

A Complete Characterization of the Gap between Convexity and SOS-Convexity

Amir Ali Ahmadi and Pablo A. Parrilo ^{*†}

Abstract

Our first contribution in this paper is to prove that three natural sum of squares (sos) based sufficient conditions for convexity of polynomials via the definition of convexity, its first order characterization, and its second order characterization are equivalent. These three equivalent algebraic conditions, henceforth referred to as sos-convexity, can be checked by semidefinite programming whereas deciding convexity is NP-hard. If we denote the set of convex and sos-convex polynomials in n variables of degree d with $\tilde{C}_{n,d}$ and $\tilde{\Sigma}C_{n,d}$ respectively, then our main contribution is to prove that $\tilde{C}_{n,d} = \tilde{\Sigma}C_{n,d}$ if and only if $n = 1$ or $d = 2$ or $(n, d) = (2, 4)$. We also present a complete characterization for forms (homogeneous polynomials) except for the case $(n, d) = (3, 4)$ which is joint work with G. Blekherman and is to be published elsewhere. Our result states that the set $C_{n,d}$ of convex forms in n variables of degree d equals the set $\Sigma C_{n,d}$ of sos-convex forms if and only if $n = 2$ or $d = 2$ or $(n, d) = (3, 4)$. To prove these results, we present in particular explicit examples of polynomials in $\tilde{C}_{2,6} \setminus \tilde{\Sigma}C_{2,6}$ and $\tilde{C}_{3,4} \setminus \tilde{\Sigma}C_{3,4}$ and forms in $C_{3,6} \setminus \Sigma C_{3,6}$ and $C_{4,4} \setminus \Sigma C_{4,4}$, and a general procedure for constructing forms in $C_{n,d+2} \setminus \Sigma C_{n,d+2}$ from nonnegative but not sos forms in n variables and degree d . Although for disparate reasons, the remarkable outcome is that convex polynomials (resp. forms) are sos-convex exactly in cases where nonnegative polynomials (resp. forms) are sums of squares, as characterized by Hilbert.

1 Introduction

1.1 Nonnegativity and sum of squares

One of the cornerstones of real algebraic geometry is Hilbert's seminal paper in 1888 [22], where he gives a complete characterization of the degrees and dimensions in which nonnegative polynomials can be written as sums of squares of polynomials. In particular, Hilbert proves in [22] that there exist nonnegative polynomials that are not sums of squares, although explicit examples of such polynomials appeared only about 80 years later and the study of the gap between nonnegative and sums of squares polynomials continues to be an active area of research to this day.

Motivated by a wealth of new applications and a modern viewpoint that emphasizes efficient computation, there has also been a great deal of recent interest from the optimization community in the representation of nonnegative polynomials as sums of squares (sos). Indeed, many fundamental problems in applied and computational mathematics can be reformulated as either deciding whether certain polynomials are nonnegative or searching over a family of nonnegative polynomials. It is

^{*}The authors are with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. Email: {a.a.a, parrilo}@mit.edu.

[†]This research was partially supported by the NSF Focused Research Group Grant on Semidefinite Optimization and Convex Algebraic Geometry DMS-0757207.

well-known however that if the degree of the polynomial is four or larger, deciding nonnegativity is an NP-hard problem. (This follows e.g. as an immediate corollary of NP-hardness of deciding matrix copositivity [30].) On the other hand, it is also well-known that deciding whether a polynomial can be written as a sum of squares can be reduced to solving a semidefinite program, for which efficient algorithms e.g. based on interior point methods is available. The general machinery of the so-called “sos relaxation” has therefore been to replace the intractable nonnegativity requirements with the more tractable sum of squares requirements that obviously provide a sufficient condition for polynomial nonnegativity.

Some relatively recent applications that sum of squares relaxations have found span areas as diverse as control theory [32], [21], quantum computation [17], polynomial games [34], combinatorial optimization [19], and many others.

1.2 Convexity and sos-convexity

Aside from nonnegativity, *convexity* is another fundamental property of polynomials that is of both theoretical and practical significance. Perhaps most notably, presence of convexity in an optimization problem often leads to tractability of finding global optimal solutions. Consider for example the problem of finding the unconstrained global minimum of a polynomial. This is an NP-hard problem in general [35], but if we know a priori that the polynomial to be minimized is convex, then every local minima is global, and even simple gradient descent methods can find a global minimum. There are other scenarios where one would like to decide convexity of polynomials. For example, it turns out that the d -th root of a degree d polynomial is a norm if and only if the polynomial is homogeneous, positive definite, and convex [42]. Therefore, if we can certify that a homogenous polynomial is convex and definite, then we can use it to define a norm, which is useful in many applications. In many other practical settings, we might want to *parameterize* a family of convex polynomials that have certain properties, e.g., that serve as a convex envelope for a non-convex function, approximate a more complicated function, or fit some data points with minimum error. In the field of robust control for example, it is common to desire a convex Lyapunov function to prove stability of an uncertain dynamical system described as a system of difference inclusions. Therefore, the ability to efficiently search over convex polynomials would lead to algorithmic ways of constructing such Lyapunov functions.

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 [31]. In a recent joint work with A. Olshevsky and J. N. Tsitsiklis, we have shown that the problem is strongly NP-hard even for polynomials of degree four [2]. 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 approximation algorithms that can deal with many instances of the problem efficiently.

The focus of this work is on an algebraic notion known as *sos-convexity* (introduced formally by Helton and Nie in [20]), which is a sufficient condition for convexity of polynomials based on an appropriately defined sum of squares decomposition of the Hessian matrix; see Definition 2.4. As we will briefly review in Section 2, the problem of deciding if a given polynomial is sos-convex amounts to solving a single semidefinite program.

Besides its computational implications, sos-convexity is an appealing concept since it bridges the geometric and algebraic aspects of convexity. Indeed, while the usual definition of convexity is concerned only with the geometry of the epigraph, in sos-convexity this geometric property (or the nonnegativity of the Hessian) must be certified through a “simple” algebraic identity, namely the sum of squares factorization of the Hessian. The original motivation of Helton and Nie for defining

sos-convexity was in relation to the question of semidefinite representability of convex sets [20]. But this notion has already appeared in the literature in a number of other settings [24], [25], [28], [11]. In particular, there has been much recent interest in the role of convexity in semialgebraic geometry [24], [7], [15], [26] and sos-convexity is a recurrent figure in this line of research.

1.3 Contributions and organization of the paper

The main contribution of this work is to establish the counterpart of Hilbert's characterization of the gap between nonnegativity and sum of squares for the notions of convexity and sos-convexity. We start by presenting some background material in Section 2. In Section 3, we prove an algebraic analogue of a classical result in convex analysis, which provides three equivalent characterizations for sos-convexity (Theorem 3.1). This result substantiates the fact that sos-convexity is *the* right sos relaxation for convexity. In Section 4, we present some examples of convex polynomials that are not sos-convex. In Section 5, we provide the characterization of the gap between convexity and sos-convexity (Theorem 5.1 and Theorem 5.2). Subsection 5.1 includes the proofs of the cases where convexity and sos-convexity are equivalent and Subsection 5.2 includes the proofs of the cases where they are not. In particular, Theorem 5.8 and Theorem 5.9 present explicit examples of convex but not sos-convex polynomials that have dimension and degree as low as possible, and Theorem 5.10 provides a general construction for producing such polynomials in higher degrees. Some concluding remarks and an open problem are presented in Section 6.

2 Preliminaries

2.1 Background on nonnegativity and sum of squares

A (multivariate) *polynomial* $p := p(x)$ in variables $x := (x_1, \dots, x_n)^T$ is a function from \mathbb{R}^n to \mathbb{R} that is a finite linear combination of monomials:

$$p(x) = \sum_{\alpha} c_{\alpha} x^{\alpha} = \sum_{\alpha_1, \dots, \alpha_n} c_{\alpha_1, \dots, \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

where the sum is over n -tuples of nonnegative integers α_i . We will be concerned throughout with polynomials with real coefficients, i.e., we will have $c_{\alpha} \in \mathbb{R}$. The ring of polynomials in n variables with real coefficients is denoted by $\mathbb{R}[x]$. The *degree* of a monomial x^{α} is equal to $\alpha_1 + \cdots + \alpha_n$. The degree of a polynomial $p \in \mathbb{R}[x]$ is defined to be the highest degree of its component monomials. A polynomial p 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 is for its degree to be even. We say that p is a *sum of squares (sos)*, if there exist polynomials q_1, \dots, q_m such that $p = \sum_{i=1}^m q_i^2$. We denote the set of psd (resp. sos) polynomials in n variables and degree d by $\tilde{P}_{n,d}$ (resp. $\tilde{\Sigma}_{n,d}$). Any sos polynomial is clearly psd, so we have $\tilde{\Sigma}_{n,d} \subseteq \tilde{P}_{n,d}$.

A *homogeneous polynomial* (or a *form*) is a polynomial where all the monomials have the same degree. A form p of degree d is a homogeneous function of degree d since it satisfies $p(\lambda x) = \lambda^d p(x)$ for any scalar $\lambda \in \mathbb{R}$. We say that a form p is *positive definite* if $p(x) > 0$ for all $x \neq 0$ in \mathbb{R}^n . Following standard notation, we denote the set of psd (resp. sos) homogeneous polynomials in n variables and degree d by $P_{n,d}$ (resp. $\Sigma_{n,d}$). Once again, we have the obvious inclusion $\Sigma_{n,d} \subseteq P_{n,d}$. All of the four sets $\Sigma_{n,d}, P_{n,d}, \tilde{\Sigma}_{n,d}, \tilde{P}_{n,d}$ are closed convex cones. The closedness of the sum of squares cone may not be so obvious. This fact was first proved by Robinson [43]. We will make crucial use of it in the proof of Theorem 3.1 in the next section.

Any form of degree d in n variables can be “dehomogenized” into a polynomial of degree $\leq d$ in $n - 1$ variables by setting $x_n = 1$. Conversely, any polynomial p of degree d in n variables can be “homogenized” into a form p_h of degree d in $n + 1$ variables, by adding a new variable y , and letting $p_h(x_1, \dots, x_n, y) := y^d p(x_1/y, \dots, x_n/y)$. The properties of being psd and sos are preserved under homogenization and dehomogenization [40].

A very natural and fundamental question that as we mentioned earlier was answered by Hilbert is to understand in what dimensions and degrees nonnegative polynomials (or forms) can be represented as sums of squares, i.e, for what values of n and d we have $\tilde{\Sigma}_{n,d} = \tilde{P}_{n,d}$ or $\Sigma_{n,d} = P_{n,d}$. Note that because of the argument in the last paragraph, we have $\tilde{\Sigma}_{n,d} = \tilde{P}_{n,d}$ if and only if $\Sigma_{n+1,d} = P_{n+1,d}$. Hence, it is enough to answer the question just for polynomials or just for forms and the answer to the other one comes for free.

Theorem 2.1 (Hilbert, [22]). $\tilde{\Sigma}_{n,d} = \tilde{P}_{n,d}$ if and only if $n = 1$ or $d = 2$ or $(n, d) = (2, 4)$. Equivalently, $\Sigma_{n,d} = P_{n,d}$ if and only if $n = 2$ or $d = 2$ or $(n, d) = (3, 4)$.

The proofs of $\tilde{\Sigma}_{1,d} = \tilde{P}_{1,d}$ and $\tilde{\Sigma}_{n,2} = \tilde{P}_{n,2}$ are relatively simple and were known before Hilbert. On the other hand, the proof of the fairly surprising fact that $\tilde{\Sigma}_{2,4} = \tilde{P}_{2,4}$ (or equivalently $\Sigma_{3,4} = P_{3,4}$) is more involved. We refer the interested reader to [37], [36], [13], and references in [40] for some modern expositions and alternative proofs of this result. Hilbert’s other main contribution was to show that these are the only cases where nonnegativity and sum of squares are equivalent by giving a nonconstructive proof of existence of polynomials in $\tilde{P}_{2,6} \setminus \tilde{\Sigma}_{2,6}$ and $\tilde{P}_{3,4} \setminus \tilde{\Sigma}_{3,4}$ (or equivalently forms in $P_{3,6} \setminus \Sigma_{3,6}$ and $P_{4,4} \setminus \Sigma_{4,4}$). From this, it follows with simple arguments that in all higher dimensions and degrees there must also be psd but not sos polynomials; see [40]. Explicit examples of such polynomials appeared in the 1960s starting from the celebrated Motzkin form [29]:

$$M(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_3^2 + x_3^6, \quad (1)$$

which belongs to $P_{3,6} \setminus \Sigma_{3,6}$, and continuing a few years later with the Robinson form [43]:

$$R(x_1, x_2, x_3, x_4) = x_1^2(x_1 - x_4)^2 + x_2^2(x_2 - x_4)^2 + x_3^2(x_3 - x_4)^2 + 2x_1x_2x_3(x_1 + x_2 + x_3 - 2x_4), \quad (2)$$

which belongs to $P_{4,4} \setminus \Sigma_{4,4}$.

Several other constructions of psd polynomials that are not sos have appeared in the literature since. An excellent survey is [40]. See also [41] and [8].

2.2 Connection to semidefinite programming and matrix generalizations

As we remarked before, what makes sum of squares an appealing concept from a computational viewpoint is its relation to semidefinite programming. It is well-known (see e.g. [32], [33]) that a polynomial p 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

$$p(x) = z^T Q z,$$

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

$$z = [1, x_1, x_2, \dots, x_n, x_1x_2, \dots, x_n^{d/2}]. \quad (3)$$

The set of all such matrices Q is the feasible set of a semidefinite program (SDP). For fixed d , the size of this semidefinite program is polynomial in n . Semidefinite programs can be solved with arbitrary accuracy in polynomial time. There are several implementations of semidefinite

programming solvers, based on interior point algorithms among others, that are very efficient in practice and widely used; see [49] and references therein.

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 $U(x) \in \mathbb{R}[x]^{m \times m}$ is positive semidefinite if $U(x)$ is positive semidefinite in the matrix sense for all $x \in \mathbb{R}^n$, i.e. if $U(x)$ has nonnegative eigenvalues for all $x \in \mathbb{R}^n$. It is straightforward to see that this condition holds if and only if the scalar polynomial $y^T U(x) y$ in $m+n$ variables $[x; y]$ is psd. A homogeneous polynomial matrix $U(x)$ is said to be positive definite, if it is positive definite in the matrix sense, i.e., has positive eigenvalues, for all $x \neq 0$ in \mathbb{R}^n . The definition of an sos-matrix is as follows [23], [18], [45].

Definition 2.2. *A symmetric polynomial matrix $U(x) \in \mathbb{R}[x]^{m \times m}$, $x \in \mathbb{R}^n$, is an sos-matrix if there exists a polynomial matrix $V(x) \in \mathbb{R}[x]^{s \times m}$ for some $s \in \mathbb{N}$, such that $P(x) = V^T(x)V(x)$.*

It turns out that a polynomial matrix $U(x) \in \mathbb{R}[x]^{m \times m}$, $x \in \mathbb{R}^n$, is an sos-matrix if and only if the scalar polynomial $y^T U(x) y$ is a sum of squares in $\mathbb{R}[x; y]$; see [23]. This is a useful fact because in particular it gives us an easy way of checking whether a polynomial matrix is an sos-matrix by solving a semidefinite program. Once again, it is obvious that being an sos-matrix is a sufficient condition for a polynomial matrix to be positive semidefinite.

2.3 Background on convexity and sos-convexity

A polynomial p is (globally) convex if for all x and y in \mathbb{R}^n and all $\lambda \in [0, 1]$, we have

$$p(\lambda x + (1 - \lambda)y) \leq \lambda p(x) + (1 - \lambda)p(y). \quad (4)$$

Since polynomials are continuous functions, the inequality in (4) holds if and only if it holds for a fixed value of $\lambda \in (0, 1)$, say, $\lambda = \frac{1}{2}$. In other words, p is convex if and only if

$$p\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}p(x) + \frac{1}{2}p(y) \quad (5)$$

for all x and y ; see e.g. [44, p. 71]. Except for the trivial case of linear polynomials, an odd degree polynomial is clearly never convex.

For the sake of direct comparison with a result that we derive in the next section (Theorem 3.1), we recall next a classical result from convex analysis on the first and second order characterization of convexity. The proof can be found in many convex optimization textbooks, e.g. [10, p. 70]. The theorem is of course true for any twice differentiable function, but for our purposes we state it for polynomials.

Theorem 2.3. *Let $p := p(x)$ be a polynomial. Let $\nabla p := \nabla p(x)$ denote its gradient and let $H := H(x)$ be its Hessian, i.e., the $n \times n$ symmetric matrix of second derivatives. Then the following are equivalent.*

- (a) $p\left(\frac{1}{2}x + \frac{1}{2}y\right) \leq \frac{1}{2}p(x) + \frac{1}{2}p(y)$, $\forall x, y \in \mathbb{R}^n$; (i.e., p is convex).
- (b) $p(y) \geq p(x) + \nabla p(x)^T(y - x)$, $\forall x, y \in \mathbb{R}^n$.
- (c) $y^T H(x) y \geq 0$, $\forall x, y \in \mathbb{R}^n$; (i.e., $H(x)$ is a positive semidefinite polynomial matrix).

Helton and Nie proposed in [20] the notion of *sos-convexity* as an sos relaxation for the second order characterization of convexity (condition (c) above).

Definition 2.4. *A polynomial p is sos-convex if its Hessian $H := H(x)$ 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 will show some other natural sos relaxations for polynomial convexity, which will turn out to be equivalent to sos-convexity.

We end this section by introducing some final notation: $\tilde{C}_{n,d}$ and $\tilde{\Sigma}C_{n,d}$ will respectively denote the set of convex and sos-convex polynomials in n variables and degree d ; $C_{n,d}$ and $\Sigma C_{n,d}$ will respectively denote set of convex and sos-convex homogeneous polynomials in n variables and degree d . Again, these four sets are closed convex cones and we have the obvious inclusions $\tilde{\Sigma}C_{n,d} \subseteq \tilde{C}_{n,d}$ and $\Sigma C_{n,d} \subseteq C_{n,d}$.

3 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 to each other. We also comment on the size of the resulting semidefinite programs.

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

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

$$\begin{aligned} g_\lambda(x, y) &= (1 - \lambda)p(x) + \lambda p(y) - p((1 - \lambda)x + \lambda y), \\ g_\nabla(x, y) &= p(y) - p(x) - \nabla p(x)^T(y - x), \\ g_{\nabla^2}(x, y) &= y^T H(x)y. \end{aligned} \tag{6}$$

Then the following are equivalent:

- (a) $g_{\frac{1}{2}}(x, y)$ is sos¹.
- (b) $g_\nabla(x, y)$ is sos.
- (c) $g_{\nabla^2}(x, y)$ is sos; (i.e., $H(x)$ is an sos-matrix).

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

$$g_{\frac{1}{2^{k+1}}}(x, y) = \frac{1}{2}g_{\frac{1}{2^k}}(x, y) + g_{\frac{1}{2}}\left(x, \frac{2^k-1}{2^k}x + \frac{1}{2^k}y\right). \tag{7}$$

The second term on the right hand side of (7) is always sos because $g_{\frac{1}{2}}$ is sos. Hence, this relation shows that for any k , if $g_{\frac{1}{2^k}}$ is sos, then so is $g_{\frac{1}{2^{k+1}}}$. Since for $k = 1$, both terms on the right hand side of (7) are sos by assumption, induction immediately gives that $g_{\frac{1}{2^k}}$ is sos for all k .

Now, let us rewrite g_λ as

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

We have

$$\frac{g_\lambda(x, y)}{\lambda} = p(y) - p(x) - \frac{p(x + \lambda(y - x)) - p(x)}{\lambda}. \tag{8}$$

Next, we take the limit of both sides of (8) by letting $\lambda = \frac{1}{2^k} \rightarrow 0$ as $k \rightarrow \infty$. Because p is differentiable, the right hand side of (8) will converge to g_∇ . On the other hand, our preceding

¹The constant $\frac{1}{2}$ in $g_{\frac{1}{2}}(x, y)$ of condition (a) is arbitrary and 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) \Rightarrow (b).

argument implies that $\frac{g_\lambda}{\lambda}$ 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 $\frac{g_\lambda}{\lambda}$ remain bounded since the limit of this sequence is g_∇ , which must have bounded coefficients (see (6)). By closedness of the sos cone, we conclude that the limit g_∇ must be sos.

(b) \Rightarrow (a): Assume g_∇ is sos. It is easy to check that

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

and hence $g_{\frac{1}{2}}$ is sos.

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

$$p(y) = p(x) + \nabla^T p(x)(y - x) + \frac{1}{2}(y - x)^T H(x)(y - x) + o(\|y - x\|^2).$$

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

$$(p(x + \epsilon z) - p(x))/\epsilon^2 - \nabla^T p(x)z/\epsilon = \frac{1}{2}z^T H(x)z + 1/\epsilon^2 o(\epsilon^2 \|z\|^2). \quad (9)$$

The left hand side of (9) is $g_\nabla(x, x + \epsilon z)/\epsilon^2$ and therefore for any fixed $\epsilon > 0$, it is an sos polynomial by assumption. As we take $\epsilon \rightarrow 0$, by closedness of the sos cone, the left hand side of (9) converges to an sos polynomial. On the other hand, as the limit is taken, the term $\frac{1}{\epsilon^2} o(\epsilon^2 \|z\|^2)$ vanishes and hence we have that $z^T H(x)z$ must be sos.

(c) \Rightarrow (b): Following the strategy of the proof of the classical case in [48, p. 165], we start by writing the Taylor expansion of p around x with the integral form of the remainder:

$$p(y) = p(x) + \nabla^T p(x)(y - x) + \int_0^1 (1 - t)(y - x)^T H(x + t(y - x))(y - x)dt. \quad (10)$$

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

$$(1 - t)(y - x)^T H(x + t(y - x))(y - x)$$

is an sos polynomial of degree d in x and y . From (10) we have

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

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 checked with a semidefinite program as explained in Subsection 2.2. 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 [18].

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 $y^T H(x)y$, only monomials of the form $x_i^k y_j$ will appear in the vector of monomials (3). 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.

Remark 3.1. There has been yet another proposal for an sos relaxation for convexity of polynomials in [11]. However, we have shown in [4] that the condition in [11] is at least as conservative as the three conditions in Theorem 3.1 and also significantly more expensive to check.

Remark 3.2. Just like convexity, the property of sos-convexity is preserved under restrictions to affine subspaces. This is perhaps most directly seen through characterization **(a)** of sos-convexity in Theorem 3.1, by also noting that sum of squares is preserved under restrictions. Unlike convexity however, if a polynomial is sos-convex on every line (or even on every proper affine subspace), this does not imply that the polynomial is sos-convex.

As an application of Theorem 3.1, we use our new characterization of sos-convexity to give a short proof of an interesting lemma of Helton and Nie.

Lemma 3.2. (Helton and Nie [20, Lemma 8]). *Every sos-convex form is sos.*

Proof. Let p be an sos-convex form of degree d . We know from Theorem 3.1 that sos-convexity of p is equivalent to the polynomial $g_{\frac{1}{2}}(x, y) = \frac{1}{2}p(x) + \frac{1}{2}p(y) - p(\frac{1}{2}x + \frac{1}{2}y)$ being sos. But since sos is preserved under restrictions and $p(0) = 0$, this implies that

$$g_{\frac{1}{2}}(x, 0) = \frac{1}{2}p(x) - p(\frac{1}{2}x) = (\frac{1}{2} - (\frac{1}{2})^d) p(x)$$

is sos. □

Note that the same argument also shows that convex forms are psd.

4 Some constructions of convex but not sos-convex polynomials

It is natural to ask whether sos-convexity is not only a sufficient condition for convexity of polynomials but also a necessary one. In other words, could it be the case that if the Hessian of a polynomial is positive semidefinite, then it must factor? To give a negative answer to this question, one has to prove existence of a convex polynomial that is not sos-convex, i.e, a polynomial p for which one (and hence all) of the three polynomials $g_{\frac{1}{2}}, g_{\nabla}$, and g_{∇^2} in (6) are psd but not sos. Note that existence of psd but not sos polynomials does not imply existence of convex but not sos-convex polynomials on its own. The reason is that the polynomials $g_{\frac{1}{2}}, g_{\nabla}$, and g_{∇^2} all possess a very special structure.² For example, $y^T H(x)y$ has the structure of being quadratic in y and a Hessian in x . (Not every polynomial matrix is a valid Hessian.) The Motzkin or the Robinson polynomials in (1) and (2) for example are clearly not of this structure.

²There are many situations where requiring a specific structure on polynomials makes psd equivalent to sos. As an example, we know that there are forms in $P_{4,4} \setminus \Sigma_{4,4}$. However, if we require the forms to have only even monomials, then all such nonnegative forms in 4 variables and degree 4 are sums of squares [16].

In an earlier paper, we presented the first example of a convex but not sos-convex polynomial [5],[3]³:

$$\begin{aligned}
p(x_1, x_2, x_3) = & 32x_1^8 + 118x_1^6x_2^2 + 40x_1^6x_3^2 + 25x_1^4x_2^4 - 43x_1^4x_2^2x_3^2 - 35x_1^4x_3^4 + 3x_1^2x_2^4x_3^2 \\
& - 16x_1^2x_2^2x_3^4 + 24x_1^2x_3^6 + 16x_2^8 + 44x_2^6x_3^2 + 70x_2^4x_3^4 + 60x_2^2x_3^6 + 30x_3^8.
\end{aligned} \tag{11}$$

As we will see later in this paper, this form which lives in $C_{3,8} \setminus \Sigma C_{3,8}$ turns out to be an example in the smallest possible number of variables but not in the smallest degree. We next present another example of a convex but not sos-convex form that has not been previously in print. The example is in $C_{6,4} \setminus \Sigma C_{6,4}$ and by contrast to the previous example, it will turn out to be minimal in the degree but not in the number of variables. What is nice about this example is that unlike the other examples in this paper it has not been derived with the assistance of a computer and semidefinite programming:

$$\begin{aligned}
q(x_1, \dots, x_6) = & x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4 \\
& + 2(x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2 + x_4^2x_5^2 + x_4^2x_6^2 + x_5^2x_6^2) \\
& + \frac{1}{2}(x_1^2x_4^2 + x_2^2x_5^2 + x_3^2x_6^2) + x_1^2x_6^2 + x_2^2x_4^2 + x_3^2x_5^2 \\
& - (x_1x_2x_4x_5 + x_1x_3x_4x_6 + x_2x_3x_5x_6).
\end{aligned} \tag{12}$$

The proof that this polynomial is convex but not sos-convex can be extracted from [2, Thm. 2.3 and Thm. 2.5]. In there, a general procedure is described for producing convex but not sos-convex quartic forms from *any* example of a psd but not sos biquadratic form. The biquadratic form that has led to the form above is that of Choi in [12].

Also note that the example in (12) shows that convex forms that possess strong symmetry properties can still fail to be sos-convex. The symmetries in this form are inherited from the rich symmetry structure of the biquadratic form of Choi (see [18]). In general, symmetries are of interest in the study of positive semidefinite and sums of squares polynomials because the gap between psd and sos can often behave very differently depending on the symmetry properties; see e.g. [9].

5 Characterization of the gap between convexity and sos-convexity

Now that we know there exist convex polynomials that are not sos-convex, our final and main goal is to give a complete characterization of the degrees and dimensions in which such polynomials can exist. This is achieved in the next theorem.

Theorem 5.1. $\tilde{\Sigma}C_{n,d} = \tilde{C}_{n,d}$ if and only if $n = 1$ or $d = 2$ or $(n, d) = (2, 4)$.

We would also like to have such a characterization for homogeneous polynomials. Although convexity is a property that is in some sense more meaningful for nonhomogeneous polynomials than for forms, one motivation for studying convexity of forms is in their relation to norms [42]. Also, in view of the fact that we have a characterization of the gap between nonnegativity and sums of squares both for polynomials and for forms, it is very natural to inquire the same result for convexity and sos-convexity. The next theorem presents this characterization for forms.

³Assuming $P \neq NP$, and given the NP-hardness of deciding polynomial convexity [2], one would expect to see convex polynomials that are not sos-convex. However, our first example in [5] appeared before the proof of NP-hardness [2]. Moreover, from complexity considerations, even assuming $P \neq NP$, one cannot conclude existence of convex but not sos-convex polynomials for any finite value of the number of variables n .

Theorem 5.2. $\Sigma C_{n,d} = C_{n,d}$ if and only if $n = 2$ or $d = 2$ or $(n, d) = (3, 4)$.

The result $\Sigma C_{3,4} = C_{3,4}$ of this theorem is joint work with G. Blekherman and is to be presented in full detail in [1]. The remainder of this paper is solely devoted to the proof of Theorem 5.1 and the proof of Theorem 5.2 except for the case $(n, d) = (3, 4)$. Before we present these proofs, we shall make two important remarks.

Remark 5.1. Difficulty with homogenization and dehomogenization. Recall from Subsection 2.1 and Theorem 2.1 that characterizing the gap between nonnegativity and sum of squares for polynomials is equivalent to accomplishing this task for forms. Unfortunately, the situation is more complicated for convexity and sos-convexity and that is the reason why we are presenting Theorems 5.1 and 5.2 as separate theorems. The difficulty arises from the fact that unlike nonnegativity and sum of squares, convexity and sos-convexity are not always preserved under homogenization. (Or equivalently, the properties of being not convex and not sos-convex are not preserved under dehomogenization.) In fact, any convex polynomial that is not psd will no longer be convex after homogenization. This is because convex forms are psd but the homogenization of a non-psd polynomial is a non-psd form. Even if a convex polynomial is psd, its homogenization may not be convex. For example the univariate polynomial $10x_1^4 - 5x_1 + 2$ is convex and psd, but its homogenization $10x_1^4 - 5x_1x_2^3 + 2x_2^4$ is not convex.⁴ To observe the same phenomenon for sos-convexity, consider the trivariate form p in (11) which is convex but not sos-convex and define $\tilde{p}(x_2, x_3) = p(1, x_2, x_3)$. Then, one can check that \tilde{p} is sos-convex (i.e., its 2×2 Hessian factors) even though its homogenization which is p is not sos-convex [5].

Remark 5.2. Resemblance to the result of Hilbert. The reader may have noticed from the statements of Theorem 2.1 and Theorems 5.1 and 5.2 that the cases where convex polynomials (forms) are sos-convex are exactly the same cases where nonnegative polynomials are sums of squares! We shall emphasize that as far as we can tell, our results do not follow (except in the simplest cases) from Hilbert’s result stated in Theorem 2.1. Note that the question of convexity or sos-convexity of a polynomial $p(x)$ in n variables and degree d is about the polynomials $g_{\frac{1}{2}}(x, y)$, $g_{\nabla}(x, y)$, or $g_{\nabla^2}(x, y)$ defined in (6) being psd or sos. Even though these polynomials still have degree d , it is important to keep in mind that they are polynomials *in $2n$ variables*. Therefore, there is no direct correspondence with the characterization of Hilbert. To make this more explicit, let us consider for example one particular claim of Theorem 5.2: $\Sigma C_{2,4} = C_{2,4}$. For a form p in 2 variables and degree 4, the polynomials $g_{\frac{1}{2}}$, g_{∇} , and g_{∇^2} will be forms in 4 variables and degree 4. We know from Hilbert’s result that in this situation psd but not sos forms do in fact exist. However, for the forms in 4 variables and degree 4 that have the special structure of $g_{\frac{1}{2}}$, g_{∇} , or g_{∇^2} , psd turns out to be equivalent to sos.

The proofs of Theorems 5.1 and 5.2 are broken into the next two subsections. In Subsection 5.1, we provide the proofs for the cases where convexity and sos-convexity are equivalent. Then in Subsection 5.2, we prove that in all other cases there exist convex polynomials that are not sos-convex.

5.1 Proofs of Theorems 5.1 and 5.2: cases where $\tilde{\Sigma}C_{n,d} = \tilde{C}_{n,d}, \Sigma C_{n,d} = C_{n,d}$

When proving equivalence of convexity and sos-convexity, it turns out to be more convenient to work with the second order characterization of sos-convexity, i.e., with the form $g_{\nabla^2}(x, y) = y^T H(x)y$ in (6). The reason for this is that this form is always quadratic in y , and this allows us to make use of the following key theorem, henceforth referred to as the “biform theorem”.

⁴What is true however is that a nonnegative form of degree d is convex if and only if the d -th root of its dehomogenization is a convex function [42, Prop. 4.4].

Theorem 5.3 (e.g. [14]). *Let $f := f(u_1, u_2, v_1, \dots, v_m)$ be a form in the variables $u := (u_1, u_2)^T$ and $v := (v_1, \dots, v_m)^T$ that is a quadratic form in v for fixed u and a form (of however large degree) in u for fixed v . Then f is psd if and only if it is sos.*⁵

The biform theorem has been proven independently by several authors. See [14] and [6] for more background on this theorem and in particular [14, Sec. 7] for an elegant proof and some refinements. We now proceed with our proofs which will follow in a rather straightforward manner from the biform theorem.

Theorem 5.4. $\tilde{\Sigma}C_{1,d} = \tilde{C}_{1,d}$ for all d . $\Sigma C_{2,d} = C_{2,d}$ for all d .

Proof. For a univariate polynomial, convexity means that the second derivative, which is another univariate polynomial, is psd. Since $\tilde{\Sigma}_{1,d} = \tilde{P}_{1,d}$, the second derivative must be sos. Therefore, $\tilde{\Sigma}C_{1,d} = \tilde{C}_{1,d}$. To prove $\Sigma C_{2,d} = C_{2,d}$, suppose we have a convex bivariate form p of degree d in variables $x := (x_1, x_2)^T$. The Hessian $H := H(x)$ of p is a 2×2 matrix whose entries are forms of degree $d - 2$. If we let $y := (y_1, y_2)^T$, convexity of p implies that the form $y^T H(x)y$ is psd. Since $y^T H(x)y$ meets the requirements of the biform theorem above with $(u_1, u_2) = (x_1, x_2)$ and $(v_1, v_2) = (y_1, y_2)$, it follows that $y^T H(x)y$ is sos. Hence, p is sos-convex. \square

Theorem 5.5. $\tilde{\Sigma}C_{n,2} = \tilde{C}_{n,2}$ for all n . $\Sigma C_{n,2} = C_{n,2}$ for all n .

Proof. Let $x := (x_1, \dots, x_n)^T$ and $y := (y_1, \dots, y_n)^T$. Let $p(x) = \frac{1}{2}x^T Qx + b^T x + c$ be a quadratic polynomial. The Hessian of p in this case is the constant symmetric matrix Q . Convexity of p implies that $y^T Qy$ is psd. But since $\Sigma_{n,2} = P_{n,2}$, $y^T Qy$ must be sos. Hence, p is sos-convex. The proof of $\Sigma C_{n,2} = C_{n,2}$ is identical. \square

Theorem 5.6. $\tilde{\Sigma}C_{2,4} = \tilde{C}_{2,4}$.

Proof. Let $p(x) := p(x_1, x_2)$ be a convex bivariate quartic polynomial. Let $H := H(x)$ denote the Hessian of p and let $y := (y_1, y_2)^T$. Note that $H(x)$ is a 2×2 matrix whose entries are (not necessarily homogeneous) quadratic polynomials. Since p is convex, $y^T H(x)y$ is psd. Let $\bar{H}(x_1, x_2, x_3)$ be a 2×2 matrix whose entries are obtained by homogenizing the entries of H . It is easy to see that $y^T \bar{H}(x_1, x_2, x_3)y$ is then the form obtained by homogenizing $y^T H(x)y$ and is therefore psd. Now we can employ the biform theorem (Theorem 5.3) with $(u_1, u_2) = (y_1, y_2)$ and $(v_1, v_2, v_3) = (x_1, x_2, x_3)$ to conclude that $y^T \bar{H}(x_1, x_2, x_3)y$ is sos. But upon dehomogenizing by setting $x_3 = 1$, we conclude that $y^T H(x)y$ is sos. Hence, p is sos-convex. \square

Theorem 5.7 (Ahmadi, Blekherman, Parrilo [1]). $\Sigma C_{3,4} = C_{3,4}$.

Unlike Hilbert's results $\tilde{\Sigma}_{2,4} = \tilde{P}_{2,4}$ and $\Sigma_{3,4} = P_{3,4}$ which are equivalent statements and essentially have identical proofs, the proof of $\Sigma C_{3,4} = C_{3,4}$ is considerably more involved than the proof of $\tilde{\Sigma}C_{2,4} = \tilde{C}_{2,4}$. Here, we briefly point out why this is the case and refer the reader to [1] for more details.

If $p(x) := p(x_1, x_2, x_3)$ is a ternary quartic form, its Hessian $H(x)$ is a 3×3 matrix whose entries are quadratic forms. In this case, we can no longer apply the biform theorem to the form $y^T H(x)y$. In fact, the matrix

$$C(x) = \begin{bmatrix} x_1^2 + 2x_2^2 & -x_1x_2 & -x_1x_3 \\ -x_1x_2 & x_2^2 + 2x_3^2 & -x_2x_3 \\ -x_1x_3 & -x_2x_3 & x_3^2 + 2x_1^2 \end{bmatrix},$$

⁵Note that the results $\Sigma_{2,d} = P_{2,d}$ and $\Sigma_{n,2} = P_{n,2}$ are both special cases of this theorem.

due to Choi [12] serves as an explicit example of a 3×3 matrix with quadratic form entries that is positive semidefinite but not an sos-matrix; i.e., $y^T C(x)y$ is psd but not sos. However, the matrix $C(x)$ above is *not* a valid Hessian, i.e., it cannot be the matrix of the second derivatives of any polynomial. If this was the case, the third partial derivatives would commute. On the other hand, we have in particular

$$\frac{\partial C_{1,1}(x)}{\partial x_3} = 0 \neq -x_3 = \frac{\partial C_{1,3}(x)}{\partial x_1}.$$

It is rather remarkable that with the additional requirement of being a valid Hessian, the form $y^T H(x)y$ turns out to be psd if and only if it is sos [1].

5.2 Proofs of Theorems 5.1 and 5.2: cases where $\tilde{\Sigma}C_{n,d} \subset \tilde{C}_{n,d}, \Sigma C_{n,d} \subset C_{n,d}$

The goal of this subsection is to establish that the cases presented in the previous subsection are the *only* cases where convexity and sos-convexity are equivalent. We will first give explicit examples of convex but not sos-convex polynomials/forms that are “minimal” jointly in the degree and dimension and then present an argument for all dimensions and degrees higher than those of the minimal cases.

5.2.1 Minimal convex but not sos-convex polynomials/forms

The minimal examples of convex but not sos-convex polynomials (resp. forms) turn out to belong to $\tilde{C}_{2,6} \setminus \tilde{\Sigma}C_{2,6}$ and $\tilde{C}_{3,4} \setminus \tilde{\Sigma}C_{3,4}$ (resp. $C_{3,6} \setminus \Sigma C_{3,6}$ and $C_{4,4} \setminus \Sigma C_{4,4}$). Recall from Remark 5.1 that we lack a general argument for going from convex but not sos-convex forms to polynomials or vice versa. Because of this, one would need to present four different polynomials in the sets mentioned above and prove that each polynomial is (i) convex and (ii) not sos-convex. This is a total of eight arguments to make which is quite cumbersome. However, as we will see in the proof of Theorem 5.8 and 5.9 below, we have been able to find examples that act “nicely” with respect to particular ways of dehomogenization. This will allow us to reduce the total number of claims we have to prove from eight to four.

The polynomials that we are about to present next have been found with the assistance of a computer and by employing some “tricks” with semidefinite programming.⁶ In this process, we have made use of software packages YALMIP [27], SOSTOOLS [38], and the SDP solver SeDuMi [47], which we acknowledge here. To make the paper relatively self-contained and to emphasize the fact that using *rational sum of squares certificates* one can make such computer assisted proofs fully formal, we present the proof of Theorem 5.8 below in the appendix. On the other hand, the proof of Theorem 5.9, which is very similar in style to the proof of Theorem 5.8, is largely omitted to save space. All of the proofs are available in electronic form and in their entirety at <http://aaa.lids.mit.edu/software>.

Theorem 5.8. $\tilde{\Sigma}C_{2,6}$ is a proper subset of $\tilde{C}_{2,6}$. $\Sigma C_{3,6}$ is a proper subset of $C_{3,6}$.

⁶We do not elaborate here on how this was exactly done. The interested reader is referred to [5, Sec. 4], where a similar technique for formulating semidefinite programs that can search over a convex subset of the (non-convex) set of convex but not sos-convex polynomials is explained. The approach in [5] however does not lead to examples that are minimal.

Proof. We claim that the form

$$\begin{aligned}
f(x_1, x_2, x_3) = & 77x_1^6 - 155x_1^5x_2 + 445x_1^4x_2^2 + 76x_1^3x_2^3 + 556x_1^2x_2^4 + 68x_1x_2^5 + 240x_2^6 - 9x_1^5x_3 \\
& - 1129x_1^3x_2^2x_3 + 62x_1^2x_2^3x_3 + 1206x_1x_2^4x_3 - 343x_2^5x_3 + 363x_1^4x_3^2 + 773x_1^3x_2x_3^2 \\
& + 891x_1^2x_2^2x_3^2 - 869x_1x_2^3x_3^2 + 1043x_2^4x_3^2 - 14x_1^3x_3^3 - 1108x_1^2x_2x_3^3 - 216x_1x_2^2x_3^3 \\
& - 839x_2^3x_3^3 + 721x_1^2x_3^4 + 436x_1x_2x_3^4 + 378x_2^2x_3^4 + 48x_1x_3^5 - 97x_2x_3^5 + 89x_3^6
\end{aligned} \tag{13}$$

belongs to $C_{3,6} \setminus \Sigma C_{3,6}$, and the polynomial⁷

$$\tilde{f}(x_1, x_2) = f(x_1, x_2, 1 - \frac{1}{2}x_2) \tag{14}$$

belongs to $\tilde{C}_{2,6} \setminus \tilde{\Sigma}C_{2,6}$. Note that since convexity and sos-convexity are both preserved under restrictions to affine subspaces (recall Remark 3.2), it suffices to show that the form f in (13) is convex and the polynomial \tilde{f} in (14) is not sos-convex. Let $x := (x_1, x_2, x_2)^T$, $y := (y_1, y_2, y_3)^T$, $\tilde{x} := (x_1, x_2)^T$, $\tilde{y} := (y_1, y_2)^T$, and denote the Hessian of f and \tilde{f} respectively by H_f and $H_{\tilde{f}}$. In the appendix, we provide rational Gram matrices which prove that the form

$$(x_1^2 + x_2^2) \cdot y^T H_f(x) y \tag{15}$$

is sos. This, together with nonnegativity of $x_1^2 + x_2^2$ and continuity of $y^T H_f(x) y$, implies that $y^T H_f(x) y$ is psd. Therefore, f is convex. The proof that \tilde{f} is not sos-convex proceeds by showing that $H_{\tilde{f}}$ is not an sos-matrix via a separation argument. In the appendix, we present a separating hyperplane that leaves the appropriate sos cone on one side and the polynomial

$$\tilde{y}^T H_{\tilde{f}}(\tilde{x}) \tilde{y} \tag{16}$$

on the other. □

Theorem 5.9. $\tilde{\Sigma}C_{3,4}$ is a proper subset of $\tilde{C}_{3,4}$. $\Sigma C_{4,4}$ is a proper subset of $C_{4,4}$.

Proof. We claim that the form

$$\begin{aligned}
h(x_1, \dots, x_4) = & 1671x_1^4 - 4134x_1^3x_2 - 3332x_1^3x_3 + 5104x_1^2x_2^2 + 4989x_1^2x_2x_3 + 3490x_1^2x_3^2 \\
& - 2203x_1x_2^3 - 3030x_1x_2^2x_3 - 3776x_1x_2x_3^2 - 1522x_1x_3^3 + 1227x_2^4 - 595x_2^3x_3 \\
& + 1859x_2^2x_3^2 + 1146x_2x_3^3 + 979x_3^4 + 1195728x_4^4 - 1932x_1x_4^3 - 2296x_2x_4^3 \\
& - 3144x_3x_4^3 + 1465x_1^2x_4^2 - 1376x_1^3x_4 - 263x_1x_2x_4^2 + 2790x_1^2x_2x_4 + 2121x_2^2x_4^2 \\
& - 292x_1x_2^2x_4 - 1224x_2^3x_4 + 2404x_1x_3x_4^2 + 2727x_2x_3x_4^2 - 2852x_1x_3^2x_4 \\
& - 388x_2x_3^2x_4 - 1520x_3^3x_4 + 2943x_1^2x_3x_4 - 5053x_1x_2x_3x_4 + 2552x_2^2x_3x_4 \\
& + 3512x_3^2x_4^2
\end{aligned} \tag{17}$$

⁷The polynomial $f(x_1, x_2, 1)$ turns out to be sos-convex, and therefore does not do the job. One can of course change coordinates, and then in the new coordinates perform the dehomogenization by setting $x_3 = 1$.

belongs to $C_{4,4} \setminus \Sigma C_{4,4}$, and the polynomial

$$\tilde{h}(x_1, x_2, x_3) = h(x_1, x_2, x_3, 1) \quad (18)$$

belongs to $\tilde{C}_{3,4} \setminus \tilde{\Sigma} C_{3,4}$. Once again, it suffices to prove that h is convex and \tilde{h} is not sos-convex. Let $x := (x_1, x_2, x_3, x_4)^T$, $y := (y_1, y_2, y_3, y_4)^T$, and denote the Hessian of h and \tilde{h} respectively by H_h and $H_{\tilde{h}}$. The proof that h is convex is done by showing that the form

$$(x_2^2 + x_3^2 + x_4^2) \cdot y^T H_h(x) y \quad (19)$$

is sos.⁸ The proof that \tilde{h} is not sos-convex is done again by means of a separating hyperplane. \square

5.2.2 Convex but not sos-convex polynomials/forms in all higher degrees and dimensions

Given a convex but not sos-convex polynomial (form) in n variables, it is very easy to argue that such a polynomial (form) must also exist in a larger number of variables. If $p(x_1, \dots, x_n)$ is a form in $C_{n,d} \setminus \Sigma C_{n,d}$, then

$$\bar{p}(x_1, \dots, x_{n+1}) = p(x_1, \dots, x_n) + x_{n+1}^d$$

belongs to $C_{n+1,d} \setminus \Sigma C_{n+1,d}$. Convexity of \bar{p} is obvious since it is a sum of convex functions. The fact that \bar{p} is not sos-convex can also easily be seen from the block diagonal structure of the Hessian of \bar{p} : if the Hessian of \bar{p} were to factor, it would imply that the Hessian of p should also factor. The argument for going from $\tilde{C}_{n,d} \setminus \tilde{\Sigma} C_{n,d}$ to $\tilde{C}_{n+1,d} \setminus \tilde{\Sigma} C_{n+1,d}$ is identical.

Unfortunately, an argument for increasing the degree of convex but not sos-convex forms seems to be significantly more difficult to obtain. In fact, we have been unable to come up with a natural operation that would produce a form in $C_{n,d+2} \setminus \Sigma C_{n,d+2}$ from a form in $C_{n,d} \setminus \Sigma C_{n,d}$. We will instead take a different route: we are going to present a general procedure for going from a form in $P_{n,d} \setminus \Sigma_{n,d}$ to a form in $C_{n,d+2} \setminus \Sigma C_{n,d+2}$. This will serve our purpose of constructing convex but not sos-convex forms in higher degrees and is perhaps also of independent interest in itself. For instance, it can be used to construct convex but not sos-convex forms that inherit structural properties (e.g. symmetry) of the known examples of psd but not sos forms. The procedure is constructive modulo the value of two positive constants (γ and α below) whose existence will be shown nonconstructively.⁹

Although the proof of the general case is no different, we present this construction for the case $n = 3$. The reason is that it suffices for us to construct forms in $C_{3,d} \setminus \Sigma C_{3,d}$ for d even and ≥ 8 . These forms together with the two forms in $C_{3,6} \setminus \Sigma C_{3,6}$ and $C_{4,4} \setminus \Sigma C_{4,4}$ presented in (13) and (17), and with the simple procedure for increasing the number of variables cover all the values of n and d for which convex but not sos-convex forms exist.

For the remainder of this section, let $x := (x_1, x_2, x_3)^T$ and $y := (y_1, y_2, y_3)^T$.

Theorem 5.10. *Let $m := m(x)$ be a ternary form of degree d (with d necessarily even and ≥ 6) satisfying the following three requirements:*

R1: m is positive definite.

R2: m is not a sum of squares.

R3: The Hessian H_m of m is positive definite at the point $(1, 0, 0)^T$.

⁸The choice of multipliers in (15) and (19) is motivated by a result of Reznick in [39].

⁹The procedure can be thought of as a generalization of the approach in our earlier work in [5].

Let $g := g(x_2, x_3)$ be any bivariate form of degree $d + 2$ whose Hessian is positive definite. Then, there exists a constant $\gamma > 0$, such that the form f of degree $d + 2$ given by

$$f(x) = \int_0^{x_1} \int_0^s m(t, x_2, x_3) dt ds + \gamma g(x_2, x_3) \quad (20)$$

is convex but not sos-convex.

Before we prove this theorem, let us comment on how one can get examples of forms m and g that satisfy the requirements of the theorem. The choice of g is in fact very easy. We can e.g. take

$$g(x_2, x_3) = (x_2^2 + x_3^2)^{\frac{d+2}{2}},$$

which has a positive definite Hessian. As for the choice of m , essentially any psd but not sos ternary form can be turned into a form that satisfies requirements **R1**, **R2**, and **R3**. Indeed if the Hessian of such a form is positive definite at just one point, then that point can be taken to $(1, 0, 0)^T$ by a change of coordinates without changing the properties of being psd and not sos. If the form is not positive definite, then it can be made so by adding a small enough multiple of a positive definite form to it. For concreteness, we construct in the next lemma a family of forms that together with the above theorem will give us convex but not sos-convex ternary forms of any degree ≥ 8 .

Lemma 5.11. *For any even degree $d \geq 6$, there exists a constant $\alpha > 0$, such that the form*

$$m(x) = x_1^{d-6}(x_1^2 x_2^4 + x_1^4 x_2^2 - 3x_1^2 x_2^2 x_3^2 + x_3^6) + \alpha(x_1^2 + x_2^2 + x_3^2)^{\frac{d}{2}} \quad (21)$$

satisfies the requirements **R1**, **R2**, and **R3** of Theorem 5.10.

Proof. The form

$$x_1^2 x_2^4 + x_1^4 x_2^2 - 3x_1^2 x_2^2 x_3^2 + x_3^6$$

is the familiar Motzkin form in (1) that is psd but not sos [29]. For any even degree $d \geq 6$, the form

$$x_1^{d-6}(x_1^2 x_2^4 + x_1^4 x_2^2 - 3x_1^2 x_2^2 x_3^2 + x_3^6)$$

is a form of degree d that is clearly still psd and less obviously still not sos; see [40]. This together with the fact that $\Sigma_{n,d}$ is a closed cone implies existence of a small positive value of α for which the form m in (21) is positive definite but not a sum of squares, hence satisfying requirements **R1** and **R2**.

Our next claim is that for any positive value of α , the Hessian H_m of the form m in (21) satisfies

$$H_m(1, 0, 0) = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix} \quad (22)$$

for some positive constants c_1, c_2, c_3 , therefore also passing requirement **R3**. To see the above equality, first note that since m is a form of degree d , its Hessian H_m will have entries that are forms of degree $d - 2$. Therefore, the only monomials that can survive in this Hessian after setting x_2 and x_3 to zero are multiples of x_1^{d-2} . It is easy to see that an x_1^{d-2} monomial in an off-diagonal entry of H_m would lead to a monomial in m that is not even. On the other hand, the form m in (21) only has even monomials. This explains why the off-diagonal entries of the right hand side of (22) are zero. Finally, we note that for any positive value of α , the form m in (21) includes positive multiples of x_1^d , $x_1^{d-2} x_2^2$, and $x_1^{d-2} x_3^2$, which lead to positive multiples of x_1^{d-2} on the diagonal of H_m . Hence, c_1, c_2 , and c_3 are positive. \square

Next, we state two lemmas that will be employed in the proof of Theorem 5.10.

Lemma 5.12. *Let m be a trivariate form satisfying the requirements **R1** and **R3** of Theorem 5.10. Let $H_{\hat{m}}$ denote the Hessian of the form $\int_0^{x_1} \int_0^s m(t, x_2, x_3) dt ds$. Then, there exists a positive constant δ , such that*

$$y^T H_{\hat{m}}(x)y > 0$$

on the set

$$\mathcal{S} := \{(x, y) \mid \|x\| = 1, \|y\| = 1, (x_2^2 + x_3^2 < \delta \text{ or } y_2^2 + y_3^2 < \delta)\}. \quad (23)$$

Proof. We observe that when $y_2^2 + y_3^2 = 0$, we have

$$y^T H_{\hat{m}}(x)y = y_1^2 m(x),$$

which by requirement **R1** is positive when $\|x\| = \|y\| = 1$. By continuity of the form $y^T H_{\hat{m}}(x)y$, we conclude that there exists a small positive constant δ_y such that $y^T H_{\hat{m}}(x)y > 0$ on the set

$$\mathcal{S}_y := \{(x, y) \mid \|x\| = 1, \|y\| = 1, y_2^2 + y_3^2 < \delta_y\}.$$

Next, we leave it to the reader to check that

$$H_{\hat{m}}(1, 0, 0) = \frac{1}{d(d-1)} H_m(1, 0, 0).$$

Therefore, when $x_2^2 + x_3^2 = 0$, requirement **R3** implies that $y^T H_{\hat{m}}(x)y$ is positive when $\|x\| = \|y\| = 1$. Appealing to continuity again, we conclude that there exists a small positive constant δ_x such that $y^T H_{\hat{m}}(x)y > 0$ on the set

$$\mathcal{S}_x := \{(x, y) \mid \|x\| = 1, \|y\| = 1, x_2^2 + x_3^2 < \delta_x\}.$$

If we now take $\delta = \min\{\delta_y, \delta_x\}$, the lemma is established. \square

The last lemma that we need has already been proven in our earlier work in [5].

Lemma 5.13 ([5]). *All principal minors of an sos-matrix are sos polynomials.*¹⁰

We are now ready to prove Theorem 5.10.

Proof of Theorem 5.10. We first prove that the form f in (20) is not sos-convex. By Lemma 5.13, if f was sos-convex, then all diagonal elements of its Hessian would have to be sos polynomials. On the other hand, we have from (20) that

$$\frac{\partial f(x)}{\partial x_1 \partial x_1} = m(x),$$

which by requirement **R2** is not sos. Therefore f is not sos-convex.

It remains to show that there exists a positive value of γ for which f becomes convex. Let us denote the Hessians of f , $\int_0^{x_1} \int_0^s m(t, x_2, x_3) dt ds$, and g , by H_f , $H_{\hat{m}}$, and H_g respectively. So, we have

$$H_f(x) = H_{\hat{m}}(x) + \gamma H_g(x_2, x_3).$$

¹⁰As a side note, we remark that the converse of Lemma 5.13 is not true even for polynomial matrices that are valid Hessians. For example, all 7 principal minors of the 3×3 Hessian of the form f in (13) are sos polynomials, even though this Hessian is not an sos-matrix.

(Here, H_g is a 3×3 matrix whose first row and column are zeros.) Convexity of f is of course equivalent to nonnegativity of the form $y^T H_f(x) y$. Since this form is bi-homogeneous in x and y , it is nonnegative if and only if $y^T H_f(x) y \geq 0$ on the bi-sphere

$$\mathcal{B} := \{(x, y) \mid \|x\| = 1, \|y\| = 1\}.$$

Let us decompose the bi-sphere as

$$\mathcal{B} = \mathcal{S} \cup \bar{\mathcal{S}},$$

where \mathcal{S} is defined in (23) and

$$\bar{\mathcal{S}} := \{(x, y) \mid \|x\| = 1, \|y\| = 1, x_2^2 + x_3^2 \geq \delta, y_2^2 + y_3^2 \geq \delta\}.$$

Lemma 5.12 together with positive definiteness of H_g imply that $y^T H_f(x) y$ is positive on \mathcal{S} . As for the set $\bar{\mathcal{S}}$, let

$$\beta_1 = \min_{x, y \in \mathcal{S}} y^T H_{\hat{m}}(x) y,$$

and

$$\beta_2 = \min_{x, y \in \bar{\mathcal{S}}} y^T H_g(x_2, x_3) y.$$

By the assumption of positive definiteness of H_g , we have $\beta_2 > 0$. If we now let

$$\gamma > \frac{|\beta_1|}{\beta_2},$$

then

$$\min_{x, y \in \mathcal{S}} y^T H_f(x) y > \beta_1 + \frac{|\beta_1|}{\beta_2} \beta_2 \geq 0.$$

Hence $y^T H_f(x) y$ is nonnegative (in fact positive) everywhere on \mathcal{B} and the proof is completed. \square

Finally, we provide an argument for existence of bivariate polynomials of degree 8, 10, 12, ... that are convex but not sos-convex.

Corollary 5.14. *Consider the form f in (20) constructed as described in Theorem 5.10. Let*

$$\tilde{f}(x_1, x_2) = f(x_1, x_2, 1).$$

Then, \tilde{f} is convex but not sos-convex.

Proof. The polynomial \tilde{f} is convex because it is the restriction of a convex function. It is not difficult to see that

$$\frac{\partial \tilde{f}(x_1, x_2)}{\partial x_1 \partial x_1} = m(x_1, x_2, 1),$$

which is not sos. Therefore from Lemma 5.13 \tilde{f} is not sos-convex. \square

Corollary 5.14 together with the two polynomials in $\tilde{\mathcal{C}}_{2,6} \setminus \tilde{\Sigma}\mathcal{C}_{2,6}$ and $\tilde{\mathcal{C}}_{3,4} \setminus \tilde{\Sigma}\mathcal{C}_{3,4}$ presented in (14) and (18), and with the simple procedure for increasing the number of variables described at the beginning of Subsection 5.2.2 cover all the values of n and d for which convex but not sos-convex polynomials exist.

6 Concluding remarks and an open problem

To conclude our paper, we would like to point out some similarities between nonnegativity and convexity that deserve attention: (i) both nonnegativity and convexity are properties that only hold for even degree polynomials, (ii) for quadratic forms, nonnegativity is in fact equivalent to convexity, (iii) both notions are NP-hard to check exactly for degree 4 and larger, and most strikingly (iv) nonnegativity is equivalent to sum of squares *exactly* in dimensions and degrees where convexity is equivalent to sos-convexity. It is unclear to us whether there can be a deeper and more unifying reason explaining these observations, in particular, the last one which was the main result of this paper.

Another intriguing question is to investigate whether one can give a direct argument proving the fact that $\tilde{\Sigma}C_{n,d} = \tilde{C}_{n,d}$ if and only if $\Sigma C_{n+1,d} = C_{n+1,d}$. This would eliminate the need for studying polynomials and forms separately, and in particular would provide a short proof of the result $\Sigma_{3,4} = C_{3,4}$ given in [1].

Finally, an open problem related to this work is to *find an explicit example of a convex form that is not a sum of squares*. Blekherman [7] has shown via volume arguments that for degree $d \geq 4$ and asymptotically for large n such forms must exist, although no examples are known. In particular, it would be interesting to determine the smallest value of n for which such a form exists. We know from Lemma 3.2 that a convex form that is not sos must necessarily be not sos-convex. Although our several constructions of convex but not sos-convex polynomials pass this necessary condition, the polynomials themselves are all sos. The question is particularly interesting from an optimization viewpoint because it implies that the well-known sum of squares relaxation for minimizing polynomials [46], [35] may not be exact even for the easy case of minimizing convex polynomials.

References

- [1] A. A. Ahmadi, G. Blekherman, and P. A. Parrilo. Convex ternary quartics are sos-convex. In preparation, 2011.
- [2] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical Programming*, 2011. Accepted for publication. Online version available at arXiv:1012.1908.
- [3] 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.
- [4] A. A. Ahmadi and P. A. Parrilo. On the equivalence of algebraic conditions for convexity and quasiconvexity of polynomials. In *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010.
- [5] A. A. Ahmadi and P. A. Parrilo. A convex polynomial that is not sos-convex. *Mathematical Programming*, 2011. DOI: 10.1007/s10107-011-0457-z.
- [6] E. M. Aylward, S. M. Itani, and P. A. Parrilo. Explicit SOS decomposition of univariate polynomial matrices and the Kalman-Yakubovich-Popov lemma. In *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007.
- [7] G. Blekherman. Convex forms that are not sums of squares. arXiv:0910.0656., 2009.

- [8] G. Blekherman. Nonnegative polynomials and sums of squares. arXiv:1010.3465, 2010.
- [9] G. Blekherman and C. B. Riener. Symmetric sums of squares in degree four. In preparation, 2011.
- [10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [11] 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.
- [12] M. D. Choi. Positive semidefinite biquadratic forms. *Linear Algebra and its Applications*, 12:95–100, 1975.
- [13] M. D. Choi and T. Y. Lam. Extremal positive semidefinite forms. *Math. Ann.*, 231:1–18, 1977.
- [14] M. D. Choi, T.-Y. Lam, and B. Reznick. Real zeros of positive semidefinite forms. I. *Math. Z.*, 171(1):1–26, 1980.
- [15] E. de Klerk and M. Laurent. On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems. Available at <http://www.optimization-online.org/DB-FILE/2010/11/2800.pdf>, 2010.
- [16] P. H. Diananda. On non-negative forms in real variables some or all of which are non-negative. *Proceedings of the Cambridge philosophical society*, 58:17–25, 1962.
- [17] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing separable and entangled states. *Physical Review Letters*, 88(18), 2002.
- [18] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192:95–128, 2004.
- [19] N. Gvozdenović and M. Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. *Mathematical Programming*, 110(1):145–173, 2007.
- [20] J. W. Helton and J. Nie. Semidefinite representation of convex sets. *Mathematical Programming*, 122(1, Ser. A):21–64, 2010.
- [21] D. Henrion and A. Garulli, editors. *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Information Sciences*. Springer, 2005.
- [22] D. Hilbert. Über die Darstellung Definiten Formen als Summe von Formenquadraten. *Math. Ann.*, 32, 1888.
- [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] J. B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization*, 19(4):1995–2014, 2008.
- [25] J. B. Lasserre. Representation of nonnegative convex polynomials. *Archiv der Mathematik*, 91(2):126–130, 2008.
- [26] J. B. Lasserre. Certificates of convexity for basic semi-algebraic sets. *Applied Mathematics Letters*, 23(8):912–916, 2010.

- [27] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, 2004. Available from <http://control.ee.ethz.ch/~joloef/yalmip.php>.
- [28] 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.
- [29] T. S. Motzkin. The arithmetic-geometric inequality. In *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*, pages 205–224. Academic Press, New York, 1967.
- [30] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.
- [31] P. M. Pardalos and S. A. Vavasis. Open questions in complexity theory for numerical optimization. *Mathematical Programming*, 57(2):337–339, 1992.
- [32] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000.
- [33] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2, Ser. B):293–320, 2003.
- [34] P. A. Parrilo. Polynomial games and sum of squares optimization. In *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006.
- [35] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. *Algorithmic and Quantitative Real Algebraic Geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 60:83–99, 2003.
- [36] A. Pfister and C. Scheiderer. An elementary proof of Hilbert’s theorem on ternary quartics. arXiv:1009.3144, 2010.
- [37] V. Powers, B. Reznick, C. Scheiderer, and F. Sottile. A new approach to Hilbert’s theorem on ternary quartics. *Comptes Rendus Mathématique*, 339(9):617 – 620, 2004.
- [38] 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>.
- [39] B. Reznick. Uniform denominators in Hilbert’s 17th problem. *Math Z.*, 220(1):75–97, 1995.
- [40] B. Reznick. Some concrete aspects of Hilbert’s 17th problem. In *Contemporary Mathematics*, volume 253, pages 251–272. American Mathematical Society, 2000.
- [41] B. Reznick. On Hilbert’s construction of positive polynomials. arXiv:0707.2156., 2007.
- [42] B. Reznick. Blenders. arXiv:1008.4533, 2010.
- [43] 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.
- [44] W. Rudin. *Real and complex analysis*. McGraw-Hill series in higher mathematics, 1987. Third edition.

- [45] C. W. Scherer and C. W. J. Hol. Matrix sum of squares relaxations for robust semidefinite programs. *Mathematical Programming*, 107:189–211, 2006.
- [46] N. Z. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23(6):731–734, 1987. (Russian orig.: Kibernetika, No. 6, (1987), 9–11).
- [47] J. Sturm. *SeDuMi version 1.05*, Oct. 2001. Latest version available at <http://sedumi.ie.lehigh.edu/>.
- [48] A. L. Tits. Lecture notes on optimal control. Available from <http://www.isr.umd.edu/~andre/664.pdf>, 2008.
- [49] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, Mar. 1996.

A Certificates complementing the proof of Theorem 5.8

Let $x := (x_1, x_2, x_2)^T$, $y := (y_1, y_2, y_3)^T$, $\tilde{x} := (x_1, x_2)^T$, $\tilde{y} := (y_1, y_2)^T$, and let f, \tilde{f}, H_f , and $H_{\tilde{f}}$ be as in the proof of Theorem 5.8. This appendix proves that the form $(x_1^2 + x_2^2) \cdot y^T H_f(x)y$ in (15) is sos and that the polynomial $\tilde{y}^T H_{\tilde{f}}(\tilde{x})\tilde{y}$ in (16) is not sos, hence proving respectively that f is convex and \tilde{f} is not sos-convex.

A rational sos decomposition of $(x_1^2 + x_2^2) \cdot y^T H_f(x)y$, which is a form in 6 variables of degree 8, is as follows:

$$(x_1^2 + x_2^2) \cdot y^T H_f(x)y = \frac{1}{84} z^T Q z,$$

where z is the vector of monomials

$$\begin{aligned} z = & [x_2 x_3^2 y_3, x_2 x_3^2 y_2, x_2 x_3^2 y_1, x_2^2 x_3 y_3, x_2^2 x_3 y_2, x_2^2 x_3 y_1, x_2^3 y_3, x_2^3 y_2, x_2^3 y_1, \\ & x_1 x_3^2 y_3, x_1 x_3^2 y_2, x_1 x_3^2 y_1, x_1 x_2 x_3 y_3, x_1 x_2 x_3 y_2, x_1 x_2 x_3 y_1, x_1 x_2^2 y_3, x_1 x_2^2 y_2, x_1 x_2^2 y_1, \\ & x_1^2 x_3 y_3, x_1^2 x_3 y_2, x_1^2 x_3 y_1, x_1^2 x_2 y_3, x_1^2 x_2 y_2, x_1^2 x_2 y_1, x_1^3 y_3, x_1^3 y_2, x_1^3 y_1]^T, \end{aligned}$$

and Q is the 27×27 positive definite matrix¹¹ presented on the next page

$$Q = [Q_1 \quad Q_2],$$

¹¹Whenever we state a matrix is positive definite, this claim is backed up by a rational LDL^T factorization of the matrix that the reader can find online at <http://aaa.lids.mit.edu/software>.

224280	-40740	20160	139692	13576	-50540	-27804	-48384	0	-29400	-32172	21252	103404
-40740	63504	36624	114324	-81480	-211428	139692	93576	-50540	-27804	-48384	0	103404
20160	36624	121128	52920	-81480	-27972	139692	-8316	-67704	0	16632	75936	15540
-81480	114324	52920	482104	-538776	36204	-211428	93072	-42252	-16632	0	214284	-57960
139692	-211428	-27972	-538776	1020600	-94416	1020600	-94416	338016	-33684	-46620	-113638	20258
93576	-8316	36204	36204	-94416	266448	1020600	266448	75348	-33432	-31080	-221606	254534
-50540	-67704	-42252	338016	-211428	338016	338016	175224	-144060	7836	-4298	-77280	108192
-48384	15372	-77196	362880	-288120	216468	-144060	604800	28560	-840	-35434	-132804	134736
0	29400	32172	19068	19068	46368	3360	-20692	101304	-40740	20160	35028	22680
-29400	0	-16632	13524	-33684	-33684	7826	-840	2786	63504	36624	51828	-196476
-32172	16632	0	-588	-46620	-31080	-4298	-35434	-11088	36624	121128	29148	-9408
21252	75936	214284	179564	-113638	-221606	-77280	-132804	-104496	51828	29148	782976	-463344
103404	15540	-57960	20258	-119112	254534	-108192	134736	-22680	196476	-9408	-463344	1167624
267456	-48132	27552	-49742	78470	124236	61404	-90384	55524	-50064	-145908	41016	-15456
60872	-25690	-142478	22848	-113820	259980	237972	20412	-95580	-47964	-27780	-438732	514500
37730	-99036	-10150	-83160	473088	34188	167244	57120	159264	10752	-93048	230832	-49980
-119210	9170	81648	244356	-41664	-194124	-9996	214368	19152	2940	-48480	204708	-85344
-116508	81564	26124	155832	-308280	-78180	-74088	14616	-49644	87108	225456	135744	8508
30660	-14952	11844	-21420	62604	14364	13608	1176	5124	-18144	-99624	31332	-178248
35700	11340	52836	-70788	86184	9396	12264	-108024	-11256	-86520	-3528	-19334	142128
102156	2856	64536	22176	-4200	77532	-70896	54348	-49616	-78876	-144998	29316	23856
95364	-504	-8412	-23100	230916	-223692	-131628	59640	-31416	-39396	-44520	158508	308196
11256	5208	32158	-33264	45444	3122	6888	-34440	-5628	19152	8988	18060	-19467
0	-1344	-3696	-34692	33768	5964	9492	-20244	5208	-9912	58884	-50883	151956
-51422	49056	32592	160370	-229068	-36792	-68796	57708	-39564	31164	-8008	141876	-126483

$Q_1 =$

267456	60872	37730	-119210	30660	102156	102156	-86100	95364	11256	0	-51422	
-48132	-25690	-99036	9170	81564	11340	2856	47148	-504	5208	-1344	49056	
27552	-142478	-10150	81648	26124	52836	64536	71820	-8412	32158	-3696	32592	
-49742	22848	-83160	244356	155832	-21420	-70788	22176	230916	-33264	-34692	160370	
78470	-113820	473088	-41664	-308280	62604	86184	-4200	-223692	28140	45444	33768	
124236	259980	34188	-194124	-78180	14368	9396	77532	-131628	81648	3122	5964	
-100464	-72996	167244	-9996	-74088	13608	12264	-70896	-72156	-26768	6888	9492	
61404	237972	57120	214368	14616	-108024	54348	59640	-31416	-13944	-20244	57708	
-90384	20412	159264	19152	-49644	5124	-11256	-49616	-31416	-3944	5208	-39564	
-145908	-27780	-183540	204708	40320	59388	259056	72744	-75096	-51002	61320	-30072	
-50064	-47964	-93048	2940	87108	-18144	-86520	-78876	-39396	-39228	-19152	-9912	
-145908	-27780	-183540	204708	40320	59388	259056	72744	-75096	-51002	61320	-30072	
41016	-438732	230832	204708	135744	8568	-178248	23856	308196	298956	-19467	151956	
-15456	514500	-49980	-85344	8568	-178248	142128	23856	308196	298956	-19467	151956	
610584	21840	127932	-65184	-323834	195636	90972	339794	-100716	96012	24864	-114219	
21840	466704	-110628	-106820	-54012	-90636	-111790	-14952	63672	107856	-67788	61404	
127932	-110628	1045968	142632	-410592	171024	86268	176820	96516	199752	13524	-70784	
-65184	-106820	142632	569856	21518	-30156	-159264	-23016	410004	-71484	-62076	-13860	
-323834	-54012	-410592	21518	604128	-229992	-75516	-297276	182385	75684	-3528	94500	
195636	-90636	171024	-30156	-229992	169512	104748	187341	-136332	-145719	35364	24612	
90972	-111790	86268	-159264	-75516	104748	381920	147168	-182595	-36876	105504	-24612	
339794	-14952	176820	-23016	-297276	187341	147168	346248	-59304	-137928	64932	-90888	
-96012	63672	96516	410004	182385	-136332	-182595	-59304	776772	48972	494536	19152	
-96012	107856	199752	-71484	75684	-145719	-36876	-137928	48972	-28392	118188	-130200	
24864	-67788	13524	-62076	-3528	35364	105504	64932	-98784	-28392	60984	0	
-114219	61404	-70784	-13860	94500	-94836	-24612	-90888	19152	118188	0	74760	
36876	-88284	-42756	74032	138432	24612	-7560	28392	180852	-130200	-3780	-65100	
												194040

$Q_2 =$

Next, we prove that the polynomial $\tilde{y}^T H_{\tilde{f}}(\tilde{x})\tilde{y}$ in (16) is not sos. Let us first present this polynomial and give it a name:

$$\begin{aligned}
t(\tilde{x}, \tilde{y}) := \tilde{y}^T H_{\tilde{f}}(\tilde{x})\tilde{y} = & 294x_1x_2y_2^2 - 6995x_2^4y_1y_2 - 10200x_1y_1y_2 - 4356x_1^2x_2y_1^2 - 2904x_1^3y_1y_2 \\
& - 11475x_1x_2^2y_1^2 + 13680x_2^3y_1y_2 + 4764x_1x_2y_1^2 + 4764x_1^2y_1y_2 + 6429x_1^2x_2^2y_1^2 \\
& + 294x_2^2y_1y_2 - 13990x_1x_2^3y_2^2 - 12123x_1^2x_2y_2^2 - 3872x_2y_1y_2 + \frac{2143}{2}x_1^4y_2^2 \\
& + 20520x_1x_2^2y_2^2 + 29076x_1x_2y_1y_2 - 24246x_1x_2^2y_1y_2 + 14901x_1x_2^3y_1y_2 \\
& + 15039x_1^2x_2^2y_1y_2 + 8572x_1^3x_2y_1y_2 + \frac{44703}{4}x_1^2x_2^2y_2^2 + 5013x_1^3x_2y_2^2 \\
& + 632y_1y_2 - 12360x_2y_2^2 - 5100x_2y_1^2 + \frac{147513}{4}x_2^2y_2^2 + 7269x_2^2y_1^2 - 45025x_2^3y_2^2 \\
& + \frac{772965}{32}x_2^4y_2^2 + \frac{14901}{8}x_2^4y_1^2 - 1936x_1y_2^2 - 84x_1y_1^2 + \frac{3817}{2}y_2^2 + 1442y_1^2 \\
& + 7269x_1^2y_2^2 + 4356x_1^2y_1^2 - 3825x_1^3y_2^2 - 180x_1^3y_1^2 + 2310x_1^4y_2^2 \\
& + 5013x_1x_2^3y_1^2 - 22950x_1^2x_2y_1y_2 - 1505x_1^4y_1y_2 - 4041x_1^3y_2^2 - 3010x_1^3x_2y_2^2.
\end{aligned}$$

Note that t is a polynomial in 4 variables of degree 6 that is quadratic in \tilde{y} . Let us denote the cone of sos polynomials in 4 variables (\tilde{x}, \tilde{y}) that have degree 6 and are quadratic in \tilde{y} by $\hat{\Sigma}_{4,6}$, and its dual cone by $\hat{\Sigma}_{4,6}^*$. Our proof will simply proceed by presenting a dual functional $\xi \in \hat{\Sigma}_{4,6}^*$ that takes a negative value on the polynomial t . We fix the following ordering of monomials in what follows:

$$\begin{aligned}
v = & [y_2^2, y_1y_2, y_1^2, x_2y_2^2, x_2y_1y_2, x_2y_1^2, x_2^2y_2^2, x_2^2y_1y_2, x_2^2y_1^2, x_2^3y_2^2, x_2^3y_1y_2, x_2^3y_1^2, x_2^4y_2^2, x_2^4y_1y_2, x_2^4y_1^2, \\
& x_1y_2^2, x_1y_1y_2, x_1y_1^2, x_1x_2y_2^2, x_1x_2y_1y_2, x_1x_2y_1^2, x_1x_2^2y_2^2, x_1x_2^2y_1y_2, x_1x_2^2y_1^2, x_1x_2^3y_2^2, x_1x_2^3y_1y_2, \\
& x_1x_2^3y_1^2, x_1^2y_2^2, x_1^2y_1y_2, x_1^2y_1^2, x_1^2x_2y_2^2, x_1^2x_2y_1y_2, x_1^2x_2y_1^2, x_1^2x_2^2y_2^2, x_1^2x_2^2y_1y_2, x_1^2x_2^2y_1^2, x_1^3y_2^2, \\
& x_1^3y_1y_2, x_1^3y_1^2, x_1^3x_2y_2^2, x_1^3x_2y_1y_2, x_1^3x_2y_1^2, x_1^4y_2^2, x_1^4y_1y_2, x_1^4y_1^2]^T.
\end{aligned} \tag{24}$$

Let \vec{t} represent the vector of coefficients of t ordered according to the list of monomials above; i.e., $t = \vec{t}^T v$. Using the same ordering, we can represent our dual functional ξ with the vector

$$\begin{aligned}
c = & [19338, -2485, 17155, 6219, -4461, 11202, 4290, -5745, 13748, 3304, -5404, 13227, 3594, \\
& -4776, 19284, 2060, 3506, 5116, 366, -2698, 6231, -487, -2324, 4607, 369, -3657, 3534, 6122, \\
& 659, 7057, 1646, 1238, 1752, 2797, -940, 4608, -200, 1577, -2030, -513, -3747, 2541, 15261, \\
& 220, 7834]^T.
\end{aligned}$$

We have

$$\langle \xi, t \rangle = c^T \vec{t} = -\frac{364547}{16} < 0.$$

On the other hand, we claim that $\xi \in \hat{\Sigma}_{4,6}^*$; i.e., for any form $w \in \hat{\Sigma}_{4,6}$, we should have

$$\langle \xi, w \rangle = c^T \vec{w} \geq 0, \tag{25}$$

where \vec{w} here denotes the coefficients of w listed according to the ordering in (24). Indeed, if w is sos, then it can be written in the form

$$w(x) = \tilde{z}^T \tilde{Q} \tilde{z} = \text{Tr } \tilde{Q} \cdot \tilde{z} \tilde{z}^T,$$

for some symmetric positive semidefinite matrix \tilde{Q} , and a vector of monomials

$$\tilde{z} = [y_2, y_1, x_2y_2, x_2y_1, x_1y_2, x_1y_1, x_2^2y_2, x_2^2y_1, x_1x_2y_2, x_1x_2y_1, x_1^2y_2, x_1^2y_1]^T.$$

It is not difficult to see that

$$c^T \vec{w} = \text{Tr } \tilde{Q} \cdot (\tilde{z}\tilde{z}^T)|_c, \quad (26)$$

where by $(\tilde{z}\tilde{z}^T)|_c$ we mean a matrix where each monomial in $\tilde{z}\tilde{z}^T$ is replaced with the corresponding element of the vector c . This yields the matrix

$$(\tilde{z}\tilde{z}^T)|_c = \begin{bmatrix} 19338 & -2485 & 6219 & -4461 & 2060 & 3506 & 4290 & -5745 & 366 & -2698 & 6122 & 659 \\ -2485 & 17155 & -4461 & 11202 & 3506 & 5116 & -5745 & 13748 & -2698 & 6231 & 659 & 7057 \\ 6219 & -4461 & 4290 & -5745 & 366 & -2698 & 3304 & -5404 & -487 & -2324 & 1646 & 1238 \\ -4461 & 11202 & -5745 & 13748 & -2698 & 6231 & -5404 & 13227 & -2324 & 4607 & 1238 & 1752 \\ 2060 & 3506 & 366 & -2698 & 6122 & 659 & -487 & -2324 & 1646 & 1238 & -200 & 1577 \\ 3506 & 5116 & -2698 & 6231 & 659 & 7057 & -2324 & 4607 & 1238 & 1752 & 1577 & -2030 \\ 4290 & -5745 & 3304 & -5404 & -487 & -2324 & 3594 & -4776 & 369 & -3657 & 2797 & -940 \\ -5745 & 13748 & -5404 & 13227 & -2324 & 4607 & -4776 & 19284 & -3657 & 3534 & -940 & 4608 \\ 366 & -2698 & -487 & -2324 & 1646 & 1238 & 369 & -3657 & 2797 & -940 & -513 & -3747 \\ -2698 & 6231 & -2324 & 4607 & 1238 & 1752 & -3657 & 3534 & -940 & 4608 & -3747 & 2541 \\ 6122 & 659 & 1646 & 1238 & -200 & 1577 & 2797 & -940 & -513 & -3747 & 15261 & 220 \\ 659 & 7057 & 1238 & 1752 & 1577 & -2030 & -940 & 4608 & -3747 & 2541 & 220 & 7834 \end{bmatrix},$$

which is positive definite. Therefore, equation (26) along with the fact that \tilde{Q} is positive semidefinite implies that (25) holds. This completes the proof.