# ALGEBRAIC DEGREE OF POLYNOMIAL OPTIMIZATION

JIAWANG NIE * AND KRISTIAN RANESTAD †

**Abstract.** Consider the polynomial optimization problem whose objective and constraints are all described by multivariate polynomials. Under some genericity assumptions, we prove that the optimality conditions always hold on optimizers, and the coordinates of optimizers are algebraic functions of the coefficients of the input polynomials. We also give a general formula for the algebraic degree of the optimal coordinates. The derivation of the algebraic degree is equivalent to counting the number of all complex critical points. As special cases, we obtain the algebraic degrees of quadratically constrained quadratic programming (QCQP), second order cone programming (SOCP) and $p$-th order cone programming (POCP), in analogy to the algebraic degree of semidefinite programming [9].

**Key words.** algebraic degree, polynomial optimization, optimality condition, quadratically constrained quadratic programming (QCQP), $p$-th order cone programming, second order cone programming (SOCP), variety

**AMS subject classifications.** 14N05, 14Q05, 14Q15, 65K05, 90C20, 90C46, 90C60

## 1. Introduction.

Consider the optimization problem

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x)=0,\, i=1,\ldots,m_e \\ & f_i(x)\geq 0,\, i=m_e+1,\ldots,m \end{cases} \tag{1.1}$$

where the $f_i$ are multivariate polynomial functions in $\mathbb{R}[x]$ (the ring of polynomials in $x=(x_1,\ldots,x_n)$ with real coefficients). The recent interest on solving polynomial optimization problems [7, 8, 11, 12] by using semidefinite relaxations or other algebraic methods motivates this study of the algebraic properties of the polynomial optimization problem (1.1). A fundamental question regarding (1.1) is how the optimal solutions depend on the input polynomials $f_i$. When the optimality condition holds and the critical equations of (1.1) have finitely many complex solutions, the optimal solutions are algebraic functions of the coefficients of the polynomials $f_i$, in particular, the coordinates of optimal solutions are roots of some univariate polynomials whose coefficients are functions of the input data. An interesting and important problem in optimization theory is to find the degrees of these algebraic functions as functions of the degrees of the $f_i$, which amounts to computing the number of complex solutions to the critical equations of (1.1). We begin our discussion with some special cases.

The simplest case of (1.1) is linear programming (LP), when all the polynomials $f_i$ have degree one. In this case, the problem (1.1) has the form (after removing the linear equality constraints)

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} & c^Tx \\ s.t. & Ax\geq b \end{cases} \tag{1.2}$$

where $c, A, b$ are matrices or vectors of appropriate dimensions. The feasible set of (1.2) is now a polyhedron described by some linear inequalities. As is well-known, if

*Department of Mathematics, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA (njw@math.ucsd.edu).

†Department of Mathematics, University of Oslo, PB 1053 Blindern, 0316 Oslo, Norway (ranestad@math.uio.no).

the polyhedron has a vertex and (1.2) has an optimal solution, then an optimizer $x^*$ of (1.2) occurs at a vertex. So $x^*$ can be determined by the linear system consisting of the active constraints. When the objective $c^T x$ is changing, the optimal solution might move from one vertex to another vertex. So the optimal solution is a piecewise linear fractional function of the input data $(c, A, b)$. When the $c, A, b$ are all rational, an optimal solution must also be rational, and hence its algebraic degree is one.

A more general convex optimization which is a proper generalization of linear programming is semidefinite programming (SDP). It has the standard form

$$\begin{cases} \min\limits_{x \in \mathbb{R}^n} & c^T x \\ s.t. & A_0 + \sum\limits_{i=1}^{n} x_i A_i \succeq 0 \end{cases} \tag{1.3}$$

where $c$ is a constant vector and the $A_i$ are constant symmetric matrices. The inequality $X \succeq 0$ means the matrix $X$ is positive semidefinite. Recently, Nie et al. [9] studied the algebraic properties of semidefinite programming. When $c$ and $A_i$ are generic, the optimal solution $x^*$ of (1.3) is shown in [9] to be a piecewise algebraic function of $c$ and $A_i$. Of course, the constraint of (1.3) can be replaced by the nonnegativity of all the principle minors of the constraint matrix, and hence (1.3) becomes a special case of (1.1). However, the problem (1.3) has very special nice properties, e.g., it is a convex program and the constraint matrix is linear with respect to $x$. Interestingly, if $c$ and $A_i$ are generic, the degree of each piece of this algebraic function only depends on the rank of the constraint matrix at the optimal solution. A formula for this degree is given in [13, Theorem 1.1].

Another optimization problem frequently used in statistics and biology is the Maximum Likelihood Estimation (MLE) problem. It has the standard form

$$\max_{x \in \Theta} \; p_1(x)^{u_1} p_2(x)^{u_2} \cdots p_m(x)^{u_m} \tag{1.4}$$

where $\Theta$ is an open subset of $\mathbb{R}^n$, the $p_i$ are polynomials such that $\sum_i p_i = 1$, and the $u_i$ are given positive integers. The optimizer $x^*$ is an algebraic function of $(u_1, \dots, u_n)$. This problem has recently been studied and a formula for the degree of this algebraic function has been found (cf. [1, 6]).

In this paper we consider the general optimization problem (1.1) when the polynomials $f_0, f_1, \dots, f_m$ define a *complete intersection*, i.e., their common set of zeros has codimension $m+1$ (see the appendix for the definition of complete intersection). We show that an optimal solution is an algebraic function of the input data. We call the degree of this algebraic function the *algebraic degree* of the polynomial optimization problem (1.1). Equivalently, the algebraic degree equals the number of complex solutions to the critical equations of (1.1) when this number is finite. Under some genericity assumptions, we give in this paper a formula for the algebraic degree of (1.1).

Throughout this paper, the words "generic" and "genericity" are frequently used. We shall use them as conditions on the input data for some property to hold, and they shall mean for all but a set of Lebesgue measure zero in the space of data.

The algebraic degree of polynomial optimization (1.1) addresses the computational complexity at a fundamental level. To solve (1.1) exactly essentially reduces to solving some univariate polynomial equations whose degrees are the algebraic degree of (1.1). As we will see, the algebraic degree grows rapidly with the degrees of the $f_i$.

The paper is organized as follows: In Section 2 we derive a general formula for the algebraic degree, and in Section 3 we give the formulae of the algebraic degrees for special cases like quadratically constrained quadratic programming, second order cone programming, and $p$-th order cone programming. The paper concludes with an appendix which introduces some basic concepts and facts in algebraic geometry that are necessary for this paper.

**2. A general formula for the algebraic degree.** In this section we will derive a formula for the algebraic degree of the polynomial optimization problem (1.1) when the polynomials define a complete intersection. Suppose the polynomial $f_i$ has degree $d_i$. Let $x^*$ be a local or global optimal solution of (1.1). At first, we assume that the polynomials are general and all the inequality constraints are active, i.e., $m_e = m$. When $m = n$, by Corollary A.2, the feasible set of (1.1) is finite and hence the algebraic degree is equal to the product of the degrees of the polynomials, $d_1 d_2 \cdots d_m$. So, from now on, assume $m < n$. If the variety

$$V = \{x \in \mathbb{C}^n : f_1(x) = \cdots = f_m(x) = 0\}$$

is smooth at $x^*$, i.e., the gradient vectors

$$\nabla f_1(x^*), \nabla f_2(x^*), \ldots, \nabla f_m(x^*)$$

are linearly independent, then the *Karush-Kuhn-Tucker* (KKT) condition holds at $x^*$ (Chapter 12 in [10]). In fact,

$$\begin{cases} \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0 \\ f_1(x^*) = \cdots = f_m(x^*) = 0 \end{cases} \tag{2.1}$$

where $\lambda_1^*, \ldots, \lambda_m^*$ are Lagrange multipliers for the constraints $f_1(x) = 0, \ldots, f_m(x) = 0$. Thus the optimal solution $x^*$ and the Lagrange multipliers $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ are determined by the polynomial system (2.1). The set of complex solutions to (2.1) forms the locus of complex critical points of (1.1). If the system (2.1) is zero-dimensional, then, by *elimination theory* (cf. [2, Chapter 3]), the coordinates of $x^*$ are algebraic functions of the coefficients of the polynomials $f_i$. Each coordinate $x_i^*$ can be determined by some univariate polynomial equation of the form

$$(x_i^*)^{\delta_i} + a_1 (x_i^*)^{\delta_i - 1} + \cdots + a_{\delta_i - 1} x_i^* + a_{\delta_i} = 0$$

where $a_j$ are rational functions of the coefficients of the $f_i$. Interestingly, when $f_1, f_2, \ldots, f_m$ are generic, the KKT condition always holds at any optimal solution, and the degrees $\delta_i$ are equal to each other. This common degree counts the number of complex solutions to (2.1), i.e., the cardinality of the complex critical locus of (1.1) or, by definition, the algebraic degree of the polynomial optimization (1.1). We will derive a general formula for this degree.

We turn to complex projective spaces, where the above question may be answered as a problem in intersection theory. For this we translate the optimization problem to a relevant intersection problem. Let $\mathbb{P}^n$ be the $n$-dimensional complex projective space. A point $\tilde{x} \in \mathbb{P}^n$ has homogeneous coordinates $[x_0, x_1, \ldots, x_n]$ unique up to multiplication by a common nonzero scalar. A variety in $\mathbb{P}^n$ is a set of points $\tilde{x}$ that satisfy a collection of homogeneous polynomial equations in $[x_0, x_1, \ldots, x_n]$. Let $\tilde{f}_i$, defined by $\tilde{f}_i(\tilde{x}) = x_0^{d_i} f_i(x_1/x_0, \ldots, x_n/x_0)$, be the homogenization of $f_i$. Define $\mathcal{U}$ to be the projective variety

$$\mathcal{U} = \{\tilde{x} \in \mathbb{P}^n : \tilde{f}_1(\tilde{x}) = \tilde{f}_2(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0\}$$

in $\mathbb{P}^n$. See the appendix for more about projective spaces and projective varieties. Next, let

$$\tilde{\nabla}\tilde{f}_i = \begin{bmatrix} \frac{\partial}{\partial x_0}\tilde{f}_i & \dots & \frac{\partial}{\partial x_n}\tilde{f}_i \end{bmatrix}^T$$

be the gradient vector with respect to the homogeneous coordinates. Notice that $(\frac{\partial}{\partial x_j}\tilde{f}_i(\tilde{x}) = x_0^{d_i-1}\frac{\partial}{\partial x_j}f_i(x_1/x_0,\dots,x_n/x_0))$, so the homogenization of $\nabla f_i$ coincides with the last $n$ coordinates $\nabla\tilde{f}_i$ in $\tilde{\nabla}\tilde{f}_i$.

In this homogeneous setting, the optimality condition for problem (1.1) with $m = m_e$ is

$$\left\{ (x,\mu) \in \mathbb{C}^n \times \mathbb{C} : \begin{array}{l} \tilde{f}_0(\tilde{x}) - \mu x_0^{d_0} = \tilde{f}_1(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0 \\ \operatorname{rank}\big[\tilde{\nabla}(\tilde{f}_0(\tilde{x}) + \mu x_0^{d_0}), \tilde{\nabla}(\tilde{f}_1(\tilde{x})), \dots, \tilde{\nabla}(\tilde{f}_m(\tilde{x}))\big] \le m \end{array} \right\} \tag{2.2}$$

where $\mu \in \mathbb{R}$ is the critical value. Let $\tilde{x}^* \in \{x_0 \neq 0\}$ be a critical point, i.e., a solution to (2.2). We may eliminate $\mu$ by asking that the matrix

$$\begin{bmatrix} \tilde{f}_0(\tilde{x}^*) & \tilde{f}_1(\tilde{x}^*) & \dots & \tilde{f}_m(\tilde{x}^*) \\ x_0^{d_0} & 0 & \dots & 0 \end{bmatrix}$$

have rank at most one and that the matrix

$$\begin{bmatrix} \frac{\partial}{\partial x_0}\tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_0}\tilde{f}_1(\tilde{x}^*) & \dots & \frac{\partial}{\partial x_0}\tilde{f}_m(\tilde{x}^*) & (d_0-1)x_0^{d_0} \\ \frac{\partial}{\partial x_1}\tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_1}\tilde{f}_1(\tilde{x}^*) & \dots & \frac{\partial}{\partial x_1}\tilde{f}_m(\tilde{x}^*) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_n}\tilde{f}_0(\tilde{x}^*) & \frac{\partial}{\partial x_n}\tilde{f}_1(\tilde{x}^*) & \dots & \frac{\partial}{\partial x_n}\tilde{f}_m(\tilde{x}^*) & 0 \end{bmatrix}$$

have rank at most $m+1$. Since $x_0 \neq 0$, the first condition implies that

$$\tilde{x}^* \in \mathcal{U} = \{\tilde{f}_1(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0\}.$$

Similarly, the rank of the second matrix equals $m+1$ only if the submatrix

$$M = \begin{bmatrix} \frac{\partial}{\partial x_1}\tilde{f}_0(\tilde{x}) & \frac{\partial}{\partial x_1}\tilde{f}_1(\tilde{x}) & \dots & \frac{\partial}{\partial x_1}\tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_n}\tilde{f}_0(\tilde{x}) & \frac{\partial}{\partial x_n}\tilde{f}_1(\tilde{x}) & \dots & \frac{\partial}{\partial x_n}\tilde{f}_m(\tilde{x}) \end{bmatrix}$$

has rank $m$. Therefore we define $\mathcal{W}$ to be the projective variety in $\mathbb{P}^n$

$$\mathcal{W} = \{\tilde{x} \in \mathbb{P}^n : \text{ all the } (m+1)\times(m+1) \text{ minors of } M \text{ vanish } \},$$

i.e., the locus of points where the rank of $[\nabla\tilde{f}_0(\tilde{x}),\dots,\nabla\tilde{f}_m(\tilde{x})]$ is less than or equal to $m$.

Proposition 2.1. *Consider the polynomial optimization problem (1.1), and assume that $m = m_e$, i.e., that all constraints are active. If the polynomials $f_1,\dots,f_m$ are generic, then we have:*

*(i) The affine variety $V = \{x \in \mathbb{C}^n : f_1(x) = \cdots = f_m(x) = 0\}$ is smooth;*

*(ii) The KKT condition holds at any optimal solution $x^*$;*

*(iii) If $f_0$ is also generic, the affine variety*

$$K = \left\{ x \in V : \exists\, \lambda_1, \ldots, \lambda_m \text{ such that } \nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0 \right\} \tag{2.3}$$

*defined by the KKT system (2.1) is finite.*

*Proof.* (i) When the polynomials $f_1, \ldots, f_m$ are generic, then their homogenizations $\tilde{f}_1, \ldots, \tilde{f}_m$ are generic, so by Corollary A.2, they define a smooth complete intersection variety $\mathcal{U}$ of codimension $m$. In particular the affine subvariety $V = \mathcal{U} \cap \{x_0 \neq 0\}$ is smooth. Therefore the Jacobian matrix

$$\begin{bmatrix} \frac{\partial}{\partial x_0}\tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_0}\tilde{f}_2(\tilde{x}) & \ldots & \frac{\partial}{\partial x_0}\tilde{f}_m(\tilde{x}) \\ \frac{\partial}{\partial x_1}\tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_1}\tilde{f}_2(\tilde{x}) & \ldots & \frac{\partial}{\partial x_1}\tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_n}\tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_n}\tilde{f}_2(\tilde{x}) & \ldots & \frac{\partial}{\partial x_n}\tilde{f}_m(\tilde{x}) \end{bmatrix}$$

has full rank at $\tilde{x}$. Furthermore, the tangent space of $V$ at $\tilde{x}$ is, of course, not contained in the hyperplane $x_0 = 0$ at infinity, so the column $\begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}^T$ is not in the column space of the matrix at $\tilde{x}$. Thus already the submatrix

$$\begin{bmatrix} \frac{\partial}{\partial x_1}\tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_1}\tilde{f}_2(\tilde{x}) & \ldots & \frac{\partial}{\partial x_1}\tilde{f}_m(\tilde{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_n}\tilde{f}_1(\tilde{x}) & \frac{\partial}{\partial x_n}\tilde{f}_2(\tilde{x}) & \ldots & \frac{\partial}{\partial x_n}\tilde{f}_m(\tilde{x}) \end{bmatrix}$$

has full rank, i.e., the gradients

$$\nabla \tilde{f}_1(\tilde{x}), \ \ldots, \ \nabla \tilde{f}_m(\tilde{x})$$

are linearly independent at $\tilde{x} \in V$.

(ii) An optimizer $x^*$ must belong to $V$, and by (i), the gradients

$$\nabla f_1(x^*), \nabla f_2(x^*), \ldots, \nabla f_m(x^*)$$

are linearly independent. Hence the KKT condition holds at $x^*$ (cf. [10, Chapter 12]).

(iii) We claim that the intersection of the complete intersection $\mathcal{U}$ and the variety $\mathcal{W}$ where the Jacobian matrix $M$ has rank at most $m$ is finite. Since our critical points $V \cap \mathcal{W}$ form a subset of $\mathcal{U} \cap \mathcal{W}$, (iii) would follow. The codimension of $\mathcal{U}$ is $m$, and this complete intersection variety is smooth, so the matrix $M$ has, by (i), rank at least $m$ at each point of $\mathcal{U}$. The variety $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$ is, by Bertini's Theorem A.1, also smooth. So as above, the matrix $M$ has full rank at the points in the affine part $V \cap \{f_0(x) = 0\}$. On the other hand, $M$ is the Jacobian matrix of the variety $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. This variety is again smooth and has codimension $m+1$ in the hyperplane $\{x_0 = 0\}$, so $M$ must have full rank $m+1$ on $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. Therefore the variety $\mathcal{W}$, where $M$ has rank at most $m$, cannot intersect $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$. But Bézout's Theorem A.3 says that if the sum of the codimensions of two varieties in $\mathbb{P}^n$ does not exceed $n$, then they must intersect. In particular, any curve in $\mathcal{U}$ intersects the hypersurface $\{\tilde{f}_0(\tilde{x}) = 0\}$. Since $\mathcal{U} \cap \{\tilde{f}_0(\tilde{x}) = 0\}$ has codimension $m+1$, we deduce that $\mathcal{W}$ must have codimension at least $n - m$. Furthermore, since any curve in $\mathcal{U} \cap \mathcal{W}$ would intersect $\{\tilde{f}_0(\tilde{x}) = 0\}$, the intersection $\mathcal{U} \cap \mathcal{W}$ must be empty or finite.

On the other hand, the variety of $n \times (m+1)$-matrices having rank no more than $m$, whose entries are homogeneous polynomials in the coordinates of points in a projective space, has codimension at most $n-m$, by Proposition A.5. So the codimension of $\mathcal{W}$ equals $n-m$. Hence $\mathcal{U}$ and $\mathcal{W}$ have complementary dimensions. Therefore the intersection $\mathcal{U} \cap \mathcal{W}$ is non-empty and (iii) follows. ∎

By Proposition 2.1, for generic polynomials $f_1, \ldots, f_m$, the optimal solutions of (1.1) can be characterized by the KKT system (2.1), and for generic objective function $f_0$ the KKT variety $K$ is finite. Geometrically, the algebraic degree of the optimization problem (1.1), under the genericity assumption, is equal to the number of distinct complex solutions of KKT, i.e., the cardinality of the variety $K$. We showed above that $K$ coincides with $V \cap \mathcal{W}$. The variety $\mathcal{U} \cap \mathcal{W}$ above clearly contains $K$. On the other hand, $\mathcal{U} \cap \mathcal{W}$ is finite and does not intersect the hyperplane $\{x_0 = 0\}$ when the polynomials $f_i$ are generic. Since $\mathcal{U} \backslash V = \mathcal{U} \cap \{x_0 = 0\}$ and $\mathcal{U} \cap \mathcal{W} \cap \{x_0 = 0\} = \emptyset$, we may conclude that $K = \mathcal{U} \cap \mathcal{W}$. Therefore the cardinality of $K$ must coincide with the degree of the variety $\mathcal{U} \cap \mathcal{W}$.

Let $S_r$ be the $r$-th complete symmetric function on $k$ letters $(n_1, n_2, \ldots, n_k)$:

$$S_r(n_1, n_2, \ldots, n_k) = \sum_{i_1+i_2+\cdots+i_k=r} n_1^{i_1} \cdots n_k^{i_k}. \tag{2.4}$$

Theorem 2.2. *Consider the polynomial optimization problem (1.1), and assume $m = m_e$, i.e., all contraints are active. If the polynomials $f_0, f_1, \ldots, f_m$ are generic, then the algebraic degree of (1.1) is equal to*

$$d_1 d_2 \cdots d_m S_{n-m}(d_0 - 1, d_1 - 1, \ldots, d_m - 1).$$

*Furthermore, if the polynomials $f_0, f_1, \ldots, f_m$ are not generic, but the system (2.1) is still zero-dimensional, then the above formula is an upper bound for the algebraic degree of (1.1).*

*Proof.* When $f_1, f_2, \ldots, f_m$ are generic, $\mathcal{U}$ is a smooth complete intersection of codimension $m$. Its degree $\deg(\mathcal{U}) = d_1 d_2 \cdots d_m$ by Corollary A.4. When $f_0$ is also generic, $\mathcal{W}$ has codimension $n-m$ and intersects $\mathcal{U}$ in a finite set of points according to Proposition 2.1. If the intersection $\mathcal{U} \cap \mathcal{W}$ is transverse (i.e., smooth) and hence consists of a collection of simple points, then the degree $\deg\big(\mathcal{U} \cap \mathcal{W}\big)$ counts the number of intersection points of $\mathcal{U} \cap \mathcal{W}$, and hence the cardinality of the KKT variety $K$, which is also the number of complex solutions to the KKT system (2.1) for problem (1.1).

To show that this intersection is transversal, we consider the subvariety $X$ in $\mathbb{P}^n \times \mathbb{P}^m$ defined by the $m$ equations $\tilde{f}_1(\tilde{x}) = \tilde{f}_2(\tilde{x}) = \cdots = \tilde{f}_m(\tilde{x}) = 0$ and the $n$ equations

$$M \cdot (\lambda_0, \ldots, \lambda_m)^T = 0,$$

where the $\lambda_i$ are homogeneous coordinate functions in the second factor. The image under the projection of the variety $X$ defined by these $m+n$ polynomials into the first factor coincides with the finite set $\mathcal{U} \cap \mathcal{W}$. Since $M$ has rank at least $m$ at every point of $\mathcal{U}$, there is a unique $\tilde{\lambda} = (\lambda_0, \ldots, \lambda_m) \in \mathbb{P}^m$ for each point $\tilde{x} \in \mathcal{U} \cap \mathcal{W}$ such that $(\tilde{x}, \tilde{\lambda})$ lies in $X$. Therefore the variety $X$ is a complete intersection in $\mathbb{P}^n \times \mathbb{P}^m$. When the coefficients of $f_0$ vary, it is easy to check that this complete intersection does not have any fixed point. This is because when the coefficients of $f_0$ vary, the common zeros of the $n$ equations $M \cdot (\lambda_0, \ldots, \lambda_m)^T = 0$ vary without fixed points. So Bertini's Theorem A.1 applies to conclude that for generic $f_0$ this complete intersection is transversal, which implies that the intersection $\mathcal{U} \cap \mathcal{W}$ in $\mathbb{P}^n$ is also transversal.

Since the intersection $\mathcal{U} \cap \mathcal{W}$ is finite, i.e., it has codimension in $\mathbb{P}^n$ equal to the sum of the codimensions of $\mathcal{U}$ and $\mathcal{W}$, Bézout's Theorem A.3 applies to compute the degree

$$\deg(\mathcal{U} \cap \mathcal{W}) = \deg(\mathcal{U}) \cdot \deg(\mathcal{W}).$$

To complete the computation, we therefore need to find $\deg(\mathcal{W})$. Since the codimension of $\mathcal{W}$ equals the codimension of the variety defined by the $(m+1) \times (m+1)$ minors of a general $n \times (m+1)$ matrix with polynomial entries, the formula in Proposition A.6 applies to compute this degree: The degree of $\mathcal{W}$ equals the degree of the determinantal variety of $n \times (m+1)$ matrices of rank at most $m$ in the space of matrices whose entries in the $i$-th column are generic forms of degree $d_i - 1$, i.e., $S_{n-m}(d_0 - 1, d_1 - 1, \ldots, d_m - 1)$. Therefore the degree formula for the critical locus $\mathcal{U} \cap \mathcal{W}$ and hence the algebraic degree of (1.1) is proved.

Assume that the polynomials $f_i$ are not generic, while the system (2.1) is still zero-dimensional. Then a perturbation argument can be applied. Let $x^*$ be one fixed optimal solution of optimization problem (1.1). Apply a generic perturbation $\Delta_\epsilon f_i$ to each $f_i$ so that $(f_i + \Delta_\epsilon f_i)(x)$ is a generic polynomial and the coefficients of $\Delta_\epsilon f_i$ tends to zero as $\epsilon \to 0$. Then one optimal solution $x^*(\epsilon)$ of the perturbed optimization problem (1.1) tends to $x^*$. By genericity of $(f_i + \Delta_\epsilon f_i)(x)$, we know

$$a_0(\epsilon)(x_i^*(\epsilon))^\delta + a_1(\epsilon)(x_i^*(\epsilon))^{\delta-1} + \cdots + a_{\delta-1}(\epsilon)x_i^*(\epsilon) + a_\delta(\epsilon) = 0.$$

Here $\delta = d_1 d_2 \cdots d_m S_{n-m}(d_0 - 1, d_1 - 1, \ldots, d_m - 1)$ and $a_j(\epsilon)$ are rational functions of the coefficients of $f_i$ and $\Delta_\epsilon f_i$. Without loss of generality, we may normalize $a_j(\epsilon)$ such that

$$\max_{0 \le j \le \delta} \; |a_j(\epsilon)| = 1.$$

When $\epsilon \to 0$, by continuity, $x_i^*$ is a root of some univariate polynomial whose degree is at most $\delta$ and coefficients are rational functions of the coefficients of polynomials $f_0, f_1, \ldots, f_m$. $\square$

Remark 2.3. *The genericity assumption in Theorem 2.2 is used to conclude that the critical locus $\mathcal{U} \cap \mathcal{W}$ is a smooth zero-dimensional variety, i.e., a set of points, by appealing to Bertini's Theorem A.1, while Bézout's Theorem A.3 counts its degree, i.e., the number of points. So both theorems are needed to get the sharp degree bound. A sufficient condition for Bertini's Theorem to apply can be expressed in terms of the sets $U_i$ of polynomials in which the polynomials $f_0, f_1, \ldots, f_m$ can be freely chosen. First, assume that the generic polynomial in each $U_i$ is reduced, and that $U_i$ intersects every Zariski open set of a complex affine space $V_i$. Second, assume that the set of common zeros of all the polynomials in $\cup_{i=0}^m V_i$ is empty. Then Bertini's Theorem applies. In fact, the polynomials $f_i$ for which the conclusion of Bertini's Theorem fails are contained in a complex subvariety of $V_i$.*

*If some of the polynomials $f_i$ are reducible, then we may replace $f_i$ by the factor of least degree that contains the optimizer. The original problem (1.1), is then modified to one with a smaller algebraic degree. This is relevant in the above context, if the generic polynomial in $U_i$ is reducible.*

Example 2.4. *Consider the following special case of problem (1.1)*

$$f_0(x) = 21x_2^2 - 92x_1x_3^2 - 70x_2^2x_3 - 95x_1^4 - 47x_1x_3^3 + 51x_2^2x_3^2 + 47x_1^5 + 5x_1x_2^4 + 33x_3^5,$$

$$f_1(x) = 88x_1 + 64x_1x_2 - 22x_1x_3 - 37x_2^2 + 68x_1x_2^2x_3 - 84x_2^4 + 80x_2^3x_3 + 23x_2^2x_3^2 - 20x_2x_3^3 - 7x_3^4,$$

$$f_2(x) = 31 - 45x_1x_2 + 24x_1x_3 - 75x_3^2 + 16x_1^3 - 44x_1^2x_3 - 70x_1x_2^2 - 23x_1x_2x_3 - 67x_2^2x_3 - 97x_2x_3^2.$$

*Here $m = m_e = 2$. By Theorem 2.2, the algebraic degree of the optimal solution is bounded by*

$$4 \cdot 3 \cdot S_1(4, 3, 2) = 12 \cdot (4 + 3 + 2) = 108.$$

*A symbolic computation over the finite field $\mathbb{Z}/17\mathbb{Z}$, using Singular [5], shows the optimal coordinate $x_1$ is a root of a univariate polynomial of degree 108. In this case the degree bound 108 is sharp. We were not able to find the exact coefficients of this univariate polynomial in the rational field $\mathbb{Q}$, since Singular could not complete the computation over $\mathbb{Q}$.*

Now we consider the more general case when $m > m_e$, i.e., there are inequality constraints. Then a similar degree formula as in Theorem 2.2 can be obtained, as soon as the active set is identified.

Corollary 2.5. *Consider the polynomial optimization problem (1.1). Let $x^*$ be an optimizer and let $j_1, \dots, j_k \subset \{m_e + 1, \dots, m\}$ be the indices of the active set of inequality constraints. If every active $f_i$ is generic, then the algebraic degree of $x^*$ is*

$$d_1 \cdots d_{m_e} d_{j_1} \cdots d_{j_k} S_{n-m_e-k}(d_0 - 1, d_1 - 1, \dots, d_{m_e} - 1, d_{j_1} - 1, \dots, d_{j_k} - 1).$$

*If at least one of polynomials $f_i$ is not generic and the system (2.1) is zero-dimensional, then the above formula is an upper bound for the algebraic degree.*

*Proof.* Note that $x^*$ is also an optimal solution of the polynomial optimization problem

$$\left.\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f_0(x) \\ s.t. & f_i(x) = 0,\ i = 1, \dots, m_e \\ & f_i(x) = 0,\ i = j_1, \dots, j_k \end{array}\right\}.$$

Hence the conclusion follows from Theorem 2.2. □

**3. Some special cases.** In this section we derive the algebraic degrees of some special polynomial optimization problems. The simplest special case is that all the polynomials $f_i$ in (1.1) have degree one, i.e., (1.1) becomes a linear programming problem of the form (1.2). If the objective $c$ is generic, precisely $n$ constraints will be active. So the algebraic degree is $S_0(0, 0, \dots, 0) = 1$. This is consistent with what we observed in the introduction. Now let us look at other special cases.

**3.1. Unconstrained optimization.** We consider the special case that the problem (1.1) has no constraints. It becomes an unconstrained optimization. The gradient of the objective vanishes at any optimal solution. By Theorem 2.2, the algebraic degree is bounded by $S_n(d_0 - 1) = (d_0 - 1)^n$, which is exactly Bézout's number for the gradient polynomial system

$$\nabla f_0(x) = 0.$$

Since $f_0$ can be chosen freely among all polynomials of degree $d_0$, Remark 2.3 applies to show that the degree bound above is sharp.

Example 3.1. *Consider the minimization of $f_0(x)$ given by*

$$\begin{aligned} f_0 = {} & x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_1^3 + x_2^3 + x_3^3 + x_4^3 - 13x_1^2 - 30x_1x_2 - 9x_1x_3 + 5x_1x_4 + 11x_2^2 \\ & - 3x_3x_2 - 3x_3^2 - 20x_3x_4 - 13x_2x_4 - 9x_4^2 + x_1 - 2x_2 + 12x_3 - 13x_4. \end{aligned}$$

*For the above polynomial, the algebraic degree of the optimal solution is* $3^4 = 81$. *A symbolic computation over the rational field* $\mathbb{Q}$, *using Singular [5], shows that the optimal coordinate* $x_1$ *of* $x^*$ *is a root of a univariate polynomial of degree* 81*:*

$$\begin{aligned} &9671406556917033397649408x_1^{81} + 19584598277756992630240051 2x_1^{80} + \cdots \\ &\quad +380685779511377249784195216850334660205275442369474081 28x_1 \\ &\quad -2957438647420262596596093352763852215662185651072180992. \end{aligned}$$

*The degree bound* 81 *is sharp for this problem.*

**3.2. Quadratically constrained quadratic programming.** Consider the special case that all the polynomials $f_0, f_1, \dots, f_m$ are quadratic. Then (1.1) becomes a quadratically constrained quadratic programming (QCQP) problem which has the standard form

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & x^T A_0 x + b_0^T x + c_0 \\ s.t. \quad & x^T A_i x + b_i^T x + c_i \geq 0,\ i = 1, \dots, \ell. \end{aligned}$$

Here $A_i, b_i, c_i$ are matrices or vectors of appropriate dimensions. The objective and all the constraints are all quadratic polynomials. At an optimal solution, suppose $m \leq \ell$ constraints are active. By Corollary 2.5, the algebraic degree is bounded by

$$2^m \cdot S_{n-m}(1, \underbrace{1, \dots, 1}_{m \text{ times}}) = 2^m \cdot \sum_{i_0+i_1+i_2+\cdots+i_m=n-m} 1 = 2^m \cdot \binom{n}{m}. \tag{3.1}$$

The polynomials $f_0, f_1, \dots, f_m$ can be chosen freely in the space of quadratic polynomials, so Remark 2.3 applies to show that the degree bound above is sharp.

Example 3.2. *Consider the polynomials*

$$\begin{aligned} f_0 &= -20 - 27x_1^2 + 89x_1x_2 + 80x_1x_3 - 45x_1x_4 + 19x_1x_5 + 42x_1 - 13x_2^2 + 31x_2x_3 - 79x_2x_4 \\ &\quad + 74x_2x_5 - 9x_2 + 56x_3^2 - 77x_3x_4 - 2x_3x_5 + 35x_3 + 40x_4^2 - 13x_4x_5 + 60x_4 + 58x_5^2 - 84x_5, \\ f_1 &= 33 + 55x_1^2 - 41x_1x_2 + 33x_1x_3 - 61x_1x_4 + 96x_1x_5 + 12x_1 + 74x_2^2 - 90x_2x_3 - 57x_2x_4 \\ &\quad - 52x_2x_5 + 51x_2 + 15x_3^2 + 81x_3x_4 + 87x_3x_5 + 75x_3 - 10x_4^2 + 58x_4x_5 + 33x_4 + 83x_5^2 - 23x_5, \\ f_2 &= 8 - 9x_1^2 + 56x_1x_2 - 24x_1x_3 + 81x_1x_4 + 85x_1x_5 - 99x_1 - 77x_2^2 - 75x_2x_3 + x_2x_4 + 38x_2x_5 \\ &\quad + 23x_2 - 97x_3^2 - 14x_3x_4 - 73x_3x_5 + 65x_3 + 3x_4^2 - 14x_4x_5 + 16x_4 + 9x_5^2 - 10x_5, \\ f_3 &= 9 + 90x_1^2 - 94x_1x_2 - 22x_1x_3 - 24x_1x_4 + 78x_1 + 32x_2^2 - 48x_2x_3 - 6x_2x_4 + 80x_2x_5 - 18x_2 \\ &\quad - 63x_3^2 + 66x_3x_4 - 13x_3x_5 + 88x_3 - 45x_4^2 - 92x_4x_5 - 69x_4 - 43x_5^2 + 32x_5. \end{aligned}$$

*For the above polynomials, the QCQP problem is nonconvex. We consider those local optimal solutions for which all the three inequalities are active. By Corollary 2.5, the algebraic degree of this problem is bounded by* $2^m\binom{n}{m} = 80$. *A symbolic computation over the finite field* $\mathbb{Z}/17\mathbb{Z}$, *using Singular [5], shows that the optimal coordinate* $x_1$ *is a root of a univariate polynomial of degree* 80. *The algebraic degree of this problem is* 80 *and the bound given by the formula (3.1) is sharp. For this example, we were also not able to find the exact coefficients of this univariate polynomial in the rational field* $\mathbb{Q}$, *since Singular could not complete the computation over* $\mathbb{Q}$.

**3.3. Second order cone programming.** The second order cone programming (SOCP) problem has the standard form

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & c^T x \\ s.t. & a_i^T x + b_i - \|C_i x + d_i\|_2 \geq 0,\ i = 1,\ldots,\ell \end{array} \tag{3.2}$$

where $c, a_i, b_i, C_i, d_i$ are matrices or vectors of appropriate dimensions. Let $x^*$ be an optimizer. Since SOCP is a convex program, the $x^*$ must also be a global solution. By removing the square root in the constraint, SOCP becomes the polynomial optimization problem

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & c^T x \\ s.t. & (a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i) \geq 0,\ i = 1,\ldots,\ell. \end{array}$$

Without loss of generality, assume that the constraints with indices $1, 2, \ldots, m$ are active at $x^*$. The objective is linear but the constraints are all quadratic. As we can see, the Hessian of the constraints have the special form $a_i a_i^T - C_i^T C_i$. Let $r_i$ be the number of rows of $C_i$. When $r_i = 1$, the constraint $a_i^T x + b_i - \|C_i x + d_i\|_2 \geq 0$ is equivalent to two linear constraints

$$-(a_i^T x + b_i) \leq C_i x + d_i \leq a_i^T x + b_i.$$

Thus, when every $r_i = 1$, the problem reduces to a linear programming problem and hence has algebraic degree one, because in this situation the polynomial $(a_i^T x + b_i)^2 - (C_i x + d_i)^2$ is reducible. When $r_i \geq 2$ and $a_i, b_i, C_i, d_i$ are generic, the polynomial $(a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i)$ is quadratic of rank $r_i + 1$ and hence irreducible. Without loss of generality, assume $1 = r_1 = r_2 = \cdots = r_k < r_{k+1} \leq \cdots \leq r_m$. Then the problem (3.2) is reduced to

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & c^T x \\ s.t. & a_i^T x + b_i + \sigma_i (C_i x + d_i) \geq 0,\ i = 1,\ldots,k \\ & (a_i^T x + b_i)^2 - (C_i x + d_i)^T (C_i x + d_i) \geq 0,\ i = k+1,\ldots,m \end{array}$$

where the scalar $\sigma_i$ is chosen such that $a_i^T x^* + b_i + \sigma_i (C_i x^* + d_i) = 0$. By Corollary 2.5, the algebraic degree of SOCP in this modified form is bounded by

$$2^{m-k} \cdot S_{n-m}(0,\ldots,0,\underbrace{1,\ldots,1}_{m-k \text{ times}}) = 2^{m-k} \cdot \sum_{i_{k+1}+i_{k+2}+\cdots+i_m = n-m} 1 = 2^{m-k} \cdot \binom{n-k-1}{m-k-1}. \tag{3.3}$$

When $k = m$, we have already seen that the algebraic degree is one.

For the sharpness of degree bound (3.3), we apply Bertini's Theorem A.1 following Remark 2.3. For every $i = k+1,\ldots,m$, define the set $U_i$ of polynomials as

$$U_i = \left\{ (a_i^T x + b_i)^2 - \sum_{1\leq j\leq r_i} \alpha_j^2 (C_i x + d_i)_j^2 :\ \alpha_1,\ldots,\alpha_{r_i} \in \mathbb{R} \right\}.$$

Next, define complex affine spaces $V_i$ as follows:

$$V_i = \left\{ (a_i^T x + b_i)^2 - \sum_{1\leq j\leq r_i} \beta_j (C_i x + d_i)_j^2 :\ \beta_1,\ldots,\beta_{r_i} \in \mathbb{C} \right\},\ i = k+1,\ldots,m.$$

Then every set $U_i$ intersects any Zariski open subset of the affine space $V_i$. On the other hand the set of common zeros of the linear polynomials

$$a_i^T x + b_i + \sigma_i(C_i x + d_i),\ i = 1, \ldots, k$$

and all the polynomials in the union $\bigcup_{i=k+1}^m V_i$ is contained in the set

$$Z = \bigcap_{i=1}^{k} \left\{x \in \mathbb{R}^n : a_i^T x + b_i + \sigma_i(C_i x + d_i) = 0\right\} \bigcap_{i=k+1}^{m} \left\{x \in \mathbb{R}^n : \begin{array}{l} a_i^T x + b_i = 0 \\ C_i x + d_i = 0 \end{array}\right\}. \tag{3.4}$$

Therefore, for generic choices $a_i, b_i, C_i, d_i$, if $r_{k+1} + \cdots + r_m + m > n$, the set $Z$ is empty. Hence Remark 2.3 applies to show that, for generic choices of $c, a_i, b_i, C_i, d_i$ with $r_{k+1} + \cdots + r_m + m > n$, the algebraic degree bound $2^{m-k} \cdot \binom{n-k-1}{m-k-1}$ is sharp.

EXAMPLE 3.3. *Consider the SOCP defined by the polynomials*

$$\begin{aligned}
f_0 &= -x_1 + 6x_2 + 13x_3 + 11x_4 + 8x_5,\\
f_1 &= (-11x_1 - 18x_2 - 4x_3 + 2x_4 - 12x_5 + 7)^2 - (-4x_1 - 10x_2 + 20x_3 - 4x_4 - 9x_5 + 3)^2\\
&\quad - (-5x_1 - 11x_2 + 8x_3 - 18x_4 + 11x_5 + 15)^2 - (21x_1 + 18x_2 - 12x_3 - 10x_4 - 8x_5 + 4)^2,\\
f_2 &= (-5x_1 - 5x_2 - 7x_3 - 6x_4 + 4x_5 + 41)^2 - (x_1 - 2x_2 + 10x_3 - 21x_4 - 11)^2\\
&\quad - (-12x_1 + 3x_2 + 16x_3 + 4x_4 + x_5 + 9)^2 - (14x_1 + 20x_2 - 13x_3 - 7x_4 + 4x_5 + 2)^2,\\
f_3 &= (x_1 - 8x_2 + 11x_3 - x_5 + 22)^2 - (2x_1 - x_2 + 3x_3 - x_4 - 25x_5 - 8)^2\\
&\quad - (-2x_1 - 17x_3 + 14x_4 + 4x_5 - 7)^2 - (x_1 + 12x_2 + 14x_3 - 6x_4 - 4x_5 - 10)^2.
\end{aligned}$$

*There are no linear constraints. For this SOCP, all the three inequalities are active at the optimizer. All the matrices $C_i$ have three rows. By formula (3.3), the algebraic degree of this problem is bounded by $2^3\binom{5-1}{3-1} = 48$. A symbolic computation over the finite field $\mathbb{Z}/17\mathbb{Z}$, using Singular [5], shows that the optimal coordinate $x_1$ is a root of a univariate polynomial of degree 48. The algebraic degree of this problem is 48, so the upper bound is sharp in this case. The exact integer coefficients of this polynomial are huge, e.g., the coefficient of $x_1^{48}$ returned by Singular is*
*20993751027404658600598159134663130280333894272179336371926053811704599113664919113955945081518362866289941284683755539037514999015805213743887244086800331841042514508227636572684706126659045130269952333391973158714518084532449779371025649171736541293437332448469583876494107694526951264585771066157333966857752253305920226530568083266479375648347403229514209064223548713844913807937173030267663957218228003769411146758482150283173793289744529268950487801811419136 00.*

**3.4. *p*-th order cone programming.** The $p$-th order cone programming (POCP) problem has the standard form

$$\begin{array}{rl} \min_{x \in \mathbb{R}^n} & c^T x \\ s.t. & a_i^T x + b_i - \|C_i x + d_i\|_p \geq 0,\ i = 1, \ldots, \ell \end{array} \tag{3.5}$$

where $c, a_i, b_i, C_i, d_i$ are matrices or vectors of appropriate dimensions. This is also a convex optimization problem. Let $x^*$ be an optimizer, and assume the constraints with indices $1, \ldots, m$ are active at $x^*$. Suppose $C_i$ has $r_i$ rows. When some $r_i = 1$, the constraint $a_i^T x + b_i - \|C_i x + d_i\|_p \geq 0$ is equivalent to two linear constraints

$$-(a_i^T x + b_i) \leq C_i x + d_i \leq a_i^T x + b_i.$$

Like the SOCP case, assume $1 = r_1 = \cdots = r_k < r_{k+1} \leq \cdots \leq r_m$. Then the problem (3.5) is equivalent to

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ s.t. \quad & a_i^T x + b_i + \sigma_i(C_i x + d_i) \geq 0,\ i = 1, \ldots, k \\ & (a_i^T x + b_i)^p - \sum_{j=1}^{r_i} (C_i x + d_i)_j^p \geq 0,\ i = k+1, \ldots, m \end{aligned}$$

where the scalar $\sigma_i$ is chosen such that $a_i^T x^* + b_i + \sigma_i(C_i x^* + d_i) = 0$. Then we have

$$S_{n-m}(0,\ldots,0,\underbrace{p-1,\ldots,p-1}_{m-k \text{ times}}) = \sum_{i_{k+1}+\cdots+i_m = n-m} (p-1)^{i_{k+1}+\cdots+i_m} = (p-1)^{n-m}\binom{n-k-1}{m-k-1}.$$

By Corollary 2.5, the algebraic degree of $x^*$ is therefore bounded by

$$p^{m-k}(p-1)^{n-m}\binom{n-k-1}{m-k-1}. \tag{3.6}$$

When $k = m$, problem (3.5) is reducible to a linear programming problem and hence its algebraic degree is one.

Now we discuss the sharpness of degree bound (3.6). Similarly to the SOCP case, for every $i = k+1, \ldots, m$, define the set of polynomials $U_i$ as

$$U_i = \left\{ (a_i^T x + b_i)^p - \sum_{1\leq j\leq r_i} \alpha_j^p (C_i x + d_i)_j^p : \alpha_1, \ldots, \alpha_{r_i} \in \mathbb{R} \right\}.$$

Then define complex affine spaces $V_i$ as follows:

$$V_i = \left\{ (a_i^T x + b_i)^p - \sum_{1\leq j\leq r_i} \beta_j (C_i x + d_i)_j^p : \beta_1, \ldots, \beta_{r_i} \in \mathbb{C} \right\},\ i = k+1, \ldots, m.$$

Then every set $U_i$ intersects any Zariski open subset of the affine space $V_i$. On the other hand, the set of common zeros of the linear polynomials

$$a_i^T x + b_i + \sigma_i(C_i x + d_i),\ i = 1, \ldots, k$$

and all the polynomials in the union $\bigcup_{i=k+1}^m V_i$ is contained in the set $Z$ defined by (3.4). Therefore, for generic choices of $a_i, b_i, C_i, d_i$ with $r_{k+1} + \cdots + r_m + m > n$, the set $Z$ is empty, and hence Remark 2.3 implies that the degree bound given by formula (3.6) is sharp.

EXAMPLE 3.4. *Consider the case $p = 4$ and the polynomials*

$$\begin{aligned} f_0 &= 9x_1 - 5x_2 + 3x_3 + 2x_4 \\ f_1 &= (1 - 6x_1 - 6x_2 + 4x_3 - 9x_4)^4 - (7 - 6x_1 + 22x_2 - x_3 + x_4)^4 \\ &\quad - (11 + x_1 - x_2 - 8x_3 + 3x_4)^4 - (-13 + 7x_1 + 16x_2 - 7x_3 + 9x_4)^4 \\ &\quad - (3 - 11x_1 + 14x_2 - 8x_3 + 5x_4)^4 - (8 + 9x_1 - 10x_2 + 2x_3 + 2x_4)^4. \end{aligned}$$

*For the above polynomials, the inequality constraint must be active since the objective is linear. By the formula (3.6), the algebraic degree of the optimal solution is bounded by* $p^m(p-1)^{n-m}\binom{n-1}{m-1} = 108$*. A symbolic computation over the finite field* $\mathbb{Z}/17\mathbb{Z}$*, using Singular [5], shows that the optimal coordinate* $x_1$ *is a root of a univariate polynomial of degree* 108*. So the algebraic degree of this problem is* 108*, and the bound given by the formula (3.6) is sharp. For this example, we were also not able to find the exact coefficients of this univariate polynomial in the rational field* $\mathbb{Q}$*, since Singular could not complete the computation over* $\mathbb{Q}$*.*

**3.5. Semidefinite programming.** In the introduction, we observed that the SDP problem of the form (1.3) can also be represented as a polynomial optimization problem of the form (1.1). Concretely, let $g_I(x)$ be the principle minor of the matrix in (1.3) with rows $I \subseteq \{1, 2, \dots, N\}$ ($N$ is the length of the matrix). Then the problem (1.3) is equivalent to

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} & c^T x \\ \quad s.t. & g_I(x) \geq 0, \quad \forall\, I \subseteq \{1, 2, \dots, N\}. \end{cases} \tag{3.7}$$

If the active constraints are known in the above, then Corollary 2.5 can be applied to get an upper bound for the algebraic degree of problem (1.3). Unfortunately, the upper bound obtained by Corollary 2.5 is usually not sharp, and often larger than the degree formula given in [13, Theorem 1.1]. This is because the polynomials $g_I(x)$ do not define a complete intersection: The codimension of $\mathcal{V}_r$ is less than the number of minors, i.e., the number of generators of the ideal of $\mathcal{V}_r$ . To see this point, suppose $r$ is the rank of the optimal matrix in (1.3). Then all $g_I(x)$ with $\mathrm{card}(I) > r$ must vanish at the optimal solution $x^*$. Let $\mathcal{V}_r$ be the variety defined by the $g_I(x)$:

$$\mathcal{V}_r = \{x \in \mathbb{R}^n : \, g_I(x) = 0, \quad \forall\, I : \, \mathrm{card}(I) > r\}$$

The ideal $I(\mathcal{V}_r)$ of $\mathcal{V}_r$ has $\binom{N}{r+1}$ generators. Since $\mathcal{V}_r$ contains the variety of matrices of rank at most $r$, which has codimension $\binom{N-r+1}{2}$, the codimension of $\mathcal{V}_r$ is smaller than the number of generators of its ideal for almost all values of $N$ and $r$. So the variety $\mathcal{V}_r$ is almost never a complete intersection.

**Acknowledgement.** The authors are grateful to Gabor Pataki, Richard Rimanyi and Bernd Sturmfels for motivating this paper, and wish to thank two referees for fruitful suggestions.

**Appendix A. Some elements of algebraic geometry.**

This section presents some basic definitions and theorems in algebraic geometry. Most of them can be found in Harris' book [4].

Let $\mathbb{C}[x_1, \dots, x_n]$ be the ring of polynomials in $x_1, \dots, x_n$. An additive subgroup $I$ of $\mathbb{C}[x_1, \dots, x_n]$ is called an *ideal* if for every $f \in I$, the product $f \cdot g \in I$ for any $g \in \mathbb{C}[x_1, \dots, x_n]$. Given an ideal $I$ of $\mathbb{C}[x_1, \dots, x_n]$, the polynomials $f_1, \dots, f_k$ are called *generators* of $I$ if for every $f \in I$ there exist $g_1, \dots, g_k \in \mathbb{C}[x_1, \dots, x_n]$ such that $f = f_1 g_1 + \ldots + f_k g_k$. We also say that the ideal $I$ is generated by $f_1, \dots, f_k$ and denote $I = \langle f_1, \dots, f_k \rangle$.

Let $\mathbb{C}^n$ be the $n$-dimensional complex affine space. A point $x \in \mathbb{C}^n$ is a complex vector $(x_1, \dots, x_n)$. A set $V$ in $\mathbb{C}^n$ is called an *affine algebraic variety* if there are polynomials $g_1, \dots, g_r$ such that

$$V = \{(x_1, \dots, x_n) \in \mathbb{C}^n : \, g_1(x_1, \dots, x_n) = \cdots = g_r(x_1, \dots, x_n) = 0\}\,.$$

In the Zariski topology on $\mathbb{C}^n$ the closed sets are precisely the affine algebraic varieties. A variety is irreducible if it is not the union of two proper closed subvarieties. Any variety is a finite union of distinct irreducible varieties. Given an affine variety $V$, the ideal of $V$ is defined to be the set of polynomials

$$I(V) = \{f \in \mathbb{C}[x_1, \dots, x_n] : \, f(x_1, \dots, x_n) = 0, \, \forall \, (x_1, \dots, x_n) \in V\} \, .$$

Let $\mathbb{P}^n$ be the $n$-dimensional complex projective space of lines through the origin in $\mathbb{C}^{n+1}$ . A point $\tilde{x} \in \mathbb{P}^n$ has homogeneous coordinates $[x_0, x_1, \dots, x_n]$, unique up to multiplication by a common nonzero scalar. A set $\mathcal{U}$ in $\mathbb{P}^n$ is called a *projective algebraic variety* if there are homogeneous polynomials $h_1(x_0, x_1, \dots, x_n), \dots, h_t(x_0, x_1, \dots, x_n)$ such that

$$\mathcal{U} = \{(x_0, x_1, \dots, x_n) \in \mathbb{P}^n : \, h_1(x_0, x_1, \dots, x_n) = \cdots = h_t(x_0, x_1, \dots, x_n) = 0\} \, .$$

In the Zariski topology on $\mathbb{P}^n$, the closed sets are precisely the projective algebraic varieties. As a special case, if $h \in \mathbb{C}[x_0, \dots, x_n]$ is a homogeneous polynomial, then the set

$$\mathcal{H} = \{\tilde{x} \in \mathbb{P}^n : \, h(\tilde{x}) = 0\} \subset \mathbb{P}^n$$

is a projective variety called a hypersurface. If furthermore $h$ has degree one, $\mathcal{H}$ is called a hyperplane.

A subset $I$ of $\mathbb{C}[x_0, \dots, x_n]$ is called a *homogeneous ideal* if $I$ is an ideal of $\mathbb{C}[x_0, \dots, x_n]$ and generated by homogeneous polynomials in the ring $\mathbb{C}[x_0, \dots, x_n]$. Given a projective variety $\mathcal{U}$, the ideal

$$I(\mathcal{U}) = \{f \in \mathbb{C}[x_0, x_1, \dots, x_n] : f(\tilde{x}) = 0, \, \forall \, \tilde{x} = (x_0, x_1, \dots, x_n) \in \mathcal{U}\}$$

is homogeneous. It is called the ideal of $\mathcal{U}$. A Zariski open subset $\mathcal{Q}$ of a projective variety $\mathcal{U}$ is called a *quasi-projective variety*, or equivalently, a quasi-projective variety is a locally closed subset of $\mathbb{P}^n$ in the Zariski topology.

The dimension of an affine (resp. projective) variety $V$ is the length $k$ of the longest chain of irreducible affine (resp. projective) subvarieties, $V = V_0 \supset V_1 \supset \cdots \supset V_k$. For a projective variety the dimension may be equivalently defined as the largest integer $k$ such that any set of $k$ hyperplanes have a common intersection point on $V$. A variety has pure dimension if all its irreducible components have the same dimension. Let $\mathcal{U}$ be a projective variety in $\mathbb{P}^n$ of pure dimension $k$. We say that $\mathcal{U}$ is a *complete intersection* if its homogeneous ideal is generated by $n-k$ homogeneous polynomials, that is, there exists homogeneous polynomials $f_1, \dots, f_{n-k} \in \mathbb{C}[x_0, x_1, \dots, x_n]$ such that $I(\mathcal{U}) = \langle f_1, \dots, f_{n-k} \rangle$.

Let $\mathcal{U}$ be an irreducible projective variety in $\mathbb{P}^n$ of dimension $k$ and $I(\mathcal{U}) = \langle f_1, \dots, f_r \rangle$. The *singular locus* $\mathcal{U}_{sing}$ is defined to be the projective algebraic variety

$$\mathcal{U}_{sing} = \{\tilde{x} \in \mathcal{U} : \, J(f_1, \dots, f_r) \text{ has rank less than } n-k \text{ at } \tilde{x}\} \, ,$$

where $J(f_1, \dots, f_r)$ denotes the Jacobian matrix of $f_1, \dots, f_r$.

The following is a fundamental theorem that we formulate in a version particularly applicable to our situation. We consider subvarieties in products of projective spaces. They are defined, as above, by ideals of polynomials that are homogeneous in two sets of variables, one set for each factor.

Theorem A.1 (Bertini's Theorem). *If $\mathcal{X} \subset \mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$ is a quasi-projective $k$-dimensional complex variety, and $\mathbb{P}^m$ a projective space parameterizing hypersurfaces in $\mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$. Let $\mathcal{Z} = \cap_{\mathcal{H} \in \mathbb{P}^m} \mathcal{H}$ be the common points of these hypersurfaces. Let $\mathcal{Y} = \mathcal{H} \cap \mathcal{X}$ for a general member $\mathcal{H} \in \mathbb{P}^m$. Then $\mathcal{Y}$ is a variety of dimension $k-1$ and*

$$\mathcal{Y}_{sing} \subset (\mathcal{X}_{sing} \cap \mathcal{Y}) \cup \mathcal{Z}.$$

*Proof.* Let $f_{\mathcal{H}}$ be a polynomial that defines $\mathcal{H}$. The polynomials $\{f_{\mathcal{H}} : \mathcal{H} \in \mathbb{P}^m\}$ generate a vector space $V$ of dimension $m+1$ of polynomials. Let $f_0, \ldots, f_m$ be a basis for this vector space. Then

$$x \mapsto [f_0(x), \ldots, f_m(x)]$$

defines a regular map

$$\mathcal{X} \setminus \mathcal{Z} \to \mathbb{P}^m.$$

The conclusion now follows from [4, Theorems 17.16 and 17.24]. □

An immediate corollary is:

Corollary A.2. *Let $f_1, \ldots, f_k \in \mathbb{C}[x_0, \ldots, x_{n_1}, y_0, \ldots, y_{n_2}]$ be generic polynomials and $k \leq n$. Then the projective variety $V(f_1, \ldots, f_k)$ defined by*

$$\Big\{(\tilde{x}, \tilde{y}) = ([x_0, x_1, \ldots, x_{n_1}], [y_0, \ldots, y_{n_2}]) \in \mathbb{P}^{n_1} \times \mathbb{P}^{n_2} : \, f_1(\tilde{x}, \tilde{y}) = \cdots = f_k(\tilde{x}, \tilde{y}) = 0\Big\}$$

*is a smooth complete intersection of dimension $(n_1 + n_2 - k)$.* The *degree* of an equidimensional projective variety $\mathcal{X}$ of dimension $k$ in $\mathbb{P}^n$ is the number of points in the intersection of $\mathcal{X}$ with $n-k$ general hyperplanes. Therefore if $\mathcal{X}$ and $\mathcal{Y}$ are subvarieties of $\mathbb{P}^n$ of the same dimension without common components, then

$$\deg(\mathcal{X} \cup \mathcal{Y}) = \deg(\mathcal{X}) + \deg(\mathcal{Y}).$$

The following is another fundamental theorem, which we also formulate in a version particularly applicable to our situation.

Theorem A.3 (Bézout's Theorem). *If $\mathcal{X}, \mathcal{Y} \subset \mathbb{P}^n$ are projective complex varieties of pure dimensions $k$ and $\ell$ with $k + \ell \geq n$, then $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$ and each irreducible component of the intersection has dimension at least $k + \ell - n$.*

*If $\mathcal{X}$ and $\mathcal{Y}$ intersect transversely, i.e., that each irreducible component of $\mathcal{X} \cap \mathcal{Y}$ has dimension $k + \ell - n$ and has a nonempty smooth open subset outside the singular loci of $\mathcal{X}$ and $\mathcal{Y}$, then*

$$\deg(\mathcal{X} \cap \mathcal{Y}) = \deg(\mathcal{X}) \cdot \deg(\mathcal{Y}).$$

*In particular, if $k + \ell = n$ and $\mathcal{X} \cap \mathcal{Y}$ is transverse, then it consists of* $\deg(\mathcal{X}) \cdot \deg(\mathcal{Y})$ *points.*

*Proof.* The first part is a secial case of [4, Theorem 17.24], except for the statement that the intersection $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$, while the second part is [4, Theorem 18.3]. To show that the intersection $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$ we pass to the affine cones $\mathcal{CX}$ and $\mathcal{CY}$ in $\mathbb{C}^{n+1}$ of the projective varieties in $\mathcal{X}$ and $\mathcal{Y}$ respectively. They, of course, both contain 0, so their intersection in $\mathbb{C}^{n+1}$ is non-empty. The cones $\mathcal{CX}$ and $\mathcal{CY}$ have dimensions $k+1$ and $\ell+1$, respectively. Therefore, by [4, Theorem 17.24], their intersection must have components of dimension at least $(k+1)+(\ell+1)-n+1 \geq 1$. The intersection $\mathcal{C}\mathcal{X} \cap \mathcal{C}\mathcal{Y}$ is clearly an affine cone over the intersection $\mathcal{X} \cap \mathcal{Y}$, so since it has dimension at least one, the intersection $\mathcal{X} \cap \mathcal{Y}$ is not empty. $\square$

Corollary A.4. *A complete intersection of hypersurfaces has degree equal to the product of the degrees of the hypersurfaces.*

Bertini's Theorem and Bézout's Theorem may be generalized to projective algebraic varieties whose ideal is given by all minors of a given size of a matrix whose entries are homogeneous polynomials. The generalization of Bertini's Theorem that we need is given by the following:

Proposition A.5. *Let $a_1 \leq \cdots \leq a_f$ be a finite sequence of positive integers, and let $M$ be an $e \times f$-matrix whose entries $m_{i,j}$ are homogeneous polynomials in $\mathbb{C}[x_0, \ldots, x_n]$ of degree $a_i$. Consider the projective variety*

$$X_r = \{\tilde{x} \in \mathbb{P}^n : M \text{ has rank at most } r \text{ at } \tilde{x}\}.$$

*Every irreducible component of $X_r$ has dimension $\geq (n-(f-r)(e-r))$. In particular, if $n-(f-r)(e-r) \geq 0$, then $X_r \neq \emptyset$. Furthermore, if the entries in $M$ are generic, then every irreducible component of $X_r$ has dimension $n-(m-r)(k-r)$, and $X_{r-1}$ is the singular locus of $X_r$.*

*Proof.* The bound on the codimension follows from [4, Proposition 17.25]. For the smoothness we consider first the case when the entries are independant variables $y_0, \ldots, y_{ef-1}$. A local parameterization of $X_r$ then shows that $X_r$ has codimension $(e-r)(f-r)$, is smooth outside $X_{r-1}$ and singular along $X_{r-1}$, (cf. [4, Example 20.5]). Substituting the variables with polynomials in $x_0, \ldots, x_n$, we may apply Bertini's Theorem A.1 to conclude. $\square$ To find the degree of $X_r$, we apply Bezout's Theorem in a slight variation of an argument in [4, Example 19.10] to get the special case $r = \min\{e-1, f-1\}$ that we need. For general $r$ the formula is given by the Thom-Porteous Formula, whose proof is more involved (cf. [3, Example 14.4.11]).

Proposition A.6. *Let $a_1 \leq \cdots \leq a_f$ be a finite sequence of positive integers, and let $M$ be an $e \times f$-matrix whose entries $m_{i,j}$ are general homogeneous polynomials in $\mathbb{C}[x_0, \ldots, x_n]$ of degree $a_j$. Consider the projective variety $X_{e-1} = \{\tilde{x} \in \mathbb{P}^n : M \text{ has rank less than } e \text{ at } \tilde{x}\}$.*

(1) *Assume that $e \leq f$ and let $E_1, \ldots, E_f$ be the elementary symmetric polynomials in the $a_i$, i.e.,*

$$(1+a_1t)\cdots(1+a_ft) = 1 + E_1t + \cdots + E_ft^f.$$

*If $n-f+e-1 \geq 0$, then*

$$\deg X_{e-1} = E_{f-e+1}.$$

(2) *Assume that $e > f$ and let $S_1, \ldots, S_k, \ldots$ be the complete symmetric polynomials in the $a_i$, i.e.,*

$$\frac{1}{(1-a_1t)(1-a_2t)\cdots(1-a_ft)} =$$

$$(1+a_1t+a_1^2t^2+\cdots)\cdots(1+a_ft+a_f^2t^2+\cdots) = 1 + S_1t + \cdots + S_kt^k + \cdots$$

*If $n-e+f-1 \geq 0$, then*

$$\deg X_{f-1} = S_{e-f+1}.$$

*Proof.* The first part we prove by specializing the matrix to one where the entries of each column in $M$ are proportional, while any $f-e+1$ polynomials from distinct columns define a complete intersection. Clearly then $X_{e-1}$ is simply the union of these complete intersections; $M$ has rank less than $e$ precisely when at least $f-e+1$ of the columns vanish. The degree of each of these complete intersections is the product of degrees of the corresponding polynomials, and their union have a degree equal to the sum $E_{f-e+1}$.

We prove the second formula by induction. Assume that $n > 1$ and that the formulas hold when $e < n$, the case $e = 1$ being trivial. Assume that $e = n > f$, and let $M_i$ be the submatrix of $M$ consisting of the first $e-i$ rows, and $N_j$ the submatrix of $M$ consisting of the rows $e-f, e-f+1, \ldots, e-j$. Let $X_i$ be the variety of points where $M_i$ has rank less than $f$, while $Y_j$ is the variety of points where $N_j$ has rank less than $f-j$. By induction $\deg X_i = S_{e-i-f+1}$, while $\deg Y_j = E_{j+1}$ by the previous argument. Notice that $X \subset X_1 \cap Y_0$ and that

$$X_f \cap Y_{f-1} \subset \cdots \subset X_i \cap Y_{i-1} \subset \cdots \subset X_1 \cap Y_0.$$

Furthermore

$$X = (X_1 \cap Y_0 \setminus (X_2 \cap Y_1 \setminus (X_3 \cap Y_2 \cdots \setminus (X_f \cap Y_{f-1})) \cdots ).$$

So by induction

$$\deg X = S_{e-f}E_1 - S_{e-f-1}E_2 + \cdots + (-1)^{f-1}S_{e-2f+1}E_f.$$

Computing the coefficient of $t^{e-f+1}$ in the identity

$$1 = \frac{(1+a_1t)\cdots(1+a_ft)}{(1+a_1t)\cdots(1+a_ft)} = (1+E_1t+\cdots+E_ft^f)(1-S_1t+\cdots+(-1)^kS_kt^k+\cdots),$$

we get

$$S_{e-f}E_1 - S_{e-f-1}E_2 + \cdots + (-1)^{f-1}S_{e-2f+1}E_f = S_{e-f+1}.$$

So the second part of the proposition also holds. □

## REFERENCES


[1] F. Catanese, S. Hoşten, A. Khetan and B. Sturmfels: The maximum likelihood degree, *American Journal of Mathematics*, **128** (2006) 671–697.

[2] D. Cox, J. Little and D. O'Shea. *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra.* Third edition. Undergraduate Texts in Mathematics. Springer, New York, 2007.

[3] W. Fulton. *Intersection Theory*. Springer Verlag, 1984.

[4] J. Harris. *Algebraic Geometry, A First Course*. Springer Verlag, 1992.

[5] G.-M. Greuel, G. Pfister and H. Schoenemann. *SINGULAR: A Computer Algebra System for Polynomial Computations.* Department of Mathematics and Centre for Computer Algebra, University of Kaiserslautern. `http://www.singular.uni-kl.de/index.html`

[6] S. Hosten, A. Khetan and B. Sturmfels. Solving the likelihood equations. *Foundations of Computational Mathematics.* 5 (2005) 389-407.

[7] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3): 796-817, 2001.

[8] J. Nie, J. Demmel and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Math. Prog.*, Series A, Vol. 106 (2006), No. 3, pp. 587-606.

[9] J. Nie, K. Ranestad and B. Sturmfels. The algebraic degree of semidefinite programming. To appear in *Mathematical Programming.*

[10] J. Nocedal and S. Wright. *Numerical Optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[11] P. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog.*, Ser. B, Vol. 96, No.2, pp. 293-320, 2003.

[12] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In S. Basu and L. Gonzalez-Vega, editors, *Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science, volume 60 of DIMACS Series in Discrete Mathematics and Computer Science*, pages 83-99. AMS, 2003.

[13] K. Ranestad and H.C. Graf von Bothmer. A general formula for the algebraic degree in semidefinite programming. 6 pages. To appear in *Bulletin of LMS*, available at http://arxiv.org/abs/math/0701877