ON THE CONSTRUCTION OF CONVERGING HIERARCHIES FOR POLYNOMIAL OPTIMIZATION BASED ON CERTIFICATES OF GLOBAL POSITIVITY*

AMIR ALI AHMADI[†] AND GEORGINA HALL[‡]

Abstract. In recent years, techniques based on convex optimization and real algebra that produce converging hierarchies of lower bounds for polynomial minimization problems have gained much popularity. At their heart, these hierarchies rely crucially on Positivstellensätze from the late 20th century (e.g., due to Stengle, Putinar, or Schmüdgen) that certify positivity of a polynomial on an arbitrary closed basic semialgebraic set. In this paper, we show that such hierarchies could in fact be designed from much more limited Positivstellensätze dating back to the early 20th century that only certify positivity of a polynomial globally. More precisely, we show that any inner approximation to the cone of positive homogeneous polynomials that is arbitrarily tight can be turned into a converging hierarchy of lower bounds for general polynomial minimization problems with compact feasible sets. This in particular leads to a semidefinite programming-based hierarchy that relies solely on Artin's solution to Hilbert's 17th problem. We also use a classical result of Polyá on global positivity of even forms to construct an "optimization-free" converging hierarchy for general polynomial minimization problems with compact feasible sets. This hierarchy only requires polynomial multiplication and checking nonnegativity of coefficients of certain fixed polynomials. As a corollary, we obtain new linear programming and second-order cone programming-based hierarchies for polynomial minimization problems that rely on the recently introduced concepts of dsos (diagonally dominant sum of squares) and sdsos (scaled diagonally dominant sum of squares) polynomials. We remark that the scope of this paper is theoretical at this stage as our hierarchies—though they involve at most two sum of squares constraints or only elementary arithmetic at each level—require the use of bisection and increase the number of variables (resp. degree) of the problem by the number of inequality constraints plus three (resp. by a factor of two).

Key words. Positivstellensätze, polynomial optimization, convex optimization

AMS subject classifications. 14Q99, 90C05, 90C26, 90C22, 90-03

1. Introduction. A polynomial optimization problem (POP) is an optimization problem of the form

$$(1) \qquad \begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & p(x) \\ \text{s.t.} \quad & g_i(x) \geq 0, \ i = 1, \ldots, m, \end{aligned}$$

where $p, g_i, \ i = 1, \ldots, m$, are polynomial functions in n variables $x := (x_1, \ldots, x_n)$ and with real coefficients. It is well-known that polynomial optimization is a hard problem to solve in general. For example, simply testing whether the optimal value of problem (1) is smaller than or equal to some rational number k is NP-hard already when the objective is quadratic and the constraints are linear [19]. Nevertheless, these problems remain topical due to their numerous applications throughout engineering, operations research, and applied mathematics (see, e.g., [15, 6, 2]). In this paper, we are interested in obtaining lower bounds on the optimal value of problem (1). We focus on a class of methods which construct hierarchies of tractable convex optimization problems whose optimal values are lower bounds on the optimal value of (1), with

[†]Department of Operations Research and Financial Engineering, Princeton, NJ 08540 (a_a_a@princeton.edu, http://aaa.princeton.edu).

[‡]Department of Operations Research and Financial Engineering, Princeton, NJ 08540 (gh4@princeton.edu, http://scholar.princeton.edu/ghall).

convergence to it as the sequence progresses. This implies that even though the original POP is nonconvex, one can obtain increasingly accurate lower bounds on its optimal value by solving convex optimization problems. One method for constructing these hierarchies of optimization problems that has gained attention in recent years relies on the use of *Positivstellensätze* (see, e.g., [16] for a survey). Positivstellensätze are algebraic identities that certify infeasibility of a set of polynomial inequalities, or equivalently[1], positivity of a polynomial on a basic semialgebraic set. (Recall that a basic semialgebraic set is a set defined by finitely many polynomial inequalities.) These Positivstellensätze can be used to prove lower bounds on POPs. Indeed, if we denote the feasible set of (1) by S, the optimal value of problem (1) is equivalent to

$$(2) \qquad \begin{aligned} \sup_{\gamma} \quad & \gamma \\ \text{s.t.} \quad & p(x) - \gamma \geq 0, \ \forall x \in S. \end{aligned}$$

Hence if γ is a strict lower bound on (1), we have that $p(x) - \gamma > 0$ on S, a fact that can be certified using Positivstellensätze. At a conceptual level, hierarchies that provide lower bounds on (1) are constructed thus: we fix the "size of the certificate" at each level of the hierarchy and search for the largest γ such that the Positivstellensätze at hand can certify positivity of $p(x) - \gamma$ over S with a certificate of this size. As the sequence progresses, we increase the size of the certificates allowed, hence obtaining increasingly accurate lower bounds on (1).

Below, we present three of the better-known Positivstellensätze, given respectively by Stengle [31], Schmüdgen [27], and Putinar [25]. These all rely on sum of squares certificates. We recall that a polynomial is a *sum of squares* (sos) if it can be written as a sum of squares of other polynomials. We start with Stengle's Positivstellensatz, which certifies infeasibility of a set of polynomial inequalities. It is sometimes referred to as "the Positivstellensatz" in related literature as it requires no assumptions, contrarily to Schmüdgen and Putinar's theorems which can be viewed as refinements of Stengle's result under additional assumptions. This Positivstellensatz was in fact discovered by Krivine in 1964 [13], and rediscovered by Stengle later[2]; see [24, Section 4.7] for a more complete history of this result.

THEOREM 1.1 (Stengle's Positivstellensatz [31]). *The closed basic semialgebraic set*

$$S = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

is empty if and only if there exist sum of squares polynomials $s_0(x), s_1(x), \ldots, s_m(x),$ $s_{12}(x), s_{13}(x), \ldots, s_{123\ldots m}(x)$ such that

$$-1 = s_0(x) + \sum_i s_i(x)g_i(x) + \sum_{\{i,j\}} s_{ij}(x)g_i(x)g_j(x) + \ldots + s_{123\ldots m}(x)g_1(x)\ldots g_m(x).$$

The next two theorems, due to Schmüdgen and Putinar, certify positivity of a polynomial p over a closed basic semialgebraic set S. They impose additional compactness assumptions comparatively to Stengle's Positivstellensatz.

THEOREM 1.2 (Schmüdgen's Positivstellensatz [27]). *Assume that the set*

$$S = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

---

[1]Note that the set $\{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$ is empty if and only if $-g_1(x) > 0$ on the set $\{x \in \mathbb{R}^n \mid g_2(x) \geq 0, \ldots, g_m(x) \geq 0\}$.

[2]We thank an anonymous referee for pointing this out to us and for providing us with the appropriate references.

is compact. If a polynomial p is positive on S, then

$$p(x) = s_0(x) + \sum_i s_i(x)g_i(x) + \sum_{\{i,j\}} s_{ij}(x)g_i(x)g_j(x) + \ldots + s_{123\ldots m}(x)g_1(x)\ldots g_m(x),$$

where $s_0(x), s_1(x), \ldots, s_m(x), s_{12}(x), s_{13}(x), \ldots, s_{123\ldots m}(x)$ are sums of squares.

THEOREM 1.3 (Putinar's Positivstellensatz [25]). *Let*

$$S = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

and assume that $\{g_1, \ldots, g_m\}$ satisfy the Archimedean property, i.e., there exists $N \in \mathbb{N}$ such that

$$N - \sum_i x_i^2 = \sigma_0(x) + \sigma_1(x)g_1(x) + \ldots + \sigma_m(x)g_m(x),$$

where $\sigma_1(x), \ldots, \sigma_m(x)$ are sums of squares. If a polynomial p is positive on S, then

$$p(x) = s_0(x) + s_1(x)g_1(x) + \ldots + s_m(x)g_m(x),$$

where $s_1(x), \ldots, s_m(x)$ are sums of squares.

Note that these three Positivstellensätze involve in their expressions sum of squares polynomials of unspecified degree. To construct hierarchies of tractable optimization problems for (2), we fix this degree: at level r, we search for the largest γ such that positivity of $p(x) - \gamma$ over S can be certified using the Positivstellensätze where the degrees of all sos polynomials are taken to be less than or equal to $2r$. Solving each level of these hierarchies is then a semidefinite program (SDP). This is a consequence of the fact that one can optimize over (or test membership to) the set of sum of squares polynomials of fixed degree using semidefinite programming [21, 20, 14]. Indeed, a polynomial p of degree $2d$ and in n variables is a sum of squares if and only if there exists a symmetric matrix $Q \succeq 0$ such that $p(x) = z(x)^T Q z(x)$, where $z(x) = (1, x_1, \ldots, x_n, \ldots, x_n^d)^T$ is the standard vector of monomials in n variables and of degree less than or equal to d. We remark that the hierarchy obtained from Stengle's Positivstellensatz was proposed and analyzed by Parrilo in [21]; the hierarchy obtained from Putinar's Positivstellensatz was proposed and analyzed by Lasserre in [14]. There have been more recent works that provide constructive proofs of Schmüdgen and Putinar's Positivstellensätze; see [5, 28, 30]. These proofs rely on other Positivstellensätze, e.g., a result by Polyá (see Theorem 1.6 below) in [28, 30], and the same result by Polyá, Farkas' lemma, and Stengle's Positivstellensatz in [5]. We would like to thank an anonymous referee for pointing out that the construction in [28] can be used to develop converging hierarchies of lower bounds for POPs with compact feasible sets. These hierarchies rely on Gröbner bases computations and linear programs involving only two variables. Some experiments with this technique were carried out by Datta [7] and Averkov has more recently shown [5] that the (potentially expensive) Gröbner bases computations can be avoided in this approach. Other recent research efforts relating to Positivstellensätze have been focused around deriving complexity bounds for Schmüdgen and Putinar's Positivstellensätze; see [18, 29].

On a historical note, Stengle, Schmüdgen, and Putinar's Positivstellensätze were derived in the latter half of the 20th century. As mentioned previously, they all certify positivity of a polynomial over an arbitrary basic semialgebraic set (modulo compactness assumptions). By contrast, there are Positivstellensätze from the early

20th century that certify positivity of a polynomial *globally.* Perhaps the most well-known Positivstellensatz of this type is due to Artin in 1927, in response to Hilbert's 17th problem. Artin shows that any nonnegative polynomial is a sum of squares of rational functions. Here is an equivalent formulation of this statement:

THEOREM 1.4 (Artin [4]). *For any nonnegative polynomial p, there exists a nonzero sos polynomial q such that $p \cdot q$ is a sum of squares.*

To the best of our knowledge, in this area, all converging hierarchies of lower bounds for POPs are based off of Positivstellensätze that certify nonnegativity of a polynomial over an arbitrary basic semialgebraic set. In this paper, we show that in fact, under compactness assumptions, it suffices to have only global certificates of nonnegativity (such as the one given by Artin) to produce a converging hierarchy for general POPs. As a matter of fact, even weaker statements that apply only to globally positive (as opposed to globally nonnegative) forms are enough to derive converging hierarchies for POPs. Examples of such statements are due to Habicht [11] and Reznick [26]. With such an additional positivity assumption, more can usually be said about the structure of the polynomial q in Artin's result. Below, we present the result by Reznick.

THEOREM 1.5 (Reznick [26]). *For any positive definite form p, there exists $r \in \mathbb{N}$ such that $p(x) \cdot (\sum_i x_i^2)^r$ is a sum of squares.*

We show in this paper that this Positivstellensatz also gives rise to a converging hierarchy for POPs with a compact feasible set similarly to the one generated by Artin's Positivstellensatz.

Through their connections to sums of squares, the two hierarchies obtained using the theorems of Reznick and Artin are semidefinite programming-based. In this paper, we also derive an "optimization-free" converging hierarchy for POPs with compact feasible sets where each level of the hierarchy only requires that we be able to test nonnegativity of the coefficients of a given fixed polynomial. To the best of our knowledge, this is the first converging hierarchy of lower bounds for POPs which does not require that convex optimization problems be solved at each of its levels. To construct this hierarchy, we use a result of Polyá [22], which just like Artin's and Reznick's Positivstellensätze, certifies global positivity of forms. However this result is restricted to even forms. Recall that a form p is *even* if each of the variables featuring in its individual monomials has an even power. This is equivalent (see [8, Lemma 2]) to p being invariant under change of sign of each of its coordinates, i.e.,

$$p(x_1, \ldots, x_n) = p(-x_1, \ldots, x_n) = \cdots = p(x_1, \ldots, -x_n).$$

THEOREM 1.6 (Polyá [22]). *For any positive definite even form p, there exists $r \in \mathbb{N}$ such that $p(x) \cdot (\sum_i x_i^2)^r$ has nonnegative coefficients.*

A perhaps better-known but equivalent formulation of this theorem is the following: for any form h that is positive on the standard simplex, there exists $r \in \mathbb{N}$ such that $h(x) \cdot (\sum_i x_i)^r$ has nonnegative coefficients. The two formulations are equivalent by simply letting $p(x) = h(x^2)$. The latter formulation has been used to derive similar optimization-free converging hierarchies of lower bounds for polynomial minimization problems over the simplex; see, e.g., [9, 10].

Our aforementioned optimization-free hierarchy also enables us to obtain linear programming (LP) and second-order cone programming (SOCP)-based hierarchies for general POPs with compact feasible sets that rely on the concepts of *dsos* and *sdsos*

polynomials. These are recently introduced inner approximations to the set of sos polynomials that have shown much better scalability properties in practice [3].

As a final remark, we wish to stress the point that the goal of this paper is first and foremost theoretical, i.e., to provide methods for constructing converging hierarchies of lower bounds for POPs using as sole building blocks certificates of global positivity. We do not make any claims that these hierarchies can outperform the popular existing hierarchies due, e.g., to Lasserre [14] and Parrilo [21]. In particular, all hierarchies that we generate increase the number of variables and the degree of the polynomials involved from n to $n + m + 3$, and from d to $2d$, respectively. They also necessitate the use of bisection, which, while not a problem in theory, increases the computational overload. We remark however that each level of our hierarchies only involves either one sum of squares constraint (the hierarchy based on the certificate of Reznick; Theorem 3.2), two sum of squares constraints (the hierarchy based on the certificate of Artin; Theorem 3.4), or nothing but elementary computations (the hierarchy based on the certificate of Polyá; Theorem 4.1). By contrast, each level of the hierarchy based on Putinar's (resp. Schmüdgen's) certificate involves $m + 1$ (resp. 2^m) sum of squares constraints, but necessitates no need to use bisection or to increase the number of variables/degree of the problem. Similarly, a hierarchy based on Stengle's certificate, which would work by showing infeasibility of the constraints $\{\gamma - p(x) \geq 0, g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, requires the use of bisection on γ and 2^{m+1} sum of squares constraints in each level, but necessitates no increase in the number of variables/degree of the problem. Of course, such comparisons would become more meaningful if one could also relate the quality of the bounds obtained from the different approaches. Some remarks on why it is nontrivial to connect our hierarchies to previous ones in this sense are made in Section 5.

1.1. Outline of the paper. The paper is structured as follows. In Section 2, we show that if one can inner approximate the cone of positive definite forms arbitrarily well (with certain basic properties), then one can produce a converging hierarchy of lower bounds for POPs with compact feasible sets (Theorem 2.4). This relies on a reduction (Theorem 2.1) that reduces the problem of certifying a strict lower bound on a POP to that of proving positivity of a certain form. In Section 3, we see how this result can be used to derive semidefinite programming-based converging hierarchies (Theorems 3.2 and 3.4) from the Positivstellensätze by Artin (Theorem 1.4) and Reznick (Theorem 1.5). In Section 4, we derive an optimization-free hierarchy (Theorem 4.1) from the Positivstellensatz of Polyá (Theorem 1.6) as well as LP and SOCP-based hierarchies which rely on dsos/sdsos polynomials (Corollary 4.8). We conclude with a few open problems in Section 5.

1.2. Notation and basic definitions. We use the standard notation $A \succeq 0$ to denote that a symmetric matrix A is positive semidefinite. Recall that a *form* is a homogeneous polynomial, i.e., a polynomial whose monomials all have the same degree. We denote the degree of a form f by $deg(f)$. We say that a form f is *nonnegative* (or positive semidefinite) if $f(x) \geq 0$, for all $x \in \mathbb{R}^n$ (we write $f \geq 0$). A form f is *positive definite* (pd) if $f(x) > 0$, for all nonzero x in $\mathbb{R}^n$ (we write $f > 0$). Throughout the paper, we denote the set of forms (resp. the set of nonnegative forms) in n variables and of degree d by $H_{n,d}$ (resp $P_{n,d}$). We denote the ball of radius R and centered at the origin by $B(0, R)$ and the unit sphere in x-space, i.e., $\{x \in \mathbb{R}^n \mid ||x||_2 = 1\}$, by S_x. We use the shorthand $f(y^2 - z^2)$ for $y, z \in \mathbb{R}^n$ to denote $f(y_1^2 - z_1^2, \ldots, y_n^2 - z_n^2)$. We say that a scalar γ is a strict lower bound on (1) if $p(x) > \gamma, \forall x \in S$. Finally, we ask the reader to carefully read Remark 2.3 which

contains the details of a notational overwriting occurring before Theorem 2.4 and valid from then on throughout the paper. This overwriting makes the paper much simpler to parse.

2. Constructing converging hierarchies for POP using global certificates of positivity. Consider the polynomial optimization problem in (1) and denote its optimal value by p^*. Let d be such that $2d$ is the smallest even integer larger than or equal to the maximum degree of $p, g_i, i = 1, \ldots, m$. We denote the feasible set of our optimization problem by

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, m\}$$

and assume that S is contained within a ball of radius R. From this, it is easy to provide (possibly very loose) upper bounds on $g_i(x)$ over the set S: as S is contained in a ball of radius R, we have $|x_i| \leq R$, for all $i = 1, \ldots, n$. We then use this to upper bound each monomial in g_i and consequently g_i itself. We use the notation η_i to denote these upper bounds, i.e., $g_i(x) \leq \eta_i$, for all $i = 1, \ldots, m$ and for all $x \in S$. Similarly, we can provide an upperbound on $-p(x)$. We denote such a bound by β, i.e., $-p(x) \leq \beta, \forall x \in S$.

The goal of this section is to produce a method for constructing converging hierarchies of lower bounds for POPs if we have access to arbitrarily accurate inner approximations of the set of positive definite forms. The first theorem (Theorem 2.1) connects lower bounds on (1) to positive definiteness of a related form. The second theorem (Theorem 2.4) shows how this can be used to derive a hierarchy for POPs.

THEOREM 2.1. *Consider the general polynomial optimization problem in (1) and recall that d is such that $2d$ is the smallest even integer larger than or equal to the maximum degree of $p, g_i, i = 1, \ldots, m$. Suppose $S \subseteq B(0, R)$ for some positive scalar R. Let $\eta_i, i = 1, \ldots, m$ (resp. β) be any finite upper bounds on $g_i(x), i = 1, \ldots, m$ (resp. $-p(x)$).*

Then, a scalar γ is a strict lower bound on (1) if and only if the homogeneous sum of squares polynomial

$$(3) \quad f_\gamma(x, s, y) := \left(\gamma y^{2d} - y^{2d}p(x/y) - s_0^2 y^{2d-2}\right)^2 + \sum_{i=1}^{m} \left(y^{2d}g_i(x/y) - s_i^2 y^{2d-2}\right)^2$$

$$+ \left((R + \sum_{i=1}^{m} \eta_i + \beta + \gamma)^d y^{2d} - (\sum_{i=1}^{n} x_i^2 + \sum_{i=0}^{m} s_i^2)^d - s_{m+1}^{2d}\right)^2$$

of degree $4d$ and in $n + m + 3$ variables $(x_1, \ldots, x_n, s_0, \ldots, s_m, s_{m+1}, y)$ is positive definite.[3]

Proof. It is easy to see that γ is a strict lower bound on (1) if and only if the set

$$T := \{x \in \mathbb{R}^n \mid \gamma - p(x) \geq 0; \quad g_i(x) \geq 0, i = 1, \ldots, m; \quad \sum_i x_i^2 \leq R\}$$

is empty. Indeed, if T is nonempty, then there exists a point $x \in S$ such that $p(x) \leq \gamma$. This implies that γ cannot be a strict lower bound on (1). Conversely, if T is empty, the intersection of S with $\{x \mid \gamma - p(x) \geq 0\}$ is empty, which implies that $\forall x \in S$, $p(x) > \gamma$.

---

[3] The reader will observe in the proof that the variables $(s_0, \ldots, s_{m+1})$ will serve as slack variables and the variable y will be used for homogenization.

We now define the set:

$$T_s = \{(x,s) \in \mathbb{R}^{n+m+2} \mid \gamma - p(x) = s_0^2; \quad g_i(x) = s_i^2, i = 1, \dots, m;$$

(4)

$$(R + \sum_{i=1}^m \eta_i + \beta + \gamma)^d - (\sum_{i=1}^n x_i^2 + \sum_{i=0}^m s_i^2)^d - s_{m+1}^{2d} = 0\}.$$

Note that T_s is empty if and only if T is empty. Indeed, if T_s is nonempty, then there exists $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^{m+2}$ such that the three sets of equations are satisfied. This obviously implies that $\gamma - p(x) \geq 0$ and that $g_i(x) \geq 0$, for all $i = 1, \dots, m$. It further implies that $\sum_i x_i^2 \leq R$ as by assumption, if $x \in S$, then x is in a ball of radius R. Conversely, suppose now that T is nonempty. There exists x such that $\gamma - p(x) \geq 0$, $g_i(x) \geq 0$ for $i = 1, \dots, m$, and $\sum_i x_i^2 \leq R$. Hence, there exist $s_0, \dots, s_m$ such that

$$\gamma - p(x) = s_0^2 \text{ and } g_i(x) = s_i{}^2, \ i = 1, \dots, m.$$

Combining the fact that $\sum_i x_i{}^2 \leq R$ and the fact that η_i, $i = 1, \dots, m$ (resp. $\gamma + \beta$) are upperbounds on g_i (resp. $\gamma - p(x)$), we obtain:

$$R + \sum_{i=1}^m \eta_i + \beta + \gamma \geq \sum_{i=1}^n x_i{}^2 + \sum_{i=0}^m s_i{}^2.$$

By raising both sides of the inequality to the power d, we show the existence of s_{m+1}.

We now show that T_s is empty if and only if $f_\gamma(x, s, y)$ is positive definite. Suppose that T_s is nonempty, i.e., there exists $(x, s) \in \mathbb{R}^{n+m+2}$ such that the equalities given in (4) hold. Note then that $f_\gamma(x, s, 1) = 0$. As $(x, s, 1)$ is nonzero, this implies that $f_\gamma(x, s, y)$ is not positive definite.

For the converse, assume that $f_\gamma(x, s, y)$ is not positive definite. As $f_\gamma(x, s, y)$ is a sum of squares and hence nonnegative, this means that there exists nonzero (x, s, y) such that $f(x, s, y) = 0$. We proceed in two cases. If $y \neq 0$, it is easy to see that $(x/y, s/y) \in T_s$ and T_s is nonempty. Consider now the case where $y = 0$. The third square in f_γ being equal to zero gives us:

$$-(\sum_i x_i^2 + \sum_{i=0}^m s_i^2)^d = s_{m+1}^{2d}.$$

This implies that $s_{m+1} = 0$ and that $x_1 = \dots = x_m = s_0 = \dots = s_m = 0$ which contradicts the fact that (x, s, y) is nonzero. □

REMARK 2.2. *Note that Theorem 2.1 implies that testing feasibility of a set of polynomial inequalities is no harder than checking whether a homogeneous polynomial that is sos has a zero. Indeed, as mentioned before, the basic semialgebraic set*

$$\{x \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$$

is empty if and only if $\gamma = 0$ is a strict lower bound on the POP

$$\inf_x \quad -g_1(x)$$
$$s.t. \quad g_2(x) \geq 0, \dots, g_m(x) \geq 0.$$

In principle, this reduction can open up new possibilities for algorithms for testing feasibility of a basic semialgebraic set. For example, the work in [1] shows that positive

definiteness of a form f is equivalent to global asymptotic stability of the polynomial vector field $\dot{x} = -\nabla f(x)$. One could as a consequence search for Lyapunov functions, as is done in [1, Example 2.1.], to certify positivity of forms. Conversely, simulating trajectories of the above vector field can be used to minimize f and potentially find its nontrivial zeros, which, by our reduction, can be turned into a point that belongs to the basic semialgebraic set at hand.

We further remark that one can always take the degree of the sos form f_γ in (3) whose positivity is under consideration to be equal to four. This can be done by changing the general POP in (1) to only have quadratic constraints and a quadratic objective via an iterative introduction of new variables and new constraints in the following fashion: $x_{ij} = x_i x_j$.

REMARK 2.3 (Notational remark). *As a consequence of Theorem 2.1, we now know that certifying lower bounds on (1) is equivalent to proving positivity of the form f_γ that appears in (3). To simplify notation, we define*

$$N := n + m + 3 \text{ and } D = 2d,$$

where n is the dimension of the decision variable of problem (1), d is such that $2d$ is the smallest even integer larger than or equal to the maximum degree of g_i and p in (1), and m is the number of constraints of problem (1). Note now that the form f_γ is a polynomial in N variables and of degree $2D$.

Our next theorem shows that, modulo some technical assumptions, if one can inner approximate the set of positive definite forms arbitrarily well (conditions (a) and (b)), then one can construct a converging hierarchy for POPs.

THEOREM 2.4. *Let $K^r_{n,2d}$ be a sequence of sets (indexed by r) of homogeneous polynomials in n variables and of degree $2d$ with the following properties:*
(a) $K^r_{n,2d} \subseteq P_{n,2d}, \forall r$, and there exists a pd form $s_{n,2d} \in K^0_{n,2d}$.
(b) If $p > 0$, then $\exists r \in \mathbb{N}$ such that $p \in K^r_{n,2d}$.
(c) $K^r_{n,2d} \subseteq K^{r+1}_{n,2d}, \forall r$.
(d) If $p \in K^r_{n,2d}$, then $\forall \epsilon \in [0,1]$, $p + \epsilon s_{n,d} \in K^r_{n,2d}$.
Recall the definition of $f_\gamma(z)$ given in (3). Consider the hierarchy of optimization problems indexed by r:

(5)
$$l_r := \sup_{\gamma} \quad \gamma$$
$$\text{s.t.} \quad f_\gamma(z) - \frac{1}{r} s_{N,2D}(z) \in K^r_{N,2D}.$$

Then, $l_r \leq p^*$ for all r, $\{l_r\}$ is nondecreasing, and $\lim_{r \to \infty} l_r = p^*$.

Proof. We first show that the sequence $\{l_r\}$ is upperbounded by p^*. Suppose that a scalar γ satisfies

$$f_\gamma(z) - \frac{1}{r} s_{N,2D}(z) \in K^r_{N,2D}.$$

We then have $f_\gamma(z) - \frac{1}{r} s_{N,2D}(z) \in P_{N,2D}$ using (a). This implies that $f_\gamma(z) \geq \frac{1}{r} s_{N,2D}(z)$, and hence f_γ is pd as $s_{N,2D}$ is pd. From Theorem 2.1, it follows that γ has to be a strict lower bound on the optimal value of (1). As $\gamma < p^*$, we have that $l_r \leq p^*$ for all r.

We now show monotonicity of the sequence $\{l_r\}$. Let γ be such that

$$f_\gamma(z) - \frac{1}{r} s_{N,2D}(z) \in K^r_{N,2D}.$$

We have the following identity:

$$f_\gamma(z) - \frac{1}{r+1}s_{N,2D}(z) = f_\gamma(z) - \frac{1}{r}s_{N,2D}(z) + \frac{1}{r(r+1)}s_{N,2D}(z).$$

Now, using the assumption and properties (c) and (d), we conclude that

$$f_\gamma(z) - \frac{1}{r+1}s_{N,2D}(z) \in K_{N,2D}^{r+1}.$$

This implies that

$$\{\gamma \mid f_\gamma(z) - \frac{1}{r}s_{N,2D}(z) \in K_{N,2D}^r\} \subseteq \{\gamma \mid f_\gamma(z) - \frac{1}{r+1}s_{N,2D}(z) \in K_{N,2D}^{r+1}\}$$

and that $l_r \leq l_{r+1}$.

Note that as the sequence $\{l_r\}$ is upper bounded and nondecreasing, it converges. Let us show that the limit of this sequence is p^*. To do this, we show that for any strict lower bound γ on (1), there exists a positive integer r such that $f_\gamma(z) - \frac{1}{r}s_{N,2D}(z) \in K_{N,2D}^r$. By Theorem 2.1, as γ is a strict lower bound, $f_\gamma(z)$ is positive definite. As a form is positive definite if and only if it is positive on the unit sphere, by continuity, there exists a positive integer r' such that $f_\gamma(z) - \frac{1}{r'}s_{N,2D}(z)$ is positive definite. Using (b), this implies that there exists a positive integer r'' such that

(6) $$f_\gamma(z) - \frac{1}{r'}s_{N,2D}(z) \in K_{N,2D}^{r''}.$$

We now proceed in two cases. If $r'' \leq r'$, we take $r = r'$ and use property (c) to conclude. If $r' \leq r''$, we have

$$f_\gamma(z) - \frac{1}{r''}s_{N,2D}(z) = f_\gamma(z) - \frac{1}{r'}s_{N,2D}(z) + \frac{r'' - r'}{r' \cdot r''}s_{N,2D}(z).$$

We take $r = r''$ and use (6) and properties (c) and (d) to conclude. $\square$

REMARK 2.5. *Note that condition (d) is subsumed by the more natural condition that $K_{n,d}^r$ be a convex cone for any n, d, and r. However, there are interesting and relevant cones which we cannot prove to be convex though they trivially satisfy condition (d) (see Theorem 3.2 for an example).*

3. Semidefinite programming-based hierarchies obtained from Artin's and Reznick's Positivstellensätze. In this section, we construct two different semidefinite programming-based hierarchies for POPs using Positivstellensätze derived by Artin (Theorem 1.4) and Reznick (Theorem 1.5). To do this, we introduce two sets of cones that we call the Artin and Reznick cones.

DEFINITION 3.1. *We define the* Reznick cone *of level r to be*

$$R_{n,2d}^r := \{p \in H_{n,2d} \mid p(x) \cdot \left(\sum_{i=1}^n x_i^2\right)^r \text{ is sos}\}.$$

Similarly, we define the Artin cone *of level r to be*

$$A_{n,2d}^r := \{p \in H_{n,2d} \mid p(x) \cdot q(x) \text{ is sos for some nonzero sos form } q \text{ of degree } 2r\}.$$

We show that both of these cones produce hierarchies of the type discussed in Theorem 2.4. Recall that p^* is the optimal value of problem (1) and that f_γ is a polynomial in N variables and of degree $2D$ as defined in (3) and Remark 2.3.

THEOREM 3.2. *Consider the hierarchy of optimization problems indexed by r:*

(7)
$$l_r := \sup_\gamma \quad \gamma$$
$$s.t. \quad f_\gamma(z) - \frac{1}{r}(\sum_{i=1}^N z_i^2)^D \in R_{N,2D}^r.$$

Then, $l_r \leq p^*$ for all r, $\{l_r\}$ is nondecreasing, and $\lim_{r\to\infty} l_r = p^*$.

Proof. It suffices to show that the Reznick cones $R_{n,2d}^r$ satisfy properties (a)-(d) in Theorem 2.4. The result will then follow from that theorem. For property (a), it is clear that, as $(\sum_i x_i^2)^r > 0$ and $p(x) \cdot (\sum_i x_i^2)^r$ is a sum of squares and hence nonnegative, $p(x)$ must be nonnegative, so $R_{n,2d}^r \subseteq P_{n,2d}$. Furthermore, the form $s_{n,2d} := (\sum_i x_i^2)^d$ belongs to $R_{n,2d}^0$ and is positive definite. Property (b) is verified as a consequence of Theorem 1.5. For (c), note that if $p(x) \cdot (\sum_i x_i^2)^r$ is sos, then $p(x) \cdot (\sum_i x_i^2)^{r+1}$ is sos since the product of two sos polynomials is sos. Finally, for property (d), note that $R_{n,2d}^r$ is a convex cone. Indeed, for any $\lambda \in [0,1]$,

$$(\lambda p(x) + (1-\lambda)q(x)) \cdot (\sum_i x_i^2)^r = \lambda p(x)(\sum_i x_i^2)^r + (1-\lambda)q(x)(\sum_i x_i^2)^r$$

is sos if p and q are in $R_{n,2d}^r$. Combining the fact that $R_{n,2d}^r$ is a convex cone and the fact that $(\sum_i x_i^2)^d \in R_{n,d}^r$, we obtain (d). $\qquad\square$

REMARK 3.3. *To solve a fixed level r of the hierarchy given in Theorem 3.2, one must proceed by* bisection *on γ (since the parameter γ appears with a power in the definition of f_γ in (3)). Bisection here would produce a sequence of upper bounds $\{U_k\}$ and lower bounds $\{L_k\}$ on l_r as follows. At iteration k, we test whether $\gamma = \frac{U_k+L_k}{2}$ is feasible for (7). If it is, then we take $L_{k+1} = \frac{U_k+L_k}{2}$ and $U_{k+1} = U_k$. If it is not, we take $U_{k+1} = \frac{U_k+L_k}{2}$ and $L_{k+1} = L_k$. We stop when $|U_{k_\epsilon} - L_{k_\epsilon}| < \epsilon$, where ϵ is a prescribed accuracy, and the algorithm returns $l_{r,\epsilon} = L_{k_\epsilon}$. Note that $l_r - \epsilon \leq l_{r,\epsilon} \leq l_r$ and that to obtain $l_{r,\epsilon}$, one needs to take a logarithmic (in $\frac{1}{\epsilon}$) number of steps using this method.*

Hence, solving the rth level of this hierarchy using bisection can be done by semidefinite programming. Indeed, for a fixed r and γ given by the bisection algorithm, one simply needs to test membership of

$$\left(f_\gamma(z) - \frac{1}{r}(\sum_i z_i^2)^D\right) \cdot (\sum_i z_i^2)^r$$

to the set of sum of squares polynomials. This amounts to solving a semidefinite program. We remark that all semidefinite programming-based hierarchies available only produce an approximate solution to the optimal value of the SDP solved at level r in polynomial time. This is independent of whether they use bisection (e.g., such as the hierarchy given in Theorem 3.2 or the one based on Stengle's Positivstellensatz) or not (e.g., the Lasserre hierarchy).

Our next theorem improves on our previous hierarchy by freeing the multiplier $(\sum_{i=1}^N z_i^2)^r$ and taking advantage of our ability to search for an optimal multiplier using semidefinite programming.

THEOREM 3.4. *Recall the definition of Artin cones from Definition 3.1. Consider the hierarchy of optimization problems indexed by r:*

$$l_r := \sup_{\gamma, q} \quad \gamma$$

(8)

$$s.t. \quad f_\gamma(z) - \frac{1}{r}(\sum_{i=1}^{N} z_i^2)^D \in A_{N,2D}^r.$$

Then, $l_r \leq p^*$ for all r, $\{l_r\}$ is nondecreasing, and $\lim_{r \to \infty} l_r = p^*$.

Proof. Just as the previous theorem, it suffices to show that the Artin cones $A_{n,2d}^r$ satisfy properties (a)-(d) of Theorem 2.4. The proof of property (a) follows the proof given for Theorem 3.2. Property (b) is satisfied as a (weaker) consequence of Artin's result (see Theorem 1.4). For (c), we have that if $p(x) \cdot q(x)$ is sos for some nonzero sos polynomial q of degree $2r$, then $p(x) \cdot q(x) \cdot (\sum_i x_i^2)$ is sos, and $q(x) \cdot (\sum_i x_i^2)$ has degree $2(r+1)$. Finally, for (d), suppose that $p \in A_{n,2d}^r$. Then there exists an sos form q such that $p(x) \cdot q(x)$ is sos. We have

$$\left(p(x) + \epsilon(\sum_i x_i^2)^d \right) \cdot q(x) = p(x) \cdot q(x) + \epsilon(\sum_i x_i^2)^d \cdot q(x),$$

which is sos as the product (resp. sum) of two sos polynomials is sos.   □

Note that again, for any fixed r, the level r of the hierarchy can be solved using bisection which leads to a sequence of semidefinite programs.

Our developments in the past two sections can be phrased in terms of a Positivstellensatz.

COROLLARY 3.5 (A new Positivstellensatz). *Consider the basic semialgebraic set*

$$S := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, m\}$$

and a polynomial $p := p(x)$. Suppose that S is contained within a ball of radius R. Let η_i and β be any finite upperbounds on $g_i(x)$ and, respectively, $-p(x)$ over the set S.[4] Let d be such that 2d is the smallest integer larger than or equal to the maximum degree of $p, g_i, i = 1, \ldots, m$. Then, $p(x) > 0$ for all $x \in S$ if and only if there exists a positive integer r such that

$$\left(h(x,s,y) - \frac{1}{r} \left(\sum_{i=1}^{n} x_i^2 + \sum_{j=0}^{m+1} s_j^2 + y^2 \right)^{2d} \right) \cdot \left(\sum_{i=1}^{n} x_i^2 + \sum_{j=0}^{m+1} s_j^2 + y^2 \right)^r$$

is a sum of squares, where the form h in variables $(x_1, \ldots, x_n, s_0, \ldots, s_{m+1}, y)$ is as follows:

$$h(x,s,y) := \left(y^{2d} p(x/y) + s_0^2 y^{2d-2} \right)^2 + \sum_{i=1}^{m} \left(y^{2d} g_i(x/y) - s_i^2 y^{2d-2} \right)^2$$

$$+ \left((R + \sum_{i=1}^{m} \eta_i + \beta)^d y^{2d} - (\sum_{i=1}^{n} x_i^2 + \sum_{i=0}^{m} s_i^2)^d - s_{m+1}^{2d} \right)^2.$$

Proof. This is an immediate corollary of arguments given in the proof of Theorem 2.1 and in the proof of Theorem 3.2 for the case where $\gamma = 0$.   □

---

[4] As discussed at the beginning of Section 2, such bounds are very easily computable.

4. Polyá's theorem and hierarchies for POPs that are optimization-free, LP-based, and SOCP-based. In this section, we use a result by Polyá on global positivity of even forms to obtain new hierarchies for polynomial optimization problems. In Section 4.1, we present a hierarchy that is *optimization-free*, in the sense that each level of the hierarchy only requires multiplication of two polynomials and checking if the coefficients of the resulting polynomial are nonnegative. In Section 4.2, we use the previous hierarchy to derive linear programming and second-order cone programming-based (converging) hierarchies that in each level produce a lower bound on the POP whose quality is at least as good as that of the optimization-free hierarchy. These rely on the recently developed concepts of dsos and sdsos polynomials (see Definition 4.7 and [3]), which are alternatives to sos polynomials that have been used in diverse applications to improve scalability; see [3, Section 4].

4.1. An optimization-free hierarchy of lower bounds for POPs. The main theorem in this section presents an optimization-free hierarchy of lower bounds for general POPs with compact feasible sets:

THEOREM 4.1. *Recall the definition of $f_\gamma(z)$ as given in (3), with $z \in \mathbb{R}^N$ and $deg(f_\gamma) = 2D$. Let $(v, w) \in \mathbb{R}^{2n}$ and define*

(9)
$$Pol_{n,2d}^r := \{p \in H_{n,2d} \mid \left(p(v^2 - w^2) + \frac{1}{2r}\left(\sum_{i=1}^n (v_i^4 + w_i^4)\right)^d\right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$

has nonnegative coefficients $\}$.

Consider the hierarchy of optimization problems indexed by r:

(10)
$$l_r := \sup_\gamma \quad \gamma$$
$$s.t. \quad f_\gamma(z) - \frac{1}{r}(\sum_{i=1}^N z_i^2)^D \in Pol_{N,2D}^r.$$

Let $m_r = \max_{i=1,\ldots,r} l_i$. Then $m_r \leq p^*$ for all r, $\{m_r\}$ is nondecreasing, and $\lim_{r \to \infty} m_r = p^*$.

As before, we use bisection to obtain the optimal value l_r of the rth level of the hierarchy up to a fixed precision ϵ (see Remark 3.3). At each step of the bisection algorithm, one simply needs to multiply two polynomials together and check nonnegativity of the coefficients of the resulting polynomial to proceed to the next step. As a consequence, this hierarchy is optimization-free as we do not need to solve (convex) optimization problems at each step of the bisection algorithm. To the best of our knowledge, no other converging hierarchy of lower bounds for general POPs (whose feasible sets are contained within a ball of known radius) dispenses altogether with the need to solve convex subprograms. We also provide a Positivstellensatz counterpart to the hierarchy given above (see Corollary 4.5). This corollary implies in particular that one can always certify infeasibility of a basic semialgebraic set by recursively multiplying polynomials together and simply checking nonnegativity of the coefficients of the resulting polynomial.

We now make a few remarks regarding the techniques used in the proof of Theorem 4.1. Unlike Theorems 3.2 and 3.4, we do not show that $Pol_{n,d}^r$ satisfies properties (a)-(d) as given in Theorem 2.4 due to some technical difficulties. It turns out however that we can avoid showing properties (c) and (d) by using a result by Reznick

and Powers [23] that we present below. Regarding properties (a) and (b), we show that a slightly modified version of (a) holds and that (b), which is the key property in Theorem 2.4, goes through as is. We note though that obtaining (b) from Polyá's result (Theorem 1.6) is not as immediate as obtaining (b) from Artin's and Reznick's results. Indeed, unlike the theorems by Artin and Reznick (see Theorems 1.4 and 1.5) which certify global positivity of *any* form, Polyá's result only certifies global positivity of *even* forms. To make this latter result a statement about general forms, we work in an appropriate lifted space. We make the simple observation that any scalar x can be written as $x = v^2 - w^2$ with $vw = 0$ (take $v = \sqrt{\max\{x,0\}}$ and $w = \sqrt{\max\{-x,0\}}$). We then replace the form $p(z)$ in variables $z \in \mathbb{R}^n$ by the even form $p(v^2 - w^2)$ in variables $(v,w) \in \mathbb{R}^{2n}$. This lifting operation preserves nonnegativity, but unfortunately it does not preserve positivity: even if $p(z)$ is pd, $p(v^2 - w^2)$ always has zeros (e.g., when $v = w$). Hence, though we now have access to an even form, we still cannot use Polyá's property as $p(v^2 - w^2)$ is not positive. This is what leads us to consider the slightly more complicated form $p(v^2 - w^2) + \frac{1}{2r}(\sum_i v_i^4 + w_i^4)^d$ in (9).

THEOREM 4.2 (Powers and Reznick [23]).   Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$, and write $|\alpha| = \alpha_1 + \ldots + \alpha_n$. Denote the standard simplex by Δ_n, i.e., $\Delta_n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | \ x_i \geq 0, i = 1, \ldots, n, \ and \ x_1 + \ldots + x_n = 1\}$. Assume that f is a form of degree $2d$ that is positive on Δ_n and let

$$\lambda = \lambda(f) := \min_{x \in \Delta_n} f(x).$$

Define $c(\alpha) = \frac{(2d)!}{\alpha_1! \ldots \alpha_n!}$. We have:

$$f(x) = \sum_{|\alpha|=2d} a_\alpha x^\alpha = \sum_{|\alpha|=2d} b_\alpha c(\alpha) x^\alpha.$$

Let $||f|| := \max_{|\alpha|=2d} |b_\alpha|$.[5]
    Then, the coefficients of

$$f(x_1, \ldots, x_n) \cdot (x_1 + \ldots + x_n)^{\bar{N}}$$

are nonnegative for $\bar{N} > d(2d-1)\frac{||f||}{\lambda} - 2d$.

Note that here the bound is given in the case where one considers the alternative (but equivalent) formulation of Polyá's Positivstellensatz to the one given in Theorem 1.6, i.e., when one is concerned with positivity of a form over the simplex. The result can easily be adapted to the formulation where one considers global positivity of an even form as shown below.

LEMMA 4.3. Let $p := p(x)$ be an even form of degree $2d$ that is positive definite and let $\beta > 0$ be its minimum on S_x. Define $q(x_1, \ldots, x_n) := p(\sqrt{x_1}, \ldots, \sqrt{x_n})$. Then,

$$p(x_1, \ldots, x_n) \cdot (\sum_i x_i^2)^{\bar{N}}$$

has nonnegative coefficients for $\bar{N} > d(2d-1)\frac{||q||}{\beta} - 2d$.

---

[5] As defined, $||f||$ is a submultiplicative norm; see [29].

Proof. Recall that $q(x_1, \ldots, x_n) = p(\sqrt{x_1}, \ldots, \sqrt{x_n})$. Since $p(x) \geq \beta$ on S_x, then $q(x) \geq \beta$ on Δ_n. Indeed, by contradiction, suppose that there exists $\hat{x} \in \Delta_n$ such that $q(\hat{x}) = \beta - \epsilon$ (where $\epsilon > 0$) and let $y = \sqrt{\hat{x}}$. Note that as $\sum_i \hat{x}_i = 1$, we have $\sum_i y_i^2 = 1$. Furthermore, $p(y) = q(\hat{x}) = \beta - \epsilon$ which contradicts the assumption. Hence, using Theorem 4.2, we have that when $\bar{N} > d(2d-1)\frac{||q||}{\beta} - 2d$,

$$q(x)(\sum_i x_i)^{\bar{N}}$$

has nonnegative coefficients. Hence,

$$q(y^2)(\sum_i y_i^2)^{\bar{N}} = p(y)(\sum_i y_i^2)^{\bar{N}}$$

also has nonnegative coefficients. □

Before we proceed with the proof of Theorem 4.1, we need the following lemma.

LEMMA 4.4. *Let*

$$(11) \quad p_{\gamma,r}(v,w) := f_\gamma(v^2 - w^2) - \frac{1}{r}\left(\sum_{i=1}^N (v_i^2 - w_i^2)^2\right)^D + \frac{1}{2r}\left(\sum_{i=1}^N (v_i^4 + w_i^4)\right)^D,$$

where f_γ is defined as in (3), let $q_{\gamma,r}(v,w) := p_{\gamma,r}(\sqrt{v}, \sqrt{w})$ and let

$$\bar{N}(r) = D(2D-1) \cdot \frac{||q_{\gamma,r}||}{\min_{S_{v,w}} p_{\gamma,r}(v,w)} - 2D.$$

If $f_\gamma(z)$ is positive definite, there exists $\hat{r}$ such that $r^2 \geq \bar{N}(r)$, for all $r \geq \hat{r}$.

Proof. As $f_\gamma(z)$ is positive definite, there exists a positive integer r_0 such that $f_\gamma(z) - \frac{1}{r}(\sum_i z_i^2)^D$ is positive definite for all $r \geq r_0$ and hence

$$(12) \quad f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D$$

is nonnegative for all $r \geq r_0$. Recall now that $||x||_p = (\sum_{i=1}^n x_i^p)^{1/p}$ is a norm for $p \geq 1$ and that

$$||x||_2 \leq n^{1/4}||x||_4.$$

This implies that

$$(\sum_i v_i^4 + \sum_i w_i^4)^D \geq \frac{1}{(2N)^D}(\sum_i v_i^2 + \sum_i w_i^2)^{2D}$$

and hence in view of (12) and the definition of $p_{\gamma,r}$, we have

$$p_{\gamma,r}(v,w) \geq \frac{1}{2^{D+1}N^D r}(\sum_i v_i^2 + \sum_i w_i^2)^{2D}, \forall r \geq r_0.$$

This enables us to conclude that

$$(13) \quad \min_{S_{v,w}} p_{\gamma,r}(v,w) \geq \frac{1}{2^{D+1}N^D r}, \quad \text{for any } r \geq r_0.$$

Further, notice that if we define $\bar{f}_\gamma(v, w) := f_\gamma(v - w)$, then using properties of the norm, we have the following chain of inequalities for any positive integer r:

$$||q_{\gamma,r}|| \leq ||\bar{f}_\gamma|| + \frac{1}{r}||(\sum_i (v_i - w_i)^2)^D|| + \frac{1}{2r}||(\sum_i (v_i^2 + w_i^2))^D||$$

$$\leq ||\bar{f}_\gamma|| + ||(\sum_i (v_i - w_i)^2)^D|| + ||(\sum_i v_i^2 + w_i^2)^D|| =: c_\gamma.$$

As a consequence, combining this with the definition of $\bar{N}(r)$ and (13), we have

$$\bar{N}(r) \leq D(2D - 1)2^{D+1} r N^D c_\gamma, \ \forall r \geq r_0.$$

Now taking $\hat{r} = \max(r_0, \lceil D(2D-1)2^{D+1}N^D c_\gamma \rceil)$, we have $r^2 \geq \bar{N}(r), \forall r \geq \hat{r}$. $\qquad \square$

We now proceed with the proof of Theorem 4.1.

Proof of Theorem 4.1. By definition, the sequence $\{m_r\}$ is nondecreasing. We show that it is upperbounded by p^* by showing that if γ is such that

$$f_\gamma(z) - \frac{1}{r}(\sum_i z_i^2)^D \in Pol_{N,2D}^r,$$

for some r, then f_γ must be positive definite. Then Theorem 2.1 gives us that γ is a strict lower bound on (1). As $p^* > \gamma$ for any such γ, we have that $l_r \leq p^*, \forall r$ and hence $m_r \leq p^*, \forall r$.

Assume that γ is such that

$$f_\gamma(z) - \frac{1}{r}(\sum_i z_i^2)^D \in Pol_{N,2D}^r$$

for some r. By definition of $Pol_{N,2D}^r$ and as $(\sum_i v_i^2 + \sum_i w_i^2)^{r^2}$ is nonnegative, we get that the form

$$f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D + \frac{1}{2r}(\sum_i v_i^4 + w_i^4)^D$$

is nonnegative. This implies that

(14)
$$f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D \geq -\frac{1}{2r} \text{ for } (v, w) \in \{(v, w) \mid \sum_i v_i^4 + \sum_i w_i^4 = 1\},$$

which gives

(15)
$$f_\gamma(z) - \frac{1}{r}(\sum_i z_i^2)^D \geq -\frac{1}{2r}, \ \forall z \in S_z.$$

Indeed, suppose that there exists $\hat{z} \in S_z$ such that (15) does not hold. Then, let $\hat{z}^+ = \max(\hat{z}, 0)$ and $\hat{z}^- = \max(-\hat{z}, 0)$. Note that both $\hat{z}^+$ and $\hat{z}^-$ are nonnegative so we can take $\hat{v} = \sqrt{\hat{z}^+}$ and $\hat{w} = \sqrt{\hat{z}^-}$. We further have that as $\hat{z} \in S_z$ and $\hat{z} = \hat{v}^2 - \hat{w}^2$, $\sum_i \hat{v}_i^4 + \sum_i \hat{w}_i^4 = 1$. Substituting $\hat{z}$ by $\hat{v}^2 - \hat{w}^2$ in (15) then violates (14). Using (15), we conclude that

$$f_\gamma(z) \geq \frac{1}{2r}, \ \forall z \in S_z$$

and that f_γ is positive definite.

We now show that the hierarchy converges, i.e., that $\lim_{r \to \infty} m_r = p^*$. To do this, we show that if γ is a strict lower bound on (1), or equivalently from Theorem 2.1, if $f_\gamma(z)$ is positive definite, then there exists r' such that

$$f_\gamma(z) - \frac{1}{r'}(\sum_i z_i^2)^D \in Pol_{N,2D}^{r'}.$$

Since f_γ is pd, there exists a positive integer r_0 such that $f_\gamma(z) - \frac{1}{r}(\sum_{i=1}^N z_i^2)^D$ is pd for any $r \geq r_0$. This implies that $f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D$ is nonnegative and

$$f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D + \frac{1}{2r}(\sum_i (v_i^4 + w_i^4))^D$$

is positive definite for $r \geq r_0$. Since this form is even, using Lemma 4.3 and the definition of $\bar{N}(r)$ in Lemma 4.4, for any $r \geq r_0$, we have that the polynomial

$$\left(f_\gamma(v^2 - w^2) - \frac{1}{r}(\sum_i (v_i^2 - w_i^2)^2)^D + \frac{1}{2r}(\sum_i (v_i^4 + w_i^4))^D \right) \cdot (\sum_i v_i^2 + \sum_i w_i^2)^{\lceil \bar{N}(r) \rceil}$$

has nonnegative coefficients. From Lemma 4.4, there exists $\hat{r}$ such that $r \geq \hat{r}$ implies $r^2 \geq \bar{N}(r)$. Taking $r' = \max\{r_0, \hat{r}\}$ and considering $p_{\gamma,r'}$ as defined in (11), we get that

$$p_{\gamma,r'}(v,w)(\sum_i v_i^2 + \sum_i w_i^2)^{r'^2}$$
$$= p_{\gamma,r'}(v,w)(\sum_i v_i^2 + \sum_i w_i^2)^{\lceil \bar{N}(r') \rceil} \cdot (\sum_i v_i^2 + \sum_i w_i^2)^{r'^2 - \lceil \bar{N}(r') \rceil}$$

has nonnegative coefficients, which is the desired result. This is because

$$p_{\gamma,r'}(v,w)(\sum_i v_i^2 + \sum_i w_i^2)^{\lceil \bar{N}(r') \rceil}$$

has nonnegative coefficients as $r' \geq r_0$, and

$$(\sum_i v_i^2 + \sum_i w_i^2)^{r'^2 - \lceil \bar{N}(r') \rceil}$$

has nonnegative coefficients as $r' \geq \hat{r}$, and that the product of two polynomials with nonnegative coefficients has nonnegative coefficients. $\square$

COROLLARY 4.5 (An optimization-free Positivstellensatz). *Consider the closed basic semialgebraic set*

$$S := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, m\}$$

and a polynomial $p := p(x)$. Suppose that S is contained within a ball of radius R. Let η_i and β be any finite upperbounds on $g_i(x)$ and, respectively, $-p(x)$ over the set S.[6] Let d be such that $2d$ is the smallest even integer larger than or equal to the maximum

---

[6] Once again, as discussed at the beginning of Section 2, such bounds are very easily computable.

degree of $p, g_i, i = 1, \ldots, m$. Then, $p(x) > 0$ for all $x \in S$ if and only if there exists a positive integer r such that

$$\left(h(v^2 - w^2) - \frac{1}{r} \left(\sum_{i=1}^{n+m+3} (v_i^2 - w_i^2)^2 \right)^d + \frac{1}{2r} \left(\sum_{i=1}^{n+m+3} (v_i^4 + w_i^4) \right)^d \right)$$

$$\cdot \left(\sum_{i=1}^{n+m+3} v_i^2 + \sum_{i=1}^{n+m+3} w_i^2 \right)^{r^2}$$

has nonnegative coefficients, where the form $h := h(z)$ in variables

$$(z_1, \ldots, z_{n+m+3}) := (x_1, \ldots, x_n, s_0, \ldots, s_{m+1}, y)$$

is as follows:

$$h(x, s, y) := \left(y^{2d} p(x/y) + s_0^2 y^{2d-2} \right)^2 + \sum_{i=1}^{m} \left(y^{2d} g_i(x/y) - s_i^2 y^{2d-2} \right)^2$$

$$+ \left((R + \sum_{i=1}^{m} \eta_i + \beta)^d y^{2d} - (\sum_{i=1}^{n} x_i^2 + \sum_{i=0}^{m} s_i^2)^d - s_{m+1}^{2d} \right)^2.$$

Proof. This is an immediate corollary of arguments given in the proof of Theorem 2.1 and in the proof of Theorem 4.1 for the case where $\gamma = 0$. □

4.2. Linear programming and second-order cone programming-based hierarchies for POPs. In this section, we present a linear programming and a second-order cone programming-based hierarchy for general POPs which by construction converge faster than the hierarchy presented in Section 4.1. These hierarchies are based on the recently-introduced concepts of dsos and sdsos polynomials [3] which we briefly revisit below to keep the presentation self-contained.

DEFINITION 4.6. *A symmetric matrix M is said to be*
- *diagonally dominant (dd) if $M_{ii} \geq \sum_{j \neq i} |M_{ij}|$ for all i.*
- *scaled diagonally dominant (sdd) if there exists a diagonal matrix D, with positive diagonal entries, such that DMD is dd.*

We have the following implications as a consequence of Gershgorin's circle theorem:

$$(16) \qquad \qquad \text{M } dd \Rightarrow \text{M } sdd \Rightarrow \text{M } \succeq 0.$$

Requiring M to be dd (resp. sdd) can be encoded via a linear program (resp. a second-order cone program) (see [3] for more details). These notions give rise to the concepts of dsos and sdsos polynomials.

DEFINITION 4.7 ([3]). *Let $z(x) = (x_1^d, x_1^{d-1} x_2, \ldots, x_n^d)^T$ be the vector of monomials in $(x_1, \ldots, x_n)$ of degree d. A form $p \in H_{n,2d}$ is said to be*
- *diagonally-dominant-sum-of-squares (dsos) if it admits a representation*

$$p(x) = z^T(x) Q z(x), \text{ where } Q \text{ is a dd matrix.}$$

- *scaled-diagonally-dominant-sum-of-squares (sdsos) if it admits a representation*

$$p(x) = z^T(x) Q z(x), \text{ where } Q \text{ is a sdd matrix.}$$

The following implications are a consequence of (16):

(17) $\quad\quad\quad\quad p(x)$ dsos $\Rightarrow p(x)$ sdsos $\Rightarrow p(x)$ sos $\Rightarrow p(x)$ nonnegative.

Given the fact that our Gram matrices and polynomials are related to each other via linear equalities, it should be clear that optimizing over the set of dsos (resp. sdsos) polynomials is an LP (resp. SOCP).

We now present our LP and SOCP-based hierarchies for POPs.

COROLLARY 4.8. *Recall the definition of $f_\gamma(z)$ as given in (3), with $z \in \mathbb{R}^N$ and $deg(f_\gamma) = 2D$, and let $p_{\gamma,r}$ be as in (11). Consider the hierarchy of optimization problems indexed by r:*

$$
\begin{aligned}
l_r := \sup_{\gamma,q} \quad & \gamma \\
\text{s.t.} \quad & p_{\gamma,r}(v,w) \cdot q(v,w) \ \text{is } s/dsos \\
& q(v,w) \ \text{is } s/dsos \ \text{and of degree } 2r^2.
\end{aligned}
$$
(18)

Let $m_r = \max_{i=1,\dots,r} l_i$. Then, $m_r \leq p^*$ for all r, $\{m_r\}$ is nondecreasing, and we have $\lim_{r \to \infty} m_r = p^*$.

Proof. This is an immediate consequence of the fact that any even form $p \in H_{n,2d}$ with nonnegative coefficients can be written as $p(x) = z(x)^T Q z(x)$ where Q is diagonal and has nonnegative (diagonal) entries. As such a Q is dd (and also sdd), we conclude that p is dsos (and also sdsos). The corollary then follows from Theorem 4.1. $\quad\square$

Note that similarly to our previous hierarchies, one must proceed by bisection on γ to solve the level r of the hierarchy. At each step of the hierarchy, we solve a linear program (resp. second-order cone program) that searches for the coefficients of q that make q dsos (resp. sdsos) and $p_{\gamma,r} \cdot q$ dsos (resp. sdsos).

There is a trade-off between the hierarchies developed in this subsection and the one developed in the previous subsection: the hierarchy of Section 4.1 is optimization-free whereas those of Section 4.2 use linear or second-order cone programming. Hence the former hierarchy is faster to run at each step. However, the latter hierarchies could potentially take fewer levels to converge. This is similar to the trade-off observed between the hierarchies presented in Theorems 3.2 and 3.4.

5. Future research directions and open problems. To conclude, we present some interesting directions for future research and some open problems.

A first research direction of interest relates to the computational performance of the methods described in this paper. As they stand, we do not believe that our hierarchies are as practically efficient as other hierarchies (e.g., those due to Lasserre and Parrilo). This is mainly because all our hierarchies require the use of bisection and increase the number of variables (from n to $n+m+3$) and the degree (from $2d$ to $4d$) of the polynomials involved. Though this blow-up is linear, one can expect that it could be problematic in practice. Out of our hierarchies, the optimization-free one may be tempting to consider for very large-scale problems as each step of it only requires polynomial multiplication. However, the caveat that one should keep in mind is that the level r required to obtain a good-quality lower bound to the POP would probably be very large. To counter-balance this, one would have to invest considerable effort in a proper implementation of the arithmetic involved in each level of the hierarchy. Such an implementation should involve among other things: (i) automation of the computation of the coefficients of f_γ, (ii) exploiting the very specific

structure of the polynomial in (9) (with p replaced by $f_\gamma(z) - \frac{1}{r}(\sum_{i=1}^{N} z_i^2)^D$) to compute its coefficients explicitly, (iii) efficient multiplication of two multivariate polynomials (e.g., by evaluating the product on random samples and obtaining its coefficients by solving a linear system, or by pursuing the ideas in [12]), and (iv) thinking about ways of parallelizing all these computations. We are unsure as to whether the optimization-free hierarchy would provide useful bounds on relevant examples—even after careful implementation—but it may be an interesting research direction to investigate.

Another interesting research direction lies in the comparison of our hierarchies to existing ones. One way to do this would be to determine the convergence rates of the hierarchies presented in this paper and compare them with those of other hierarchies. The main difficulty in obtaining such rates is the absence of a quantitative version of Theorem 2.1, which would link the quality of the lower bound on the optimal value of the POP (i.e., the value of the gap $p^* - \gamma$) to the minimum of f_γ on the unit sphere in z-space. If such a result were available, one could have hopes, e.g., to obtain convergence rates for our optimization-free hierarchy by applying Theorem 4.2 and Lemma 4.3 to f_γ. (A similar approach was undertaken by de Klerk, Laurent, and Parrilo in [9] to provide convergence rates for an analogous hierarchy for the problem of minimizing a polynomial on the simplex.) Likewise, one could hope to relate Lasserre's hierarchy (for example) to the Reznick or Artin hierarchies presented in this paper and derive convergence rates for them by using previously known results on the complexity of the Lasserre hierarchy [18]. As it stands, we are unable to show that if level r of Lasserre's hierarchy certifies positivity of $p(x) - \gamma$ on the feasible set, then so does a level r' of our Reznick or Artin hierarchies. This is due to the fact that, from a sum of squares certificate of positivity of $p(x) - \gamma$, one does not a priori know how to pick r' in such a way that $f_\gamma(z) - \frac{1}{r'}(\sum_i z_i^2)^{2d}$ becomes positive definite (let alone admits a particular certificate of positivity). This question would again be answered if we had a quantitative version of Theorem 2.1.

Finally, we present two more concrete open problems spawned by the writing of this paper. The first one concerns the assumptions needed to construct our hierarchies.

Open problem 1. Theorems 2.1 and 2.4 require that the feasible set S of the POP given in (1) be contained in a ball of radius R. Can these theorems be extended to the case where there is no compactness assumption on S?

The second open problem is linked to the Artin and Reznick cones presented in Definition 3.1.

Open problem 2. As mentioned before, Reznick cones $R_{n,2d}^r$ are convex for all r. We are unable to prove however that Artin cones $A_{n,2d}^r$ are convex (even though they satisfy properties (a)-(d) of Theorem 2.4 like Reznick cones do). Are Artin cones convex for all r? We know that they are convex for $r = 0$ and for r large enough as they give respectively the sos and psd cones (see [17] for the latter claim). However, we do not know the answer already for $r = 1$. Our preliminary attempts to find forms that would disprove convexity of these cones have failed so far.

REFERENCES

[1] A. A. Ahmadi, *On the difficulty of deciding asymptotic stability of cubic homogeneous vector fields*, in Proceedings of the American Control Conference, 2012.

[2] A. A. Ahmadi and A. Majumdar, *Some applications of polynomial optimization in operations research and real-time decision making*, Optimization Letters, 10 (2016), pp. 709–729.

[3] A. A. Ahmadi and A. Majumdar, *DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization*, (2017). Available on arXiv at https://arxiv.org/abs/1706.02586.

[4] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, in Abhandlungen aus dem mathematischen Seminar der Universität Hamburg, vol. 5, Springer, 1927, pp. 100–115.

[5] G. Averkov, *Constructive proofs of some Positivstellensätze for compact semialgebraic subsets of $\mathbb{R}^d$*, Journal of Optimization Theory and Applications, 158 (2013), pp. 410–418.

[6] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*, SIAM, 2012.

[7] R. S. Datta, *Computing with polynomials: a personal odyssey.* https://people.mbi.ohio-state.edu/datta.53/cs282paper.pdf, 2002.

[8] E. de Klerk, M. Laurent, and P. Parrilo, *On the equivalence of algebraic approaches to the minimization of forms on the simplex*, Positive Polynomials in Control, (2005), pp. 580–580.

[9] E. De Klerk, M. Laurent, and P. A. Parrilo, *A PTAS for the minimization of polynomials of fixed degree over the simplex*, Theoretical Computer Science, 361 (2006), pp. 210–225.

[10] E. de Klerk and D. V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM Journal on Optimization, 12 (2002), pp. 875–892.

[11] W. Habicht, *Über die Zerlegung strikte definiter Formen in Quadrate*, Commentarii Mathematici Helvetici, 12 (1939), pp. 317–322.

[12] S. C. Johnson, *Sparse polynomial arithmetic*, ACM SIGSAM Bulletin, 8 (1974), pp. 63–71.

[13] J.-L. Krivine, *Anneaux préordonnés*, J. d'analyse mathématique, 12 (1964), pp. 307–326.

[14] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization, 11 (2001), pp. 796–817.

[15] J. B. Lasserre, *Moments, Positive Polynomials and their Applications*, World Sci., 2009.

[16] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging applications of algebraic geometry, Springer, 2009, pp. 157–270.

[17] H. Lombardi, D. Perrucci, and M.-F. Roy, *An elementary recursive bound for effective Positivstellensatz and Hilbert 17-th problem*, arXiv preprint arXiv:1404.2338, (2014).

[18] J. Nie and M. Schweighofer, *On the complexity of Putinar's Positivstellensatz*, Journal of Complexity, 23 (2007), pp. 135–150.

[19] P. M. Pardalos and S. A. Vavasis, *Quadratic programming with one negative eigenvalue is NP-hard*, Journal of Global Optimization, 1 (1991), pp. 15–22.

[20] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, PhD thesis, California Institute of Technology, May 2000.

[21] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Mathematical Programming, 96 (2003), pp. 293–320.

[22] G. Pólya, *Über positive Darstellung von Polynomen*, Vierteljschr. Naturforsch. Ges. Zürich, 73 (1928), pp. 141–145.

[23] V. Powers and B. Reznick, *A new bound for Pólya's theorem with applications to polynomials positive on polyhedra*, Journal of pure and applied algebra, 164 (2001), pp. 221–229.

[24] A. Prestel and C. Delzell, *Positive Polynomials: from Hilbert's 17th Problem to Real Algebra*, Springer Science & Business Media, 2013.

[25] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal, 42 (1993), pp. 969–984.

[26] B. Reznick, *Uniform denominators in Hilbert's 17th problem*, Math Z., 220 (1995), pp. 75–97.

[27] K. Schmüdgen, *The k-moment problem for compact semi-algebraic sets*, Mathematische Annalen, 289 (1991), pp. 203–206.

[28] M. Schweighofer, *An algorithmic approach to Schmüdgen's Positivstellensatz*, Journal of Pure and Applied Algebra, 166 (2002), pp. 307–319.

[29] M. Schweighofer, *On the complexity of Schmüdgen's Positivstellensatz*, Journal of Complexity, 20 (2004), pp. 529–543.

[30] M. Schweighofer, *Optimization of polynomials on compact semialgebraic sets*, SIAM Journal on Optimization, 15 (2005), pp. 805–825.

[31] G. Stengle, *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry*, Mathematische Annalen, 207 (1974), pp. 87–97.