A Second Order Cone Characterization for Sums of Nonnegative Circuits

Jie Wang and Victor Magron
jwang,vmagron@laas.fr
Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS)
Toulouse, France

ABSTRACT

The second-order cone (SOC) is a class of simple convex cones and optimizing over them can be done more efficiently than with semidefinite programming. It is interesting both in theory and in practice to investigate which convex cones admit a representation using SOCs, given that they have a strong expressive ability. In this paper, we prove constructively that the cone of sums of nonnegative circuits (SONC) admits an SOC representation. Based on this, we give a new algorithm to compute SONC decompositions for certain classes of nonnegative polynomials via SOC programming. Numerical experiments demonstrate the efficiency of our algorithm for polynomials with a fairly large size (both size of degree and number of variables).

CCS CONCEPTS

• **Mathematics of computing** → **Semidefinite programming**; • **Computing methodologies** → **Algebraic algorithms**; **Optimization algorithms**.

KEYWORDS

sum of nonnegative circuit polynomials, second-order cone representation, second-order cone programming, polynomial optimization, sum of binomial squares

ACM Reference Format:
Jie Wang and Victor Magron. 2020. A Second Order Cone Characterization for Sums of Nonnegative Circuits. In *International Symposium on Symbolic and Algebraic Computation (ISSAC '20), July 20–23, 2020, Kalamata, Greece.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3373207.3404033

1 INTRODUCTION

A *circuit polynomial* is of the form $\sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1,\dots,x_n]$, where $c_{\boldsymbol{\alpha}} > 0$ for all $\boldsymbol{\alpha} \in \mathscr{A}$, $\mathscr{A} \subseteq (2\mathbb{N})^n$ comprises the vertices of a simplex and $\boldsymbol{\beta}$ lies in the interior of this simplex. The set of *sums of nonnegative circuit polynomials (SONC)* was introduced by Iliman and Wolff in [10] as a new certificate of nonnegativity for sparse polynomials, which is independent of the well-known set of sums of squares (SOS). Another recently introduced alternative certificates [6] are sums of arithmetic-geometric-exponentials (SAGE), which can be obtained via relative entropy programming. The connection between SONC and SAGE polynomials have been recently studied in [13, 20, 27]. It happens that SONC polynomials and SAGE polynomials are actually equivalent [20], and that both have a cancellation-free representation in terms of generators [20, 27].

One of the significant differences between SONC and SOS is that SONC decompositions preserve sparsity of polynomials while SOS decompositions do not in general [27]. The set of SONC polynomials with a given support forms a convex cone, called a *SONC cone*. Optimization problems over SONC cones can be formulated as geometric programs or more generally relative entropy programs (see [11] for the unconstrained case and [7] for the constrained case). Numerical experiments for unconstrained POPs (polynomial optimization problems) in [25] have demonstrated the advantage of the SONC-based methods compared to the SOS-based methods, especially in the high-degree but fairly sparse case.

In the SOS case, there have been several attempts to exploit sparsity occurring in (un-)constrained POPs. The sparse variant [26] of the moment-SOS hierarchy exploits the correlative sparsity pattern among the input variables to reduce the support of the resulting SOS decompositions. Such sparse representation results have been successfully applied in many fields, such as optimal power-flow [12], roundoff error bounds [15] and recently extended to the noncommutative case [14]. Another way to exploit sparsity is to consider patterns based on terms (rather than variables), yielding an alternative sparse variant of Lasserre's hierarchy [28].

One of the similar features shared by SOS/SONC-based frameworks is their intrinsic connections with conic programming: SOS decompositions are computed via semidefinite programming and SONC decompositions via geometric programming. In both cases, the resulting optimization problems are solved with interior-point algorithms, thus output approximate nonnegativity certificates. However, one can still obtain an exact certificate from such output via hybrid numerical-symbolic algorithms when the input polynomial lies in the interior of the SOS/SONC cone. One way is to rely on rounding-projection algorithms adapted to the SOS cone [22] and the SONC cone [19], or alternatively on perturbation-compensation schemes [16, 18] available within the `RealCertify` [17] library.

In this paper, we study the second-order cone representation of SONC cones. An n-dimensional *(rotated) second-order cone (SOC)* is defined as $\mathbf{K}^n := \{\boldsymbol{a} \in \mathbb{R}^n \mid 2a_1a_2 \geq \sum_{i=3}^n a_i^2, a_1 \geq 0, a_2 \geq 0\}$. The SOC is well-studied and has mature solvers. Optimizing via second-order cone programming (SOCP) can be handled more efficiently than with semidefinite programming [1, 2]. On the other hand, despite the simplicity of SOCs, they have a strong ability to express other convex cones (many such examples can be found in [5, Section 3.3]). Therefore, it is interesting in theory and also important from

ISSAC '20, July 20–23, 2020, Kalamata, Greece

ACM ISBN 978-1-4503-7100-1/20/07…$15.00
https://doi.org/10.1145/3373207.3404033

the view of applications to investigate which convex cones can be expressed by SOCs.

Given sets of lattice points $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \text{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \text{conv}(\mathscr{A}) \cap (\mathbb{N}^n \backslash (2\mathbb{N})^n)$ ($\text{conv}(\mathscr{A})$ is the convex hull of $\mathscr{A}$) with $\mathscr{A} \cap \mathscr{B}_1 = \emptyset$, let $\text{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ be the SONC cone supported on $\mathscr{A}, \mathscr{B}_1, \mathscr{B}_2$ (see Definition 4.2). The first main result of this paper is the following theorem.

Theorem 1.1. *For $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \text{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \text{conv}(\mathscr{A}) \cap (\mathbb{N}^n \backslash (2\mathbb{N})^n)$ with $\mathscr{A} \cap \mathscr{B}_1 = \emptyset$, the convex cone $\text{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ admits an SOC representation.*

The fact that SONC cones admit an SOC characterization was firstly proven by Averkov [4, Theorem 17]. However, Averkov's result is more theoretical. Even though Averkov's proof theoretically allows one to construct an SOC representation for a SONC cone, the construction is complicated and wasn't explicitly given in Averkov's paper. Our proof of Theorem 1.1, which involves writing a SONC polynomial as a sum of binomial squares with rational exponents (Theorem 3.9), is totally different from Averkov's and leads to a more concise (hence more efficient) SOC representation for SONC cones. This enables us to propose a new algorithm, based on SOCP, providing SONC decompositions for a certain class of nonnegative polynomials, which in turn yields lower bounds for unconstrained POPs. We test the algorithm on various randomly generated polynomials up to a fairly large size, involving $n \sim 40$ variables and of degree $d \sim 60$. The numerical results demonstrate the efficiency of our algorithm.

2 PRELIMINARIES

Let $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$ be the ring of real n-variate polynomials, and let $\mathbb{R}_+$ be the set of positive real numbers. For a finite set $\mathscr{A} \subseteq \mathbb{N}^n$, we denote by $\text{conv}(\mathscr{A})$ the convex hull of $\mathscr{A}$. Given a finite set $\mathscr{A} \subseteq \mathbb{N}^n$, we consider polynomials $f \in \mathbb{R}[\mathbf{x}]$ supported on $\mathscr{A} \subseteq \mathbb{N}^n$, i.e., f is of the form $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ with $c_{\boldsymbol{\alpha}} \in \mathbb{R}, \mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of f is $\text{supp}(f) := \{\boldsymbol{\alpha} \in \mathscr{A} \mid c_{\boldsymbol{\alpha}} \neq 0\}$ and the Newton polytope of f is defined as $\text{New}(f) := \text{conv}(\text{supp}(f))$. For a polytope P, we use $V(P)$ to denote the vertex set of P and use P° to denote the interior of P. For a set A, we use $\#A$ to denote the cardinality of A. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ which is nonnegative over $\mathbb{R}^n$ is called a *nonnegative polynomial*, or a *positive semi-definite (PSD) polynomial*. The following definition of circuit polynomials was proposed by Iliman and De Wolff in [10].

Definition 2.1. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is called a *circuit polynomial* if it is of the form $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}}$ and satisfies the following conditions: (i) $\mathscr{A} \subseteq (2\mathbb{N})^n$ comprises the vertices of a simplex, (ii) $c_{\boldsymbol{\alpha}} > 0$ for each $\boldsymbol{\alpha} \in \mathscr{A}$, (iii) $\boldsymbol{\beta} \in \text{conv}(\mathscr{A})^\circ \cap \mathbb{N}^n$.

If $f = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}}$ is a circuit polynomial, then from the definition we can uniquely write $\boldsymbol{\beta} = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} \lambda_{\boldsymbol{\alpha}} \boldsymbol{\alpha}$ with $\lambda_{\boldsymbol{\alpha}} > 0$ and $\sum_{\boldsymbol{\alpha} \in \mathscr{A}} \lambda_{\boldsymbol{\alpha}} = 1$. We define the corresponding *circuit number* as $\Theta_f := \prod_{\boldsymbol{\alpha} \in \mathscr{A}} (c_{\boldsymbol{\alpha}}/\lambda_{\boldsymbol{\alpha}})^{\lambda_{\boldsymbol{\alpha}}}$. The nonnegativity of the circuit polynomial f is decided by its circuit number alone, that is, f is nonnegative if and only if either $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ and $|d| \leq \Theta_f$, or $\boldsymbol{\beta} \in (2\mathbb{N})^n$ and $d \leq \Theta_f$ ([10, Theorem 3.8]). To provide a concise narrative, we refer to a nonnegative circuit polynomial by a nonnegative circuit and also view a monomial square as a nonnegative circuit. An explicit representation of a polynomial being a *sum of nonnegative circuits*, or *SONC* for short, provides a certificate for its nonnegativity. Such a certificate is called a *SONC decomposition*. For simplicity, we denote the set of SONC polynomials by SONC.

For a polynomial $f \in \mathbb{R}[\mathbf{x}]$, let $\Lambda(f) := \{\boldsymbol{\alpha} \in \text{supp}(f) \mid \boldsymbol{\alpha} \in (2\mathbb{N})^n \text{ and } c_{\boldsymbol{\alpha}} > 0\}$ and $\Gamma(f) := \text{supp}(f) \backslash \Lambda(f)$. Then we can write f as $f = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$. For each $\boldsymbol{\beta} \in \Gamma(f)$, let

$$\mathscr{F}(\boldsymbol{\beta}) := \{\Delta \mid \Delta \text{ is a simplex}, \boldsymbol{\beta} \in \Delta^\circ, V(\Delta) \subseteq \Lambda(f)\}. \quad (1)$$

By [27, Theorem 5.5], if $f \in \text{SONC}$, then it has a decomposition

$$f = \sum_{\boldsymbol{\beta} \in \Gamma(f)} \sum_{\Delta \in \mathscr{F}(\boldsymbol{\beta})} f_{\boldsymbol{\beta}\Delta} + \sum_{\boldsymbol{\alpha} \in \tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}, \quad (2)$$

where $f_{\boldsymbol{\beta}\Delta}$ is a nonnegative circuit supported on $V(\Delta) \cup \{\boldsymbol{\beta}\}$ for each Δ and $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta} \in \Gamma(f)} \cup_{\Delta \in \mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$.

3 SONC AND SUMS OF BINOMIAL SQUARES

In this section, we give a characterization of SONC polynomials in terms of sums of binomial squares with rational exponents.

3.1 Rational mediated sets

A lattice point $\boldsymbol{\alpha} \in \mathbb{N}^n$ is *even* if it is in $(2\mathbb{N})^n$. For a subset $M \subseteq \mathbb{N}^n$, define $\overline{A}(M) := \{\frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) \mid \boldsymbol{v} \neq \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{w} \in M \cap (2\mathbb{N})^n\}$ as the set of averages of distinct even points in M. A subset $\mathscr{A} \subseteq (2\mathbb{N})^n$ is called a *trellis* if $\mathscr{A}$ comprises the vertices of a simplex. For a trellis $\mathscr{A}$, we call M an *$\mathscr{A}$-mediated set* if $\mathscr{A} \subseteq M \subseteq \overline{A}(M) \cup \mathscr{A}$ ([9, 23, 24]).

Theorem 3.1. *Let $f = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$ with $d \neq 0$ be a nonnegative circuit. Then f is a sum of binomial squares iff there exists an $\mathscr{A}$-mediated set containing $\boldsymbol{\beta}$. Moreover, suppose that $\boldsymbol{\beta}$ belongs to an $\mathscr{A}$-mediated set M and for each $\boldsymbol{u} \in M \backslash \mathscr{A}$, let us write $\boldsymbol{u} = \frac{1}{2}(\boldsymbol{v}_{\boldsymbol{u}} + \boldsymbol{w}_{\boldsymbol{u}})$ for some $\boldsymbol{v}_{\boldsymbol{u}} \neq \boldsymbol{w}_{\boldsymbol{u}} \in M \cap (2\mathbb{N})^n$. Then one has the decomposition $f = \sum_{\boldsymbol{u} \in M \backslash \mathscr{A}} (a_{\boldsymbol{u}} \mathbf{x}^{\frac{1}{2}\boldsymbol{v}_{\boldsymbol{u}}} - b_{\boldsymbol{u}} \mathbf{x}^{\frac{1}{2}\boldsymbol{w}_{\boldsymbol{u}}})^2$, with $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$.*

Proof. It follows from Theorem 5.2 in [10]. □

By Theorem 3.1, if we want to represent a nonnegative circuit polynomial as a sum of binomial squares, we need to first decide if there exists an $\mathscr{A}$-mediated set containing a given lattice point and then to compute one if there exists. However, there are obstacles for each of these two steps: (1) there may not exist such an $\mathscr{A}$-mediated set containing a given lattice point; (2) even if such a set exists, there is no efficient algorithm to compute it. In order to overcome these two difficulties, we introduce the concept of $\mathscr{A}$-rational mediated sets as a replacement of $\mathscr{A}$-mediated sets by admitting rational numbers in coordinates.

Concretely, for a subset $M \subseteq \mathbb{Q}^n$, let us define $\widetilde{A}(M) := \{\frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) \mid \boldsymbol{v} \neq \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{w} \in M\}$ as the set of averages of distinct rational points in M. Let us assume that $\mathscr{A} \subseteq \mathbb{Q}^n$ comprises the vertices of a simplex. We say that M is an *$\mathscr{A}$-rational mediated set* if $\mathscr{A} \subseteq M \subseteq \widetilde{A}(M) \cup \mathscr{A}$. We shall see that for a trellis $\mathscr{A}$ and a lattice point $\boldsymbol{\beta} \in \text{conv}(\mathscr{A})^\circ$, an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$ always exists and moreover, there is an effective algorithm to compute it.

First, let us consider the one dimensional case. For a sequence of integer numbers $A = \{s, q_1, \ldots, q_m, p\}$ (arranged from small to large), if every q_i is an average of two distinct numbers in A, then we say A is an *(s, p)-mediated sequence*. Note that the property of (s, p)-mediated sequences is preserved under translations, that is, there is a one-to-one correspondence between (s, p)-mediated sequences and $(s + r, p + r)$-mediated sequences for any integer number r. So it suffices to consider the case of $s = 0$.

For a fixed p and an integer $0 < q < p$, a *minimal* $(0, p)$*-mediated sequence* containing q is a $(0, p)$-mediated sequence containing q with the least number of elements. Denote the number of elements in a minimal $(0, p)$-mediated sequence containing q by $N(\frac{q}{p})$. One can then easily show that $N(\frac{1}{p}) = \lceil \log_2(p) \rceil + 2$ by induction on p. We conjecture that this formula holds for general q, i.e.,

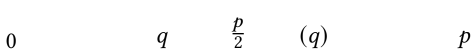
Conjecture 3.2. *If* $gcd(p, q) = 1$, *then* $N(\frac{q}{p}) = \lceil \log_2(p) \rceil + 2$.

Generally we do not know how to compute a minimal $(0, p)$-mediated sequence containing a given q. However, we have an algorithm to compute an approximately minimal $(0, p)$-mediated sequence containing a given q as the following lemma shows.

Lemma 3.3. *For* $0 < q < p \in \mathbb{N}$, *there exists a* $(0, p)$*-mediated sequence containing* q *with the cardinality less than* $\frac{1}{2}(\log_2(p) + \frac{3}{2})^2$.

Proof. We can assume $\gcd(p, q) = 1$ (otherwise one can consider $p/\gcd(p, q), q/\gcd(p, q)$ instead). Let us do induction on p. Assume that for any $p', q' \in \mathbb{N}, 0 < q' < p' < p$, there exists a $(0, p')$-mediated sequence containing q' with the number of elements less than $\frac{1}{2}(\log_2(p') + \frac{3}{2})^2$.

Case 1: Suppose that p is an even number. If $q = \frac{p}{2}$, then by $\gcd(p, q) = 1$, we have $q = 1$ and $A = \{0, 1, 2\}$ is a $(0, p)$-mediated sequence containing q. Otherwise, we have either $0 < q < \frac{p}{2}$ or $\frac{p}{2} < q < p$. For $0 < q < \frac{p}{2}$, by the induction hypothesis, there exists a $(0, \frac{p}{2})$-mediated sequence A' containing q. For $\frac{p}{2} < q < p$, since the property of mediated sequences is preserved under translations, one can first subtract $\frac{p}{2}$ and obtain a $(0, \frac{p}{2})$-mediated sequence containing $q - \frac{p}{2}$ by the induction hypothesis. Then by adding $\frac{p}{2}$, one obtains a $(\frac{p}{2}, p)$-mediated sequence A' containing q. It follows that $A = A' \cup \{p\}$ or $A = \{0\} \cup A'$ is a $(0, p)$-mediated sequence containing q.

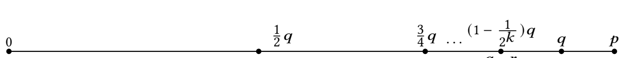


Moreover, we have

$$\#A = 1 + \#A' < 1 + \frac{1}{2}(\log_2(\frac{p}{2}) + \frac{3}{2})^2 < \frac{1}{2}(\log_2(p) + \frac{3}{2})^2.$$

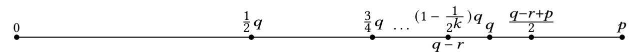
Case 2: Suppose that p is an odd number. Without loss of generality, assume that q is an even number (otherwise one can consider $p - q$ instead and then obtain a $(0, p)$-mediated sequence containing q through the map $x \mapsto p - x$ which clearly preserves the property of mediated sequences).

Let $q = 2^k r$ for some $k, r \in \mathbb{N}\backslash\{0\}$ and $2 \nmid r$. If $q = p - r$, then $q = \frac{q-r+p}{2}$. Since $\gcd(p, q) = 1$, we have $r = 1$. Let $A = \{0, \frac{1}{2}q, \frac{3}{4}q, \ldots, (1 - \frac{1}{2^k})q, q, p\}$. For $1 \le i \le k$, we have $(1 - \frac{1}{2^i})q = \frac{1}{2}(1 - \frac{1}{2^{i-1}})q + \frac{1}{2}q$. Therefore, A is a $(0, p)$-mediated sequence containing q.



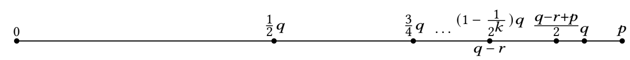
$$\#A = k + 3 < \frac{1}{2}(\log_2(2^k + 1) + \frac{3}{2})^2 = \frac{1}{2}(\log_2(p) + \frac{3}{2})^2.$$

If $q < p - r$, then q lies on the line segment between $q - r$ and $\frac{q-r+p}{2}$. Since $\frac{q-r+p}{2} - (q - r) = \frac{p+r-q}{2} < p$, then by the induction hypothesis, there exists a $(q - r, \frac{q-r+p}{2})$-mediated sequence A' containing q (using translations). It follows that $A = \{0, \frac{1}{2}q, \frac{3}{4}q, \ldots, (1 - \frac{1}{2^{k-1}})q, p\} \cup A'$ is a $(0, p)$-mediated sequence containing q.



$$\begin{aligned}\#A = k + 1 + \#A' &< \log_2(\frac{q}{r}) + 1 + \frac{1}{2}(\log_2(\frac{p + r - q}{2}) + \frac{3}{2})^2 \\ &< \log_2(p) + 1 + \frac{1}{2}(\log_2(\frac{p}{2}) + \frac{3}{2})^2 \\ &= \frac{1}{2}(\log_2(p) + \frac{3}{2})^2.\end{aligned}$$

If $q > p - r$, then q lies on the line segment between $\frac{q-r+p}{2}$ and p. Since $p - \frac{q-r+p}{2} = \frac{p+r-q}{2} < p$, then by the induction hypothesis, there exists a $(\frac{q-r+p}{2}, p)$-mediated sequence A' containing q (using translations). It follows that the set $A = \{0, \frac{1}{2}q, \frac{3}{4}q, \ldots, (1 - \frac{1}{2^k})q\} \cup A'$ is a $(0, p)$-mediated sequence containing q.



As previously, we have $\#A = k + 1 + \#A' < \frac{1}{2}(\log_2(p) + \frac{3}{2})^2$. □

Lemma 3.4. *Suppose that* $\boldsymbol{\alpha}_1$ *and* $\boldsymbol{\alpha}_2$ *are two rational points, and* $\boldsymbol{\beta}$ *is any rational point on the line segment between* $\boldsymbol{\alpha}_1$ *and* $\boldsymbol{\alpha}_2$. *Then there exists an* $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$*-rational mediated set* M *containing* $\boldsymbol{\beta}$. *Furthermore, if the denominators of coordinates of* $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}$ *are odd numbers, and the numerators of coordinates of* $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ *are even numbers, then we can ensure that the denominators of coordinates of points in* M *are odd numbers and the numerators of coordinates of points in* $M\backslash\{\boldsymbol{\beta}\}$ *are even numbers.*

Proof. Suppose $\boldsymbol{\beta} = (1 - \frac{q}{p})\boldsymbol{\alpha}_1 + \frac{q}{p}\boldsymbol{\alpha}_2$, $p, q \in \mathbb{N}, 0 < q < p, \gcd(p, q) = 1$. We then construct a one-to-one correspondence between the points on the one-dimensional number axis and the points on the line across $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ via the map: $s \mapsto (1 - \frac{s}{p})\boldsymbol{\alpha}_1 + \frac{s}{p}\boldsymbol{\alpha}_2$, such that $\boldsymbol{\alpha}_1$ corresponds to the origin, $\boldsymbol{\alpha}_2$ corresponds to p and $\boldsymbol{\beta}$ corresponds to q. Then it is clear that a $(0, p)$-mediated sequence containing q corresponds to a $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$-rational mediated set containing $\boldsymbol{\beta}$. Hence by Lemma 3.3, there exists a $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$-rational mediated set M containing $\boldsymbol{\beta}$ with the number of elements less than $\frac{1}{2}(\log_2(p) + \frac{3}{2})^2$. Moreover, we can see that if $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}$ are lattice points, then the elements in M are also lattice points.

If the denominators of coordinates of $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}$ are odd numbers, and the numerators of coordinates of $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ are even numbers, assume that the least common multiple of denominators appearing in the coordinates of $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}$ is r and then remove the denominators by multiplying the coordinates of $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}$ by r such that $r\boldsymbol{\alpha}_1, r\boldsymbol{\alpha}_2$ are even lattice points. If $r\boldsymbol{\beta}$ is even, let M' be the $\{\frac{r}{2}\boldsymbol{\alpha}_1, \frac{r}{2}\boldsymbol{\alpha}_2\}$-rational mediated set containing $\frac{r}{2}\boldsymbol{\beta}$ obtained as above (the elements in M' are lattice points). Then $M = \frac{2}{r}M' := \{\frac{2}{r}\boldsymbol{u} \mid \boldsymbol{u} \in M'\}$

is an $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$-rational mediated set containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in M are odd numbers and the numerators of coordinates of points in $M\backslash\{\boldsymbol{\beta}\}$ are even numbers.

If $r\boldsymbol{\beta}$ is not even, assume without loss of generality that $\boldsymbol{\beta}$ lies on the line segment between $\boldsymbol{\alpha}_1$ and $\frac{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2}{2}$. Let $\boldsymbol{\beta}' = 2\boldsymbol{\beta}-\boldsymbol{\alpha}_1$ with $r\boldsymbol{\beta}'$ an even lattice point. Let M' be the $\{\frac{r}{2}\boldsymbol{\alpha}_1, \frac{r}{2}\boldsymbol{\alpha}_2\}$-rational mediated set containing $\frac{r}{2}\boldsymbol{\beta}'$ obtained as above (note that the elements in M' are lattice points). Then $M = \frac{2}{r}M' \cup \{\boldsymbol{\beta}\}$ is an $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$-rational mediated set containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in M are odd numbers and the numerators of coordinates of points in $M\backslash\{\boldsymbol{\beta}\}$ are even numbers as desired. □

Lemma 3.5. *For a trellis $\mathscr{A} = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m\}$ and a lattice point $\boldsymbol{\beta} \in \operatorname{conv}(\mathscr{A})^\circ$, there exists an $\mathscr{A}$-rational mediated set $M_{\mathscr{A}\boldsymbol{\beta}}$ containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}$ are odd numbers and the numerators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}\backslash\{\boldsymbol{\beta}\}$ are even numbers.*

Proof. Suppose $\boldsymbol{\beta} = \sum_{i=1}^m \frac{q_i}{p}\boldsymbol{\alpha}_i$, where $p = \sum_{i=1}^m q_i$, $p, q_i \in \mathbb{N}\backslash\{0\}$, $(p, q_1, \ldots, q_m) = 1$. If p is an even number, then because $(p, q_1, \ldots, q_m) = 1$, there must exist an odd number among the q_i's. Without loss of generality assume q_1 is an odd number. If p is an odd number and there exists an even number among the q_i's, then without loss of generality assume q_1 is an even number. In any of these two cases, we have

$$\boldsymbol{\beta} = \frac{q_1}{p}\boldsymbol{\alpha}_1 + \frac{p-q_1}{p}\left(\frac{q_2}{p-q_1}\boldsymbol{\alpha}_2 + \cdots + \frac{q_m}{p-q_1}\boldsymbol{\alpha}_m\right).$$

Let $\boldsymbol{\beta}_1 = \frac{q_2}{p-q_1}\boldsymbol{\alpha}_2 + \cdots + \frac{q_m}{p-q_1}\boldsymbol{\alpha}_m$. Then $\boldsymbol{\beta} = \frac{q_1}{p}\boldsymbol{\alpha}_1 + \frac{p-q_1}{p}\boldsymbol{\beta}_1$.

If p is an odd number and all q_i's are odd numbers, then we have

$$\begin{aligned}\boldsymbol{\beta} = &\frac{q_1}{q_1+q_2}\left(\frac{q_1+q_2}{p}\boldsymbol{\alpha}_1 + \frac{q_3}{p}\boldsymbol{\alpha}_3 + \cdots + \frac{q_m}{p}\boldsymbol{\alpha}_m\right)\\ &+ \frac{q_2}{q_1+q_2}\left(\frac{q_1+q_2}{p}\boldsymbol{\alpha}_2 + \frac{q_3}{p}\boldsymbol{\alpha}_3 + \cdots + \frac{q_m}{p}\boldsymbol{\alpha}_m\right).\end{aligned}$$

Let $\boldsymbol{\beta}_1 = \frac{q_1+q_2}{p}\boldsymbol{\alpha}_1 + \frac{q_3}{p}\boldsymbol{\alpha}_3 + \cdots + \frac{q_m}{p}\boldsymbol{\alpha}_m$ and $\boldsymbol{\beta}_2 = \frac{q_1+q_2}{p}\boldsymbol{\alpha}_2 + \frac{q_3}{p}\boldsymbol{\alpha}_3 + \cdots + \frac{q_m}{p}\boldsymbol{\alpha}_m$. Then $\boldsymbol{\beta} = \frac{q_1}{q_1+q_2}\boldsymbol{\beta}_1 + \frac{q_2}{q_1+q_2}\boldsymbol{\beta}_2$.

Apply the same procedure for $\boldsymbol{\beta}_1$ (and $\boldsymbol{\beta}_2$), and continue iteratively. Eventually we obtain a set of points $\{\boldsymbol{\beta}_i\}_{i=1}^l$ such that for each i, $\boldsymbol{\beta}_i = \lambda_i\boldsymbol{\beta}_j + \mu_i\boldsymbol{\beta}_k$ or $\boldsymbol{\beta}_i = \lambda_i\boldsymbol{\beta}_j + \mu_i\boldsymbol{\alpha}_k$ or $\boldsymbol{\beta}_i = \lambda_i\boldsymbol{\alpha}_j + \mu_i\boldsymbol{\alpha}_k$, where $\lambda_i + \mu_i = 1, \lambda_i, \mu_i > 0$. We claim the denominators of coordinates of $\boldsymbol{\beta}_i$ are odd numbers, and the numerators of coordinates of $\boldsymbol{\beta}_i$ are even numbers. This is because for each $\boldsymbol{\beta}_i$, we have the expression $\boldsymbol{\beta}_i = \sum_j \frac{s_j}{r}\boldsymbol{\alpha}_j$, where r is an odd number and all $\boldsymbol{\alpha}_j$'s are even lattice points. For $\boldsymbol{\beta}_i = \lambda\boldsymbol{\beta}_j + \mu\boldsymbol{\beta}_k$ (or $\boldsymbol{\beta}_i = \lambda\boldsymbol{\beta}_j + \mu\boldsymbol{\alpha}_k$, $\boldsymbol{\beta}_i = \lambda\boldsymbol{\alpha}_j + \mu\boldsymbol{\alpha}_k$ respectively), let M_i be the $\{\boldsymbol{\beta}_j, \boldsymbol{\beta}_k\}$- (or $\{\boldsymbol{\beta}_j, \boldsymbol{\alpha}_k\}$-, $\{\boldsymbol{\alpha}_j, \boldsymbol{\alpha}_k\}$- respectively) rational mediated set containing $\boldsymbol{\beta}_i$ obtained by Lemma 3.4 such that the denominators of coordinates of points in M_i are odd numbers and the numerators of coordinates of points in $M_i\backslash\{\boldsymbol{\beta}\}$ are even numbers for $i = 0, \ldots, l$ (set $\boldsymbol{\beta}_0 = \boldsymbol{\beta}$). Let $M_{\mathscr{A}\boldsymbol{\beta}} = \cup_{i=0}^l M_i$. Then $M_{\mathscr{A}\boldsymbol{\beta}}$ is clearly an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$ with the desired property. □

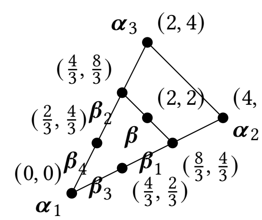
3.2 Decomposing SONC with binomial squares

For $r \in \mathbb{N}$ and $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$, let $f(\mathbf{x}^r) := f(x_1^r, \ldots, x_n^r)$. For any odd $r \in \mathbb{N}$, $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}}$ is a nonnegative circuit iff $f(\mathbf{x}^r) = \sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{r\boldsymbol{\alpha}} - d\mathbf{x}^{r\boldsymbol{\beta}}$ is a nonnegative circuit.

Theorem 3.6. *Let $f = \sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$ with $d \neq 0$ be a circuit polynomial. Assume that $M_{\mathscr{A}\boldsymbol{\beta}}$ is the $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$ provided by Lemma 3.5. and for each $\boldsymbol{u} \in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}$, let $\boldsymbol{u} = \frac{1}{2}(\boldsymbol{v_u}+\boldsymbol{w_u}), \boldsymbol{v_u} \neq \boldsymbol{w_u} \in M_{\mathscr{A}\boldsymbol{\beta}}$. Then f is nonnegative iff f can be written as $f = \sum_{\boldsymbol{u}\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w_u}})^2$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$.*

Proof. Assume that the least common multiple of denominators appearing in the coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}$ is r, which is odd. Then $f(\mathbf{x})$ is nonnegative if and only if $f(\mathbf{x}^r)$ is nonnegative. Multiply all coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}$ by r to remove the denominators, and the obtained $rM_{\mathscr{A}\boldsymbol{\beta}} := \{r\boldsymbol{u} \mid \boldsymbol{u} \in M_{\mathscr{A}\boldsymbol{\beta}}\}$ is an $r\mathscr{A}$-mediated set containing $r\boldsymbol{\beta}$. Hence by Theorem 3.1, $f(\mathbf{x}^r)$ is nonnegative if and only if $f(\mathbf{x}^r)$ can be written as $f(\mathbf{x}^r) = \sum_{\boldsymbol{u}\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{r}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{r}{2}\boldsymbol{w_u}})^2$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$, which is equivalent to $f(\mathbf{x}) = \sum_{\boldsymbol{u}\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w_u}})^2$. □

Example 3.7. Let $f = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ and $\mathscr{A} = \{\boldsymbol{\alpha}_1 = (0,0), \boldsymbol{\alpha}_2 = (4,2), \boldsymbol{\alpha}_3 = (2,4)\}$, $\boldsymbol{\beta} = (2,2)$. Let $\boldsymbol{\beta}_1 = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}\boldsymbol{\alpha}_2$ and $\boldsymbol{\beta}_2 = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}\boldsymbol{\alpha}_3$ such that $\boldsymbol{\beta} = \frac{1}{2}\boldsymbol{\beta}_1 + \frac{1}{2}\boldsymbol{\beta}_2$. Let $\boldsymbol{\beta}_3 = \frac{2}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\alpha}_2$ and $\boldsymbol{\beta}_4 = \frac{2}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\alpha}_3$. Then $M = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4\}$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$.



By Theorem 3.6, one has $f = x^4y^2+x^2y^4+1-3x^2y^2 = (a_1x^{\frac{2}{3}}y^{\frac{4}{3}} - b_1x^{\frac{4}{3}}y^{\frac{2}{3}})^2 + (a_2xy^2 - b_2x^{\frac{1}{3}}y^{\frac{2}{3}})^2 + (a_3x^{\frac{2}{3}}y^{\frac{4}{3}} - b_3)^2 + (a_4x^2y - b_4x^{\frac{2}{3}}y^{\frac{1}{3}})^2 + (a_5x^{\frac{4}{3}}y^{\frac{2}{3}} - b_5)^2$. Comparing coefficients yields $f = \frac{3}{2}(x^{\frac{2}{3}}y^{\frac{4}{3}} - x^{\frac{4}{3}}y^{\frac{2}{3}})^2 + (xy^2 - x^{\frac{1}{3}}y^{\frac{2}{3}})^2 + \frac{1}{2}(x^{\frac{2}{3}}y^{\frac{4}{3}} - 1)^2 + (x^2y - x^{\frac{2}{3}}y^{\frac{1}{3}})^2 + \frac{1}{2}(x^{\frac{4}{3}}y^{\frac{2}{3}} - 1)^2$, a sum of five binomial squares with rational exponents.

Lemma 3.8. *Let $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$. For an odd number r, $f(\mathbf{x}) \in$ SONC if and only if $f(\mathbf{x}^r) \in$ SONC.*

Proof. It comes from the fact that $f(\mathbf{x})$ is a nonnegative circuit iff $f(\mathbf{x}^r)$ is a nonnegative circuit for an odd number r. □

Theorem 3.9. *Let $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. Let $\mathscr{F}(\boldsymbol{\beta})$ be as in (1). For every $\boldsymbol{\beta} \in \Gamma(f)$ and every $\Delta \in \mathscr{F}(\boldsymbol{\beta})$, let $M_{\boldsymbol{\beta}\Delta}$ be the $V(\Delta)$-rational mediated set containing $\boldsymbol{\beta}$ provided by Lemma 3.5. Let $M = \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} M_{\boldsymbol{\beta}\Delta}$. For each $\boldsymbol{u} \in M\backslash\Lambda(f)$, let $\boldsymbol{u} = \frac{1}{2}(\boldsymbol{v_u} + \boldsymbol{w_u}), \boldsymbol{v_u} \neq \boldsymbol{w_u} \in M$. Let $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$. Then $f \in$ SONC iff f can be written as $f = \sum_{\boldsymbol{u}\in M\backslash\Lambda(f)} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w_u}})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$.*

Proof. Suppose $f \in$ SONC. By Theorem 5.5 in [27], we can write f as $f = \sum_{\boldsymbol{\beta}\in\Gamma(f)} \sum_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} f_{\boldsymbol{\beta}\Delta} + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$ such that every $f_{\boldsymbol{\beta}\Delta} = \sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}$ is a nonnegative circuit polynomial. We have $f_{\boldsymbol{\beta}\Delta} = \sum_{\boldsymbol{u}\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w_u}})^2$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$ by Theorem 3.6. Thus $f = \sum_{\boldsymbol{u}\in M\backslash\Lambda(f)} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w_u}})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$. Suppose $f = \sum_{\boldsymbol{u}\in M\backslash\Lambda(f)} (a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v_u}} -$

$b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w}_{\boldsymbol{u}}})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$. Assume that the least common multiple of denominators appearing in the coordinates of points in M is r, which is odd. Then $f(\mathbf{x}^r) = \sum_{\boldsymbol{u}\in M\backslash\Lambda(f)}(a_{\boldsymbol{u}}\mathbf{x}^{\frac{r}{2}\boldsymbol{v}_{\boldsymbol{u}}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{r}{2}\boldsymbol{w}_{\boldsymbol{u}}})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{r\boldsymbol{\alpha}}$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$, which is a SONC since every binomial square (and monomial square) is a nonnegative circuit. Hence by Lemma 3.8, $f(\mathbf{x}) \in$ SONC. □

4 SOC REPRESENTATIONS OF SONC CONES

SOCP plays an important role in convex optimization and can be handled via very efficient algorithms. If an SOC representation exists for a given convex cone, then it is possible to design efficient algorithms for optimization problems over the convex cone. In [8], Fawzi proved that PSD cones do not admit any SOC representations in general, which implies that SOS cones do not admit any SOC representations in general. In this section, we prove that dramatically unlike the SOS cones, SONC cones always admit SOC representations. Let $\boldsymbol{Q}^k := \boldsymbol{Q} \times \cdots \times \boldsymbol{Q}$ be the Cartesian product of k copies of an SOC $\boldsymbol{Q}$. A *linear slice* of $\boldsymbol{Q}^k$ is an intersection of $\boldsymbol{Q}^k$ with a linear subspace.

Definition 4.1. A convex cone $C \subseteq \mathbb{R}^m$ has a *SOC lift of size* k (or simply a *$\boldsymbol{Q}^k$-lift*) if it can be written as the projection of a slice of $\boldsymbol{Q}^k$, that is, there is a subspace L of $\boldsymbol{Q}^k$ and a linear map $\pi\colon \boldsymbol{Q}^k \to \mathbb{R}^m$ such that $C = \pi(\boldsymbol{Q}^k \cap L)$.

Definition 4.2. Given sets of lattice points $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \mathrm{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \mathrm{conv}(\mathscr{A}) \cap (\mathbb{N}^n\backslash(2\mathbb{N})^n)$ such that $\mathscr{A} \cap \mathscr{B}_1 = \varnothing$, define the SONC cone supported on $\mathscr{A}, \mathscr{B}_1, \mathscr{B}_2$ as

$$\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2} := \{(\mathbf{c}_{\mathscr{A}}, \mathbf{d}_{\mathscr{B}_1}, \mathbf{d}_{\mathscr{B}_2}) \in \mathbb{R}_+^{|\mathscr{A}|} \times \mathbb{R}_+^{|\mathscr{B}_1|} \times \mathbb{R}^{|\mathscr{B}_2|} \mid \sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\mathscr{B}_1\cup\mathscr{B}_2} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathrm{SONC}\},$$

where $\mathbf{c}_{\mathscr{A}} = (c_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathscr{A}}$, $\mathbf{d}_{\mathscr{B}_1} = (d_{\boldsymbol{\beta}})_{\boldsymbol{\beta}\in\mathscr{B}_1}$ and $\mathbf{d}_{\mathscr{B}_2} = (d_{\boldsymbol{\beta}})_{\boldsymbol{\beta}\in\mathscr{B}_2}$. It is easy to check that $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ is indeed a convex cone.

Let $\mathbb{S}_+^2$ be the convex cone of 2×2 positive semidefinite matrices

$$\mathbb{S}_+^2 := \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2} \mid \begin{bmatrix} a & b \\ b & c \end{bmatrix} \text{ is positive semidefinite}\right\}.$$

Lemma 4.3. *$\mathbb{S}_+^2$ is a 3-dimensional rotated SOC.*

Proof. It is immediate from the definition. □

Theorem 4.4. *For $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \mathrm{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \mathrm{conv}(\mathscr{A}) \cap (\mathbb{N}^n\backslash(2\mathbb{N})^n)$ such that $\mathscr{A} \cap \mathscr{B}_1 = \varnothing$, the convex cone $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ has an $(\mathbb{S}_+^2)^k$-lift for some $k \in \mathbb{N}$.*

Proof. For every $\boldsymbol{\beta} \in \mathscr{B}_1 \cup \mathscr{B}_2$, let $\mathscr{F}(\boldsymbol{\beta})$ be as in (1). Then for every $\boldsymbol{\beta} \in \mathscr{B}_1 \cup \mathscr{B}_2$ and every $\Delta \in \mathscr{F}(\boldsymbol{\beta})$, let $M_{\boldsymbol{\beta}\Delta}$ be the $V(\Delta)$-rational mediated set containing $\boldsymbol{\beta}$ provided by Lemma 3.5. Let $M = \cup_{\boldsymbol{\beta}\in\mathscr{B}_1\cup\mathscr{B}_2} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} M_{\boldsymbol{\beta}\Delta}$. For each $\boldsymbol{u}_i \in M\backslash\mathscr{A}$, let us write $\boldsymbol{u}_i = \frac{1}{2}(\boldsymbol{v}_i + \boldsymbol{w}_i)$. Let $B = \cup_{\boldsymbol{u}_i\in M\backslash\mathscr{A}}\{\frac{1}{2}\boldsymbol{v}_i, \frac{1}{2}\boldsymbol{w}_i\}$, $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$ and $k = \#M\backslash\mathscr{A} + \#\tilde{\mathscr{A}}$.

Then by Theorem 3.9, a polynomial f is in $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ if and only if f can be written as $f = \sum_{\boldsymbol{u}_i\in M\backslash\mathscr{A}}(a_i\mathbf{x}^{\frac{1}{2}\boldsymbol{v}_i} - b_i\mathbf{x}^{\frac{1}{2}\boldsymbol{w}_i})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$, $a_i, b_i \in \mathbb{R}$, which is equivalent to the existence of a symmetric matrix $Q = \sum_{i=1}^k Q_i$ such that $f = (\mathbf{x}^B)^T Q\mathbf{x}^B$ with $\mathbf{x}^B := (\mathbf{x}^{\boldsymbol{\beta}})_{\boldsymbol{\beta}\in B}$, where Q_i is a symmetric matrix with zeros everywhere except either at the four positions corresponding to the monomials $\mathbf{x}^{\frac{1}{2}\boldsymbol{v}_i}, \mathbf{x}^{\frac{1}{2}\boldsymbol{w}_i}$ or at the position corresponding to a monomial $\mathbf{x}^{\frac{1}{2}\boldsymbol{\alpha}}$ for some $\boldsymbol{\alpha} \in \tilde{\mathscr{A}}$. This leads respectively to either four entries forming a 2×2 positive semidefinite submatrix or one single positive entry.

Let $\pi : (\mathbb{S}_+^2)^k \to \mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ be the linear map that maps an element in $Q_1 \times \cdots \times Q_k$ to the coefficient vector of f which is in $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ via the equality $f = (\mathbf{x}^B)^T Q\mathbf{x}^B$ with $Q = \sum_{i=1}^k Q_i$. So we obtain an $(\mathbb{S}_+^2)^k$-lift for $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$. □

5 SONC OPTIMIZATION VIA SOCP

In this section, we tackle the following unconstrained polynomial optimization problem via SOCP, based on the representation of SONC cones derived in the previous section:

$$(\mathrm{P}): \quad \sup\{\xi : f(\mathbf{x}) - \xi \geq 0, \quad \mathbf{x} \in \mathbb{R}^n\}. \tag{3}$$

Let us denote by ξ^* the optimal value of (3). Replace the nonnegativity constraint in (3) by the following one to obtain a SONC relaxation with optimal value ξ_{sonc}:

$$(\mathrm{SONC}): \quad \sup\{\xi : f(\mathbf{x}) - \xi \in \mathrm{SONC}\}. \tag{4}$$

5.1 Conversion to PN-polynomials

Suppose $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. If $d_{\boldsymbol{\beta}} > 0$ for all $\boldsymbol{\beta} \in \Gamma(f)$, then we call f a *PN-polynomial*. The "PN" in PN-polynomial is short for "positive part plus negative part". For a PN-polynomial $f(\mathbf{x})$, it is clear that $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ iff $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}_+^n$.

Lemma 5.1. *Let $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ be a PN-polynomial. Then for any positive integer k, $f(\mathbf{x}) \in$ SONC if and only if $f(\mathbf{x}^k) \in$ SONC.*

Proof. It comes from the fact that a polynomial $f(\mathbf{x})$ with exactly one negative term is a nonnegative circuit iff $f(\mathbf{x}^k)$ is a nonnegative circuit for any positive integer $k \in \mathbb{N}$. □

Theorem 5.2. *Let $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$ be a PN-polynomial. Let $\mathscr{F}(\boldsymbol{\beta})$ be as in (1). For every $\boldsymbol{\beta} \in \Gamma(f)$ and every $\Delta \in \mathscr{F}(\boldsymbol{\beta})$, let $M_{\boldsymbol{\beta}\Delta}$ be a $V(\Delta)$-rational mediated set containing $\boldsymbol{\beta}$. Let $M = \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} M_{\boldsymbol{\beta}\Delta}$ and $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$. For each $\boldsymbol{u} \in M\backslash\Lambda(f)$, let $\boldsymbol{u} = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{w})$. Then $f \in$ SONC if and only if f can be written as $f = \sum_{\boldsymbol{u}\in M\backslash\Lambda(f)}(a_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v}} - b_{\boldsymbol{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{w}})^2 + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$, $a_{\boldsymbol{u}}, b_{\boldsymbol{u}} \in \mathbb{R}$.*

Proof. It follows easily from Lemma 5.1 and Theorem 3.1. □

The significant difference between Theorem 3.9 and Theorem 5.2 is that to represent a SONC PN-polynomial as a sum of binomial squares, we do not require the denominators of coordinates of points in $\mathscr{A}$-rational mediated sets to be odd. By virtue of this fact, for given trellis $\mathscr{A} = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m\}$ and lattice point $\boldsymbol{\beta} \in \mathrm{conv}(\mathscr{A})^\circ$, we can then construct an $\mathscr{A}$-rational mediated set $M_{\mathscr{A}\boldsymbol{\beta}}$ containing $\boldsymbol{\beta}$ which is smaller than that the one from Lemma 3.5.

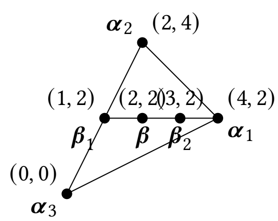
Lemma 5.3. *For a trellis $\mathscr{A}$ and a lattice point $\boldsymbol{\beta} \in \mathrm{conv}(\mathscr{A})^\circ$, there is an $\mathscr{A}$-rational mediated set $M_{\mathscr{A}\boldsymbol{\beta}}$ containing $\boldsymbol{\beta}$.*

Proof. Suppose that $\boldsymbol{\beta} = \sum_{i=1}^{m} \frac{q_i}{p} \boldsymbol{\alpha}_i$, where $p = \sum_{i=1}^{m} q_i$, $p, q_i \in \mathbb{N}^*$, $(p, q_1, \ldots, q_m) = 1$. We can write

$$\boldsymbol{\beta} = \frac{q_1}{p}\boldsymbol{\alpha}_1 + \frac{p-q_1}{p}\left(\frac{q_2}{p-q_1}\boldsymbol{\alpha}_2 + \cdots + \frac{q_m}{p-q_1}\boldsymbol{\alpha}_m\right).$$

Let $\boldsymbol{\beta}_1 = \frac{q_2}{p-q_1}\boldsymbol{\alpha}_2 + \cdots + \frac{q_m}{p-q_1}\boldsymbol{\alpha}_m$. Then $\boldsymbol{\beta} = \frac{q_1}{p}\boldsymbol{\alpha}_1 + \frac{p-q_1}{p}\boldsymbol{\beta}_1$. Apply the same procedure for $\boldsymbol{\beta}_1$, and continue like this. Eventually we obtain a set of points $\{\boldsymbol{\beta}_i\}_{i=0}^{m-2}$ (set $\boldsymbol{\beta}_0 = \boldsymbol{\beta}$) such that $\boldsymbol{\beta}_i = \lambda_i\boldsymbol{\alpha}_{i+1} + \mu_i\boldsymbol{\beta}_{i+1}$, $i = 0, \ldots, m-3$ and $\boldsymbol{\beta}_{m-2} = \lambda_{m-2}\boldsymbol{\alpha}_{m-1} + \mu_{m-2}\boldsymbol{\alpha}_m$, where $\lambda_i + \mu_i = 1, \lambda_i, \mu_i > 0, i = 0, \ldots, m-2$. For $\boldsymbol{\beta}_i = \lambda_i\boldsymbol{\alpha}_{i+1} + \mu_i\boldsymbol{\beta}_{i+1}$ (resp. $\boldsymbol{\beta}_{m-2} = \lambda_{m-2}\boldsymbol{\alpha}_{m-1} + \mu_{m-2}\boldsymbol{\alpha}_m$), let M_i be the $\{\boldsymbol{\alpha}_{i+1}, \boldsymbol{\beta}_{i+1}\}$- (resp. $\{\boldsymbol{\alpha}_{m-1}, \boldsymbol{\alpha}_m\}$-) rational mediated set containing $\boldsymbol{\beta}_i$ obtained by Lemma 3.4, $i = 0, \ldots, m-2$. Let $M_{\mathscr{A}\boldsymbol{\beta}} = \cup_{i=0}^{m-2} M_i$. Then clearly $M_{\mathscr{A}\boldsymbol{\beta}}$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$. □

Example 5.4. Let $f = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ be the Motzkin's polynomial and $\mathscr{A} = \{\boldsymbol{\alpha}_1 = (4,2), \boldsymbol{\alpha}_2 = (2,4), \boldsymbol{\alpha}_3 = (0,0)\}$, $\boldsymbol{\beta} = (2,2)$. Then $\boldsymbol{\beta} = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\alpha}_2 + \frac{1}{3}\boldsymbol{\alpha}_3 = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}(\frac{1}{2}\boldsymbol{\alpha}_2 + \frac{1}{2}\boldsymbol{\alpha}_3)$. Let $\boldsymbol{\beta}_1 = \frac{1}{2}\boldsymbol{\alpha}_2 + \frac{1}{2}\boldsymbol{\alpha}_3$ such that $\boldsymbol{\beta} = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}\boldsymbol{\beta}_1$. Let $\boldsymbol{\beta}_2 = \frac{2}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\beta}_1$. Then it is easy to check that $M = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$.



By a simple computation, we have $f = (1 - xy^2)^2 + 2(x^{\frac{1}{2}}y - x^{\frac{3}{2}}y)^2 + (xy - x^2y)^2$. Here we represent f as a sum of three binomial squares with rational exponents.

We associate to a polynomial $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$, the PN-polynomial $\tilde{f} = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} |d_{\boldsymbol{\beta}}|\mathbf{x}^{\boldsymbol{\beta}}$.

Lemma 5.5. *Suppose $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. If $\tilde{f}$ is nonnegative, then f is nonnegative. Moreover, $\tilde{f} \in$ SONC if and only if $f \in$ SONC.*

Proof. For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} f(\mathbf{x}) &= \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \\ &\geq \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}|\mathbf{x}|^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} |d_{\boldsymbol{\beta}}||\mathbf{x}|^{\boldsymbol{\beta}} = \tilde{f}(|\mathbf{x}|), \end{aligned}$$

where $|\mathbf{x}| = (|x_1|, \ldots, |x_n|)$. It follows that the nonnegativity of $\tilde{f}$ implies the nonnegativity of f.

For every $\boldsymbol{\beta} \in \Gamma(f)$, let $\mathscr{F}(\boldsymbol{\beta})$ be as in (1). Let $\mathscr{B} = \{\boldsymbol{\beta} \in \Gamma(f) \mid \boldsymbol{\beta} \notin (2\mathbb{N})^n \text{ and } d_{\boldsymbol{\beta}} < 0\}$ and $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$. Assume $\tilde{f} \in$ SONC. Then we can write

$$\begin{aligned} \tilde{f} = &\sum_{\boldsymbol{\beta}\in\Gamma(f)\backslash\mathscr{B}} \sum_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} \Big(\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}\Big) \\ &+ \sum_{\boldsymbol{\beta}\in\mathscr{B}} \sum_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} \Big(\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \tilde{d}_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}\Big) + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \end{aligned}$$

s.t. each $\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}$ and each $\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \tilde{d}_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}$ are nonnegative circuit polynomials. Note that $\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} + \tilde{d}_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}$ is also a nonnegative circuit polynomial and $\sum_{\Delta\in\Delta(\boldsymbol{\beta})} \tilde{d}_{\boldsymbol{\beta}\Delta} = |d_{\boldsymbol{\beta}}| = -d_{\boldsymbol{\beta}}$ for any $\boldsymbol{\beta} \in \mathscr{B}$. Hence,

$$\begin{aligned} f = &\sum_{\boldsymbol{\beta}\in\Gamma(f)\backslash\mathscr{B}} \sum_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} \Big(\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}\Big) \\ &+ \sum_{\boldsymbol{\beta}\in\mathscr{B}} \sum_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} \Big(\sum_{\boldsymbol{\alpha}\in V(\Delta)} c_{\boldsymbol{\beta}\Delta\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} + \tilde{d}_{\boldsymbol{\beta}\Delta}\mathbf{x}^{\boldsymbol{\beta}}\Big) + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \in \text{SONC}. \end{aligned}$$

The inverse follows similarly. □

Hence by Lemma 5.5, if we replace the polynomial f in (4) by its associated PN-polynomial $\tilde{f}$, then this does not affect the optimal value of (4):

$$\text{(SONC-PN)}: \quad \sup\{\xi : \tilde{f}(\mathbf{x}) - \xi \in \text{SONC}\}. \tag{5}$$

5.2 Compute a simplex cover

Given a polynomial $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$, in order to obtain a SONC decomposition of f, we use all simplices containing $\boldsymbol{\beta}$ for each $\boldsymbol{\beta} \in \Gamma(f)$ in Theorem 3.9. In practice, we do not need that many simplices. A recent study [21] proposes a systematic method to compute an optimal simplex cover. It would be worth trying to combine this framework with our SOC characterization for SONC cones to achieve a more accurate algorithm. Here we rely on a heuristics to compute a set of simplices with vertices coming from $\Lambda(f)$ and that covers $\Gamma(f)$. For $\boldsymbol{\beta} \in \Gamma(f)$ and $\boldsymbol{\alpha}_0 \in \Lambda(f)$, define an auxiliary linear program:

$$\begin{aligned} &\text{SimSel}(\boldsymbol{\beta}, \Lambda(f), \boldsymbol{\alpha}_0) = \text{Argmax} \quad \lambda_{\boldsymbol{\alpha}_0} \\ &\text{s.t.}\Big\{\sum_{\boldsymbol{\alpha}\in\Lambda(f)} \lambda_{\boldsymbol{\alpha}} \cdot \boldsymbol{\alpha} = \boldsymbol{\beta}, \sum_{\boldsymbol{\alpha}\in\Lambda(f)} \lambda_{\boldsymbol{\alpha}} = 1, \lambda_{\boldsymbol{\alpha}} \geq 0, \forall \boldsymbol{\alpha} \in \Lambda(f)\Big\}. \end{aligned}$$

Following [25], we can ensure the output of $\text{SimSel}(\boldsymbol{\beta}, \Lambda(f), \boldsymbol{\alpha}_0)$ corresponds to a trellis which contains $\boldsymbol{\alpha}_0$ and covers $\boldsymbol{\beta}$. The so-called `SimplexCover`[1] procedure computes such a simplex cover.

Let $\mathbf{K}$ be the 3-dimensional rotated SOC, i.e.,

$$\mathbf{K} := \{(a, b, c) \in \mathbb{R}^3 \mid 2ab \geq c^2, a \geq 0, b \geq 0\}. \tag{6}$$

Suppose $\tilde{f} = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. By algorithm `SimplexCover`, we compute a simplex cover $\{(\mathscr{A}_k, \boldsymbol{\beta}_k)\}_{k=1}^{l}$. For each k, let M_k be an $\mathscr{A}_k$-rational mediated set containing $\boldsymbol{\beta}_k$ and $s_k = \#M_k\backslash\mathscr{A}_k$. For each $\boldsymbol{u}_i^k \in M_k\backslash\mathscr{A}_k$, let us write $\boldsymbol{u}_i^k = \frac{1}{2}(\boldsymbol{v}_i^k + \boldsymbol{w}_i^k)$. Let $\tilde{\mathscr{A}} = \{\boldsymbol{\alpha} \in \Lambda(f) \mid \boldsymbol{\alpha} \notin \cup_{\boldsymbol{\beta}\in\Gamma(f)} \cup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})} V(\Delta)\}$. Then we can relax (SONC-PN) to an SOCP problem (SONC-SOCP) as follows:

$$\begin{cases} \sup & \xi \\ \text{s.t.} & \tilde{f}(\mathbf{x}) - \xi = \sum_{k=1}^{l}\sum_{i=1}^{s_k}(2a_i^k\mathbf{x}^{\boldsymbol{v}_i^k} + b_i^k\mathbf{x}^{\boldsymbol{w}_i^k} - 2c_i^k\mathbf{x}^{\boldsymbol{u}_i^k}) + \sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}, \\ & (a_i^k, b_i^k, c_i^k) \in \mathbf{K}, \quad \forall i, k. \end{cases} \tag{7}$$

Let us denote by ξ_{socp} the optimal value of (7). Then, we have $\xi_{socp} \leq \xi_{sonc} \leq \xi^*$.

Remark 5.6. *The quality of obtained SONC lower bounds depends on two successive steps: the relaxation to the corresponding PN-polynomial (from ξ^* to ξ_{sonc}) and the relaxation to a specific simplex cover (from ξ_{sonc} to ξ_{socp}). The loss of bound-quality at the*

[1] Algorithm 4 in https://arxiv.org/abs/1906.06179

second step can be improved by choosing a more optimal simplex cover. Nevertheless, it may happen that the loss of bound-quality at the first step is already big, as shown in Example 5.7, which indicates that the gap between nonnegative polynomials and SONC PN-polynomials may greatly affect the quality of SONC lower bounds.

Example 5.7. Let $f = 1 + x_1^4 + x_2^4 - x_1x_2^2 - x_1^2x_2 + 5x_1x_2$. Since $\Lambda(f)$ forms a trellis, the simplex cover for f is unique. One obtains $\xi_{socp} = \xi_{sonc} \approx -6.916501$ while $\xi^* \approx -2.203372$. Hence the relative optimality gap is near 214%.

6 NUMERICAL EXPERIMENTS

Here, we present numerical results of the proposed algorithms for unconstrained POPs. Our tool, called `SONCSOCP`, implements the simplex cover algorithm as well as a procedure `MedSet`[2] computing the rational mediated set and computes the optimal value ξ_{socp} of the SOCP (7) with Mosek [3]. All experiments were performed on an Intel Core i5-8265U@1.60GHz CPU with 8GB RAM memory and WINDOWS 10 system. `SONCSOCP` is available at github:`SONCSOCP`.

Our benchmarks are issued from the database of randomly generated polynomials provided by Seidler and de Wolff in [25]. Depending on the Newton polytope, these benchmarks are divided into three classes: the ones with standard simplices, the ones with general simplices and the ones with arbitrary Newton polytopes. (We use n, d, t, l to denote the number of variables, the degree, the number of terms and the lower bound on the number of inner terms respectively. See [25] for the details on the construction of these polynomials). We compare the performance of `SONCSOCP` with the ones of `POEM`, which relies on the ECOS solver to solve geometric programs (see [25] for more details). To measure the quality of a given lower bound ξ_{lb}, we rely on the '`local_min`' function available in `POEM` which computes an upper bound ξ_{min} on the minimum of a polynomial. The relative optimality gap is defined by $\frac{|\xi_{min}-\xi_{lb}|}{|\xi_{min}|}$. In the following tables, the column 'time' is the running time in seconds and the column 'opt' the optimal value.

Standard simplex. For the standard simplex case, we take 10 polynomials of different types (labeled by N). Running time and lower bounds obtained with `SONCSOCP` and `POEM` are displayed in Table 1. Note that for polynomials with $\Lambda(\cdot)$ forming a trellis, the simplex cover is unique, thus the bounds obtained by `SONCSOCP` and `POEM` are the same theoretically, which is also reflected in Table 1. For each polynomial, the relative optimality gap is less than 1% and for 8 out of 10 polynomials, it is less than 0.1% (see Figure 2).

	N	1	2	3	4	5	6	7	8	9	10
	n	10	10	10	20	20	20	30	30	40	40
	d	40	50	60	40	50	60	50	60	50	60
	t	20	20	20	30	30	30	50	50	100	100
time	SONCSOCP	0.04	0.04	0.04	0.14	0.14	0.13	0.43	0.40	2.23	2.21
	POEM	0.26	0.27	0.26	0.43	0.44	0.42	1.78	1.79	2.20	2.25
opt	SONCSOCP	3.52	3.52	3.52	2.64	2.64	2.64	2.94	2.94	4.41	4.41
	POEM	3.52	3.52	3.52	2.64	2.64	2.64	2.94	2.94	4.41	4.41

Table 1: Results for the standard simplex case

General simplex. Here, we take 10 polynomials of different types (labeled by N). Running time and lower bounds obtained with `SONCSOCP` and `POEM` are displayed in Table 2. As before, the SONC lower bounds obtained by `SONCSOCP` and `POEM` are the same.

[2] Algorithm 3 in https://arxiv.org/abs/1906.06179

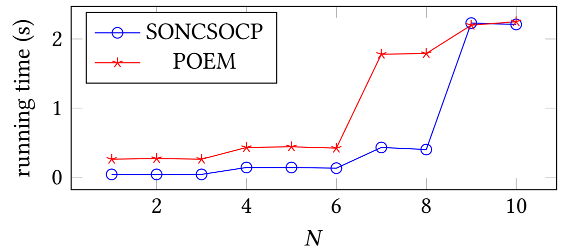


Figure 1: Running time for the standard simplex case

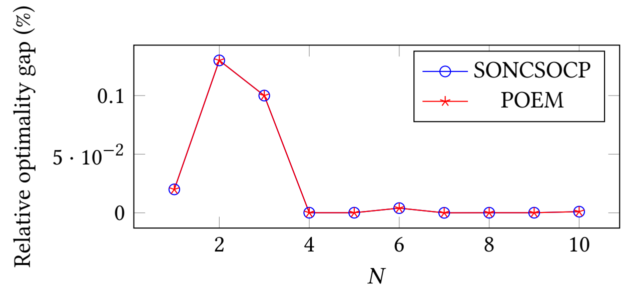


Figure 2: Relative optimality gap for the standard simplex case

For each polynomial except for the one corresponding to $N = 7$, the relative optimality gap is within 30%, and for 6 out of 10 polynomials, the gap is below 1% (see Figure 4). `POEM` fails to obtain a lower bound for the instance $N = 10$ by returning −Inf. Figure 3 shows that, overall, the running times of `SONCSOCP` and `POEM` are close. `SONCSOCP` is faster than `POEM` for the instance $N = 6$, possibly because better performance are obtained when the degree is relatively low.

	N	1	2	3	4	5	6	7	8	9	10
	n	10	10	10	10	10	10	10	10	10	10
	d	20	30	40	50	60	20	30	40	50	60
	t	20	20	20	20	20	30	30	30	30	30
time	SONCSOCP	0.32	0.29	0.36	0.48	0.54	0.56	0.73	0.88	1.04	1.04
	POEM	0.28	0.31	0.31	0.31	0.43	0.74	0.75	0.74	0.72	0.76
opt	SONCSOCP	1.18	0.22	0.38	0.90	0.06	4.00	−4.64	1.62	2.95	5.40
	POEM	1.18	0.22	0.38	0.90	0.06	4.00	−4.64	1.62	2.95	−Inf

Table 2: Results for the general simplex case

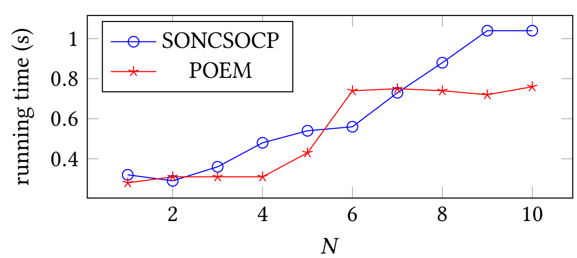


Figure 3: Running time for the general simplex case

Arbitrary polytope. We take 20 polynomials of different types (labeled by N). `POEM` always throws an error "expected square matrix". Running time and lower bounds obtained with `SONCSOCP` are displayed in Table 3. The relative optimality gap is always within 25% and within 1% for 17 out of 20 polynomials (see Figure 5).

7 CONCLUSIONS

In this paper, we provide a constructive proof that each SONC cone admits an SOC representation. Based on this, we propose an

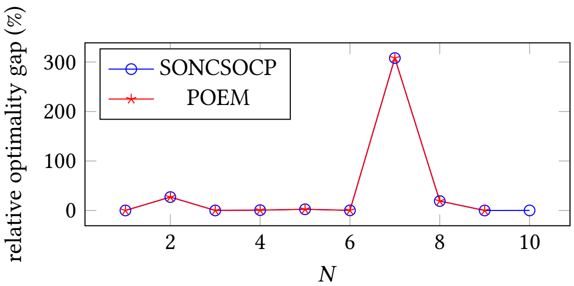


Figure 4: Relative optimality gap for the general simplex case

N		1	2	3	4	5	6	7	8	9	10
n		10	10	10	10	10	10	10	10	10	10
d		20	20	20	30	30	30	40	40	40	50
t		30	100	300	30	100	300	30	100	300	30
l		15	71	231	15	71	231	15	71	231	15
SONCSOCP	time	0.38	1.75	6.86	0.64	3.13	11.3	0.72	4.01	14.6	0.76
	opt	0.70	3.32	31.7	3.31	15.3	3.31	0.47	5.42	38.7	1.56
N		11	12	13	14	15	16	17	18	19	20
n		10	10	10	10	10	20	20	20	20	20
d		50	50	60	60	60	30	30	40	40	40
t		100	300	30	100	300	50	100	50	100	200
l		71	231	15	71	231	5	15	5	15	35
SONCSOCP	time	4.41	16.8	1.84	11.2	42.4	3.20	8.84	2.60	10.5	38.7
	opt	0.20	7.00	3.31	2.52	23.4	0.70	4.91	4.13	2.81	9.97

Table 3: Results for the arbitrary polytope case

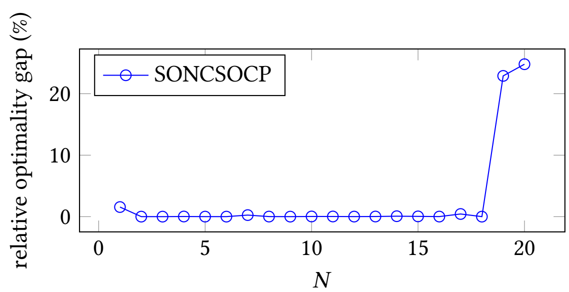


Figure 5: Relative optimality gap for the arbitrary polytope case

algorithm to compute a lower bound for unconstrained POPs via SOCP. Numerical experiments demonstrate the efficiency of our algorithm even when the number of variables and the degree are fairly large. Even though the complexity of our algorithm depends on the degree in theory, it turns out that this dependency is rather mild. For all numerical examples tested in this paper, the running time is below one minute even for polynomials of degree up to 60. Since the running time is satisfactory, the main concern of SONC-based algorithms for sparse polynomial optimization may be the quality of obtained lower bounds. For many examples tested in this paper, the relative optimality gap is within 1%. However, it can happen that the SONC lower bound is not accurate and this cannot be avoided by choosing an optimal simplex cover. To improve the quality of such bounds, it is mandatory to find more complex representations of nonnegative polynomials, which involve SONC polynomials. We also plan to design a rounding-projection procedure, in the spirit of [22], to obtain exact nonnegativity certificates for polynomials lying in the interior of the SONC cone. A related investigation track is the complexity analysis and software implementation of the resulting hybrid numeric-symbolic scheme, as well as performance comparisons with concurrent methods based on semidefinite programming [16] or geometric programming [19].

REFERENCES

[1] Amir Ali Ahmadi and Anirudha Majumdar. 2019. DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. *SIAM Journal on Applied Algebra and Geometry* 3, 2 (2019), 193–230.

[2] Farid Alizadeh and Donald Goldfarb. 2003. Second-order cone programming. *Mathematical programming* 95, 1 (2003), 3–51.

[3] E. D. Andersen and K. D. Andersen. 2000. The Mosek Interior Point Optimizer for Linear Programming: An Implementation of the Homogeneous Algorithm. In *High Performance Optimization*, Hans Frenk, Kees Roos, Tamás Terlaky, and Shuzhong Zhang (Eds.). Applied Optimization, Vol. 33. Springer US, 197–232. https://doi.org/10.1007/978-1-4757-3216-0_8

[4] Gennadiy Averkov. 2019. Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization. *SIAM Journal on Applied Algebra and Geometry* 3, 1 (2019), 128–151.

[5] Ahron Ben-Tal and Arkadi Nemirovski. 2001. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications.* Vol. 2. Siam.

[6] V. Chandrasekaran and P. Shah. 2016. Relative Entropy Relaxations for Signomial Optimization. *SIAM J. Optim.* 26, 2 (2016), 1147–1173.

[7] Mareike Dressler, Sadik Iliman, and Timo De Wolff. 2019. An approach to constrained polynomial optimization via nonnegative circuit polynomials and geometric programming. *Journal of Symbolic Computation* 91 (2019), 149–172.

[8] Hamza Fawzi. 2019. On representing the positive semidefinite cone using the second-order cone. *Mathematical Programming* 175, 1-2 (2019), 109–118.

[9] Jacob Hartzer, Olivia Röhrig, Timo de Wolff, and Oğuzhan Yürük. 2019. Initial Steps in the Classification of Maximal Mediated Sets. *arXiv preprint arXiv:1910.00502* (2019).

[10] Sadik Iliman and Timo De Wolff. 2016. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Research in the Mathematical Sciences* 3, 1 (2016), 9.

[11] Sadik Iliman and Timo De Wolff. 2016. Lower bounds for polynomials with simplex newton polytopes based on geometric programming. *SIAM Journal on Optimization* 26, 2 (2016), 1128–1146.

[12] C. Josz. 2016. *Application of polynomial optimization to electricity transmission networks*. Theses. Université Pierre et Marie Curie - Paris VI. https://tel.archives-ouvertes.fr/tel-01478431

[13] L. Katthan, H. Naumann, and T. Theobald. 2019. A Unified framework of SAGE and SONC polynomials and its duality theory. *arXiv preprint arXiv:1903.08966* (2019).

[14] I. Klep, V. Magron, and J. Povh. 2019. Sparse Noncommutative Polynomial Optimization. *arXiv preprint arXiv:1909.00569* (2019).

[15] V. Magron, G. Constantinides, and A. Donaldson. 2017. Certified Roundoff Error Bounds Using Semidefinite Programming. *ACM Trans. Math. Softw.* 43, 4, Article 34 (2017), 34 pages.

[16] V. Magron and M. Safey El Din. 2018. On Exact Polya and Putinar's Representations. In *ISSAC'18: Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*. ACM, New York, NY, USA.

[17] V. Magron and M. Safey El Din. 2018. RealCertify: a Maple package for certifying non-negativity. In *ISSAC'18: Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*. ACM, New York, NY, USA.

[18] Victor Magron, Mohab Safey El Din, and Markus Schweighofer. 2019. Algorithms for weighted sum of squares decomposition of non-negative univariate polynomials. *Journal of Symbolic Computation* 93 (2019), 200–220.

[19] Victor Magron, Henning Seidler, and Timo de Wolff. 2019. Exact Optimization via Sums of Nonnegative Circuits and Arithmetic-Geometric-Mean-Exponentials. In *Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation* (Beijing, China) *(ISSAC '19)*. New York, NY, USA, 291–298.

[20] Riley Murray, Venkat Chandrasekaran, and Adam Wierman. 2018. Newton polytopes and relative entropy optimization. *arXiv preprint arXiv:1810.01614* (2018).

[21] Dávid Papp. 2019. Duality of sum of nonnegative circuit polynomials and optimal SONC bounds. *arXiv preprint arXiv:1912.04718* (2019).

[22] H. Peyrl and P.A. Parrilo. 2008. Computing sum of squares decompositions with rational coefficients. *Theoretical Computer Science* 409, 2 (2008), 269–281.

[23] Victoria Powers and Bruce Reznick. 2019. A note on mediated simplices. *arXiv preprint arXiv:1909.11008* (2019).

[24] Bruce Reznick. 1989. Forms derived from the arithmetic-geometric inequality. *Math. Ann.* 283, 3 (1989), 431–464.

[25] Henning Seidler and Timo de Wolff. 2018. An experimental comparison of sonc and sos certificates for unconstrained optimization. *arXiv preprint arXiv:1808.08431* (2018).

[26] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. 2006. Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity. *SIAM Journal on Optimization* 17, 1 (2006), 218–242.

[27] J. Wang. 2018. Nonnegative polynomials and circuit polynomials. *arXiv preprint arXiv:1804.09455* (2018).

[28] J. Wang, V. Magron, and J.-B. Lasserre. 2019. TSSOS: a moment-SOS hierarchy that exploits term sparsity. *arXiv preprint arXiv:1912.08899* (2019).