

# Disjunctive Sum of Squares

Amir Ali Ahmadi\*, Sanjeeb Dash†, Yixuan Hua\*, Bartolomeo Stellato\*

## Abstract

We introduce the concept of *disjunctive sum of squares* for certifying nonnegativity of polynomials. Unlike the popular sum of squares approach where nonnegativity is certified by a single algebraic identity, the disjunctive sum of squares approach certifies nonnegativity with multiple algebraic identities which can be found in parallel. Our main result is a disjunctive Positivstellensatz proving that we can keep the degree of each algebraic identity as low as the degree of the polynomial whose nonnegativity is in question. Based on this result, we construct a semidefinite programming based converging hierarchy of lower bounds for the problem of minimizing a polynomial over a compact basic semialgebraic set, where the size of the largest semidefinite constraint is fixed throughout the hierarchy. We further prove a second disjunctive Positivstellensatz which leads to an optimization-free hierarchy for polynomial optimization. We specialize this result to the problem of proving copositivity of matrices. Finally, we describe how the disjunctive sum of squares approach can be combined with a branch-and-bound algorithm and we present numerical experiments on polynomial, copositive, and combinatorial optimization problems.

## 1 Introduction

A polynomial  $p : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be *nonnegative* if  $p(x) \geq 0$  for all  $x \in \mathbb{R}^n$ . The problem of optimizing over the intersection of the cone of nonnegative polynomials of a given degree with an affine subspace is fundamental to optimization and algebraic geometry, with broad applications in polynomial optimization [27], control systems and robotics [23], [51], probability theory [6], and statistics [15], to name a few; see the textbooks [28], [8], [29], or survey papers [31], [22] for many other applications and further context.

A polynomial  $p : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be a *sum of squares* (sos) if it can be expressed as  $p(x) = \sum_i q_i^2(x)$  for some polynomials  $q_i : \mathbb{R}^n \rightarrow \mathbb{R}$ . A sum of squares decomposition of a polynomial provides a single algebraic identity that certifies its nonnegativity. Moreover, as

---

\*Princeton University, Operations Research and Financial Engineering. AAA was partially supported by the Sloan Fellowship, the Princeton AI Lab Seed Grant, the Princeton SEAS Innovation Grant, and a Research Gift in Mathematical Optimization. YH was partially supported by the Princeton AI Lab Seed Grant, the Princeton SEAS Innovation Grant, and the IBM PhD Fellowship.

†IBM TJ Watson Research Center.

is well known by now, such a decomposition (assuming it exists) can be obtained by solving a semidefinite program (SDP) [39], [27], [40].

Throughout the paper, we denote the set of nonnegative polynomials and sos polynomials in  $n$  variables and degree at most  $d$  by  $P_{n,d}$  and  $\Sigma_{n,d}$ , respectively. It has been known since the seminal work of Hilbert [25] that not every nonnegative polynomial is a sum of squares. In fact, Hilbert established that  $P_{n,d}$  and  $\Sigma_{n,d}$  coincide only in three specific cases: when  $n = 1$ ,  $d = 2$ , or  $(n, d) = (2, 4)$ . The first example of a nonnegative polynomial which is not sos was constructed by Motzkin many years later [37]:

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

Despite the existence of such examples, it turns out that a proof of nonnegativity based on a single algebraic identity is always possible though at the expense of increasing the degree of the sos polynomials that appear in the identity. More specifically, it follows from Artin’s solution to Hilbert’s 17th problem [4] that for every nonnegative polynomial  $p$ , there exists a nonzero sos polynomial  $q$ , such that the product  $pq$  is sos (and hence nonnegativity of  $p$  is algebraically certified by a representation as a ratio of two sos polynomials). For example, for the Motzkin polynomial, we have the following degree-8 identity:

$$(x_1^2 + x_2^2)M(x_1, x_2) = x_1^2(1 - x_2^2)^2 + x_2^2(1 - x_1^2)^2 + x_1^2 x_2^2 (x_1^2 + x_2^2 - 2)^2. \quad (2)$$

By utilizing Artin’s theorem, one can design a complete SDP-based hierarchy for proving nonnegativity of a polynomial  $p$ . In level  $j$  of this hierarchy, one searches for a nonzero sos polynomial  $q$  of degree  $2j$  that makes the product  $pq$  also sos.<sup>1</sup> However, the degree of the polynomial  $pq$  can be much higher than the degree of  $p$  (see, e.g., [7]) and is upper bounded currently only by a tower of five exponentials in the number of variables and degree of  $p$  [34]. Since the size of the SDP underlying an sos constraint grows rapidly with the degree of the sos polynomial, the SDP hierarchy based on Artin’s theorem can quickly become impractical in certain applications. This challenge also arises in the context of other sum of squares based SDP hierarchies (e.g., in [39], [27], [40]) which rely on related seminal theorems of algebraic geometry such as the Positivstellensätze of Stengle [49], Schmüdgen [47], or Putinar [42].

With this context in mind, the central question that motivates our research is the following: “*Can one lower the complexity of a sum of squares based proof of polynomial nonnegativity by constructing several algebraic identities instead of a single one?*” Our primary notion of complexity here is the degree of the sos polynomials that appear in the proof of nonnegativity. However, our results also simplify the structure of these sos polynomials making the search for them amenable to cheaper classes of conic programs, such as linear programs. As an example, consider the following proof of nonnegativity of the Motzkin polynomial using two degree-6 algebraic identities:

$$\begin{aligned} M(x_1, x_2) &= (1 - x_1 x_2)^2 + (x_1^2 x_2 - x_1 x_2^2)^2 + 2x_1 x_2 (1 - x_1 x_2)^2 \\ &= (1 + x_1 x_2)^2 + (x_1^2 x_2 + x_1 x_2^2)^2 - 2x_1 x_2 (1 + x_1 x_2)^2. \end{aligned} \quad (3)$$

---

<sup>1</sup>Unlike the approach we later propose in this paper, one cannot use this SDP hierarchy to *impose* a nonnegativity constraint on a polynomial  $p$ . This is because of the bilinear matrix inequality that arises from the sos constraint on the product of the two unknown polynomials  $p$  and  $q$ .

Indeed, for every point  $(x_1, x_2) \in \mathbb{R}^2$ , either  $x_1x_2 \geq 0$  or  $x_1x_2 < 0$ . When  $x_1x_2 \geq 0$ , nonnegativity of  $M(x_1, x_2)$  follows from the first identity. Similarly, when  $x_1x_2 < 0$  or equivalently  $-x_1x_2 > 0$ , nonnegativity of  $M(x_1, x_2)$  follows from the second identity. We call a proof of this type a *disjunctive sum of squares* proof of nonnegativity. Comparing the two algebraic proofs in (2) and (3), we see that the former involves an sos polynomial of degree 8 while the latter involves two sos polynomials of degree 6, which is the same as the degree of the Motzkin polynomial. From a computational standpoint, this means that the size of the semidefinite constraint underlying the second proof is smaller.<sup>2</sup> Moreover, the sos polynomials appearing in (3) have a special structure (known as “diagonal dominance”; see [3]), that enables the search for this particular disjunctive sum of squares proof to be carried out by solving a linear program. This is, in general, much cheaper than solving a semidefinite program.

Our main contribution in this paper is to formalize the concept of disjunctive sum of squares proofs of nonnegativity and show that *low-degree* proofs of this type always exist under standard assumptions. We outline our contributions in some more detail next.

## 1.1 Organization and Main Contributions

The remainder of this paper is organized as follows. In Section 2, we define the notion of a disjunctive sum of squares proof of nonnegativity of a polynomial. We also show that a natural alternating maximization approach based on SDP can find low-degree disjunctive sum of squares proofs of nonnegativity of several classical non-sos polynomials.

In Section 3, we prove a “disjunctive Positivstellensatz”, which shows that the degree of the sos polynomials appearing in a disjunctive sum of squares proof can be bounded above by the degree of the polynomial whose nonnegativity is being certified. Sections 3.1 and 3.2 are devoted to the proof of this statement. Based on this result, Section 3.3 introduces an SDP-based hierarchy, with semidefinite constraints of *fixed size*, which yields a converging sequence of lower bounds on the minimum value of a homogeneous polynomial over the unit sphere. Our analysis also yields a convergence rate for this hierarchy.

In Section 4, we prove a second disjunctive Positivstellensatz, which shows that a proof of nonnegativity of a polynomial can be given by simply checking nonnegativity of the coefficients of the same polynomial under several linear changes of coordinates. These linear transformations are explicit and derived from disjunctions based on simplicial partitions of the Euclidean space. An upper bound on the number of linear transformations is provided.

In Section 5, we demonstrate how our disjunctive sum of squares approach can be applied to constrained polynomial optimization problems. In Section 5.1, using a reduction from [2], we construct a semidefinite programming based converging hierarchy of lower bounds for the general problem of minimizing a polynomial over a compact basic semialgebraic set, where the size of the largest semidefinite constraint is fixed throughout the hierarchy. In Section 5.2, we focus on the case where the constraint set is the unit simplex, and we provide

---

<sup>2</sup>We recall that, imposing an sos constraint on a polynomial in  $n$  variables and degree  $2d$  results in a positive semidefinite constraint on a matrix of size  $\binom{n+d}{d} \times \binom{n+d}{d}$  (see, e.g., [39], [40]).

a specialized disjunctive Positivstellensatz for certifying copositivity of matrices.

In Section 6, we show how one can more practically search for disjunctive sum of squares proofs by dynamically partitioning the Euclidean space through a spatial branch-and-bound framework. As examples, we propose branch-and-bound algorithms for minimizing a homogeneous polynomial over the unit sphere and for certifying copositivity of matrices. In Section 6.3, we present some numerical experiments. Finally, Section 7 concludes our paper and highlights potential future directions.

## 2 SOS with Respect to an Algebraic Disjunction

In this section, we introduce the concept of sum of squares with respect to an algebraic disjunction, which is central to our framework. Throughout the paper, we denote the degree of a polynomial  $p$  by  $\deg p$ .

**Definition 1** (Algebraic Disjunction). *Let  $r, n_1, \dots, n_r$  be positive integers. We say that a collection of polynomials  $\mathcal{D} = \{\{q_{k,j}\}_{j=1, \dots, n_k}, k = 1, \dots, r\}$  forms an algebraic disjunction of  $\mathbb{R}^n$  if*

$$\mathbb{R}^n = \bigcup_{k=1}^r \Omega_k, \text{ where } \Omega_k = \{x \in \mathbb{R}^n \mid q_{k,j}(x) \geq 0, j = 1, \dots, n_k\}.$$

**Definition 2.** *Let  $p$  be a polynomial in  $n$  variables and degree at most  $d$ ,  $\bar{d} \geq d$  be an integer, and  $\mathcal{D} = \{\{q_{k,j}\}_{j=1, \dots, n_k}, k = 1, \dots, r\}$  be an algebraic disjunction of  $\mathbb{R}^n$ . We say that  $p$  is sos of order  $\bar{d}$  with respect to  $\mathcal{D}$  (or  $p$  is  $\text{sos}_{\mathcal{D}}^{\bar{d}}$ ) if there exist sos polynomials  $s_{k,j}$  for  $k = 1, \dots, r$  and  $j = 0, 1, \dots, n_k$ , such that*

$$p(x) = s_{k,0}(x) + \sum_{j=1}^{n_k} s_{k,j}(x)q_{k,j}(x), \quad k = 1, \dots, r, \quad (4)$$

where  $\deg s_{k,0} \leq \bar{d}$  and  $\deg(s_{k,j}q_{k,j}) \leq \bar{d}$  for  $k = 1, \dots, r$  and  $j = 1, \dots, n_k$ . We denote the set of  $\text{sos}_{\mathcal{D}}^{\bar{d}}$  polynomials in  $n$  variables and degree at most  $d$  as  $\Sigma_{n,d}^{\bar{d}}(\mathcal{D})$ .

Observe that for any algebraic disjunction  $\mathcal{D}$  and integer  $\bar{d} \geq d$ , the set  $\Sigma_{n,d}^{\bar{d}}(\mathcal{D})$  is a convex cone and lies between  $\Sigma_{n,d}$  and  $P_{n,d}$ . Indeed, by setting  $s_{k,j} = 0$  for  $k = 1, \dots, r$  and  $j = 1, \dots, n_k$ , we note that  $\Sigma_{n,d} \subseteq \Sigma_{n,d}^{\bar{d}}(\mathcal{D})$ . This inclusion can be strict even when  $\bar{d} = d$ ; e.g., the Motzkin polynomial in (1) belongs to  $\Sigma_{2,6}^6(\{\{+x_1x_2\}, \{-x_1x_2\}\})$ , as shown by the identities in (3). To see that  $\Sigma_{n,d}^{\bar{d}}(\mathcal{D}) \subseteq P_{n,d}$ , note that for any point  $x \in \Omega_k$ , we have  $q_{k,j}(x) \geq 0$  for all  $j = 1, \dots, n_k$ , and therefore, in view of (4),  $p(x) \geq 0$ . Consequently,  $p$  is nonnegative over each subregion  $\Omega_k$ , and thus nonnegative over  $\mathbb{R}^n$ . We call the  $r$  algebraic identities in (4) a disjunctive sos proof of nonnegativity of  $p$ .

From a computational standpoint, for any fixed  $\bar{d}$ , linear optimization over the intersection of an affine subspace with the cone  $\Sigma_{n,d}^{\bar{d}}(\mathcal{D})$  can be carried out by an SDP of size polynomial in  $n$ . We also note that for implementation purposes, there is no

need to include  $s_{k,0}$  as a decision variable; instead, one can impose the constraint that  $p(x) - \sum_{j=1}^{n_k} s_{k,j}(x)q_{k,j}(x) \in \Sigma_{n,\bar{d}}$  for  $k = 1, \dots, r$ .

Unlike the well-known hierarchies for proving nonnegativity of a polynomial (e.g., based on the Positivstellensätze of Artin [4], Reznick [43], Stengle [49], [39], or Putinar [42], [27]) that increase the degrees of the sos polynomials involved in the proof at each level of the hierarchy, our approach keeps the degrees of these sos polynomials fixed (and low). We instead increase the number of subsets in the algebraic disjunction  $\mathcal{D}$  (see Theorem 1 and Theorem 8).

A natural question is how to search for disjunctive proofs of nonnegativity of a polynomial  $p$ ; i.e., for the identities in (4). Note that as written, the joint search for the polynomials  $q_{k,j}$  that define the algebraic disjunction  $\mathcal{D}$  and the sos multipliers  $s_{k,j}$  cannot be formulated as a convex optimization problem. An obvious workaround to this issue is to alternate the search for these decision variables. The details of this approach are described in the rest of this section together with some examples. In Section 3 and Section 4, we present a more theoretically sound approach to finding disjunctive proofs of nonnegativity. We show that by appropriately fixing a family of algebraic disjunctions, we can always construct low-degree proofs of nonnegativity using convex optimization.

**An alternating maximization algorithm.** In the joint search for the polynomials  $q_{k,j}$  and the sos multipliers  $s_{k,j}$  in (4), a challenge is to ensure that the polynomials  $q_{k,j}$  form a valid algebraic disjunction (see Definition 1). A simple way to guarantee this is to take  $\ell$  polynomials  $h_1, \dots, h_\ell$ , let  $r = 2^\ell$ ,  $\varepsilon \in \{-1, 1\}^{r \times \ell}$  with all rows of  $\varepsilon$  distinct, and let

$$q_{k,j} = \varepsilon_{k,j} h_j, \quad k = 1, \dots, r, \quad j = 1, \dots, \ell.$$

With this construction, it is easy to see that the union of the corresponding  $r$  subregions  $\Omega_1, \dots, \Omega_r$  as defined in Definition 1 is  $\mathbb{R}^n$ . Building on this, given a polynomial  $p$  of degree  $d$  and an integer  $\bar{d} \geq d$ , the problem boils down to finding polynomials  $h_j$  and sos multipliers  $s_{k,j}$  such that

$$p(x) = s_{k,0}(x) + \sum_{j=1}^{\ell} \varepsilon_{k,j} h_j(x) s_{k,j}(x), \quad k = 1, \dots, r, \quad (5)$$

where  $\deg s_{k,0} \leq \bar{d}$ ,  $\deg(h_j s_{k,j}) \leq \bar{d}$  for  $k = 1, \dots, r$  and  $j = 1, \dots, \ell$ . To approach this problem using convex optimization, we alternate between the search for the polynomials  $h_j$  and the sos multipliers  $s_{k,j}$ . To track progress across iterations, we always maximize a lower bound on  $p$  that can be certified with identities of the form (5). More concretely, we alternate between the following two steps, which can both be reformulated as a semidefinite program:

1. Given polynomials  $h_j^*$  for  $j = 1, \dots, \ell$ , solve the SDP

$$\begin{aligned} & \max_{\gamma, s_{k,j}} \quad \gamma \\ & \text{s.t.} \quad p(x) - \gamma = s_{k,0}(x) + \sum_{j=1}^{\ell} \varepsilon_{k,j} h_j^*(x) s_{k,j}(x), \quad k = 1, \dots, r \\ & \quad \deg s_{k,0} \leq \bar{d}, \quad \deg(h_j^* s_{k,j}) \leq \bar{d}, \quad k = 1, \dots, r, \quad j = 1, \dots, \ell, \\ & \quad s_{k,j} \text{ is sos}, \quad k = 1, \dots, r, \quad j = 0, \dots, \ell. \end{aligned}$$

2. Given sos polynomials  $s_{k,j}^*$  for  $k = 1, \dots, r$  and  $j = 1, \dots, \ell$ , solve the SDP<sup>3</sup>

$$\begin{aligned} & \max_{\gamma, h_j, s_{k,0}} \quad \gamma \\ & \text{s.t.} \quad p(x) - \gamma = s_{k,0}(x) + \sum_{j=1}^{\ell} \varepsilon_{k,j} h_j(x) s_{k,j}^*(x), \quad k = 1, \dots, r \\ & \quad \deg s_{k,0} \leq \bar{d}, \quad \deg(h_j s_{k,j}^*) \leq \bar{d}, \quad k = 1, \dots, r, \quad j = 1, \dots, \ell, \\ & \quad s_{k,0} \text{ is sos}, \quad k = 1, \dots, r. \end{aligned}$$

To run Step 1, we let the polynomials  $h_j^*$  for  $j = 1, \dots, \ell$  be arbitrary (e.g., chosen at random) if we are in the first iteration, or an optimal solution to Step 2 if we are in subsequent odd iterations. Similarly, to run Step 2, we let  $s_{k,j}^*$  for  $k = 1, \dots, r$  and  $j = 1, \dots, \ell$  be an optimal solution to Step 1.

Applying the alternating maximization algorithm to the Motzkin polynomial in (1) yields two additional disjunctive sos proofs of nonnegativity. In this example, we set  $\ell = 1$  and  $\bar{d} = 6$ . To initialize the algorithm, we generate  $h_1$  randomly, once of degree 1 and once of degree 4. The alternating maximization algorithm converges in fewer than 20 iterations and produces (after some minor post-processing to polish the coefficients to rational numbers) the following two algebraic identities, respectively:

$$\begin{aligned} M(x_1, x_2) &= (1 - x_1 x_2^2)^2 + (x_1^2 x_2 - x_1 x_2)^2 + 2x_1(x_1 x_2 - x_2)^2 \\ &= (1 + x_1 x_2^2)^2 + (x_1^2 x_2 + x_1 x_2)^2 - 2x_1(x_1 x_2 + x_2)^2, \end{aligned} \tag{6}$$

$$\begin{aligned} M(x_1, x_2) &= \frac{1}{2}(x_1^2 x_2 + x_2^3 - 2x_2)^2 + (x_2^2 - 1)^2 + \frac{1}{2}(x_1^4 - x_2^4 - 2x_1^2 + 2x_2^2)x_2^2 \\ &= \frac{1}{2}(x_1^3 + x_1 x_2^2 - 2x_1)^2 + (x_1^2 - 1)^2 - \frac{1}{2}(x_1^4 - x_2^4 - 2x_1^2 + 2x_2^2)x_1^2. \end{aligned} \tag{7}$$

Figure 1 demonstrates the subregions associated with the three degree-6 disjunctive sos proofs of nonnegativity of the Motzkin polynomial in (3), (6), and (7).

Similarly, we are able to find disjunctive proofs of nonnegativity of several polynomials that are known to *not* be a sum of squares. We present three such certificates below.

---

<sup>3</sup>In fact, both Step 1 and Step 2 can be decomposed into  $r$  independent SDPs and solved in parallel; see Remark 3 in Section 3.3.

- Choi-Lam-1 [13]. Let  $CL_1(x_1, x_2, x_3, x_4) = -4x_1x_2x_3x_4 + x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2 + x_4^4$ . Then  $CL_1 = s_1 + h s_2 = s_3 - h s_4$ , with  $h = x_1x_2$ ,

$$\begin{aligned} s_1 &= (x_1x_3 - x_2x_3)^2 + (x_1x_2 - x_4^2)^2, & s_2 &= 2(-x_3 + x_4)^2, \\ s_3 &= (x_1x_3 + x_2x_3)^2 + (x_1x_2 + x_4^2)^2, & s_4 &= 2(x_3 + x_4)^2. \end{aligned}$$

- Choi-Lam-2 [14]. Let  $CL_2(x_1, x_2, x_3) = -3x_1^2x_2^2x_3^2 + x_1^2x_3^4 + x_1^4x_2^2 + x_2^4x_3^2$ . Then  $CL_2 = s_1 + h s_2 = s_3 - h s_4$ , with  $h = x_1x_2$ ,

$$\begin{aligned} s_1 &= (-x_1x_3^2 + x_2^2x_2)^2 + (-x_1x_2x_3 + x_2^2x_3)^2, & s_2 &= 2(-x_1x_3 + x_2x_3)^2, \\ s_3 &= (x_1x_3^2 + x_1^2x_2)^2 + (x_1x_2x_3 + x_2^2x_3)^2, & s_4 &= 2(x_1x_3 + x_2x_3)^2. \end{aligned}$$

- Stengle-1 [50]. Let  $S_1(x_1, x_2, x_3) = -2x_1x_2^2x_3^3 + x_1^2x_3^4 - 2x_1^3x_2^2x_3 + x_1^3x_3^3 + 2x_1^4x_2^2 + x_1^6 + x_2^4x_3^2$ . Then  $S_1 = s_1 + h s_2 = s_3 - h s_4$ , with  $h = x_1x_3$ ,

$$\begin{aligned} s_1 &= (-x_1x_3^2 - x_1^3 + x_2^2x_3)^2, & s_2 &= (x_1x_3)^2, \\ s_3 &= (x_1x_3^2 + \frac{1}{2}x_1^2x_3 + \frac{7}{10}x_1^3 - \frac{6}{7}x_2^2x_3)^2 + \frac{1}{14}(x_1x_2x_3)^2 \\ &\quad + \frac{7}{20}(x_1^2x_3 - x_1^3 + \frac{5}{7}x_2^2x_3)^2 + \frac{4}{25}(x_1^3 - \frac{5}{7}x_2^2x_3)^2 + \frac{1}{196}(x_2^2x_3)^2, \\ s_4 &= \frac{2}{7}(-\frac{1}{2}x_1x_2 + x_2x_3)^2. \end{aligned}$$

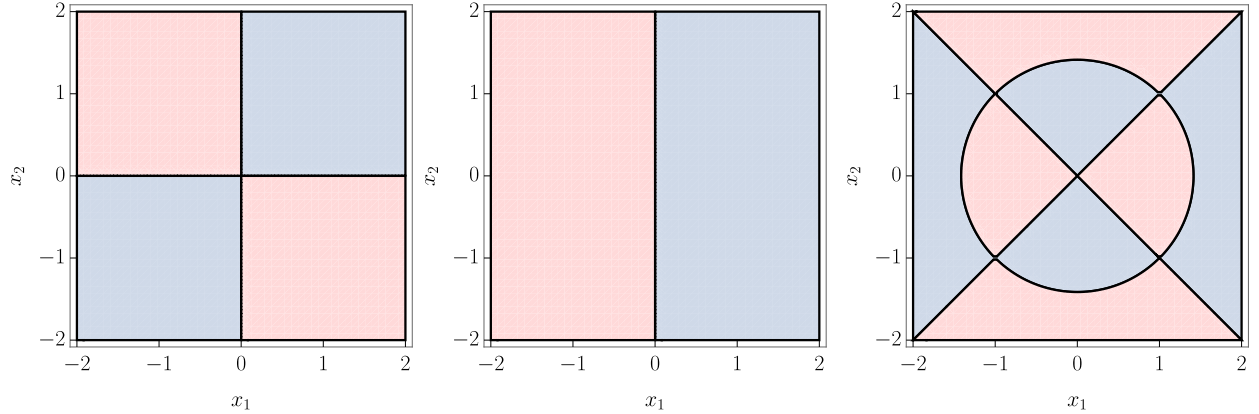


Figure 1: Three different partitions of  $\mathbb{R}^2$  corresponding to three different degree-6 disjunctive sos proofs of nonnegativity of the Motzkin polynomial (see (1)). Left: subregions correspond to the identity in (3) and are defined by  $x_1x_2 \geq 0$  and  $-x_1x_2 \geq 0$ . Middle: subregions correspond to the identity in (6) and are defined by  $x_1 \geq 0$  and  $-x_1 \geq 0$ . Right: subregions correspond to the identity in (7) and are defined by  $x_1^4 - x_2^4 - 2x_1^2 + 2x_2^2 \geq 0$  and  $-(x_1^4 - x_2^4 - 2x_1^2 + 2x_2^2) \geq 0$ .

### 3 A Disjunctive Positivstellensatz

In this section, we introduce our main result in Theorem 1 which shows that one can a priori fix a family of low-degree algebraic disjunctions with respect to which there always exists a low-degree disjunctive sos proof of nonnegativity.

Let us establish some basic notation related to *forms*, i.e., homogeneous polynomials, which will be useful in the remainder of this paper. We denote the set of forms, nonnegative forms, and sum of squares forms in  $n$  variables and degree  $d$  by  $H_{n,d}$ ,  $\widehat{P}_{n,d}$ , and  $\widehat{\Sigma}_{n,d}$  respectively. Observe that  $\widehat{P}_{n,d} = P_{n,d} \cap H_{n,d}$  and  $\widehat{\Sigma}_{n,d} = \Sigma_{n,d} \cap H_{n,d}$ . Additionally, we say that a form  $p$  in  $n$  variables is *positive definite* if  $p(x) > 0$  for all  $x \in \mathbb{R}^n$  and  $x \neq 0$ . We are now ready to present our main result.

**Theorem 1** (Disjunctive Positivstellensatz). *For any fixed dimension  $n$  and even degree  $d$ , there exists an explicit family of algebraic disjunctions  $\mathcal{D}^m$ , indexed by  $m$ , such that for any positive definite form  $p \in H_{n,d}$ , we have  $p \in \Sigma_{n,d}^d(\mathcal{D}^m)$  for some  $m$ . Furthermore, each subset in  $\mathcal{D}^m$  contains exactly one form of degree at most  $d$ .<sup>4</sup>*

We prove Theorem 1 as follows. In Section 3.1, we introduce local Positivstellensätze, which provide a single algebraic identity proving nonnegativity of a polynomial in the neighborhood of a given point where it is known to be positive. In Section 3.2, we combine these neighborhoods to cover the entire Euclidean space and use the corresponding algebraic identities to construct a disjunctive sos proof of (global) nonnegativity. In Section 3.3, we show how this theorem leads to a hierarchy of SDPs—with semidefinite constraints of fixed size—that produces arbitrarily tight lower bounds on the minimum value of a homogeneous polynomial over the unit sphere.

### 3.1 Local Positivstellensätze

For a polynomial  $p$  in  $n$  variables and degree  $d$ , we denote its homogenization by

$$p_h(x, x_{n+1}) := x_{n+1}^d p(x/x_{n+1}).$$

It is straightforward to check that homogenization preserves both nonnegativity and the sos property; that is,  $p \in P_{n,d}$  (resp.,  $p \in \Sigma_{n,d}$ ) if and only if  $p_h \in \widehat{P}_{n+1,d}$  (resp.,  $p_h \in \widehat{\Sigma}_{n+1,d}$ ). Hence, given a general (not necessarily homogeneous) polynomial  $p$ , we can construct a disjunctive sos proof of nonnegativity of  $p$  by first obtaining one for  $p_h$  and then setting the last variable to 1 in the identities. This justifies our focus on homogeneous polynomials. Moreover, since odd-degree polynomials cannot be nonnegative on the entire Euclidean space, it suffices to consider even-degree forms.

We now turn to establishing local certificates of nonnegativity. Suppose a polynomial  $p$  takes a positive value at some point  $x^* \in \mathbb{R}^n$ . By continuity,  $p$  remains nonnegative within a sufficiently small neighborhood of  $x^*$ . This observation motivates the development of a local Positivstellensatz: an algebraic identity demonstrating that if the angle between a point  $x$  and  $x^*$  is sufficiently small, then  $p(x) \geq 0$ .

We introduce some notation for the theorems that follow. We denote the unit sphere in  $\mathbb{R}^n$  by  $\mathbb{S}^{n-1}$ , and the  $n \times n$  identity matrix by  $I_n$ . For a polynomial  $p$ , we denote its largest coefficient in absolute value (resp. the sum of the absolute values of its coefficients) in the standard monomial basis by  $\|p\|_\infty$  (resp. by  $\|p\|_1$ ).

---

<sup>4</sup>Our proof in fact provides several constructions: the forms in the algebraic disjunction may be chosen to have any even degree between 2 and  $d$ .

**Theorem 2.** For a polynomial  $p \in H_{n,d}$  with  $d$  even, if  $p(x^*) > 0$  for some  $x^* \in \mathbb{S}^{n-1}$ , then for  $t > 0$  sufficiently large,  $p(x)$  can be written as

$$p(x) = \left( \frac{p(x^*)}{2} (x^T x^*)^d - th_d(x; x^*) \right) + \sigma(x), \quad (8)$$

where  $\sigma(x) \in \widehat{\Sigma}_{n,d}$  and

$$h_d(x; x^*) := \sum_{k=1}^{d/2} (\|x\|_2^2 - (x^T x^*)^2)^k (x^T x^*)^{d-2k}. \quad (9)$$

In particular, the above statement holds for any

$$t \geq T(p; x^*) := \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \left( \frac{n}{2p(x^*)} \|p(Ux)\|_\infty^2 + \|p(Ux)\|_1 \right). \quad (10)$$

*Remark 1.* The identity in (8) shows the implication that

$$\frac{p(x^*)}{2} (x^T x^*)^d - th_d(x; x^*) \geq 0 \quad \Rightarrow \quad p(x) \geq 0. \quad (11)$$

To gain geometric intuition, consider the region defined by

$$A(x^T x^*)^d - th_d(x; x^*) \geq 0, \quad (12)$$

where  $A$  and  $t$  are positive scalars. Let  $\theta$  denote the angle between  $x$  and  $x^*$ . Since  $x^T x^* = \|x\|_2 \cos \theta$ , the inequality above is equivalent to  $\sum_{k=1}^{d/2} (\tan \theta)^{2k} \leq A/t$ . Therefore, the subregion corresponds to  $|\theta| \leq \arctan z$ , where  $z$  is the unique positive real root of  $\sum_{k=1}^{d/2} z^{2k} = A/t$ . Consequently, identity (8) shows that a sufficiently small angle between  $x$  and  $x^*$  implies nonnegativity of  $p(x)$ . Figure 2 illustrates the intersection of the region defined by (12) and  $\mathbb{S}^{n-1}$ , which is a spherical cap.

To prove Theorem 2, we first establish the following lemma. Let us recall the multi-index notation that will be used in its proof. For an  $n$ -tuple of nonnegative integers  $\alpha := (\alpha_1, \dots, \alpha_n)$ , and a vector  $x \in \mathbb{R}^n$ , we write  $x^\alpha$  to denote  $x_1^{\alpha_1} \dots x_n^{\alpha_n}$  and  $|\alpha|$  to denote  $\sum_{i=1}^n \alpha_i$ . For two multi-indices  $\alpha$  and  $\beta$ , we denote by  $\alpha + \beta$  their componentwise sum.

**Lemma 3.** Consider a (not necessarily homogeneous) polynomial  $p$  in  $n$  variables and degree at most  $d$ , where  $d$  is even. If  $p(x^*) > 0$  for some  $x^* \in \mathbb{R}^n$ , then for  $t > 0$  sufficiently large,  $p(x)$  can be written as

$$p(x) = \left( \frac{p(x^*)}{2} - tg_d(x - x^*) \right) + \sigma(x),$$

where  $\sigma(x) \in \Sigma_{n,d}$  and

$$g_d(x) := \sum_{k=1}^{d/2} \|x\|_2^{2k}. \quad (13)$$

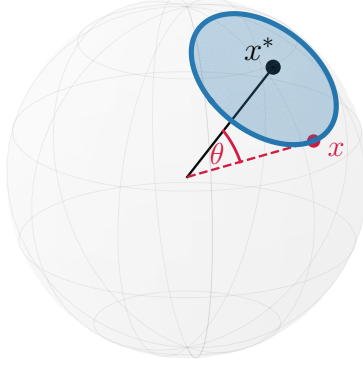


Figure 2: A spherical cap centered at  $x^*$ : the local identity (8) certifies  $p(x) \geq 0$  for every  $x \in \mathbb{S}^{n-1}$  within angle  $\theta$  of  $x^*$ .

In particular, the above statement holds for any

$$t \geq \tilde{T}(p; x^*) := \frac{n}{2p(x^*)} \|q\|_\infty^2 + \|q\|_1,$$

where  $q(x) = p(x + x^*) - p(x^*)$ .

*Proof.* Without loss of generality, we assume  $x^* = 0$ ; otherwise, applying the statement of the lemma to the polynomial  $p(x + x^*)$  and recalling that the existence of an sos decomposition is invariant under an affine change of variables, yields the desired result. Let

$$r_t(x) := p(x) + tg_d(x) - \frac{p(0)}{2} = q(x) + tg_d(x) + \frac{p(0)}{2}.$$

Thus, our task is to show that  $r_t(x)$  is sos for  $t \geq \tilde{T}(p; 0)$ . To do this, we work with the following representation of the summands:

$$\begin{aligned} q(x) + \frac{p(0)}{2} &= z(x)^T Q z(x) = \sum_{\alpha: |\alpha| \leq d/2} \sum_{\beta: |\beta| \leq d/2} Q_{\alpha, \beta} x^{\alpha + \beta}, \\ g_d(x) &= z(x)^T D z(x) = \sum_{\alpha: |\alpha| \leq d/2} \sum_{\beta: |\beta| \leq d/2} D_{\alpha, \beta} x^{\alpha + \beta}. \end{aligned}$$

Here,  $z(x) := (1, x_1, \dots, x_n, \dots, x_1^{d/2}, \dots, x_n^{d/2})$  is the vector of monomials of degree at most  $d/2$ , and  $Q$  and  $D$  are symmetric matrices. We use the multi-index notation  $Q_{\alpha, \beta}$  (resp.  $D_{\alpha, \beta}$ ) to refer to the appropriate entry of the matrix  $Q$  (resp.  $D$ ).

Since  $g_d(0) = 0$ , we have  $D_{0,0} = 0$ . Observe that  $g_d(x)$  is a positive linear combination of all monomials of the form  $x^{2\alpha}$ , where the multi-index  $\alpha \neq 0$  and  $|\alpha| \leq d/2$ . Thus, we can take  $D$  to be diagonal with  $D_{\alpha, \alpha} \geq 1$  for every nonzero multi-index  $\alpha$ . For the matrix  $Q$ , note that  $q(0) = 0$ , so  $Q_{0,0} = p(0)/2$ . Furthermore, the entries of  $Q$  can be rearranged such that  $Q_{0,\alpha} = Q_{\alpha,0} = 0$  for all  $\alpha$  with  $|\alpha| \in \{2, \dots, d/2\}$ . This is possible because each

monomial  $x^\alpha$  with  $|\alpha| \geq 2$  can be written as the product  $x^\beta x^\gamma$  for some nonzero multi-indices  $\beta, \gamma$  with  $\beta + \gamma = \alpha$ . With this rearrangement, we can write

$$\begin{aligned} r_t(x) &= z(x)^T(Q + tD)z(x) \\ &= \frac{p(0)}{2n} \sum_{i=1}^n \left(1 + \frac{nq_i}{p(0)}x_i\right)^2 + z(x)^T(\hat{Q} + tD)z(x), \end{aligned} \quad (14)$$

where  $q_i$  denotes the coefficient of the term  $x_i$  in  $q(x)$ , and  $\hat{Q}$  is defined entrywise as

$$\hat{Q}_{\alpha,\beta} = \begin{cases} 0 & \text{if } \alpha = 0 \text{ or } \beta = 0, \\ Q_{\alpha,\beta} - \frac{nq_i^2}{2p(0)} & \text{if } (\alpha, \beta) = (e_i, e_i), \quad i = 1, \dots, n, \\ Q_{\alpha,\beta} & \text{otherwise.} \end{cases}$$

Here,  $e_i$  is the  $i$ -th standard basis vector of  $\mathbb{R}^n$  (so that  $x^{e_i} = x_i$  in multi-index notation). Finally, one can verify that when  $t \geq \tilde{T}(p; 0)$ , the matrix  $\hat{Q} + tD$  is diagonally dominant<sup>5</sup> and thus positive semidefinite. Therefore, when  $t \geq \tilde{T}(p; 0)$ , the polynomial  $z(x)^T(\hat{Q} + tD)z(x)$  is sos. In view of (14), when  $t \geq \tilde{T}(p; 0)$ ,  $r_t(x)$  is also sos. ■

We are now ready to prove Theorem 2.

*Proof of Theorem 2.* We first consider the special case where  $x^* = e_1$ . Let  $z := (x_2, \dots, x_n)$ . We dehomogenize  $p$  by defining  $\tilde{p}(z) := p(1, z)$ . Then  $\tilde{p}(z)$  is a polynomial with  $\tilde{p}(0) = p(e_1) > 0$ . By Lemma 3, whenever  $t \geq \tilde{T}(\tilde{p}; 0)$ ,  $\tilde{p}(z)$  admits the representation

$$\tilde{p}(z) = \left( \frac{p(e_1)}{2} - tg_d(z) \right) + \tilde{\sigma}(z),$$

where  $\tilde{\sigma}(z) \in \Sigma_{n-1,d}$  and  $g_d$  is defined as in (13). Next, homogenizing  $\tilde{p}(z)$  yields the identity:

$$\begin{aligned} p(x) &= p(x_1, z) = x_1^d \tilde{p}(z/x_1) \\ &= \left( \frac{p(e_1)}{2} x_1^d - t \sum_{k=1}^{d/2} \|z\|_2^{2k} x_1^{d-2k} \right) + x_1^d \tilde{\sigma}(z/x_1) \\ &= \left( \frac{p(e_1)}{2} x_1^d - t \sum_{k=1}^{d/2} (\|x\|_2^2 - x_1^2)^k x_1^{d-2k} \right) + \sigma(x) \\ &= \left( \frac{p(e_1)}{2} x_1^d - th_d(x; e_1) \right) + \sigma(x), \end{aligned}$$

<sup>5</sup>Recall that a symmetric matrix  $A = (a_{ij})$  is diagonally dominant if  $a_{ii} \geq \sum_{j \neq i} |a_{ij}|$  for all  $i$ . It follows from Gershgorin's circle theorem [21] that diagonally dominant matrices are positive semidefinite.

where  $\sigma(x) = x_1^d \tilde{\sigma}(z/x_1) \in \widehat{\Sigma}_{n,d}$ , since homogenization preserves the sos property. Note that

$$\begin{aligned} T(p; e_1) &\geq \frac{n}{2p(e_1)} \|p\|_\infty^2 + \|p\|_1 \\ &= \frac{n}{2\tilde{p}(0)} \|\tilde{p}\|_\infty^2 + \|\tilde{p}\|_1 \\ &= \tilde{T}(\tilde{p}; 0), \end{aligned}$$

where the second line follows from the fact that the coefficients of  $p$  remain the same after dehomogenization. Therefore, in view of the fact that  $h_d(x; e_1)$  is sos, the claim of Theorem 2 follows in the case where  $x^* = e_1$ .

To establish the claim for a general point  $x^* \in \mathbb{S}^{n-1}$ , observe that for any such point there exists a matrix  $U \in \mathbb{R}^{n \times n}$  with  $U^T U = I$ , such that  $U e_1 = x^*$ . Define  $p_U(x) := p(Ux)$ . Then  $p_U$  is a polynomial with  $p_U(e_1) = p(x^*) > 0$ . Applying the statement of the theorem to  $p_U$ , whenever  $t \geq T(p_U; e_1)$ , we have

$$p_U(x) = \left( \frac{p_U(e_1)}{2} (x^T e_1)^d - t \sum_{k=1}^{d/2} (\|x\|_2^2 - (x^T e_1)^2)^k (x^T e_1)^{d-2k} \right) + \hat{\sigma}(x),$$

where  $\hat{\sigma}(x) \in \widehat{\Sigma}_{n,d}$ . Substituting  $x$  with  $U^T x$  gives

$$\begin{aligned} p(x) &= p_U(U^T x) \\ &= \left( \frac{p(x^*)}{2} (x^T U e_1)^d - t \sum_{k=1}^{d/2} (\|U^T x\|_2^2 - (x^T U e_1)^2)^k (x^T U e_1)^{d-2k} \right) + \hat{\sigma}(U^T x) \\ &= \left( \frac{p(x^*)}{2} (x^T x^*)^d - t \sum_{k=1}^{d/2} (\|x\|_2^2 - (x^T x^*)^2)^k (x^T x^*)^{d-2k} \right) + \hat{\sigma}(U^T x) \\ &= \left( \frac{p(x^*)}{2} (x^T x^*)^d - t h_d(x; x^*) \right) + \hat{\sigma}(U^T x), \end{aligned}$$

where in the second line we use the fact that the 2-norm is unitarily invariant. Since  $\hat{\sigma}(x) \in \widehat{\Sigma}_{n,d}$  and the sos property is preserved under a linear change of variables, we have  $\hat{\sigma}(U^T x) \in \widehat{\Sigma}_{n,d}$ . Finally, note that  $T(p_U; e_1) = T(p; x^*)$ , since the set of orthogonal matrices is invariant under multiplication by any of its elements.  $\blacksquare$

*Remark 2.* A closer examination of the proof of Theorem 2 reveals that the polynomial  $\sigma(x)$  in identity (8) is not only sos, but in fact “scaled-diagonally-dominant-sum-of-squares” (sdsos) with respect to a specific basis [1]. Indeed, from equation (14) in the proof of Lemma 3, we see that the polynomial  $\hat{\sigma}(x)$  is sdsos. Consequently,  $\sigma(x) = \hat{\sigma}(U^T x)$  is sdsos with respect to a rotated monomial basis induced by the linear transformation  $U^T$ . This implies that the search for an algebraic proof of polynomial nonnegativity in the form of identity (8) can be achieved by a second-order cone program (SOCP). This observation

connects to Algorithm 1 in Section 3.3, which can potentially be implemented as an SOCP-based method.

We now present another local Positivstellensatz, analogous to Theorem 2, which provides an algebraic certificate of nonnegativity for a polynomial  $p$  in a neighborhood of a point  $x^*$  where  $p$  is positive. The neighborhood is defined by a single polynomial of even degree  $d' \in \{2, \dots, d-2\}$ . This formulation offers additional flexibility in the choice of the sos polynomial multiplied by the polynomial defining the neighborhood.

**Theorem 4.** *Let  $d$  be an even integer,  $d'$  be an even integer in  $\{2, \dots, d-2\}$ , and  $p \in H_{n,d}$ . If  $p(x^*) > 0$  for some  $x^* \in \mathbb{S}^{n-1}$ , then for  $t > 0$  sufficiently large,  $p(x)$  can be written as*

$$p(x) = \left( \frac{p(x^*)}{2} (x^T x^*)^{d'} - t h_{d'}(x; x^*) \right) \left( (x^T x^*)^{d-d'} + h_{d-d'}(x; x^*) \right) + \sigma(x),$$

where  $\sigma(x) \in \widehat{\Sigma}_{n,d}$ , and  $h_{d'}(x; x^*)$ ,  $h_{d-d'}(x; x^*)$  are as in (9). In particular, the above statement holds for any

$$t \geq T'(p; x^*) := T(p; x^*) + \frac{p(x^*)}{2},$$

where  $T(p; x^*)$  is defined as in (10).

The main step in the proof of Theorem 4 is the following lemma, which builds on Lemma 3.

**Lemma 5.** *Consider a (not necessarily homogeneous) polynomial  $p$  in  $n$  variables and degree at most  $d$ , where  $d$  is even. Let  $d'$  be an even integer in  $\{2, \dots, d-2\}$ . If  $p(x^*) > 0$  for some  $x^* \in \mathbb{R}^n$ , then for  $\hat{t} > 0$  sufficiently large,  $p(x)$  can be written as*

$$p(x) = \left( \frac{p(x^*)}{2} - \hat{t} g_{d'}(x - x^*) \right) (1 + g_{d-d'}(x - x^*)) + \hat{\sigma}(x),$$

where  $\hat{\sigma}(x) \in \Sigma_{n,d}$ . In particular, the above statement holds for any

$$\hat{t} \geq \tilde{T}(p; x^*) + \frac{p(x^*)}{2}.$$

Here,  $g_{d'}(x)$ ,  $g_{d-d'}(x)$ , and  $\tilde{T}(p; x^*)$  are defined as in Lemma 3.

*Proof.* Without loss of generality, we assume  $x^* = 0$ ; otherwise, applying the statement of the lemma to the polynomial  $p(x + x^*)$  and recalling that the existence of an sos decomposition is invariant under an affine change of variables, yields the desired result. By Lemma 3, for any  $t \geq \tilde{T}(p; 0)$ ,  $p(x)$  admits the representation

$$p(x) = \frac{p(0)}{2} - t g_d(x) + \sigma(x),$$

where  $\sigma(x) \in \Sigma_{n,d}$ . Let  $A = p(0)/2$  and  $\hat{t} = t + A$ . Then

$$\begin{aligned}
p(x) &= A - tg_d(x) + \sigma(x) \\
&= (A - \hat{t}g_{d'}(x))(1 + g_{d-d'}(x)) - Ag_{d-d'}(x) + \hat{t}g_{d'}(x)(1 + g_{d-d'}(x)) - tg_d(x) + \sigma(x) \\
&= (A - \hat{t}g_{d'}(x))(1 + g_{d-d'}(x)) + A(g_d(x) - g_{d-d'}(x)) \\
&\quad + \hat{t}(g_{d'}(x)(1 + g_{d-d'}(x)) - g_d(x)) + \sigma(x) \\
&= (A - \hat{t}g_{d'}(x))(1 + g_{d-d'}(x)) + \hat{\sigma}(x),
\end{aligned}$$

where

$$\hat{\sigma}(x) = \sigma(x) + A(g_d(x) - g_{d-d'}(x)) + \hat{t}(g_{d'}(x)(1 + g_{d-d'}(x)) - g_d(x)).$$

It remains to show that  $\hat{\sigma}(x) \in \Sigma_{n,d}$ . Observe that

$$g_d(x) - g_{d-d'}(x) = \sum_{k=(d-d')/2+1}^{d/2} \|x\|_2^{2k},$$

and

$$\begin{aligned}
g_{d'}(x)(1 + g_{d-d'}(x)) - g_d(x) &= g_{d'}(x) + \left(g_{d'-2}(x) + \|x\|_2^{d'}\right)g_{d-d'}(x) - g_d(x) \\
&= g_{d'-2}(x)g_{d-d'}(x).
\end{aligned}$$

Hence, both expressions are sos, which implies  $\hat{\sigma}(x)$  is sos.  $\blacksquare$

The proof of Theorem 4 follows the same approach as the proof of Theorem 2 and is therefore only outlined. We first consider the special case where  $x^* = e_1$ . We dehomogenize  $p$  by defining  $\tilde{p}(z) := p(1, z)$ , apply Lemma 5 to  $\tilde{p}$ , and homogenize back to show that whenever  $t \geq T'(p; e_1)$ ,  $p(x)$  can be written as

$$p(x) = \left(\frac{p(e_1)}{2}(x^T e_1)^{d'} - th_{d'}(x; e_1)\right) \left((x^T e_1)^{d-d'} + h_{d-d'}(x; e_1)\right) + \sigma(x),$$

where  $\sigma(x) \in \widehat{\Sigma}_{n,d}$ . We then establish the claim for a general point  $x^* \in \mathbb{S}^{n-1}$  via an orthogonal change of variables, as in the proof of Theorem 2.

### 3.2 From Local Positivstellensätze to a Global Positivstellensatz

Building on the results established in the previous subsection, we now present the proof of Theorem 1.

*Proof of Theorem 1.* For any integer  $m \geq 2$ , we construct a net  $\mathcal{N}_m = \{x^1, \dots, x^{r_m}\} \subset \mathbb{S}^{n-1}$  such that for any  $y \in \mathbb{S}^{n-1}$ , there exists  $x \in \mathcal{N}_m$  with  $\langle x, y \rangle \geq m/\sqrt{m^2 + 1}$ . Fix an even degree  $d' \in \{2, \dots, d\}$  and let the initial algebraic disjunction be  $\mathcal{D}_{d'}^1 = \{\{0\}\}$ . For  $m \geq 2$ , we construct the subsequent algebraic disjunctions as

$$\mathcal{D}_{d'}^m = \{\{(x^T x^i)^{d'} - t_m h_{d'}(x; x^i)\}, \quad i = 1, \dots, r_m\}, \quad (15)$$

where  $t_m = \left(\sum_{k=1}^{d'/2} m^{-2k}\right)^{-1}$  and  $h_d(x; x^i)$  is defined as in (9). For any  $x \in \mathbb{S}^{n-1}$  and any  $x^i \in \mathcal{N}_m$ , let  $\theta \in [0, \pi]$  be the unique angle satisfying  $\langle x, x^i \rangle = \cos \theta$ . Then, it follows that

$$\begin{aligned} (x^T x^i)^{d'} - t_m h_{d'}(x; x^i) &= (\cos \theta)^{d'} - t_m \sum_{k=1}^{d'/2} (\sin \theta)^{2k} (\cos \theta)^{d'-2k} \\ &= (\cos \theta)^{d'} \left( 1 - t_m \sum_{k=1}^{d'/2} (\tan \theta)^{2k} \right). \end{aligned}$$

Therefore, the inequality  $(x^T x^i)^{d'} - t_m h_{d'}(x; x^i) \geq 0$  holds if and only if

$$\sum_{k=1}^{d'/2} (\tan \theta)^{2k} \leq t_m^{-1} \quad \Leftrightarrow \quad |\tan \theta| \leq m^{-1} \quad \Leftrightarrow \quad \langle x, x^i \rangle \geq \frac{m}{\sqrt{m^2 + 1}}.$$

Hence  $\mathcal{D}_d^m$  forms a valid algebraic disjunction of  $\mathbb{R}^n$  by the construction of the net.

Finally, note that  $t_m \rightarrow \infty$  as  $m \rightarrow \infty$ . Therefore, keeping in mind that  $\mathbb{S}^{n-1}$  is a compact set, for any positive definite form  $p \in H_{n,d}$ , letting  $p_{\min} := \min_{x \in \mathbb{S}^{n-1}} p(x)$ , there exists  $m$  such that  $t_m$  exceeds  $2T(p; x^*)/p_{\min}$  (resp.  $2T'(p; x^*)/p_{\min}$ ) for any  $x^* \in \mathbb{S}^{n-1}$ , where  $T(p; x^*)$  (resp.  $T'(p; x^*)$ ) is defined in Theorem 2 (resp. Theorem 4). This guarantees that  $p \in \Sigma_{n,d}^d(\mathcal{D}_d^m)$  for any fixed even degree  $d' \in \{2, \dots, d\}$ .  $\blacksquare$

The preceding proof in fact yields an explicit bound on the number of subregions required to certify nonnegativity of a polynomial with a disjunctive sos certificate. As an illustration, we present such a bound for the case  $d' = d$  in Theorem 6 below. Given the NP-hardness of certifying polynomial nonnegativity, any such bound is expected to grow exponentially with the dimension.

**Theorem 6.** *Consider a positive definite form  $p \in H_{n,d}$ . Let  $p_{\min} := \min_{x \in \mathbb{S}^{n-1}} p(x)$ , and define*

$$T^*(p) := \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \left( \frac{n}{2p_{\min}} \|p(Ux)\|_{\infty}^2 + \|p(Ux)\|_1 \right), \quad M(p) := \left\lceil \sqrt{\frac{T^*(p)d}{p_{\min}}} \right\rceil. \quad (16)$$

Let  $\mathcal{D}_d^m$  be the family of algebraic disjunctions constructed in (15), where each subset of  $\mathcal{D}_d^m$  contains exactly one degree- $d$  form. Then  $p \in \Sigma_{n,d}^d(\mathcal{D}_d^m)$  for all  $m \geq M(p)$ . In particular,  $(1 + 2M(p))^n$  subregions, each defined by a single degree- $d$  form, suffice to establish a disjunctive sos proof of nonnegativity of  $p$ .

*Proof.* Recall the definition of the algebraic disjunction

$$\mathcal{D}_d^m = \{ \{ (x^T x^i)^d - t_m h_d(x; x^i) \}, \quad i = 1, \dots, r_m \},$$

where  $t_m = \left(\sum_{k=1}^{d/2} m^{-2k}\right)^{-1}$ . For each  $i = 1, \dots, r_m$ , by Theorem 2, there exists a polynomial  $\sigma_i \in \widehat{\Sigma}_{n,d}$  such that

$$p(x) = \left( \frac{p(x^i)}{2} (x^T x^i)^d - T(p; x^i) h_d(x; x^i) \right) + \sigma_i(x),$$

where  $T(p; \cdot)$  is defined as in (10). Rewriting gives

$$\begin{aligned} p(x) &= \frac{p(x^i)}{2} (x^T x^i)^d + \sigma_i(x) - T(p; x^i) h_d(x; x^i) \\ &= \frac{p(x^i)}{2} \left( (x^T x^i)^d - t_m h_d(x; x^i) \right) + \left( \sigma_i(x) + \left( \frac{p(x^i)}{2} t_m - T(p; x^i) \right) h_d(x; x^i) \right). \end{aligned}$$

Note that  $t_m \geq 2m^2/d$ . Therefore, whenever  $m \geq M(p)$ ,

$$\frac{p(x^i)}{2} t_m \geq \frac{p_{\min}}{2} \cdot \frac{2M^2(p)}{d} \geq T^*(p) \geq T(p; x^i), \quad i = 1, \dots, r_m.$$

Since  $h_d(x; x^*) \in \widehat{\Sigma}_{n,d}$  for any  $x^* \in \mathbb{R}^n$ , it follows that

$$\sigma_i(x) + \left( \frac{p(x^i)}{2} t_m - T(p; x^i) \right) h_d(x; x^i) \in \widehat{\Sigma}_{n,d}.$$

Thus,  $p \in \Sigma_{n,d}^d(\mathcal{D}_d^m)$  whenever  $m \geq M(p)$ . We now bound the size of  $\mathcal{N}_m$  when  $m = M(p)$ . Recall that  $\mathcal{N}_m = \{x^1, \dots, x^{r_m}\} \subset \mathbb{S}^{n-1}$  is a net such that for any  $y \in \mathbb{S}^{n-1}$ , there exists  $x \in \mathcal{N}_m$  with  $\langle x, y \rangle \geq m/\sqrt{m^2 + 1}$ . This implies that

$$\begin{aligned} \|x - y\|_2^2 &= 2 - 2\langle x, y \rangle \leq 2 \left( 1 - \frac{m}{\sqrt{m^2 + 1}} \right) \\ &= \frac{2(\sqrt{m^2 + 1} - m)}{\sqrt{m^2 + 1}} = \frac{2}{(\sqrt{m^2 + 1} + m)\sqrt{m^2 + 1}} \\ &\leq \frac{1}{m^2}. \end{aligned}$$

By [52, Corollary 4.2.11], we conclude that  $r_m \leq (1 + 2m)^n = (1 + 2M(p))^n$ . ■

### 3.3 An Algorithm Based on Spherical-Cap Subregions

Building on Theorem 1, we propose an algorithm to approximate the minimum value  $p_{\min}$  of a form  $p \in H_{n,d}$  over the unit sphere  $\mathbb{S}^{n-1}$ . This quantity is also equal to the optimal value of the following optimization problem:

$$\begin{aligned} p_{\min} &= \max_{\gamma \in \mathbb{R}} \gamma \\ \text{s.t.} \quad & p(x) - \gamma \|x\|_2^d \geq 0, \quad \forall x \in \mathbb{R}^n. \end{aligned} \tag{17}$$

For any even degree  $d' \in \{2, \dots, d\}$ , we can present a complete SDP-based hierarchy of lower bounds on  $p_{\min}$ . At level  $m$  of this hierarchy, we construct a net  $\mathcal{N}_m$  of the unit sphere<sup>6</sup> and the corresponding algebraic disjunction  $\mathcal{D}_{d'}^m$  defined in (15). Using the notation introduced in Definition 2, we compute a lower bound  $p^{(m)}$  by solving the following SDP:

$$\begin{aligned} p^{(m)} &:= \max_{\gamma \in \mathbb{R}} \gamma \\ \text{s.t.} \quad & p(x) - \gamma \|x\|_2^d \in \Sigma_{n,d}^d(\mathcal{D}_{d'}^m). \end{aligned} \tag{18}$$

This SDP can be decomposed into  $r_m$  independent SDPs, one for each subset in  $\mathcal{D}_{d'}^m$ , and solved in parallel; see Remark 3 for details. The next theorem utilizes Theorem 6 to establish a convergence rate for the sequence of lower bounds  $p^{(m)}$  produced by the SDP hierarchy in (18) (in the case  $d' = d$  as a representative example).

**Theorem 7.** *Let  $p \in H_{n,d}$  and  $p_{\min} := \min_{x \in \mathbb{S}^{n-1}} p(x)$ . Let  $\mathcal{D}_d^m$  be the family of algebraic disjunctions constructed in (15), where each subset of  $\mathcal{D}_d^m$  contains exactly one degree- $d$  form. Then, we have*

$$p_{\min} - p^{(m)} = O(1/m).$$

*Proof.* Let  $q(x) := \|x\|_2^d$  and fix  $\epsilon \in (0, p_{\min})$ . By Theorem 6, we have  $p - \epsilon q \in \Sigma_{n,d}^d(\mathcal{D}_d^m)$  for all  $m \geq M(p - \epsilon q)$ , where  $M(\cdot)$  is defined in (16). Thus,  $p^{(m)} \geq \epsilon$  whenever  $m \geq M(p - \epsilon q)$ . Let  $c_\infty := \|q\|_\infty$ ,  $c_1 := \|q\|_1$ , which are constants depending only on  $n$  and  $d$ . We now bound  $T^*(p - \epsilon q)$  from above using the definition in (16):

$$\begin{aligned} T^*(p - \epsilon q) &= \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \left( \frac{n}{2(p_{\min} - \epsilon)} \|(p - \epsilon q)(Ux)\|_\infty^2 + \|(p - \epsilon q)(Ux)\|_1 \right) \\ &\leq \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \left( \frac{n}{p_{\min} - \epsilon} (\|p(Ux)\|_\infty^2 + \epsilon^2 \|q(Ux)\|_\infty^2) + \|p(Ux)\|_1 + \epsilon \|q(Ux)\|_1 \right) \\ &= \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \left( \frac{n}{p_{\min} - \epsilon} (\|p(Ux)\|_\infty^2 + c_\infty^2 \epsilon^2) + \|p(Ux)\|_1 + c_1 \epsilon \right) \\ &\leq \frac{2p_{\min}}{p_{\min} - \epsilon} T^*(p) + \frac{c_\infty^2 n \epsilon^2}{p_{\min} - \epsilon} + c_1 \epsilon \\ &= \frac{2p_{\min} T^*(p) + c_\infty^2 n \epsilon^2 + c_1 \epsilon (p_{\min} - \epsilon)}{p_{\min} - \epsilon} \\ &\leq \frac{2p_{\min} T^*(p) + (c_\infty^2 n + c_1/4) p_{\min}^2}{p_{\min} - \epsilon}. \end{aligned}$$

Here, in the first inequality we used the triangle inequality for the polynomial norms. Therefore, for any  $\epsilon \in (0, p_{\min})$ ,

$$M(p - \epsilon q) \leq \sqrt{\frac{2T^*(p - \epsilon q)d}{p_{\min} - \epsilon}} + 1 \leq \frac{\sqrt{2(2p_{\min} T^*(p) + (c_\infty^2 n + c_1/4) p_{\min}^2)d}}{p_{\min} - \epsilon} + 1.$$

<sup>6</sup>There are different techniques for constructing such a net; see, e.g., [32], [12].

---

**Algorithm 1** Spherical-Cap-Based Algorithm for Minimizing a Form
 

---

**Input:** A polynomial  $p \in H_{n,d}$ , an even number  $d' \in \{2, \dots, d\}$ , a tolerance  $\epsilon_{\text{tol}} > 0$ .

**Output:** A lower bound  $L$  on the minimum value of  $p$  over  $\mathbb{S}^{n-1}$  and a point  $\bar{x} \in \mathbb{S}^{n-1}$  such that  $p(\bar{x}) - L \leq \epsilon_{\text{tol}}$ .

**Initialization:** Let  $L = -\infty, U = \infty, m = 1$ .

**while**  $U - L > \epsilon_{\text{tol}}$  **do**

Generate a net  $\mathcal{N}_m = \{x^1, \dots, x^{r_m}\} \subset \mathbb{S}^{n-1}$  such that for all  $y \in \mathbb{S}^{n-1}$ , there exists  $x \in \mathcal{N}_m$  with  $\langle x, y \rangle \geq m/\sqrt{m^2 + 1}$ .

$$t_m \leftarrow \left( \sum_{k=1}^{d'/2} m^{-2k} \right)^{-1}.$$

$$L^m \leftarrow \max_{\gamma, s_i, \hat{s}_i} \gamma$$

$$\text{s.t. } p(x) - \gamma \|x\|_2^d = s_i(x) + ((x^T x^i)^{d'} - t_m h_{d'}(x; x^i)) \hat{s}_i(x), \quad i = 1, \dots, r_m$$

$$s_i \in \widehat{\Sigma}_{n,d}, \hat{s}_i \in \widehat{\Sigma}_{n,d-d'}, \quad i = 1, \dots, r_m.$$

$$L \leftarrow \max\{L, L^m\}.$$

$$U^m \leftarrow \min_{1 \leq i \leq r_m} \{p(x^i)\}, j \leftarrow \operatorname{argmin}_{1 \leq i \leq r_m} \{p(x^i)\}$$

▷ Break ties arbitrarily.

**if**  $U^m < U$  **then**  $U \leftarrow U^m, \bar{x} \leftarrow x^j$ .

$$m \leftarrow m + 1.$$

**return**  $L, \bar{x}$

---

Finally, since the above bound holds for every  $\epsilon \in (0, p_{\min})$ , when  $m$  is sufficiently large, we may choose  $\epsilon \in (0, p_{\min})$  such that

$$m = \frac{\sqrt{2(2p_{\min}T^*(p) + (c_\infty^2 n + c_1/4)p_{\min}^2)d}}{p_{\min} - \epsilon} + 1.$$

With this choice, it follows that

$$p_{\min} - p^{(m)} \leq p_{\min} - \epsilon = \frac{\sqrt{2(2p_{\min}T^*(p) + (c_\infty^2 n + c_1/4)p_{\min}^2)d}}{m - 1} = O(1/m).$$

■

Since the net  $\mathcal{N}_m$  already appears in the formulation of our SDP, it is natural to use the points in  $\mathcal{N}_m$  to obtain upper bounds on  $p_{\min}$  via direct point evaluation. As the level  $m$  increases, the resolution of the net improves and its covering radius tends to zero, yielding increasingly tight upper bounds on  $p_{\min}$ . In particular, by adapting the argument in the comments following [53, Proposition 4], one can construct a net such that the error between the resulting upper bound at level  $m$  and  $p_{\min}$  is of order  $O(1/m^2)$ . Combined with Theorem 7, this implies that the gap between the lower and upper bounds on  $p_{\min}$  converges to zero as  $m$  increases, with an overall rate of  $O(1/m)$ . Therefore, for any prescribed tolerance  $\epsilon_{\text{tol}} > 0$  to approximate  $p_{\min}$ , Algorithm 1 is guaranteed to terminate within  $O(1/\epsilon_{\text{tol}})$  iterations.

*Remark 3.* The computation of  $L^m$  in Algorithm 1 can be decomposed into  $r_m$  independent SDPs that can be solved in parallel. Indeed, for each  $k \in \{1, \dots, r_m\}$ , one can compute

$$\begin{aligned} \gamma_k^* := & \max_{\gamma_k, s_k, \hat{s}_k} \gamma_k \\ \text{s.t.} & p(x) - \gamma_k \|x\|_2^d = s_k(x) + ((x^T x^k)^{d'} - t_m h_{d'}(x; x^k)) \hat{s}_k(x) \\ & s_k \in \widehat{\Sigma}_{n,d}, \hat{s}_k \in \widehat{\Sigma}_{n,d-d'}, \end{aligned}$$

and set  $L^m \leftarrow \min_{1 \leq k \leq r_m} \gamma_k^*$ .

In practice, it is often more efficient to refine the subregions of the unit sphere adaptively instead of increasing the resolution of the net uniformly. Such an adaptive algorithm can be implemented within a spatial branch-and-bound framework, as discussed in Section 6.

## 4 An Optimization-Free Disjunctive Positivstellensatz

In this section, we introduce a new disjunctive approach to certify polynomial nonnegativity. Unlike our previous approach, which relies on sos polynomials and solutions to SDPs, the new approach only requires checking nonnegativity of the coefficients of some polynomials we can explicitly write down. In this framework, *checking* polynomial nonnegativity does not require any optimization solver at all, and *imposing* a nonnegativity constraint on a polynomial leads to a linear program. Our main result in this section is an optimization-free disjunctive Positivstellensatz (Theorem 8), analogous to Theorem 1.

### 4.1 Simplicial Disjunctions and Optimization-Free Certificates

In this section, we work with algebraic disjunctions where the corresponding subregions of the space are polyhedral cones. Recall that a polyhedral cone in  $\mathbb{R}^n$  can be represented as a conic combination of a finite number of vectors  $v_1, \dots, v_\ell \in \mathbb{R}^n$ :

$$\text{cone}\{v_1, \dots, v_\ell\} = \{x \in \mathbb{R}^n \mid x = \alpha_1 v_1 + \dots + \alpha_\ell v_\ell \text{ for some } \alpha_1, \dots, \alpha_\ell \geq 0\}.$$

We say that a polyhedral cone is simplicial if it is generated by linearly independent vectors.

**Definition 3.** Let  $r$  be a positive integer. We say that a collection of  $n \times n$  invertible matrices  $\mathcal{V} = \{V_k\}_{k=1, \dots, r}$  forms a simplicial disjunction of  $\mathbb{R}^n$  if

$$\mathbb{R}^n = \bigcup_{k=1}^r \Omega_k, \quad \text{where} \quad \Omega_k = \text{cone}\{V_k\} := \text{cone}\{v_{k,1}, \dots, v_{k,n}\},$$

and  $v_{k,j}$  is the  $j$ -th column of  $V_k$ .

Note that any vector  $x \in \text{cone}\{V_k\}$  can be written as  $x = V_k z$  for some elementwise nonnegative vector  $z \in \mathbb{R}^n$ . Since  $V_k$  is invertible, this implies that

$$\text{cone}\{V_k\} = \{x \in \mathbb{R}^n \mid V_k^{-1} x \geq 0\}.$$

Therefore, the simplicial disjunction  $\mathcal{V}$  can be viewed as a special case of an algebraic disjunction introduced in Definition 1 with  $n_k = n$  for  $k = 1, \dots, r$ , and polynomials  $q_{k,j}(x) = c_{k,j}^T x$  for  $k = 1, \dots, r, j = 1, \dots, n$ , where  $c_{k,j} \in \mathbb{R}^n$  are the columns of  $V_k^{-T}$ . We now explain how simplicial disjunctions can help certify polynomial nonnegativity. Because of the homogenization argument given in Section 3.1, we restrict attention to forms.

**Definition 4.** A polynomial  $p \in H_{n,d}$  is said to have nonnegative coefficients with respect to a simplicial disjunction  $\mathcal{V} = \{V_k\}_{k=1,\dots,r}$  if  $p(V_k x)$  has nonnegative coefficients for all  $k = 1, \dots, r$ . We denote the set of such polynomials by  $NC_{n,d}(\mathcal{V})$ .

Analogous to the cone  $\Sigma_{n,d}^{\bar{d}}(\mathcal{D})$  in Definition 2, the set  $NC_{n,d}(\mathcal{V})$  is a convex cone. Moreover, any polynomial  $p \in NC_{n,d}(\mathcal{V})$  has an explicit disjunctive proof of nonnegativity. Indeed, for  $k = 1, \dots, r$ , let  $\sigma_k(x) := p(V_k x)$ , which, by definition, has nonnegative coefficients. We can therefore write

$$p(x) = \sigma_k(V_k^{-1}x), \quad k = 1, \dots, r.$$

Since  $\sigma_k$  has nonnegative coefficients, we have the implication

$$V_k^{-1}x \geq 0 \Rightarrow \sigma_k(V_k^{-1}x) \geq 0,$$

for all  $k = 1, \dots, r$ . Equivalently,  $x \in \text{cone}\{V_k\} \Rightarrow p(x) \geq 0$ . Since  $\bigcup_{k=1}^r \text{cone}\{V_k\} = \mathbb{R}^n$  by the definition of a simplicial disjunction,  $p$  is nonnegative over  $\mathbb{R}^n$ .

We now establish the following optimization-free disjunctive Positivstellensatz, which runs parallel to the disjunctive Positivstellensatz in Theorem 1. It shows that every positive definite form has a disjunctive proof of nonnegativity of the above type.

**Theorem 8** (Optimization-Free Disjunctive Positivstellensatz). *For any fixed dimension  $n$  and degree  $d$ , there exists an explicit family of simplicial disjunctions  $\mathcal{V}^m$ , indexed by  $m$ , such that for any positive definite form  $p \in H_{n,d}$ , we have  $p \in NC_{n,d}(\mathcal{V}^m)$  for some  $m$ .*

## 4.2 A Local Optimization-Free Positivstellensatz

To prove Theorem 8, we first establish a local optimization-free Positivstellensatz in Theorem 9, which runs parallel to Theorem 2. Recall that  $\|\cdot\|_\infty$  denotes the largest coefficient in absolute value of a polynomial in the standard monomial basis. For an  $n$ -dimensional multi-index  $\alpha$  with  $|\alpha| = d$ , we define

$$\binom{d}{\alpha} := \frac{d!}{\alpha_1! \dots \alpha_n!}.$$

We denote the  $n$ -dimensional vector of all ones by  $\mathbf{1}$ .

**Theorem 9.** *For a polynomial  $p \in H_{n,d}$  with  $d$  even, if  $p(x^*) > 0$  for some  $x^* \in \mathbb{S}^{n-1}$ , then there exists  $\delta > 0$  such that for any matrix  $V \in \mathbb{R}^{n \times n}$  with  $\|V - x^* \mathbf{1}^T\|_F \leq \delta$ , the polynomial  $p(Vx)$  has nonnegative coefficients. In particular, the above statement holds for any*

$$\delta \leq \delta(p; x^*) := \sup_{\epsilon > 1/p(x^*)} \min \left\{ \frac{(p(x^*)\epsilon)^{1/d} - 1}{(p(x^*)\epsilon)^{1/d} + 1}, \frac{1}{(1 + \eta(p)\epsilon)n} \right\},$$

where

$$\eta(p) := \sup_{U \in \mathbb{R}^{n \times n}: U^T U = I_n} \|p(Ux)\|_\infty.$$

*Remark 4.* The polynomial  $p(Vx)$  having nonnegative coefficients implies that  $p(x) \geq 0$  for all  $x \in \text{cone}\{V\}$ . In particular, when  $V$  is invertible,  $\text{cone}\{V\}$  is simplicial and full-dimensional, which provides a proof of nonnegativity of  $p$  over a full-dimensional neighborhood around  $x^*$ .

*Proof.* We first consider the special case where  $x^* = e_1$ . Let  $p(x) = \sum_{\alpha: |\alpha|=d} p_\alpha x^\alpha$ . For each multi-index  $\beta$  with  $|\beta| = d$ , we estimate the coefficient of the monomial  $x^\beta$  in the expansion  $p(Vx) = \sum_{\alpha: |\alpha|=d} p_\alpha (Vx)^\alpha$ . Let  $c(\alpha, \beta)$  denote the coefficient of  $x^\beta$  arising from the expansion of  $(Vx)^\alpha$ . Then, the coefficient of  $x^\beta$  in  $p(Vx)$  is given by

$$\sum_{\alpha: |\alpha|=d} p_\alpha c(\alpha, \beta).$$

In the case where  $\alpha \neq de_1$  (i.e.,  $x^\alpha \neq x_1^d$ ), we bound  $c(\alpha, \beta)$  by first noting that

$$(Vx)^\alpha = \prod_{k=1}^n (V_{k1}x_1 + \dots + V_{kn}x_n)^{\alpha_k}.$$

Since  $\|V - e_1 \mathbf{1}^T\|_F \leq \delta$ , it follows that  $V_{1j} \in [1 - \delta, 1 + \delta]$  for  $j = 1, \dots, n$ , and  $V_{kj} \in [-\delta, \delta]$  for  $k = 2, \dots, n$  and  $j = 1, \dots, n$ . Hence,  $\max_{1 \leq j \leq n} |V_{1j}| \leq 1 + \delta$  and  $\max_{1 \leq j \leq n} |V_{kj}| \leq \delta$  for  $k = 2, \dots, n$ . Therefore, recalling the multinomial theorem, we have

$$\begin{aligned} |c(\alpha, \beta)| &\leq \sum_{\substack{\gamma_1 + \dots + \gamma_n = \beta \\ |\gamma_i| = \alpha_i, i=1, \dots, n}} \prod_{k=1}^n \left( \max_{1 \leq j \leq n} |V_{kj}| \right)^{\alpha_k} \binom{\alpha_k}{\gamma_k} \\ &\leq (1 + \delta)^{\alpha_1} \delta^{d - \alpha_1} \sum_{\substack{\gamma_1 + \dots + \gamma_n = \beta \\ |\gamma_i| = \alpha_i, i=1, \dots, n}} \prod_{k=1}^n \binom{\alpha_k}{\gamma_k} \\ &= (1 + \delta)^{\alpha_1} \delta^{d - \alpha_1} \binom{d}{\beta}, \end{aligned} \tag{19}$$

The last equality follows from matching the coefficients of  $x^\beta$  on both sides of the identity  $(x_1 + \dots + x_n)^d = \prod_{k=1}^n (x_1 + \dots + x_n)^{\alpha_k}$  and employing the multinomial theorem.

We now turn to the case  $\alpha = de_1$ , where we have  $(Vx)^\alpha = (V_{11}x_1 + \dots + V_{1n}x_n)^d$ . For any  $\delta \leq \delta(p; e_1)$ , we have  $\delta < 1/n \leq 1$ . Thus,  $\min_{1 \leq j \leq n} V_{1j} \geq 1 - \delta \geq 0$ , and

$$c(\alpha, \beta) \geq \left( \min_{1 \leq j \leq n} V_{1j} \right)^d \binom{d}{\beta} \geq (1 - \delta)^d \binom{d}{\beta}. \tag{20}$$

Combining the bounds (19) for  $\alpha \neq de_1$  and (20) for  $\alpha = de_1$ , the coefficient of  $x^\beta$  in  $p(Vx)$  can be bounded from below as

$$\begin{aligned}
\sum_{\alpha:|\alpha|=d} p_\alpha c(\alpha, \beta) &= p_{de_1} c(de_1, \beta) + \sum_{\alpha:|\alpha|=d, \alpha \neq de_1} p_\alpha c(\alpha, \beta) \\
&= p(e_1) c(de_1, \beta) + \sum_{l=0}^{d-1} \sum_{\alpha:|\alpha|=d, \alpha_1=l} p_\alpha c(\alpha, \beta) \\
&\geq \binom{d}{\beta} \left( p(e_1)(1-\delta)^d - \|p\|_\infty \sum_{l=0}^{d-1} \sum_{\alpha:|\alpha|=d, \alpha_1=l} (1+\delta)^{\alpha_1} \delta^{d-\alpha_1} \right) \\
&\geq \binom{d}{\beta} \left( p(e_1)(1-\delta)^d - \|p\|_\infty \sum_{l=0}^{d-1} n^{d-l} (1+\delta)^l \delta^{d-l} \right) \\
&= \binom{d}{\beta} \left( p(e_1)(1-\delta)^d - \|p\|_\infty \left( \frac{1+\delta}{n\delta} - 1 \right)^{-1} ((1+\delta)^d - n^d \delta^d) \right) \\
&\geq \binom{d}{\beta} \left( p(e_1)(1-\delta)^d - \|p\|_\infty \left( \frac{1}{n\delta} - 1 \right)^{-1} (1+\delta)^d \right)
\end{aligned}$$

where the second inequality follows from the fact that the number of distinct  $n$ -dimensional multi-indices  $\alpha$  satisfying  $|\alpha| = d$  and  $\alpha_1 = l$  is given by  $\binom{d-l+n-2}{n-2} \leq n^{d-l}$ .

For any  $\epsilon > 1/p(e_1) > 0$ , since

$$\delta \leq \frac{1}{(1 + \eta(p)\epsilon)n} \leq \frac{1}{(1 + \|p\|_\infty \epsilon)n},$$

we have

$$\sum_{\alpha:|\alpha|=d} p_\alpha c(\alpha, \beta) \geq \binom{d}{\beta} (p(e_1)(1-\delta)^d - \epsilon^{-1}(1+\delta)^d).$$

This expression is nonnegative because

$$\delta \leq \frac{(p(e_1)\epsilon)^{1/d} - 1}{(p(e_1)\epsilon)^{1/d} + 1}.$$

Therefore, the coefficient of  $x^\beta$  in  $p(Vx)$  remains nonnegative for every multi-index  $\beta$  with  $|\beta| = d$  whenever  $\delta \leq \delta(p; e_1)$ . Therefore, the theorem follows for the case where  $x^* = e_1$ .

For a general point  $x^* \in \mathbb{S}^{n-1}$ , there exists an orthogonal matrix  $U \in \mathbb{R}^{n \times n}$ , such that  $Ue_1 = x^*$ . Let  $p_U(x) = p(Ux)$ . Then  $p_U$  is a polynomial with  $p_U(e_1) = p(x^*) > 0$ . Applying the statement of the theorem to  $p_U$ , we have that for every matrix  $\tilde{V} \in \mathbb{R}^{n \times n}$  with  $\|\tilde{V} - e_1 \mathbf{1}^T\|_F \leq \delta(p_U; e_1)$ , the polynomial  $p_U(\tilde{V}x) = p(U\tilde{V}x)$  has nonnegative coefficients. Hence, for every matrix  $V \in \mathbb{R}^{n \times n}$  with  $\|V - x^* \mathbf{1}^T\|_F \leq \delta(p_U; e_1)$ , the matrix  $U^T V$  satisfies

$$\|U^T V - e_1 \mathbf{1}^T\|_F = \|U(U^T V - e_1 \mathbf{1}^T)\|_F = \|V - x^* \mathbf{1}^T\|_F \leq \delta(p_U; e_1).$$

This implies that  $p(Vx) = p(U(U^T V)x)$  has nonnegative coefficients. Finally, observe that  $\eta(p)$  is unitarily invariant, i.e.,  $\eta(p) = \eta(p_U)$  for any orthogonal matrix  $U$ . Consequently,

$$\begin{aligned} \delta(p_U; e_1) &= \sup_{\epsilon > 1/p_U(e_1)} \min \left\{ \frac{(p_U(e_1)\epsilon)^{1/d} - 1}{(p_U(e_1)\epsilon)^{1/d} + 1}, \frac{1}{(1 + \eta(p_U)\epsilon)n} \right\} \\ &= \sup_{\epsilon > 1/p(x^*)} \min \left\{ \frac{(p(x^*)\epsilon)^{1/d} - 1}{(p(x^*)\epsilon)^{1/d} + 1}, \frac{1}{(1 + \eta(p)\epsilon)n} \right\} \\ &= \delta(p; x^*), \end{aligned}$$

which completes the proof. ■

### 4.3 Proof of the Optimization-Free Disjunctive Positivstellensatz

With the theorem from the previous subsection in place, we are now ready to prove Theorem 8.

*Proof of Theorem 8.* We construct  $\mathcal{V}^1$  using  $2^n$  matrices, each of the form  $[\pm e_1 \dots \pm e_n]$ , so that the conic hulls of their columns correspond to the  $2^n$  orthants of  $\mathbb{R}^n$ . We describe how to further refine the partitioning of the space by focusing on the nonnegative orthant (the other orthants can be handled analogously by flipping signs). The columns of the matrix  $[e_1 \dots e_n]$  form the vertices of the unit simplex. For  $m \geq 2$ , we divide each edge of the unit simplex into  $m$  equal segments of length  $\sqrt{2}/m$  and partition the simplex into  $m^{n-1}$  smaller simplices. Taking the conic hulls of these simplices partitions the nonnegative orthant into  $m^{n-1}$  simplicial cones, whose extreme rays associate with the columns of  $m^{n-1}$  invertible  $n \times n$  matrices. Specifically, if  $w_1, \dots, w_n$  are the vertices of one of the smaller simplices, we include the invertible matrix

$$V = \left[ \frac{w_1}{\|w_1\|_2} \quad \dots \quad \frac{w_n}{\|w_n\|_2} \right]$$

as a member of the simplicial disjunction  $\mathcal{V}^m$ . Repeating this procedure for each of the  $2^n$  orthants, we obtain a total of  $m^{n-1}2^n$  matrices that together define  $\mathcal{V}^m$ .

We now estimate how fast the maximum pairwise distance between the columns of each matrix  $V$  constructed as above tends to zero as  $m \rightarrow \infty$ . Note that at level  $m$ , the vertices of the small simplices have the form  $\alpha/m$ , where  $\alpha = (\alpha_1, \dots, \alpha_n)$  is an  $n$ -tuple of nonnegative integers such that  $\sum_{i=1}^n \alpha_i = m$ . Let  $\alpha/m$  and  $\beta/m$  be two vertices of the same simplex. Since  $\|\alpha/m - \beta/m\|_2 = \sqrt{2}/m$ , we have  $\|\alpha - \beta\|_2 = \sqrt{2}$ . Note that their corresponding columns in  $V$  are  $\alpha/\|\alpha\|_2$  and  $\beta/\|\beta\|_2$ . The distance between these two columns can be

upper bounded as

$$\begin{aligned}
\left\| \frac{\alpha}{\|\alpha\|_2} - \frac{\beta}{\|\beta\|_2} \right\|_2 &= \left\| \frac{\alpha - \beta}{\|\alpha\|_2} + \beta \left( \frac{1}{\|\alpha\|_2} - \frac{1}{\|\beta\|_2} \right) \right\|_2 \\
&= \left\| \frac{\alpha - \beta}{\|\alpha\|_2} + \frac{\beta}{\|\beta\|_2} \cdot \frac{\|\beta\|_2 - \|\alpha\|_2}{\|\alpha\|_2} \right\|_2 \\
&\leq \frac{\|\alpha - \beta\|_2}{\|\alpha\|_2} + \frac{|\|\beta\|_2 - \|\alpha\|_2|}{\|\alpha\|_2} \\
&\leq \frac{2\|\alpha - \beta\|_2}{\|\alpha\|_2} \leq \frac{2\sqrt{2n}}{m},
\end{aligned} \tag{21}$$

where we have used the triangle inequality and the fact that

$$\|\alpha\|_2^2 \geq \frac{1}{n} \left( \sum_{i=1}^n \alpha_i \right)^2 = \frac{m^2}{n}.$$

Let  $v \in \mathbb{S}^{n-1}$  denote any column of a matrix  $V$  constructed as above. Then, it follows from (21) that  $\|V - v\mathbf{1}^T\|_F \leq 2\sqrt{2n}/m$ . Let  $p_{\min} = \min_{x \in \mathbb{S}^{n-1}} p(x)$ , and let  $\eta(p)$  and  $\delta(p; v)$  be as in Theorem 9. Define

$$\tilde{M}(p) := \left[ 2\sqrt{2n} \left( \sup_{\epsilon > 1/p_{\min}} \min \left\{ \frac{(p_{\min}\epsilon)^{1/d} - 1}{(p_{\min}\epsilon)^{1/d} + 1}, \frac{1}{(1 + \eta(p)\epsilon)n} \right\} \right)^{-1} \right].$$

It follows that  $\|V - v\mathbf{1}^T\|_F \leq 2\sqrt{2n}/\tilde{M}(p) \leq \delta(p; v)$  whenever  $m \geq \tilde{M}(p)$ . By Theorem 9, the polynomial  $p(Vx)$  has nonnegative coefficients. Since the same argument applies to every matrix  $V$  constructed as above, we conclude that  $p \in \text{NC}_{n,d}(\mathcal{V}^m)$  whenever  $m \geq \tilde{M}(p)$ . ■

Analogous to Algorithm 1, which is based on Theorem 1, the optimization-free disjunctive Positivstellensatz can be used to construct an optimization-free algorithm for certifying polynomial nonnegativity. In practice, instead of uniformly refining each simplicial cone as in the proof, it is more efficient to adaptively refine these cones within a spatial branch-and-bound framework. Even with such a modification, we expect the optimization-free approach to require a significantly larger number of regions to obtain a disjunctive certificate of nonnegativity compared to the disjunctive sum of squares approach based on semidefinite programming (i.e., the approach of Section 3). Studying these tradeoffs computationally is a direction for future research.

## 5 Extensions to Constrained Optimization Problems

In this section, we extend the disjunctive sum of squares method from certifying global nonnegativity of polynomials to certifying nonnegativity over a closed basic semialgebraic set, i.e., a subset of the Euclidean space defined by finitely many polynomial inequalities.

We present two approaches. The first, discussed in Section 5.1, builds on a polynomial-time reduction from [2] that converts constrained polynomial optimization problems (POPs) into unconstrained ones, to which our disjunctive sum of squares method then applies. The second, given in Section 5.2, extends Definition 1 so that the subregions defined by the algebraic disjunction cover the feasible set of interest rather than all of  $\mathbb{R}^n$ . We illustrate this approach for the problem of testing matrix copositivity, which reduces to minimizing a quadratic form over the nonnegative orthant.

## 5.1 Lower Bounds on the Optimal Values of Polynomial Optimization Problems

Consider a POP of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & p(x) \\ \text{s.t.} \quad & g_j(x) \geq 0, \quad j = 1, \dots, J, \end{aligned} \tag{22}$$

with a compact feasible set. Throughout this section, we let  $p^*$  denote the optimal value of problem (22). We show that our disjunctive Positivstellensätze can be combined with the reduction from [2, Theorem 2.1] to produce a converging hierarchy of lower bounds on  $p^*$ .

We first describe the reduction from [2]. Let  $d$  be the smallest even integer larger or equal to the maximum degrees of the polynomials  $p$  and  $g_1, \dots, g_J$ . Following [2, Theorem 2.1], one can explicitly construct a form  $f_\gamma(z)$  in  $N := n + J + 3$  variables  $z$  and of degree  $2d$  using a given scalar  $\gamma$ , the coefficients of  $p, g_1, \dots, g_J$ , and an upper bound on the radius of a ball containing the feasible set of problem (22). This construction has the property that  $\gamma$  is a strict lower bound on  $p^*$  if and only if  $f_\gamma(z)$  is positive definite.

Then, [2, Theorem 2.4] shows how to use this reduction to solve problem (22). Recall that  $H_{n,d}$  denotes the set of homogeneous polynomials in  $n$  variables and degree  $d$ , and  $\widehat{P}_{n,d}$  denotes the set of nonnegative forms in  $H_{n,d}$ . Consider a sequence of subsets of  $H_{n,d}$ , denoted by  $K_{n,d}^i$  and indexed by  $i$ , that satisfies the following properties:

- (1)  $K_{n,d}^i \subseteq \widehat{P}_{n,d}$  for all  $i$ , and there exists a positive definite form  $s_{n,d} \in K_{n,d}^0$ ,
- (2) if  $p \in H_{n,d}$  is positive definite, then there exists an integer  $i$  such that  $p \in K_{n,d}^i$ ,
- (3)  $K_{n,d}^i \subseteq K_{n,d}^{i+1}$  for all  $i$ ,
- (4) if  $p \in K_{n,d}^i$ , then  $p + \epsilon s_{n,d} \in K_{n,d}^i$  for all  $\epsilon \in [0, 1]$ .

By [2, Theorem 2.4], for any such sequence  $K_{n,d}^i$  satisfying properties (1)–(4), the sequence  $l_i$  defined by

$$\begin{aligned} l_i := \max_{\gamma \in \mathbb{R}} \quad & \gamma \\ \text{s.t.} \quad & f_\gamma(z) - \frac{1}{i} s_{N,2d}(z) \in K_{N,2d}^i, \end{aligned} \tag{23}$$

is non-decreasing, satisfies  $l_i \leq p^*$  for all  $i$ , and  $\lim_{i \rightarrow \infty} l_i = p^*$ . As shown in [2], Artin, Reznick, and Polyá's Positivstellensätze all lead to sequences  $K_{n,d}^i$  satisfying properties (1)–(4). In Theorem 10 below, we show that our disjunctive sum of squares method fits into this framework as well. Moreover, in contrast to the above Positivstellensätze, our method has the advantage that the degrees of the polynomials required to certify membership in  $K_{n,d}^i$  remain fixed (and low) throughout the hierarchy.

We recall that  $\widehat{\Sigma}_{n,d}$  denotes the set of sos forms in  $n$  variables and degree  $d$ .

**Theorem 10.** *Fix a dimension  $n$  and an even degree  $d$ . Let  $\mathcal{D}^m$  be any family of algebraic disjunctions, indexed by  $m$ , that satisfies the guarantees of Theorem 1. Let  $\widehat{K}_{n,d}^0 := \widehat{\Sigma}_{n,d}$ . For each integer  $i \geq 1$ , define*

$$\widehat{K}_{n,d}^i := \bigcup_{m=1}^i \Sigma_{n,d}^d(\mathcal{D}^m). \quad (24)$$

Then, the sequence  $\widehat{K}_{n,d}^i \cap H_{n,d}$  satisfies properties (1)–(4) stated above.

*Proof.* We verify properties (1)–(4).

- (1) When  $i = 0$ ,  $\widehat{K}_{n,d}^0 = \widehat{\Sigma}_{n,d} \subseteq \widehat{P}_{n,d}$ . Let  $i \geq 1$ . Since  $\Sigma_{n,d} \subseteq \Sigma_{n,d}^d(\mathcal{D}^m) \subseteq P_{n,d}$  for all  $m$ , we have  $\widehat{\Sigma}_{n,d} \subseteq \Sigma_{n,d}^d(\mathcal{D}^m) \cap H_{n,d} \subseteq \widehat{P}_{n,d}$  for all  $m$ . Taking unions gives  $\widehat{\Sigma}_{n,d} \subseteq \widehat{K}_{n,d}^i \cap H_{n,d} \subseteq \widehat{P}_{n,d}$  for all  $i \geq 1$ . Moreover, the positive definite form  $\|x\|_2^d \in \widehat{\Sigma}_{n,d} = \widehat{K}_{n,d}^0$ .
- (2) By Theorem 1, for any positive definite form  $p$ , there exists an integer  $\bar{m}$  such that  $p \in \Sigma_{n,d}^d(\mathcal{D}^{\bar{m}})$ , and therefore  $p \in \widehat{K}_{n,d}^{\bar{m}}$ .
- (3) The sequence  $\widehat{K}_{n,d}^i$  is nested by construction.
- (4) When  $i = 0$ , the property holds trivially. When  $i \geq 1$ , suppose  $p \in \widehat{K}_{n,d}^i$ . Then, there exists  $m \in \{1, \dots, i\}$  such that  $p \in \Sigma_{n,d}^d(\mathcal{D}^m)$ . Recall from Theorem 1 that each subset in  $\mathcal{D}^m$  contains exactly one form, i.e., we can write

$$\mathcal{D}^m = \{\{q_j(x)\}, \quad j = 1, \dots, r_m\}.$$

For each  $j \in \{1, \dots, r_m\}$ , there exist sos polynomials  $s_j$  and  $\sigma_j$  such that

$$p(x) = s_j(x) + \sigma_j(x)q_j(x),$$

with  $\deg s_j \leq d$  and  $\deg(\sigma_j q_j) \leq d$ . Consequently,

$$p(x) + \epsilon \|x\|_2^d = (s_j(x) + \epsilon \|x\|_2^d) + \sigma_j(x)q_j(x).$$

Since  $\|x\|_2^d$  is sos, the polynomial  $s_j(x) + \epsilon \|x\|_2^d$  remains sos for all  $\epsilon \in [0, 1]$  and  $j = 1, \dots, r_m$ . Therefore,  $p(x) + \epsilon \|x\|_2^d \in \Sigma_{n,d}^d(\mathcal{D}^m) \subseteq \widehat{K}_{n,d}^i$  for all  $\epsilon \in [0, 1]$ . ■

Since the sequence  $\widehat{K}_{n,d}^i$  satisfies properties (1)–(4), replacing<sup>7</sup>  $K_{N,2d}^i$  with  $\widehat{K}_{N,2d}^i$  in (23) and applying [2, Theorem 2.4] yields a non-decreasing sequence  $l_i$  that satisfies  $l_i \leq p^*$  for all  $i$  and  $\lim_{i \rightarrow \infty} l_i = p^*$ . Furthermore, one can show that the theorem remains valid if the cones  $\Sigma_{n,d}^d(\mathcal{D}^m)$  in (24) are replaced by the cones  $\text{NC}_{n,d}(\mathcal{V}^m)$  introduced in Theorem 8, yielding an optimization-free hierarchy for general POPs with a compact feasible set.

## 5.2 An Application to Copositive Programming

In this section, we give an example of how the disjunctive sum of squares approach can be adapted to certify nonnegativity over a subset of the Euclidean space. We take this subset to be the nonnegative orthant, which arises naturally in copositive programming.

A symmetric matrix  $Q \in \mathbb{R}^{n \times n}$  is said to be *copositive* if  $x^T Q x \geq 0$  for all  $x \geq 0$ . Let  $\mathcal{C}_n$  denote the cone of  $n \times n$  copositive matrices. A copositive program is a linear optimization problem over the intersection of  $\mathcal{C}_n$  with an affine subspace. In recent years, copositive programming has found many applications in both discrete and continuous optimization; see, e.g., [11], [18]. Since checking copositivity of a given matrix is NP-hard [38], tractable inner approximations to  $\mathcal{C}_n$  are of interest in most applications of copositive programming.

A well-known SDP-based sufficient condition for copositivity is the so-called “ $P + N$ ” criterion. For a symmetric matrix  $A \in \mathbb{R}^{n \times n}$ , let us use the standard notation  $A \succeq 0$  (resp.  $A \geq 0$ ) to denote that  $A$  is positive semidefinite (resp. elementwise nonnegative). Observe that if a matrix  $Q \in \mathbb{R}^{n \times n}$  can be written as  $Q = P + N$ , where  $P \succeq 0$  and  $N \geq 0$ , then  $Q$  is copositive. This condition is known to also be necessary for copositivity if  $n \leq 4$  [17]. For  $n \geq 5$ , however, there are copositive matrices that do not admit a  $P + N$  decomposition. A classical example is the Horn matrix [17]:

$$H = \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}. \quad (25)$$

To bridge the gap between the  $P + N$  criterion and copositivity, a complete SDP-based hierarchy of sufficient conditions for copositivity was introduced in [39]. Let  $x.^2 := (x_1^2, \dots, x_n^2)$ . At level  $m$  of this hierarchy, one checks if the polynomial  $(x.^2)^T Q (x.^2) \cdot (x_1^2 + \dots + x_n^2)^m$  is sos, and if so, copositivity is certified. The zeroth level of this hierarchy recovers exactly the  $P + N$  criterion described above [39]. Moreover, this hierarchy is complete: for any strictly copositive matrix  $Q$  (i.e., a matrix  $Q$  satisfying  $x^T Q x > 0$  for all  $x \geq 0$  with  $x \neq 0$ ), there exists a level  $m$  at which copositivity is certified [39], [43]. However, the size of the semidefinite constraint at level  $m$  of this hierarchy grows as  $\binom{n+m+2}{n} \times \binom{n+m+2}{n}$ . While one

---

<sup>7</sup>To compute the optimal value  $l_i$  of problem (23) with  $K_{N,2d}^i$  replaced by  $\widehat{K}_{N,2d}^i$ , one can take the maximum between  $l_{i-1}$  and the optimal value of problem (23) with  $K_{N,2d}^i$  replaced by  $\Sigma_{N,2d}^{2d}(\mathcal{D}^i)$ . As discussed earlier, a membership constraint to  $\Sigma_{N,2d}^{2d}(\mathcal{D}^i)$  amounts to solving a semidefinite program.

can reduce this size using symmetry reduction techniques [20, Section 8.1], the higher levels of this hierarchy often remain too expensive to solve.

Here, we utilize the disjunctive sos approach to propose an alternative SDP-based complete hierarchy for certifying copositivity of an  $n \times n$  matrix, where the size of the semidefinite constraint remains  $n \times n$  throughout the hierarchy. To this end, we extend the notion of a simplicial disjunction in Definition 4 and require the subregions to cover only the nonnegative orthant rather than all of  $\mathbb{R}^n$ .

**Definition 5.** *Let  $r$  be a positive integer. We say that a collection of  $n \times n$  invertible matrices  $\mathcal{V} = \{V_k\}_{k=1, \dots, r}$  forms a simplicial disjunction of a region  $\Omega \subseteq \mathbb{R}^n$  if*

$$\Omega \subseteq \bigcup_{k=1}^r \Omega_k, \quad \text{where} \quad \Omega_k = \text{cone}\{V_k\} := \text{cone}\{v_{k,1}, \dots, v_{k,n}\},$$

and  $v_{k,j}$  is the  $j$ -th column of  $V_k$ .

**Definition 6.** *A symmetric matrix  $Q \in \mathbb{R}^{n \times n}$  is disjunctive  $P+N$  with respect to a simplicial disjunction  $\mathcal{V} = \{V_k\}_{k=1, \dots, r}$  of the nonnegative orthant if there exist matrices  $P_k$  and  $N_k$ , for  $k = 1, \dots, r$ , such that*

$$V_k^T Q V_k = P_k + N_k, \quad P_k \succeq 0, \quad N_k \geq 0, \quad k = 1, \dots, r.$$

We denote the set of such matrices by  $\mathcal{PN}_n(\mathcal{V})$ .

Observe that for any simplicial disjunction  $\mathcal{V}$  of the nonnegative orthant, the set  $\mathcal{PN}_n(\mathcal{V})$  is a semidefinite representable convex cone. Moreover, it is straightforward to see that  $\mathcal{PN}_n(\mathcal{V}) \subseteq \mathcal{C}_n$ . Indeed, if  $Q \in \mathcal{PN}_n(\mathcal{V})$ , the polynomial  $y^T V_k^T Q V_k y$  is nonnegative for all  $y \geq 0$  and  $k = 1, \dots, r$ . This implies that  $x^T Q x$  is nonnegative on each subregion  $\Omega_k = \text{cone}\{V_k\}$  as any point  $x \in \text{cone}\{V_k\}$  can be written as  $x = V_k y$  with  $y \geq 0$ . Since these subregions cover the nonnegative orthant, we have  $x^T Q x \geq 0$  for all  $x \geq 0$ , and hence  $Q$  is copositive.

Let us revisit the Horn matrix  $H$  in (25) in view of Definition 6. Consider the simplicial disjunction  $\mathcal{V} = \{V_1, V_2\}$  of the nonnegative orthant, where

$$V_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 1 \end{bmatrix}.$$

One can readily verify that

$$V_1^T H V_1 = \begin{bmatrix} 1 & -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

$$V_2^T H V_2 = \begin{bmatrix} 1 & -1 & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & -1 \\ -1 & 1 & -1 & 0 & 1 \\ 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{bmatrix}.$$

Let  $e_1, \dots, e_5$  be the standard basis vectors in  $\mathbb{R}^5$ . The first identity proves nonnegativity of  $x^T H x$  over  $\text{cone}\{V_1\} = \text{cone}\{e_1, e_2, e_3, e_4, (e_4 + e_5)/2\}$ . Similarly, the second identity proves nonnegativity of  $x^T H x$  over  $\text{cone}\{V_2\} = \text{cone}\{e_1, e_2, e_3, (e_4 + e_5)/2, e_5\}$ . Together, these two identities certify that the Horn matrix  $H \in \mathcal{PN}_5(\mathcal{V})$  and is hence copositive.

We show in Theorem 11 below that one can a priori fix a family of simplicial disjunctions of the nonnegative orthant such that any strictly copositive matrix is disjunctive  $P + N$  with respect to some simplicial disjunction in the family. In fact, the proof shows that the disjunctive proofs do not necessarily need to involve positive semidefinite matrices; that is, one can set  $P_k = 0$  in Definition 6. This result follows as a simple corollary of what we have proven in the previous section.

In the proof of the following theorem, we use the standard notation  $\Delta^{n-1}$  to denote the unit simplex in  $\mathbb{R}^n$ , i.e.,  $\Delta^{n-1} = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, x \geq 0\}$ , and we let  $\text{conv}\{V\}$  denote the convex hull of the columns of a matrix  $V \in \mathbb{R}^{n \times n}$ .

**Theorem 11** (Disjunctive Positivstellensatz for Matrix Copositivity). *For any fixed dimension  $n$ , there exists an explicit family of simplicial disjunctions  $\mathcal{V}^m$  of the nonnegative orthant, indexed by  $m$ , such that for any strictly copositive matrix  $Q \in \mathbb{R}^{n \times n}$ , we have  $Q \in \mathcal{PN}_n(\mathcal{V}^m)$  for some  $m$ .*

*Proof.* The construction of the family of simplicial disjunctions  $\mathcal{V}^m$  is similar to the construction for Theorem 8. We let  $\mathcal{V}^1 = \{I_n\}$ . For  $m \geq 2$ , we divide each edge of the unit simplex  $\Delta^{n-1}$  into  $m$  equal segments of length  $\sqrt{2}/m$  and partition the simplex into  $m^{n-1}$  smaller simplices. For each smaller simplex, we form a matrix whose columns are the vertices of this simplex and include this matrix as a member of  $\mathcal{V}^m$ .

Let  $q(x) := x^T Q x$ . Since  $q$  is homogeneous and  $Q$  is strictly copositive, we have  $q(x^*/\|x^*\|_2) > 0$  for all  $x^* \in \Delta^{n-1}$ . Letting  $\delta(\cdot; \cdot)$  be defined as in Theorem 9, it follows that  $\delta(q; x^*/\|x^*\|_2) > 0$  for all  $x^* \in \Delta^{n-1}$ . It is straightforward to check that  $\delta_{\min} := \inf_{x^* \in \Delta^{n-1}} \|x^*\|_2 \delta(q; x^*/\|x^*\|_2) > 0$ . For any integer  $m \geq \lceil \sqrt{2n}/\delta_{\min} \rceil$  and any matrix  $V \in \mathcal{V}^m$ ,  $\text{conv}\{V\}$  is a polytope with edge length  $\sqrt{2}/m$ , so for any column  $v \in \Delta^{n-1}$  of  $V$ , we have

$$\|V - v\mathbf{1}^T\|_F \leq \sqrt{n} \cdot \frac{\sqrt{2}}{m} \leq \delta_{\min} \leq \|v\|_2 \delta(q; v/\|v\|_2).$$

Thus,  $\|V/\|v\|_2 - (v/\|v\|_2)\mathbf{1}^T\|_F \leq \delta(q; v/\|v\|_2)$ , and by Theorem 9, the polynomial  $q(Vx/\|v\|_2) = x^T V^T Q V x / \|v\|_2^2$  has nonnegative coefficients, hence  $V^T Q V \geq 0$ . Since the same argument holds for each matrix in  $\mathcal{V}^m$ , we conclude that  $Q \in \mathcal{PN}_n(\mathcal{V}^m)$ .  $\blacksquare$

*Remark 5.* We make two observations regarding Theorem 11 and Definition 6:

- (i) By the construction in the proof of Theorem 11, the first simplicial disjunction  $\mathcal{V}^1$  contains only the identity matrix, and hence checking membership in  $\mathcal{PN}_n(\mathcal{V}^1)$  reduces exactly to the  $P + N$  test. Moreover, the tests for copositivity in the higher levels of the hierarchy all dominate the  $P + N$  test. Indeed, consider an arbitrary integer  $m \geq 1$ . If  $Q = P + N$  with  $P \succeq 0$  and  $N \geq 0$ , then for each matrix  $V$  in  $\mathcal{V}^m$ ,  $V^T Q V = V^T P V + V^T N V$ , where  $V^T P V \succeq 0$ , and  $V^T N V \geq 0$  as the entries of  $V$  are nonnegative by construction. Hence,  $Q \in \mathcal{PN}_n(\mathcal{V}^m)$ .
- (ii) Setting  $P_k = 0$  in Definition 6 reduces the condition  $V_k^T Q V_k = P_k + N_k$  to  $V_k^T Q V_k \geq 0$ , which coincides with the sufficient condition for copositivity in [9, Theorem 1]. However, by doing so, no level of the resulting hierarchy would dominate the  $P + N$  test. See also Section 6.3.2 for some related comparisons in numerical experiments.

## 6 A Spatial Branch-and-Bound Framework

While the Positivstellensätze in previous sections show that one can always find low-degree disjunctive sos proofs of nonnegativity of a polynomial by uniformly dividing the space, in practice it is more efficient to perform this division *adaptively*. In this section, we describe how this can be done by integrating our framework into a spatial branch-and-bound (SBB) algorithm. We present SBB algorithms to certify polynomial nonnegativity and matrix copositivity in Section 6.1 and Section 6.2, respectively, and evaluate their numerical performance in Section 6.3.

### 6.1 A SBB Framework for Polynomial Minimization

In this section, we apply a branch-and-bound algorithm to approximate the minimum value  $p_{\min}$  of a form  $p \in H_{n,d}$  over the unit sphere  $\mathbb{S}^{n-1}$  by adaptively dividing the sphere into subregions using simplicial disjunctions. We use simplicial disjunctions because it is straightforward to partition a simplicial cone into two smaller ones. We consider two choices for the initial simplicial disjunction:

- (1)  $2^{n-1}$  matrices, each of the form<sup>8</sup>  $\begin{bmatrix} \pm e_1 & \dots & \pm e_{n-1} & e_n \end{bmatrix}$ .
- (2)  $n + 1$  matrices, constructed as follows: Consider  $n + 1$  vertices of a regular  $n$ -simplex

$$c_i = \sqrt{1 + n^{-1}} e_i - n^{-3/2} (\sqrt{n + 1} - 1) \mathbf{1}, \quad i = 1, \dots, n, \quad c_{n+1} = -n^{-1/2} \mathbf{1}.$$

Each matrix is formed by taking  $n$  of the  $n + 1$  vectors above as its columns.

---

<sup>8</sup>Since even-degree homogeneous polynomials are symmetric with respect to the origin, it suffices to certify nonnegativity on a hemisphere.

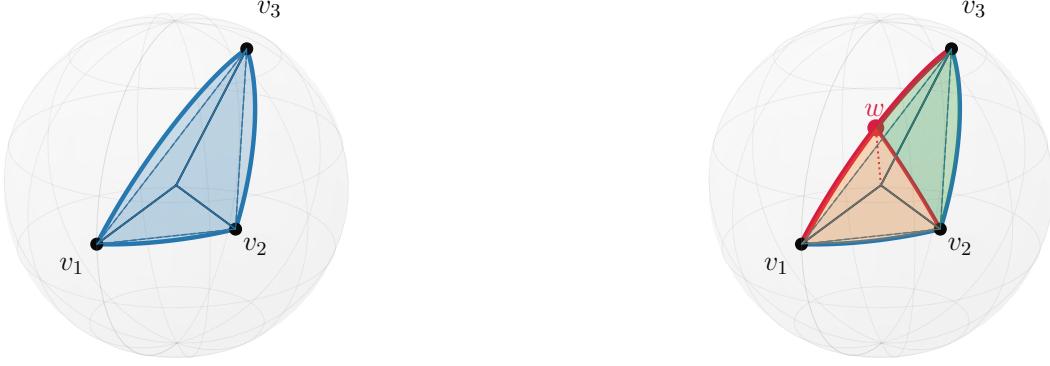


Figure 3: Branching step for  $\text{cone}\{V\} = \text{cone}\{v_1, v_2, v_3\}$ . Left:  $\mathcal{Q} = \text{cone}\{V\} \cap \mathbb{S}^2$  is the spherical triangle bounded by geodesic arcs (solid); planar chords of  $\text{cone}\{V\}$  shown dashed. Right: the longest arc  $(v_1, v_3)$  is bisected at its geodesic midpoint  $w$ , splitting  $\mathcal{Q}$  into  $\mathcal{Q}^+$  and  $\mathcal{Q}^-$ .

**Bounding.** Each subregion of the sphere takes the form  $\mathcal{Q} = \text{cone}\{V\} \cap \mathbb{S}^{n-1}$  for some invertible matrix  $V \in \mathbb{R}^{n \times n}$ . We compute a lower bound on  $p$  over  $\mathcal{Q}$  by solving

$$\begin{aligned} \phi(V) := \max_{\gamma \in \mathbb{R}} \quad & \gamma \\ \text{s.t.} \quad & p(V(x.^2)) - \gamma \|V(x.^2)\|_2^d \in \widehat{\Sigma}_{n,2d}. \end{aligned} \quad (26)$$

Note that any feasible  $\gamma$  in problem (26) is a valid lower bound on  $p$  over  $\mathcal{Q}$ . Indeed, any point  $z \in \text{cone}\{V\}$  can be written as  $z = V(x.^2)$  for some  $x \in \mathbb{R}^n$ , so the constraint in (26) implies  $p(z) - \gamma \|z\|_2^d \geq 0$  for all  $z \in \text{cone}\{V\}$ . In particular,  $p(z) - \gamma \geq 0$  for all  $z \in \text{cone}\{V\} \cap \mathbb{S}^{n-1}$ .

To obtain an upper bound on  $p$  over  $\mathcal{Q}$ , we run  $K$  projected gradient descent steps (where  $K$  is chosen by the user): starting from a column of  $V$ , we iterate

$$y^k = \Pi_{\text{cone}\{V\}}(x^k - \beta \nabla p(x^k)), \quad x^{k+1} = \frac{y^k}{\|y^k\|_2}, \quad k = 0, \dots, K-1, \quad (27)$$

where  $\Pi_{\text{cone}\{V\}}$  denotes the projection onto  $\text{cone}\{V\}$  and  $\beta > 0$  is a chosen stepsize.

**Branching.** At each iteration, we select a subregion

$$\mathcal{Q} = \text{cone}\{V\} \cap \mathbb{S}^{n-1}$$

with the smallest lower bound, where  $V = [v_1 \dots v_n] \in \mathbb{R}^{n \times n}$  is some invertible matrix. Letting  $(i, j) \in \text{argmax}_{1 \leq k, l \leq n} \|v_k - v_l\|_2$ , we set  $w = (v_i + v_j) / \|v_i + v_j\|_2$  and form two further subregions

$$\mathcal{Q}^+ = \text{cone}\{V^+\} \cap \mathbb{S}^{n-1}, \quad \mathcal{Q}^- = \text{cone}\{V^-\} \cap \mathbb{S}^{n-1},$$

where  $V^+ = [v_1 \dots v_{i-1} v_{i+1} \dots v_n w]$  and  $V^- = [v_1 \dots v_{j-1} v_{j+1} \dots v_n w]$ . Figure 3 illustrates an example of this construction when  $n = 3$ .

The overall branch-and-bound scheme is presented in Algorithm 2. One can utilize Theorem 9 together with convergence results for branch-and-bound methods (see, e.g. [33, Theorem 5.26] or [26, Chapter IV]) to show that for any prescribed tolerance  $\epsilon_{\text{tol}} > 0$ , Algorithm 2 is guaranteed to terminate.

---

**Algorithm 2** Spatial Branch-and-Bound Algorithm for Minimizing a Form

---

**Input:** A polynomial  $p \in H_{n,d}$ , a tolerance  $\epsilon_{\text{tol}} > 0$ .

**Output:** A lower bound  $L$  on the minimum value of  $p$  over  $\mathbb{S}^{n-1}$  and a point  $\bar{x} \in \mathbb{S}^{n-1}$  such that  $p(\bar{x}) - L \leq \epsilon_{\text{tol}}(1 + |L| + |p(\bar{x})|)$ .

**Initialization:** Generate an initial simplicial disjunction  $\{V_i\}_{i=1,\dots,r_0}$ .

$\mathcal{T} \leftarrow \{V_i \mid i = 1, \dots, r_0\}$ . ▷ Create a branch-and-bound tree  
 $l_i \leftarrow \phi(V_i), \quad i = 1, \dots, r_0$ . ▷ Local lower bound computed by solving (26)  
 $u_i \leftarrow \min_{1 \leq j \leq n} p(V_i e_j), \quad i = 1, \dots, r_0$ . ▷ Local upper bound  
 $L \leftarrow \min_{1 \leq i \leq r_0} \{l_i\}, \quad U \leftarrow \min_{1 \leq i \leq r_0} \{u_i\}$ . ▷ Global lower and upper bounds

**while**  $U - L > \epsilon_{\text{tol}}(1 + |L| + |U|)$  **do**

$V \leftarrow \operatorname{argmin}_{\hat{V} \in \mathcal{T}} \phi(\hat{V}), \quad \mathcal{T} \leftarrow \mathcal{T} \setminus \{V\}$ .

$v_1, \dots, v_n \leftarrow$  columns of  $V$ .

$(i, j) \leftarrow \operatorname{argmax}_{1 \leq k, l \leq n} \|v_k - v_l\|_2, \quad w \leftarrow (v_i + v_j) / \|v_i + v_j\|_2$ .

$V^+ \leftarrow [v_1 \dots v_{i-1} v_{i+1} \dots v_n w], \quad V^- \leftarrow [v_1 \dots v_{j-1} v_{j+1} \dots v_n w]$ . ▷ Branch

$\mathcal{T} \leftarrow \mathcal{T} \cup \{V^+, V^-\}$ .

$L \leftarrow \min_{\hat{V} \in \mathcal{T}} \phi(\hat{V})$ . ▷ Update global lower bound

**for**  $\hat{V} \in \{V^+, V^-\}$  **do** ▷ Update global upper bound

$x^K \leftarrow K$ -th iterate of (27) with  $x^0 = w$ .

$u \leftarrow p(x^K)$ .

**if**  $u < U$  **then**  $U \leftarrow u, \quad \bar{x} \leftarrow x^K$ .

**return**  $L, \bar{x}$

---

## 6.2 A SBB Framework for Matrix Copositivity

In this section, we focus on the problem of testing copositivity of a matrix  $Q \in \mathbb{R}^{n \times n}$ , which reduces to computing the minimum value  $p_{\min}$  of a quadratic form  $p(x) = x^T Q x$  over the unit simplex  $\Delta^{n-1}$ . We develop a branch-and-bound algorithm analogous to Algorithm 2, tailored to the unit simplex. This algorithm can be used more generally to solve copositive programs. We initialize the algorithm with a single subregion, the unit simplex  $\Delta^{n-1} = \operatorname{conv}\{I_n\}$ , corresponding to the simplicial disjunction  $\{I_n\}$ .

**Bounding.** Each subregion of the unit simplex takes the form  $\mathcal{Q} = \operatorname{conv}\{V\}$  for some invertible matrix  $V = [v_1 \dots v_n] \in \mathbb{R}^{n \times n}$ . Note that the minimum of  $p$  over  $\mathcal{Q}$  equals the optimal value of

$$\begin{aligned} \max_{t \in \mathbb{R}} \quad & t \\ \text{s.t.} \quad & V^T(Q - tJ_n)V \in \mathcal{C}_n, \end{aligned}$$

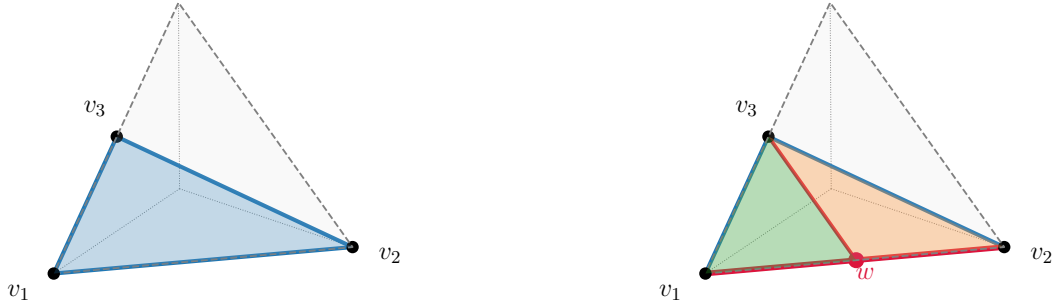


Figure 4: Branching step for  $\text{conv}\{V\} = \text{conv}\{v_1, v_2, v_3\}$ . Left:  $\mathcal{Q} = \text{conv}\{V\}$  is the blue triangle; the unit simplex  $\Delta^2$  shown dashed. Right: the longest edge  $(v_1, v_2)$  is bisected at its midpoint  $w = (v_1 + v_2)/2$ , splitting  $\mathcal{Q}$  into  $\mathcal{Q}^+$  and  $\mathcal{Q}^-$ .

where  $J_n$  is the  $n \times n$  all-one matrix and  $\mathcal{C}_n$  is the cone of  $n \times n$  copositive matrices. Using the inner approximations of  $\mathcal{C}_n$  from Section 5.2, we compute a lower bound on  $p$  over  $\mathcal{Q}$  by solving the following semidefinite program:

$$\begin{aligned} \psi(V) := & \max_{t \in \mathbb{R}, N \in \mathbb{R}^{n \times n}} t \\ \text{s.t.} & \quad V^T(Q - tJ_n)V - N \succeq 0, \\ & \quad N \succeq 0. \end{aligned} \tag{28}$$

To obtain an upper bound on  $p$  over  $\mathcal{Q}$ , we run  $K$  projected gradient descent steps starting from a column of  $V$ :

$$x^{k+1} = \Pi_{\text{conv}\{V\}}(x^k - 2\beta Qx^k), \quad k = 0, \dots, K-1. \tag{29}$$

Here,  $\Pi_{\text{conv}\{V\}}$  denotes the projection onto  $\text{conv}\{V\}$  and  $\beta > 0$  is the stepsize.

**Branching.** At each iteration, we select a subregion with the smallest lower bound and split it by bisecting its longest edge. More specifically, if

$$\mathcal{Q} = \text{conv}\{V\}$$

denotes the selected subregion, where  $V = [v_1 \dots v_n] \in \mathbb{R}^{n \times n}$  is some invertible matrix, we let  $(i, j) \in \text{argmax}_{1 \leq k, l \leq n} \|v_k - v_l\|_2$ , and set  $w = (v_i + v_j)/2$ , forming two further subregions

$$\mathcal{Q}^+ = \text{conv}\{V^+\}, \quad \mathcal{Q}^- = \text{conv}\{V^-\},$$

where  $V^+ = [v_1 \dots v_{i-1} v_{i+1} \dots v_n w]$  and  $V^- = [v_1 \dots v_{j-1} v_{j+1} \dots v_n w]$ . Figure 4 illustrates an example of this construction when  $n = 3$ .

The overall branch-and-bound scheme is presented in Algorithm 3. By the same argument as in Section 6.1, one can show that for any prescribed tolerance  $\epsilon_{\text{tol}} > 0$ , Algorithm 3 terminates. If the algorithm returns a nonnegative lower bound  $L$ , then  $\mathcal{Q}$  is certified to be

copositive. If it returns a point  $\bar{x} \in \Delta^{n-1}$  with  $\bar{x}^T Q \bar{x} < 0$ , then  $Q$  is not copositive. In the remaining case, copositivity is inconclusive with the prescribed tolerance.

---

**Algorithm 3** Spatial Branch-and-Bound Algorithm for Testing Matrix Copositivity

---

**Input:** A symmetric matrix  $Q \in \mathbb{R}^{n \times n}$ , a tolerance  $\epsilon_{\text{tol}} > 0$ .

**Output:** A lower bound  $L$  on the minimum value of  $x^T Q x$  over  $\Delta^{n-1}$  and a point  $\bar{x} \in \Delta^{n-1}$  such that  $\bar{x}^T Q \bar{x} - L \leq \epsilon_{\text{tol}}(1 + |L| + |\bar{x}^T Q \bar{x}|)$ .

**Initialization:**  $\mathcal{T} \leftarrow \{I_n\}$ . ▷ Create a branch-and-bound tree  
 $L \leftarrow \psi(I_n)$ ,  $U \leftarrow \min_{1 \leq i \leq n} e_i^T Q e_i$ . ▷ Initialize global lower and upper bounds  
**while**  $U - L > \epsilon_{\text{tol}}(1 + |L| + |U|)$  **do**  
 $V \leftarrow \operatorname{argmin}_{\hat{V} \in \mathcal{T}} \psi(\hat{V})$ ,  $\mathcal{T} \leftarrow \mathcal{T} \setminus \{V\}$ .  
 $v_1, \dots, v_n \leftarrow$  columns of  $V$ .  
 $(i, j) \leftarrow \operatorname{argmax}_{1 \leq k, l \leq n} \|v_k - v_l\|_2$ ,  $w \leftarrow (v_i + v_j)/2$ .  
 $V^+ \leftarrow [v_1 \dots v_{i-1} v_{i+1} \dots v_n w]$ ,  $V^- \leftarrow [v_1 \dots v_{j-1} v_{j+1} \dots v_n w]$ . ▷ Branch  
 $\mathcal{T} \leftarrow \mathcal{T} \cup \{V^+, V^-\}$ .  
 $L \leftarrow \min_{\hat{V} \in \mathcal{T}} \psi(\hat{V})$ . ▷ Update global lower bound  
**for**  $\hat{V} \in \{V^+, V^-\}$  **do** ▷ Update global upper bound  
 $x^K \leftarrow$   $K$ -th iterate of (29) with  $x^0 = w$ .  
 $u \leftarrow (x^K)^T Q x^K$ .  
**if**  $u < U$  **then**  $U \leftarrow u$ ,  $\bar{x} \leftarrow x^K$ .  
**return**  $L, \bar{x}$

---

## 6.3 Numerical Experiments

In this subsection, we present some numerical experiments focusing on certifying nonnegativity of non-sos polynomials, and solving copositive programs with applications to nonconvex quadratic programming and the maximum clique problem in combinatorial optimization. The code and data used for the numerical experiments are available in our GitHub repository.<sup>9</sup>

### 6.3.1 Polynomial Minimization

We apply Algorithm 2 to minimize several classical polynomials for which the standard sos test (see, e.g., [41]) fails to return the minimum value. The examples, drawn from [44], [16], [1], are listed in the following table.

We record the number of subregions generated by Algorithm 2 with  $\epsilon_{\text{tol}} = 10^{-4}$ , using a single projected gradient descent step in (27) for computing an upper bound at each node of the branch-and-bound tree. We also experiment with the two types of initialization described in Section 6.1. Table 1 shows that with both initializations, the algorithm terminates with a small number of subregions in most cases.

---

<sup>9</sup><https://github.com/YH7422/DisjunctiveSOS>

Name	$n$	$d$	Expression
Motzkin	3	6	$M_h(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_3^2 + x_3^6$
Robinson-1	3	6	$R_1(x_1, x_2, x_3) = x_1^6 + x_2^6 + x_3^6 - (x_1^4 x_2^2 + x_1^2 x_2^4 + x_1^4 x_3^2 + x_1^2 x_3^4 + x_2^4 x_3^2 + x_2^2 x_3^4) + 3x_1^2 x_2^2 x_3^2$
Robinson-2	4	4	$R_2(x_1, \dots, x_4) = x_1^2(x_1 - x_4)^2 + x_2^2(x_2 - x_4)^2 + x_3^2(x_3 - x_4)^2 + 2x_1 x_2 x_3(x_1 + x_2 + x_3 - 2x_4)$
Choi-Lam-1	4	4	$CL_1(x_1, \dots, x_4) = x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 + x_4^4 - 4x_4 x_1 x_2 x_3$
Choi-Lam-2	3	6	$CL_2(x_1, x_2, x_3) = x_1^4 x_2^2 + x_2^4 x_3^2 + x_3^4 x_1^2 - 3x_1^2 x_2^2 x_3^2$
Lax	5	4	$L(x_1, \dots, x_5) = \sum_{i=1}^5 \prod_{j \neq i} (x_i - x_j)$
Schmüdgen	3	6	$SC(x_1, x_2, x_3) = 200(x_1^3 - 4x_1 x_3^2)^2 + 200(x_2^3 - 4x_2 x_3^2)^2 + (x_2^2 - x_1^2)x_1(x_1 + 2x_3)(x_1^2 - 2x_1 x_3 + 2x_2^2 - 8x_3^2)$
PARTITION	6	4	$P(x_1, \dots, x_6) = \left(\sum_{i=1}^5 x_i\right)^2 x_6^2 + \sum_{i=1}^5 (x_i^2 - x_6^2)^2$
Delzell	4	8	$D(x_1, \dots, x_4) = x_1^4 x_2^2 x_4^2 + x_2^4 x_3^2 x_4^2 + x_1^2 x_3^4 x_4^2 - 3x_1^2 x_2^2 x_3^2 x_4^2 + x_3^8$
Stengle- $k$	3	$4k + 2$	$S_k(x_1, x_2, x_3) = x_1^{2k+1} x_3^{2k+1} + (x_2^2 x_3^{2k-1} - x_1^{2k+1} - x_1 x_3^{2k})^2$

Table 1: The number of subregions produced by Algorithm 2 with two types of initialization described in Section 6.1.

Name	# subregions		Name	# subregions	
	Init 1	Init 2		Init 1	Init 2
Motzkin	4	7	PARTITION	69	161
Robinson-1	4	8	Delzell	8	5
Robinson-2	8	19	Stengle-1	4	10
Choi-Lam-1	5	15	Stengle-2	4	4
Choi-Lam-2	4	8	Stengle-3	4	4
Lax	98	149	Stengle-4	4	4
Schmüdgen	4	5	Stengle-5	4	4

*Remark:* The number of subregions can be lowered, sometimes significantly, by a better implementation of the branch-and-bound algorithm. See <https://github.com/YH7422/DisjunctiveSOS> for heuristic improvements to our implementation made by Claude Code, which result, e.g., in 64 (Init 1) and 49 (Init 2) regions for the Lax polynomial, and 32 (Init 1) and 35 (Init 2) regions for the PARTITION polynomial.

Figures 5 and 6 show the lower and upper bounds on the minimum values of the Robinson-2 polynomial and the PARTITION polynomial versus the number of subregions produced by Algorithm 2.

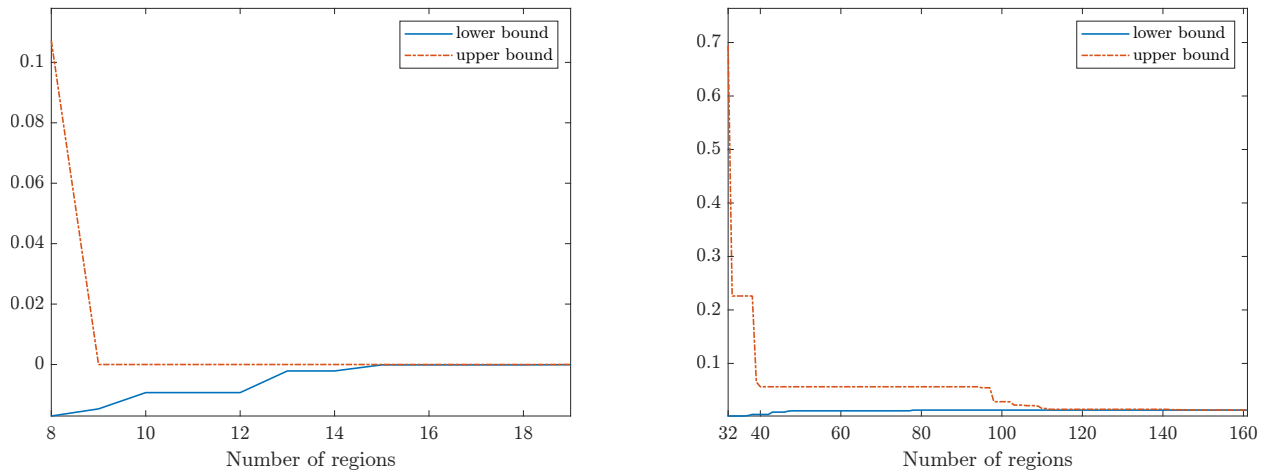


Figure 5: Lower and upper bounds on the minimum value of a given polynomial versus the number of subregions produced by Algorithm 2 with the first type of initialization. Left: Robinson-2 polynomial. Right: PARTITION polynomial.

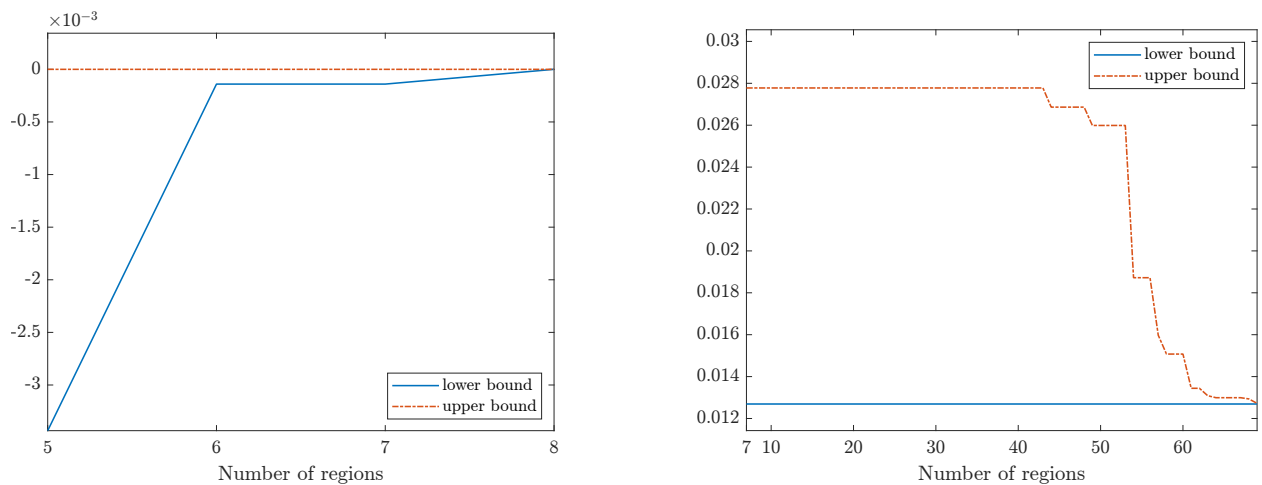


Figure 6: Lower and upper bounds on the minimum value of a given polynomial versus the number of subregions produced by Algorithm 2 with the second type of initialization. Left: Robinson-2 polynomial. Right: PARTITION polynomial.

### 6.3.2 Applications to Copositive Programming

We now present some numerical experiments with Algorithm 3, which can solve copositive programs or the closely related problem of minimizing a quadratic form over the unit simplex.

Table 2: The number of subregions produced by Algorithm 3 to solve nonconvex QPs.

Instance	$Q_1$	$Q_2$	$Q_3$	$Q_4$
Adaptive Linear Approximation [10]	6	158	44	27
Algorithm 3	2	42	5	17

**Solving nonconvex quadratic programs (QPs).** We apply Algorithm 3 to four nonconvex QP instances of the form  $\min_{x \in \Delta^{n-1}} x^T Q x$  taken from [10]. The four matrices defining the objective functions of these QPs are as follows:

$$\begin{aligned}
 Q_1 &= \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}, & Q_2 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \\
Q_3 &= \begin{bmatrix} -14 & -15 & -16 & 0 & 0 \\ -15 & -14 & -12.5 & -22.5 & -15 \\ -16 & -12.5 & -10 & -26.5 & -16 \\ 0 & -22.5 & -26.5 & 0 & 0 \\ 0 & -15 & -16 & 0 & -14 \end{bmatrix}, & Q_4 &= \begin{bmatrix} 0.9044 & 0.1054 & 0.5140 & 0.3322 & 0 \\ 0.1054 & 0.8715 & 0.7385 & 0.5866 & 0.9751 \\ 0.5140 & 0.7385 & 0.6936 & 0.5368 & 0.8086 \\ 0.3322 & 0.5866 & 0.5368 & 0.5633 & 0.7478 \\ 0 & 0.9751 & 0.8086 & 0.7478 & 1.2932 \end{bmatrix}.
 \end{aligned}$$

We record the number of subregions generated by Algorithm 3 with  $\epsilon_{\text{tol}} = 10^{-6}$ . We perform 5 projected gradient descent steps in (29) to obtain upper bounds at each node of the branch-and-bound tree. Table 2 compares our subregion counts against those reported in [10] using an LP-based adaptive approximation algorithm.

Figure 7 shows the lower and upper bounds on the optimal value of the four QPs versus the number of subregions.

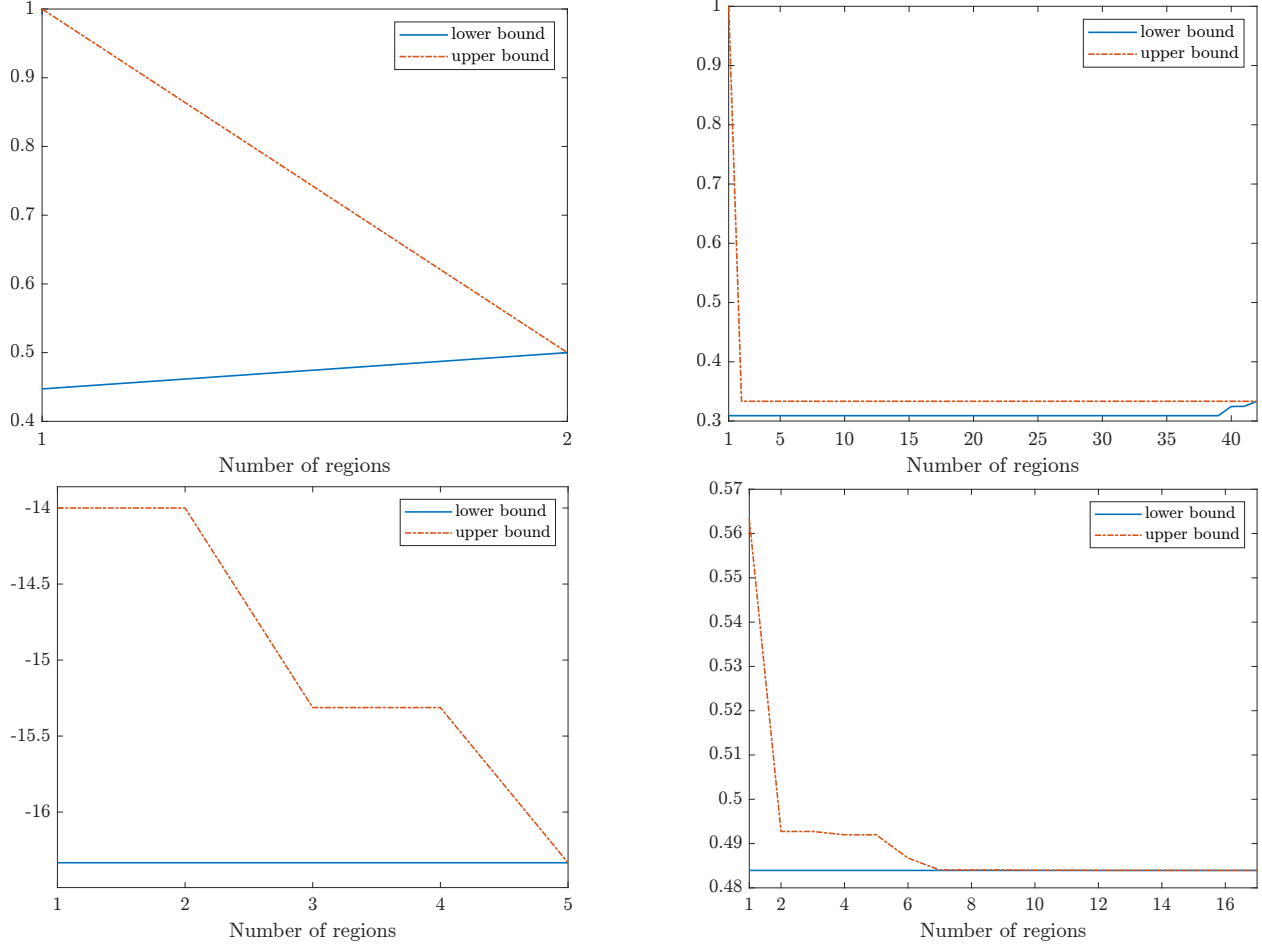


Figure 7: Lower and upper bounds on the optimal value of the QP  $\min_{x \in \Delta^{n-1}} x^T Q x$  produced by Algorithm 3 versus the number of subregions. Top left:  $Q = Q_1$ . Top right:  $Q = Q_2$ . Bottom left:  $Q = Q_3$ . Bottom right:  $Q = Q_4$ .

**Computing the clique number of a graph.** A clique in a graph is a set of pairwise adjacent vertices. The clique number of a graph  $G$ , denoted by  $\omega(G)$ , is the size of a maximum clique in  $G$ . By a theorem of Motzkin and Straus [36], we have

$$\frac{1}{\omega(G)} = \min_{x \in \Delta^{n-1}} x^T (I_n + \bar{A}) x,$$

where  $\bar{A}$  is the adjacency matrix of the complement graph  $\bar{G}$ ; i.e., the graph obtained by swapping the edges and non-edges of  $G$ . Equivalently, the clique number can be written as the optimal value of the copositive program

$$\begin{aligned} \omega(G) = \min_{k \in \mathbb{R}} k \\ \text{s.t. } k(I_n + \bar{A}) - J_n \in \mathcal{C}_n. \end{aligned} \tag{30}$$

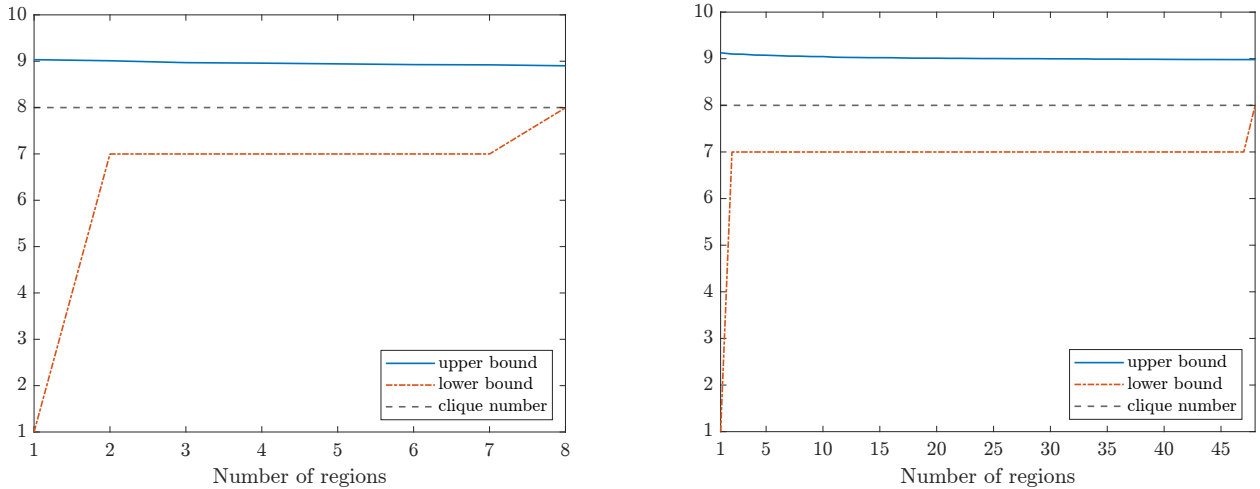
Table 3: Lower and upper bounds on the clique number, and the number of subregions produced by Algorithm 3 for four random instances of  $G(75, 0.5)$ .

Instance	$\omega(G)$	$\vartheta'(\bar{G})$	Upper bound	Lower bound	# subregions
1	8	9.0336	8.9014	7.9956	8
2	8	9.1274	8.9815	7.9967	48
3	8	8.9770	8.8686	7.9950	4
4	8	9.1400	8.9952	7.9973	16

To estimate the clique number by Algorithm 3, at each node of the branch-and-bound tree, we perform 10 projected gradient descent steps in (29) with  $Q = I_n + \bar{A}$  to obtain a lower bound on  $\omega(G)$ , and apply our inner approximation of  $\mathcal{C}_n$  to (30) to obtain an upper bound on  $\omega(G)$ . Since the clique number is an integer, the algorithm terminates when the ceiled lower bound and the floored upper bound coincide.

We present some experiments on Erdős-Rényi random graphs [19]  $G(n, p)$ , where  $n$  vertices are connected pairwise and independently with probability  $p \in (0, 1)$ . We generate four random instances with  $n = 75$  and  $p = 0.5$ , and compare our bounds against the Schrijver upper bound  $\vartheta'(\bar{G})$  on  $\omega(G)$  [48].<sup>10</sup> Table 3 shows that our algorithm successfully determines the exact clique number in all four instances using a moderate number of subregions.

Figure 8 shows the lower and upper bounds, produced by Algorithm 3, on the clique number of the four random instances of  $G(75, 0.5)$  versus the number of subregions.



<sup>10</sup>The Schrijver bound  $\vartheta'(\bar{G})$  satisfies  $\omega(G) \leq \vartheta'(\bar{G}) \leq \vartheta(\bar{G})$ , where  $\vartheta$  is the more familiar Lovász theta function [35].

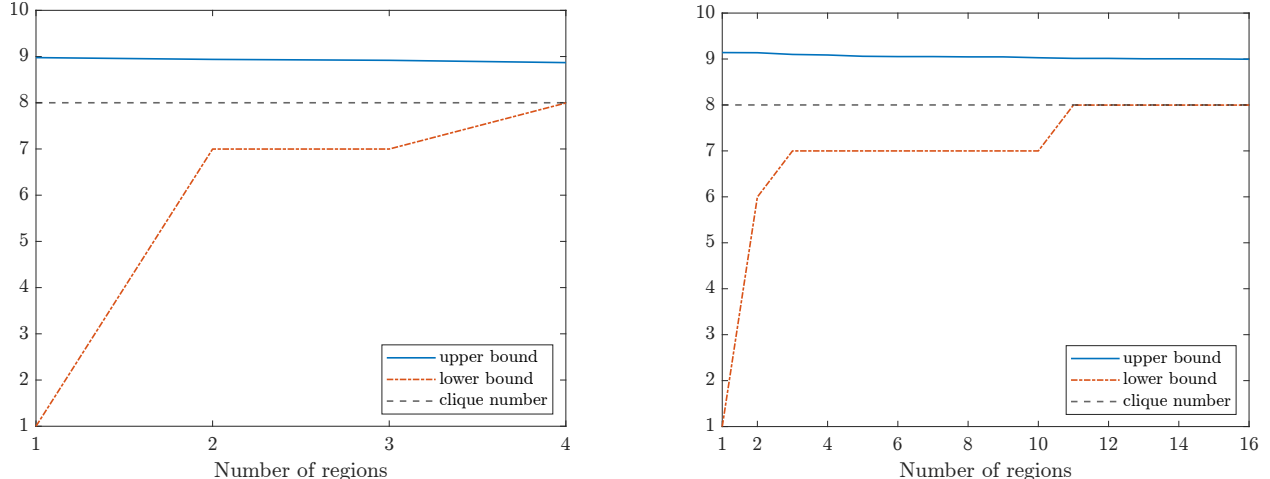


Figure 8: Lower and upper bounds on the clique numbers of four random instances of  $G(75, 0.5)$  produced by Algorithm 3 versus the number of subregions.

## 7 Conclusions and Future Directions

In this paper, we introduced the disjunctive sum of squares method for certifying nonnegativity of polynomials. Unlike the classical sum of squares approach, which certifies nonnegativity via a single algebraic identity, our method constructs multiple algebraic identities, which together certify nonnegativity. We showed that by doing so, the degrees of the sum of squares polynomials involved in the proof (and hence the size of the underlying semidefinite constraints) can be kept low and fixed. We proposed algorithms that can search for disjunctive proofs of nonnegativity, and studied extensions of our method to copositive programs and polynomial optimization problems.

This work leaves several interesting directions for future research. While we proposed a general reduction from constrained polynomial optimization problems to unconstrained ones in Section 5.1, future work could tailor the design of algebraic disjunctions to the geometry of specific feasible sets (as we did, e.g., in the case of the nonnegative orthant in Section 5.2). Another interesting direction is to develop more efficient implementations of the disjunctive sum of squares approach that in addition to utilizing spatial branch-and-bound (cf. Section 6), potentially exploit the natural parallelism across subregions. Along these lines, it would be interesting to combine the disjunctive sum of squares method with techniques for exploiting sparsity [54]–[56] and symmetry [20], [46] in sum of squares programs. Another avenue to study is the dual side of our framework through the theory of moments [27], [30], which has proven useful for extracting solutions in polynomial optimization problems [24]. Finally, a promising direction is the design of instance-specific algebraic disjunctions that require fewer subregions than the universal constructions provided in this paper (cf. Theorems 1 and 8). The branch-and-bound framework of Section 6 constructs such disjunctions adaptively, but there could be techniques that in addition exploit algebraic properties of the given problem instance.

We end with an open problem of theoretical interest: does there exist a nonnegative form that admits no disjunctive sum of squares proof of nonnegativity with any number of identities? This question has two natural variants: one where the degree of the sum of squares proof is required to be the same as that of the given form, and one where it is allowed to be larger. By Theorem 1, any such form must have a zero. A natural candidate would be a polynomial with a “bad point”; see, e.g., [45], [5] and references therein. An example of such a polynomial is Delzell’s polynomial  $D$  (see [16])

$$D(x_1, \dots, x_4) = x_1^4 x_2^2 x_4^2 + x_2^4 x_3^2 x_4^2 + x_1^2 x_3^4 x_4^2 - 3x_1^2 x_2^2 x_3^2 x_4^2 + x_3^8,$$

which was also a test case for us in Section 6.3.1. It is known that while  $D$  is nonnegative,  $(\sum_{i=1}^4 x_i^2)^k D(x_1, \dots, x_4)$  is not sos for any integer  $k \geq 0$ , which implies that the hierarchy of SDPs based on Reznick’s Positivstellensatz [43] cannot certify nonnegativity of  $D$ . It turns out, however, that the disjunctive sum of squares approach can produce a low-degree proof of nonnegativity of  $D$  only with two subregions:

$$\begin{aligned} D(x_1, \dots, x_4) &= (x_1^2 x_2 x_4 - x_1 x_3^2 x_4)^2 + (x_2^2 x_3 x_4 - x_1 x_2 x_3 x_4)^2 + (x_3^4)^2 + 2x_1 x_2 (x_1 x_3 x_4 - x_2 x_3 x_4)^2 \\ &= (x_1^2 x_2 x_4 + x_1 x_3^2 x_4)^2 + (x_2^2 x_3 x_4 + x_1 x_2 x_3 x_4)^2 + (x_3^4)^2 - 2x_1 x_2 (x_1 x_3 x_4 + x_2 x_3 x_4)^2. \end{aligned}$$

## Acknowledgments

We are grateful to Abraar Chaudhry, Georgina Hall, and Jeffrey Zhang for insightful discussions.

## References

- [1] A. A. Ahmadi and G. Hall, “Sum of squares basis pursuit with linear and second-order cone programming,” in *Contemporary Mathematics*, Providence, RI: American Mathematical Society, pp. 27–53, 2017.
- [2] A. A. Ahmadi and G. Hall, “On the construction of converging hierarchies for polynomial optimization based on certificates of global positivity,” *Mathematics of Operations Research*, vol. 44, no. 4, pp. 1192–1207, 2019.
- [3] A. A. Ahmadi and A. Majumdar, “DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization,” *SIAM Journal on Applied Algebra and Geometry*, vol. 3, no. 2, pp. 193–230, 2019.
- [4] E. Artin, “Über die Zerlegung definiter Funktionen in Quadrate,” *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 5, pp. 100–115, 1927.
- [5] O. Benoist, “On the bad points of positive semidefinite polynomials,” *Mathematische Zeitschrift*, vol. 300, pp. 3383–3403, 2022.
- [6] D. Bertsimas and I. Popescu, “Optimal inequalities in probability theory: A convex optimization approach,” *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 780–804, 2005.

- [7] G. Blekherman, J. Gouveia, and J. Pfeiffer, “Sums of squares on the hypercube,” *Mathematische Zeitschrift*, vol. 284, pp. 41–54, 2016.
- [8] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*. Philadelphia: SIAM, 2012.
- [9] S. Bundfuss and M. Dür, “Algorithmic copositivity detection by simplicial partition,” *Linear Algebra and its Applications*, vol. 428, no. 7, pp. 1511–1523, 2008.
- [10] S. Bundfuss and M. Dür, “An adaptive linear approximation algorithm for copositive programs,” *SIAM Journal on Optimization*, vol. 20, no. 1, pp. 30–53, 2009.
- [11] S. Burer, “Copositive programming,” in *Handbook on Semidefinite, Conic and Polynomial Optimization*, New York: Springer, pp. 201–218, 2012.
- [12] S. Chakraborty and H. Narayanan, “Generating an equidistributed net on a sphere using random rotations,” *Discrete & Computational Geometry*, vol. 67, pp. 231–257, 2022.
- [13] M. D. Choi and T. Y. Lam, “An old question of Hilbert,” in *Proceedings of Quadratic Forms Conference*, ser. Queen’s Papers in Pure and Applied Mathematics, G. Orzech, Ed., vol. 46, Kingston, Ontario: Queen’s University, pp. 385–405, 1976.
- [14] M. D. Choi and T. Y. Lam, “Extremal positive semidefinite forms,” *Mathematische Annalen*, vol. 231, pp. 1–18, 1977.
- [15] M. Curmei and G. Hall, “Shape-constrained regression using sum of squares polynomials,” *Operations Research*, vol. 73, no. 1, pp. 543–559, 2025.
- [16] C. N. Delzell, “A constructive, continuous solution to Hilbert’s 17th problem, and other results in semi-algebraic geometry,” Ph.D. dissertation, Stanford University, 1980.
- [17] P. H. Diananda, “On non-negative forms in real variables some or all of which are non-negative,” *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, no. 1, pp. 17–25, 1962.
- [18] M. Dür, “Copositive programming — a survey,” in *Recent Advances in Optimization and its Applications in Engineering*, Berlin, Heidelberg: Springer, pp. 3–20, 2010.
- [19] P. Erdős and A. Rényi, “On random graphs I,” *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.
- [20] K. Gatermann and P. A. Parrilo, “Symmetry groups, semidefinite programs, and sums of squares,” *Journal of Pure and Applied Algebra*, vol. 192, no. 1–3, pp. 95–128, 2004.
- [21] S. A. Gershgorin, “Über die Abgrenzung der Eigenwerte einer Matrix,” *Bulletin de l’Académie des Sciences de l’URSS, Classe des sciences mathématiques et naturelles*, no. 6, pp. 749–754, 1931.
- [22] G. Hall, “Applications of sums of squares,” in *Sum of Squares: Theory and Applications*, ser. Proceedings of Symposia in Applied Mathematics, vol. 77, Providence, RI: American Mathematical Society, pp. 103–140, 2020.

- [23] D. Henrion and A. Garulli, *Positive polynomials in control*. Berlin, Heidelberg: Springer, vol. 312, 2005.
- [24] D. Henrion and J. B. Lasserre, “Detecting global optimality and extracting solutions in GloptiPoly,” in *Positive Polynomials in Control*, ser. Lecture Notes in Control and Information Sciences, vol. 312, Berlin, Heidelberg: Springer, pp. 293–310, 2005.
- [25] D. Hilbert, “Über die Darstellung Definitiver Formen als Summe von Formenquadraten,” *Mathematische Annalen*, vol. 32, no. 3, pp. 342–350, 1888.
- [26] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*. Berlin, Heidelberg: Springer, 1996.
- [27] J. B. Lasserre, “Global optimization with polynomials and the problem of moments,” *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.
- [28] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*. Singapore: World Scientific, vol. 1, 2009.
- [29] J. B. Lasserre, *An Introduction to Polynomial and Semi-algebraic Optimization*. Cambridge: Cambridge University Press, vol. 52, 2015.
- [30] J. B. Lasserre, “The moment-SOS hierarchy,” in *Proceedings of the International Congress of Mathematicians*, 2019.
- [31] M. Laurent, “Sums of squares, moment matrices and optimization over polynomials,” in *Emerging Applications of Algebraic Geometry*, New York: Springer, pp. 157–270, 2009.
- [32] P. C. Leopardi, “A partition of the unit sphere into regions of equal area and small diameter,” *Electronic Transactions on Numerical Analysis*, vol. 25, pp. 309–327, 2006.
- [33] M. Locatelli and F. Schoen, *Global Optimization: Theory, Algorithms, and Applications* (MOS-SIAM Series on Optimization). Philadelphia: SIAM, vol. 15, 2013.
- [34] H. Lombardi, D. Perrucci, and M.-F. Roy, “An elementary recursive bound for effective Positivstellensatz and Hilbert 17-th problem,” *Memoirs of the American Mathematical Society*, vol. 263, no. 1277, 2020.
- [35] L. Lovász, “On the Shannon capacity of a graph,” *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, 1979.
- [36] T. S. Motzkin and E. G. Straus, “Maxima for graphs and a new proof of a theorem of turán,” *Canadian Journal of Mathematics*, vol. 17, pp. 533–540, 1965.
- [37] T. S. Motzkin, “The arithmetic-geometric inequality,” in *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*, New York: Academic Press, pp. 205–224, 1967.
- [38] K. G. Murty and S. N. Kabadi, “Some NP-complete problems in quadratic and non-linear programming,” *Mathematical Programming*, vol. 39, no. 2, pp. 117–129, 1987.

- [39] P. A. Parrilo, “Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization,” Ph.D. dissertation, California Institute of Technology, 2000.
- [40] P. A. Parrilo, “Semidefinite programming relaxations for semialgebraic problems,” *Mathematical Programming*, vol. 96, pp. 293–320, 2003.
- [41] P. A. Parrilo and B. Sturmfels, “Minimizing polynomial functions,” in *Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science*, ser. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 60, Providence, RI: American Mathematical Society, pp. 83–99, 2003.
- [42] M. Putinar, “Positive polynomials on compact semi-algebraic sets,” *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.
- [43] B. Reznick, “Uniform denominators in Hilbert’s seventeenth problem,” *Mathematische Zeitschrift*, vol. 220, no. 1, pp. 75–97, 1995.
- [44] B. Reznick, “Some concrete aspects of Hilbert’s 17th problem,” *Contemporary mathematics*, vol. 253, pp. 251–272, 2000.
- [45] B. Reznick, “On the absence of uniform denominators in Hilbert’s 17th problem,” *Proceedings of the American Mathematical Society*, vol. 133, no. 10, pp. 2829–2834, 2005.
- [46] C. Riener, T. Theobald, L. Jansson Andrén, and J. B. Lasserre, “Exploiting symmetries in SDP-relaxations for polynomial optimization,” *Mathematics of Operations Research*, vol. 38, no. 1, pp. 122–141, 2013.
- [47] K. Schmüdgen, “The K-moment problem for compact semi-algebraic sets.,” *Mathematische Annalen*, vol. 289, no. 2, pp. 203–206, 1991.
- [48] A. Schrijver, “A comparison of the Delsarte and Lovász bounds,” *IEEE Transactions on Information Theory*, vol. 25, no. 4, pp. 425–429, 1979.
- [49] G. Stengle, “A Nullstellensatz and a Positivstellensatz in semialgebraic geometry,” *Mathematische Annalen*, vol. 207, pp. 87–97, 1974.
- [50] G. Stengle, “Integral solution of Hilbert’s seventeenth problem,” *Mathematische Annalen*, vol. 246, pp. 33–39, 1979.
- [51] R. Tedrake, I. R. Manchester, M. Tobenkin, and J. W. Roberts, “LQR-trees: Feedback motion planning via sums-of-squares verification,” *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1038–1052, 2010.
- [52] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge: Cambridge University Press, vol. 47, 2018.
- [53] M. Vianello, “Subperiodic Dubiner distance, norming meshes and trigonometric polynomial optimization,” *Optimization Letters*, vol. 12, pp. 1659–1667, 2018.

- [54] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity,” *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 218–242, 2006.
- [55] Y. Zheng and G. Fantuzzi, “Sum-of-squares chordal decomposition of polynomial matrix inequalities,” *Mathematical Programming*, vol. 197, pp. 71–108, 2022.
- [56] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, “Chordal and factor-width decompositions for scalable semidefinite and polynomial optimization,” *Annual Reviews in Control*, vol. 52, pp. 243–279, 2021.