



Convex hulls of monomial curves, and a sparse positivstellensatz

Gennadiy Averkov¹ · Claus Scheiderer²

Received: 7 March 2023 / Accepted: 11 January 2024
© The Author(s) 2024

Abstract

Consider the closed convex hull K of a monomial curve given parametrically as $(t^{m_1}, \dots, t^{m_n})$, with the parameter t varying in an interval I . We show, using constructive arguments, that K admits a lifted semidefinite description by $\mathcal{O}(d)$ linear matrix inequalities (LMIs), each of size $\lfloor \frac{n}{2} \rfloor + 1$, where $d = \max\{m_1, \dots, m_n\}$ is the degree of the curve. On the dual side, we show that if a univariate polynomial $p(t)$ of degree d with at most $2k + 1$ monomials is non-negative on \mathbb{R}_+ , then p admits a representation $p = t^0\sigma_0 + \dots + t^{d-k}\sigma_{d-k}$, where the polynomials $\sigma_0, \dots, \sigma_{d-k}$ are sums of squares and $\deg(\sigma_i) \leq 2k$. The latter is a univariate positivstellensatz for sparse polynomials, with non-negativity of p being certified by sos polynomials whose degree only depends on the sparsity of p . Our results fit into the general attempt of formulating polynomial optimization problems as semidefinite problems with LMIs of small size. Such small-size descriptions are much more tractable from a computational viewpoint.

Mathematics Subject Classification 90C23 · 90C22 · 13J30 · 06E05

1 Introduction

Polynomial optimization studies the task of minimizing an n -variate polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ over a semialgebraic set $S \subseteq \mathbb{R}^n$. The convexification approach to this problem consists of picking finite sets $A \supseteq \text{supp}(p)$ of \mathbb{Z}_+^n , considering the monomials $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ with $\alpha \in A$ and finding exact or approximate descriptions of the

✉ Gennadiy Averkov
averkov@b-tu.de

Claus Scheiderer
claus.scheiderer@uni-konstanz.de

¹ Brandenburgische Technische Universität Cottbus-Senftenberg, Platz der Deutschen Einheit 1, 03046 Cottbus, Germany

² Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany

closed convex hulls

$$K = \overline{\text{conv}\{(x^\alpha)_{\alpha \in A} : x \in S\}}$$

within some optimization paradigm. The basic idea is that optimizing a polynomial $p = c_0 + \sum_{\alpha \in A} c_\alpha x^\alpha \in \mathbb{R}[x_1, \dots, x_n]$ means to optimize the affine-linear function $\lambda : (y_\alpha)_{\alpha \in A} \mapsto c_0 + \sum_{\alpha \in A} c_\alpha y_\alpha$ over K . Usually, descriptions of the sets K arise from *positivstellensätze* from real algebra, since non-negativity of λ on K corresponds to non-negativity of p on S . Most *positivstellensätze* in real algebra employ sum-of-squares (sos) certificates. This fact establishes a connection from real algebra to semidefinite optimization, since the cone $\Sigma_{n,2d}$ of n -variate sos polynomials of degree $\leq 2d$ is a linear image of the cone of symmetric positive semidefinite matrices of size $\binom{n+d}{n}$. As suggested by Lasserre in his seminal work [13] and following this line of ideas, Putinar's *positivstellensatz* [17] leads to a hierarchy of outer approximations of K in terms of semidefinite constraints. Although Lasserre's approach provides a universal template for converting polynomial problems to semidefinite ones, further adjustment is usually needed to make it computationally tractable. Despite the fact that semidefinite programs are solvable in polynomial time under mild assumptions within a given error tolerance [3, 16], solving a semidefinite program can quickly become extremely challenging in practice since the size of the semidefinite constraints is a critical parameter. We view a semidefinite problem as the problem of optimization of a linear function subject to finitely many linear matrix inequalities (LMIs). An LMI of size d is the condition $M(y) \geq 0$ that imposes semidefiniteness of a symmetric $d \times d$ matrix $M(y)$ whose entries are affine-linear functions in the variables $y = (y_1, \dots, y_N)$. To get a first impression of how the workload is increased by large size of the LMIs, consider the approximation of a semidefinite problem $\inf\{L(y) : M_1(y) \geq 0, \dots, M_\ell(y) \geq 0\}$ by the convex problem $L(y) - \epsilon \sum_{i=1}^\ell \log \det M_i(y)$ that uses the logarithmic barrier and a small parameter $\epsilon > 0$. Solving the convex problem with the gradient descent method would involve computing the gradients of the barriers, which requires to invert the matrices $M_i(y)$ [6, 4.3.1]. But inverting $M_i(y)$ of a large size is known to be expensive. So, while in theoretical considerations it is customary to model ℓ LMIs as a single LMI of size ℓd , using a block-diagonal matrix, this reduction conceals the aspect of efficiency, since a general LMI of size ℓd has a much higher computational cost than ℓ LMIs of size d .

To address this issue, the following terminology was introduced in [4]:

Definition 1.1 Let $C \subseteq \mathbb{R}^n$ be a set. A description of C in the form

$$C = \{x \in \mathbb{R}^n : M_1(x, y) \geq 0, \dots, M_\ell(x, y) \geq 0 \text{ for some } y \in \mathbb{R}^m\} \quad (1)$$

where the M_i are LMIs in (x, y) , is called a *lifted* (or *extended*) *semidefinite representation* of C . If every M_i has size $\leq d$, we say that (1) is a lifted representation of size $\leq d$. The minimal d for which C admits a lifted representation of size $\leq d$ is the *semidefinite extension degree* of C , denoted $\text{sxdeg}(C)$. If no such d exists one puts $\text{sxdeg}(C) = \infty$.

By studying $\text{sxddeg}(C)$ one keeps track of the size of the LMIs needed to optimize linear functions over C , disregarding their number. For computations, both size and number of the LMIs play a role, but size is more critical as a parameter. Clearly $\text{sxddeg}(C) = 1$ if and only if C is a polyhedron, and it is easy to see that $\text{sxddeg}(C) \leq 2$ if and only if C is second-order cone representable [9]. In [4] it was shown that the cone $\Sigma_{n,2d}$ of n -variate sos polynomials of degree $\leq 2d$ has $\text{sxddeg} = \binom{n+d}{n}$, matching the size of the LMIs in Lasserre's approach. The rapid growth of $\binom{n+d}{n}$ in n and d explains why the computational time needed to solve Lasserre's relaxations is extremely sensitive to the choice of n and d . This issue with the size of the LMIs was detected by a number of researchers. Some suggestions for how to cope with it in practice were made in [1, 5, 23, 24]. The common general idea in [1, 5, 23, 24] is to deliberately choose a cone C , which has small $\text{sxddeg}(C)$ by its construction, to serve as a tractable outer approximation of K . Or, from the dual viewpoint, to choose a cone P of non-negative polynomials of more specific structure than in Putinar's positivstellensatz, such that P has small $\text{sxddeg}(P)$. While these approaches seem to help in practice, they are purely heuristic, since positivstellensätze are still missing that would guarantee that positivity can indeed always be certified in the intended, computationally less expensive, manner.

In this paper, we make a first step in filling this gap by studying the size of semidefinite representations for sparse polynomial optimization problems in one variable. Consider a set $A = \{m_1, \dots, m_n\}$ of positive integers $m_1 > \dots > m_n > 0$ and a non-degenerate interval $I \subseteq \mathbb{R}$. For optimizing the univariate polynomial $p = \sum_{a \in A} c_a t^a \in \mathbb{R}[t]$ on the interval I , we are interested in finding a lifted semidefinite representation of $K = \overline{\text{conv}\{(t^{m_1}, \dots, t^{m_n}) : t \in I\}} \subseteq \mathbb{R}^n$. Since the curve $(t^{m_1}, \dots, t^{m_n})$ is a projection of the rational normal curve $(t^i)_{1 \leq i \leq m_1}$ of degree m_1 , Lasserre's approach gives a description of size $\lfloor \frac{m_1}{2} \rfloor + 1$. When the number $n = |A|$ of monomials is small compared to the degree m_1 , it is desirable to find an alternative description of smaller size. Our main result (Theorem 3.19) shows that the semidefinite extension degree of such K is at most $\lfloor \frac{n}{2} \rfloor + 1$, which is the best possible bound. Consequently, the size of the description depends only on the number of monomials, and not on the degree. The description is completely explicit and follows from a sparse positivstellensatz (Theorem 3.9). The latter characterizes non-negativity of a degree d polynomial p with at most $2k + 1$ terms on an interval $I \subseteq \mathbb{R}_+$, by using the cone $t^0 \Sigma_{1,2k} + \dots + t^{d-2k} \Sigma_{1,2k}$, which is of semidefinite extension degree $k + 1$. The key technical ingredient in our proof is Jacobi's bialternant formula for Schur polynomials from the theory of symmetric polynomials.

Our result about non-negativity of univariate sparse polynomials should help to understand the impact of sparsity-based approaches to optimization problems of arbitrary dimension. Univariate problems may, on the one hand, demonstrate phenomena that occur in every dimension. On the other, whenever some limitation can be identified in the one-dimensional case, it will also be present in a certain form in an n -dimensional setting as well. Furthermore, our main result can serve as a starting point for further studies in several directions. It would obviously be interesting to obtain similar results in the multivariate case. While our K is given as the closed convex hull of a monomial curve, one may ask for bounds on the semidefinite extension degree of closed convex

hulls K of arbitrary semialgebraic curves $S \subseteq \mathbb{R}^n$. For $n = 2$, it was proved in [19] that every closed convex semialgebraic set K in the plane is second-order cone representable, i.e. has $\text{sxdeg}(K) \leq 2$. It can be shown that the bound $\text{sxdeg}(K) \leq \lfloor \frac{n}{2} \rfloor + 1$ holds for the closed convex hull of an arbitrary semialgebraic curve $S \subseteq \mathbb{R}^n$, but the proof is much more involved [20].

The paper is organized as follows. After a few preliminaries in Sect. 2, the main result is obtained in Sect. 3. In Sect. 4 we introduce the cones of sums of copositive fewnomials, which form a sparse counterpart of the cones of non-negative polynomials. We show that in the one-dimensional case these cones admit semidefinite descriptions with LMIs of small size, and explain that such descriptions lead to variations of Lasserre's relaxations that are based on LMIs of small size.

After this paper was submitted, Philipp di Dio pointed out to us that our Lemma 3.4 is contained, in much greater generality, in work of Karlin and Studden on T -systems. We are grateful for this hint and refer to [11] Chapter 2 for more details.

2 Preliminaries

2.1. Let $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{R}_+ = [0, \infty[$. The cardinality of a set A is denoted by $|A|$. For a tuple $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n$ we use the notation $|\alpha| := \alpha_1 + \dots + \alpha_n$. For univariate polynomials we mostly use t to denote the variable, and we write $\mathbb{R}[t]_d = \{f \in \mathbb{R}[t] : \deg(f) \leq d\}$. The *support* of a polynomial $p = \sum_{\alpha} c_{\alpha} x^{\alpha} \in \mathbb{R}[x_1, \dots, x_n]$ is the set $\text{supp}(p) := \{\alpha \in \mathbb{Z}_+^n : c_{\alpha} \neq 0\}$. By S^m we denote the space of real symmetric $m \times m$ matrices, and S_+^m is the cone of positive semidefinite matrices in S^m . The linear span of a subset M of a vector space is denoted $\text{lin}(M)$.

2.2. Let $C \subseteq \mathbb{R}^n$ be a convex cone. The *dual cone* of C is $C^* = \{y \in \mathbb{R}^n : \forall x \in C \langle x, y \rangle \geq 0\}$, where $\langle x, y \rangle$ denotes the standard inner product on \mathbb{R}^n . The bi-dual $C^{**} := (C^*)^*$ of C equals the closure of C , i.e. $C^{**} = \overline{C}$. The cone C is *pointed* if $C \cap (-C) = \{0\}$. When $C \subseteq V$ is a convex cone in an arbitrary real vector space V , the dual cone is $C^* := \{y \in V^{\vee} : \forall x \in C \ y(x) \geq 0\}$, where V^{\vee} is the dual vector space of V .

A *face* of C is a convex cone F with $F \subseteq C$ such that $x, y \in C$ and $x + y \in F$ imply $x, y \in F$. For any $x \in C$ there is a unique inclusion-minimal face F of C with $x \in F$, called the *supporting face* of x in C . One-dimensional faces of convex cones are called *extreme rays*. It is well-known that a finite-dimensional closed and pointed convex cone is the Minkowski sum of its extreme rays.

2.3. We briefly explain the conic duality behind the approaches in polynomial optimization by providing a generic version of the discussion in [14] Ch. 10. We are given a subset S of \mathbb{R}^n and polynomials p, q in a finite-dimensional vector subspace V of $\mathbb{R}[x_1, \dots, x_n]$ such that $q > 0$ on S . A common choice for q is the constant $q = 1$. The problem of minimizing the quotient p/q over S can be relaxed (i.e., lower bounded) by making use of a closed convex cone $C \subseteq V$ that satisfies $g \geq 0$ on S for every

$g \in C$. One has

$$\inf_s \frac{p}{q} \geq \sup \{ \lambda \in \mathbb{R} : p - \lambda q \in C \},$$

where the supremum is a conic optimization problem dual to the problem

$$\inf \{ v(p) : v \in C^*, v(q) = 1 \}.$$

In the following proposition, the mentioned duality is phrased without any reference to polynomial by identifying V and its dual space V^\vee with \mathbb{R}^N , where $N := \dim(V)$.

Proposition 2.4 *Let N be a positive integer and $C \subseteq \mathbb{R}^N$ be a closed and pointed convex cone. Then for every $p \in C$ and $q \in C \setminus \{0\}$ one has*

$$\sup \{ \lambda \in \mathbb{R} : p - \lambda q \in C \} = \inf \{ \langle p, v \rangle : v \in C^*, \langle v, q \rangle = 1 \}. \tag{2}$$

This is a special case of the duality of conic optimization problems, cf. the discussion in [7] 2.1.4. We omit the proof which is quite straightforward.

3 Convex hulls of monomial curves

Let $I \subseteq \mathbb{R}$ be a non-degenerate closed interval, let $V \subseteq \mathbb{R}[t]$ be a linear subspace of finite dimension, and let $P = \{f \in V : f \geq 0 \text{ on } I\}$, a closed and pointed convex cone in V . We start by showing that every face of the cone P is described by suitable vanishing conditions at points of I , or at infinity when I is unbounded. Let $\text{ord}_s(f)$ denote the order of vanishing of $f \in \mathbb{R}[t]$ at $s \in \mathbb{R}$.

Proposition 3.1 *For $0 \neq f \in P$ put $W_f = \{g \in V : \forall s \in I \text{ ord}_s(g) \geq \text{ord}_s(f)\}$. Let $U_f = \{g \in W_f : \deg(g) \leq \deg(f)\}$ if I is unbounded, and put $U_f = W_f$ if I is compact. Then U_f is the linear span of the supporting face F_f of f . In particular, $\dim(F_f) = \dim(U_f)$.*

Before starting with the proof, we record two consequences. For extreme rays of P , Proposition 3.1 implies:

Corollary 3.2 *Assume that I is compact, let $f \in V$ span an extreme ray of P . Then f has at least $\dim(V) - 1$ roots in I , counting with multiplicity.*

Proof By assumption we have $F_f = \mathbb{R}_+ f$, so $W_f = \mathbb{R} f$ by Proposition 3.1. Since W_f consists of the elements in V with at least the same roots in I as f , there have to be at least $\text{codim}(W_f) = \dim(V) - 1$ many roots of f in I , counting with multiplicities. □

In the non-compact case, the result reads as follows:

Corollary 3.3 *Let I be unbounded, let $f \in P$ span an extreme ray of P . Write $d = \deg(f)$ and $V_d := \{g \in V : \deg(g) \leq d\}$. Then f has at least $\dim(V_d) - 1$ roots in I , counting with multiplicity.*

Proof Same argument as for Corollary 3.2, since $\mathbb{R}f$ consists of the elements in V_d with at least the same roots in I as f . □

Proof of Proposition 3.1 The supporting face of f is $F_f = \{p \in P : \exists \gamma > 0 \text{ with } f - \gamma p \in P\}$. Clearly, if $p, q \in P$ with $f = p + q$ then $p, q \in W_f$, and also $\deg(p), \deg(q) \leq \deg(f)$ when I is unbounded. Therefore $F_f \subseteq U_f$ holds. To prove $U_f \subseteq \text{lin}(F_f)$, note that for every $g \in U_f$ there exists a constant $c > 0$ such that $c|g| \leq f$ on I . Indeed, for every $s \in I$ there is a constant $c_s > 0$ with $c_s|g(t)| \leq f(t)$ on some neighborhood J_s of s . If I contains a right half-line, there are $a, c > 0$ with $c|g(t)| \leq f(t)$ for all $t \in J_\infty = [a, \infty[$, and similarly in the case of a left half-line. By passing to a finite subcovering of I we find a constant c as required.

By the preceding remark we find a linear basis g_1, \dots, g_r of U_f with the property that $|g_i| \leq f$ on I for $i = 1, \dots, r$. It follows that $f - g_i \in F_f$ ($i = 1, \dots, r$), and so $U_f = \text{lin}(f - g_1, \dots, f - g_r, f)$ is contained in $\text{lin}(F_f)$. □

We continue to assume that $V \subseteq \mathbb{R}[t]$ is a linear subspace of finite dimension, and we put $\dim(V) =: n + 1$. Assume we are given n vanishing conditions for elements of V at points on the real line, where we allow conditions of higher order vanishing. If these conditions are independent on V , the following lemma gives an explicit formula for the essentially unique solution of these conditions in V . Given a polynomial $p = p(t)$ in $\mathbb{R}[t]$, we denote the j -th derivative of $p(t)$ by $p^{(j)}(t) = \frac{d^j}{dt^j} p(t)$.

Lemma 3.4 *Let p_0, \dots, p_n be a linear basis of V . Let $\xi = (\xi_1, \dots, \xi_r)$ be a tuple of pairwise different real numbers, and let $b_1, \dots, b_r \geq 1$ be integers with $\sum_{i=1}^r b_i = n$. The following conditions are equivalent:*

(i) *The matrix*

$$A(t; \xi) = \begin{pmatrix} p_0(t) & \cdots & p_n(t) \\ p_0(\xi_1) & \cdots & p_n(\xi_1) \\ \vdots & & \vdots \\ p_0^{(b_1-1)}(\xi_1) & \cdots & p_n^{(b_1-1)}(\xi_1) \\ \vdots & & \vdots \\ p_0(\xi_r) & \cdots & p_n(\xi_r) \\ \vdots & & \vdots \\ p_0^{(b_r-1)}(\xi_r) & \cdots & p_n^{(b_r-1)}(\xi_r) \end{pmatrix} \tag{3}$$

of size $(n + 1) \times (n + 1)$ and with entries in $\mathbb{R}[t]$ has non-zero determinant;
(ii) *the subspace $\{f \in V : \text{ord}_{\xi_i}(f) \geq b_i \text{ for } i = 1, \dots, r\}$ of V has dimension one.*
When (i) and (ii) hold, the polynomial $\det A(t; \xi)$ is the unique (up to scaling) element of V that vanishes in ξ_i of order at least b_i , for $i = 1, \dots, r$.

Proof Note that the dimension in (ii) is always ≥ 1 . The matrix B formed by the lower n rows of (3) is the matrix of the linear map $\phi : V \rightarrow \mathbb{R}^n$,

$$p \mapsto \left(p(\xi_1), \dots, p^{(b_1-1)}(\xi_1), \dots, p(\xi_r), \dots, p^{(b_r-1)}(\xi_r) \right)$$

with respect to the basis p_0, \dots, p_n of V . Subspace (ii) is just the kernel of ϕ . The determinant of (3) is an element of V , and is non-zero if and only if B has a non-vanishing $n \times n$ -minor. This is equivalent to ϕ being surjective, and hence also to (ii). The last assertion is clear since $\det A(t; \xi)$ is an element of V that has a zero of multiplicity $\geq b_i$ at ξ_i , for $i = 1, \dots, r$. \square

3.5. Next we recall some background on Schur polynomials. Let $n \in \mathbb{N}$ be a fixed integer and let $x = (x_0, \dots, x_n)$ be a tuple of indeterminates. We consider partitions $\mu = (m_0, \dots, m_n)$ into $n + 1$ pieces, with $m_0 \geq \dots \geq m_n \geq 0$. A particular partition is $\delta = (n, n - 1, \dots, 1, 0)$. Given a partition μ as above, the determinant

$$F_\mu(x) := \det \begin{pmatrix} x_0^{m_0} & \dots & x_0^{m_n} \\ \vdots & & \vdots \\ x_n^{m_0} & \dots & x_n^{m_n} \end{pmatrix}$$

is identically zero unless the m_i are pairwise distinct. In this latter case, $\lambda = \mu - \delta = (m_0 - n, \dots, m_{n-1} - 1, m_n)$ is another partition. Clearly, the Vandermonde product

$$v(x) := \prod_{0 \leq i < j \leq n} (x_i - x_j) = F_\delta(x)$$

divides $F_\mu(x)$. The co-factor is the *Schur polynomial* $s_\lambda(x)$ of the partition λ . In other words, the *bialternant formula*

$$F_\mu(x) = v(x) \cdot s_\lambda(x) \tag{4}$$

holds. Depending on how Schur polynomials are introduced, identity (4) is either the definition of $s_\lambda(x)$ or a theorem, see [21] Theorem 7.15.1.

The Schur polynomial $s_\lambda(x)$ is symmetric as a polynomial in the variables x_i , homogeneous of degree $|\lambda|$, and of degree λ_0 with respect to each variable x_i . Schur polynomials have the remarkable property that all their coefficients are non-negative integers. In fact there exists a combinatorial description of the coefficients, see Section 7.10 in [21]. In our context, the integrality of the coefficients plays no role, but the non-negativity is a crucial property.

3.6. We use Schur polynomials to deduce a product formula for the determinant in Lemma 3.4, in the case where the p_i are monomials. Let $\mu = (m_0, \dots, m_n)$ be a partition into different parts, i.e. with $m_0 > \dots > m_n \geq 0$. We write $p_i(t) = t^{m_i}$ for $i = 0, \dots, n$, and introduce the tuples $p(t) = (p_0(t), \dots, p_n(t))$ and $p^{(j)}(t) = (p_0^{(j)}(t), \dots, p_n^{(j)}(t))$ for $j \geq 0$, where again $p_i^{(j)}(t) = \frac{d^j}{dt^j} p_i(t)$.

Let $b = (b_0, \dots, b_r)$ be a tuple of integers $b_i \geq 1$ such that $\sum_i b_i = n + 1$, let $y = (y_0, \dots, y_r)$ be a tuple of $r + 1$ variables. We consider the determinant $F_{\mu,b}(y)$ of size $n + 1$ that contains, for each $i = 0, \dots, r$, the rows

$$p^{(j)}(y_i) = (p_0^{(j)}(y_i), \dots, p_n^{(j)}(y_i))$$

for $j = 0, 1, \dots, b_i - 1$. In other words, let

$$F_{\mu,b}(y) := \det \begin{pmatrix} p_0(y_0) & \cdots & p_n(y_0) \\ p'_0(y_0) & \cdots & p'_n(y_0) \\ \vdots & & \vdots \\ p_0^{(b_0-1)}(y_0) & \cdots & p_n^{(b_0-1)}(y_0) \\ \vdots & & \vdots \\ p_0(y_r) & \cdots & p_n(y_r) \\ \vdots & & \vdots \\ p_0^{(b_r-1)}(y_r) & \cdots & p_n^{(b_r-1)}(y_r) \end{pmatrix} \quad (5)$$

and let again $\lambda = \mu - \delta$.

Proposition 3.7 *The determinant $F_{\mu,b}(y)$ has the product decomposition*

$$F_{\mu,b}(y) = c \cdot v_b(y) \cdot s_{\lambda,b}(y)$$

where

$$v_b(y) := \prod_{0 \leq i < j \leq r} (y_i - y_j)^{b_i b_j}$$

and

$$s_{\lambda,b}(y) := s_{\lambda}(\underbrace{y_0, \dots, y_0}_{b_0}, \dots, \underbrace{y_r, \dots, y_r}_{b_r})$$

and c is a constant, equal to

$$c = \prod_{i=0}^r (-1)^{b_i(b_i-1)/2} \cdot \prod_{j=0}^{b_i-1} j!$$

Proof We inductively derive the assertion from the bialternant formula. Let A_0 be the matrix with rows $p(x_0), \dots, p(x_n)$, so

$$\det(A_0) = s_{\lambda}(x) \cdot \prod_{0 \leq i < j \leq n} (x_i - x_j)$$

by (4). Replacing the second row $p(x_1)$ by $(p(x_1) - p(x_0))/(x_1 - x_0)$, the determinant gets divided by $x_1 - x_0$. If we now specialize $x_1 := x_0$, the resulting matrix A_1 has second row $p'(x_0)$ and has determinant

$$\det(A_1) = -s_{\lambda}(x_0, x_0, x_2, \dots, x_n) \cdot \prod_{j=2}^n (x_0 - x_j)^2 \cdot \prod_{2 \leq i < j \leq n} (x_i - x_j)$$

By pulling out $x_1 - x_0$, we have therefore used up the factor $x_0 - x_1$ from $v(x)$ and got a minus sign. Now iterate this step. Next we replace the third row of A_1 , which is $p(x_2)$, by

$$\frac{p(x_2) - p(x_0) - (x_2 - x_0)p'(x_0)}{(x_2 - x_0)^2}$$

Specializing $x_2 := x_0$, the resulting matrix A_2 has rows

$$p(x_0), p'(x_0), \frac{1}{2}p''(x_0), p(x_3), \dots$$

and has determinant

$$\det(A_2) = -s_\lambda(x_0, x_0, x_0, x_3, \dots, x_n) \cdot \prod_{j=3}^n (x_0 - x_j)^3 \cdot \prod_{3 \leq i < j \leq n} (x_i - x_j)$$

And so on. After b_0 many steps, the rows have become

$$p(x_0), p'(x_0), \dots, \frac{1}{(b_0 - 1)!} p^{(b_0-1)}(x_0), p(x_{b_0}), \dots, p(x_n),$$

and at that point we have thrown in $0 + 1 + \dots + (b_0 - 1) = \binom{b_0}{2}$ many minus signs.

We can now repeat this procedure. The next step consists in doing b_1 many steps on the variables $x_{b_0}, \dots, x_{b_0+b_1-1}$. And so on. Finally, relabel the variables that have survived as y_0, \dots, y_r . □

3.8. After these preparations we come to the main result of our paper. Let $\mu = (m_0, \dots, m_n)$ with $m_0 > \dots > m_n \geq 0$, write $k = 1 + \lfloor \frac{n}{2} \rfloor$. Let $\mathbb{R}[t]$ be the ring of polynomials in the variable t . We consider the subspace $V = V_\mu$ of $\mathbb{R}[t]$ that is spanned by the monomials t^{m_0}, \dots, t^{m_n} . Let $I \subseteq \mathbb{R}$ be an interval, and let

$$P := \{f \in V : f \geq 0 \text{ on } I\}.$$

Note that P is a closed convex cone in V . The key result is the following nichtnegativstellensatz:

Theorem 3.9 *Let (a) $I = [0, 1]$ or (b) $I = \mathbb{R}_+ = [0, \infty[$. Then every $f \in P$ can be written as a finite sum*

$$\begin{aligned} \text{(a) } f &= \sum_i g_i(t)^2 q_i(t) + (1-t) \sum_j h_j(t)^2 r_j(t), \\ \text{(b) } f &= \sum_i g_i(t)^2 q_i(t), \end{aligned} \tag{6}$$

where g_i, q_i, h_j, r_j are polynomials with $\deg(g_i), \deg(h_j) \leq \lfloor \frac{n}{2} \rfloor = k - 1$, such that the degree of every summand is $\leq m_0$, and such that the coefficients of q_i, r_j are all non-negative.

Conversely, when f has a representation as in the theorem, it is obvious that $f \geq 0$ on I . See Remark 3.18 for why the case $I = \mathbb{R}$ is not included here.

3.10. To start the proof of Theorem 3.9 (in either case (a) or (b)), let Q be the set of all $f \in V$ that have a representation (6). Then Q is a convex cone, and $Q \subseteq P$. Let $E \subseteq P$ be the set of all polynomials f that generate an extreme ray of P . Every element of P is a sum of finitely many elements of E , since the cone P is closed and pointed. To prove $Q = P$, it therefore suffices to show $f \in Q$ for any given $f \in E$.

3.11. Let us make two more reduction steps. Assume that $f \in Q$ has been shown (for all monomial subspaces) whenever $f \in E$ satisfies $f(0) > 0$. Let $f \in E$ with $f(0) = 0$. Writing $f = t^w \tilde{f}$ where $w = \text{ord}_0(f)$, there is an index $0 \leq s \leq n$ with $m_s = w$. The subspace $\tilde{V} = \text{lin}(t^{m_0-w}, \dots, t^{m_{s-1}-w}, 1)$ of $\mathbb{R}[t]$ has dimension $s + 1 \leq n + 1$ and contains \tilde{f} . Moreover $\tilde{f} \geq 0$ on I and $\tilde{f}(0) > 0$. To fix ideas, assume we are in case (a), so $I = [0, 1]$. By assumption, the theorem holds for \tilde{f} and the subspace \tilde{V} . This means that there is an identity

$$\tilde{f} = \sum_i g_i(t)^2 q_i(t) + (1 - t) \sum_j h_j(t)^2 r_j(t) \tag{7}$$

where $\deg(g_i), \deg(h_j) \leq \lfloor \frac{s}{2} \rfloor$ and the q_i, r_j have non-negative coefficients, such that every summand in (7) has degree $\leq m_0 - w$. Multiplying the identity with t^w , we get an identity for f as claimed in Theorem 3.9(b). In case (b), the argument is exactly the same.

3.12. So we need to establish an identity (6) for every $f \in E$ with $f(0) > 0$. For this, we clearly may discard all monomials of degree greater than $d = \deg(f)$. In other words, we may assume that $\deg(f) = m_0$. This reduction step plays a role only in case (b). (In fact, $\deg(f) = m_0$ is automatic if $I = [0, 1]$, as will be seen from the proof below.)

3.13. From Descartes' rule of signs (e.g. [12] Cor. 1.10.3) it follows that every non-zero $f \in V$ has at most n strictly positive roots, counting with multiplicity. On the other hand, Corollaries 3.2 and 3.3 show that every $f \in E$ with $\deg(f) = m_0$ has at least n roots in I . If in addition $f(0) > 0$, it follows that f has precisely n strictly positive roots, and they all lie in I .

3.14. Let $f \in E$ with $f(0) > 0$ and $\deg(f) = m_0$. Note that $m_n = 0$ now since $f(0) \neq 0$. By 3.13, f has precisely n positive roots, and they all lie in I . It suffices to show that every such f satisfies an identity (6).

Let ξ_1, \dots, ξ_r be the different positive roots of f , and let $b_i = \text{ord}_{\xi_i}(f)$ ($i = 1, \dots, r$). Each b_i is an even integer, except possibly in case (a) when $\xi_i = 1$. We have $\sum_{i=1}^r b_i = n$ since f has n roots in I .

Consider the determinant (5) in 3.6, with $b_0 := 1$ and $p_i = t^{m_i}$ ($i = 0, \dots, n$). After substituting $y_0 = t$ and $y_i = \xi_i$ ($i = 1, \dots, r$), Proposition 3.7 shows that the

determinant has a factorization

$$F_{\mu,b}(t, \xi_1, \dots, \xi_r) = \gamma \prod_{i=1}^r (t - \xi_i)^{b_i} \cdot s_\lambda(t, \underbrace{\xi_1, \dots, \xi_1}_{b_1}, \dots, \underbrace{\xi_r, \dots, \xi_r}_{b_r})$$

with $\gamma \neq 0$ a constant. The last factor $s_\lambda(t, \xi_1, \dots, \xi_r)$ is a polynomial in t . Since all coefficients of the Schur polynomial s_λ are ≥ 0 , and since $\xi_i > 0$ for all i , this polynomial is not identically zero (and has degree $\lambda_0 = m_0 - n$). Hence the determinant is non-zero, and Lemma 3.4 implies that it agrees with f up to scaling. This means that f has a factorization

$$f = \gamma' \prod_{i=1}^r (t - \xi_i)^{b_i} \cdot s_\lambda(t, \underbrace{\xi_1, \dots, \xi_1}_{b_1}, \dots, \underbrace{\xi_r, \dots, \xi_r}_{b_r}) \tag{8}$$

with a constant $\gamma' \neq 0$. The first factor after the constant has the form

$$\prod_{i=1}^r (t - \xi_i)^{b_i} = g(t)^2 u(t)$$

where $\deg(g) = k - 1 = \lfloor \frac{n}{2} \rfloor$ and $u(t) = 1$ or $t - 1$. The case $u(t) = t - 1$ occurs only in case (a) ($I = [0, 1]$) when n is odd. The last factor $s_\lambda(t, \xi_1, \dots)$ in (8) is a polynomial in t with non-negative coefficients. So, up to a non-zero constant factor, the right hand side has the form claimed in Theorem 3.9. Since $f|_I \geq 0$ we see that $\gamma' > 0$ in case (a) with n even, and also in case (b). When n is odd in case (a), we have $\xi_i = 1$ for some i , and b_i is odd, so $\gamma' < 0$ in this case. The theorem is proved. \square

Recall our convention $k = 1 + \lfloor \frac{n}{2} \rfloor$. Let $\Sigma = \Sigma_{2k} := \{g \in \mathbb{R}[t] : \deg(g) \leq 2k, g \geq 0 \text{ on } \mathbb{R}\}$. Each $g \in \Sigma_{2k}$ can be written $g = g_1^2 + g_2^2$ with polynomials g_1, g_2 of degree $\leq k$. Theorem 3.9 can be stated in the following alternative form:

Corollary 3.15 *Again let (a) $I = [0, 1]$ or (b) $I = \mathbb{R}_+$. With assumptions as in Theorem 3.8, the inclusion*

$$(a) \ P \subseteq \left(\Sigma + t\Sigma + \dots + t^{m_0-2k}\Sigma \right) + (1-t)\left(\Sigma + t\Sigma + \dots + t^{m_0-2k-1}\Sigma \right)$$

or

$$(b) \ P \subseteq \left(\Sigma + t\Sigma + \dots + t^{m_0-2k}\Sigma \right)$$

holds, respectively.

Remark 3.16 Part (b) of Corollary 3.15 can be considered to be a *sparse* (or *fewnomial*) *positivstellensatz* (or rather, *nichtnegativstellensatz*) for univariate polynomials: Every polynomial $f \in \mathbb{R}[t]$ with m monomials that is non-negative on \mathbb{R}_+ can be written as a (finite) sum

$$f = \sum_{i \geq 0} t^i (p_i(t)^2 + q_i(t)^2)$$

where p_i, q_i are polynomials of degree $\leq \lfloor \frac{m-1}{2} \rfloor$. This point of view will be expanded in more detail in the next section.

Remark 3.17 As before, let $V \subseteq \mathbb{R}[t]$ be a linear subspace generated by $n + 1$ monomials, and let $P = \{f \in V : f \geq 0 \text{ on } I\}$ where (a) $I = [0, 1]$ or (b) $I = \mathbb{R}_+$. In either case, Corollary 3.15 allows us to read off a block semidefinite representation of P of block size at most $k + 1$. Let $\varphi : S^{k+1} \rightarrow \mathbb{R}[t]_{2k}$ be the linear map that sends a symmetric matrix $M = (a_{ij})_{0 \leq i, j \leq k}$ to

$$\varphi(M) := (1, t, \dots, t^k) \cdot M \cdot (1, t, \dots, t^k)^\top = \sum_{i, j=0}^k a_{ij} t^{i+j}$$

Then $\varphi(S_+^{k+1}) = \Sigma_{2k}$. In case (a), consider the linear map

$$\phi : (S^{k+1})^{2(m_0-2k)+1} \rightarrow \mathbb{R}[t]$$

given by

$$\begin{aligned} & (M_0, \dots, M_{m_0-2k}; N_0, \dots, N_{m_0-2k-1}) \\ \mapsto & \sum_{i=0}^{m_0-2k} t^i \varphi(M_i) + (1-t) \sum_{j=0}^{m_0-2k-1} t^j \varphi(N_j) \end{aligned}$$

For (b), consider the linear map $\phi : (S^{k+1})^{m_0-2k+1} \rightarrow \mathbb{R}[t]$ given by

$$(M_0, \dots, M_{m_0-2k}) \mapsto \sum_{i=0}^{m_0-2k} t^i \varphi(M_i)$$

In either case we have $P = V \cap \phi(S_+)$, according to Corollary 3.15, where $S_+ = (S_+^{k+1})^N$ with $N = 2(m_0 - 2k) + 1$ in case (a) and $N = m_0 - 2k + 1$ in case (b). This is an explicit block diagonal semidefinite representation of P of block size $k + 1$.

Remark 3.18 For sparse polynomials f in $\mathbb{R}[t]$ that are non-negative on the whole real axis, there does not in general exist a sparse decomposition

$$f = \sum_{i \in 2\mathbb{Z}_+} t^i (p_i(t)^2 + q_i(t)^2)$$

similar to the one described in Remark 3.16. For example, $f = 3t^4 - 4t^3 + 1 = (t - 1)^2(3t^2 + 2t + 1)$ is non-negative on \mathbb{R} but cannot be written $f = g_1(t) + t^2 g_2(t)$ with g_1, g_2 sums of squares of linear polynomials.

Still, we may easily produce a block semidefinite representation of block size $k + 1$ for $P = \{f \in V : f \geq 0 \text{ on } \mathbb{R}\}$, from the case $I = \mathbb{R}_+$ in 3.17. It suffices to remark that $f(t) \geq 0$ on \mathbb{R} is equivalent to $f(t)$ and $f(-t)$ both being ≥ 0 on \mathbb{R}_+ .

By cone duality we get the desired theorem for convex hulls of monomial space curves:

Theorem 3.19 *Let $m_1, \dots, m_n \geq 1$ be integers, and let $I \subseteq \mathbb{R}$ be a semialgebraic set. The closed convex hull K of the set*

$$S = \{(t^{m_1}, \dots, t^{m_n}) : t \in I\}$$

in \mathbb{R}^n has semidefinite extension degree at most $\lfloor \frac{n}{2} \rfloor + 1$.

Proof The set K is identified with an affine-linear slice of the dual cone P^* of P . Since $\text{sxdeg}(P^*) = \text{sxdeg}(P)$ ([19] Prop. 1.7), it suffices to prove $\text{sxdeg}(P) \leq \lfloor \frac{n}{2} \rfloor + 1$. We may assume that the m_i are pairwise distinct. When $I = [0, 1]$ or $I = \mathbb{R}_+$, the claim $\text{sxdeg}(P) \leq \lfloor \frac{n}{2} \rfloor + 1$ was shown in Remark 3.17. We sketch how the case of other sets I can essentially be reduced to these two cases, without going into full details. Clearly I can be assumed to be closed. If $I = I_1 \cup I_2$ with I_1, I_2 semialgebraic, it is enough to prove the claim for both I_1 and I_2 , in view of [19] Prop. 1.6. In this way we reduce to considering $I = [a, b]$ or $I = [a, \infty[$ where $0 \leq a < b < \infty$. The cases $a = 0$ are already done (assuming $b = 1$ in the compact case was nowhere essential). Assume $I = [a, b]$ with $0 < a < b < \infty$. Then every $f \in P$ that generates an extreme ray of P has exactly n roots in I , counting with multiplicities. So the reduction step 3.11 in the proof of Theorem 3.9 is not needed. Otherwise we may just follow the proof of this theorem. In this way we arrive at a representation of every element of P in a form similar to 3.9(a), but with weights $1, t - a, b - t$ and $(t - a)(b - t)$ instead of only 1 and $1 - t$. When $I = [a, \infty[$ with $a > 0$, we may proceed in a similar way. \square

Remark 3.20 For $I = [0, 1]$ or $I = \mathbb{R}_+$, an explicit block semidefinite representation of K of block size $\lfloor \frac{n}{2} \rfloor + 1$ can be obtained from Remarks 3.17, 3.18 by dualizing.

Remark 3.21 In Theorem 3.9, the upper bound $\lfloor \frac{n}{2} \rfloor = k - 1$ for the degrees of g_i, h_j cannot be made smaller in general. This can be seen by considering the tuple $\mu = (n, n - 1, \dots, 1, 0)$ and the corresponding subspace $V = \mathbb{R}[t]_{\leq n}$ of $\mathbb{R}[t]$: The closed convex hull K of the set $S = \{(t, t^2, \dots, t^n) : t \in I\}$ (with $I = \mathbb{R}_+$ or $I = [0, 1]$) has $\text{sxdeg}(K) = k = \lfloor \frac{n}{2} \rfloor + 1$, by [4] Corollary 2.14 (see the reasoning at the beginning of the proof of Theorem 3.19).

Remark 3.22 It is natural to ask whether the bound $\text{sxdeg}(K) \leq \lfloor \frac{n}{2} \rfloor + 1$ extends to cases more general than convex hulls of monomial curves. In fact, the same bound is true in general whenever $K \subseteq \mathbb{R}^n$ is the closed convex hull of a one-dimensional semialgebraic set in \mathbb{R}^n [20]. However the proof gets much more difficult than in the monomial case. Even when the curve is parametrized by polynomials instead of monomials, one has to argue locally on sufficiently small intervals on the curve, and there does not seem to be an explicit form for a block semidefinite representation. When the curve is non-rational, the proof becomes technically even much more complicated.

Remark 3.23 In Sections 4–5 of [18], the authors explain how the convex cone $\mathcal{M} = \text{cone}\{(t^{m_0}, \dots, t^{m_n}) : t \in I\}$ is useful in population genetics, in the case $I = [0, 1]$ and $m_i = \binom{i+2}{2} - 1$ ($0 \leq i \leq n$). There one needs to determine the Euclidean

distance from a given point in \mathbb{R}^{n+1} to the hyperplane section $\mathcal{C} = \{(x_0, \dots, x_n) \in \mathcal{M} : x_0 + \dots + x_n = 1\}$ of \mathcal{M} . Corollary 20 of [18] relies on the fact that \mathcal{C} has a lifted description by two LMIs, each of size $\mathcal{O}(n^2)$. This may be compared with our main result, which implies that \mathcal{C} has a lifted description by $\mathcal{O}(n^2)$ many LMIs each of size $\mathcal{O}(n)$. Interestingly, the authors of [18] are also using Descartes' rule of sign, Schur polynomials and the bialternant formula. Other than in our paper, their purpose is to understand the facial structure of \mathcal{M} (see Theorem 3 and Corollary 6 in [18]).

4 Sparse non-negativity certificates in polynomial optimization

4.1 Sums of copositive fewnomials and sparse semidefinite relaxations

We are going to take a second look at Corollary 3.15, concentrating on part (b). Let $k \geq 1$ be an integer. A univariate polynomial $f \in \mathbb{R}[t]$ will be called a k -nomial if f is a linear combination of at most k monomials t^i . We say that f is *copositive* if $f \geq 0$ on \mathbb{R}_+ . Given a finite set $J \subseteq \mathbb{Z}_+ := \{0, 1, 2, \dots\}$, let

$$\text{CP}(J) := \{p \in \mathbb{R}[t] : p \geq 0 \text{ on } \mathbb{R}_+ \text{ and } \text{supp}(p) \subseteq J\}.$$

This is the cone P considered in Sect. 3, for the monomial curve corresponding to J . Let

$$\text{SOCF}_{k,d} := \sum_{\substack{J \subseteq \{0, \dots, d\} \\ |J|=k}} \text{CP}(J)$$

The acronym SOCF stands for *sums of copositive fewnomials*. So $f \in \mathbb{R}[t]_d$ lies in $\text{SOCF}_{k,d}$ if, and only if, f can be written as a finite sum of copositive k -nomials of degree at most d . It is obvious how to generalize this setup from univariate polynomials to a multivariate setting. This could be an interesting setup to explore in the future.

As before, let $\Sigma_{2k} \subseteq \mathbb{R}[t]$ denote the set of sums of squares polynomials of degree $\leq 2k$. The main result of Sect. 3 implies the following positivstellensatz for the cones SOCF:

Corollary 4.1 *For all integers $k, d \geq 1$ with $d > 2k$, we have*

$$\text{SOCF}_{2k+1,d} = t^0 \Sigma_{2k} + \dots + t^{d-2k} \Sigma_{2k}. \quad (9)$$

Moreover $\text{sxdeg}(\text{SOCF}_{2k+1,d}) = k + 1$.

For $d = 2k$, note that $\text{SOCF}_{2k+1,2k}$ is just the cone of all copositive univariate polynomials of degree at most $2k$, i.e., it is the cone

$$\{f \in \mathbb{R}[t]_{2k} : f \geq 0 \text{ on } \mathbb{R}_+\} = \Sigma_{2k} + t \Sigma_{2k-2},$$

where the latter description of copositivity in the univariate case is well known.

Proof The inclusion “ \subseteq ” holds since, for every set $J \subseteq \{0, \dots, d\}$ with $|J| \leq 2k + 1$, the cone $\text{CP}(J)$ is contained in the right hand side of (9) (Corollary 3.15(b)). The reverse inclusion is obvious. The last statement follows from identity (9) by using [19] Lemma 1.4(d). \square

To describe the duals of the SOCF cones, we identify $v \in (\mathbb{R}[t]_d)^\vee$ and $(v_0, \dots, v_d) \in \mathbb{R}^{d+1}$ via $v_i = v(t^i)$.

Lemma 4.2 *Let $d, k \geq 1$ with $d > 2k$. The dual of $\text{SOCF}_{2k+1,d}$ is the cone*

$$\{v = (v_0, \dots, v_d) \in \mathbb{R}^{d+1} : M_{0,k}(v) \geq 0, \dots, M_{d-2k,k}(v) \geq 0\},$$

where

$$M_{s,k}(v) := (v_{s+i+j})_{i,j=0,\dots,k} \in \mathbb{R}^{(k+1) \times (k+1)}.$$

Proof This follows from (9) in Corollary 4.1, since $(C_1 + C_2)^* = C_1^* \cap C_2^*$ and since $(\Sigma_{2k})^* = \{u = (u_0, \dots, u_k) : M_{0,k}(u) \geq 0\}$; see, for example, [7] Sect. 4.6. \square

The notation $M_{s,k}(v)$ from Lemma 4.2 is used in the following proposition, which provides primal and dual sparse relaxations for the problem $\min_{\mathbb{R}_+} f$, with $f \in \mathbb{R}[t]_d$, with a flexible choice of the sparsity threshold $2k + 1$.

Proposition 4.3 *Let k, d be positive integers with $2k < d$ and $f \in \mathbb{R}[t]_d$. Then the following elements in $\mathbb{R} \cup \{\pm\infty\}$ coincide:*

- (1) $\sup\{\lambda \in \mathbb{R} : f - \lambda \in \text{SOCF}_{2k+1,d}\},$
- (2) $\sup\{\lambda \in \mathbb{R} : f - \lambda \in t^0 \Sigma_{2k} + \dots + t^{d-2k} \Sigma_{2k}\},$
- (3) $\inf\{\langle f, v \rangle : v = (v_0, \dots, v_d) \in \mathbb{R}^{d+1}, v_0 = 1, M_{0,k}(v), \dots, M_{d-2k,k}(v) \geq 0\}.$

Proof This follows from (9) in Corollary 4.1, using Proposition 2.4. \square

Remark 4.4 For the case $k = 0$, relaxations in Proposition 4.3 are trivial with the optimal value $f(0)$ or $-\infty$, depending on whether or not all coefficients of $f - f(0)$ are non-negative. But already the simplest non-trivial case $k = 1$ exhibits connections to active ongoing research on sparse relaxations in polynomial optimization, namely to the relaxations that rely on the so-called *sage* and *sonc* polynomials; see [8] and [10, 15], respectively. Both kinds of polynomials emerge from the same idea applied to non-negativity on \mathbb{R}_+^n and \mathbb{R}^n , respectively. A *sage* polynomial is a sum of polynomials that are non-negative on \mathbb{R}_+^n and whose supports are contained in a simplicial circuit. Here, a simplicial circuit is an inclusion-minimal affinely dependent set whose convex hull is a simplex. Since simplicial circuits within \mathbb{R} are merely three-element sets, we see that the cone $\text{SOCF}_{3,d}$ is the cone of all univariate *sage* polynomials of degree at most d . Thus, the case $k = 1$ provides a primal and a dual *sage* relaxation of $\min_{\mathbb{R}_+} f$. See, for example, [10] for the discussion of the duality for the *sonc* and *sage* cones and the applications in optimization. We also note that the special case $k = 1$ of Corollary 4.1 amounts to a description of non-negative *sage* polynomials in terms of the so-called reduced circuits; see [10]. That is, in the univariate situation, an analog

of a reduced circuit in the setting of copositive $(2k + 1)$ -nomials is a set of $2k + 1$ consecutive integer values, since the support of polynomials in $t^i \Sigma_{2k}$ is a subset of $\{i, \dots, i + 2k\}$.

We stress that the sparsity threshold of sage and sonic polynomials is bound to the dimension, since a circuit in dimension n cannot have more than $n + 2$ elements. In contrast to this, in our setting the sparsity threshold $2k + 1$ may vary arbitrarily, bridging the sage relaxations with the dense standard relaxations of Lasserre. While being aware that our results are limited to the univariate case, we hope that such results might serve as an inspiration for similar studies in an arbitrary dimension n . That is, it would be interesting to understand properties of sos and moment relaxations that have a variable sparsity threshold.

Remark 4.5 The above discussion of sparse polynomial optimization over \mathbb{R}_+ can also be used to handle sparse polynomial optimization on an arbitrary interval $I \subseteq \mathbb{R}_+$. For example, when $2k < d - 1$ and $I = [0, 1]$, in view of Corollaries 3.15(a) and 4.1, the cone of sums of $(2k + 1)$ -nomials of degree at most d that are non-negative on $[0, 1]$ can be described as $\text{SOCF}_{2k+1,d} + (1 - t) \text{SOCF}_{2k,d-1}$. For both of the involved SOCF cones we have semidefinite descriptions of size $k + 1$. This leads to a semidefinite formulation of the optimization problem $\inf_{t \in [0,1]} p(t)$ for p a polynomial with $|\text{supp}(p)| \leq 2k + 1$ using $\mathcal{O}(d)$ LMIs of size $k + 1$, analogous to the formulation (3) given in Proposition 4.3.

4.2 Sparse semidefinite relaxations with a chordal-graph sparsity

We consider graphs $G = (V, E)$ where $|V| < \infty$ and edges $e \in E$ are two-element subsets of V . A *cycle* of length k in G is a connected subgraph of G with k nodes and each node having degree two. A *chord* of a cycle C in G is an edge of G that connects two nodes of C but is not an edge of C . A graph G is said to be *chordal* if every cycle of G of length at least four has a chord in G . A *clique* $W \subseteq V$ of $G = (V, E)$ is a set of nodes with any two distinct nodes in W connected by an edge. The following result provides a convenient representation of the cone of psd matrices with the chordal sparsity pattern:

Theorem 4.6 (see [2] and [22] Sect. 9.2) *Let $G = (V, E)$ be a graph with $n = |V| < \infty$. With G we associate the cone*

$$S_{+,E}^V := \{(a_{ij})_{i,j \in V} \in S_+^n : a_{ij} = 0 \text{ when } i \neq j \text{ and } \{i, j\} \notin E\}.$$

If G is chordal, then this cone admits the decomposition

$$S_{+,E}^V = S_+^{V_1} + \dots + S_+^{V_N},$$

where $V_1, \dots, V_N \subseteq V$ are all inclusion-maximal cliques of G and, for $W \subseteq V$,

$$S_+^W := \{(a_{ij})_{i,j \in V} \in S_+^n : a_{ij} = 0 \text{ for } (i, j) \notin W \times W\}.$$

Remark 4.7 $S_{+,E}^V$ is the so-called cone of psd matrices with the sparsity pattern of $G = (V, E)$. If G is chordal, $S_{+,E}^V$ is known to have favorable theoretical and computational properties, which allow one to solve conic optimization problems with respect to the cones $S_{+,E}^V$ more efficiently than the problems with respect to S_+^n . For example, computations with barrier functions can be done more efficiently on $S_{+,E}^V$ since there are more efficient ways to carry out Cholesky factorization for matrices in $S_{+,E}^V$ when the graph $G = (V, E)$ is chordal. We refer to [25] Appendix A for further details.

Remark 4.8 As a consequence of Theorem 4.6 we see that, for a chordal graph $G = (V, E)$, the semidefinite extension degree of $S_{+,E}^V$ is the clique number of G , i.e., the maximum clique size of G .

For positive integers n and k with $k < n$, the cone

$$S_{+,k}^n = \{(a_{ij})_{i,j=0,\dots,n-1} \in S_+^n : a_{ij} = 0 \text{ for } |i - j| > k\}$$

is the cone $S_{+,E}^V$ for the graph $G = (V, E)$ with vertex set $V = \{0, \dots, n - 1\}$ and edge set E consisting of $\{i, j\}$ that satisfy $0 < |i - j| \leq k$. In [22] Sect. 8.2, $S_{+,k}^n$ is called the cone of psd matrices with the *band sparsity pattern of the band width* $2k + 1$. It is clear that the graph G defining $S_{+,k}^n$ is chordal and that inclusion-maximal cliques of G are sets of the form $\{i, \dots, i + k\}$ with $0 \leq i < n - k$. We show that, similarly to how S_+^n is used for representing sos cones and truncated quadratic modules, the cone $S_{+,k}^n$ can be used for representing the cone $\text{SOCF}_{2k+1,n}$. That is, optimization problem (1) from Proposition 4.3 can be formulated as a conic problem with respect to the cones $S_{+,k}^n$. Such formulations may have computational advantages if used in solvers that can exploit the chordal sparsity.

Proposition 4.9 *Let d, k be positive integers with $2k < d$. Then the cone $\text{SOCF}_{2k+1,d}$ admits a representation as a linear image of $S_{+,k}^{e+1} \times S_{+,k}^e$, if $d = 2e$ is even, and as a linear image of $S_{+,k}^{e+1} \times S_{+,k}^{e+1}$, if $d = 2e + 1$ is odd.*

Proof We only consider the case of an even d with $d = 2e$, because the case of an odd d is completely analogous. We use the linear map $\varphi : S^{e+1} \rightarrow \mathbb{R}[t]_d$ with $\varphi(A) := \sum_{i,j=0}^e a_{ij} t^{i+j}$ for $A = (a_{ij})_{i,j=0,\dots,e} \in S^{e+1}$ from Remark 3.17. As mentioned in Remark 3.17, $\varphi(S_+^W) = \Sigma_{2k}$ for $W = \{0, \dots, k\}$, where the notation S_+^W is borrowed from Theorem 4.6. The latter implies $\varphi(S_+^{V_i}) = t^{2i} \Sigma_{2k}$ for $V_i := \{i, \dots, i + k\}$ with $0 \leq i \leq d - 2k$. Consequently, splitting the sum representing $\text{SOCF}_{2k+1,d}$ in (9) into two according to the parity of the exponents in t^0, \dots, t^{d-2k} , we obtain

$$\begin{aligned} \text{SOCF}_{2k+1,d} &= \sum_{i=0}^{e-k} t^{2i} \Sigma_{2k} + t \sum_{i=0}^{e-k-1} t^{2i} \Sigma_{2k} \\ &= \phi \left(\underbrace{\sum_{i=0}^{e-k} S_+^{V_i}}_{=:K} \right) + t \phi \left(\underbrace{\sum_{i=0}^{e-k-1} S_+^{V_i}}_{=:L} \right). \end{aligned}$$

Thus, $\text{SOCF}_{2k+1,d}$ is a linear image of $K \times L$ under the linear map $(A, B) \mapsto \varphi(A) + t \varphi(B)$. In view of Theorem 4.6, K and L are copies of $S_{+,k}^{e+1}$ and $S_{+,k}^e$ respectively. This gives the desired assertion. \square

Remark 4.10 Wang et al. [23, 24] suggest to use the cones $S_{+,E}^V$ from chordal graphs $G = (V, E)$ to approximate sparse sos polynomials. For a given finite set $A \subseteq \mathbb{Z}_+^n$, they consider the cone

$$\Sigma(A) := \{f \in \mathbb{R}[x_1, \dots, x_n] : f \text{ sos and } \text{supp}(f) \subseteq A\}$$

They suggest two iterative algorithms (one in [23] and another one in [24]) that take A as an input and produce a chordal graph $G = (V, E)$ with $V \subseteq \mathbb{Z}_+^n$ in order to use the image of the respective cone $S_{+,E}^V$ under the map

$$\varphi : (a_{\alpha,\beta})_{\alpha,\beta \in V} \mapsto \sum_{\alpha,\beta \in V} a_{\alpha,\beta} x^{\alpha+\beta}$$

as an approximation of $\Sigma(A)$. The algorithm in [23] produces a $G = (V, E)$ being a disjoint union of cliques. The graph $G = (V, E)$ in [23] is guaranteed to satisfy the equality

$$\Sigma(A) = \left\{ f \in \varphi(S_{+,E}^V) : \text{supp}(f) \subseteq A \right\} \quad (10)$$

(see Thm. 3.3 in [23]), but there is no guarantee that the clique number of G is small, or rather, the dependence of the clique number of G on the properties of A remains unexplored. Since the graph $G = (V, E)$ generated by an algorithm from [23] may have large cliques, in [24] another approach is suggested that generates a graph $G = (V, E)$ with a smaller number of edges. This other approach is heuristic in the sense that there are no theoretical guarantees for the equality $\Sigma(A) = \varphi(S_{+,E}^V)$ (see Example 3.5. in [24]). It would be interesting to study the semidefinite extension degree of the cones $\Sigma(A)$ and try to relate these cones to the cones $S_{+,E}^V$ in a non-heuristic way.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ahmadi, A.A., Majumdar, A.: DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. *SIAM J. Appl. Algebra Geom.* **3**, 193–230 (2019)
2. Agler, J., Helton, W., McCullough, S., Rodman, L.: Positive semidefinite matrices with a given sparsity pattern. *Linear Algebra Appl.* **107**, 101–149 (1988)
3. Alizadeh, F.: Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.* **5**, 13–51 (1995)
4. Averkov, G.: Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization. *SIAM J. Appl. Algebra Geom.* **3**, 128–151 (2019)
5. Averkov, G., Peters, B., Sager, S.: Convexification of polynomial optimization problems by means of monomial patterns. Preprint, [arXiv:1901.05675](https://arxiv.org/abs/1901.05675)
6. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications. SIAM, MPS, Philadelphia (2001)
7. Blekherman, G., Parrilo, P.A., Thomas, R.R. (eds.): Semidefinite Optimization and Convex Algebraic Geometry. SIAM, MOS, Philadelphia (2013)
8. Dressler, M., Ilman, S., de Wolff, T.: A positivstellensatz for sums of nonnegative circuit polynomials. *SIAM J. Appl. Algebra Geom.* **1**, 536–555 (2017)
9. Fawzi, H.: On representing the positive semidefinite cone using the second-order cone. *Math. Program.* **175**, 109–118 (2019)
10. Kathän, L., Naumann, H., Theobald, T.: A unified framework of SAGE and SONC polynomials and its duality theory. *Math. Comput.* **90**, 1297–1322 (2021)
11. Karlin, S., Studden, W.J.: Tchebycheff Systems: With Applications in Analysis and Statistics. Interscience, Wiley, New York (1966)
12. Knebusch, M., Scheiderer, C.: Real Algebra. A First Course. Universitext, Springer, Berlin (2022)
13. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11** 796–817 (2000/01)
14. Lasserre, J.B.: An Introduction to Polynomial and Semi-algebraic Optimization. Cambridge University Press, Cambridge (2015)
15. Murray, R., Chandrasekaran, V., Wierman, A.: Newton polytopes and relative entropy optimization. *Found. Comput. Math.* **21**, 1703–1737 (2021)
16. Nesterov, Y., Nemirovskii, A.: Interior-point Polynomial Algorithms in Convex Programming. SIAM Studies in Applied Mathematics 13. SIAM, Philadelphia (1994)
17. Putinar, M.: Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* **42**, 969–984 (1993)
18. Rosen, Z., Scholten, G., Vinzant, C.: Sparse moments of univariate step functions and allele frequency spectra. *Vietnam J. Math.* **50**, 523–544 (2022)
19. Scheiderer, C.: Second-order cone representation for convex sets in the plane. *SIAM J. Appl. Algebra Geom.* **5**, 114–139 (2021)
20. Scheiderer, C.: Work in progress (2023)
21. Stanley, R.: Enumerative Combinatorics, Volume 2 Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1999)
22. Vandenberghe, L., Andersen, M.S.: Chordal graphs and semidefinite optimization. *Found. Trends Optim.* **1**, 241–433 (2015)
23. Wang, J., Magron, V., Lasserre, J.B.: TSSOS: A moment-SOS hierarchy that exploits term sparsity. *SIAM J. Optim.* **31**, 30–58 (2021)
24. Wang, J., Magron, V., Lasserre, J.B.: Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. *SIAM J. Optim.* **31**, 114–141 (2021)
25. Zheng Y., Fantuzzi, G., Papachristodoulou, A., Wynn, A.: Fast ADMM for semidefinite programs with chordal sparsity. In: 2017 American Control Conference (ACC), pp. 3335–3340. IEEE