Check for
updates

Optimization Over the Boolean Hypercube Via Sums of Nonnegative Circuit Polynomials

Mareike Dressler[1] · Adam Kurpisz[2] · Timo de Wolff[3]

Abstract

Various key problems from theoretical computer science can be expressed as polynomial optimization problems over the boolean hypercube. One particularly successful way to prove complexity bounds for these types of problems is based on sums of squares (SOS) as nonnegativity certificates. In this article, we initiate optimization problems over the boolean hypercube via a recent, alternative certificate called sums of nonnegative circuit polynomials (SONC). We show that key results for SOS-based certificates remain valid: First, for polynomials, which are nonnegative over the n-variate boolean hypercube with constraints of degree d there exists a SONC certificate of degree at most $n + d$. Second, if there exists a degree d SONC certificate for nonnegativity of a polynomial over the boolean hypercube, then there also exists a short degree d SONC certificate that includes at most $n^{O(d)}$ nonnegative circuit polynomials. Moreover, we prove that, in opposite to SOS, the SONC cone is not closed under taking affine transformation of variables and that for SONC there does not exist an equivalent to Putinar's Positivstellensatz for SOS. We discuss these results from both the algebraic and the optimization perspective.

✉ Timo de Wolff
    t.de-wolff@tu-braunschweig.de

    Mareike Dressler
    mdressler@ucsd.edu

    Adam Kurpisz
    adam.kurpisz@ifor.math.ethz.ch

[1] Department of Mathematics, University of California, San Diego, 9500 Gilman Drive 0112, La Jolla, CA 92093-0112, USA

[2] Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

[3] Institut für Analysis und Algebra, AG Algebra, Technische Universität Braunschweig, Universitätsplatz 2, 38106 Braunschweig, Germany

1 Introduction

An *optimization problem over a boolean hypercube* is an n-variate (constrained) polynomial optimization problem where the feasibility set is restricted to some vertices of an n-dimensional hypercube. This class of optimization problems belongs to the core of theoretical computer science. However, it is known that solving them is NP-hard in general, since one can easily cast, e.g., the Independent Set Problem in this framework.

One of the most promising approaches in constructing efficient algorithms is the *sum of squares (SOS) hierarchy* [26,49,52,60], also known as *Lasserre relaxation* [39]. The method is based on a Positivstellensatz result [54] saying that the polynomial f, which is positive over the feasibility set, can be expressed as a sum of squares times the constraints defining the set. Bounding a maximum degree of a polynomial used in a representation of f provides a family of algorithms parametrized by an integer d. Finding a degree d SOS certificate for nonnegativity of f can be performed by solving a *semidefinite programming* (SDP) formulation of size $n^{O(d)}$. Finally, for every (feasible) n-variate unconstrained hypercube optimization problem there exists a degree $2n$ SOS certificate.

The SOS algorithm is a frontier method in algorithm design. It provides the best available algorithms used for a wide variety of optimization problems. The degree 2 SOS for the INDEPENDENT SET problem implies the Lovász θ-function [44] and gives the Goemans–Williamson relaxation [23] for the MAX CUT problem. The Goemans–Linial relaxation for the SPARSEST CUT problem (analyzed in [2]) can be captured by the SOS of degree 6. Finally, the subexponential time algorithm for UNIQUE GAMES [1] is implied by a SOS of sublinear degree [6,28]. Moreover, it has been shown that SOS is equivalent in power to any SDP extended formulation of comparable size in approximating maximum constraint satisfaction problems (CSP) [43]. Recently, SOS has been also applied to problems in dictionary learning [4,59], tensor completion and decomposition [5,30,53], and robust estimation [34]. Other applications of the SOS method can be found in [6,9,14,15,18,20,28,45,46,55], see also the surveys [16,40,42].

On the other hand, it is known that the SOS algorithm admits certain weaknesses. For example, Grigoriev shows in [24] that a $\Omega(n)$ degree SOS certificate is needed to detect a simple integrality argument for the Knapsack Problem, see also [25,36,41]. Other SOS degree lower bounds for Knapsack problems appeared in [13,37]. Some lower bounds on the effectiveness of SOS have been shown for CSP problems [33,61] and for the Planted Clique Problem [3,47]. For a polynomially solvable problem of scheduling unit size jobs on a single machine to minimize the number of late jobs, a degree $\Omega(\sqrt{n})$ SOS, applied on the natural, widely used, formulation of the problem, does *not* provide a relaxation with bounded integrality gap [38]. This excludes an existence of a polynomial time, constant factor approximation algorithm based on the SOS method for this formulation.

Finally, in the case of global nonnegativity, not every nonnegative polynomial is SOS, as first proven by Hilbert [29]. The first explicit example was given by Motzkin [48]; see also [58] for an overview. Even worse, as shown by Blekherman [11], for fixed degree $d \geqslant 4$ asymptotically, i.e., $n \to \infty$, almost every nonnegative polynomial is not a sum of squares.

Moreover, it remains open if finding a degree d SOS certificate can be performed in time $n^{O(d)}$. Indeed, as noted in the recent paper by O'Donnell [50], and further discussed by Raghavendra and Weitz in [56], it is not obviously true that the search can be done so efficiently. Namely, even if a small degree SOS certificate exists, the polynomials in the certificate do not have necessarily small coefficients. O'Donnell in [50] gives an example of polynomial optimization problem that admits a degree 2 SOS certificate. However, every degree 2 SOS certificate for this problem has exponential bit complexity. Next, in [56] the example is modified and cast into a hypercube optimization problem again having a degree 2 SOS certificate, but with super-polynomial bit complexity for certificates up to the degree $O(\sqrt{n})$. This excludes the possibility that known optimization tools used for solving SDP problems like the ellipsoid method [27,32] are able to find a degree d certificates in time $n^{O(d)}$.

The above arguments motivate the search of new nonnegativity certificates for solving optimization problems efficiently.

In this article, we initiate an analysis of hypercube optimization problems via *sums of nonnegative circuit polynomials (SONC)*. SONCs are a nonnegativity certificate introduced in [31], which are independent of sums of squares; see Definition 2.1 and Theorem 2.5 for further details. This means particularly that certain polynomials like the Motzkin polynomial, which have no SOS certificate for global nonnegativity, can be certified as nonnegative via SONCs. Moreover, SONCs generalize polynomials which are certified to be nonnegative via the arithmetic-geometric mean inequality [57]. Similarly as Lasserre's relaxation for SOS, a Schmüdgen-like Positivstellensatz yields a converging hierarchy of lower bounds for polynomial optimization problems with compact constraint set; see [22, Theorem 4.8] and Theorem 2.6. These bounds can be computed via a convex optimization program called *relative entropy programming (REP)* [22, Theorem 5.3]. Our main question in this article is:

> *Can SONC certificates be an alternative for SOS methods for optimization problems over the hypercube?*

We answer this question affirmatively in the sense that we prove SONC complexity bounds for boolean hypercube optimization analogous to the SOS bounds mentioned above. More specifically, we show:

1. For every polynomial which is nonnegative over an n-variate hypercube with constraints of degree at most d, there exists a SONC certificate of nonnegativity of degree at most $n + d$; see Theorem 4.8 and Corollary 4.9.
2. If a polynomial f admits a degree d SONC certificate of nonnegativity over an n-variate hypercube, then the polynomial f admits also a *short* degree d SONC certificate that includes at most $n^{O(d)}$ nonnegative circuit polynomials; see Theorem 4.10.

Furthermore, we show some structural properties of SONCs:

1. We give a simple, constructive example showing that the SONC cone is not closed under multiplication. Subsequently, we use this construction to show that the SONC cone is not closed under taking affine transformations of variables, either; see Lemma 3.1 and Corollary 3.2 and the discussion afterwards.

2. We address an open problem raised in [22] asking whether the Schmüdgen-like Positivstellensatz for SONCs (Theorem 2.6) can be improved to an equivalent of Putinar's Positivstellensatz [54]. We answer this question negatively by showing an explicit hypercube optimization example, which provably does not admit a Putinar representation for SONCs; see Theorem 5.1 and the discussion afterwards.

Our article is organized as follows: In Sect. 2 we introduce the necessary background from theoretical computer sciences and about SONCs. In Sect. 3 we show that the SONC cone is closed neither under multiplication nor under affine transformations. In Sect. 4 we provide our two main results regarding the degree bounds for SONC certificates over the hypercube. In Sect. 5 we prove the non-existence of an equivalent of Putinar's Positivstellensatz for SONCs and discuss this result.

2 Preliminaries

In this section, we collect basic notions and statements on sums of nonnegative circuit polynomials (SONC).

Throughout the paper, we use bold letters for vectors, e.g., $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Let $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ and $\mathbb{R}_{\geqslant 0}$ ($\mathbb{R}_{>0}$) be the set of nonnegative (positive) real numbers. Furthermore, let $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$ be the ring of real n-variate polynomials and the set of all n-variate polynomials of degree less than or equal to $2d$ is denoted by $\mathbb{R}[\mathbf{x}]_{n,2d}$. We use $[n]$ for the set $\{1, \ldots, n\}$, and we denote by $\mathbf{e_1}, \ldots, \mathbf{e_n}$ the canonical basis vectors in $\mathbb{R}^n$.

2.1 Sums of Nonnegative Circuit Polynomials

Let $A \subset \mathbb{N}^n$ be a finite set. In what follows, we consider polynomials $f \in \mathbb{R}[\mathbf{x}]$ supported on A. Thus, f is of the form $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in A} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ with $f_{\boldsymbol{\alpha}} \in \mathbb{R}$ and $\mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. A lattice point is called *even* if it is in $(2\mathbb{N})^n$ and a term $f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ is called a *monomial square* if $f_{\boldsymbol{\alpha}} > 0$ and $\boldsymbol{\alpha}$ is even. We denote by $\mathrm{New}(f) = \mathrm{conv}\{\boldsymbol{\alpha} \in \mathbb{N}^n : f_{\boldsymbol{\alpha}} \neq 0\}$ the Newton polytope of f.

Initially, we introduce the foundation of SONC polynomials, namely *circuit polynomials*; see also [31]:

Definition 2.1 A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is called a *circuit polynomial* if it is of the form

$$f(\mathbf{x}) := \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}, \qquad (2.1)$$

with $r \leqslant n$, exponents $\boldsymbol{\alpha}(j), \boldsymbol{\beta} \in A$, and coefficients $f_{\boldsymbol{\alpha}(j)} \in \mathbb{R}_{>0}$, $f_{\boldsymbol{\beta}} \in \mathbb{R}$, such that the following conditions hold:

(C1) New(f) is a simplex with even vertices $\boldsymbol{\alpha}(0), \boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(r)$.
(C2) The exponent $\boldsymbol{\beta}$ is in the strict interior of New(f). Hence, there exist unique *barycentric coordinates* λ_j relative to the vertices $\boldsymbol{\alpha}(j)$ with $j = 0, \ldots, r$ satisfying

$$\boldsymbol{\beta} = \sum_{j=0}^{r} \lambda_j \boldsymbol{\alpha}(j) \text{ with } \lambda_j > 0 \text{ and } \sum_{j=0}^{r} \lambda_j = 1.$$

We call the terms $f_{\boldsymbol{\alpha}(0)}\mathbf{x}^{\boldsymbol{\alpha}(0)}, \ldots, f_{\boldsymbol{\alpha}(r)}\mathbf{x}^{\boldsymbol{\alpha}(r)}$ the *outer terms* and $f_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ the *inner term* of f.

For every circuit polynomial, we define the corresponding *circuit number* as

$$\Theta_f := \prod_{j=0}^{r} \left(\frac{f_{\boldsymbol{\alpha}(j)}}{\lambda_j} \right)^{\lambda_j}. \tag{2.2}$$

□

Note that the name of these polynomials is motivated by the fact that their support set forms a *circuit*, i.e., a minimal affine dependent set, see, e.g., [51]. The first fundamental statement about circuit polynomials is that its nonnegativity is determined by its circuit number Θ_f and $f_{\boldsymbol{\beta}}$ entirely:

Theorem 2.2 ([31], Theorem 3.8) *Let f be a circuit polynomial with inner term $f_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ and let Θ_f be the corresponding circuit number, as defined in* (2.2). *Then, the following statements are equivalent:*

1. f *is nonnegative.*
2. $|f_{\boldsymbol{\beta}}| \leqslant \Theta_f$ *and* $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ *or* $f_{\boldsymbol{\beta}} \geqslant -\Theta_f$ *and* $\boldsymbol{\beta} \in (2\mathbb{N})^n$.

We illustrate the previous definition and theorem by an example:

Example 2.3 The *Motzkin polynomial* [48] is given by

$$M(x_1, x_2) := x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1.$$

It is a circuit polynomial since New$(f) = \text{conv}\{(4, 2), (2, 4), (0, 0)\}$, and $\boldsymbol{\beta} = (2, 2)$ with $\lambda_0, \lambda_1, \lambda_2 = 1/3$. We have $|f_{\boldsymbol{\beta}}| = 3$ and compute $\Theta_f = \sqrt[3]{\left(\frac{1}{1/3} \right)^3} = 3$. Hence, we can conclude that $M(x_1, x_2)$ is nonnegative by Theorem 2.2.

Definition 2.4 We define for every $n, d \in \mathbb{N}^*$ the set of *sums of nonnegative circuit polynomials* (SONC) in n variables of degree $2d$ as

$$C_{n,2d} := \left\{ f \in \mathbb{R}[\mathbf{x}]_{n,2d} \ : \ f = \sum_{i=1}^{k} \mu_i p_i, \ \begin{array}{c} p_i \text{ is a nonnegative circuit polynomial,} \\ \mu_i \geqslant 0, k \in \mathbb{N}^* \end{array} \right\}$$

□

We denote by SONC both the set of SONC polynomials and the property of a polynomial to be a sum of nonnegative circuit polynomials.

In what follows, let $P_{n,2d}$ be the cone of nonnegative n-variate polynomials of degree at most $2d$ and $\Sigma_{n,2d}$ be the corresponding cone of sums of squares, respectively. An important observation is that SONC polynomials form a convex cone independent of the SOS cone:

Theorem 2.5 ([31], Proposition 7.2) *$C_{n,2d}$ is a convex cone satisfying:*

1. *$C_{n,2d} \subseteq P_{n,2d}$ for all $n, d \in \mathbb{N}^*$,*
2. *$C_{n,2d} \subseteq \Sigma_{n,2d}$ if and only if $(n, 2d) \in \{(1, 2d), (n, 2), (2, 4)\}$,*
3. *$\Sigma_{n,2d} \nsubseteq C_{n,2d}$ for all $(n, 2d)$ with $2d \geqslant 6$.*

For further details about the SONC cone, see [21,22,31].

2.2 SONC Certificates Over a Constrained Set

In [22, Theorem 4.8], Iliman, the first, and the third author showed that for an arbitrary real polynomial which is strictly positive on a compact, basic closed semialgebraic set K there exists a SONC certificate of nonnegativity. We recall this result in what follows.

We assume that K is given by polynomial inequalities $g_i(\mathbf{x}) \geqslant 0$ for $i = 1, \ldots, s$ and is compact. For technical reason we add $2n$ redundant box constraints $l_j^+(\mathbf{x}) := N + x_j \geqslant 0$ and $l_j^-(\mathbf{x}) := N - x_j \geqslant 0$ for $j \in [n]$ for some sufficiently large $N \in \mathbb{N}$, which exists due to our assumption of compactness of K; see [22] for further details. For convenience we use $l_j^\pm(\mathbf{x})$ to refer to both $N + x_j$ and $N - x_j$. Hence, we have

$$K := \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geqslant 0 \text{ for } i \in [s] \text{ and } l_j^\pm(\mathbf{x}) \geqslant 0 \text{ for } j \in [n]\}. \tag{2.3}$$

In what follows, we consider polynomials $H^{(q)}(\mathbf{x})$ defined as products of at most $q \in \mathbb{N}^*$ of the polynomials $g_i, l_j^\pm$ and 1, i.e.,

$$H^{(q)}(\mathbf{x}) := \prod_{k=1}^{q} h_k(\mathbf{x}), \tag{2.4}$$

where $h_k \in \{1, g_1, \ldots, g_s, l_1^+, \ldots, l_n^+, l_1^-, \ldots, l_n^-\}$. Now we can state:

Theorem 2.6 ([22], Theorem 4.8) *Let $f, g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$ be real polynomials and K be a compact, basic closed semialgebraic set as in (2.3). If $f > 0$ on K, then there*

exist $d, q \in \mathbb{N}^*$ such that we have an explicit representation of f of the following form:

$$f(\mathbf{x}) = \sum_{finite} s(\mathbf{x}) H^{(q)}(\mathbf{x}),$$

where the $s(\mathbf{x})$ are contained in $C_{n,2d}$ and every $H^{(q)}(\mathbf{x})$ is a product as in (2.4).

The central object of interest is the smallest value of d and q that allows f a decomposition as in Theorem 2.6. This motivates the following definition of a *degree d SONC certificate*.

Definition 2.7 Let $f \in \mathbb{R}[\mathbf{x}]$ such that f is positive on the set K given in (2.3). Then, f has a *degree d SONC certificate* if it admits for some $q \in \mathbb{N}^*$ the following decomposition:

$$f(\mathbf{x}) = \sum_{finite} s(\mathbf{x}) H^{(q)}(\mathbf{x}),$$

where the $s(\mathbf{x})$ are SONCs, the $H^{(q)}(\mathbf{x})$ are products as in (2.4), and

$$\deg\left(\sum_{finite} s(\mathbf{x}) H^{(q)}(\mathbf{x})\right) \leqslant d.$$

$\square$

Note that for a given set $A \subseteq \mathbb{N}^n$, searching through the space of degree d certificates can be computed via a *relative entropy program (REP)* [22] of size $n^{O(d)}$. REPs are generalizations of geometric programs but still are convex optimization programs [12]. Convex optimization problems are efficiently solvable under relatively mild conditions. For example, using the ellipsoid method, a convex optimization problem is solvable in polynomial time under the assumption of polynomial computability, polynomial growth, and polynomial boundedness of its feasible set; see, e.g., [10, Section 5.3] for more details. The violation of polynomial boundedness is the reason why solving constant degree SOS in polynomial time is an open problem; see our discussion in the introduction and [50,56]. Finally, there exist convex optimization problems, which are known to be not solvable in polynomial time at all; see, e.g., [19].

There are two main bottlenecks that might affect its complexity of finding a degree d SONC certificate. The first one is to guarantee the existence of a sufficiently short degree d SONC certificate. The answer to the equivalent question for the SOS degree d certificates follows from the fact that a polynomial is SOS if and only if the corresponding matrix of coefficients of size $n^{O(d)}$, called the *Gram matrix*, is positive semidefinite. Since every real, symmetric matrix M that is positive semidefinite admits a decomposition $M = V V^\top$, this yields an explicit SOS certificate including at most $n^{O(d)}$ squared polynomials. For more details, we refer the reader to the excellent lecture notes in [8].

The second bottleneck: even if there exists a short degree d SONC certificate, then it is not clear a priori whether there exists an efficient algorithm, i.e., particularly of bit complexity $n^{O(d)}$, to find it in the space of n-variate circuit polynomials of degree d. The number of circuits grows exponentially in n and d.

In this paper, we resolve the first bottleneck regarding the existence of short SONC certificates affirmatively. Namely, we show that one can always restrict oneself to SONC certificates including at most $n^{O(d)}$ nonnegative circuit polynomials, see Sect. 4.2 for details and further discussion on this topic.

3 Properties of the SONC Cone

In this section, we show that the SONC cone is neither closed under multiplication nor under affine transformations. First, we give a constructive proof for the fact that the SONC cone is not closed under multiplication, which is simpler than the initial proof of this fact in [22, Lemma 4.1]. Second, we use our construction to show that the SONC cone is not closed under affine transformation of variables.

Lemma 3.1 *For every $n, d \in \mathbb{N}^*$ with $n \geqslant 2$ the SONC cone $C_{n,2d}$ is not closed under multiplication in the following sense: if $p_1, p_2 \in C_{n,2d}$, then $p_1 \cdot p_2 \notin C_{n,4d}$ in general.*

Proof For every $n, d \in \mathbb{N}^*$ with $n \geqslant 2$ we construct two SONC polynomials p_1, $p_2 \in C_{n,2d}$ such that the product $p_1 p_2$ is an n-variate, degree $4d$ polynomial that is not inside $C_{n,4d}$.

Let $n = 2$. We construct the following two polynomials $p_1, p_2 \in \mathbb{R}[x_1, x_2]$:

$$p_1(x_1, x_2) := (1 - x_1)^2, \qquad p_2(x_1, x_2) := (1 - x_2)^2.$$

First, observe that p_1, p_2 are nonnegative circuit polynomials, since, in both cases, $\lambda_1 = \lambda_2 = 1/2$, $f_{\boldsymbol{\alpha}(1)} = f_{\boldsymbol{\alpha}(2)} = 1$, and $f_{\boldsymbol{\beta}} = -2$, thus $2 = \Theta_f \geqslant |f_{\boldsymbol{\beta}}|$.

Now consider the polynomial $r(x_1, x_2) = p_1 p_2 = ((1 - x_1)(1 - x_2))^2$. We show that this polynomial, even though it is nonnegative, is not a SONC polynomial. Note that $r(x_1, x_2) = 1 - 2x_1 - 2x_2 + 4x_1 x_2 + x_1^2 + x_2^2 - 2x_1^2 x_2 - 2x_1 x_2^2 + x_1^2 x_2^2$; the support of r is shown in Fig. 1. Assume that $r \in C_{2,4}$, i.e., r has a SONC decomposition. This implies that the term $-2x_1$ has to be an inner term of some nonnegative circuit polynomial r_1 in this representation. Such a circuit polynomial necessarily has the terms 1 and x_1^2 as outer terms, that is,

$$r_1(x_1) = p_1(x_1, x_2) = 1 + x_1^2 - 2x_1.$$

Since $\Theta_{r_1} = 2$ the polynomial r_1 is indeed nonnegative and, in addition, we cannot choose a smaller constant term to construct r_1. Next, also the term $-2x_2$ has to be an inner term of some nonnegative circuit polynomial r_2. Since this term again is on the boundary of New(r), the only option for such an r_2 is: $r_2(x_2) = p_2(x_1, x_2) = 1 + x_2^2 - 2x_2$. However, the term 1 has been already used in the above polynomial r_1, which
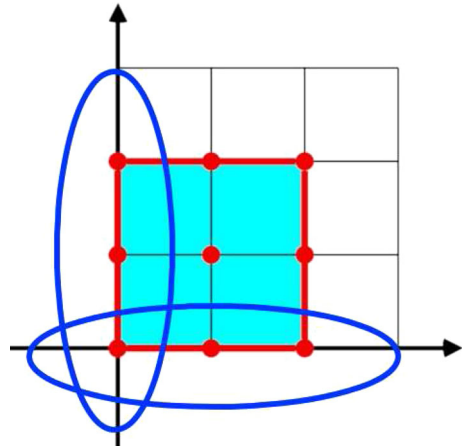
leads to a contradiction, i.e., $r \notin C_{2,4}$. Since $C_{n,4d} \subseteq C_{n+1,4d}$ and $C_{n,4d} \subseteq C_{n,4d+2}$, the general statement follows. $\qquad\square$

In what follows, we show another operation, which behaves differently for SONC than it does for SOS: Similarly as in the case of multiplications, affine transformations also do not preserve the SONC structure. This observation is important for possible degree bounds on SONC certificates, when considering optimization problems over distinct descriptions of the hypercube.

Corollary 3.2 *For every $n, d \geqslant 2$, the SONC cone $C_{n,2d}$ is not closed under affine transformation of variables.*

Proof Consider the polynomial $f(x_1, x_2) = x_1^2 x_2^2$. Clearly, the polynomial f is a nonnegative circuit polynomial since it is a monomial square, hence $f \in C_{n,2d}$ with $n, d \geqslant 2$. Now consider the following affine transformation of the variables x_1 and x_2:

$$x_1 \mapsto 1 - x_1, \qquad x_k \mapsto 1 - x_2.$$

After applying the transformation, the polynomial f equals the polynomial $p_1 p_2$ from the proof of Lemma 3.1 and thus is not inside $C_{n,2d}$. $\qquad\square$

Corollary 3.2, from optimization perspective, implies that problem formulations obtained by applying affine transformations of variables can lead to problems of different tractability when using the SONC method. This means, on the one hand, that a choice of representation has to be done carefully, which makes the process of algorithm design more demanding. On the other hand, even a small change of representation might allow to find a SONC certificate or simplify an existing one. Note that every affine transformation of variables applied to the Motzkin polynomial *never* has a SOS certificate over reals, as the SOS cone is closed under affine transformations. The affine closure of the SONC cone, however, strictly contains the SONC cone and still yields a

certificate of nonnegativity. In this sense, Corollary 3.2 motivates the following future research question:

Find an efficient algorithm to determine whether an affine transformation of a given polynomial f admits a SONC representation.

4 An Upper Bound on the Degree of SONC Certificates Over the Hypercube

In the previous section, we showed that the SONC cone is not closed under taking an affine transformation of variables, Corollary 3.2. Thus, if a polynomial f admits a degree d SONC certificate proving that it is nonnegative on a given compact semi-algebraic set K, then it is a priori not clear whether a polynomial g, obtained from f via an affine transformation of variables, admits a degree d SONC certificate of nonnegativity on K, too. The degree needed to prove nonnegativity of g might be much larger than d according to the argumentation in the proof of Corollary 3.2.

In this section, we prove that every n-variate polynomial which is nonnegative over the boolean hypercube has a degree n SONC certificate. Moreover, if the hypercube is additionally constrained with some polynomials of degree at most d, then the nonnegative polynomial over such a set has degree $n + d$ SONC certificate. We show this fact for all hypercubes $\{a_i, b_i\}^n$; see Theorem 4.3 for further details.

Formally, we consider the following setting: We investigate real multivariate polynomials in $\mathbb{R}[\mathbf{x}]$. For $j \in [n]$, and $a_j, b_j \in \mathbb{R}$, such that $a_j < b_j$ let

$$g_j(\mathbf{x}) := (x_j - a_j)(x_j - b_j)$$

be a *quadratic polynomial with two distinct real roots*. Let $\mathcal{H} \subset \mathbb{R}^n$ denote the *n-dimensional hypercube* given by $\prod_{j=1}^n \{a_j, b_j\}$. Moreover, let

$$\mathcal{P} := \{p_0, p_1, \ldots, p_m \ : \ p_0 = 1 \text{ and } p_i \in \mathbb{R}[\mathbf{x}] \text{ for all } i \in [m]\}$$

be a set of polynomials, which we consider as constraints $p_i(\mathbf{x}) \geqslant 0$ with $\deg(p_i(\mathbf{x})) \leqslant d$ for all $i \in [n]$ as follows. We define

$$\mathcal{H}_{\mathcal{P}} := \{\mathbf{x} \in \mathbb{R}^n : \ g_j(\mathbf{x}) = 0, \ j \in [n], \ p(\mathbf{x}) \geqslant 0, \ p \in \mathcal{P}\}$$

as the *n-dimensional hypercube $\mathcal{H}$ constrained by polynomial inequalities given by $\mathcal{P}$*.

Throughout the paper, we assume that $|\mathcal{P}| = \mathrm{poly}(n)$, i.e., the size of the constraint set $\mathcal{P}$ is polynomial in n. This is usually the case, since otherwise the problem gets less tractable from the optimization point of view.

As a first step, we introduce a *Kronecker function*:

Definition 4.1 For every $\mathbf{v} \in \mathcal{H}$, the function

$$\delta_{\mathbf{v}}(\mathbf{x}) := \prod_{j \in [n]:\ v_j = a_j} \left(\frac{-x_j + b_j}{b_j - a_j} \right) \cdot \prod_{j \in [n]:\ v_j = b_j} \left(\frac{x_j - a_j}{b_j - a_j} \right) \qquad (4.1)$$

is called the *Kronecker delta (function)* of the vector $\mathbf{v}$.

Next we show that the term "Kronecker delta" is justified, i.e., we show that for every $\mathbf{v} \in \mathcal{H}$ the function $\delta_{\mathbf{v}}(\mathbf{x})$ takes the value zero for all $\mathbf{x} \in \mathcal{H}$ except for $\mathbf{x} = \mathbf{v}$ where it takes the value one.

Lemma 4.2 *For every $\mathbf{v} \in \mathcal{H}$, it holds that:*

$$\delta_{\mathbf{v}}(\mathbf{x}) = \begin{cases} 0, & \text{for every } \mathbf{x} \in \mathcal{H} \setminus \{\mathbf{v}\}, \\ 1, & \text{for } \mathbf{x} = \mathbf{v}. \end{cases}$$

Proof On the one hand, if $\mathbf{x} \in \mathcal{H} \setminus \{\mathbf{v}\}$, then there exists an index k such that $\mathbf{x}_k \neq \mathbf{v}_k$. This implies that there exists at least one multiplicative factor in $\delta_{\mathbf{v}}$ which attains the value zero due to (4.1). On the other hand, if $\mathbf{x} = \mathbf{v}$, then we have

$$\delta_{\mathbf{v}}(\mathbf{x}) = \prod_{j \in [n]:\ v_j = a_j} \left(\frac{-a_j + b_j}{b_j - a_j} \right) \prod_{j \in [n]:\ v_j = b_j} \left(\frac{b_j - a_j}{b_j - a_j} \right) = 1.$$

$\square$

The main result of this section is the following theorem.

Theorem 4.3 *Let $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,n}$. Then, $f(\mathbf{x}) \geqslant 0$ for every $\mathbf{x} \in \mathcal{H}_{\mathcal{P}}$ if and only if f has the following representation:*

$$f(\mathbf{x}) = \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} c_{\mathbf{v}} \delta_{\mathbf{v}}(\mathbf{x}) + \sum_{\mathbf{v} \in \mathcal{H} \setminus \mathcal{H}_{\mathcal{P}}} c_{\mathbf{v}} \delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{x}) + \sum_{j=1}^{n} s_j(\mathbf{x}) g_j(\mathbf{x}) +$$

$$\sum_{j=1}^{n} s_{n+j}(\mathbf{x})(-g_j(\mathbf{x})), \qquad (4.2)$$

where $s_1, \ldots, s_{2n} \in C_{n,n-2}$, $c_{\mathbf{v}} \in \mathbb{R}_{\geqslant 0}$, and $p_{\mathbf{v}} \in \mathcal{P}$.

Remark 4.4 Note that the choice of the $p_{\mathbf{v}}$ in the theorem can be stated precisely: If $f(\mathbf{v}) < 0$, then we choose $p_{\mathbf{v}}$ as one of the constraints satisfying $p_{\mathbf{v}} < 0$ (which has to exist since $\mathbf{v} \in \mathcal{H} \setminus \mathcal{H}_{\mathcal{P}}$). If, however, $f(\mathbf{v}) \geqslant 0$, then we choose $p_{\mathbf{v}} = 1$. For further details, see the proof of Theorem 4.3 in Sect. 4.1.

Since we are interested in optimization on the boolean hypercube $\mathcal{H}$, we assume without loss of generality that the polynomial f considered in Theorem 4.3 has degree at most n. Indeed, if f has degree bigger than n, one can efficiently reduce the degree

of f by applying iteratively the polynomial division with respect to polynomials g_j with $j \in [n]$. The remainder of the division process is a polynomial with degree at most n that agrees with f on all the vertices of $\mathcal{H}$.

We begin with proving the easy direction of the equivalence stated in Theorem 4.3.

Lemma 4.5 *If f admits a decomposition of the form* (4.2), *then $f(\mathbf{x})$ is nonnegative for all $\mathbf{x} \in \mathcal{H}_\mathcal{P}$.*

Proof The coefficients $c_\mathbf{v}$ are nonnegative; all $s_j(\mathbf{x})$ are SONC and hence nonnegative on $\mathbb{R}^n$. We have $\pm g_j(\mathbf{x}) \geqslant 0$ for all $\mathbf{x} \in \mathcal{H}$, and for all choices of $\mathbf{v} \in \mathcal{H}$ we have $p_\mathbf{v}(\mathbf{x}) \geqslant 0$ for all $\mathbf{x} \in \mathcal{H}_\mathcal{P}$, and $\delta_\mathbf{v}(\mathbf{x}) \in \{0, 1\}$ for all $\mathbf{x} \in \mathcal{H}$. Thus, the right-hand side of (4.2) is a sum of nonnegative terms for all $\mathbf{x} \in \mathcal{H}_\mathcal{P}$. □

We postpone the rest of the proof of Theorem 4.3 to the end of the section. Now, we state a result about the presentation of the Kronecker delta function $\delta_\mathbf{v}$. In what follows let K be the basic closed semialgebraic set defined by $g_1, \ldots, g_n$ and $l_1^\pm, \ldots, l_n^\pm$ as in (2.3).

Lemma 4.6 *For every $\mathbf{v} \in \mathcal{H}$ the Kronecker delta function can be written as*

$$\delta_\mathbf{v} = \sum_{j=1}^{2^n} s_j H_j^{(n)},$$

for $s_1, \ldots, s_{2^n} \in \mathbb{R}_{\geqslant 0}$ and every $H_j^{(n)}$ given as in (2.4) *with $q = n$.*

Proof First, note that the function $\delta_\mathbf{v}$ can be rewritten as

$$\delta_\mathbf{v}(\mathbf{x}) = \prod_{j=1}^n \frac{1}{b_j - a_j} \prod_{j \in [n]:\ v_j = a_j} \left(-x_j + b_j\right) \prod_{j \in [n]:\ v_j = b_j} \left(x_j - a_j\right),$$

where $\prod_{j=1}^n \frac{1}{b_j - a_j} \in \mathbb{R}_{\geqslant 0}$. Now, the proof follows just by noting that for every $j \in [n]$ both inequalities $-x_j + b_j \geqslant 0$ and $x_j - a_j \geqslant 0$ are in K. □

The following statement is well known in similar variations; see, e.g., [7, Lemma 2.2 and its proof]. For clarity, we provide an own proof here.

Proposition 4.7 *Let $f \in \mathbb{R}[\mathbf{x}]_{n,2d}$ be a polynomial vanishing on $\mathcal{H}$. Then, we have $f = \sum_{j=1}^n p_j g_j$ for some polynomials $p_j \in \mathbb{R}[\mathbf{x}]_{n,2d-2}$.*

Proof Let $\mathcal{J} := \langle g_1, \ldots, g_n \rangle$ be the ideal generated by the g_j's. Let $\mathcal{V}(\mathcal{J})$ denote the affine variety corresponding to $\mathcal{J}$, let $\mathcal{I}(\mathcal{V}(\mathcal{J}))$ denote its radical ideal, and let $\mathcal{I}(\mathcal{H})$ denote the ideal of $\mathcal{H}$. It follows from $\prod_{j=1}^n g_j \in \mathcal{J}$ that $\mathcal{V}(\mathcal{J}) \subseteq \mathcal{H}$ and hence $\mathcal{I}(\mathcal{H}) \subseteq \mathcal{I}(\mathcal{V}(\mathcal{J})) = \mathcal{J}$. The last equality holds since $\mathcal{J}$ itself is a radical ideal. This results from Seidenberg's Lemma; see [35, Proposition 3.7.15] by means of the following observations. The affine variety $\mathcal{V}(\mathcal{J})$ consists exactly of the points defining $\mathcal{H}$, therefore we know that $\mathcal{J}$ is a zero-dimensional ideal. Furthermore, for

every $j \in [n]$ the polynomials g_j satisfy $g_j \in \mathcal{J} \cap \mathbb{R}[x_j]$ and $\gcd(g_j, g'_j) = 1$. Thus, every $f \in \mathcal{I}(\mathcal{H})$ is of the form $f = \sum_{j=1}^{n} p_j g_j$.

Moreover $G := \{g_1, \ldots, g_n\}$ is a Gröbner basis for $\mathcal{J}$ with respect to the graded lexicographic order $\prec_{\text{glex}}$. This follows from Buchberger's Criterion, which says that G is a Gröbner basis for $\mathcal{J}$ if and only if for all pairs $i \neq j$ the remainder on the division of the S-polynomials $S(g_i, g_j)$ by G with respect to $\prec_{\text{glex}}$ is zero. Consider an arbitrary pair g_i, g_j with $i > j$. Then, the corresponding S-polynomial is given by

$$S(g_i, g_j) = (a_j + b_j)x_i^2 x_j - (a_i + b_i)x_i x_j^2 - a_j b_j x_i^2 + a_i b_i x_j^2 .$$

Applying polynomial division with respect to $\prec_{\text{glex}}$ yields the remainder 0 and hence G is a Gröbner basis for $\mathcal{J}$ with respect to $\prec_{\text{glex}}$. Therefore, we conclude that if $f \in \mathbb{R}[\mathbf{x}]_{n,2d}$, then $\deg(p_j) \leqslant 2d - 2$. $\qquad \square$

For an introduction to Gröbner bases, see, for example, [17].

Theorem 4.8 *Let $d \in \mathbb{N}$ and $f \in \mathbb{R}[\mathbf{x}]_{n,2d+2}$ such that f vanishes on $\mathcal{H}$. Then, there exist $s_1, \ldots, s_{2n} \in C_{n,2d}$ such that $f = \sum_{j=1}^{n} s_j g_j + \sum_{j=1}^{n} s_{n+j}(-g_j)$.*

Proof By Proposition 4.7 we know that $f = \sum_{j=1}^{n} p_j g_j$ for some polynomials p_j of degree $\leqslant 2d$. Hence, it is sufficient to show that every single summand $p_j g_j$ is of the form $\sum_{j=1}^{n} s_j g_j - \sum_{j=1}^{n} s_{n+j} g_j$ for some $s_1, \ldots, s_{2n} \in C_{n,2d}$. Let $p_j = \sum_{i=1}^{\ell} a_{ji} m_{ji}$ where every $a_{ji} \in \mathbb{R}$ and every m_{ji} is a single monomial. We show that $p_j g_j$ has the desired form by investigating an arbitrary individual term $a_{ji} m_{ji} g_j$.

Case 1 Assume that the exponent of m_{ji} is contained in $(2\mathbb{N})^n$. If $a_{ji} m_{ji}$ is a monomial square, then $a_{ji} m_{ji}$ is a circuit polynomial. If $a_{ji} < 0$, then $-a_{ji} m_{ji}$ is a monomial square. In both cases we obtain a representation $s_{ji}(\pm g_j)$, where $s_{ji} \in C_{n,2d}$.

Case 2 Assume the exponent $\boldsymbol{\beta}$ of m_{ji} contains odd numbers. Without loss of generality, assume that $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k, \beta_{k+1}, \ldots, \beta_n)$ such that the first k entries are odd and the remaining $n - k$ entries are even. We construct a SONC polynomial $s_{ji} = a_{\boldsymbol{\alpha}(1)} \mathbf{x}^{\boldsymbol{\alpha}(1)} + a_{\boldsymbol{\alpha}(2)} \mathbf{x}^{\boldsymbol{\alpha}(2)} + a_{ji} \mathbf{x}^{\boldsymbol{\beta}}$ such that

$$\boldsymbol{\alpha}(1) = \boldsymbol{\beta} + \sum_{j=1}^{\lceil k/2 \rceil} \mathbf{e}_j - \sum_{j=\lceil k/2 \rceil+1}^{k} \mathbf{e}_j, \quad \boldsymbol{\alpha}(2) = \boldsymbol{\beta} - \sum_{j=1}^{\lceil k/2 \rceil} \mathbf{e}_j + \sum_{j=\lceil k/2 \rceil+1}^{k} \mathbf{e}_j, \quad (4.3)$$

$$|a_{ji}| \leqslant 2\sqrt{a_{\boldsymbol{\alpha}(1)} a_{\boldsymbol{\alpha}(2)}}. \tag{4.4}$$

By the construction (4.3), it follows that $\boldsymbol{\alpha}(1), \boldsymbol{\alpha}(2) \in (2\mathbb{N})^n$ and $\boldsymbol{\beta} = 1/2(\boldsymbol{\alpha}(1) + \boldsymbol{\alpha}(2))$. Thus, s_{ji} is a circuit polynomial and by (4.4) the coefficients $a_{\boldsymbol{\alpha}(1)}, a_{\boldsymbol{\alpha}(2)}$ are chosen large enough such that $|a_{ji}|$ is bounded by the circuit number $2\sqrt{a_{\boldsymbol{\alpha}(1)} a_{\boldsymbol{\alpha}(2)}}$ corresponding to s_{ji}. Therefore, s_{ji} is nonnegative by [31, Theorem 1.1]. Hence, we obtain

$$a_{ji} m_{ji} g_j = s_{ji} g_j + (a_{\boldsymbol{\alpha}(1)} \mathbf{x}^{\boldsymbol{\alpha}(1)} + a_{\boldsymbol{\alpha}(2)} \mathbf{x}^{\boldsymbol{\alpha}(2)})(-g_j),$$

where $s_{ji}, a_{\boldsymbol{\alpha}(1)} \mathbf{x}^{\boldsymbol{\alpha}(1)}$, and $a_{\boldsymbol{\alpha}(2)} \mathbf{x}^{\boldsymbol{\alpha}(2)}$ are nonnegative circuit polynomials.

Degree: All involved nonnegative circuit polynomials are of degree at most $2d$. In Case 1 this follows by construction. In Case 2 we have for the circuit polynomial s_{ji} that $\deg(\boldsymbol{\alpha}(1)), \deg(\boldsymbol{\alpha}(2)) = \deg(\boldsymbol{\beta})$ if k is even, and $\deg(\boldsymbol{\alpha}(1)) = \deg(\boldsymbol{\beta}) + 1$, $\deg(\boldsymbol{\alpha}(2)) = \deg(\boldsymbol{\beta})$ if k is odd. Since $\boldsymbol{\beta}$ is an exponent of the polynomial f, we know that $\deg(\boldsymbol{\beta}) \leqslant 2d$. If k is odd, however, then

$$
\deg(\boldsymbol{\beta}) = \sum_{j=1}^{k} \underbrace{\beta_j}_{\text{odd number}} + \sum_{j=k+1}^{n} \underbrace{\beta_j}_{\text{even number}} ,
$$

i.e., $\deg(\boldsymbol{\beta})$ is a sum of k many odd numbers, with k being odd, plus a sum of even numbers. Thus, $\deg(\boldsymbol{\beta})$ has to be an odd number and hence $\deg(\boldsymbol{\beta}) < 2d$. Therefore, all degrees of terms in s_{ji} are bounded by $2d$ and thus $s_{ji} \in C_{n,2d}$.

Conclusion: Combining the Cases 1 and 2, and the degree argument, we obtain a representation

$$
f = \sum_{j=1}^{n} p_j g_j = \sum_{j=1}^{n} \sum_{i=1}^{\ell_j} a_{ji} m_{ji} g_j = \sum_{j=1}^{n} \sum_{i=1}^{\ell_j} s_{ji} g_j + \bar{s}_{ji}(-g_j),
$$

with $s_{ji}, \bar{s}_{ji} \in C_{n,2d}$ for every $i, j \in [n]$. By defining $s_j = \sum_{i=1}^{\ell_j} s_{ji} \in C_{n,2d}$ and $s_{n+j} = \sum_{i=1}^{\ell_j} \bar{s}_{ji} \in C_{n,2d}$ we obtain the desired representation of f. $\square$

4.1 Proof of Theorem 4.3

In this section, we combine the results of this section and finish the proof of Theorem 4.3.

Due to Lemma 4.5, it remains to show that if $f(\mathbf{x}) \geqslant 0$ for every $\mathbf{x} \in \mathcal{H}_{\mathcal{P}}$, then $f(\mathbf{x})$ admits a decomposition of the form (4.2).

When restricting the domain of the polynomial f to the boolean hypercube $\mathcal{H}$, then f can be represented in the following way:

$$
\begin{aligned}
f(\mathbf{x}) &= f(\mathbf{x}) \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) + f(\mathbf{x}) \sum_{\mathbf{v} \in \mathcal{H} \setminus \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) && \text{for all } \mathbf{x} \in \mathcal{H} \\
&= \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) f(\mathbf{v}) + \sum_{\mathbf{v} \in \mathcal{H} \setminus \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) f(\mathbf{v}) && \text{for all } \mathbf{x} \in \mathcal{H}, \quad (4.5)
\end{aligned}
$$

where the last equality follows by Lemma 4.2 .

Now, our goal is to obtain the desired representation (4.2) from (4.5) via applying Theorem 4.8. The challenge is that the representation (4.2) requires nonnegative coefficients $c_{\mathbf{v}}$ (to obtain a Schmüdgen-like representation). If f is nonnegative over the entire hypercube $\mathcal{H}$, then we can simply put $c_{\mathbf{v}} = f(\mathbf{v})$ and $p_{\mathbf{v}} = 1 = p_0 \in \mathcal{P}$. However, in our construction (4.5) there might exist a vector $\mathbf{v} \in \mathcal{H} \setminus \mathcal{H}_{\mathcal{P}}$ such that f attains a negative value at $\mathbf{v}$. If $f(\mathbf{v}) < 0$, then let $p_{\mathbf{v}} \in \mathcal{P}$ be one of the poly-

nomials among the constraints satisfying $p_{\mathbf{v}}(\mathbf{v}) < 0$. Since by Lemma 4.2 we have $\delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{x}) = \delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{v})$ for every $\mathbf{v}, \mathbf{x} \in \mathcal{H}$, we can now write:

$$f(\mathbf{x}) = \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) f(\mathbf{v}) + \sum_{\mathbf{v} \in \mathcal{H} \backslash \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{x}) \frac{f(\mathbf{v})}{p_{\mathbf{v}}(\mathbf{v})} \qquad \text{for all } \mathbf{x} \in \mathcal{H}.$$

Thus, the polynomial $f(\mathbf{x}) - \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) f(\mathbf{v}) - \sum_{\mathbf{v} \in \mathcal{H} \backslash \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{x}) \frac{f(\mathbf{v})}{p_{\mathbf{v}}(\mathbf{v})}$ has degree at most $n + d$ and vanishes on $\mathcal{H}$. By Theorem 4.8 we finally get

$$f(\mathbf{x}) = \sum_{j=1}^{n} s_j(\mathbf{x}) g_j(\mathbf{x}) + \sum_{j=1}^{n} s_{n+j}(\mathbf{x})(-g_j(\mathbf{x})) + \sum_{\mathbf{v} \in \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) f(\mathbf{v}) +$$
$$\sum_{\mathbf{v} \in \mathcal{H} \backslash \mathcal{H}_{\mathcal{P}}} \delta_{\mathbf{v}}(\mathbf{x}) p_{\mathbf{v}}(\mathbf{x}) \frac{f(\mathbf{v})}{p_{\mathbf{v}}(\mathbf{v})} \; ,$$

for some $s_1, \ldots, s_{2n} \in C_{n,n-2}$ and $p_{\mathbf{v}} \in \mathcal{P}$. This finishes the proof together with Lemma 4.6. ▫

Corollary 4.9 *For every polynomial f which is nonnegative over the boolean hypercube constrained with polynomial inequalities of degree at most d, there exists a degree $n + d$ SONC certificate.*

Proof The argument follows directly from Theorem 4.3 by noting that the right hand side of (4.2) is a SONC certificate of degree $n + d$ (see Definition 2.7). ▫

4.2 Degree d SONC Certificates

In this section, we show that if a polynomial f admits a degree d SONC certificate, then f also admits a short degree d certificate that involves at most $n^{O(d)}$ terms. We conclude the section with a discussion regarding the time complexity of finding a degree d SONC certificate.

Theorem 4.10 *Let f be an n-variate polynomial, nonnegative on the constrained hypercube $\mathcal{H}_{\mathcal{P}}$ with $|\mathcal{P}| = \mathrm{poly}(n)$. Assume that there exists a degree d SONC certificate for f, then there exists a degree d SONC certificate for f involving at most $n^{O(d)}$ many nonnegative circuit polynomials.*

Proof Since there exists a degree d SONC proof of the nonnegativity of f on $\mathcal{H}_{\mathcal{P}}$, we know that

$$f(\mathbf{x}) = \sum_j s_j H_j^{(q)},$$

where the summation is finite, the s_j's are SONCs, and every $H_j^{(q)}$ is a product as defined in (2.4).

Step 1: We analyze the terms s_j. Since every s_j is a SONC, we know that there exists a representation

$$s_j \; = \; \kappa_j \cdot \sum_{i=1}^{k_j} \mu_{ij} \cdot q_{ij}$$

such that $\kappa_j, \mu_{1j}, \ldots, \mu_{k_j j} \in \mathbb{R}_{>0}$, $\sum_{i=1}^{k_j} \mu_{ij} = 1$, and the q_{ij} are nonnegative circuit polynomials. Since s_j is of degree at most d, we know that $Q_j := \{q_{1j}, \ldots, q_{k_j j}\}$ is contained in $\mathbb{R}[\mathbf{x}]_{n,d}$, which is a real vector space of dimension $\binom{n+d}{d}$. Since s_j/κ_j is a convex combination of the q_{ij}, i.e., in the convex hull of Q_j, and $\dim(Q_j) \leqslant \binom{n+d}{d}$, applying Carathéodory's Theorem, see, e.g., [62], yields that s_j/κ_j can be written as a convex combination of at most $\binom{n+d}{d} + 1$ many of the q_{ij}.

Step 2: We analyze the terms $H_j^{(q)}$. By definition of $\mathcal{H}_\mathcal{P}$ and the terms $H_j^{(q)}$, we have

$$H_j^{(q)} \; = \; g_{j_1} \cdots g_{j_s} \cdot l_{r_1}^+ \cdots l_{r_t}^+ \cdot l_{r_1}^- \cdots l_{r_t}^- \cdot p_{\ell_1} \cdots p_{\ell_v}$$

with $j_1, \ldots, j_s \in [n]$, $r_1, \ldots, r_t \in [n]$, and $\ell_1, \ldots, \ell_v \in [m]$. Since the maximal degree of $H_j^{(q)}$ is d, the number of different $H_j^{(q)}$'s is bounded from above by $\binom{n+2n+m}{d}$.

Conclusion: In summary, we obtain a representation:

$$f(\mathbf{x}) \; = \; \sum_{i=1}^{\binom{n+2n+m}{d}} H_j^{(q)} s_j \; = \; \sum_{i=1}^{\binom{n+2n+m}{d}} H_j^{(q)} \kappa_j \sum_{j=1}^{\binom{n+d}{d}+1} \mu_{ij} q_{ij}.$$

Since, as assumed, m can bounded by poly(n), the total number of summands is poly$(n)^{O(d)} = n^{O(d)}$, and we found a desired representation with at most $n^{O(d)}$ nonnegative circuit polynomials of degree at most d. □

Theorem 4.10 states that when searching for a degree d SONC certificate it is enough to restrict to certificates containing at most $n^{O(d)}$ nonnegative circuit polynomials. Moreover, as proved in [22, Theorem 3.2] for a given set $A \subseteq \mathbb{N}^n$, searching through the space of degree d SONC certificates supported on set A can be computed via a relative entropy program (REP) of size $n^{O(d)}$, see, e.g., [22] for more information about REP. However, the above arguments do *not* necessarily imply that the search through the space of degree d SONC certificates can be performed in time $n^{O(d)}$. The difficulty is that one needs to restrict the configuration space of n-variate degree d SONCs to a subset of order $n^{O(d)}$ to be able to *formulate* the corresponding REP in time $n^{O(d)}$. Since the current proof of Theorem 4.10 just guarantees the *existence* of a short SONC certificate, it is currently not clear how to search for a short certificate efficiently. We leave this as an open problem.

5 There Exists No Equivalent to Putinar's Positivstellensatz for SONCs

In this section, we address the open problem raised in [22] asking whether the Theorem 2.6 can be strengthened by requiring $q = 1$. Such a strengthening, for a positive polynomial over some basic closed semialgebraic set, would provide a SONC decomposition equivalent to Putinar's Positivstellensatz for SOS. The advantage of Putinar's Positivstellensatz over Schmüdgen's Positivstellensatz is that for every fixed degree d the cardinality of possible degree d certificates is smaller; for background see, e.g., [42,54] however, asymptotically still in both cases it is $n^{O(d)}$.

We answer this question in a negative way. More precisely, we provide a polynomial f which is strictly positive over the hypercube $\{\pm 1\}^n$ such that there does not exist a SONC decomposition of f for $q = 1$. Moreover, we prove it not only for the most natural choice of the box constraints that is $l_i^{\pm} = 1 \pm x_i$, but for a generic type of box constraints of the form $\ell_i^{\pm} := 1 + c_i \pm x_i$, for $c_i \in \mathbb{R}_{\geqslant 0}$. We close the section with a short discussion.

Let $\mathcal{H} = \{\pm 1\}^n$ and consider the following set of polynomials parametrized by a natural number a:

$$f_a(\mathbf{x}) := (a - 1) \prod_{i=1}^{n} \left(\frac{x_i + 1}{2} \right) + 1.$$

These polynomials satisfy $f_a(\mathbf{e}) = a$ for a vector $\mathbf{e} = \sum_{i=1}^{n} \mathbf{e}_i$ and $f_a(\mathbf{x}) = 1$ for every other $\mathbf{x} \in \mathcal{H} \setminus \{\mathbf{e}\}$. We define for every $d \in \mathbb{N}$

$$S_d := \left\{ \sum_{\text{finite}} s \cdot h \; : \; s \in C_{n,2d}, \; h \in \left\{ 1, \pm(x_i^2 - 1), 1 + c_i \pm x_i \; : \; i \in [n], c_i \in \mathbb{R}_{\geqslant 0} \right\} \right\}$$

as the set of polynomials admitting a SONC decomposition over $\mathcal{H}$ given by Theorem 2.6 for $q = 1$. The main result of this section is the following theorem.

Theorem 5.1 *For every $a > \frac{2^n - 1}{2^{n-2} - 1}$, we have $f_a \notin S_d$ for all $d \in \mathbb{N}$.*

Before we prove this theorem, we show the following structural results. Note that similar structural observations were already made for AGIforms by Reznick in [57] using a different notation.

Lemma 5.2 *Every $s(\mathbf{x}) \in C_{n,2d}$ attains at most two different values on $\mathcal{H} = \{\pm 1\}^n$. Moreover, if $s(\mathbf{x})$ attains two different values, then each value is attained for exactly the half of the hypercube vertices.*

Proof By Definition 2.1 every nonnegative circuit polynomial is of the form:

$$s(\mathbf{x}) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}.$$

Note that for $j = 0, \ldots, r$, we have $\boldsymbol{\alpha}(j) \in (2\mathbb{N})^n$. Hence, when evaluated over the hypercube $\mathbf{x} \in \mathcal{H} = \{\pm 1\}^n$, $s(\mathbf{x})$ can take only one of at most two different values $\sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \pm f_{\boldsymbol{\beta}}$.

If $s(\mathbf{x})$ attains two different values over $\mathcal{H}$, then there has to exist a non-empty subset of variables that have an odd entry in $\boldsymbol{\beta}$. Let $I \subseteq [n]$ be this subset. Then, $s(\mathbf{x}) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)}(\mathbf{x}) - f_{\boldsymbol{\beta}}(\mathbf{x})$, for $\mathbf{x} \in \mathcal{H}$ if and only if $\mathbf{x}$ has an odd number of -1 entries in the set I. The number of such vectors is equal to

$$2^{n-|I|} \sum_{\substack{i=0, \\ i \text{ odd}}}^{|I|} 2^i = 2^{n-|I|} 2^{|I|-1} = 2^{n-1}.$$

$\square$

Lemma 5.3 *Every polynomial $s(\mathbf{x})\ell_i^{\pm}(\mathbf{x})$, with $s \in C_{n,2d}$ and $\ell_i^{\pm} = 1 + c_i \pm x_i$ being a box constraint, attains at most four different values on $\mathcal{H} = \{\pm 1\}^n$. Moreover, each value is attained for at least one-fourth of the hypercube vertices.*

Proof By Lemma 5.2, $s(\mathbf{x})$ attains at most the two values $\left(\sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \pm f_{\boldsymbol{\beta}} \right)$ on $\mathcal{H}$. Similarly, $\ell_i^{\pm}(\mathbf{x})$ attains at most the two values $1 + c_i \pm x_i$ over $\mathcal{H}$. Thus, a polynomial $s(\mathbf{x})\ell_i^{\pm}(\mathbf{x})$ attains at most the four different values $\left(\sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \pm f_{\boldsymbol{\beta}} \right)(1 + c_i \pm x_i)$ on $\mathcal{H}$.

Let I be as in the proof of Lemma 5.2, i.e., the subset of variables that have an odd entry in $\boldsymbol{\beta}$. If $I = \emptyset$, then the first term $\sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}}$ is constant over the hypercube $\mathcal{H}$, thus $s(\mathbf{x})\ell_i^{\pm}(\mathbf{x})$ takes two different values depending on the i-th entry of the vector. Each value is attained for exactly half of the vectors.

If $I \neq \emptyset$ and $i \notin I$, the claim holds since the value of the first term depends only on the entries in I and the value of the second term depends on the i-th entry. Hence, the polynomial $s(\mathbf{x})\ell_i^{\pm}(\mathbf{x})$ attains four values, such that each value is attained on exactly one-fourth of the vertices of $\mathcal{H}$.

Finally, let $I \neq \emptyset$ and $i \in I$. We partition the hypercube vertices into two sets depending on the i-th entry. Each set has cardinality 2^{n-1}. Consider the set with $x_i = 1$. For the vectors in this set the second term takes a constant value $2 + c$. Over this set the polynomial s takes one of the values $\sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)}(\mathbf{x}) \pm f_{\boldsymbol{\beta}}(\mathbf{x})$, depending on whether $\mathbf{x}$ has an odd or even number of -1 entries in the set $I \setminus \{-1\}$. In both cases the number of such vectors is equal to

$$2^{n-|I|} \sum_{\substack{i=0, \\ i \text{ odd}}}^{|I|-1} 2^i = 2^{n-|I|} 2^{|I|-2} = 2^{n-2}.$$

The analysis for the case $x_i = -1$ is analogous. $\square$

Now we can provide the proof of Theorem 5.1.

Proof (Proof of Theorem 5.1)

Assume $f_a \in S_d$ for some $a \in \mathbb{N}$ and $d \in \mathbb{N}$. We prove that a has to be smaller or equal than $\frac{2^n - 1}{2^{n-2} - 1}$. Since $f_a \in S_d$, we know that

$$f_a(\mathbf{x}) = s_0(\mathbf{x}) + \sum_{i=1}^{n} s_i(\mathbf{x})\ell_i^{\pm}(\mathbf{x}) + \sum_{j=1}^{n} \tilde{s}_j(\mathbf{x})(x_j^2 - 1) + \tilde{s}_{j+n}(\mathbf{x})(1 - x_j^2)$$

with $s_0, \ldots, s_n, \tilde{s}_1, \ldots, \tilde{s}_{2n} \in C_{n,2d}$. Since $\pm(x_j^2 - 1)$ for $j \in [n]$ vanishes over the hypercube $\mathcal{H}$, for some $s_0, s_1, \ldots, s_n \in C_{n,2d}$, we can conclude

$$f_a(\mathbf{x}) = s_0(\mathbf{x}) + \sum_{i=1}^{n} s_i(\mathbf{x})\ell_i^{\pm}(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \mathcal{H}. \tag{5.1}$$

Let $s_{0,k}$, and $s_{i,j}$ be some nonnegative circuit polynomials such that $s_0 = \sum_k s_{0,k}$, and $s_i = \sum_j s_{i,j}$ for $i \in [n]$. Thus, we get

$$\sum_{\mathbf{x} \in \mathcal{H}} \left(s_0(\mathbf{x}) + \sum_{i=1}^{n} s_i(\mathbf{x})\ell_i^{\pm}(\mathbf{x}) \right) = \sum_k \sum_{\mathbf{x} \in \mathcal{H}} s_{0,k}(\mathbf{x}) + \sum_{i=1}^{n} \sum_j \sum_{\mathbf{x} \in \mathcal{H}} s_{i,j}(\mathbf{x})\ell_{i,j}^{\pm}(\mathbf{x})$$

$$\geqslant \sum_k 2^{n-1} s_{0,k}(\mathbf{e}) + \sum_{i=1}^{n} \sum_j 2^{n-2} s_{i,j}(\mathbf{e})\ell_{i,j}^{\pm}(\mathbf{e})$$

$$\geqslant 2^{n-2} \left(s_0(\mathbf{e}) + \sum_{i=1}^{n} s_i(\mathbf{e})\ell_i^{\pm}(\mathbf{e}) \right) = 2^{n-2}a,$$

where the first inequality comes directly from Lemma 5.2 and 5.3 and the last equality from the fact that $f_a(\mathbf{e}) = a$. On the other hand, by the properties of the function f_a and the equality (5.1), we know that

$$\sum_{\mathbf{x} \in \mathcal{H}} \left(s_0(\mathbf{x}) + \sum_{i=1}^{n} s_i(\mathbf{x})\ell_i^{\pm}(\mathbf{x}) \right) = 2^n - 1 + a,$$

which makes the subsequent inequality a necessary requirement for $f_a \in S_d$:

$$a \leqslant \frac{2^n - 1}{2^{n-2} - 1}.$$

$\square$

Speaking from a broader perspective, we interpret Theorem 5.1 as an indication that the real algebraic structures, which we use to handle sums of squares, do not apply in the same generality to SONCs. We find this not at all surprising from the point of view that in the nineteenth century Hilbert initially used SOS as a certificate for nonnegativity and many of the algebraic structures in question where developed afterwards with Hilbert's results in mind; see [58] for a historic overview. Our previous work shows that SONCs, in contrast, can, e.g., very well be analyzed with combinatorial methods. We thus see Theorem 5.1 as further evidence about the very different behavior of

SONCs and SOS and as an encouragement to take methods beside the traditional real algebraic ones into account for the successful application of SONCs in the future.

Proving that SONC is an alternative for SOS over the boolean hypercube in the sense of the main results of this article motivates two directions of follow-up research. First, one could specialize further on the boolean hypercube. For example, as we discussed in the introduction, there are many (constrained) boolean hypercube optimization problems known for which SOS provides an efficient algorithm. Given our results, it is a natural question to ask whether SONC can provide comparable algorithmic solutions for these problems. Second, one could generalize the region of feasibility, asking whether the results for SONC on the boolean hypercube can be (partially) extended to more general finite varieties.

References

1. S. Arora, B. Barak, and D. Steurer, *Subexponential algorithms for unique games and related problems*, FOCS, 2010, pp. 563–572.
2. S. Arora, S. Rao, and U. V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, J. ACM **56** (2009), no. 2, 5:1–5:37.
3. B. Barak, S. B. Hopkins, J. A. Kelner, P. Kothari, A. Moitra, and A. Potechin, *A nearly tight sum-of-squares lower bound for the planted clique problem*, IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA, 2016, pp. 428–437.
4. B. Barak, J. A. Kelner, and D. Steurer, *Dictionary learning and tensor decomposition via the sum-of-squares method*, Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015, 2015, pp. 143–151.
5. B. Barak and A. Moitra, *Noisy tensor completion via the sum-of-squares hierarchy*, Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016, 2016, pp. 417–445.
6. B. Barak, P. Raghavendra, and D. Steurer, *Rounding semidefinite programming hierarchies via global correlation*, FOCS, 2011, pp. 472–481.
7. B. Barak and D. Steurer, *Sum-of-squares proofs and the quest toward optimal algorithms*, Electronic Colloquium on Computational Complexity (ECCC) **21** (2014), 59.
8. B. Barak and D. Steurer, *Proofs, beliefs, and algorithms through the lens of sum-of-squares*, 2016, Published online.

9. M. H. Bateni, M. Charikar, and V. Guruswami, *Maxmin allocation via degree lower-bounded arborescences*, STOC, 2009, pp. 543–552.

10. A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization. Analysis, algorithms, and engineering applications.*, vol. 2, Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics; Philadelphia, PA: MPS, Mathematical Programming Society, 2001.

11. G. Blekherman, *There are significantly more nonnegative polynomials than sums of squares*, Israel Journal of Mathematics **153** (2006), no. 1, 355–380.

12. V. Chandrasekaran and P. Shah, *Relative entropy optimization and its applications*, Math. Program. **161** (2017), no. 1-2, 1–32.

13. K. K. H. Cheung, *Computation of the Lasserre ranks of some polytopes*, Math. Oper. Res. **32** (2007), no. 1, 88–94.

14. E. Chlamtac, *Approximation algorithms using hierarchies of semidefinite programming relaxations*, FOCS, 2007, pp. 691–701.

15. E. Chlamtac and G. Singh, *Improved approximation guarantees through higher levels of SDP hierarchies*, APPROX-RANDOM, 2008, pp. 49–62.

16. E. Chlamtac and M. Tulsiani, *Convex relaxations and integrality gaps*, to appear in Handbook on semidefinite, conic and polynomial optimization, Springer, 2012.

17. D.A. Cox and J. Little D. O'Shea, *Ideals, varieties, and algorithms*, fourth ed., Undergraduate Texts in Mathematics, Springer, Cham, 2015, An introduction to computational algebraic geometry and commutative algebra.

18. M. Cygan, F. Grandoni, and M. Mastrolilli, *How to sell hyperedges: The hypermatching assignment problem*, SODA, 2013, pp. 342–351.

19. E. de Klerk and D.V. Pasechnik, *Approximation of the stability number of a graph via copositive programming.*, SIAM J. Optim. **12** (2002), no. 4, 875–892.

20. W. F. de la Vega and C. Kenyon-Mathieu, *Linear programming relaxations of maxcut*, SODA, 2007, pp. 53–61.

21. T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Oberwolfach Rep. (2015), no. 23, 53–56.

22. M. Dressler, S. Iliman, and T. de Wolff, *A Positivstellensatz for Sums of Nonnegative Circuit Polynomials*, SIAM J. Appl. Algebra Geom. **1** (2017), no. 1, 536–555.

23. M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach. **42** (1995), no. 6, 1115–1145.

24. D. Grigoriev, *Complexity of positivstellensatz proofs for the knapsack*, Comput. Complexity **10** (2001), no. 2, 139–154.

25. D. Grigoriev, E. A. Hirsch, and D. V. Pasechnik, *Complexity of semi-algebraic proofs*, STACS, 2002, pp. 419–430.

26. D. Grigoriev and N. Vorobjov, *Complexity of null-and positivstellensatz proofs*, Ann. Pure App. Logic **113** (2001), no. 1-3, 153–160.

27. M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, vol. 2, Springer, 1988.

28. V. Guruswami and A. K. Sinop, *Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives*, FOCS, 2011, pp. 482–491.

29. D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten*, Annals of Mathematics **32** (1888), 342–350.

30. S. B. Hopkins, T. Schramm, J. Shi, and D. Steurer, *Fast spectral algorithms from sum-of-squares proofs: tensor decomposition and planted sparse vectors*, Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016, 2016, pp. 178–191.

31. S. Iliman and T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Res. Math. Sci. **3** (2016), 3:9.

32. L.G. Khachiyan, *Polynomial algorithms in linear programming*, USSR Computational Mathematics and Mathematical Physics **20** (1980), no. 1, 53 – 72.

33. P. K. Kothari, R. Mori, R. O'Donnell, and D. Witmer, *Sum of squares lower bounds for refuting any CSP*, Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017, 2017, pp. 132–145.

34. P. K. Kothari, J. Steinhardt, and D. Steurer, *Robust moment estimation and improved clustering via sum of squares*, Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, 2018.

35. M. Kreuzer and L. Robbiano, *Computational commutative algebra. 1*, Springer-Verlag, Berlin, 2000.

36. A. Kurpisz, S. Leppänen, and M. Mastrolilli, *Sum-of-squares hierarchy lower bounds for symmetric formulations*, Integer Programming and Combinatorial Optimization - 18th International Conference, IPCO 2016, Liège, Belgium, June 1-3, 2016, Proceedings, 2016, pp. 362–374.

37. A. Kurpisz, S. Leppänen, and M. Mastrolilli, *On the hardest problem formulations for the 0/1 lasserre hierarchy*, Math. Oper. Res. **42** (2017), no. 1, 135–143.

38. A. Kurpisz, S. Leppänen, and M. Mastrolilli, *An unbounded sum-of-squares hierarchy integrality gap for a polynomially solvable problem*, Math. Program. **166** (2017), no. 1-2, 1–17.

39. J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2000/01), no. 3, 796–817.

40. M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming*, Math. Oper. Res. **28** (2003), no. 3, 470–496.

41. M. Laurent, *Lower bound for the number of iterations in semidefinite hierarchies for the cut polytope*, Math. Oper. Res. **28** (2003), no. 4, 871–883.

42. M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.

43. J. R. Lee, P. Raghavendra, and D. Steurer, *Lower bounds on the size of semidefinite programming relaxations*, STOC, 2015, pp. 567–576.

44. L. Lovász, *On the shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), 1–7.

45. A. Magen and M. Moharrami, *Robust algorithms for on minor-free graphs based on the Sherali-Adams hierarchy*, APPROX-RANDOM, 2009, pp. 258–271.

46. M. Mastrolilli, *High degree sum of squares proofs, bienstock-zuckerberg hierarchy and CG cuts*, Integer Programming and Combinatorial Optimization - 19th International Conference, IPCO 2017, Waterloo, ON, Canada, June 26-28, 2017, Proceedings, 2017, pp. 405–416.

47. R. Meka, A. Potechin, and A. Wigderson, *Sum-of-squares lower bounds for planted clique*, Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015, 2015, pp. 87–96.

48. T.S. Motzkin, *The arithmetic-geometric inequality*, Symposium on Inequalities (1967), 205–224, cited By 1.

49. Y. Nesterov, *Global quadratic optimization via conic relaxation*, pp. 363–384, Kluwer Academic Publishers, 2000.

50. R. O'Donnell, *SOS is not obviously automatizable, even approximately*, 8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA, 2017, pp. 59:1–59:10.

51. J. Oxley, *Matroid theory*, Oxford Graduate Texts in Mathematics, vol. 2, Oxford University Press, 2011.

52. P. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, PhD thesis, California Institute of Technology, 2000.

53. A. Potechin and D. Steurer, *Exact tensor completion with sum-of-squares*, Proceedings of the 30th Conference on Learning Theory, COLT 2017, Amsterdam, The Netherlands, 7-10 July 2017, 2017, pp. 1619–1673.

54. M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. **42** (1993), no. 3, 969–984.

55. P. Raghavendra and N. Tan, *Approximating csps with global cardinality constraints using sdp hierarchies*, SODA, 2012, pp. 373–387.

56. P. Raghavendra and B. Weitz, *On the bit complexity of sum-of-squares proofs*, 44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland, 2017, pp. 80:1–80:13.

57. B. Reznick, *Forms derived from the arithmetic-geometric inequality*, Math. Ann. **283** (1989), no. 3, 431–464.

58. B. Reznick, *Some concrete aspects of Hilbert's 17th Problem*, Real algebraic geometry and ordered structures (Baton Rouge, LA, 1996), Contemp. Math., vol. 253, Amer. Math. Soc., Providence, RI, 2000, pp. 251–272.

59. T. Schramm and D. Steurer, *Fast and robust tensor decomposition with applications to dictionary learning*, Proceedings of the 30th Conference on Learning Theory, COLT 2017, Amsterdam, The Netherlands, 7-10 July 2017, 2017, pp. 1760–1793.
60. N. Shor, *Class of global minimum bounds of polynomial functions*, Cybernetics **23** (1987), no. 6, 731–734.
61. J. Thapper and S. Zivny, *The limits of SDP relaxations for general-valued csps*, 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, Reykjavik, Iceland, June 20-23, 2017, 2017, pp. 1–12.
62. G.M. Ziegler, *Lectures on polytopes*, Springer Verlag, 2007.