LE MATEMATICHE
Vol. LXXVI (2021) – Issue II, pp. 483–499
doi: 10.4418/2021.76.2.12

THE DEGREE OF THE CENTRAL CURVE IN SEMIDEFINITE, LINEAR, AND QUADRATIC PROGRAMMING

S. HOŞTEN - I. SHANKAR - A. TORRES

The Zariski closure of the central path which interior point algorithms track in convex optimization problems such as linear, quadratic, and semidefinite programs is an algebraic curve. The degree of this curve has been studied in relation to the complexity of these interior point algorithms, and for linear programs it was computed by De Loera, Sturmfels, and Vinzant in 2012. We show that the degree of the central curve for generic semidefinite programs is equal to the maximum likelihood degree of linear concentration models. New results from the intersection theory of the space of complete quadrics imply that this is a polynomial in the size of semidefinite matrices with degree equal to the number of constraints. Besides its degree we explore the arithmetic genus of the same curve. We also compute the degree of the central curve for generic linear programs with different techniques which extend to bounding the same degree for generic quadratic programs.

1. Introduction

Let $\mathcal{S}^m_{\mathbb{R}}$ and $\mathcal{S}^m_{\mathbb{C}}$ be the vector spaces of $m \times m$ symmetric matrices with real and complex entries, respectively. Our starting point is semidefinite programs of the form

Received on November 30, 2020

AMS 2010 Subject Classification: 90C25,14Q05,62R01

Keywords: Central curve, semidefinite programming, degree of central curve, linear concentration models.

$$
\begin{aligned}
&\text{minimize } \langle C, X\rangle \\
&\text{subject to } \langle A_i, X\rangle = b_i, \quad i = 1,\ldots,d \\
&\qquad\qquad X \succeq 0
\end{aligned}
\tag{1}
$$

where C and $A_i, i = 1,\ldots,d$, are in $\mathcal{S}^m_{\mathbb{R}}$, and $b_i \in \mathbb{R}$ for $i = 1,\ldots,d$. We use the standard Euclidean inner product $\langle Y,Z\rangle = \mathrm{Tr}(YZ)$ on $\mathcal{S}^m_{\mathbb{R}}$, and $X \succeq 0$ means that X belongs to the cone of $m \times m$ positive semidefinite matrices. Typically, we will assume that the cost matrix C, the constraint matrices $A_1,\ldots,A_d$, and $b = (b_1,\ldots,b_d)^t$ are generic. This assures, among other things, that if (1) is feasible, it is strictly feasible.

The central curve of the above semidefinite program is obtained from the Karush-Kuhn-Tucker (KKT) conditions to an auxiliary optimization problem with a logarithmic barrier function. The KKT conditions are

$$
\begin{aligned}
&C - \lambda X^{-1} - \Sigma_{i=1}^d y_i A_i = 0, \\
&\langle A_i, X\rangle = b_i, \quad i = 1,\ldots,d, \\
&X \succeq 0
\end{aligned}
\tag{2}
$$

where $y_1,\ldots,y_d$ are the dual variables to the dual semidefinite program.

Definition 1.1. Let $(X^*(\lambda), y^*(\lambda))$ be the unique solution of the system (2) for a fixed $\lambda > 0$. The (primal) *central curve* $\mathcal{C}_{SDP}(C,\{A_i\},b)$ is the projection onto $\mathcal{S}^m_{\mathbb{C}}$ of the Zariski closure in $\mathcal{S}^m_{\mathbb{C}} \times \mathbb{C}^d$ of $\{(X^*(\lambda), y^*(\lambda)) : \lambda > 0\}$.

The central curve contains the *central path* $\{X^*(\lambda) : \lambda > 0\}$. Interior point algorithms follow a piecewise linear approximation to the central path to obtain an optimal solution to (1) as λ approaches zero [4, 8, 9, 16, 17]. The degree of $\mathcal{C}_{SDP}(C,\{A_i\},b)$ can be used to give an upper bound on the total curvature of the central path which is a heuristic measure on the number of steps interior point algorithms will take to find an optimal solution.

Interior point methods were first developed for linear programming problems, and the study of the central curve for linear programming from the perspective of algebraic geometry was initiated by Bayer and Lagarias in [3] and [2]. Dedieu, Malajovich, and Shub [7] studied the total curvature of the central path for linear programs in relation to bounding the number of iterations interior point algorithms take. By now we know that the total curvature can be exponential in the dimension of the ambient space [1]. Most relevant to our work, De Loera, Sturmfels, and Vinzant [6] obtained a breakthrough by computing the degree of the linear programming central curve. Given the linear program

$$\begin{aligned} &\text{minimize } cx \\ &\text{subject to } Ax = b \\ &\qquad\qquad x \geq 0, \end{aligned} \tag{3}$$

where $c \in \mathbb{R}^m$ is a row vector, A is $d \times m$ matrix of rank d, and $b \in \mathbb{R}^d$ is a column vector, they have related this degree to the degree of a *reciprocal variety* and a matroid invariant.

Theorem 1.2. *[6, Lemma 11] For generic b and c, the degree of the central curve of the linear program* (3) *is equal to the degree of the reciprocal variety*

$$\mathcal{L}_{A,c}^{-1} := \overline{\left\{(u_1,\dots,u_m) \in \mathbb{C}^m : \left(\frac{1}{u_1},\dots,\frac{1}{u_m}\right) \in \operatorname{rowspan}\begin{pmatrix} A \\ c \end{pmatrix} \text{ and } u_i \neq 0, i = 1,\dots,m\right\}}$$

as well as the Möbius number $|\mu(A,c)|$ *of the rank* $d+1$ *matroid associated to the row span of* $\begin{pmatrix} A \\ c \end{pmatrix}$*. When A is also generic, the degree of the central curve is equal to*

$$\binom{m-1}{d}.$$

Our main contribution is Theorem 2.2 where we prove that the degree of the central curve for the SDP (1) when C and b are generic is equal to the maximum likelihood degree (ML degree) of the linear concentration model generated by $\{A_i\}$ and C. When $\{A_i\}$ are also generic, this degree is equal to the degree of the reciprocal variety associated to the linear subspace $\mathcal{L}_{\{A_i\},C} = \operatorname{span}\{A_1,\dots,A_d,C\}$:

$$\mathcal{L}_{\{A_i\},C}^{-1} := \overline{\left\{X \in \mathcal{S}_{\mathbb{C}}^m : X^{-1} \in \mathcal{L}_{\{A_i\},C}\right\}}.$$

We further show in Corollary 2.3 that, when $\{A_i\}, C$, and b are generic, the degree of $\mathcal{C}_{SDP}(C,\{A_i\},b)$ is symmetric in the number of the linear equations defining (1). Corollary 2.4 concludes that in this case the degree of the central curve is a polynomial in m of degree d. This theorem and the two corollaries complete the work started in [18], proving Conjectures 4.3 and 4.4 in the same work.

In the remainder of Section 2 we will report our observations on the arithmetic genus of $\mathcal{C}_{SDP}(C,\{A_i\},b)$. We will also discuss semidefinite programs and the degree of their central curves associated to sum of squares (SOS) polynomials. In Section 3 we will revisit the degree of the central curve of the linear program (3) when A, c, and b are generic. Besides relating this degree to the ML degree of linear concentration models generated by diagonal matrices, in Theorem 3.3 we will provide a different proof that this degree is equal to $\binom{m-1}{d}$. Section 4 extends this result and its proof technique to convex quadratic programs with linear constraints. Theorem 4.1 bounds the degree of the central curve of such programs when the objective function and the constraints are generic.

2. Semidefinite Programs and Linear Concentration Models

In this section we consider the central curve $\mathcal{C}_{SDP}(C,\{A_i\},b)$ when C and b are generic. In what follows, we describe the degree of this curve as the ML degree of a linear concentration model. When $\{A_i\}$ are also generic, we denote $\deg(\mathcal{C}_{SDP}(C,\{A_i\},b))$ by $\psi_{SDP}(m,d)$.

2.1. Linear concentration models and degree of the central curve

Let $\mathcal{L}$ be a linear subspace of $\mathcal{S}^m_{\mathbb{R}}$ spanned by d linearly independent symmetric matrices $\{K_1,\dots,K_d\}$. A linear concentration model is the set

$$\mathcal{L}^{-1}_{\geq 0} := \{\Sigma \in \mathcal{S}^m_{\geq 0} : \Sigma^{-1} \in \mathcal{L}\}$$

where $\mathcal{S}^m_{\geq 0}$ is the cone of positive semidefinite matrices. Every matrix Σ in $\mathcal{L}^{-1}_{\geq 0}$ is the covariance matrix of a multivariate normal distribution on $\mathbb{R}^m$, and the elements of $\mathcal{L}$ are concentration matrices.

Given a sample covariance matrix S, the maximum likelihood estimate $\hat{K}$ of S with respect to the linear concentration model defined by $\mathcal{L}$ is the unique positive semidefinite solution to the zero-dimensional polynomial equations

$$\Sigma K = Id_m, \qquad K \in \mathcal{L}, \qquad \Sigma - S \in \mathcal{L}^{\perp}. \tag{4}$$

The ML degree of this linear concentration model is defined as the number of solutions to (4) in $\mathcal{S}^m_{\mathbb{C}}$.

In [20] it was proven that when the matrices $K_1,\dots,K_d$ are generic, the ML degree of the linear concentration model is precisely the degree of the reciprocal variety $\mathcal{L}^{-1}$.

Theorem 2.1. *[20, Theorem 2.3] The ML degree $\phi(m,d)$ of a linear concentration model defined by a generic linear subspace $\mathcal{L}$ of dimension d in $\mathcal{S}^m$ equals the degree of the projective variety $\mathcal{L}^{-1}$. This degree further satisfies*

$$\phi(m,d) = \phi\left(m, \binom{m+1}{2} + 1 - d\right).$$

Now we are ready to prove our main theorem.

Theorem 2.2. *Given an SDP as in* (1) *with C and b generic,* $\deg(\mathcal{C}_{SDP}(C,\{A_i\},b))$ *is equal to the ML degree of the linear concentration model generated by* $\mathcal{L} = \operatorname{span}\{C, A_1, \ldots, A_d\}$. *If in addition* $A_1,\ldots,A_d$ *are generic,* $\psi_{SDP}(m,d)$ *is equal to the degree of* $\mathcal{L}^{-1}$, *and hence* $\psi_{SDP}(m,d) = \phi(m,d+1)$.

Proof. By definition

$$\deg(\mathcal{C}_{SDP}(C,\{A_i\},b)) = |\mathcal{C}_{SDP}(C,\{A_i\},b) \cap \mathcal{H}|$$

where $\mathcal{H}$ is a generic hyperplane in $\mathcal{S}^m_{\mathbb{C}}$. Using the KKT conditions (2), the equations defining $\mathcal{C}_{SDP}(C,\{A_i\},b) \cap \mathcal{H}$ are

$$\begin{aligned} X^{-1} &= \frac{1}{\lambda}C - \frac{1}{\lambda}\Sigma_{i=1}^d y_i A_i \\ \langle A_i, X\rangle - b_i &= 0, \quad i = 1,\ldots,d \\ \langle B, X\rangle - b_{d+1} &= 0, \end{aligned} \tag{5}$$

for some generic $B \in \mathcal{S}^m_{\mathbb{C}}$ and $b_{d+1} \in \mathbb{C}$.

The first equation in (5) means that $X^{-1} \in \mathcal{L}$, where $\mathcal{L} = \operatorname{span}\{C, A_1, \ldots, A_d\}$. Since C is generic, in the last equation of (5) we can take $B = C$. Additionally, if we define S as a matrix such that $\langle A_i, S\rangle = b_i$, for $i = 1,\ldots,d$, and $\langle C, S\rangle = b_{d+1}$, the last $d+1$ equations in (5) mean that $X - S \in \mathcal{L}^{\perp}$. Note that these are precisely the likelihood equations of the linear concentration model determined by $\mathcal{L}$. This proves that $\deg(\mathcal{C}_{SDP}(C,\{A_i\},b))$ is equal to the ML degree of the linear concentration model defined by $\mathcal{L}$. Additionally, if $A_1,\ldots,A_d$ are generic, Theorem 2.1 guarantees that $\phi(m,d+1)$ coincides with the degree of $\mathcal{L}^{-1}$, which means that $\psi(m,d)$ is equal to the degree of $\mathcal{L}^{-1}$ as well. □

Corollary 2.3. *The degree of the central curve for a generic SDP satisfies*

$$\psi_{SDP}(m,d) = \psi_{SDP}\left(m, \binom{m+1}{2} - d - 1\right).$$

Proof.

$$\begin{aligned} \psi_{SDP}(m,d) &= \phi(m,d+1) \\ &= \phi(m, \binom{m+1}{2} + 1 - (d+1)) \\ &= \phi(m, \binom{m+1}{2} - d)) \\ &= \psi_{SDP}\left(m, \binom{m+1}{2} - d - 1\right). \end{aligned}$$

□

Corollary 2.4. *$\psi_{SDP}(m,d)$ is a polynomial in m of degree d.*

Proof. This result follows from the work of Michałek, Monin, Wiśniewski, Manivel, Seynnaeve, and Vodička who employed the space of complete quadrics and intersection theory to prove the polynomiality of $\phi(m,d)$ ([15] and [14, Theorem 1.3]) and from the seperate work of Cid-Ruiz [5, Corollary C]. □

2.2. Arithmetic Genus

The ideal of polynomials $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ in $\mathbb{C}[x_{ij} : 1 \le i \le j \le m]$ vanishing on the reciprocal variety $\mathcal{L}^{-1}_{\{A_i\},C}$ is a prime ideal since this variety is irreducible. The proof of Theorem 2.1 (see [20, Theorem 2.3]) relies on the fact that $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ is Cohen-Macaulay when $\{A_i\}$ and C are generic [11, 12]. Since the central curve $\mathcal{C}_{SDP}(C,\{A_i\},b)$ is obtained from intersecting the reciprocal variety with d generic linear equations in (2), the numerator of the Hilbert series of $I_{\mathcal{L}^{-1}_{\{A_i\},C}}$ and that of the defining ideal of the the central curve are identical. The Hilbert series for the central curve will be of the form

$$\frac{h_0 + h_1 t + h_2 t^2 + \cdots + h_k t^k}{(1-t)^2}$$

where the coefficients h_j are nonnegative integers with $h_0 = 1$ and $h_k \neq 0$. The arithmetic genus of the central curve can be calculated as

$$\text{genus}(m,d) := \text{genus}(\mathcal{C}_{SDP}(C,\{A_i\},b)) = 1 - \sum_{j=0}^{k} (1-j)h_j.$$

The following table shows genus(m,d) for all values we can compute with Macaulay2 [10] and/or using the two propositions that follow.

$m \backslash d$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	0	0												
3	0	0	1	0	0									
4	0	1	10	20	22	20	10	1	0					
5	0	3										33	3	0

Proposition 2.1. For $m \ge 2$,

$$\text{genus}(m,1) = \text{genus}\left(m, \binom{m+1}{2} - 1\right) = 0.$$

In these cases, the central curve is a rational curve. Furthermore, when $d = 1$ the numerator of the Hilbert series is $1 + (m-2)t$, and when $d = \binom{m+1}{2} - 1$ it is 1.

Proof. In the case $d = \binom{m+1}{2} - 1$, the reciprocal variety is equal to $\mathbb{P}^d$, and therefore the central curve is $\mathbb{P}^1$. In the case $d = 1$, the reciprocal variety is the image of $\operatorname{span}\{C, A_1\} \simeq \mathbb{P}^1$ under the rational map given by the $(m-1)$-minors of a generic $m \times m$ symmetric matrix. Hence it is a rational curve of degree $m-1$. This implies that the numerator of the Hilbert series of the ideal defining the reciprocal variety, and therefore that of the central curve, is $1 + (m-2)t$. This means that the central curve is also a rational curve, i.e., its genus is equal to zero. □

Proposition 2.2. For $m \geq 2$,

$$\operatorname{genus}\left(m, \binom{m+1}{2} - 2\right) = \binom{m-2}{2} \text{ and } \operatorname{genus}\left(m, \binom{m+1}{2} - 3\right) = 1 + (m-1)^2(m-3).$$

Proof. In the first case, the reciprocal variety is a hypersurface defined by a single polynomial of degree $m-1$. Therefore the numerator of the Hilbert series is equal to $1 + t + \cdots + t^{m-2}$. Therefore the arithmetic genus of the central curve is

$$1 - \sum_{j=0}^{m-2} (1-j) = \sum_{j=1}^{m-3} j = \binom{m-2}{2}.$$

In the second case, the reciprocal variety is of codimension two, and it is a complete intersection generated by two degree $m-1$ generators; see [20, p. 611] and Lemma 2.5 below. Therefore the numerator of the Hilbert series is equal to $(1 + t + \cdots + t^{m-2})^2 = 1 + 2t + \cdots + (m-2)t^{m-3} + (m-1)t^{m-2} + (m-2)t^{m-1} + \cdots + 2t^{2m-5} + t^{2m-4}$. Using the formula for the arithmetic genus first yields $1 + (2m-6)\binom{m-1}{2} + (m-3)(m-1)$. This in turn is equal to $1 + (m-3)(m-1)^2$. □

Lemma 2.5. *When $d = \binom{m+1}{2} - 3$, the reciprocal variety $\mathcal{L}^{-1}_{\{A_i\},C}$ associated to a generic linear subspace $\mathcal{L}_{\{A_i\},C}$ is a complete intersection of codimension two generated by two polynomials of degree $m-1$.*

Proof. Let V be the variety of codimension 3 in $\mathbb{P}^{\binom{m+1}{2}-1}$ defined by the $(m-1)$-minors of a generic $m \times m$ symmetric matrix, and let X be the quasiprojective variety $\mathbb{P}^{\binom{m+1}{2}-1} \setminus V$. Consider the regular map $F : X \longmapsto \mathbb{P}^{\binom{m+1}{2}-1}$ given by the $(m-1)$-minors of a generic $m \times m$ symmetric matrix. Given the generic codimension two subspace $\mathcal{L}_{\{A_i\},C}$, the inverse image $F^{-1}(\mathcal{L}_{\{A_i\},C})$ is an irreducible subvariety of X by Bertini's theorem [13, Theorem 3.3.1]. This subvariety is defined by two generic linear combinations of $(m-1)$-minors, f_1 and f_2, which are of degree $m-1$. The variety in $\mathbb{P}^{\binom{m+1}{2}-1}$ defined by the same two polynomials is a complete intersection of codimension two. This variety contains the reciprocal variety which is irreducible and has also codimension two. Therefore if the ideal $\langle f_1, f_2 \rangle$ is prime it has to be the defining ideal of the reciprocal variety. But this is the case, since it is a complete intersection and hence all its components have the same codimension. Any component other than the one coming from $F^{-1}(\mathcal{L}_{\{A_i\},C})$ is associated to V, but V has codimension three. $\square$

We note that in the above table the entry for $m = 5$ and $d = 12$ is computed using Proposition 2.2. However, the entry for $m = 5$ and $d = 3$, which is conjecturally equal to 33 is missing. Nevertheless, we venture to state the following conjecture.

Conjecture 2.3. $\operatorname{genus}(m,d) = \operatorname{genus}\left(m, \binom{m+1}{2} - d\right)$.

Although we cannot prove this conjecture, we can prove the analogous statement for the central curve of linear programs (3) when A, c and b are generic. The central curve for linear programs is defined as in Definition 1.1 but using the KKT conditions for linear programs; see (9) below.

Theorem 2.6. *Let A_d and A_{m-d} be generic matrices of size $d \times m$ and $(m-d) \times m$ and of rank d and $m-d$, respectively. Let b_d and b_{m-d} be two generic vectors in $\mathbb{R}^d$ and $\mathbb{R}^{m-d}$. The central curve of the linear program defined by A_d, b_d, and a generic vector c has the same arithmetic genus as the central curve of the linear program defined by A_{m-d}, b_{m-d} and c.*

Proof. Let $\mathcal{C}_{LP}(d)$ and $\mathcal{C}_{LP}(m-d)$ denote the central curve of the generic linear programs as in the statement. In this generic case, from [6] we have

$$\operatorname{genus}(\mathcal{C}_{LP}(d)) = 1 - \sum_{j=0}^{d} (1-j) \binom{m-d+j-2}{j}, \tag{6}$$

$$\operatorname{genus}(\mathcal{C}_{LP}(m-d)) = 1 - \sum_{j=0}^{m-d} (1-j) \binom{d+j-2}{j}, \tag{7}$$

where the binomial coefficients in each equation come from the coefficients of the Hilbert series computed in [6]. To check that both computations have the same value, we need the identities

$$\sum_{j=0}^{n} \binom{r+j}{j} = \binom{r+n+1}{n} \quad \text{and} \quad \sum_{j=0}^{d} j \binom{m-d+j-2}{j} = (m-d-1)\binom{m-1}{d-1}.$$

First we get

$$\begin{aligned}
\operatorname{genus}(\mathcal{C}_{LP}(d)) &= 1 - \sum_{j=0}^{d} \binom{m-d+j-2}{j} + \sum_{j=0}^{d} j\binom{m-d+j-2}{j} \\
&= 1 - \binom{m-d-2+d+1}{d} + (m-d-1)\binom{m-1}{d-1} \\
&= 1 - \binom{m-1}{d} + (m-d-1)\binom{m-1}{d-1} \\
&= 1 - \frac{(m-1)!}{(m-1-d)!d!} + (m-d-1)\frac{(m-1)!}{(m-1-d+1)!(d-1)!} \\
&= 1 - \frac{(m-1)!(m-md+d^2)}{(m-d)!d!}
\end{aligned}$$

where the second line comes from the identities mentioned above with $r = m - d - 2$. Doing a similar computation for $\operatorname{genus}(\mathcal{C}_{LP}(m-d))$ we get

$$\begin{aligned}
\operatorname{genus}(\mathcal{C}_{LP}(m-d)) &= 1 - \sum_{j=0}^{m-d} \binom{m-(m-d)+j-2}{j} + \sum_{j=0}^{m-d} j\binom{m-(m-d)+j-2}{j} \\
&= 1 - \binom{d-2+m-d+1}{m-d} + (d-1)\binom{m-1}{m-1-d} \\
&= 1 - \binom{m-1}{m-d} + (d-1)\binom{m-1}{m-1-d} \\
&= 1 - \frac{(m-1)!}{(d-1)!(m-d)!} + (d-1)\frac{(m-1)!}{d!(m-1-d)!} \\
&= 1 - \frac{(m-1)!(m-md+d^2)}{(m-d)!d!}.
\end{aligned}$$

□

2.3. Sum of Squares Polynomials

We conclude Section 2 by considering semidefinite programs for sums of squares problems. For this, let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a homogeneous polynomial of degree $2D$ and let L_p be the affine subspace of symmetric matrices Q satisfying the identity

$$p = [x]^T Q[x] \tag{8}$$

where $[x]$ is a vector of all monomials of degree D in n variables. The intersection of L_p with the cone of positive semidefinite matrices is the Gram spectrahedron of p, and it is nonempty if and only if p is a sum of squares (SOS) polynomial. That is, certifying that a polynomial is SOS reduces to checking the feasibility of an SDP. This can be achieved by solving an SDP using a random (generic) cost matrix C.

Example 2.4. Suppose we wish to show that a generic ternary quartic is an SOS. The A_i's and b_i's come from equating coefficients in (8). For example, if we let $[x] = [x^2, xy, xz, y^2, yz, z^2]$, the linear equation for the x^2y^2 term will be

$$p_{(2,2,0)} = \langle A_{(2,2,0)}, Q\rangle$$

where $p_{(2,2,0)}$ is the x^2y^2 coefficient of the random ternary quartic p,

$$A_{(2,2,0)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and Q is the decision variable of the SDP which will have $d = \binom{3+4-1}{4} = 15$ constraints with matrices of size $m = \binom{3+2-1}{2} = 6$.

In general, the matrices $\{A_i\}$ for the linear constraints will be sparse and far from generic. However, if the polynomial p that we want to certify to be an SOS polynomial is generic, then the b_i's in the corresponding SDP will be also generic. Using a generic cost matrix C in this SDP allows us to consider the degree of the central curve for a generic SOS polynomial.

We wish to report our computations in three instances: binary sextics $(n = 2, 2D = 6)$, binary octics $(n = 2, 2D = 8)$, and ternary quartics $(n = 3, 2D = 4)$. The corresponding SDPs are given by input data with $m = 4, d = 7$ for binary sextics, $m = 5, d = 9$ for binary octics, and $m = 6, d = 15$ for ternary quartics. We note that for the same size SDPs with generic $\{A_i\}$ we will obtain $\psi_{SDP}(4,7) = 9$, $\psi_{SDP}(5,9) = 137$, and $\psi_{SDP}(6,15) = 528$. We believe that studying this invariant for various families of SOS polynomials is an interesting future project.

Proposition 2.5. The degrees of the central curves for SDPs associated to generic binary sextics, binary octics, and ternary quartics, where generic cost matrices are used, are 7, 45, and 66, respectively.

As a last remark about the SDP arising from sums of squares, we would like to mention that since C and b are generic, the computations in the previous proposition, also correspond to the ML degree of a linear concentration model. Namely, the concentration model defined by catalecticants and an additional generic matrix corresponding to the cost matrix. Exploring this relation is also an interesting future project.

3. Linear Programs

By choosing C and $\{A_i\}$ in (1) to be diagonal matrices we recover linear programs (3).

The central curve $\mathcal{C}_{LP}(c,A,b)$ for such a linear program can be defined as in the case of the central curve for a semidefinite program using the corresponding KKT conditions:

$$\begin{aligned} &c-\lambda\left(\frac{1}{x_1},\ldots,\frac{1}{x_m}\right)-y^tA=0\\ &Ax=b,\\ &x\geq 0. \end{aligned} \tag{9}$$

When in the data defining (3), c and b are generic the degree of the central curve $\mathcal{C}_{LP}(c,A,b)$ is equal to the degree of the reciprocal variety $\mathcal{L}^{-1}_{A,c}$ [6, Lemma 11]. Further, if A is also generic, this degree is equal to $\binom{m-1}{d}$. For the case when all the data is generic, we will denote the degree of the linear programming central curve by $\psi_{LP}(m,d)$.

The observations that connect the ML degree of generic linear concentration models to the degree of the central curve of generic semidefinite programs have their counterpart here as well. One can consider the ML degree of linear concentration models generated by diagonal matrices as in [20, Section 3]. For generic models we denote the ML degree by $\phi_{\text{diag}}(m,d)$. A consequence of Corollary 3 in [20] is the following.

Corollary 3.1.

$$\phi_{\text{diag}}(m,d) \quad = \quad \binom{m-1}{d-1}.$$

An argument parallel to the one used in the proof of Theorem 2.2 gives

Corollary 3.2. $\psi_{LP}(m,d)=\phi_{\text{diag}}(m,d+1)$.

In the rest of this section we will develop another method to prove that $\psi_{LP}(m,d)=\binom{m-1}{d}$. This method will be extended for bounding the degree of the central curve for generic convex quadratic programs with linear constraints in the next section. We note that our techniques which are based on counting solutions to polynomial systems were employed for a similar purpose in [7].

First we consider the polynomial system obtained by clearing denominators and dropping the $x\geq 0$ condition in (9):

$$\begin{aligned} &c_ix_i-\lambda-(y^ta_i)x_i=0, \quad i=1,\ldots,m\\ &Ax=b \end{aligned} \tag{10}$$

where a_i is the ith column of the matrix A. For generic data, the central curve is obtained as the Zariski closure in $\mathbb{C}^m$ of the projection of the solution set in $(\mathbb{C}^*)^{m+d+1}$ to the equations (10). Further, the degree of this central curve would be equal to the number of points in $(\mathbb{C}^*)^m$ obtained as the intersection of the central curve with a generic hyperplane defined by $ex = f$.

Lemma 3.1. The degree of $\mathcal{C}_{LP}(c,A,b)$ for generic c, A, and b is equal to the number of solutions in $(\mathbb{C}^*)^{m+d+1}$ to the system (10) together with an extra equation of the form $ex = f$ where the coefficients of this equation are generic.

Proof. Clearly, every solution to (10) plus $ex = f$ in $(\mathbb{C}^*)^{m+d+1}$ projects to a point in $\mathcal{C}_{LP}(c,A,b) \cap \{x : ex = f\}$. Conversely, the genericity of $ex = f$ implies that the points in $\mathcal{C}_{LP}(c,A,b) \cap \{x : ex = f\}$ come from points in $(\mathbb{C}^*)^{m+d+1}$ that satisfy (10) and $ex = f$. We show that for each point x^* "downstairs" there is a unique point "upstairs". Suppose there are at least two points (x^*, y^*, λ^*) and (x^*, z^*, μ^*) with these properties. Then it is easy to check that $(x^*, ty^* + (1-t)z^*, t\lambda^* + (1-t)\mu^*)$ is also a solution with the same properties for any t. But this is a contradiction since we have only finitely many preimages by the genericity of the data. □

This lemma implies that in order to compute the degree of $\mathcal{C}_{LP}(c,A,b)$ for generic A, c, and b we need to count the solutions in $(\mathbb{C}^*)^{m+d+1}$ to

$$\begin{aligned} c_i x_i - \lambda - (y^t a_i) x_i &= 0, \quad i = 1, \ldots, m \\ Ax &= b \\ ex &= f \end{aligned} \tag{11}$$

where $ex = f$ is also generic. Note that the rank of the matrix $\begin{pmatrix} A \\ e \end{pmatrix}$ is $d+1$ and the solutions to the last $d+1$ equations in (11) can be parametrized by

$$x = v_0 + t_1 v_1 + \cdots + t_{m-d-1} v_{m-d-1}$$

where $v_0, v_1, \ldots, v_{m-d-1}$ are generic vectors. Substituting this into the first m equations in (11) we obtain m equations in m variables $\lambda, y_1, \ldots, y_d, t_1, \ldots, t_{m-d-1}$. Furthermore, the genericity assumptions guarantee that each equation will have support equal to

$$\lambda, 1, t_1, \ldots, t_{m-d-1}, y_1, y_1 t_1, \ldots, y_1 t_{m-d-1}, \ldots, y_d, y_d t_1, \ldots, y_d t_{m-d-1}.$$

The Newton polytope of a polynomial with this support is a pyramid of height one with base equal to the product of simplices $\Delta_{m-d-1} \times \Delta_d$.

Theorem 3.3. *$\psi_{LP}(m,d)$ is equal to the volume of $\Delta_{m-d-1}\times\Delta_d$:*

$$\binom{m-1}{d} = \sum_{k=0}^{m-d-1}\binom{m-k-2}{d-1}$$

Proof. The above lemma and the previous discussion imply that $\psi_{LP}(m,d)$ is equal to the number solutions in $(\mathbb{C}^*)^m$ to m equations in m variables, where each equation has support equal to the set of monomials listed above. Bernstein's Theorem implies that this number is bounded above by the normalized volume of the Newton polytope of these monomials. Since this polytope is a pyramid of height one over $\Delta_{m-d-1}\times\Delta_d$, we just need to compute the normalized volume of the product of simplices. Further, because every triangulation of $\Delta_{m-d-1}\times\Delta_d$ is unimodular we just need to count the number of simplices in any triangulation. One such triangulation is the staircase triangulation. The maximal simplices in this triangulation are described as follows. Consider a $(m-d)\times(d+1)$ rectangular grid. The simplices in the staircase triangulation of $\Delta_{m-d-1}\times\Delta_d$ are in bijection with paths from the northwest corner of this grid to the southeast corner where a path consists of steps in the east or south direction. The total number of steps in each path is $m-1$, and out of these steps d have to be south steps. Therefore there are a total of $\binom{m-1}{d}$ such paths. These paths can be partitioned into those which reach the south edge of the grid k steps before the southeast corner where $k=0,\ldots,m-d-1$. The number of these kinds of paths for each k is $\binom{m-k-2}{d-1}$. Finally, the proof of Lemma 11 in [6] implies that $\psi_{LP}(m,d)\geq\binom{m-1}{d}$, and this concludes the proof. $\square$

4. Quadratic Programs

To complete our study of central curves in optimization problems we will now consider convex quadratic programs with linear constraints.

$$\begin{aligned}&\text{minimize } \frac{1}{2}x^tQx+cx\\ &\text{subject to } Ax=b\\ &\qquad\qquad\quad x\geq 0,\end{aligned}\tag{12}$$

where Q is an $m\times m$ positive definite matrix, $c\in\mathbb{R}^m$, A is $d\times m$ matrix of rank d, and $b\in\mathbb{R}^d$. The KKT conditions that lead to the definition of the central curve are

$$\begin{aligned}&x^tQ+c-\lambda\left(\frac{1}{x_1},\ldots,\frac{1}{x_m}\right)-y^tA=0\\ &Ax=b,\\ &x\geq 0.\end{aligned}\tag{13}$$

When Q, c, A, and b are generic, we denote by $\psi_{QP}(m,d)$ the degree of the central curve for generic quadratic programs. One can show by a homotopy continuation argument that it is sufficient to assume Q to be a generic diagonal matrix. For precise details of this result, we refer the reader to [19, Section 3.2]. With $Q = \mathrm{diag}(q_1,\dots,q_m)$, after clearing denominators and ignoring the nonnegativity constraints $x \geq 0$ in (13), we arrive to the following system of polynomial equations:

$$\begin{aligned} &q_i x_i^2 + c_i x_i - \lambda - (y^t a_i) x_i = 0 \quad i = 1,\dots,m \\ &Ax = b, \end{aligned} \tag{14}$$

where a_i is the ith column of the matrix A. As in the linear programming case we have the following lemma.

Lemma 4.1. $\psi_{QP}(m,d)$, the degree of the central curve of a generic quadratic program is equal to the number of solutions in $(\mathbb{C}^*)^{m+d+1}$ to the system (14) together with an extra equation of the form $ex = f$ where the coefficients of this equation are also generic.

Proof. The proof of this lemma is identical to the proof of Lemma 3.1. □

Theorem 4.1.
$$\psi_{QP}(m,d) \leq \sum_{k=0}^{m-d-1} \binom{m-k-2}{d-1} 2^k.$$

This is the volume of the Newton polytope of a polynomial with support in monomials

$$\lambda, 1, t_1, \dots, t_{m-d-1}, t_1^2, t_1 t_2, \dots, t_{m-d-1}^2$$

$$y_1, y_1 t_1, \dots, y_1 t_{m-d-1}, \dots, y_d, y_d t_1, \dots, y_d t_{m-d-1}$$

Proof. By Lemma 4.1 and as in the proof of Theorem 3.3 we need to count solutions to (14) plus a generic linear equation $ex = f$ in the torus $(\mathbb{C}^*)^{m+d+1}$. The solutions to the equations $Ax = b$ and $ex = f$ can again be parametrized as

$$x = v_0 + t_1 v_1 + \cdots + t_{m-d-1} v_{m-d-1}$$

where $v_0,\dots,v_{m-d-1}$ are generic vectors. Substituting this into the first m equations in (14) we obtain m equations in m variables $\lambda, y_1,\dots,y_d, t_1,\dots,t_{m-d-1}$. Furthermore, the genericity assumptions guarantee that each equation will have support equal to

$$\lambda, 1, t_1, \dots, t_{m-d-1}, t_1^2, t_1 t_2, \dots, t_{m-d-1}^2$$

$$y_1, y_1 t_1, \dots, y_1 t_{m-d-1}, \dots, y_d, y_d t_1, \dots, y_d t_{m-d-1}$$

The number of solutions to these m equations in $(\mathbb{C}^*)^m$ is bounded by the normalized volume of the Newton polytope of the above monomials. Since this is a pyramid of height one, we just need to compute the volume of the Newton polytope of the monomials except λ. This polytope has a staircase triangulation as for $\Delta_{m-d-1} \times \Delta_d$ where each simplex corresponds to a path as we described in the proof of Theorem 3.3, except that the volume of a simplex corresponding to a path which reaches the south edge of the grid k steps before the southeast corner is 2^k. Therefore $\psi_{QP}(m,d)$ is at most

$$\sum_{k=0}^{m-d-1} \binom{m-k-2}{d-1} 2^k.$$

□

5. Acknowledgements

We are grateful to Bernd Sturmfels for his help in Lemma 2.5, and to Frank Sottile for pointing out an error in the original version of Theorem 4.1.

REFERENCES

[1] Xavier Allamigeon - Pascal Benchimol - Stéphane Gaubert - Michael Joswig, *Log-barrier interior point methods are not strongly polynomial*, SIAM J. Appl. Algebra Geom. 2 no. 1 (2018), 140–178.

[2] D. A. Bayer - J. C. Lagarias, *The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories*, Trans. Amer. Math. Soc. 314 no. 2 (1989), 499–526.

[3] D. A. Bayer - J. C. Lagarias, *The nonlinear geometry of linear programming. II. Legendre transform coordinates and central trajectories*, Trans. Amer. Math. Soc. 314 no. 2 (1989), 527–581.

[4] Stephen Boyd - Lieven Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.

[5] Yairon Cid-Ruiz, *Equations and Multidegree for Inverse Symmetric Matrix Pairs*, 2020, To appear in Le Matematiche on Linear Spaces of Symmetric Matrices. arXiv preprint arxiv.org/abs/2011.04616.

[6] Jesús A. De Loera - Bernd Sturmfels - Cynthia Vinzant, *The central curve in linear programming*, Found. Comput. Math. 12 no. 4 (2012), 509–540.

[7] Jean-Pierre Dedieu - Gregorio Malajovich - Mike Shub, *On the curvature of the central path of linear programming theory*, Found. Comput. Math. 5 no. 2 (2005), 145–171.

[8] Anthony V. Fiacco - Garth P. McCormick, *Nonlinear programming: Sequential unconstrained minimization techniques*, John Wiley and Sons, Inc., New York-London-Sydney, 1968.

[9] Anders Forsgren - Philip E. Gill - Margaret H. Wright, *Interior Methods for Nonlinear Optimization*, SIAM Review 44 no. 4 (2002), 525–597.

[10] Daniel R. Grayson - Michael E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, Available at `http://www.math.uiuc.edu/Macaulay2/`.

[11] J. Herzog - W. V. Vasconcelos - R. Villarreal, *Ideals with sliding depth*, Nagoya Math. J. 99 (1985), 159–172.

[12] Boris V. Kotzev, *Determinantal ideals of linear type of a generic symmetric matrix*, J. Algebra 139 no. 2 (1991), 484–504.

[13] Robert Lazarsfeld, *Positivity in algebraic geometry. I*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge., vol. 48, Springer-Verlag, Berlin, 2004.

[14] Laurent Manivel - Mateusz Michałek - Leonid Monin - Tim Seynnaeve - Martin Vodička, *Complete Quadrics: Schubert Calculus for Gaussian Models and Semidefinite Programming*, 2020, To appear in Le Matematiche on Linear Spaces of Symmetric Matrices. arXiv preprint arxiv.org/abs/2011.08791.

[15] Mateusz Michałek - Leonid Monin - Jarosław A. Wiśniewski, *Maximum Likelihood Degree, Complete Quadrics and* $\mathbb{C}^*$*–Action*, 2020, arXiv preprint arxiv.org/abs/2004.07735.

[16] Yurii Nesterov - Arkadii Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM Studies in Applied Mathematics, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[17] Jorge Nocedal - Stephen J. Wright, *Numerical optimization*, second ed., Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.

[18] Joshua D. Rhodes, *Computing the Central Sheet in Linear, Quadratic, and Semidefinite Programs*, August 2016, Masters Thesis.

[19] Dennis Schlief, *Degree for the Central Curve of Quadratic, Programing*, July 2014, Masters Thesis.

[20] Bernd Sturmfels - Caroline Uhler, *Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry*, Ann. Inst. Statist. Math. 62 no. 4 (2010), 603–638.

S. HOŞTEN
San Francisco State University, California, USA
e-mail: serkan@sfsu.edu

I. SHANKAR
University of California, Berkeley, USA
e-mail: isabelle_shankar@berkeley.edu

A. TORRES
KTH Royal Institute of Technology, Sweden
e-mail: amtb@kth.se