clauses $A \lor \overline{x}_{i,1}$, $A \lor \overline{x}_{i,2}$, $\ldots$, $A \lor \overline{x}_{i,m_i}$.

We start by resolving the axioms $y_{i,0}$ and $\overline{y}_{i,0} \lor x_{i,1} \lor y_{i,1}$, and then we apply the resolution rule again on this resolvent and the clause $A \lor \overline{x}_{i,1}$ (available by the induction hypothesis) to get $A \lor y_{i,1}$. We now deduce $A \lor y_{i,j}$ for increasing j. Suppose we have already obtained $A \lor y_{i,j-1}$. Using the inductively derived clause $A \lor \overline{x}_{i,j}$ and the axiom $\overline{y}_{i,j-1} \lor x_{i,j} \lor y_{i,j}$, we can resolve on variables $y_{i,j-1}$ and $x_{i,j}$ to obtain $A \lor y_{i,j}$. Once $A \lor y_{i,m_i}$ has been derived, we resolve it with the axiom $\overline{y}_{i,m_i}$ to get A. By backward induction we reach the empty clause for $i = 1$, which concludes the resolution refutation. Since $i \leq k$, the refutation has width $k + 1$. It is easy to verify that all axioms and intermediate clauses in the refutation are used exactly once. Thus, the refutation is tree-like, and has size exactly twice the number of axioms clauses minus one, which, in particular, is $\mathrm{O}\bigl(\prod_{i=1}^{k} m_i\bigr)$. □

When we construct formulas to be relativized as described in Section 1.2, it is convenient to use variables $x_{i,\vec{\jmath}}$, where i ranges over some specific domain D and $\vec{\jmath}$ is a collection of other indices. We say that the variable $x_{i,\vec{\jmath}}$ *mentions* the element $i \in D$. The *domain-width* of a clause is the number of distinct elements of D mentioned by its variables. The domain-width of a CNF formula or resolution proof is defined by taking the maximum domain-width over all its clauses, and the domain-width of refuting a CNF formula F is the minimal domain-width of any resolution refutation of F. Similarly, the *domain-degree* of a monomial is the number of distinct elements in D mentioned by its variables, the domain-degree of a polynomial or SOS proof is the maximal domain-degree of any monomial in it, and the domain-degree of refuting an unsatisfiable system of polynomial constraints is defined by taking the minimum over all refutations.

3 A Degree Lower Bound for Clique Formulas

In this section we state and prove the formal version of Theorem 1.2, namely a lower bound for the domain-degree needed in SOS to prove that a graph G has no k-clique. Let us start by describing how we encode the k-clique problem as a CNF formula.

Definition 3.1 (k-clique formula). Let k be a positive integer, $G = (V, E)$ be an undirected graph on N vertices, and $(v_1, v_2, \ldots, v_N)$ be an enumeration of $V(G) = V$. Then the formula k-Clique(G) consists of the clauses

$$\overline{x}_{i,u} \vee \overline{x}_{i',v} \qquad i, i' \in [k], i \neq i', \{u, v\} \notin E(G), \tag{3.1a}$$

$$\overline{x}_{i,u} \vee \overline{x}_{i,v}, \qquad i \in [k], u, v \in V(G), u \neq v, \tag{3.1b}$$

$$z_{i,0} \qquad i \in [k], \tag{3.1c}$$

$$\overline{z}_{i,(j-1)} \vee x_{i,v_j} \vee z_{i,j} \qquad i \in [k], j \in [N], \tag{3.1d}$$

$$\overline{z}_{i,N} \qquad i \in [k]. \tag{3.1e}$$

The formula k-Clique(G) encodes the claim that G has a clique of size k. The intended meaning of the variable $x_{i,v}$ for $v \in V(G)$ is that v is the ith vertex of the clique. The clauses in (3.1a) enforce that any two members of the clique are distinct and are connected by an edge. The clauses in (3.1b) enforce that at most one vertex is chosen for each $i \in [k]$. The clauses in (3.1c)–(3.1e) are simply the 3-CNF encoding (using extension variables) of the clause $\bigvee_{j=1}^{N} x_{i,v_j}$ enforcing that at least one vertex is chosen for each $i \in [k]$.

The variables of k-Clique(G) are indexed by i over the domain $[k]$ and the domain-width of the formula is 2. The next proposition shows that the naive brute-force approach to decide k-Clique(G) can be carried on in resolution (and hence by Lemma 2.6 also in SOS).

Proposition 3.2. *If G has no clique of size k, then k-Clique(G) has a resolution refutation of size $\mathrm{O}(|V|^k)$ and width $k+1$.*

Proof. We first use the weakening rule to derive all clauses of the form

$$\overline{x}_{1,u_1} \vee \overline{x}_{2,u_2} \vee \cdots \vee \overline{x}_{k,u_k} \tag{3.2}$$

for every sequence of vertices $(u_1, u_2, \ldots, u_k)$. This is possible since either the sequence contains a repetition or it includes two vertices with no edge between them, and in both cases this means that the clause (3.2) is a superclause of some clause of the form (3.1a). Then we derive the empty clause by applying Lemma 2.7 to the clauses (3.1c)–(3.1e) and (3.2). □

In order to obtain suitably hard instances of k-Clique(G) we construct a reduction from 3-XORs to k-partite graphs. It is convenient for us to describe the special case of k-clique on k-partite graphs directly as an encoding as polynomial equations and inequalities as follows next.

Definition 3.3 (Polynomial encoding of k-clique on k-partite graphs). For a k-partite graph G with $V(G) = V_1 \dot\cup V_2 \dot\cup \cdots \dot\cup V_k$ we let k-Block(G) denotes the following collection of polynomial constraints:

$$\sum_{v \in V_i} x_v = 1 \qquad i \in [k], \tag{3.3a}$$

$$x_u + x_v \leq 1 \qquad u \in V_i, v \in V_{i'}, i \neq i', \{u, v\} \notin E(G). \tag{3.3b}$$

It is straightforward to verify that these constrants encode the claim that G has a clique with one element in each block V_i, since exactly one element is chosen from each block by (3.3a) and all the chosen elements have to be pairwise connected by (3.3b).

Any lower bound on degree that we establish for k-Block(G) will hold also for k-Clique(G) as stated in the following proposition.

Proposition 3.4. *Consider a k-partite graph G, where $V(G) = V_1 \,\dot\cup\, V_2 \,\dot\cup \cdots \dot\cup\, V_k$. If k-*Clique(G) *has an SOS refutation in domain-degree d, then k-*Block(G) *has an SOS refutation in domain-degree d.*

Proof. The proof is by transforming a refutation of k-Clique(G) into a refutation of k-Block(G) of the same domain-degree. To give an overview, we start with a refutation of k-Clique(G) of domain-degree d and replace its variables with polynomials of degree at most 1 mentioning only variables from k-Block(G). In this way we get an SOS refutation of domain-degree at most d from the substituted axioms of k-Clique(G). The latter polynomials are not necessarily axioms of k-Block(G), but we show that they have SOS derivations of domain-degree 1 from the axioms of k-Block(G). This concludes the proof.

The variable substitution has two steps: first we substitute every variable $z_{i,j}$ with the linear form $\sum_{t=j+1}^{N} x_{i,v_t}$, where $\{v_j\}_{j=1}^{N}$ is the enumeration of $V(G)$ in Definition 3.1, and then we set x_{i,v_j} to 0 whenever $v_j \notin V_i$.

As mentioned above, we now need to give SOS derivations of domain-degree 1 of all transformed axioms in k-Clique(G) from k-Block(G). For the axioms (3.1c)–(3.1e), the SOS encoding is

$$z_{i,0} \geq 1 \qquad i \in [k], \tag{3.4a}$$

$$\left(1 - z_{i,(j-1)}\right) + x_{i,v_j} + z_{i,j} \geq 1 \qquad i \in [k], j \in [N], \tag{3.4b}$$

$$\left(1 - z_{i,N}\right) \geq 1 \qquad i \in [k]. \tag{3.4c}$$

After the first step of the substitution the inequalities (3.4a), (3.4b) and (3.4c) become, respectively, the inequality $\sum_{j=1}^{N} x_{i,v_j} \geq 1$, and two occurrences of tautology $1 \geq 1$. Furthermore, after the second step of the substitution the inequality (3.4a) becomes $\sum_{v \in V_i} x_{i,v} \geq 1$, which is subsumed by Equation (3.3a). Each of the axioms (3.1a) and (3.1b) is encoded as

$$1 - x_{i,u} - x_{i',v} \geq 0 \tag{3.5}$$

for some pair of indices i, i' and vertices u, v. We assume that $u \in V_i$ and $v \in V_{i'}$, because otherwise the variable substitution turns the inequality into either a tautology or into $1 - x_{i,u} \geq 0$, where the latter follows from $(1 - x_{i,u})^2 \geq 0$ by multilinearity. If $i \neq i'$ then the inequality (3.5) is an axiom of k-Block(G). If that is not the case, then we can obtain $1 - x_{i,u} - x_{i,v}$ in domain-degree 1 using the derivation

$$\underbrace{1 - \sum_{v \in V_i} x_{i,w}}_{\text{from Equation (3.3a)}} + \underbrace{\sum_{w \notin \{u,v\}} (x_{i,w})^2}_{\text{sum of squares}} = 1 - \sum_{v \in V_i} x_{i,w} + \sum_{w \notin \{u,v\}} x_{i,w} = 1 - x_{i,u} - x_{i,v} \tag{3.6}$$

where the first identity holds by multilinearity. The proposition follows. □

What we want to do now is to prove a domain-degree lower bound for instances of k-Block(G) where the graph G is obtained by a reduction from (unsatisfiable) sets of $\mathbb{F}_2$-linear equations. We rely on the version of Grigoriev's degree lower bound [Gri01b] shown by Schoenebeck [Sch08], which is conveniently stated for random 3-XOR formulas as encoded next.

Definition 3.5 (Polynomial encoding of random 3-XOR). A random 3-XOR formula ϕ represents a system of Δn linear equations modulo 2 defined over n variables. Each equation is sampled at random among all equations of the form $x \oplus y \oplus z = b$ as follows: x, y, z are sampled uniformly without replacement from the set of n variables and b is sampled uniformly in $\{0, 1\}$. The polynomial encoding of any such linear equation modulo 2 is

$$(1 - x)(1 - y)z = 0 \tag{3.7a}$$

$$(1 - x)y(1 - z) = 0 \tag{3.7b}$$

$$x(1 - y)(1 - z) = 0 \tag{3.7c}$$

$$xyz = 0 \tag{3.7d}$$

when $b = 0$ and

$$(1-x)(1-y)(1-z) = 0 \tag{3.7e}$$
$$xy(1-z) = 0 \tag{3.7f}$$
$$x(1-y)z = 0 \tag{3.7g}$$
$$(1-x)yz = 0 \tag{3.7h}$$

when $b = 1$.

Fixing $\delta = 1/4$ and $\Delta = 8$ in [Sch08] we have the following theorem.

Theorem 3.6 ([Sch08]). *There exists an α, $0 < \alpha < 1$, such that for every $\epsilon > 0$ there exists an $n_\epsilon \in \mathbb{N}$ such that a random 3-XOR formula ϕ in $n \geq n_\epsilon$ variables and $8n$ constraints has the following properties with probability at least $1-\epsilon$.*

1. *At most $6n$ parity constraints of ϕ can be simultaneously satisfied.*
2. *Any sums-of-squares refutation of ϕ requires degree αn.*

Now we are ready to describe how to transform a 3-XOR formula ϕ into a k-partite graph G_ϕ^k that has a clique of size k if and only if ϕ is satisfiable.

Definition 3.7 (3-XOR graph). Given $k \in \mathbb{N}$ and a 3-XOR formula ϕ with $8n$ constraints over n variables, where we assume for simplicity that k divides $8n$, we construct a 3-*XOR graph* G_ϕ^k as follows.

We arbitrarily split the formula ϕ into k linear systems with $8n/k$ constraints each, denoted as $\phi_1, \phi_2, \ldots \phi_k$. For each ϕ_i we let V_i be a set of at most $N \leq 2^{24n/k}$ vertices labelled by all possible assignments to the at most $24n/k$ variables appearing in ϕ_i. For two distinct vertices $u \in V_i$ and $v \in V_{i'}$ there is an edge between u and v in G_ϕ^k if the two assignments corresponding to u and v are compatible, i.e., when they assign the same values to the common variables, and also the union of the two assignments does not violate any constraint in ϕ. (In particular, each V_i is an independent set, since two distinct assignments to the same set of variables are not compatible.)

The key property of the reduction in Definition 3.7 is that it allows small domain-degree refutations of k-Block$\big(G_\phi^k\big)$ to be converted into small degree refutations of ϕ.

Lemma 3.8. *If k-*Block$\big(G_\phi^k\big)$ *has an SOS refutation of domain-degree d, then ϕ has an SOS refutation of degree $24dn/k$.*

Proof. Again we start by giving an overview of the proof, which works by transforming a refutation of k-Block$\big(G_\phi^k\big)$ of domain-degree d into a refutation of ϕ of degree $24dn/k$.

Given a refutation of k-Block$\big(G_\phi^k\big)$ of domain-degree d, we replace every variable x_v with a polynomial over the variables of ϕ. In this way we get an SOS refutation from the polynomials corresponding to the substituted axioms of k-Block$\big(G_\phi^k\big)$. The latter polynomials need not be axioms of ϕ, but we show that they can be efficiently derived in SOS from ϕ. We thus obtain an SOS refutation of ϕ, the degree of which is easily verified to be as in the statement of the lemma.

We now describe the substitution in detail. Consider a block V_i and suppose that the corresponding 3-XOR formula ϕ_i mentions t variables. Let us write $\vec{x}$ to denote this set of variables. Then every vertex $v \in V_i$ represents an assignment $\beta \in \{0,1\}^t$ to $\vec{x}$. In what follows, we denote the indicator polynomial $\delta_{\vec{x}=\beta}$ in (2.4) by δ_v for brevity, and we substitute for each variable x_v the polynomial δ_v of degree $t \leq 24n/k$.

Before the substitution each monomial in the original refutation has domain-degree at most d by assumption. Two important observations are that $(\delta_v)^2 = \delta_v$ for every $v \in V_i$ and that $\delta_u \delta_v = 0$ for every two distinct u, v in the same block V_i. Therefore, after the substitution each monomial is either identically zero or the product of at most d indicator polynomials, and hence its degree is at most $24dn/k$. To verify these observations, note that the identity $(\delta_v)^2 = \delta_v$ holds by Proposition 2.4. The equality

$\delta_u\delta_v = 0$ holds because δ_u and δ_v are the indicator polynomials of two incompatible assignments, and so their product always evaluates to zero. Applying Proposition 2.4 again, we conclude that the (multilinear) polynomial $\delta_u\delta_v$ is identically zero.

In order to complete the proof outline above, we now need to present SOS derivations starting from the 3-XOR constraints of ϕ of all polynomial constraints resulting from the substitutions in the axioms of $k\text{-Block}\big(G^k_\phi\big)$ described above, and to do so in degree at most $24n/k$.

Let us first look at the axioms (3.3a). By Fact 2.1, the identity

$$\sum_{v \in V_i} \delta_v = \sum_{\beta \in \{0,1\}^t} \delta_{\vec{x}=\beta} = 1 \tag{3.8}$$

holds syntactically, so substitutions in axioms of the form (3.3a) result in tautologies $1 = 1$.

The remaining axioms of $k\text{-Block}\big(G^k_\phi\big)$ in (3.3b) have the form $x_u + x_v \leq 1$ for non-edges (u, v) between vertices in different blocks. By construction of G^k_ϕ the reason u and v are not connected is either that the partial assignments corresponding to the two vertices are incompatible, or that their union violates some constraint in ϕ.

In the first case, $1 - \delta_u - \delta_v \geq 0$ is an SOS axiom because of the identity

$$(1 - \delta_u - \delta_v)^2 = 1 - \delta_u - \delta_v \ , \tag{3.9}$$

which follows from the observation that δ_u and δ_v are the indicator polynomials of two incompatible assignments and cannot evaluate to 1 simultaneously, and so $(1 - \delta_u - \delta_v)$ evaluates to either 0 or 1 and is identical to its square by Proposition 2.4. The degree of (3.9) is $24n/k$.

In the second case, the two assignments corresponding to u and v are compatible but their union violates some initial equation $f = 0$ of the form (3.7a)–(3.7h). Any such f is a degree-3 indicator polynomial which evaluates to 1 whenever the assignment satisfies the equations $\delta_u\delta_v = 1$. This means that $\delta_u\delta_v$ contains f as a factor. We factorize f as $f_u f_v$ so that $\delta_u = f_u\delta'_u$ and $\delta_v = f_v\delta'_v$. Given this notation, we can derive $0 \leq 1 - \delta_u - \delta_v$ using the indentity

$$(1 - f_u - f_v)^2 + (f_u - \delta_u)^2 + (f_v - \delta_v)^2 - 2f_u f_v = 1 - \delta_u - \delta_v \tag{3.10}$$

of degree at most $24n/k$. To verify (3.10), observe that the left-hand side is the sum of some squared polynomials and $-2f_u f_v = -2f = 0$. Expanding the squared polynomials and using Proposition 2.4 repeatedly we have that $(f_u)^2 = f_u$, $(f_v)^2 = f_v$, $(\delta_u)^2 = \delta_u$, and $(\delta_v)^2 = \delta_v$, from which we also conclude that

$$f_u\delta_u = f_u\big(f_u\delta'_u\big) = \big(f_u\big)^2\delta'_u = f_u\delta'_u = \delta_u \tag{3.11}$$

and

$$f_v\delta_v = f_v\big(f_v\delta'_v\big) = \big(f_v\big)^2\delta'_v = f_v\delta'_v = \delta_v \tag{3.12}$$

which establishes that (3.10) holds. The lemma follows. □

Now we can put together all the material in this section to prove a formal version of Theorem 1.2 as stated next.

Theorem 3.9. *There are universal constants $N_0 \in \mathbb{N}^+$ and α_0, $0 < \alpha_0 < 1$, such that for every $k \geq 1$ there exists a graph G_k with at most $kN_0 = \mathrm{O}(k)$ vertices and a 3-CNF formula $k\text{-Clique}(G_k)$ of size polynomial in k with the following properties:*

1. *Resolution can refute $k\text{-Clique}(G_k)$ in size $2^{\mathrm{O}(k \log k)}$ and width $k + 1$.*
2. *Any SOS refutation of $k\text{-Clique}(G_k)$ requires domain-degree $\alpha_0 k$.*

Proof. Fix any positive $\epsilon < 1$ and let $N_0 = 2^{24n_\epsilon}$, $\alpha_0 = \frac{\alpha}{24}$ and $n = kn_\epsilon$, where n_ϵ and α are the universal constants from Theorem 3.6. To build the graph G_k we take a 3-XOR formula ϕ on n variables and $8n$ equations from the distribution in Definition 3.5. Since $n \geq n_\epsilon$, Theorem 3.6 implies that there is a formula in the support of the distribution that is unsatisfiable and that requires degree αn to be refuted in SOS. We fix ϕ to be that formula and let G_k be the graph G^k_ϕ constructed as in Definition 3.7. Then G^k_ϕ is k-partite, with each part having at most $2^{24n/k} = N_0$ vertices, and the graph has no k-clique because otherwise ϕ would be satisfiable.

Suppose that there is an SOS refutation of $k\text{-Clique}\bigl(G^k_\phi\bigr)$ of domain-degree d. We want to argue that $d \geq \alpha_0 k$. Since G^k_ϕ is k-partite, by Proposition 3.4 the formula $k\text{-Block}\bigl(G^k_\phi\bigr)$ also has an SOS refutation in domain-degree d. By Lemma 3.8, this in turn yields an SOS refutation of ϕ in degree $24dn/k$. Now Theorem 3.6 implies that $24dn/k \geq \alpha n$, and hence $d \geq \frac{\alpha}{24}k = \alpha_0 k$.

To conclude the proof, we can just observe that the resolution width and size upper bounds are a direct application of Proposition 3.2. □

4 Size Lower Bounds from Relativization

Using the material developed in Section 3, we can now describe how to *relativize* formulas in order to to amplify degree lower bounds to size lower bounds in SOS . This method works for formulas that are "symmetric" in a certain sense, and so we start by explaining exactly what is meant by this.

Definition 4.1 (Symmetric formula). Consider a CNF formula F on variables $x_{i,\vec{\jmath}}$, where i is an index in some domain D and $\vec{\jmath}$ denotes a collection of other indices. For every subset of indices $\vec{\imath} = \{i_1, i_2, \ldots, i_s\} \subseteq D$ we identify the subformula $F_{\vec{\imath}}$ of F such that each clause $C \in F_{\vec{\imath}}$ mentions *exactly* the indices in $\vec{\imath}$, so that a formula F of domain-width d can be written as

$$F = \bigwedge_{s=0}^{d} \bigwedge_{\substack{\vec{\imath}\subseteq D \\ |\vec{\imath}|=s}} F_{\vec{\imath}} \ . \tag{4.1}$$

We say that F is *symmetric with respect to* D if it is invariant with respect to permutations of D, i.e., if for every $F_{\vec{\imath}} \subseteq F$ it also holds that $F_{\pi(\vec{\imath})} \subseteq F$, where π is any permutation on D and $\pi\left(\vec{\imath}\right)$ is the set of images of the indices in $\vec{\imath}$. Phrased differently, F is symmetric with respect to D if for any permutation π on D the *syntactic* equality $F = \bigwedge_{\vec{\imath}\subseteq D} F_{\pi(\vec{\imath})}$ holds (where we recall that we treat CNF formulas as sets of clauses). We apply this terminology for systems of polynomial equations and inequalities in the same way.

Let us illustrate Definition 4.1 by giving perhaps the most canonical example of a formula that is symmetric in this sense.

Example 4.2. Recall that the CNF encoding of the pigeonhole principle with a set of pigeons D and holes $[n]$ claims that there is a mapping from pigeons in D to holes such that no hole gets two pigeons. For every pigeon $i \in D$ there is a clause $\bigvee_{j\in[n]} x_{i,j}$ and for every two distinct pigeons i, i' and hole j there is a clause $\overline{x}_{i,j} \vee \overline{x}_{i',j}$. Since any permutation of the set of pigeons D gives us back exactly the same set of clauses (only listed in a different order) the pigeonhole principle formula is symmetric with respect to D.

By now, the reader will already have guessed that another example of a symmetric formula, which will be more interesting to us in the currect context, is the k-clique formula discussed in Section 3.

Observation 4.3. *The* $k\text{-Clique}(G)$ *formula in Definition 3.1 over variables* $x_{i,v}$ *is symmetric with respect to the indices* $i \in [k]$.

Starting with any formula F symmetric with respect to a domain D, we can build a family of similar formulas by varying the size of the domain. If F has domain-width d, then for each s, $0 \leq s \leq d$, the subformulas $F_{\vec{\imath}}$ with $|\vec{\imath}| = s$ in (4.1) are the same up to renaming of the domain indices in $\vec{\imath}$. Hence, we can arbitrarily pick one such subformula to represent them all, and denote it as F_s. The formulas $\{F_s\}_{s=0}^{d}$ are completely determined by F, and together with D they in turn completely determine F. Using this observation, we can generalize the formula F over domain D to any domain D' with $|D'| \geq d$ by defining $F[D']$ to be the formula

$$F[D'] = \bigwedge_{s=0}^{d} \bigwedge_{\substack{\vec{\imath} \subseteq D \\ |\vec{\imath}|=s}} F_{\vec{\imath}} \quad , \tag{4.2}$$

where each $F_{\vec{\imath}}$ for $|\vec{\imath}| = s$ is an isomorphic copy of F_s with its domain indices renamed according to $\vec{\imath}$. Let us state some simple but useful facts that can be read off directly from (4.2):

1. For any formula F of domain-width d symmetric with respect to domain D, it holds that $F[D]$ is (syntactically) equal to F.
2. For any domains D', D'' with $|D'| = |D''| \geq d$, the two formulas $F[D']$ and $F[D'']$ are isomorphic.
3. For any $D'' \supsetneq D'$ with $|D'| \geq d$, the formula $F[D'']$ contains many isomorphic copies of $F[D']$.

When we want to emphasize the domain D of a formula F in what follows, we will denote the formula F as $F[D]$. When the domain is $D = [t]$, we abuse notation slightly and write $F[t]$ instead of $F[[t]]$. As discussed above, from a symmetric formula F of domain-width d we can obtain a well-defined sequence of formulas $F[t]$ for all $t \geq d$. We say that the *unsatisfiability threshold* of such a sequence of formulas is the least t such that $F[t]$ is unsatisfiable. For instance, the pigeonhole principle formula in Example 4.2 has unsatisfiability threshold $n + 1$.

4.1 Relativization of Symmetric Formulas

Given a formula $F = F[m]$ symmetric with respect to $[m]$ and a parameter $k < m$, we now want to define the k*-relativization* of $F[m]$, which is intended to encode the claim that that there exists a subset $D \subseteq [m]$ of size $|D| \geq k$ such that the subformula $F[D] \subseteq F[m]$ is satisfiable. We remark that a CNF formula encoding such a claim will be unsatisfiable when k is at least the unsatisfiability threshold of F.

In order to express the existence of the subset D we use *selectors* $s_1, s_2, \ldots, s_m$ as indicators of membership in the subset and encode the constraint on the subset size $|D| = \sum_{i=1}^{m} s_i \geq k$ as described in the next definition.

Definition 4.4. The *threshold-k formula* for variables $\vec{s} = \{s_1, \ldots, s_m\}$ is the 3-CNF formula $\mathsf{Thr}_k(\vec{s})$ that consists of the clauses

$$y_{\ell,0} \qquad \ell \in [k], \tag{4.3a}$$

$$\overline{y}_{\ell,i-1} \vee p_{\ell,i} \vee y_{\ell,i} \qquad \ell \in [k], i \in [m], \tag{4.3b}$$

$$\overline{y}_{\ell,m} \qquad i \in [m], \tag{4.3c}$$

$$\overline{p}_{\ell,i} \vee \overline{p}_{\ell',i} \qquad \ell, \ell' \in [k], \ell \neq \ell', i \in [m], \tag{4.3d}$$

$$\overline{p}_{\ell,i} \vee s_i \qquad \ell \in [k], i \in [m] \quad . \tag{4.3e}$$

To see that $\mathsf{Thr}_k(\vec{s})$ indeed enforces a cardinality constraint, note that the variables $p_{\ell,i}$ encode a mapping between $[k]$ and $[m]$ (with $p_{\ell,i}$ being true if and only if ℓ maps to i). The clauses (4.3a)–(4.3c) force every $\ell \in [k]$ to have an image in $[m]$, since they form the 3-CNF representation of clauses $\bigvee_i p_{\ell,i}$. The clauses (4.3d) forbid two distinct elements of $[k]$ to have the same image, so there must be at least k elements in the range of the map, and for each of them the corresponding selector must be true because of the clauses (4.3e). We will need the following properties of the threshold formula.

Observation 4.5. *The formula* $\mathsf{Thr}_k(\vec{s})$ *in Definition 4.4 has the following properties:*

1. $\mathsf{Thr}_k(\vec{s})$ *has size polynomial in both* k *and* m*.*
2. *For any partial assignment to* $\vec{s}$ *with at least* k *ones there is an assignment to the extension variables that satisfies* $\mathsf{Thr}_k(\vec{s})$*.*
3. *There is a resolution refutation of the set of clauses* $\mathsf{Thr}_k(\vec{s}) \cup \left\{\bigvee_{i\in D} \overline{s}_i \middle| D \subseteq [m],\ |D| = k\right\}$ *of size* $\mathrm{O}\left(km^k\right)$ *and width* $k+1$*.*

Proof. The first two items are immediate. In order to show the third item we can first derive each clause $\overline{p}_{1,i_1} \vee \ldots \vee \overline{p}_{k,i_k}$ by resolving $\overline{s}_{i_1} \vee \ldots \vee \overline{s}_{i_k}$ with clauses of the form (4.3e), and then apply Lemma 2.7. □

Using the formula in Definition 4.4 to encode cardinality constraints on subsets, we can now define formally what we mean by the relativization of a symmetric formula.

Definition 4.6 (Relativization). Given a CNF formula F symmetric with respect to a domain $[m]$ and a parameter $k < m$, the *k-relativization* (or *k-relativized formula*) $F[k;m]$ is the formula consisting of

1. the threshold formula $\mathsf{Thr}_k(\vec{s})$ over selectors $\vec{s} = \{s_1, \ldots, s_m\}$;
2. a *selectable clause* $\overline{s}_{i_1} \vee \ldots \vee \overline{s}_{i_s} \vee C$ for each clause $C \in F[m]$, where $\{i_1, i_2, \ldots, i_s\}$ are the indices mentioned by C.

Since we are dealing with refutations of unsatisfiable formulas, it will always be the case that the parameter k in Definition 4.6 is at least the unsatisfiability threshold of F. An important property of relativized formulas is that the hardness of $F[k;m]$ scales nicely with m. In particular, if $F[k]$ is not too hard, then the relativization $F[k;m]$ also is not too hard.

Proposition 4.7. *If* $F[k]$ *has a resolution refutation of size* S *and width* w*, then* $F[k;m]$ *has a resolution refutation of size* $S \cdot \binom{m}{k} + \mathrm{O}\left(km^k\right)$ *and width* $w + k$*.*

Proof. For every set $D \subseteq [m]$ with $|D| = k$ we show how to derive

$$\bigvee_{i\in D} \overline{s}_i \tag{4.4}$$

in size $S+1$ and width $w+k$ from $F[k;m]$. Without loss of generality (because of symmetry) we assume that $D = [k]$, so that we want to derive $\overline{s}_1 \vee \cdots \vee \overline{s}_k$. Consider the assignment $\rho = \{s_1 = 1, \ldots, s_k = 1\}$. In the restricted formula $F[k;m]{\restriction_\rho}$ the selectable clauses in Definition 4.6, item 2, with all indices in $[k]$ become the clauses of $F[k]$, which has a refutation of size S and width w. Thus the clause $\overline{s}_1 \vee \cdots \vee \overline{s}_k$ can be derived in size $S+1$ and width $w+k$ from $F[k;m]$ by Fact 2.5. After we have derived all clauses of the form (4.4) in this way, we can obtain the empty clause in width $k+1$ and in size at most $\mathrm{O}\left(km^k\right)$ using Observation 4.5. □

4.2 Random Restrictions and Size Lower Bounds

To prove size lower bounds on refutations of relativized formulas $F[k;m]$ we use random restrictions sampled as follows.

Definition 4.8 (Random restrictions for relativized formulas). Given a relativized formula $F[k;m]$, we define a distribution $\mathcal{R}$ of partial assignments over the variables of this formula by the following process.

1. Pick uniformly at random a set $D \subseteq [m]$ of size k.
2. Fix s_i to 1 if $i \in D$ and to 0 otherwise.

3. Extend this to any assignment to the remaining variables of the formula $\mathsf{Thr}_k(\vec{s})$ that satisfies this threshold formula.

4. For every variable $x_{i,\vec{\jmath}}$ that has index $i \notin D$, fix $x_{i,\vec{\jmath}}$ to 0 or 1 uniformly and independently at random.

5. All remaining variables $x_{i,\vec{\jmath}}$ for the indices $i \in D$ are left unset.

It is straightforward to verify that the distribution $\mathcal{R}$ is constructed in such a way as to give us back $F[k]$ from $F[k;m]$.

Observation 4.9. *For any relativized formula $F[k;m]$ and any $\rho \in \mathcal{R}$ it holds that $F[k;m]\restriction_\rho$ is equal to $F[k]$ up to renaming of variables.*

The key technical ingredient in the size lower bound on sums-of-squares proofs is the following property of the distribution $\mathcal{R}$, which was proven in [AMO13, ALN14] but is rephrased below using the notation and terminology in this paper. We also provide a brief proof sketch just to give the reader a sense of how the argument goes.

Lemma 4.10 ([AMO13, ALN14]). *Let k, ℓ, m be positive integers such that $m \geq 16$ and $\ell \leq k \leq m/(4\log m)$. Let M be a monomial over the variables of $F[k;m]$ and let ρ be a random restriction sampled from the distribution $\mathcal{R}$ in Definition 4.8. Then the domain-degree of $M\restriction_\rho$ is less than ℓ with probability at least $1 - (4k\log m)^k/m^\ell$.*

Proof sketch. Ley ℓ' be the domain-degree of M. The restriction ρ will set independently and uniformly at random at least $\ell' - k$ of its variables, so if $(\ell' - k)$ is larger than $\ell \log m$, the restricted monomial $M\restriction_\rho$ is non zero with probability at most $1/m^\ell$. Otherwise we upper bound the probability that $M\restriction_\rho$ has domain-degree ℓ with the probability that the ℓ' indices in M contain ℓ of the k surviving indices. By a union bound this probability is at most $(4k\log m)^k/m^\ell$. □

Using Lemma 4.10, it is now straightforward to show that relativization amplifies degree lower bounds to size lower bounds.

Theorem 4.11. *Let k, ℓ, m be positive integers such that $m \geq 16$ and $\ell \leq k \leq m/(4\log m)$. If the CNF formula $F[k]$ requires sums-of-squares refutations of domain-degree ℓ, then the relativized formula $F[k;m]$ requires sums-of-squares refutations of size $m^\ell/(4k\log m)^k$.*

Proof. Suppose that there is a sums-of-squares refutation of $F[k;m]$ in size S, i.e., containing S monomials. For ρ sampled from $\mathcal{R}$, we see that the probability that some monomial in the refutation restricted by ρ has domain-degree at least ℓ is at most

$$S \cdot \frac{(4k\log m)^k}{m^\ell} \tag{4.5}$$

by appealing to Lemma 4.10 and taking a union bound.

As noted in Observation 4.9, the formula $F[k;m]\restriction_\rho$ is equal to $F[k]$ up to renaming of variables, and so it cannot have a refutation of domain-degree ℓ or less. This implies that the bound on the probability (4.5) is greater than one, and thus we obtain

$$S > \frac{m^\ell}{(4k\log m)^k} \ , \tag{4.6}$$

which proves the theorem. □

4.3 Statement of Main Result and Discussion of Possible Improvements

Putting everything together, we can establish the formal version of our main results in Theorem 1.1 as follows.

Theorem 4.12. *Let $k = k(m)$ be any monotone non-decreasing integer-valued function such that $k(m) \leq m/(4\log m)$. Then there is a family of 4-CNF formulas $\{F_{m,k}\}_{m\geq 1}$ with $\mathrm{O}\bigl(km^2\bigr)$ clauses over $\mathrm{O}(km)$ variables such that:*

1. *Resolution can refute $F_{m,k}$ in size $k^{\mathrm{O}(k)}m^k$ and width $2k+1$.*
2. *Any sums-of-squares refutation of $F_{m,k}$ requires size $\Omega\bigl(m^{\alpha_0 k}/(4k\log m)^k\bigr)$, where α_0 is a universal constant.*

Proof. Let G be a graph with properties as in Theorem 3.9 and let $F[k]$ be the CNF formula k-Clique(G) in Definition 3.1. Since $F[k]$ is symmetric, we can relativize it as in Definition 4.6 to obtain $F[k;m]$, which will be our 4-CNF formula $F_{m,k}$. Theorem 3.9 says that $F[k]$ has a resolution refutation of size $k^{\mathrm{O}(k)}$ and width $k+1$, and appealing to Proposition 4.7 we get a resolution refutation of $F_{m,k}$ in size $k^{\mathrm{O}(k)}m^k$ and width $2k+1$. Since we have a domain-degree lower bound of $\alpha_0 k$ for refuting $F[k]$ according to Theorem 3.9, we can use Theorem 4.11 to deduce that the required size to refute $F_{m,k}$ in sums-of-squares is at least $\Omega\bigl(m^{\alpha_0 k}/(4k\log m)^k\bigr)$. The theorem follows. □

We remark that straightforward calculations show that when $k(m) = \mathrm{O}\bigl(m^\delta\bigr)$ for $\delta < \alpha_0$ the upper bound in Theorem 4.12 is $m^{\mathrm{O}(k)}$ and the lower bound is $m^{\Omega(k)}$.

Let us now discuss a couple of the parameters in Theorem 4.12 and how they could be improved slightly. We stated our main theorem for 4-CNF formulas, since that is the clause size that results naturally from our construction. However, if one wants to minimize the clause width and obtain an analogous result for 3-CNF formulas this is also possible to achieve, just as was done in [ALN14] for other proof systems. To prove a version of Theorem 4.12 for 3-CNF formulas we need a simple but rather ad-hoc variation of the relativization argument presented above. Let us briefly describe what modifications are needed.

The way we presented the construction above, we started with the 3-CNF formula k-Clique(G) and then applied relativization, which turned the clauses (3.1c)–(3.1e) into the 4-CNF formula

$$\overline{s}_i \lor z_{i,0} \qquad i \in [k], \tag{4.7a}$$

$$\overline{s}_i \lor \overline{z}_{i,(j-1)} \lor x_{i,v_j} \lor z_{i,j} \qquad i \in [k],\, j \in [N], \tag{4.7b}$$

$$\overline{s}_i \lor \overline{z}_{i,N} \qquad i \in [k]. \tag{4.7c}$$

An alternative approach would be to first encode k-Clique(G) with wide clauses $\bigvee_{j=1}^{N} x_{i,v_j}$ instead of clauses of the form (3.1c)–(3.1e), relativize this new, wide formula, and then convert the relativized formula into 3-CNF using extension variables. Instead of clauses (4.7c)–(4.7c), this would yield the collection of clauses

$$\overline{s}_i \lor z_{i,0} \qquad i \in [k], \tag{4.8a}$$

$$\overline{z}_{i,(j-1)} \lor x_{i,v_j} \lor z_{i,j} \qquad i \in [k],\, j \in [N], \tag{4.8b}$$

$$\overline{z}_{i,N} \qquad i \in [k]. \tag{4.8c}$$

This causes a small technical problem in that some of these clauses mention $i \in [m]$ but lack the literal $\overline{s}_i$, and so a random restriction sampled as in Definition 4.8 may actually falsify these clauses. The solution to this is to change the random assignment so that when $s_i = 0$, we fix each x_{i,v_j} uniformly at random in $\{0, 1\}$, set each $z_{i,(j-1)}$ equal to the value assigned to x_{i,v_j}, and finally fix $z_{i,N}$ to 0. The new restriction satisfies all clauses (4.8a)–(4.8c), and the proof of Lemma 4.10 still goes through.

Another parameter in Theorem 4.12 that could be improved is the value of α_0, which determines how tightly the size lower bound matches the upper bound implied by width/degree and also how high we can push $k(m)$. In our reduction from a 3-XOR formula ϕ to the clique formula k-Clique$\left(G^k_\phi\right)$ we start by splitting the $8n$ constraints into k blocks. The vertices in each block correspond to assignments to $24n/k$ variables, and because of this an SOS refutation in domain-degree d of k-Clique$\left(G^k_\phi\right)$ can be converted to a refutation in degree $24dn/k$ of ϕ.

If we want to obtain a more efficient reduction, we could instead split the n *variables*, rather than the $8n$ constraints, into k parts. In this way each vertex in G^k_ϕ would correspond to an assigment to n/k variables, and an SOS refutation in domain-degree d would translate to a refutation of ϕ in degree dn/k. But now we cannot reduce to the clique problem anymore. Splitting with respect to constraints allows us to enforce pairwise consistency between vertices in different blocks referring to common variables. When splitting with respect to variables, the vertices in different blocks correspond to partial assigments on disjoint domains and so are always pairwise compatible. However, we must still require that these partial assignments are consistent with the constraints in ϕ. Each such constraint refers to up to three blocks. Thus, any satisfying assignment to ϕ corresponds to k vertices such that no triple of vertices violates an 3-XOR constraint. This reduces to the problem of finding a k-hyperclique in a 3-uniform hypergraph. The rest of the reduction can be made to work as in Lemma 3.8. In the end we get an analogous result of that in Theorem 3.9 but with α_0 equal to α instead of $\frac{\alpha}{24}$, which also improves Theorem 4.12. In this paper we instead presented a reduction to the k-clique problem for standard graphs, partly because we believe that a degree lower bound for this problem can be considered to be of independent interest.

5 Concluding Remarks

In this paper, we show that using Lasserre semidefinite programming relaxations to find degree-d sums-of-squares proofs is optimal up to constant factors in the exponent of the running time. More precisely, we show that there are constant-width CNF formulas on n variables that are refutable in sums-of-squares in degree d but require proofs of size $n^{\Omega(d)}$.

As for so many other results for the sums-of-squares proof system, in the end our proof boils down to a reduction from 3-XOR using Schoenebeck's version [Sch08] of Grigoriev's degree lower bound [Gri01b]. It would be very interesting to obtain other SOS degree lower bounds by different means than by reducing from Grigoriev's results for 3-XOR and knapsack.

Another interesting problem would be to prove average-case SOS degree lower bound for k-clique formulas over Erdős–Rényi random graphs, or size lower bounds for (non-relativized) k-clique formulas over any graphs. In this context, it might be worth to point out that the problem of establishing proof size lower bounds for k-clique formulas for constant k, which has been discussed, for instance, in [BGLR12], still remains open even for the resolution proof system (although lower bounds have been shown for tree-like resolution in [BGL13] and for full resolution for a version of clique formulas using a different encoding more amenable to lower bound techniques in [LPRT13]).

Acknowledgements

We are grateful to Albert Atserias for numerous discussions about (and explanations of) Lasserre/SOS and other LP and SDP hierarchies, as well as for help with correcting some references in a preliminary version of this manuscript. We thank Per Austrin for valuable suggestions and feedback during the initial stages of this work, and Michael Forbes for comments on an early version of the overall proof construction. Finally, we are thankful for the comments from the anonymous reviewers, which helped improve the exposition in this paper considerably.

The authors were funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) / ERC grant agreement no. 279611. The second author was also supported by Swedish Research Council grants 621-2010-4797 and 621-2012-5645.

References

[ALN14] Albert Atserias, Massimo Lauria, and Jakob Nordström. Narrow proofs may be maximally long. Technical Report TR14-118, Electronic Colloquium on Computational Complexity (ECCC), September 2014. Preliminary version appeared in *CCC '14*.

[AMO13] Albert Atserias, Moritz Müller, and Sergi Oliva. Lower bounds for DNF-refutations of a relativized weak pigeonhole principle. In *Proceedings of the 28th Annual IEEE Conference on Computational Complexity (CCC '13)*, pages 109–120, June 2013.

[BBH$^+$12] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC '12)*, pages 307–326, May 2012.

[BGL13] Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. Parameterized complexity of DPLL search procedures. *ACM Transactions on Computational Logic*, 14(3):20:1–20:21, August 2013. Preliminary version appeared in *SAT '11*.

[BGLR12] Olaf Beyersdorff, Nicola Galesi, Massimo Lauria, and Alexander A. Razborov. Parameterized bounded-depth Frege is not optimal. *ACM Transactions on Computation Theory*, 4:7:1–7:16, September 2012. Preliminary version appeared in *ICALP '11*.

[BIK$^+$94] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower bounds on Hilbert's Nullstellensatz and propositional proofs. In *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science (FOCS '94)*, pages 794–806, November 1994.

[BPS07] Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for Lovász–Schrijver systems and beyond follow from multiparty communication complexity. *SIAM Journal on Computing*, 37(3):845–869, 2007. Preliminary version appeared in *ICALP '05*.

[BS14] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. Technical Report TR14-059, Electronic Colloquium on Computational Complexity (ECCC), April 2014.

[Dan06] Stefan Dantchev. Relativisation provides natural separations for resolution-based proof systems. In *Proceedings of the 1st International Computer Science Symposium in Russia (CSR '06)*, volume 3967 of *Lecture Notes in Computer Science*, pages 147–158. Springer, June 2006.

[DM14] Stefan Dantchev and Barnaby Martin. Relativization makes contradictions harder for resolution. *Annals of Pure and Applied Logic*, 165(3):837–857, March 2014.

[DR03] Stefan Dantchev and Søren Riis. On relativisation and complexity gap for resolution-based proof systems. In *Proceedings of the 17th International Workshop on Computer Science Logic (CSL '03)*, volume 2803 of *Lecture Notes in Computer Science*, pages 142–154. Springer, August 2003.

[GHP02] Dima Grigoriev, Edward A. Hirsch, and Dmitrii V. Pasechnik. Exponential lower bound for static semi-algebraic proofs. In *Proceedings of the 29th International Colloquium on Automata, Languages and Programming (ICALP '02)*, volume 2380 of *Lecture Notes in Computer Science*, pages 257–268. Springer, July 2002.

[GP14] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC '14)*, pages 847–856, May 2014.

[Gri01a] Dima Grigoriev. Complexity of Positivstellensatz proofs for the knapsack. *Computational Complexity*, 10(2):139–154, December 2001.

[Gri01b] Dima Grigoriev. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1–2):613–622, May 2001.

[GV01] Dima Grigoriev and Nicolai Vorobjov. Complexity of Null- and Positivstellensatz proofs. *Annals of Pure and Applied Logic*, 113(1–3):153–160, December 2001.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC '02)*, pages 767–775, May 2002.

[KI06] Arist Kojevnikov and Dmitry Itsykson. Lower bounds of static Lovász–Schrijver calculus proofs for Tseitin tautologies. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP '06)*, volume 4051 of *Lecture Notes in Computer Science*, pages 323–334. Springer, July 2006.

[Kra04] Jan Krajíček. Combinatorics of first order structures and propositional proof systems. *Archive for Mathematical Logic*, 43(4):427–441, May 2004.

[Kri64] Jean-Louis Krivine. Anneaux préordonnés. *Journal d'Analyse Mathématique*, 12(1):307–326, 1964.

[Las01] Jean B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. In *Proceedings of the 8th International Conference on Integer Programming and Combinatorial Optimization (IPCO '01)*, volume 2081 of *Lecture Notes in Computer Science*, pages 293–303. Springer, June 2001.

[LPRT13] Massimo Lauria, Pavel Pudlák, Vojtěch Rödl, and Neil Thapen. The complexity of proving that a graph is Ramsey. In *Proceedings of the 40th International Colloquium on Automata, Languages and Programming (ICALP '13)*, volume 7965 of *Lecture Notes in Computer Science*, pages 684–695. Springer, July 2013.

[LRST14] James R. Lee, Prasad Raghavendra, David Steurer, and Ning Tan. On the power of symmetric LP and SDP relaxations. In *Proceedings of the 29th Annual IEEE Conference on Computational Complexity (CCC '14)*, pages 13–21, June 2014.

[Nes00] Yurii Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, pages 405–440. Kluwer Academic Publisher, 2000.

[OZ13] Ryan O'Donnell and Yuan Zhou. Approximability and proof complexity. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '13)*, pages 1537–1556, January 2013.

[Par00] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, May 2000. Available at http://resolver.caltech.edu/CaltechETD:etd-05062004-055516.

[PS12] Toniann Pitassi and Nathan Segerlind. Exponential lower bounds and integrality gaps for tree-like Lovász–Schrijver procedures. *SIAM Journal on Computing*, 41(1):128–159, 2012. Preliminary version appeared in *SODA '09*.

[Sch08] Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08)*, pages 593–602, October 2008.

[Sho87] N. Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23(5):695–700, 1987. Translated from *Kibernetika*, No. 5, pages 102-–106, 1987.

[Ste73] Gilbert Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207(2):87–97, 1973.

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC '09)*, pages 303–312, June 2009.