On the Equivalence of Algebraic Conditions for Convexity and Quasiconvexity of Polynomials

Amir Ali Ahmadi and Pablo A. Parrilo

***Abstract*— This paper is concerned with algebraic relaxations, based on the concept of sum of squares decomposition, that give sufficient conditions for convexity of polynomials and can be checked efficiently with semidefinite programming. We propose three natural sum of squares relaxations for polynomial convexity based on respectively, the definition of convexity, the first order characterization of convexity, and the second order characterization of convexity. The main contribution of the paper is to show that all three formulations are equivalent; (though the condition based on the second order characterization leads to semidefinite programs that can be solved much more efficiently). This equivalence result serves as a direct algebraic analogue of a classical result in convex analysis.**

We also discuss recent related work in the control literature that introduces different sum of squares relaxations for polynomial convexity. We show that the relaxations proposed in our paper will always do at least as well the ones introduced in that work, with significantly less computational effort. Finally, we show that contrary to a claim made in the same related work, if an even degree polynomial is homogeneous, then it is quasiconvex if and only if it is convex. An example is given.

I. Introduction

Algebraic relaxations for polynomial optimization problems have seen a lot of attention from the control community in the past decade. The primary reason for this interest is the fact that several fundamental problems in controls, such as Lyapunov analysis and synthesis of nonlinear or hybrid systems with polynomial or rational dynamics, can often be reduced to the problem of deciding whether certain polynomials are nonnegative either globally or on some regions of the space. This problem is unfortunately known to be computationally intractable (already NP-hard for polynomials of degree four [1]). On the other hand, algebraic conditions based on the decomposition of polynomials as sum of squares (sos) of polynomials have shown to provide very effective sufficient conditions for polynomial nonnegativity. The success of sums of squares techniques stems from several facts such as their computational tractability for reasonable size problems, capability of being incorporated in search problems, explicitness of certificates of nonnegativity, and availability of a hierarchical procedure that often converges to necessary conditions for polynomial nonnegativity. For a range of applications of sum of squares programs in systems and control, we refer the reader to the by no means exhaustive list of works [2], [3], [4], [5], [6], [7], [8], [9], [10], and references therein.

Amir Ali Ahmadi and Pablo A. Parrilo are with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. E-mail: a_a_a@mit.edu, parrilo@mit.edu.

Aside from nonnegativity, convexity is another important property of polynomials that one would like to establish. Here again one might want to *decide* if a given polynomial is convex, or *parameterize* a family of polynomials with certain properties. For example, if we are trying to find a global minimum of a given polynomial, it could be very useful to decide first if the polynomial is convex. If convexity of the polynomial is *a priori* guaranteed, then very basic techniques (e.g. gradient descent) can find a global minimum, a task that is in general intractable in absence of convexity. As another example, if we can certify that a homogenous polynomial is convex, then we can use it to define a norm (see Theorem 5.1 in Section V), which may be useful in many applications. To name an instance where parametrization of convex polynomials may come up, we can mention the problem of finding a convex polynomial that closely approximates the convex envelope of a more complicated non-convex function. Or in systems analysis, we could be searching for a convex polynomial Lyapunov function whose convex sublevel sets contain relevant information about the trajectories of the system.

The question of determining the computational complexity of deciding convexity of polynomials appeared in 1992 on a list of seven open problems in complexity theory for numerical optimization [11]. Very recently, in joint work with A. Olshevsky and J. N. Tsitsiklis, we have shown that the problem is NP-hard for polynomials of degree four or larger [12]. If testing membership in the set of convex polynomials is hard, searching or optimizing over them is obviously also hard. This result, like any other hardness result, stresses the need for good relaxation algorithms that can deal with many instances of the problem efficiently.

The focus of this work is on algebraic sum of squares based relaxations for establishing convexity of polynomials. The motivation for doing this is that all the attractive features of sum of squares programming carry over to this problem immediately. In fact, the idea of using sum of squares to establish convexity of polynomials has already appeared in several places in the literature in different contexts. In [13], Helton and Nie formally introduce the term *sos-convexity*, which is based on an appropriately defined sos decomposition of the Hessian matrix of polynomials; (see Definition 2 in Section II-B.) They use sos-convexity in their work to give sufficient conditions for semidefinite representability of semialgebraic sets. In [14], Lasserre uses sos-convexity to extend Jensen's inequality in probability theory to linear functionals that are not necessarily probability measures. Predating the paper of Helton and Nie, Magnani, Lall, and Boyd

[15] use the same relaxation to find convex polynomials that best fit a set of data points or to find minimum volume convex sets, given by sublevel sets of convex polynomials, that contain a set of points in space.

Another work in the control literature that tackles exactly this problem and to which we will make several references in this paper is the work of Chesi and Hung [16]. Rather than working with the Hessian matrix, Chesi and Hung use an sos relaxation on an equivalent formulation of the definition of convexity.

Our main contribution in this paper is to revisit the following basic questions: what are some natural ways to use the sos relaxation for formulating sufficient conditions for convexity of polynomials? How do the different formulations relate? We will show in Section III that three most natural sos relaxations that are based on (i) the definition of convexity, (ii) the first order characterization of convexity, and (iii) the second order characterization of convexity all turn out to be exactly equivalent (Theorem 3.1). We will comment on how the computational efficiency of solving the resulting SDPs compare. In Section IV, we show that any of these three equivalent formulations always performs at least as well as the formulations proposed in [16] (Theorem 4.1), at a considerably lower computational cost.

Another problem that we will touch upon in this paper is that of establishing quasiconvexity (convexity of sublevel sets) of polynomials. Even though every convex function is quasiconvex, the converse is not true in general. In [16] and [17], the authors argue for several applications of establishing quasiconvexity of Lyapunov functions and develop some sos relaxations for parameterizing quasiconvex polynomials. However, it is claimed (incorrectly) in these works that quasiconvexity is a strictly weaker condition than convexity even for positive homogenous polynomials. In fact, separate conditions for quasiconvexity of homogenous polynomials are proposed in [17]. We will give a simple proof in Section V that shows that a positive homogenous polynomial is quasiconvex if and only if it is convex. We will reaffirm this point by revisiting an example of [16] in Section VI.

The organization of our paper is as follows. In Section II we give preliminary material on nonnegativity, sum of squares, convexity, and sos-convexity of polynomials and the connection to semidefinite programming. In Section III, we present our main result that establishes the equivalence of three natural algebraic relaxations for convexity of polynomials. We compare these relaxations with those given in [16] in Section IV. The result on equivalence of convexity and quasiconvexity for homogenous polynomials is presented in Section V. An example is given in Section VI. Finally, our conclusions are discussed in Section VII.

II. PRELIMINARIES

A. *Nonnegativity, sum of squares, and semidefinite programming*

We denote by $\mathbb{K}[x] := \mathbb{K}[x_1, \ldots, x_n]$ the ring of polynomials in n variables with coefficients in the field $\mathbb{K}$. A polynomial $p(x) \in \mathbb{R}[x]$ is said to be nonnegative or positive semidefinite (psd) if $p(x) \geq 0$ for all $x \in \mathbb{R}^n$. Clearly, a necessary condition for a polynomial to be psd or to be globally convex is for its total degree to be even. So, for the rest of the paper we will always be dealing with polynomials of even degree. We say that $p(x)$ is a sum of squares (sos), if there exist polynomials $q_1(x), \ldots, q_m(x)$ such that $p(x) = \sum_{i=1}^{m} q_i^2(x)$. It is clear that $p(x)$ being sos implies that $p(x)$ is psd. Hilbert [18] showed that the converse is not true except for specific degrees and dimensions; see [19].

A polynomial where all the monomials have the same degree is called a *form*. A form $p(x)$ of degree d is a homogenous function of degree d (since it satisfies $p(\lambda x) = \lambda^d p(x)$). It is easy to show that if a form of degree d is sos, then d is even and the polynomials q_i in the sos decomposition are forms of degree $d/2$.

For a fixed dimension and degree the set of psd polynomials and the set of sos polynomials are both closed convex cones. The fact that the sos cone is closed is not so obvious and was first proven by Robinson [20]. His result will be key to our main theorem.

Unlike nonnegativity, the problem of deciding whether a given polynomial admits an sos decomposition turns out to be a tractable problem. The reason is that the problem can be reformulated as a semidefinite program (SDP) (perhaps better known to the control community as a linear matrix inequality (LMI)) [21]. The following theorem illustrates this.

Theorem 2.1: ([2], [22]) A multivariate polynomial $p(x)$ in n variables and of even degree d is a sum of squares if and only if there exists a positive semidefinite matrix Q (often called the Gram matrix) such that

$$p(x) = z^T Q z, \tag{1}$$

where z is the vector of monomials of degree up to $d/2$

$$z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^{d/2}]. \tag{2}$$

The notions of positive semidefiniteness and sum of squares of scalar polynomials can be naturally extended to polynomial matrices, i.e., matrices with entries in $\mathbb{R}[x]$. We say that a symmetric polynomial matrix $P(x) \in \mathbb{R}[x]^{m \times m}$ is PSD if $P(x)$ is PSD for all $x \in \mathbb{R}^n$. It is straightforward to see that this condition holds if and only if the polynomial $y^T H(x) y$ in $m + n$ variables $[x; y]$ is psd. The definition of an sos-matrix is as follows [23], [24], [25].

Definition 1: A symmetric polynomial matrix $P(x) \in \mathbb{R}[x]^{m \times m}$, $x \in \mathbb{R}^n$ is an sos-matrix if there exists a polynomial matrix $M(x) \in \mathbb{R}[x]^{s \times m}$ for some $s \in \mathbb{N}$, such that $P(x) = M^T(x) M(x)$.

It turns out that a polynomial matrix $P(x) \in \mathbb{R}[x]^{m \times m}$, $x \in \mathbb{R}^n$ is an sos-matrix if and only if the scalar polynomial $y^T P(x) y$ is a sum of squares in $\mathbb{R}[x; y]$; see [23]. This is a useful fact because it enables us to check whether a polynomial matrix is an sos-matrix with the machinery of Theorem 2.1. Once again, it is obvious that being an sos-matrix is a sufficient condition for a polynomial matrix to be PSD.

B. Convexity and sos-convexity

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain is a convex set and for all x and y in the domain and all $\lambda \in [0,1]$, we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y). \tag{3}$$

Our focus in this paper is on establishing convexity of polynomials on their entire domain $\mathbb{R}^n$. The following theorem states that for continuous functions (and in particular for polynomials) it is enough to check condition (3) only for $\lambda = 1/2$ to establish convexity [26, p. 71].

Theorem 2.2: A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if

$$f\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right) \leq \tfrac{1}{2}f(x) + \tfrac{1}{2}f(y) \tag{4}$$

for all x and y in the domain.
The choice of $\lambda = \frac{1}{2}$ in this theorem is arbitrary and any other fixed $\lambda \in (0,1)$ would work as well.

Next, we recall a classical result from convex analysis. The proof can be found in many convex analysis textbooks, e.g. [27, p. 70]. The theorem is true for any twice differentiable function, but for our purposes we state it for polynomials.

Theorem 2.3: Let $p : \mathbb{R}^n \to \mathbb{R}$ be a multivariate polynomial. Let $\nabla p(x)$ denote its gradient and let $\nabla^2 p(x)$ be its Hessian, i.e., the $n \times n$ symmetric matrix of second derivatives. Then the following are equivalent.

(a) p is convex.
(b) $p(y) \geq p(x) + \nabla p(x)^T (y-x), \quad \forall x, y \in \mathbb{R}^n$.
(c) $y^T \nabla^2 p(x) y$ is psd; (i.e., $\nabla^2 p(x)$ is a PSD matrix).

Helton and Nie proposed in [13] the notion of sos-convexity as an sos relaxation for the second order characterization of convexity (condition **(c)** above).

Definition 2: ([13]) A polynomial $p(x)$ is sos-convex if its Hessian $\nabla^2 p(x)$ is an sos-matrix.

With what we have discussed so far, it should be clear that sos-convexity is a sufficient condition for convexity of polynomials that can be checked with semidefinite programming. In the next section we show some other natural sos relaxations for polynomial convexity, which will turn out to be equivalent to sos-convexity.

III. EQUIVALENT ALGEBRAIC RELAXATIONS FOR CONVEXITY OF POLYNOMIALS

An obvious way to formulate alternative sos relaxations for convexity of polynomials is to replace every inequality in Theorem 2.3 with its sos version. In this section we examine how these relaxations relate in terms of conservatism. We also comment on the size of the resulting semidefinite programs.

Our results below can be thought of as an algebraic analogue of Theorem 2.3.

Theorem 3.1: Let $p(x)$ be a polynomial of degree d in n variables with its gradient and Hessian denoted respectively by $\nabla p(x)$ and $\nabla^2 p(x)$. Let g_λ, g_∇, and g_{∇^2} be defined as

$$\begin{aligned} g_\lambda(x,y) &= (1-\lambda)p(x) + \lambda p(y) - p((1-\lambda)x + \lambda y), \\ g_\nabla(x,y) &= p(y) - p(x) - \nabla p(x)^T (y-x), \\ g_{\nabla^2}(x,y) &= y^T \nabla^2 p(x) y. \end{aligned} \tag{5}$$

Then the following are equivalent:
(a) $g_{\frac{1}{2}}(x,y)$ is sos[1].
(b) $g_\nabla(x,y)$ is sos.
(c) $g_{\nabla^2}(x,y)$ is sos; (i.e., $\nabla^2 p(x)$ is an sos-matrix).

Proof: **(a)⇒(b)**: Assume $g_{\frac{1}{2}}(x,y)$ is sos. We start by proving that $g_{\frac{1}{2^k}}(x,y)$ will also be sos for any integer $k \geq 2$. A little bit of straightforward algebra yields the relation

$$g_{\frac{1}{2^k}}(x,y) = \tfrac{1}{2} g_{\frac{1}{2^{k-1}}}(x,y) + g_{\frac{1}{2}}\left(x, \tfrac{2^{k-1}-1}{2^{k-1}}x + \tfrac{1}{2^{k-1}}y\right). \tag{6}$$

The second term on the right hand side of (6) is always sos because $g_{\frac{1}{2}}(x,y)$ is sos. When $k=2$, both terms on the right hand side of this equation are sos by assumption and hence (6) implies that $g_{\frac{1}{4}}(x,y)$ must be sos. Once this is established, we can reapply equation (6) with $k=3$ to conclude that $g_{\frac{1}{8}}(x,y)$ is sos, and so on.

Now, let us rewrite $g_\lambda(x,y)$ as

$$g_\lambda(x,y) = p(x) + \lambda(p(y) - p(x)) - p(x + \lambda(y-x)).$$

We have

$$\frac{g_\lambda(x,y)}{\lambda} = p(y) - p(x) - \frac{p(x+\lambda(y-x)) - p(x)}{\lambda}. \tag{7}$$

Next, we take the limit of both sides of (7) by letting $\lambda = \frac{1}{2^k} \to 0$ as $k \to \infty$. Because p is differentiable, the right hand side of (7) will converge to $g_\nabla(x,y)$. On the other hand, our preceding argument implies that $\frac{g_\lambda(x,y)}{\lambda}$ is an sos polynomial (of degree d in $2n$ variables) for any $\lambda = \frac{1}{2^k}$. Moreover, as λ goes to zero, the coefficients of $\frac{g_\lambda(x,y)}{\lambda}$ remain bounded since the limit of this sequence is $g_\nabla(x,y)$, which must have bounded coefficients (see (5)). By closedness of the sos cone, we conclude that the limit $g_\nabla(x,y)$ must be sos.

(b)⇒(a): Assume $g_\nabla(x,y)$ is sos. It is easy to check that

$$g_{\frac{1}{2}}(x,y) = \tfrac{1}{2} g_\nabla\left(\tfrac{1}{2}x + \tfrac{1}{2}y, x\right) + \tfrac{1}{2} g_\nabla\left(\tfrac{1}{2}x + \tfrac{1}{2}y, y\right),$$

and hence $g_{\frac{1}{2}}(x,y)$ is sos.

(b)⇒(c): Let us write the second order Taylor approximation of p around x:

$$\begin{aligned} p(y) = \; & p(x) + \nabla^T p(x)(y-x) \\ & + \tfrac{1}{2}(y-x)^T \nabla^2 p(x)(y-x) + o(||y-x||^2). \end{aligned}$$

After rearranging terms, letting $y = x + \varepsilon z$ (for $\varepsilon > 0$), and dividing both sides by ε^2 we get:

$$\begin{aligned} (p(x+\varepsilon z) - p(x))/\varepsilon^2 - \nabla^T p(x) z/\varepsilon = \; & \tfrac{1}{2} z^T \nabla^2 p(x) z \\ & + 1/\varepsilon^2 o(\varepsilon^2 ||z||^2). \end{aligned} \tag{8}$$

[1] The constant $\frac{1}{2}$ in $g_{\frac{1}{2}}(x,y)$ of condition **(a)** is arbitrary and is chosen for convenience. One can show that $g_{\frac{1}{2}}$ being sos implies that g_λ is sos for any fixed $\lambda \in [0,1]$. Conversely, if g_λ is sos for some $\lambda \in (0,1)$, then $g_{\frac{1}{2}}$ is sos. The proofs are similar to the proof of **(a)⇒(b)**.

The left hand side of (8) is $g_\nabla(x, x+\varepsilon z)/\varepsilon^2$ and therefore for any fixed $\varepsilon > 0$, it is an sos polynomial by assumption. As we take $\varepsilon \to 0$, by closedness of the sos cone, the left hand side of (8) converges to an sos polynomial. On the other hand, as the limit is taken, the term $\frac{1}{\varepsilon^2} o(\varepsilon^2 ||z||^2)$ vanishes and hence we have that $z^T \nabla^2 p(x) z$ must be sos.

(c)⇒(b): Following the strategy of the proof of the classical case in [28, p. 165], we start by writing the Taylor expansion of p around x with the integral form of the remainder:

$$\begin{aligned} p(y) = \quad & p(x) + \nabla^T p(x)(y-x) \\ & + \int_0^1 (1-t)(y-x)^T \nabla^2 p(x + t(y-x))(y-x)dt. \end{aligned} \tag{9}$$

Since $y^T \nabla^2 p(x) y$ is sos by assumption, for any $t \in [0,1]$ the integrand

$$(1-t)(y-x)^T \nabla^2 p(x + t(y-x))(y-x)$$

is an sos polynomial of degree d in x and y. From (9) we have

$$g_\nabla = \int_0^1 (1-t)(y-x)^T \nabla^2 p(x + t(y-x))(y-x)dt.$$

It then follows that g_∇ is sos because integrals of sos polynomials, if they exist, are sos. ■

We conclude that conditions **(a)**, **(b)**, and **(c)** are equivalent sufficient conditions for convexity of polynomials, and can each be turned into a set of LMIs using Theorem 2.1. It is easy to see that all three polynomials $g_{\frac{1}{2}}(x,y)$, $g_\nabla(x,y)$, and $g_{\nabla^2}(x,y)$ are polynomials in $2n$ variables and of degree d. (Note that each differentiation reduces the degree by one.) Each of these polynomials have a specific structure that can be exploited for formulating smaller SDPs. For example, the symmetries $g_{\frac{1}{2}}(x,y) = g_{\frac{1}{2}}(y,x)$ and $g_{\nabla^2}(x,-y) = g_{\nabla^2}(x,y)$ can be taken advantage of via symmetry reduction techniques developed in [24].

The issue of symmetry reduction aside, we would like to point out that formulation **(c)** (which was the original definition of sos-convexity) can be significantly more efficient than the other two conditions. The reason is that the polynomial $g_{\nabla^2}(x,y)$ is always quadratic and homogeneous in y and of degree $d-2$ in x. This makes $g_{\nabla^2}(x,y)$ much more sparse than $g_\nabla(x,y)$ and $g_{\nabla^2}(x,y)$, which have degree d both in x and in y. Furthermore, because of the special bipartite structure of $y^T \nabla^2 p(x) y$, only monomials of the form $x_i^k y_j$ will appear in the vector of monomials (2) of Theorem 2.1. This in turn reduces the size of the Gram matrix, and hence the size of the SDP. It is perhaps not too surprising that the characterization of convexity based on the Hessian matrix is a more efficient condition to check. After all, this is a local condition (curvature at every point in every direction must be nonnegative), whereas conditions **(a)** and **(b)** are both global.

We shall end this section by commenting on the necessity of these conditions. In [29], [30], the authors showed with an explicit example that there exist convex polynomials that are not sos-convex. In view of Theorems 2.3 and 3.1, it follows immediately that the same example rejects necessity of conditions **(a)** and **(b)**. However, standard Positivstellensatz results related to Hilbert's 17th problem can be used to make these conditions necessary under mild assumptions but at the cost of solving larger and larger SDPs; see e.g. [31].

IV. COMPARISON WITH THE FORMULATION OF CHESI AND HUNG

In [16], Chesi and Hung have developed a sum of squares based sufficient condition for polynomial convexity that is based on a reformulation of the definition of convexity in (3). Below, we restate their result and then compare it to the sufficient conditions of the previous section.

Let $p(x)$ be a polynomial of degree d in n variables. Let $z = [x, y]^T \in \mathbb{R}^{2n}$ and let

$$w(z,\alpha) = p(\alpha_1 x + \alpha_2 y) - \alpha_1 p(x) - \alpha_2 p(y),$$

where $\alpha = [\alpha_1, \alpha_2]^T \in \mathbb{R}^2$ is an auxiliary vector. Rewrite $w(z,\alpha)$ as

$$w(z,\alpha) = \sum_{i=1}^{d} \sum_{j=0}^{i} \alpha_1^{i-j} \alpha_2^j w_{i,j}(z),$$

where $w_{i,j}(z)$ are suitable polynomials. Define

$$t(z,\alpha) = \sum_{i=1}^{d} (\alpha_1 + \alpha_2)^{d-i} \sum_{j=0}^{i} \alpha_1^{i-j} \alpha_2^j w_{i,j}(z). \tag{10}$$

Observe that $t(z,\alpha) = w(z,\alpha)$ for all α such that $\alpha_1 + \alpha_2 = 1$. The result of Chesi and Hung (Theorem 1 in [16]) states that $p(x)$ is convex if and only if

$$t(z,\alpha^2) \leq 0 \quad \forall (z,\alpha) \in \mathbb{R}^{2n} \times \mathbb{R}^2,$$

where $\alpha^2 \triangleq [\alpha_1^2, \alpha_2^2]^T$. The sos relaxation that they propose is then

$$-t(z,\alpha^2) \quad \text{sos}. \tag{11}$$

The following easy theorem shows that this formulation is at least as conservative as the formulations of Theorem 3.1.

Theorem 4.1: For a polynomial $p(x)$, define the polynomial $t(z,\alpha)$ as in (10). If $-t(z,\alpha^2)$ is sos, then all of the conditions **(a)**, **(b)**, and **(c)** in Theorem 3.1 hold.

Proof: Let $\alpha = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T$. Because $-t(z,\alpha^2)$ is sos, it follows that

$$\begin{aligned} -t(z,[\tfrac{1}{2},\tfrac{1}{2}]^T) &= -\textstyle\sum_{i=1}^{d} (\frac{1}{2}+\frac{1}{2})^{d-i} \sum_{j=0}^{i} \frac{1}{2}^{i-j} \frac{1}{2}^{j} w_{i,j}(z) \\ &= -w(z,[\tfrac{1}{2},\tfrac{1}{2}]^T) \\ &= \tfrac{1}{2}p(x) + \tfrac{1}{2}p(y) - p(\tfrac{1}{2}x + \tfrac{1}{2}y) \end{aligned}$$

is sos. Hence $g_{\frac{1}{2}}(x,y)$ is sos. ■

The polynomial $-t(z,\alpha^2)$ in (11) has $2n+2$ variables $(x, y, \alpha_1, \alpha_2)$. It has a total degree of $3d$ (degree d in x and y and degree $2d$ in α). Checking whether this polynomial is sos is drastically more expensive than checking our preferred formulation based on the Hessian matrix. For example, when $p(x)$ is a homogenous polynomial in $n = 2$ variables of degree $d = 4$, checking condition (11) reduces to an LMI with a 50×50 Gram matrix and 960 scalar parameters [16], whereas the LMI resulting from setting $y^T \nabla^2 p(x) y$ to be sos has a 4×4 Gram matrix and only a single scalar parameter.

In [16], the authors realize the computational burden involved with their formulation and propose a further relaxation of their result; see Theorem 2 in [16]. We do not discuss their second result because their examples suggest that the new formulation is strictly more conservative than the first one. Moreover, for the example of the previous paragraph, their further relaxed formulation still reduces to three LMIs, each with a 10×10 Gram matrix and 20 scalar parameters [16].

V. QUASICONVEXITY OF HOMOGENOUS POLYNOMIALS

A function $p(x)$ is said to be *quasiconvex* if all its sublevel sets

$$\mathcal{S}_p(\alpha) := \{x|\ p(x) \leq \alpha\}, \tag{12}$$

for $\alpha \in \mathbb{R}$, are convex sets. It is easy to see that every convex function is quasiconvex but the converse is not true. In this section we show that the converse is true for positive continuous functions that are homogenous. For our purposes, we state the theorem for polynomials. This result is contrary to the developments in [16] and [17]; see also the example in the next section.

Note that for homogeneous polynomials quasiconvexity is completely determined by convexity of the unit sublevel set $\mathcal{S}_p(1)$, which we will simply denote by $\mathcal{S}_p$.

Theorem 5.1: Let $p(x)$ be a positive[2] and homogeneous polynomial of even degree d. If $p(x)$ is quasiconvex, then it must also be convex.

Proof: By assumption, the sublevel set

$$\mathcal{S}_p = \{x|\ p(x) \leq 1\}$$

is convex. Let $q(x) := p(x)^{1/d}$. We have

$$\begin{array}{rcl} \mathcal{S}_q & = & \{x|\ q(x) \leq 1\} \\ & = & \{x|\ p(x)^{1/d} \leq 1\} \\ & = & \mathcal{S}_p. \end{array}$$

So, $q(x)$ has a convex unit sublevel set. We also have that

$$q(\lambda x) = p(\lambda x)^{1/d} = (\lambda^d p(x))^{1/d} = |\lambda| q(x),$$

for any $\lambda \in \mathbb{R}$, and $q(x) > 0$ for all $x \neq 0$. Therefore $q(x)$ is a positive homogenous function of degree 1 that is quasiconvex. We will show that $q(x)$ must be convex.

Because $q(x)$ is continuous, by Theorem 2.2 it suffices to show

$$q\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right) \leq \tfrac{1}{2}q(x) + \tfrac{1}{2}q(y),$$

which because of homogeneity of degree 1 reduces to

$$q(x+y) \leq q(x) + q(y). \tag{13}$$

Pick any nonzero $x, y \in \mathbb{R}^n$. We have $q(\frac{x}{q(x)}) = q(\frac{y}{q(y)}) = 1$. Therefore $\frac{x}{q(x)} \in \mathcal{S}_q$ and $\frac{y}{q(y)} \in \mathcal{S}_q$. By convexity of $\mathcal{S}_q$ we have

$$q\left(\frac{q(x)}{q(x)+q(y)}\frac{x}{q(x)} + \frac{q(y)}{q(x)+q(y)}\frac{y}{q(y)}\right) \leq 1.$$

Therefore

$$\frac{1}{q(x)+q(y)} q(x+y) \leq 1,$$

and hence (13).[3] Convexity of $p(x) = q(x)^d$ then follows from the fact that convex nonnegative functions raised to a power larger than one remain convex [27, p. 86]. ■

Theorem 5.1 suggests that for homogenous polynomials, no separate conditions are needed for quasiconvexity. In particular, the algebraic relaxations for convexity that we presented in Section III can be used to efficiently test quasiconvexity.[4]

VI. AN EXAMPLE

We revisit an example of Chesi and Hung [16, Example 2].

Example 6.1: Consider the homogenous polynomial[5]

$$p(x) = 5000x_1^4 + 4x_1^2x_2^2 - cx_1x_2^3 + x_2^4. \tag{14}$$

Our first task is to find the largest value of c, denoted by c_*, for which the polynomial is convex.

In [16], a lower bound of 3.271 on c_* is claimed using the sos program (11) discussed in Section IV. We will give a lower bound based on sos-convexity. The Hessian of $p(x)$ is given by

$$\nabla^2 p(x) = \begin{bmatrix} 60000x_1^2 + 8x_2^2 & 16x_1x_2 - 3cx_2^2 \\ 16x_1x_2 - 3cx_2^2 & 8x_1^2 + 12x_2^2 - 6cx_1x_2 \end{bmatrix}.$$

Requiring $\nabla^2 p(x)$ to be an sos-matrix reduces to the following SDP:

$$\begin{array}{ll} \max & c \\ \text{s.t.} & y^T \nabla^2 p(x) y = z^T Q z \\ & Q \succeq 0, \end{array} \tag{15}$$

where $y = (y_1, y_2)^T$, and z is the vector of monomials

$$z = [x_1y_1, x_1y_2, x_2y_1, x_2y_2]^T.$$

The decision variables of this SDP are the scalar c and one free parameter in the symmetric 4×4 Gram matrix Q. Let $\hat{c}_*$ denote the optimal value of SDP (15). We have used SOSTOOLS [32] and the SDP solver SeDuMi [33] to get a numerical value of $\hat{c}_* = 3.2658$ to four significant digits.

Our lower bound is in fact tight for this example.[6] Because the example is simple enough, we can compute the exact value of c_* to be

$$c_* = \sqrt{\frac{179984}{16875}} \approx 3.2658. \tag{16}$$

This is the cutoff value for the determinant of $\nabla^2 p(x)$ to remain nonnegative, and also the exact optimal value of (15).

We now turn to the problem of finding the largest value of c, denoted by $c_\sharp$, for which $p(x)$ has convex sublevel sets.

[2] With little more work, the positivity assumption can be dropped; see [12].

[3] In fact it can be shown that $q(x)$ is a gauge (or Minkowski) norm defined by the symmetric convex set $\mathcal{S}_q$; i.e., $q(x) = \inf\{t > 0|\ x \in t\mathcal{S}_q\}$.

[4] As for quasiconvexity of non-homogeneous polynomials, the conditions of [17] and [16] can also be simplified. See midpoint convexity [27, p. 60]. For quasiconvexity of odd degree polynomials, see [12].

[5] There is a typo in the exponent of the monomial $x_1x_2^3$ in [16]. We have confirmed this with the authors.

[6] More generally, we can show that convexity and sos-convexity are equivalent for bivariate forms.

In [16], a lower bound of 3.785 on $c_\sharp$ is asserted and it is claimed that even though the polynomial is not convex at $c = 3.785$, the sos relaxations proposed in [16] prove convexity of the sublevel sets. In view of our Theorem 5.1, this claim has to be wrong. Because the polynomial $p(x)$ is positive for $c = c_*$, Theorem 5.1 implies that the cutoff of convexity and quasiconvexity must be the same; i.e., we must have

$$c_\sharp = c_* = \sqrt{\frac{179984}{16875}}.$$

Let us confirm this directly with a curvature argument. To reject quasiconvexity, it is enough to show a point x and a direction y such that

$$y^T \nabla p(x) = 0 \quad \text{and} \quad y^T \nabla^2 p(x) y < 0. \qquad (17)$$

See [27, p. 101]. Let

$$\bar{x} = (33\sqrt{31} + \sqrt{33747}, 300)^T,$$

and let

$$\bar{y} = \left(-\frac{\partial p}{\partial x_2}(\bar{x}), \frac{\partial p}{\partial x_1}(\bar{x}) \right)^T$$

be the direction tangent to the level set at $\bar{x}$. Then around $c_\sharp$, the polynomial $\bar{y}^T \nabla^2 p(\bar{x}) \bar{y}$ as a function of c has a Taylor expansion

$$\bar{y}^T \nabla^2 p(\bar{x}) \bar{y} = k_1 (c - c_\sharp) + \cdots ,$$

with $k_1 < 0$. Therefore, for $c = c_\sharp + \varepsilon$ and ε positive and arbitrarily small, $\bar{x}$ and $\bar{y}$ satisfy (17).

VII. CONCLUSIONS

We proposed three natural approaches for using sum of squares relaxations to impose polynomial convexity. Our main result, which was an algebraic analogue of a classical result in convex analysis, implied that all these relaxations are equivalent. We also clarified the connection to previous works in the literature and argued that a relaxation for positive semidefiniteness of the Hessian matrix leads to the most computationally efficient SDPs. Finally, we showed that for positive homogenous polynomials, convexity of sublevel sets is equivalent to convexity of the function.

REFERENCES

[1] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

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

[3] D. Henrion and A. Garulli, editors. *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Information Sciences*. Springer, 2005.

[4] Special issue on positive polynomials in control. *IEEE Trans. Automat. Control*, 54(5), 2009.

[5] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. *Homogeneous polynomial forms for robustness analysis of uncertain systems*. Number 390 in Lecture Notes in Control and Information Sciences. Springer, 2009.

[6] Z. Jarvis-Wloszek, R. Feeley, W. Tan, K. Sun, and A. Packard. Some controls applications of sum of squares programming. In *Proceedings of the 42th IEEE Conference on Decision and Control*, 2003.

[7] A. A. Ahmadi and P. A. Parrilo. Non-monotonic Lyapunov functions for stability of discrete time nonlinear and switched systems. In *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008.

[8] A. A. Ahmadi. Non-monotonic Lyapunov functions for stability of nonlinear and switched systems: theory and computation. Master's Thesis, Massachusetts Institute of Technology, June 2008. Available from http://dspace.mit.edu/handle/1721.1/44206.

[9] S. Prajna and A. Papachristodoulou. Analysis of switched and hybrid systems – beyond piecewise quadratic methods. In *Proceedings of the American Control Conference*, 2003.

[10] S. Prajna, P. A. Parrilo, and A. Rantzer. Nonlinear control synthesis by convex optimization. *IEEE Trans. Automat. Control*, 49(2):310–314, 2004.

[11] P. M. Pardalos and S. A. Vavasis. Open questions in complexity theory for numerical optimization. *Mathematical Programming*, 57(2):337–339, 1992.

[12] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. MIT LIDS technical report, 2010.

[13] J. W. Helton and J. Nie. Semidefinite representation of convex sets. *Mathematical Programming*, 122(1, Ser. A):21–64, 2010.

[14] J. B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization*, 19(4):1995–2014, 2008.

[15] A. Magnani, S. Lall, and S. Boyd. Tractable fitting with convex polynomials via sum of squares. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[16] G. Chesi and Y. S. Hung. Establishing convexity of polynomial Lyapunov functions and their sublevel sets. *IEEE Trans. Automat. Control*, 53(10):2431–2436, 2008.

[17] G. Chesi and Y. S. Hung. On the convexity of sublevel sets of polynomial and homogeneous polynomial Lyapunov functions. In *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006.

[18] D. Hilbert. Über die Darstellung Definiter Formen als Summe von Formenquadraten. *Math. Ann.*, 32, 1888.

[19] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In *Contemporary Mathematics*, volume 253, pages 251–272. American Mathematical Society, 2000.

[20] R. M. Robinson. Some definite polynomials which are not sums of squares of real polynomials. In *Selected questions of algebra and logic (collection dedicated to the memory of A. I. Mal′cev) (Russian)*, pages 264–282. Izdat. "Nauka" Sibirsk. Otdel., Novosibirsk, 1973.

[21] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[22] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2, Ser. B):293–320, 2003.

[23] M. Kojima. Sums of squares relaxations of polynomial semidefinite programs. *Research report B-397, Dept. of Mathematical and Computing Sciences. Tokyo Institute of Technology*, 2003.

[24] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192:95–128, 2004.

[25] C. W. Scherer and C. W. J. Hol. Matrix sum of squares relaxations for robust semidefinite programs. *Mathematical Programming*, 107:189–211, 2006.

[26] W. Rudin. *Real and complex analysis*. McGraw-Hill series in higher mathematics, 1987. Third edition.

[27] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[28] A. L. Tits. Lecture notes on optimal control. Available from http://www.isr.umd.edu/˜andre/664.pdf, 2008.

[29] A. A. Ahmadi and P. A. Parrilo. A convex polynomial that is not sos-convex. MIT LIDS technical report 2810. arXiv:0903.1287., 2009.

[30] A. A. Ahmadi and P. A. Parrilo. A positive definite polynomial Hessian that does not factor. In *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009.

[31] B. Reznick. Uniform denominators in Hilbert's 17th problem. *Math Z.*, 220(1):75–97, 1995.

[32] S. Prajna, A. Papachristodoulou, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, 2002-05. Available from http://www.cds.caltech.edu/sostools and http://www.mit.edu/˜parrilo/sostools.

[33] J. Sturm. *SeDuMi version 1.05*, October 2001. Latest version available at http://sedumi.ie.lehigh.edu/.