The Euclidean Distance Degree of an Algebraic Variety

Jan Draisma, Emil Horobeţ, Giorgio Ottaviani,
Bernd Sturmfels and Rekha R. Thomas

Abstract

The nearest point map of a real algebraic variety with respect to Euclidean distance is an algebraic function. For instance, for varieties of low rank matrices, the Eckart-Young Theorem states that this map is given by the singular value decomposition. This article develops a theory of such nearest point maps from the perspective of computational algebraic geometry. The Euclidean distance degree of a variety is the number of critical points of the squared distance to a generic point outside the variety. Focusing on varieties seen in applications, we present numerous tools for exact computations.

1 Introduction

Many models in the sciences and engineering are expressed as sets of real solutions to systems of polynomial equations in n unknowns. For such a real algebraic variety $X \subset \mathbb{R}^n$, we consider the following problem: given $u \in \mathbb{R}^n$, compute $u^* \in X$ that minimizes the squared Euclidean distance $d_u(x) = \sum_{i=1}^n (u_i - x_i)^2$ from the given point u. This optimization problem arises in a wide range of applications. For instance, if u is a noisy sample from X, where the error model is a standard Gaussian in $\mathbb{R}^n$, then u^* is the maximum likelihood estimate for u.

In order to find u^* algebraically, we consider the set of solutions in $\mathbb{C}^n$ to the equations defining X. In this manner, we regard X as a complex variety in $\mathbb{C}^n$, and we examine all complex critical points of the squared distance function $d_u(x) = \sum_{i=1}^n (u_i - x_i)^2$ on X. Here we only allow those critical points x that are non-singular on X. The number of such critical points is constant on a dense open subset of data $u \in \mathbb{R}^n$. That number is called the *Euclidean distance degree* (or ED degree) of the variety X, and denoted as EDdegree(X).

Using Lagrange multipliers, and the observation that $\nabla d_u = 2(u - x)$, our problem amounts to computing all regular points $x \in X$ such that $u - x = (u_1 - x_1, \ldots, u_n - x_n)$ is perpendicular to the tangent space $T_x X$ of X at x. Thus, we seek to solve the constraints

$$x \in X\,,\ x \notin X_{\text{sing}} \quad \text{and} \quad u - x \perp T_x X, \tag{1.1}$$

where X_{sing} denotes the singular locus of X. The ED degree of X counts the solutions x.

Example 1.1. We illustrate our problem for a plane curve. Figure 1 shows the *cardioid*

$$X \;=\; \left\{(x, y) \in \mathbb{R}^2 \,:\, (x^2 + y^2 + x)^2 = x^2 + y^2\right\}.$$

For general data (u,v) in $\mathbb{R}^2$, the cardioid X contains precisely three points (x,y) whose tangent line is perpendicular to $(u-x, v-y)$. Thus $\text{EDdegree}(X) = 3$. All three critical points (x,y) are real, provided (u,v) lies outside the *evolute*, which is the small inner cardioid

$$\{ (u,v) \in \mathbb{R}^2 \, : \, 27u^4 + 54u^2v^2 + 27v^4 + 54u^3 + 54uv^2 + 36u^2 + 9v^2 + 8u \, = \, 0 \}. \qquad (1.2)$$

The evolute is called *ED discriminant* in this paper. If (u,v) lies inside the evolute then two of the critical points are complex, and the unique real solution maximizes d_u. ♢

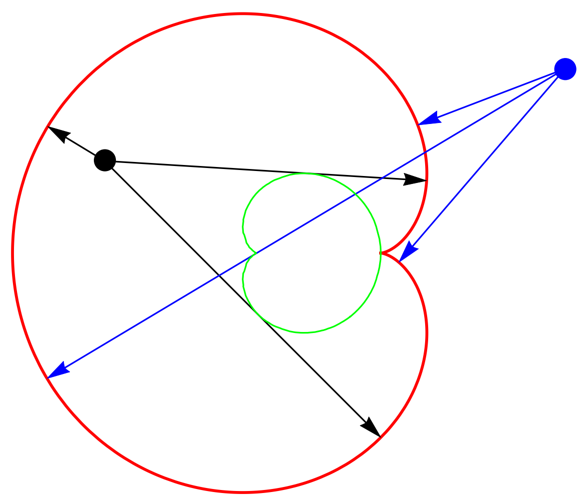

Figure 1: The cardioid has ED degree three. The inner cardioid is the ED discriminant.

Readers familiar with algebraic statistics [11] may note that the ED degree of a variety X is an additive analogue of its *ML degree* (maximum likelihood degree). Indeed, if X represents a statistical model for discrete data then maximum likelihood estimation leads to polynomial equations which we can write in a form that looks like (1.1), with $u/x = (u_1/x_1, \ldots, u_n/x_n)$:

$$x \in X\,, \;\; x \notin X_{\text{sing}} \quad \text{and} \quad u/x \perp T_x(X). \qquad (1.3)$$

See [23, 24] for details. Here, the optimal solution $\hat{u}$ minimizes the Kullback-Leibler distance from the distribution u to the model X. Thus, ED degree and ML degree are close cousins.

Example 1.2. To compare these two paradigms, ED versus ML, we consider the algebraic function that takes a 2×2-matrix u to its closest rank one matrix. For this problem we have $\text{MLdegree}(X) = 1$ and $\text{EDdegree}(X) = 2$. To see what this means, consider the instance

$$u = \begin{pmatrix} 3 & 5 \\ 7 & 11 \end{pmatrix}.$$

The closest rank 1 matrix in the maximum likelihood sense of [11, 24] has rational entries:

$$\hat{u} \quad = \quad \frac{1}{3{+}5{+}7{+}11} \begin{pmatrix} (3{+}5)(3{+}7) & (3{+}5)(5{+}11) \\ (7{+}11)(3{+}7) & (7{+}11)(5{+}11) \end{pmatrix}.$$

By contrast, when minimizing the Euclidean distance, we must solve a quadratic equation:

$$u^* = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \text{ where } v_{11}^2 - 3v_{11} - \tfrac{437}{1300} = 0, v_{12} = \tfrac{62}{41}v_{11} + \tfrac{19}{82},\ v_{21} = \tfrac{88}{41}v_{11} + \tfrac{23}{82},\ v_{22} = \tfrac{141}{41}v_{11} + \tfrac{14}{41}.$$

This rank 1 matrix arises from the Singular Value Decomposition, as seen in Example 2.3. ♢

In Example 1.2, for real data u, both solutions u^* are real. For most varieties X and data u, however, the number of real critical points is much smaller than $\mathrm{EDdegree}(X)$. To quantify that difference we also study the expected number of real critical points of d_u on X. This number, denoted $\mathrm{aEDdegree}(X)$ and called the *average ED degree*, depends on the underlying probability distribution on $\mathbb{R}^n$. For instance, for the cardioid X in Example 1.1, $\mathrm{aEDdegree}(X)$ can be any real number between 1 and 3, depending on how we sample (u, v).

This paper is organized as follows. In Section 2 we rigorously define ED degree for affine and projective varieties, and show how the ED degree of X and all critical points of d_u can be computed in practice. The projective case is important because many varieties in applications are defined by homogenous equations. For the most part, our exposition assumes no prerequisites beyond undergraduate mathematics. We follow the book by Cox, Little and O'Shea [8], and we illustrate the main concepts with code in `Macaulay2` [18].

Section 3 is devoted to case studies in control theory, geometric modeling, computer vision, and low rank matrix completion. New results include formulas for the ED degree for the Hurwitz stability problem and for the number of critical formations on the line, as in [2].

In Section 4 we introduce the *ED correspondence* $\mathcal{E}_X$, which is the variety of pairs (x, u) with $x \in X$ is critical for d_u. The ED correspondence is of vital importance for the computation of average ED degrees, in that same section. We show how to derive parametric representations of $\mathcal{E}_X$, and how these translate into integral representations for $\mathrm{aEDdegree}(X)$.

Duality plays a key role in both algebraic geometry and optimization theory [31]. Every projective variety $X \subset \mathbb{P}^n$ has a dual variety $X^* \subset \mathbb{P}^n$, whose points are the hyperplanes tangent to X. In Section 5 we prove that $\mathrm{EDdegree}(X) = \mathrm{EDdegree}(X^*)$, we express this number as the sum of the classical polar classes [22], and we lift the ED correspondence to the conormal variety of (X, X^*). When X is smooth and toric, we obtain a combinatorial formula for $\mathrm{EDdegree}(X)$ in terms of the volumes of faces of the corresponding polytope.

In Section 6 we study the behavior of the ED degree under linear projections and under intersections with linear subspaces. We also examine the fact that the ED degree can go up or can go down when passing from an affine variety in $\mathbb{C}^n$ to its projective closure in $\mathbb{P}^n$.

In Section 7 we express $\mathrm{EDdegree}(X)$ in terms of Chern classes when X is smooth and projective, and we apply this to classical Segre and Veronese varieties. We also study the *ED discriminant* which is the locus of all data points u where two critical points of d_u coincide. For instance, in Example 1.1, the ED discriminant is the inner cardioid. Work of Catanese and Trifogli [6] offers degree formulas for ED discriminants in various situations.

Section 8 covers the approximation of tensors by rank one tensors, following [9, 12, 13].

2 Equations defining critical points

An algebraic variety X in $\mathbb{R}^n$ can be described either implicitly, by a system of polynomial equations in n variables, or parametrically, as the closure of the image of a polynomial map

$\psi : \mathbb{R}^m \to \mathbb{R}^n$. The second representation arises frequently in applications, but it is restricted to varieties X that are unirational. The first representation exists for any variety X. In what follows we start with the implicit representation, and we derive the polynomial equations that characterize the critical points of the squared distance function $d_u = \sum_{i=1}^n (x_i - u_i)^2$ on X. The function d_u extends to a polynomial function on $\mathbb{C}^n$. So, if x is a complex point in X then $d_u(x)$ is usually a complex number, and that number can be zero even if $x \neq u$. The Hermitian inner product and its induced metric on $\mathbb{C}^n$ will not appear in this paper.

Fix a radical ideal $I_X = \langle f_1, \ldots, f_s \rangle \subset \mathbb{R}[x_1, \ldots, x_n]$ and $X = V(I_X)$ its variety in $\mathbb{C}^n$. Since ED degree is additive over the components of X, we may assume that X is irreducible and that I_X is a prime ideal. The formulation (1.1) translates into a system of polynomial equations as follows. We write $J(f)$ for the $s \times n$ Jacobian matrix, whose entry in row i and column j is the partial derivative $\partial f_i / \partial x_j$. The singular locus X_{sing} of X is defined by

$$I_{X_{\text{sing}}} \;=\; I_X + \big\langle c \times c\text{-minors of } J(f) \big\rangle,$$

where c is the codimension of X. That ideal is often hard to compute, unless c is small, and it is usually not radical. We now augment the Jacobian matrix $J(f)$ with the row vector $u - x$ to get an $(s+1) \times n$-matrix. That matrix has rank $\leq c$ on the critical points of d_u on X. From the subvariety of X defined by these rank constraints we must remove contributions from the singular locus X_{sing}. Thus the *critical ideal* for $u \in \mathbb{C}^n$ is the following saturation:

$$\left(I_X + \left\langle (c+1) \times (c+1)\text{-minors of } \begin{pmatrix} u - x \\ J(f) \end{pmatrix} \right\rangle \right) : \big(I_{X_{\text{sing}}}\big)^{\infty}. \tag{2.1}$$

Note that if I_X were not radical, then the above ideal could have an empty variety.

Lemma 2.1. *For generic $u \in \mathbb{C}^n$, the variety of the critical ideal in $\mathbb{C}^n$ is finite. It consists precisely of the critical points of the squared distance function d_u on the manifold $X \backslash X_{\text{sing}}$.*

Proof. For fixed $x \in X \backslash X_{\text{sing}}$, the Jacobian $J(f)$ has rank c, so the $(c+1) \times (c+1)$-minors of $\begin{pmatrix} u - x \\ J(f) \end{pmatrix}$ define an affine-linear subspace of dimension c worth of u's. Hence the variety of pairs $(x, u) \in X \times \mathbb{C}^n$ that are zeros of (2.1) is irreducible of dimension n. The fiber of its projection into the second factor over a generic point $u \in \mathbb{C}^n$ must hence be finite. □

The *ED degree* of X is defined to be the number of critical points in Lemma 2.1. We start with two examples that are familiar to all students of applied mathematics.

Example 2.2 (Linear Regression)**.** Every linear space X has ED degree 1. Here the critical equations (1.1) take the form $x \in X$ and $u - x \perp X$. These linear equations have a unique solution u^*. If u and X are real then u^* is the unique point in X that is closest to u. ♢

Example 2.3 (The Eckart-Young Theorem)**.** Fix positive integers $r \leq s \leq t$ and set $n = st$. Let X be the variety of $s \times t$-matrices of rank $\leq r$. This determinantal variety has

$$\text{EDdegree}(X) \;=\; \binom{s}{r}. \tag{2.2}$$

To see this, we consider a generic real $s \times t$-matrix U and its *singular value decomposition*

$$U \;=\; T_1 \cdot \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_s) \cdot T_2. \tag{2.3}$$

Here $\sigma_1 > \sigma_2 > \cdots > \sigma_s$ are the singular values of U, and T_1 and T_2 are orthogonal matrices of format $s \times s$ and $t \times t$ respectively. According to the Eckart-Young Theorem,

$$U^* \;=\; T_1 \cdot \mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) \cdot T_2$$

is the closest rank r matrix to U. More generally, the critical points of d_U are

$$T_1 \cdot \mathrm{diag}(0, \ldots, 0, \sigma_{i_1}, 0, \ldots, 0, \sigma_{i_r}, 0, \ldots, 0) \cdot T_2$$

where $I = \{i_1 < \ldots < i_r\}$ runs over all r-element subsets of $\{1, \ldots, s\}$. This yields the formula (2.2). The case $r = 1, s = t = 2$ was featured in Example 1.2. ♢

Example 2.4. The following `Macaulay2` code computes the ED degree of a variety in $\mathbb{R}^3$:

```
R = QQ[x1,x2,x3]; I = ideal(x1^5+x2^5+x3^5); u = {5,7,13};
sing = I + minors(codim I,jacobian(I));
M = (matrix{apply(# gens R,i->(gens R)_i-u_i)})||(transpose(jacobian I));
J = saturate(I + minors((codim I)+1,M), sing);
dim J, degree J
```

We chose a random vector `u` as input for the above computation. The output reveals that the Fermat quintic cone $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \,:\, x_1^5 + x_2^5 + x_3^5 = 0\}$ has ED degree 23. ♢

Here is a general upper bound on the ED degree in terms of the given polynomials f_i.

Proposition 2.5. *Let X be a variety of codimension c in $\mathbb{C}^n$ that is cut out by polynomials $f_1, f_2, \ldots, f_c, \ldots, f_s$ of degrees $d_1 \geq d_2 \geq \cdots \geq d_c \geq \cdots \geq d_s$. Then*

$$\mathrm{EDdegree}(X) \;\leq\; d_1 d_2 \cdots d_c \cdot \sum_{i_1+i_2+\cdots+i_c \leq n-c} (d_1 - 1)^{i_1} (d_2 - 1)^{i_2} \cdots (d_c - 1)^{i_c}.$$

Equality holds when X is a generic complete intersection of codimension c (hence $c = s$).

This result can be derived from our Chern class formula given in Theorem 7.7, and from Theorem 6.11 which relates the ED degree of an affine variety and of its projective closure. For details see Example 7.8. A similar bound for the ML degree appears in [23, Theorem 5].

Many varieties arising in applications are rational and they are presented by a parametrization $\psi : \mathbb{R}^m \to \mathbb{R}^n$ whose coordinates ψ_i are rational functions in m unknowns $t = (t_1, \ldots, t_m)$. Instead of first computing the ideal of X by implicitization and then following the approach above, we can use the parametrization directly to compute the ED degree of X.

The squared distance function in terms of the parameters equals

$$D_u(t) \;=\; \sum_{i=1}^{n} (\psi_i(t) - u_i)^2.$$

The equations we need to solve are given by m rational functions in m unknowns:

$$\frac{\partial D_u}{\partial t_1} \;=\; \cdots \;=\; \frac{\partial D_u}{\partial t_m} \;=\; 0. \tag{2.4}$$

The critical locus in $\mathbb{C}^m$ is the set of all solutions to (2.4) at which the Jacobian of ψ has maximal rank. The closure of the image of this set under ψ coincides with the variety of (2.1). Hence, if the parametrization ψ is generically finite-to-one of degree k, then the critical locus in $\mathbb{C}^m$ is finite, by Lemma 2.1, and its cardinality equals $k \cdot \mathrm{EDdegree}(X)$.

In analogy to Proposition 2.5, we can ask for the ED degree when generic polynomials are used in the parametrization of X. Suppose that $n-m$ of the n polynomials $\psi_i(t)$ have degree $\leq d$, while the remaining m polynomials are generic of degree d. Then Bézout's Theorem implies

$$\mathrm{EDdegree}(X) \;=\; (2d-1)^m. \tag{2.5}$$

The following example is non-generic. It demonstrates the effect of scaling coordinates.

Example 2.6. Let $m=2, n=4$ and consider the map $\psi(t_1,t_2)=(t_1^3, t_1^2t_2, t_1t_2^2, t_2^3)$, which has degree $k=3$. Its image $X \subset \mathbb{C}^4$ is the cone over the twisted cubic curve. The system (2.4) consists of two quintics in t_1, t_2, so Bézout's Theorem predicts $25 = 5\times 5$ solutions. The origin is a zero of multiplicity 4 and maps to a singular point of X. The critical locus in $\mathbb{C}^2$ consists of $21 = 25-4$ points. We conclude that the toric surface X has $\mathrm{EDdegree}(X) = 21/k = 7$.

Next we change the parametrization by scaling the middle two monomials as follows:

$$\widetilde{\psi}(t_1,t_2) \;\;=\;\; \big(\, t_1^3\,,\, \sqrt{3}t_1^2t_2\,,\, \sqrt{3}t_1t_2^2\,,\, t_3^3\,\big). \tag{2.6}$$

We still have $k=3$. Now, the function whose critical points we are counting has the form

$$\widetilde{D}(t_1,t_2) \;=\; (t_1^3-a)^2 + 3(t_1^2t_2-b)^2 + 3(t_1t_2^2-c)^2 + (t_2^3-d)^2,$$

where a,b,c,d are random scalars. A computation shows that the number of complex critical points of $\widetilde{D}$ equals 9. So, the corresponding toric surface $\widetilde{X}$ has $\mathrm{EDdegree}(\widetilde{X}) = 9/k = 3$. $\diamondsuit$

The variety $X \subset \mathbb{C}^n$ is an *affine cone* if $x \in X$ implies $\lambda x \in X$ for all $\lambda \in \mathbb{C}$. This means that I_X is a homogeneous ideal in $\mathbb{R}[x_1,\ldots,x_n]$. By slight abuse of notation, we identify X with the projective variety given by I_X in $\mathbb{P}^{n-1}$. The former is the affine cone over the latter.

We define the *ED degree of a projective variety* in $\mathbb{P}^{n-1}$ to be the ED degree of the corresponding affine cone in $\mathbb{C}^n$. For instance, in Example 2.6 we considered two twisted cubic curves X and $\widetilde{X}$ that lie in $\mathbb{P}^3$. These curves have ED degrees 3 and 7 respectively.

To take advantage of the homogeneity of the generators of I_X, and of the geometry of projective space $\mathbb{P}^{n-1}$, we replace (2.1) with the following homogeneous ideal in $\mathbb{R}[x_1,\ldots,x_n]$:

$$\left(I_X + \left\langle (c+2)\times(c+2)\text{-minors of } \begin{pmatrix} u \\ x \\ J(f) \end{pmatrix} \right\rangle\right) : \big(I_{X_{\mathrm{sing}}} \cdot \langle x_1^2+\cdots+x_n^2\rangle\,\big)^{\infty}. \tag{2.7}$$

The singular locus of an affine cone is the cone over the singular locus of the projective variety. They are defined by the same ideal $I_{X_{\mathrm{sing}}}$. The *isotropic quadric* $Q = \{x \in \mathbb{P}^{n-1} :$

$x_1^2 + \cdots + x_n^2 = 0\}$ plays a special role, seen clearly in the proof of Lemma 2.7. In particular, the role of Q exhibits that the computation of ED degree is a metric problem. Note that Q has no real points. The `Macaulay2` code in Example 2.4 can be adapted to verify $\mathrm{EDdegree}(Q) = 0$.

The following lemma concerns the transition between affine cones and projective varieties.

Lemma 2.7. *Fix an affine cone $X \subset \mathbb{C}^n$ and a data point $u \in \mathbb{C}^n \backslash X$. Let $x \in X \backslash \{0\}$ such that the corresponding point $[x]$ in $\mathbb{P}^{n-1}$ does not lie in the isotropic quadric Q. Then $[x]$ lies in the projective variety of (2.7) if and only if some scalar multiple λx of x lies in the affine variety of (2.1). In that case, the scalar λ is unique.*

Proof. Since both ideals are saturated with respect to $I_{X_{\mathrm{sing}}}$, it suffices to prove this under the assumption that $x \in X \backslash X_{\mathrm{sing}}$, so that the Jacobian $J(f)$ at x has rank c. If $u - \lambda x$ lies in the row space of $J(f)$, then the span of u, x, and the rows of $J(f)$ has dimension at most $c+1$. This proves the only-if direction. Conversely, suppose that $[x]$ lies in the variety of (2.7). First assume that x lies in the row span of $J(f)$. Then $x = \sum \lambda_i \nabla f_i(x)$ for some $\lambda_i \in \mathbb{C}$. Now recall that if f is a homogeneous polynomial in $\mathbb{R}[x_1, \ldots, x_n]$ of degree d, then $x \cdot \nabla f(x) = d\, f(x)$. Since $f_i(x) = 0$ for all i, we find that $x \cdot \nabla f_i(x) = 0$ for all i, which implies that $x \cdot x = 0$, i.e., $[x] \in Q$. This contradicts the hypothesis, so the matrix $\begin{pmatrix} x \\ J(f) \end{pmatrix}$ has rank $c+1$. But then $u - \lambda x$ lies in the row span of $J(f)$ for a unique $\lambda \in \mathbb{C}$. □

Every real algebraic variety satisfies the assumption of the following corollary.

Corollary 2.8. *Let X be a variety in $\mathbb{P}^{n-1}$ that is not contained in the isotropic quadric Q, and let u be generic. Then $\mathrm{EDdegree}(X)$ is equal to the number of zeros of (2.7) in $\mathbb{P}^{n-1}$.*

Proof. Since $X \not\subseteq Q$ and u is generic, none of the critical points of d_u in $X \backslash X_{\mathrm{sing}}$ will lie in Q. The claim follows from Lemma 2.7. For further details see Theorems 4.1 and 4.4. □

Corollary 2.8 implies that Proposition 2.5 holds almost verbatim for projective varieties.

Corollary 2.9. *Let X be a variety of codimension c in $\mathbb{P}^{n-1}$ that is cut out by homogeneous polynomials $F_1, F_2, \ldots, F_c, \ldots, F_s$ of degrees $d_1 \geq d_2 \geq \cdots \geq d_c \geq \cdots \geq d_s$. Then*

$$\mathrm{EDdegree}(X) \;\leq\; d_1 d_2 \cdots d_c \cdot \sum_{i_1 + i_2 + \cdots + i_c \leq n-c-1} (d_1 - 1)^{i_1} (d_2 - 1)^{i_2} \cdots (d_c - 1)^{i_c}. \qquad (2.8)$$

Equality holds when X is a generic complete intersection of codimension c in $\mathbb{P}^{n-1}$.

Fixing the codimension c of X is essential in Proposition 2.5 and Corollary 2.9. Without this hypothesis, the bounds do not hold. In Example 5.10, we display homogeneous polynomials $F_1, \ldots, F_c$ of degrees $d_1, \ldots, d_c$ whose variety has ED degree larger than (2.8).

Example 2.10. The following `Macaulay2` code computes the ED degree of a curve in $\mathbb{P}^2$:

```
R = QQ[x1,x2,x3]; I = ideal(x1^5+x2^5+x3^5);  u = {5,7,13};
sing = minors(codim I,jacobian(I));
M = matrix {u}||matrix {gens R}||(transpose(jacobian I));
J = saturate(I+minors((codim I)+2,M), sing*ideal(x1^2+x2^2+x3^2));
dim J, degree J
```

The output confirms that the Fermat quintic curve given by $x_1^5+x_2^5+x_3^5=0$ has ED degree 23. By contrast, as seen from Corollary 2.9, a general curve of degree five in $\mathbb{P}^2$ has ED degree 25. Saturating with $I_{X_{\rm sing}}$ alone in the fourth line of the code would yield 25. $\diamondsuit$

It should be stressed that the ideals (2.1) and (2.7), and our two `Macaulay2` code fragments, are blueprints for first computations. In order to succeed with larger examples, it is essential that these formulations be refined. For instance, to express rank conditions on a polynomial matrix M, the determinantal constraints are often too large, and it is better to add a matrix equation of the form $\Lambda \cdot M = 0$, where Λ is a matrix filled with new unknowns. This leads to a system of bilinear equations, so the methods of Faugère *et al.* [14] can be used. We also recommend trying tools from numerical algebraic geometry, such as `Bertini` [3].

3 First applications

The problem of computing the closest point on a variety arises in numerous applications. In this section we discuss some concrete instances, and we explore the ED degree in each case.

Example 3.1 (Geometric modeling)**.** Thomassen *et al.* [36] study the nearest point problem for a parametrized surface X in $\mathbb{R}^3$. The three coordinates of their parametrization $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ are polynomials in the parameters (t_1, t_2) that have degree d_1 in t_1 and degree d_2 in t_2. The image $X = \overline{\psi(\mathbb{R}^2)}$ is a *Bézier surface* of bidegree (d_1, d_2). It is shown in [36, §3] that

$$\text{EDdegree}(X) \;=\; 4d_1d_2 + (2d_1-1)(2d_2-1).$$

This refines the Bézout bound in (2.5). The authors of [36] demonstrate how to solve the critical equations $\partial D_u/\partial t_1 = \partial D_u/\partial t_2 = 0$ with resultants based on moving surfaces. $\diamondsuit$

Example 3.2 (The closest symmetric matrix)**.** Let X denote the variety of symmetric $s\times s$-matrices of rank $\leq r$. The nearest point problem for X asks the following question: given a symmetric $s \times s$-matrix $U = (U_{ij})$, find the rank r matrix U^* that is closest to U. There are two natural interpretations of this question in the Euclidean distance context, depending on whether $n = s^2$ or $n = \binom{s+1}{2}$. In the first case, we regard X as a subvariety of the space $\mathbb{R}^{s\times s}$ of all $s \times s$-matrices, and in the second case, the coordinates are the entries U_{ij} for $1 \leq i \leq j \leq s$. The difference lies in which of the following two functions we are minimizing:

$$D_U \;=\; \sum_{i=1}^{s}\sum_{j=1}^{s}\bigl(U_{ij} - \sum_{k=1}^{r} t_{ik}t_{kj}\bigr)^2 \qquad \text{or} \qquad D_U \;=\; \sum_{1\leq i\leq j\leq s}\bigl(U_{ij} - \sum_{k=1}^{r} t_{ik}t_{kj}\bigr)^2. \tag{3.1}$$

These unconstrained optimization problems use the parametrization of symmetric $s \times s$-matrices of rank r that comes from multiplying an $s \times r$ matrix $T = (t_{ij})$ with its transpose. The two formulations are dramatically different as far as the ED degree is concerned. On the left side, the Eckart-Young Theorem applies, and $\text{EDdegree}(X) = \binom{s}{r}$ as in Example 2.3. On the right side, $\text{EDdegree}(X)$ is much larger than $\binom{s}{r}$. For instance, for $s = 3$ and $r = 1$ or 2,

$$\text{EDdegree}(X) = 3 \qquad \text{and} \qquad \text{EDdegree}(X) = 13. \tag{3.2}$$

The two ideals that represent the constrained optimization problems equivalent to (3.1) are

$$\left\langle 2\times 2\text{-minors of } \begin{pmatrix} \sqrt{2}x_{11} & x_{12} & x_{13} \\ x_{12} & \sqrt{2}x_{22} & x_{23} \\ x_{13} & x_{23} & \sqrt{2}x_{33} \end{pmatrix} \right\rangle \text{ and } \left\langle 2\times 2\text{-minors of } \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} \right\rangle. \tag{3.3}$$

These equivalences can be seen via a change of variables. For example, for the left ideal in (3.3), the constrained optimization problem is to minimize $\sum_{1\le i\le j\le 3}(u_{ij}-x_{ij})^2$ subject to the nine quadratic equations $2x_{11}x_{22}=x_{12}^2, \sqrt{2}x_{11}x_{23}=x_{12}x_{13},\ldots,2x_{22}x_{33}=x_{23}^2$. Now making the change of variables $x_{ii}=X_{ii}$ for $i=1,2,3$ and $x_{ij}=\sqrt{2}X_{ij}$ for $1\le i<j\le 3$, and similarly, $u_{ii}=U_{ii}$ for $i=1,2,3$ and $u_{ij}=\sqrt{2}U_{ij}$ for $1\le i<j\le 3$, we get the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{3}(U_{ii}-X_{ii})^2 + \sum_{1\le i<j\le 3} 2(U_{ij}-X_{ij})^2 \\ \text{subject to} & X_{ik}X_{jl}=X_{il}X_{jk} \ \text{ for } \ 1\le i<j\le 3, 1\le k<l\le 3. \end{array}$$

This is equivalent to the left problem in (3.1) for $r=1$ via the parametrization $X_{ij}=t_it_j$. The appearance of $\sqrt{2}$ in the left matrix in (3.3) is analogous to the appearance of $\sqrt{3}$ in Example 2.6. In Example 5.6 we discuss a general ED degree formula for symmetric $s\times s$-matrices of rank $\le r$ that works for the version on the right. The same issue for ML degrees is the difference between "scaled" and "unscaled" in the table at the end of [23, §5]. ♢

Example 3.3 (Computer vision)**.** This article got started with the following problem from [1, 19, 34]. Let $A_1,A_2,\ldots,A_n$ be generic real 3×4 matrices that model n cameras. The associated *multiview variety* is the closure of the image of the map $\mathbb{P}^3 \dashrightarrow (\mathbb{P}^2)^n, y\mapsto (A_1y,A_2y,\ldots,A_ny)$. Its defining prime ideal I_n is multi-homogeneous and lives in the polynomial ring $\mathbb{R}[x_{ij}: i=1,\ldots,n, j=0,1,2]$, where $(x_{i0}:x_{i1}:x_{i2})$ are homogeneous coordinates of the i-th plane $\mathbb{P}^2$. Explicit determinantal generators and Gröbner bases for I_n are derived in [1]. If we dehomogenize I_n by setting $x_{i0}=1$ for $i=1,2,\ldots,n$, then we get a 3-dimensional affine variety X_n in $\mathbb{R}^{2n}=(\mathbb{R}^2)^n$. Note that I_n and X_n depend on the choice of the matrices $A_1,A_2,\ldots,A_n$. This dependence is governed by the Hilbert scheme in [1].

The Euclidean distance problem for X_n is known in computer vision as *n-view triangulation*. Following [19] and [34], the data $u\in\mathbb{R}^{2n}$ are n noisy images of a point in $\mathbb{R}^3$ taken by the n cameras. The maximum likelihood solution of the recovery problem with Gaussian noise is the configuration $u^*\in X_n$ of minimum distance to u. For $n=2$, the variety X_2 is a hypersurface cut out by a bilinear polynomial $(1,x_{11},x_{12})M(1,x_{21},x_{22})^T$, where M is a 3×3-matrix of rank 2. Hartley and Sturm [19] studied the critical equations and found that $\text{EDdegree}(X_2)=6$. Their computations were extended by Stewénius *et al.* [34] up to $n=7$:

n	2	3	4	5	6	7
$\text{EDdegree}(X_n)$	6	47	148	336	638	1081

This table suggests the conjecture that these ED degrees grow as a cubic polynomial:

Conjecture 3.4. *The Euclidean distance degree of the affine multiview variety X_n equals*

$$\text{EDdegree}(X_n) \quad = \quad \frac{9}{2}n^3-\frac{21}{2}n^2+8n-4.$$

At present we do not know how to prove this. Our first idea was to replace the affine threefold X_n by a projective variety. For instance, consider the closure $\overline{X}_n$ of X_n in $\mathbb{P}^{2n}$. Alternatively, we can regard I_n as a homogeneous ideal in the usual $\mathbb{Z}$-grading, thus defining a projective variety Y_n in $\mathbb{P}^{3n-1}$. However, for $n \geq 3$, the ED degrees of both $\overline{X}_n$ and Y_n are larger than the ED degree of X_n. For instance, in the case of three cameras we have

$$\text{EDdegree}(X_3) = 47 \quad < \quad \text{EDdegree}(\overline{X}_3) \;=\; 112 \quad < \quad \text{EDdegree}(Y_3) \;=\; 148.$$

Can one find a natural reformulation of Conjecture 3.4 in terms of projective geometry? ♢

Many problems in engineering lead to minimizing the distance from a given point u to an algebraic variety. One such problem is detecting voltage collapse and blackouts in electrical power systems [29, page 94]. It is typical to model a power system as a differential equation $\dot{x} = f(x, \lambda)$ where x is the state and λ is the parameter vector of load powers. As λ varies, the state moves around. At critical load powers, the system can lose equilibrium and this results in a blackout due to voltage collapse. The set of critical λ's form an algebraic variety X that one wants to stay away from. This is done by calculating the closest point on X to the current set of parameters λ_0 used by the power system. A similar, and very well-known, problem from *control theory* is to ensure the *stability* of a univariate polynomial.

n	EDdegree(Γ_n)	EDdegree($\bar{\Gamma}_n$)	aEDdegree(Γ_n)	aEDdegree($\bar{\Gamma}_n$)
3	5	2	1.162...	2
4	5	10	1.883...	2.068...
5	13	6	2.142...	3.052...
6	9	18	2.416...	3.53...
7	21	10	2.66...	3.742...

Table 1: ED degrees and average and ED degrees of small Hurwitz determinants.

Example 3.5 (Hurwitz stability)**.** Consider a univariate polynomial with real coefficients,

$$u(z) \;=\; u_0 z^n + u_1 z^{n-1} + u_2 z^{n-2} + \cdots + u_{n-1} z + u_n.$$

We say that $u(z)$ is *stable* if each of its n complex zeros has negative real part. It is an important problem in control theory to check whether a given polynomial $u(z)$ is stable, and, if not, to find a polynomial $x(z)$ in the closure of the stable locus that is closest to $u(z)$.

The stability of $x(z) = \sum_{i=0}^{n} x_i z^i$ is characterized by the following *Hurwitz test*. The nth *Hurwitz matrix* is an $n \times n$ matrix with $x_1, \ldots, x_n$ on the diagonal. Above the diagonal entry x_i in column i, we stack as much of $x_{i+1}, x_{i+2}, \ldots, x_n$ consecutively, followed by zeros if there is extra room. Similarly, below x_i, we stack as much of $x_{i-1}, x_{i-2}, \ldots, x_1, x_0$ consecutively, followed by zeros if there is extra room. The Hurwitz test says that $x(z)$ is stable if and only if every leading principal minor of H_n is positive. For instance, for $n = 5$ we have

$$H_5 \;=\; \begin{pmatrix} x_1 & x_3 & x_5 & 0 & 0 \\ x_0 & x_2 & x_4 & 0 & 0 \\ 0 & x_1 & x_3 & x_5 & 0 \\ 0 & x_0 & x_2 & x_4 & 0 \\ 0 & 0 & x_1 & x_3 & x_5 \end{pmatrix}.$$

The ratio $\bar{\Gamma}_n = \det(H_n)/x_n$, which is the $(n-1)$th leading principal minor of H_n, is a homogeneous polynomial in the variables $x_0, \dots, x_{n-1}$ of degree $n-1$. Let Γ_n denote the non-homogeneous polynomial obtained by setting $x_0 = 1$ in $\bar{\Gamma}_n$. We refer to Γ_n resp. $\bar{\Gamma}_n$ as the non-homogeneous resp. homogeneous *Hurwitz determinant*. Table 1 shows the ED degrees and the average ED degrees of both Γ_n and $\bar{\Gamma}_n$ for some small values of n. The average ED degree was computed with respect to the standard multivariate Gaussian distribution in $\mathbb{R}^n$ or $\mathbb{R}^{n+1}$ centered at the origin. For the formal definition of aEDdegree($\cdot$) see Section 4. The first two columns in Table 1 seem to be oscillating by parity. Theorem 3.6 explains this. Interestingly, the oscillating behavior does not occur for average ED degree. ♢

Theorem 3.6. *The ED degrees of the Hurwitz determinants are given by the following table:*

	EDdegree(Γ_n)	EDdegree($\bar{\Gamma}_n$)
$n = 2m+1$	$8m-3$	$4m-2$
$n = 2m$	$4m-3$	$8m-6$

Proof. The hypersurface $X = V(\bar{\Gamma}_n)$ defines the boundary of the stability region. If a polynomial $x(z)$ lies on X, then it has a complex root on the imaginary axis, so it admits a factorization $x(z) = (cz^2 + d)(b_0 z^{n-2} + \cdots + b_{n-2})$. This representation yields a parametrization of the hypersurface $X \subset \mathbb{P}^n$ with parameters $b_0, \dots, b_{n-2}, c, d$. We can rewrite this as

$$x := \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c & & & & \\ 0 & c & & & \\ d & 0 & c & & \\ & \ddots & \ddots & \ddots & \\ & & d & 0 & c \\ & & & d & 0 \\ & & & & d \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-2} \end{bmatrix} =: C \cdot b.$$

Where this parametrization is regular and x is a smooth point of X, the tangent space $T_x X$ is spanned by the columns of C and the vectors b', b'' obtained by appending or prepending two zeros to b, respectively. Thus, for $u \in \mathbb{C}^{n+1}$, the condition $u - Cb \perp T_x X$ translates into

$$C^T(u - Cb) = 0 \quad \text{and} \quad (b')^T(u - Cb) = 0 \quad \text{and} \quad (b'')^T(u - Cb) = 0.$$

The first equation expresses b as a rational homogenous function in c, d, namely, $b = b(c,d) = (C^T C)^{-1} C^T u$. By Cramer's rule, the entries of the inverse of a matrix are homogeneous rational functions of degree -1. Hence the degree of $b(c,d)$ equals $-2 + 1 = -1$. Let $\gamma = \gamma(c,d)$ be the denominator of $(C^T C)^{-1}$, i.e., the lowest-degree polynomial in c, d for which $\gamma \cdot (C^T C)^{-1}$ is polynomial; and let N be the degree of γ. Then $\gamma b'$ has entries that are homogeneous polynomials in c, d of degree $N - 1$. Similarly, $\gamma u - \gamma C b$ has degree N. Hence

$$p(c,d) \ := \ (\gamma b') \cdot (\gamma u - \gamma C b) \quad \text{and} \quad q(c,d) \ := \ (\gamma b'') \cdot (\gamma u - \gamma C b)$$

are homogeneous polynomials of degree $2N - 1$ that vanish on the desired points $(c : d) \in \mathbb{P}^1$. Indeed, if p and q vanish on $(c : d)$ and $\gamma(c,d)$ is non-zero, then there is a unique b that makes (b, c, d) critical for the data u. It turns out that p is divisible by d, that q is divisible by c, and that $p/d = q/c$. Thus $2N-2$ is an upper bound for EDdegree(X).

To compute γ, note that $C^T C$ decomposes into two blocks, corresponding to even and odd indices. When $n = 2m+1$ is odd, these two blocks are identical, and γ equals their determinant, which is $c^{2m} + c^{2m-2}d^2 + \cdots + d^{2m}$. Hence $N = 2m$. When $n = 2m$ is even, the two blocks are distinct, and γ equals the product of their determinants, which is $(c^{2m} + c^{2m-2}d^2 + \cdots + d^{2m})(c^{2m-2} + \cdots + d^{2m-2})$. Hence $N = 4m-2$. In both cases one can check that p/d is irreducible, and this implies that EDdegree$(X) = 2N-2$. This establishes the stated formula for EDdegree$(\bar{\Gamma}_n)$. A similar computation can be performed in the non-homogeneous case, by setting $x_0 = b_0 = c = 1$, leading to the formula for EDdegree(Γ_n). $\square$

Example 3.7 (Interacting agents)**.** This concerns a problem we learned from work of Anderson and Helmke [2]. Let X denote the variety in $\mathbb{R}^{\binom{p}{2}}$ with parametric representation

$$d_{ij} \;=\; (z_i - z_j)^2 \quad \text{for} \quad 1 \leq i < j \leq p. \tag{3.4}$$

Thus, the points in X record the squared distances among p interacting agents with coordinates $z_1, z_2, \ldots, z_p$ on the line $\mathbb{R}^1$. Note that X is the cone over a projective variety in $\mathbb{P}^{\binom{p}{2}-1}$. The prime ideal of X is given by the 2×2-minors of the *Cayley-Menger matrix*

$$\begin{bmatrix} 2d_{1p} & d_{1p}{+}d_{2p}{-}d_{12} & d_{1p}{+}d_{3p}{-}d_{13} & \cdots & d_{1p}{+}d_{p-1,p}{-}d_{1,p-1} \\ d_{1p}{+}d_{2p}{-}d_{12} & 2d_{2p} & d_{2p}{+}d_{3p}{-}d_{23} & \cdots & d_{2p}{+}d_{p-1,p}{-}d_{2,p-1} \\ d_{1p}{+}d_{3p}{-}d_{13} & d_{2p}{+}d_{3p}{-}d_{23} & 2d_{3p} & \cdots & d_{3p}{+}d_{p-1,p}{-}d_{3,p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1p}{+}d_{p-1,p}{-}d_{1,p-1} & d_{2p}{+}d_{p-1,p}{-}d_{2,p-1} & d_{3p}{+}d_{p-1,p}{-}d_{3,p-1} & \cdots & 2d_{p-1,p} \end{bmatrix} \tag{3.5}$$

Indeed, under the parametrization (3.4), the $(p-1) \times (p-1)$ matrix (3.5) factors as $2Z^T Z$, where Z is the row vector $(z_1{-}z_p, z_2{-}z_p, z_3{-}z_p, \ldots, z_{p-1}{-}z_p)$. We can form the Cayley-Menger matrix (3.5) for any finite metric space on p points. The metric space can be embedded in a Euclidean space if and only if (3.5) is positive semidefinite [25, (8)]. That Euclidean embedding is possible in dimension r if and only if the rank of (3.5) is at most r.

The following theorem is inspired by [2] and provides a refinement of results therein. In particular, it explains the findings in [2, §4] for $p \leq 4$. There is an extra factor of $1/2$ because of the involution $z \mapsto -z$ on the fibers of the map (3.4). For instance, for $p = 4$, our formula gives EDdegree$(X) = 13$ while [2, Theorem 13] reports 26 non-zero critical points. The most interesting case occurs when p is divisible by 3, and this will be explained in the proof.

Theorem 3.8. *The ED degree of the Cayley-Menger variety* $X \subset \mathbb{P}^{\binom{p}{2}-1}$ *equals*

$$\text{EDdegree}(X) \;=\; \begin{cases} \frac{3^{p-1}-1}{2} & \text{if } p \equiv 1,2 \mod 3 \\ \frac{3^{p-1}-1}{2} - \frac{p!}{3((p/3)!)^3} & \text{if } p \equiv 0 \mod 3 \end{cases} \tag{3.6}$$

Proof. After the linear change of coordinates given by $x_{ii} = 2d_{ip}$ and $x_{ij} = d_{ip} + d_{jp} - d_{ij}$, the Cayley-Menger variety X agrees with the symmetric $(p-1) \times (p-1)$-matrices of rank 1. This is the Veronese variety for $d = 2$. The number $(3^{p-1}-1)/2$ is a special instance of the formula in Proposition 7.9. To show that it is valid here, we need to prove that X intersects the isotropic quadric Q transversally, i.e., the intersection $X \cap Q$ is non-singular. If there are isolated nodal singular points, then their number gets subtracted.

The parametrization (3.4) defines the second Veronese embedding $\mathbb{P}^{p-2} \to X \subset \mathbb{P}^{\binom{p}{2}-1}$, where $\mathbb{P}^{p-2}$ is the projective space of the quotient $\mathbb{C}^p/\mathbb{C}\cdot(1,\ldots,1)$. So $X \cap Q$ is isomorphic to its inverse image in $\mathbb{P}^{p-2}$ under this map. That inverse image is the hypersurface in $\mathbb{P}^{p-2}$ defined by the homogeneous quartic $f = \sum_{1\leq i<j\leq p}(z_i - z_j)^4$. We need to analyze the singular locus of the hypersurface $V(f)$ in $\mathbb{P}^{p-2}$, which is the variety defined by all partial derivatives of f. Arguing modulo 3 one finds that if p is not divisible by 3 then $V(f)$ is smooth, and then we have $\text{EDdegree}(X) = (3^{p-1}-1)/2$. If p is divisible by 3 then $V(f)$ is not smooth, but $V(f)_{sing}$ consists of isolated nodes that form one orbit under permuting coordinates. One representative is the point in $\mathbb{P}^{p-2}$ represented by the vector

$$\big(0,0,\ldots,0,\,1,1,\ldots,1,\,\xi,\xi,\ldots,\xi\big) \in \mathbb{C}^p \quad \text{where} \;\; \xi^2 - \xi + 1 = 0.$$

The number of singular points of the quartic hypersurface $V(f)$ is equal to

$$\frac{p!}{3\cdot((p/3)!)^3}.$$

This is the number of words either empty or beginning with the first letter of the ternary alphabet $\{0,1,\xi\}$, where each letter of the alphabet occurs p times; see [28, A208881]. □

4 ED correspondence and average ED degree

The ED correspondence arises when the variety X is fixed but the data point u varies. After studying this, we restrict to the real numbers, and we introduce the average ED degree, making precise a notion that was hinted at in Example 3.5. The ED correspondence yields an integral formula for $\text{aEDdegree}(X)$. This integral can sometimes be evaluated in closed form. In other cases, experiments show that evaluating the integral numerically is more efficient than estimating $\text{aEDdegree}(X)$ by sampling u and counting real critical points.

We start with an irreducible affine variety $X \subset \mathbb{C}^n$ of codimension c with prime ideal $I_X = \langle f_1,\ldots,f_s\rangle$ in $\mathbb{R}[x_1,\ldots,x_n]$. The *ED correspondence* $\mathcal{E}_X$ is the subvariety of $\mathbb{C}^n \times \mathbb{C}^n$ defined by the ideal (2.1) in the polynomial ring $\mathbb{R}[x_1,\ldots,x_n,u_1,\ldots,u_n]$. Now, the u_i are unknowns that serve as coordinates on the second factor in $\mathbb{C}^n \times \mathbb{C}^n$. Geometrically, $\mathcal{E}_X$ is the topological closure in $\mathbb{C}^n \times \mathbb{C}^n$ of the set of pairs (x,u) such that $x \in X\backslash X_{\text{sing}}$ is a critical point of d_u. The following theorem implies and enriches Lemma 2.1.

Theorem 4.1. *The ED correspondence $\mathcal{E}_X$ is an irreducible variety of dimension n inside $\mathbb{C}^n \times \mathbb{C}^n$. The first projection $\pi_1 : \mathcal{E}_X \to X \subset \mathbb{C}^n$ is an affine vector bundle of rank c over $X\backslash X_{\text{sing}}$. Over generic data points $u \in \mathbb{C}^n$, the second projection $\pi_2 : \mathcal{E}_X \to \mathbb{C}^n$ has finite fibers $\pi_2^{-1}(u)$ of cardinality equal to* $\text{EDdegree}(X)$*. If, moreover, we have $T_xX \cap (T_xX)^\perp = \{0\}$ at some point $x \in X\backslash X_{\text{sing}}$, then π_2 is a dominant map and* $\text{EDdegree}(X)$ *is positive.*

In our applications, the variety X always has real points that are smooth, i.e. in $X\backslash X_{\text{sing}}$. If this holds, then the last condition in Theorem 4.1 is automatically satisfied: the tangent space at such a point is real and intersects its orthogonal complement trivially. But, for instance, the hypersurface $Q = V(x_1^2+\cdots+x_n^2)$ does not satisfy this condition: at any point $x \in Q$ the tangent space $T_xQ = x^\perp$ intersects its orthogonal complement $\mathbb{C}x$ in all of $\mathbb{C}x$.

Proof. The affine vector bundle property follows directly from the system (1.1) or, alternatively, from the matrix representation (2.1): fixing $x \in X\backslash X_{\text{sing}}$, the fiber $\pi_1^{-1}(x)$ equals $\{x\}\times(x+(T_xX)^\perp)$, where the second factor is an affine space of dimension c varying smoothly with x. Since X is irreducible, so is $\mathcal{E}_X$, and its dimension equals $(n-c)+c = n$. For dimension reasons, the projection π_2 cannot have positive-dimensional fibers over generic data points u, so those fibers are generically finite sets, of cardinality equal to EDdegree(X).

For the last statement, note that the diagonal $\Delta(X) := \{(x,x) \in \mathbb{C}^n \times \mathbb{C}^n \mid x \in X\}$ is contained in $\mathcal{E}_X$. Fix a point $x \in X\backslash X_{\text{sing}}$ for which $T_xX \cap (T_xX)^\perp = \{0\}$. Being an affine bundle over $X\backslash X_{\text{sing}}$, $\mathcal{E}_X$ is smooth at the point (x,x). The tangent space $T_{(x,x)}\mathcal{E}_X$ contains both the tangent space $T_{(x,x)}\Delta(X) = \Delta(T_xX)$ and $\{0\}\times(T_xX)^\perp$. Thus the image of the derivative at x of $\pi_2 : \mathcal{E}_X \to \mathbb{C}^2$ contains both T_xX and $(T_xX)^\perp$. Since these spaces have complementary dimensions and intersect trivially by assumption, they span all of $\mathbb{C}^n$. Thus the derivative of π_2 at (x,x) is surjective onto $\mathbb{C}^n$, and this implies that π_2 is dominant. □

Corollary 4.2. *If X is (uni-)rational then so is the ED correspondence $\mathcal{E}_X$.*

Proof. Let $\psi : \mathbb{C}^m \to \mathbb{C}^n$ be a rational map that parametrizes X, where $m = \dim X = n-c$. Its Jacobian $J(\psi)$ is an $n \times m$-matrix of rational functions in the standard coordinates $t_1,\ldots,t_m$ on $\mathbb{C}^m$. The columns of $J(\psi)$ span the tangent space of X at the point $\psi(t)$ for generic $t \in \mathbb{C}^m$. The left kernel of $J(\psi)$ is a linear space of dimension c. We can write down a basis $\{\beta_1(t),\ldots,\beta_c(t)\}$ of that kernel by applying Cramer's rule to the matrix $J(\psi)$. In particular, the β_j will also be rational functions in the t_i. Now the map

$$\mathbb{C}^m \times \mathbb{C}^c \to \mathcal{E}_X,\ (t,s) \mapsto \left(\psi(t), \psi(t) + \sum_{i=1}^{c} s_i\beta_i(t)\right)$$

is a parametrization of $\mathcal{E}_X$, which is birational if and only if ψ is birational. □

Example 4.3. The twisted cubic cone X from Example 2.6 has the parametrization $\psi : \mathbb{C}^2 \to \mathbb{C}^4$, $(t_1,t_2) \mapsto (t_1^3, t_1^2t_2, t_1t_2^2, t_2^3)$. We saw that EDdegree$(X) = 7$. Here is a parametrization of the ED correspondence $\mathcal{E}_X$ that is induced by the construction in the proof above:

$$\mathbb{C}^2\times\mathbb{C}^2 \to \mathbb{C}^4\times\mathbb{C}^4\,, \quad ((t_1,t_2),(s_1,s_2)) \mapsto$$
$$\big((t_1^3,t_1^2t_2,t_1t_2^2,t_2^3),\ (t_1^3+s_1t_2^2,\ t_1^2t_2-2s_1t_1t_2+s_2t_2^2,\ t_1t_2^2+s_1t_1^2-2s_2t_1t_2,\ t_2^3+s_2t_1^2)\big).$$

The prime ideal of $\mathcal{E}_X$ in $\mathbb{R}[x_1,x_2,x_3,x_4,u_1,u_2,u_3,u_4]$ can be computed from (2.1). It is minimally generated by seven quadrics and one quartic. It is important to note that these generators are homogeneous with respect to the usual $\mathbb{Z}$-grading but not bi-homogeneous.

The formulation (2.7) leads to the subideal generated by all bi-homogeneous polynomials that vanish on $\mathcal{E}_X$. It has six minimal generators, three of degree $(2,0)$ and three of degree $(3,1)$. Geometrically, this corresponds to the variety $\mathcal{PE}_X \subset \mathbb{P}^3\times\mathbb{C}^4$ we introduce next. ♢

If X is an affine cone in $\mathbb{C}^n$, we consider the closure of the image of $\mathcal{E}_X\cap((\mathbb{C}^n\backslash\{0\})\times\mathbb{C}^n)$ under the map $(\mathbb{C}^n\backslash\{0\})\times\mathbb{C}^n \to \mathbb{P}^{n-1}\times\mathbb{C}^n$, $(x,u)\mapsto([x],u)$. This closure is called the *projective ED correspondence* of X, and it is denoted $\mathcal{PE}_X$. It has the following properties.

Theorem 4.4. *Let $X \subseteq \mathbb{C}^n$ be an affine cone not contained in the isotropic quadric Q. Then the projective ED correspondence $\mathcal{PE}_X$ of X is an n-dimensional irreducible variety in $\mathbb{P}^{n-1} \times \mathbb{C}^n$. It is the zero set of the ideal (2.7). Its projection onto the projective variety in $\mathbb{P}^{n-1}$ given by X is a vector bundle over $X \backslash (X_{\text{sing}} \cup Q)$ of rank $c+1$. The fibers over generic data points u of its projection onto $\mathbb{C}^n$ are finite of cardinality equal to* EDdegree(X).

Proof. The first statement follows from Lemma 2.7: let $x \in X \backslash (X_{\text{sing}} \cup Q)$ and $u \in \mathbb{C}^n$. First, if $(x, u) \in \mathcal{E}_X$, then certainly $([x], u)$ lies in the variety of the ideal (2.1). Conversely, if $([x], u)$ lies in the variety of that ideal, then there exists a (unique) λ such that $(\lambda x, u) \in \mathcal{E}_X$. If λ is non-zero, then this means that $([x], u)$ lies in the projection of $\mathcal{E}_X$. If λ is zero, then $u \perp T_x X$ and hence $(\epsilon x, \epsilon x + u) \in \mathcal{E}_X$ for all $\epsilon \in \mathbb{C}$. The limit of $([\epsilon x], \epsilon x + u)$ for $\epsilon \to 0$ equals $([x], u)$, so the latter point still lies in the closure of the image of $\mathcal{E}_X$, i.e., in the projective ED correspondence. The remaining statements are proved as in the proof of Theorem 4.1. □

We now turn our attention to the average ED degree of a real affine variety X in $\mathbb{R}^n$. In applications, the data point u also lies in $\mathbb{R}^n$, and u^* is the unique closest point to u in X. The quantity EDdegree(X) measures the algebraic complexity of writing the optimal solution u^* as a function of the data u. But when applying other, non-algebraic methods for finding u^*, the number of *real-valued* critical points of d_u for randomly sampled data u is of equal interest. In contrast with the number of complex-valued critical points, this number is typically not constant for all generic u, but rather constant on the connected components of the complement of an algebraic hypersurface $\Sigma_X \subset \mathbb{R}^n$, which we call the *ED discriminant*. To get, nevertheless, a meaningful count of the critical points, we propose to average over all u with respect to a measure on $\mathbb{R}^n$. In the remainder of this section, we describe how to compute that average using the ED correspondence. Our method is particularly useful in the setting of Corollary 4.2, i.e., when X and hence $\mathcal{E}_X$ have rational parametrizations.

We equip data space $\mathbb{R}^n$ with a volume form ω whose associated density $|\omega|$ satisfies $\int_{\mathbb{R}^n} |\omega| = 1$. A common choice for ω is the standard multivariate Gaussian $\frac{1}{(2\pi)^{n/2}} e^{-||x||^2/2}\, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n$. This choice is natural when X is an affine cone: in that case, the origin 0 is a distinguished point in $\mathbb{R}^n$, and the number of real critical points will be invariant under scaling u. Now we ask for the *expected number* of critical points of d_u when u is drawn from the probability distribution on $\mathbb{R}^n$ with density $|\omega|$. This *average ED degree* of the pair (X, ω) is

$$\text{aEDdegree}(X, \omega) \ := \ \int_{\mathbb{R}^n} \#\{\text{real critical points of } d_u \text{ on } X\} \cdot |\omega|. \tag{4.1}$$

In the formulas below, we write $\mathcal{E}_X$ for the set of real points of the ED correspondence. Using the substitution rule from multivariate calculus, we rewrite the integral in (4.1) as follows:

$$\text{aEDdegree}(X, \omega) \ = \ \int_{\mathbb{R}^n} \#\pi_2^{-1}(u) \cdot |\omega| \ = \ \int_{\mathcal{E}_X} |\pi_2^*(\omega)|, \tag{4.2}$$

where $\pi_2^*(\omega)$ is the pull-back of the volume form ω along the derivative of the map π_2.

See Figure 2 for a cartoon illustrating the computation in (4.2). Note that $\pi_2^*(\omega)$ need not be a volume form since it may vanish at some points—namely, at the *ramification locus* of π_2, i.e., at points where the derivative of π_2 is not of full rank. This ramification locus is typically an algebraic hypersurface in $\mathcal{E}_X$, and equal to the inverse image of the ED discriminant Σ_X.

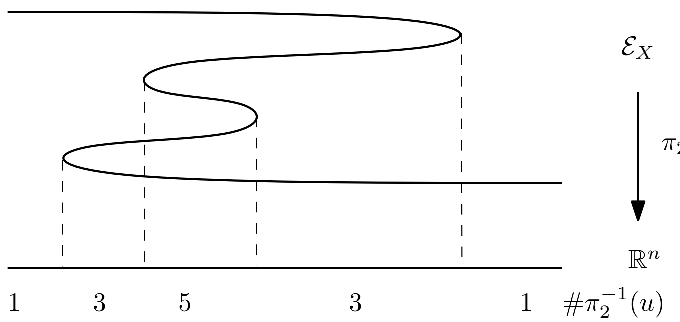

Figure 2: The map from the ED correspondence $\mathcal{E}_X$ to data space has four branch points. The weighted average of the fiber sizes $1, 3, 5, 3, 1$ can be expressed as an integral over $\mathcal{E}_X$.

The usefulness of the formula (4.2), and a more explicit version of it to be derived below, will depend on whether the strata in the complement of the branch locus of π_2 are easy to describe. We need such a description because the integrand will be a function "without absolute values in it" only on such open strata that lie over the complement of Σ_X.

Suppose that we have a parametrization $\phi : \mathbb{R}^n \to \mathcal{E}_X$ of the ED correspondence that is generically one-to-one. For instance, if X itself is given by a birational parametrization ψ, then ϕ can be derived from ψ using the method in the proof of Corollary 4.2. We can then write the integral over $\mathcal{E}_X$ in (4.2) more concretely as

$$\int_{\mathcal{E}_X} |\pi_2^*(\omega)| = \int_{\mathbb{R}^n} |\phi^*\pi_2^*(\omega)| = \int_{\mathbb{R}^n} |\det J_t(\pi_2 \circ \phi)| \cdot f(\pi_2(\phi(t))) \cdot \mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_n. \quad (4.3)$$

Here f is the smooth (density) function on $\mathbb{R}^n$ such that $\omega_u = f(u) \cdot \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_n$. In the standard Gaussian case, this would be $f(u) = e^{-||u||^2/2}/(2\pi)^{n/2}$. The determinant in (4.3) is taken of the differential of $\pi_2 \circ \phi$. To be fully explicit, the composition $\pi_2 \circ \phi$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $J_t(\pi_2 \circ \phi)$ denotes its $n \times n$ Jacobian matrix at a point t in the domain of ϕ.

Example 4.5 (ED and average ED degree of an ellipse)**.** For an illustrative simple example, let X denote the ellipse in $\mathbb{R}^2$ with equation $x^2 + 4y^2 = 4$. We first compute EDdegree(X). Let $(u, v) \in \mathbb{R}^2$ be a data point. The tangent line to the ellipse X at (x, y) has direction $(-4y, x)$. Hence the condition that $(x, y) \in X$ is critical for $d_{(u,v)}$ translates into the equation $(u-x, v-y)\cdot(-4y, x) = 0$, i.e., into $3xy+vx-4uy = 0$. For generic (u, v), the curve defined by the latter equation and the ellipse intersect in 4 points in $\mathbb{C}^2$, so EDdegree$(X) = 4$.

Now we consider aEDdegree(X, ω) where $\omega = \frac{1}{2\pi}e^{-(u^2+v^2)/2}\mathrm{d}u \wedge \mathrm{d}v$ is the standard Gaussians centered at the midpoint $(0, 0)$ of the ellipse. Given $(x, y) \in X$, the (u, v) for which (x, y) is critical are precisely those on the normal line. This is the line through (x, y) with direction $(x, 4y)$. In Figure 3 we plotted some of these normal lines. A colorful dynamic version of the same picture can be seen at `http://en.wikipedia.org/wiki/Evolute`. The *evolute* of the ellipse X is what we named the ED discriminant. It is the sextic *Lamé curve*

$$\Sigma_X = V(64u^6 + 48u^4v^2 + 12u^2v^4 + v^6 - 432u^4 + 756u^2v^2 - 27v^4 + 972u^2 + 243v^2 - 729).$$

Consider the rational parametrization of X given by $\psi(t) = \left(\frac{8t}{1+4t^2}, \frac{4t^2-1}{1+4t^2}\right)$, $t \in \mathbb{R}$. From ψ

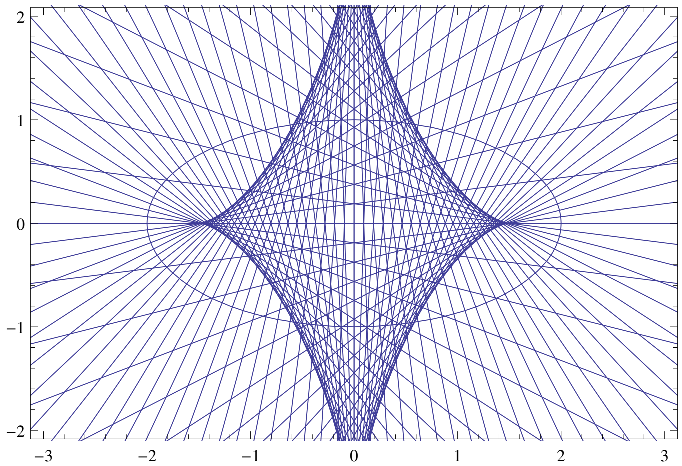

Figure 3: Computing the average ED of an ellipse: the evolute divides the plane into an inside region, where fibers or π_2 have cardinality 4, and an outside region, where fibers of π_2 have cardinality 2. The average ED of the ellipse is a weighted average of these numbers.

we construct a parametrization ϕ of the surface $\mathcal{E}_X$ as in Corollary 4.2, so that

$$\pi_2 \circ \phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2, (t, s) \mapsto \left((s+1)\frac{8t}{1+4t^2}, (4s+1)\frac{4t^2-1}{1+4t^2} \right).$$

The Jacobian determinant of $\pi_2 \circ \phi$ equals $\frac{-32(1+s+4(2s-1)t^2+16(1+s)t^4)}{(1+4t^2)^3}$, so aEDdegree$(X)$ is

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\left|\frac{-32(1+s+4(2s-1)t^2+16(1+s)t^4)}{(1+4t^2)^3}\right| e^{\frac{-(1+4s)^2-8(7-8(-1+s)s)t^2-16(1+4s)^2t^4}{2(1+4t^2)^2}}\mathrm{d}t\right)\mathrm{d}s.$$

Numerical integration (using `Mathematica 9`) finds the value 3.04658... in 0.202 seconds.

The following experiment independently validates this average ED degree calculation. We sample data points (u, v) randomly from Gaussian distribution. For each (u, v) we compute the number of real critical points, which is either 2 or 4, and we average these numbers. The average value approaches 3.05..., but it requires 10^5 samples to get two digits of accuracy. The total running time is 38.73 seconds, so much longer than the numerical integration. ♢

Example 4.6. The cardioid X from Example 1.1 can be parametrized by

$$\psi : \mathbb{R} \to \mathbb{R}^2,\ t \mapsto \left(\frac{2t^2-2}{(1+t^2)^2}, \frac{-4t}{(1+t^2)^2} \right).$$

From this we derive the following parametrization of the ED-correspondence $\mathcal{E}_X$:

$$\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2 \times \mathbb{R}^2,\ (t, s) \mapsto \left(\psi(t), \frac{2(t^4-1+4s(3t^2-1)}{(1+t^2)^3}, \frac{4t(-1-6s+(2s-1)t^2)}{(1+t^2)^3} \right).$$

Fixing the standard Gaussian centered at $(0, 0)$, the integral (4.3) now evaluates as follows:

$$\text{aEDdegree}(X, \omega) \;=\; \frac{1}{2\pi}\int_{\mathbb{R}^2} |\det J_{t,s}(\pi_2 \circ \phi)| e^{-\frac{||\pi_2\circ\phi(t,s)||^2}{2}} \mathrm{d}t\mathrm{d}s \;\approx\; 2.8375.$$

Thus, our measure gives little mass to the region inside the smaller cardioid in Figure 1. ♢

We next present a family of examples where the integral (4.3) can be computed exactly.

Example 4.7. We take X as the cone over the *rational normal curve*, in a special coordinate system, as in Example 2.6 and Corollary 8.7. Fix $\mathbb{R}^2$ with the standard orthonormal basis e_1, e_2. Let $S^n\mathbb{R}^2$ be the space of homogeneous polynomials of degree n in the variables e_1, e_2. We identify this space with $\mathbb{R}^{n+1}$ by fixing the basis $f_i := \sqrt{\binom{n}{i}} \cdot e_1^i e_2^{n-i}$ for $i = 0, \ldots, n$. This ensures that the natural action of the orthogonal group $\mathrm{O}_2(\mathbb{R})$ on polynomials in e_1, e_2 is by transformations that are orthogonal with respect to the standard inner product on $\mathbb{R}^{n+1}$.

Define $v, w : \mathbb{R}^2 \to \mathbb{R}^2$ by $v(t_1, t_2) := t_1 e_1 + t_2 e_2$ and $w(t_1, t_2) := t_2 e_1 - t_1 e_2$. These two vectors form an orthogonal basis of $\mathbb{R}^2$ for $(t_1, t_2) \neq (0, 0)$. Our surface X is parametrized by

$$\psi : \mathbb{R}^2 \;\to\; S^n\mathbb{R}^2 = \mathbb{R}^{n+1}, \quad (t_1, t_2) \;\mapsto\; v(t_1, t_2)^n = \sum_{i=0}^{n} t_1^i t_2^{n-i} \sqrt{\binom{n}{i}} f_i.$$

For $n = 3$, this parametrization specializes to the second parametrization in Example 2.6. Fix the standard Gaussian centered at the origin in $\mathbb{R}^{n+1}$. In what follows, we shall prove

$$\mathrm{aEDdegree}(X) \;=\; \sqrt{3n - 2}. \tag{4.4}$$

We begin by parametrizing the ED correspondence, as suggested in the proof of Corollary 4.2. For $(t_1, t_2) \neq (0, 0)$, the tangent space $T_{\psi(t_1,t_2)}X$ is spanned by $v(t_1, t_2)^n$ and $v(t_1, t_2)^{n-1} \cdot w(t_1, t_2)$. Since, by the choice of scaling, the vectors $v^n, v^{n-1}w, \ldots, w^n$ form an orthogonal basis of $\mathbb{R}^{n+1}$, we find that the orthogonal complement $(T_{\psi(t_1,t_2)}X)^\perp$ has the orthogonal basis

$$w(t_1, t_2)^n,\; v(t_1, t_2) \cdot w(t_1, t_2)^{n-1}, \ldots,\; v(t_1, t_2)^{n-2} \cdot w(t_1, t_2)^2.$$

The resulting parametrization $\phi : \mathbb{R}^2 \times \mathbb{R}^{n-1} \to \mathcal{E}_X$ of the ED correspondence equals

$$(t_1, t_2, s_0, ..., s_{n-2}) \;\mapsto\; \big(\psi(t_1, t_2),\; \psi(t_1, t_2)^n + s_0 v(t_1, t_2)^n + \cdots + s_{n-2} w(t_1, t_2)^n\big).$$

Next we determine the Jacobian $J = J(\pi_2 \circ \psi)$ at the point $\psi(t_1, t_2)$. It is most convenient to do so relative to the orthogonal basis $v(t_1, t_2),\ w(t_1, t_2),\ (1, 0, \ldots, 0),\ \ldots, (0, \ldots, 0, 1)$ of $\mathbb{R}^2 \times \mathbb{R}^{n-1}$ and the orthogonal basis $w(t_1, t_2)^n, \ldots, v(t_1, t_2)^n$ of $\mathbb{R}^{n+1}$. Relative to these bases,

$$J \quad = \quad \begin{bmatrix} * & * & 1 & 0 & \cdots & 0 \\ * & * & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & 0 & 0 & \cdots & 1 \\ 0 & n - 2s_{n-2} & 0 & 0 & \cdots & 0 \\ n & * & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where the stars are irrelevant for $\det(J)$. For instance, an infinitesimal change $v(t_1, t_2) \mapsto v(t_1, t_2) + \epsilon w(t_1, t_2)$ leads to a change $w(t_1, t_2) \mapsto w(t_1, t_2) - \epsilon v(t_1, t_2)$ and to a change of $\pi_2 \circ \psi$ in which the coefficient of $\epsilon v(t_1, t_2)^{n-1} \cdot w(t_1, t_2)$ equals $n - 2s_{n-2}$. When computing the determinant of J, we must consider that the chosen bases are orthogonal but not orthonormal: the norm of $v(t_1, t_2)^i \cdot w(t_1, t_2)^{n-i}$, corresponding to the i-th row, equals $\sqrt{(t_1^2 + t_2^2)^n} \binom{n}{i}^{-1/2}$;

and the norm of $v(t_1,t_2)$ and $w(t_1,t_2)$, corresponding to the first and second column, equals $\sqrt{t_1^2+t_2^2}$. Multiplying the determinant of the matrix above with the product of these scalars, and dividing by the square of $\sqrt{t_1^2+t_2^2}$ for the first two columns, we obtain the formula

$$|\det J(\pi_2\circ\psi)| \;=\; n\cdot|n-2s_{n-2}|\cdot(t_1^2+t_2^2)^{n(n+1)/2-1}\cdot\prod_{i=0}^{n}\binom{n}{i}^{-1/2}.$$

Next, the squared norm of $u=\pi_2\circ\psi(t_1,t_2,s_0,...,s_{n-2})$ equals

$$||u||^2=(t_1^2+t_2^2)^n\cdot\left(1+\sum_{i=0}^{n-2}s_i^2\binom{n}{i}^{-1}\right).$$

The average ED degree of X relative to the standard Gaussian equals

$$\text{aEDdegree}(X)=\frac{1}{(2\pi)^{(n+1)/2}}\int|\det J(\pi_2\circ\psi)|e^{-||u||^2/2}\mathrm{d}v_1\mathrm{d}v_2\mathrm{d}s_0\cdots\mathrm{d}s_{n-2}.$$

parametrizing the regions where $\det J(\pi_2\circ\psi)$ is positive or negative by $s_{n-2}\in(-\infty,n/2)$ or $s_{n-2}\in(n/2,\infty)$, this integral can be computed in closed form. Its value equals $\sqrt{3n-2}$. Interestingly, this value is the square root of the generic ED degree in Example 5.12. ♢

We close this section with the remark that different applications require different choices of the measure $|\omega|$ on data space. For instance, one might want to draw u from a product of intervals equipped with the uniform distribution, or to concentrate the measure near X.

5 Duality

This section deals exclusively with irreducible affine cones $X\subset\mathbb{C}^n$, or, equivalently, with their corresponding projective varieties $X\subset\mathbb{P}^{n-1}$. Such a variety has a *dual variety* $Y:=X^*\subset\mathbb{C}^n$, which is defined as follows, where the line indicates the topological closure:

$$Y \;:=\; \overline{\left\{y\in\mathbb{C}^n\mid\exists x\in X\backslash X_{\text{sing}}:y\perp T_xX\right\}}.$$

See [31, Section 5.2.4] for an introduction to this duality in the context of optimization. Algorithm 5.1 in [31] explains how to compute the ideal of Y from that of X.

The variety Y is an irreducible affine cone, so we can regard it as an irreducible projective variety in $\mathbb{P}^{n-1}$. That projective variety parametrizes hyperplanes tangent to X at non-singular points, if one uses the standard bilinear form on $\mathbb{C}^n$ to identify hyperplanes with points in $\mathbb{P}^{n-1}$. We will prove EDdegree(X) = EDdegree(Y). Moreover, for generic data $u\in\mathbb{C}^n$, there is a natural bijection between the critical points of d_u on the cone X and the critical points of d_u on the cone Y. We then link this result to the literature on the *conormal variety* (cf. [22]) which gives powerful techniques for computing ED degrees of smooth varieties that intersect the isotropic quadric $Q=V(x_1^2+\cdots+x_n^2)$ transversally. Before dealing with the general case, we revisit the example of the Eckart-Young Theorem.

Example 5.1. For the variety X_r of $s \times t$ matrices ($s \le t$) of rank $\le r$, we have $X_r^* = X_{s-r}$ [17, Chap. 1, Prop. 4.11]. From Example 2.3 we see that $\mathrm{EDdegree}(X_r) = \mathrm{EDdegree}(X_{s-r})$. There is a bijection between the critical points of d_U on X_r and on X_{s-r}. To see this, consider the singular value decomposition (2.3). For a subset $I = \{i_1, \ldots, i_r\}$ of $\{1, \ldots, s\}$, we set

$$U_I \;=\; T_1 \cdot \mathrm{diag}(\ldots, \sigma_{i_1}, \ldots, \sigma_{i_2}, \ldots, \sigma_{i_r}, \ldots) \cdot T_2,$$

where the places of σ_j for $j \notin I$ have been filled with zeros in the diagonal matrix. Writing I^c for the complementary subset of size $s - r$, we have $U = U_I + U_{I^c}$. This decomposition is orthogonal in the sense that $\langle U_I, U_{I^c} \rangle = \mathrm{tr}(U_I^t U_{I^c}) = 0$. It follows that, if U is real, then $|U|^2 = |U_I|^2 + |U_{I^c}|^2$, where $|U|^2 = tr(U^t U)$. As I ranges over all r-subsets, U_I runs through the critical points of d_U on the variety X_r, and U_{I^c} runs through the critical points of d_U on the dual variety X_{s-r}. Since the formula above reads as $|U|^2 = |U - U_{I^c}|^2 + |U - U_I|^2$, we conclude that the proximity of the real critical points reverses under this bijection. For instance, if U_I is the real point on X_r closest to U, then U_{I^c} is the real point on X_{s-r} farthest from U. For a similar result in the multiplicative context of maximum likelihood see [10]. $\diamondsuit$

The following theorem shows that the duality seen in Example 5.1 holds in general.

Theorem 5.2. *Let $X \subset \mathbb{C}^n$ be an irreducible affine cone, $Y \subset \mathbb{C}^n$ its dual variety, and $u \in \mathbb{C}^n$ a generic data point. The map $x \mapsto u - x$ gives a bijection from the critical points of d_u on X to the critical points of d_u on Y. Consequently,* $\mathrm{EDdegree}(X) = \mathrm{EDdegree}(Y)$. *Moreover, if u is real, then the map sends real critical points to real critical points, and hence* $\mathrm{aEDdegree}(X, \omega) = \mathrm{aEDdegree}(Y, \omega)$ *for any volume form ω. The map is proximity-reversing: the closer a real critical point x is to the data point u, the further $u - x$ is from u.*

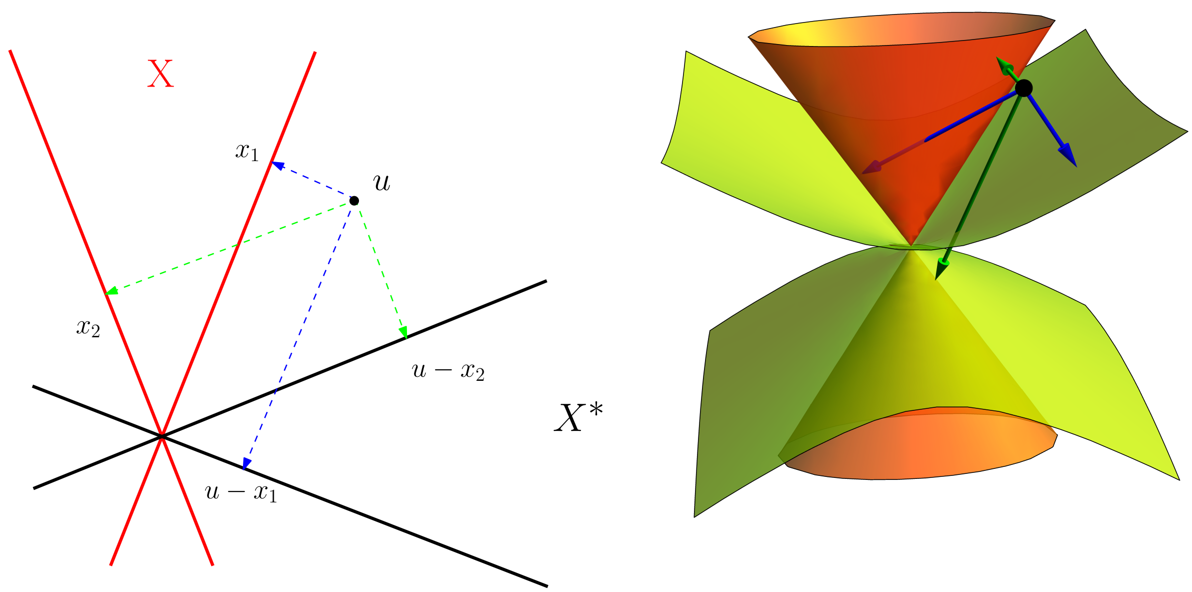

Figure 4: The bijection between critical points on X and critical points on X^*.

The statement of Theorem 5.2 is illustrated in Figure 4. On the left, the variety X is a 1-dimensional affine cone in $\mathbb{R}^2$. This X is not irreducible but it visualizes our duality in the simplest possible case. The right picture shows the same scenario in one dimension higher. Here X and X^* are quadratic cones in $\mathbb{R}^3$, corresponding to a dual pair of conics in $\mathbb{P}^2$.

The proof of Theorem 5.2 uses properties of the *conormal variety*, which is defined as

$$\mathcal{N}_X \;:=\; \overline{\{(x,y) \in \mathbb{C}^n \times \mathbb{C}^n \mid x \in X \backslash X_{\text{sing}} \text{ and } y \perp T_x X\}}.$$

The conormal variety is the zero set of the following ideal in $\mathbb{R}[x,y]$:

$$N_X \;:=\; \left(I_X + \left\langle (c+1)\times(c+1)\text{-minors of } \begin{pmatrix} y \\ J(f) \end{pmatrix} \right\rangle \right) : \big(I_{X_{\text{sing}}} \big)^{\infty}, \tag{5.1}$$

where $f = (f_1, \ldots, f_s)$ is a system of homogeneous generators of I_X. It is known that $\mathcal{N}_X$ is irreducible of dimension $n-1$. The projection of $\mathcal{N}_X$ into the second factor $\mathbb{C}^n$ is the dual variety $Y = X^*$. Its ideal I_Y is computed by elimination, namely, by intersecting (5.1) with $\mathbb{R}[y]$. An important property of the conormal variety is the *Biduality Theorem* [17, Chapter 1], which states that $\mathcal{N}_X$ equals $\mathcal{N}_Y$ up to swapping the two factors. In symbols, we have

$$\mathcal{N}_X \;=\; \mathcal{N}_Y \;=\; \overline{\{(x,y) \in \mathbb{C}^n \times \mathbb{C}^n \mid \; y \in Y \backslash Y_{\text{sing}} \text{ and } x \perp T_y Y\}}.$$

This implies $(X^*)^* = Y^* = X$. Thus the biduality relation in [31, Theorem 5.13] holds. To keep the symmetry in our notation, we will henceforth write $\mathcal{N}_{X,Y}$ for $\mathcal{N}_X$ and $N_{X,Y}$ for N_X.

Proof of Theorem 5.2. The following is illustrated in Figure 4. If x is a critical point of d_u on X, then $y := u - x$ is orthogonal to $T_x X$, and hence $(x,y) \in \mathcal{N}_{X,Y}$. By the genericity of u, all y thus obtained from critical points x of d_u are non-singular points on Y. By the Biduality Theorem, we have $u - y = x \perp T_y Y$, i.e., y is a critical point of d_u on Y. This shows that $x \mapsto u - x$ maps critical points of d_u on X into critical points of d_u on Y. Applying the same argument to Y, and using that $Y^* = X$, we find that, conversely, $y \mapsto u - y$ maps critical points of d_u on Y to critical points of d_u on X. This establishes the bijection.

The consequences for EDdegree(X) and aEDdegree(X, ω) are straightforward. For the last statement we observe that $u - x \perp x \in T_x X$ for critical x. For $y = u - x$, this implies

$$||u-x||^2 + ||u-y||^2 \;=\; ||u-x||^2 + ||x||^2 \;=\; ||u||^2.$$

Hence the assignments that take real data points u to X and X^* are proximity-reversing. □

Duality leads us to define the *joint ED correspondence* of the cone X and its dual Y as

$$\begin{aligned} \mathcal{E}_{X,Y} \;:=&\; \overline{\{(x, u-x, u) \in \mathbb{C}^n_x \times \mathbb{C}^n_y \times \mathbb{C}^n_u \mid x \in X \backslash X_{\text{sing}} \text{ and } u - x \perp T_x X\}} \\ =&\; \overline{\{(u-y, y, u) \in \mathbb{C}^n_x \times \mathbb{C}^n_y \times \mathbb{C}^n_u \mid \; y \in Y \backslash Y_{\text{sing}} \text{ and } u - y \perp T_y Y\}}. \end{aligned}$$

The projection of $\mathcal{E}_{X,Y}$ into $\mathbb{C}^n_x \times \mathbb{C}^n_u$ is the ED correspondence $\mathcal{E}_X$ of X, its projection into $\mathbb{C}^n_y \times \mathbb{C}^n_u$ is $\mathcal{E}_Y$, and its projection into $\mathbb{C}^n_x \times \mathbb{C}^n_y$ is the conormal variety $\mathcal{N}_{X,Y}$. The affine variety $\mathcal{E}_{X,Y}$ is irreducible of dimension n, since $\mathcal{E}_X$ has these properties (by Theorem 4.1), and the projection $\mathcal{E}_{X,Y} \to \mathcal{E}_X$ is birational with inverse $(x,u) \mapsto (x, u-x, u)$.

Following Theorem 4.4, we also introduce the *projective joint ED correspondence* $\mathcal{PE}_{X,Y}$. By definition, $\mathcal{PE}_{X,Y}$ is the closure of the image of $\mathcal{E}_{X,Y} \cap \big((\mathbb{C}^n \backslash \{0\})^2 \times \mathbb{C}^n\big)$ in $\mathbb{P}^{n-1}_x \times \mathbb{P}^{n-1}_y \times \mathbb{C}^n_u$.

Proposition 5.3. *Let $X \subset \mathbb{C}^n$ be an irreducible affine cone, let $Y \subset \mathbb{C}^n$ be the dual variety of X, and assume that neither X nor Y is contained in $Q = V(q)$, where $q = x_1^2 + \cdots + x_n^2$. Then $\mathcal{PE}_{X,Y}$ is an irreducible n-dimensional variety in $\mathbb{P}_x^{n-1} \times \mathbb{P}_y^{n-1} \times \mathbb{C}_u^n$, defined by the ideal*

$$\left(N_{X,Y} + \left\langle 3 \times 3\text{-minors of the } 3 \times n\text{-matrix} \begin{pmatrix} u \\ x \\ y \end{pmatrix} \right\rangle\right) : \big\langle q(x) \cdot q(y) \big\rangle^{\infty} \subset \mathbb{R}[x,y,u]. \quad (5.2)$$

Proof. The irreducibility of $\mathcal{PE}_{X,Y}$ follows from that of $\mathcal{E}_{X,Y}$ which has the same dimension.

To see that $\mathcal{PE}_{X,Y}$ is defined by the ideal (5.2), note first that any point (x, y, u) with $x \in X \backslash X_{\text{sing}}$ and $y \perp T_x X$ and $x + y = u$ has $(x, y) \in \mathcal{N}_{X,Y}$ and $\dim\langle x, y, u\rangle \leq 2$, so that $([x], [y], u)$ is a zero of (5.2). This shows that $\mathcal{PE}_{X,Y}$ is contained in the variety of (5.2).

Conversely, let $([x], [y], u)$ be in the variety of (5.2). The points with $q(x)q(y) \neq 0$ are dense in the variety of (5.2), so we may assume $x, y \notin Q$. Moreover, since $(x, y) \in \mathcal{N}_{X,Y}$, we may assume that x, y are non-singular points of X and Y, and that $x \perp T_y Y$ and $y \perp T_x X$. This implies $x \perp y$. Since x, y are not isotropic, they are linearly independent. Then $u = cx + dy$ for unique constants $c, d \in \mathbb{C}$. If $c, d \neq 0$, then we find that $(cx, dy, u) \in \mathcal{E}_{X,Y} \cap ((\mathbb{C}^n \backslash \{0\})^2 \times \mathbb{C}_u^n)$ and hence $([x], [y], u) \in \mathcal{PE}_{X,Y}$. If $c \neq 0$ but $d = 0$, then $(cx, \epsilon y, u + \epsilon y) \in \mathcal{E}_{X,Y}$ for all $\epsilon \neq 0$, so that the limit of $([cx], [\epsilon y], u + \epsilon y)$ for $\epsilon \to 0$, which is $([x], [y], u)$, lies in $\mathcal{PE}_{X,Y}$. Similar arguments apply when $d \neq 0$ but $c = 0$ or when $c = d = 0$. □

Our next result gives a formula for EDdegree(X) in terms of the *polar classes* of classical algebraic geometry [30]. These non-negative integers $\delta_i(X)$ are the coefficients of the class

$$[\mathcal{N}_{X,Y}] \quad = \quad \delta_0(X)s^{n-1}t + \delta_1(X)s^{n-2}t^2 + \cdots + \delta_{n-2}(X)st^{n-1} \quad (5.3)$$

of the conormal variety, when regarded as a subvariety of $\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$. For topologists, the polynomial (5.3) is the class representing $\mathcal{N}_{X,Y}$ in the cohomology ring $H^*(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}) = \mathbb{Z}[s,t]/\langle s^n, t^n\rangle$. For commutative algebraists, it is the *multidegree* of the $\mathbb{Z}^2$-graded ring $\mathbb{R}[x,y]/N_{X,Y}$. This is explained in [27, Section 8.5], and is implemented in `Macaulay2` with the command `multidegree`. For geometers, the polar classes $\delta_i(X)$ have the following definition: intersecting the $(n-2)$-dimensional subvariety $\mathcal{N}_{X,Y} \subset \mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$ with an n-dimensional subvariety $L \times M$ where L, M are general linear subspaces of $\mathbb{P}^{n-1}$ of dimensions $n - j$ and j, respectively, one gets a finite number of simple points. The number $\delta_{j-1}(X)$ counts these points. The shift by one is to ensure compatibility with Holme's paper [22].

So, for example, $\delta_0(X)$ counts the number of intersections of $\mathcal{N}_{X,Y}$ with $\mathbb{P}^{n-1} \times M$ where M is a general projective line. These are the intersections of the dual variety Y with M. Thus, if Y is a hypersurface, then $\delta_0(X)$ is the degree of Y, and otherwise $\delta_0(X)$ is zero. In general, the first non-zero coefficient of (5.3) is the degree of Y and the last non-zero coefficient is the degree of X. For all i, we have $\delta_i(Y) = \delta_{n-2-i}(X)$; see [22, Theorem 2.3].

Theorem 5.4. *If $\mathcal{N}_{X,Y}$ does not intersect the diagonal $\Delta(\mathbb{P}^{n-1}) \subset \mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$, then*

$$\text{EDdegree}(X) \;=\; \delta_0(X) + \cdots + \delta_{n-2}(X) \;=\; \delta_{n-2}(Y) + \cdots + \delta_0(Y) \;=\; \text{EDdegree}(Y).$$

A sufficient condition for $\mathcal{N}_{X,Y}$ not to intersect $\Delta(\mathbb{P}^{n-1})$ is that $X \cap Q$ is a transversal intersection everywhere (i.e. $X \cap Q$ is smooth) and disjoint from X_{sing}. Indeed, suppose that $(x,x) \in \mathcal{N}_{X,Y}$ for some $x \in X$. There exists a sequence of points $(x_i, y_i) \in \mathcal{N}_{X,Y}$ with $x_i \in X \backslash X_{\text{sing}}$, $y_i \perp T_{x_i}X$, such that $\lim_{i\to\infty}(x_i,y_i) \to (x,x)$. Then $y_i \perp x_i$, so taking the limit we find $x \in Q$. If, moreover, X is smooth at x, then $T_{x_i}X$ converges to the tangent space T_xX. We conclude that $x \perp T_xX$, which means that X is tangent to Q at x.

Proof of Theorem 5.4. Denote by Z the variety of linearly dependent triples $(x,y,u) \in \mathbb{P}^{n-1}_x \times \mathbb{P}^{n-1}_y \times \mathbb{C}^n_u$. By Proposition 5.3, the intersection $(\mathcal{N}_{X,Y} \times \mathbb{C}^n) \cap Z$ contains the projective ED correspondence $\mathcal{PE}_{X,Y}$ as a component. The two are equal because $(\mathcal{N}_{X,Y} \times \mathbb{C}^n) \cap Z$ is swept out by the 2-dimensional vector spaces $\{(x,y)\} \times \langle x,y \rangle$, as (x,y) runs through the irreducible variety $\mathcal{N}_{X,Y}$, and hence it is irreducible. Here we are using that $\mathcal{N}_{X,Y} \cap \Delta(\mathbb{P}^{n-1}) = \emptyset$.

Hence EDdegree(X) is the length of a general fiber of the map $\pi_3 : (\mathcal{N}_{X,Y} \times \mathbb{C}^n) \cap Z \to \mathbb{C}^n$. Next, a tangent space computation shows that the intersection $(\mathcal{N}_{X,Y} \times \mathbb{C}^n) \cap Z$ is transversal, so an open dense subset of it is a smooth scheme. By generic smoothness [20, Corollary III.10.7], the fiber $\pi_3^{-1}(u)$ over a generic data point u consists of simple points only. This fiber is scheme-theoretically the same as $\mathcal{N}_{X,Y} \cap Z_u$, where Z_u is the fiber in Z over u. The cardinality of this intersection is the coefficient of $s^{n-1}t^{n-1}$ in the product $[\mathcal{N}_{X,Y}] \cdot [Z_u]$ in $H^*(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1}) = \mathbb{Z}[s,t]/\langle s^n, t^n \rangle$. The determinantal variety Z_u has codimension $n-2$, and

$$[Z_u] \;=\; s^{n-2} + s^{n-3}t + s^{n-4}t^2 + \cdots + st^{n-3} + t^{n-2}.$$

This is a very special case of [27, Corollary 16.27]. By computing modulo $\langle s^n, t^n \rangle$, we find

$$[\mathcal{N}_{X,Y}] \cdot [Z_u] \;=\; (\delta_0(X)s^{n-1}t + \cdots + \delta_{n-2}(X)st^{n-1}) \cdot [Z] \;=\; (\delta_0(X) + \cdots + \delta_{n-2}(X))s^{n-1}t^{n-1}.$$

This establishes the desired identity. □

Remark 5.5. If X and Y are smooth then $X \cap Q$ is smooth if and only if $\Delta(\mathbb{P}^{n-1}) \cap \mathcal{N}_{X,Y} = \emptyset$ if and only if $Y \cap Q$ is smooth. We do not know whether this holds when X or Y is singular.

Example 5.6. Let X be the variety of symmetric $s \times s$-matrices x of rank $\leq r$ and Y the variety of symmetric $s \times s$-matrices y of rank $\leq s-r$. These two determinantal varieties form a dual pair [31, Example 5.15]. Their conormal ideal $N_{X,Y}$ is generated by the relevant minors of x and y and the entries of the matrix product xy. The class $[N_{X,Y}]$ records the *algebraic degree of semidefinite programming.* A formula was found by von Bothmer and Ranestad in [4]. Using the package `Schubert2` in `Macaulay2` [18], and summing over the index m in [4, Proposition 4.1], we obtain the following table of values for EDdegree(X):

$s =$	2	3	4	5	6	7
$r=1$	4	13	40	121	364	1093
$r=2$		13	122	1042	8683	72271
$r=3$			40	1042	23544	510835
$r=4$				121	8683	510835
$r=5$					364	72271
$r=6$						1093

In order for X to satisfy the hypothesis in Theorem 5.4, it is essential that the coordinates are generic enough, so that $X \cap Q$ is smooth. Luckily, the usual coordinates in $\mathbb{C}^{\binom{s+1}{2}}$ enjoy this property, and the table above records the ED degree for the second interpretation in Example 3.2. Specifically, our number 13 for $s=3$ and $r=2$ appeared on the right in (3.2). The symmetry in the columns of our table reflects the duality result in Theorem 5.2. ♢

Example 5.7. Following [31, Ex. 5.44], *Cayley's cubic surface* $X = V(f) \subset \mathbb{P}^3_x$ is given by

$$f(x) \;=\; \det \begin{pmatrix} x_0 & x_1 & x_2 \\ x_1 & x_0 & x_3 \\ x_2 & x_3 & x_0 \end{pmatrix}.$$

Its dual in $\mathbb{P}^3_y$ is the *quartic Steiner surface* $Y = V(g)$, with $g \;=\; y_1^2y_2^2+y_1^2y_3^2+y_2^2y_3^2-2y_0y_1y_2y_3$. The conormal ideal $N_{X,Y}$ is minimally generated by 18 bihomogeneous polynomials in $\mathbb{R}[x,y]$:

f of degree $(3,0)$; g of degree $(0,4)$; $q(x,y) = x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3$ of degree $(1,1)$;
six generators of degree $(1,2)$, such as $x_2y_1y_2 + x_3y_1y_3 + x_0y_2y_3$; and
nine generators of degree $(2,1)$, such as $x_0x_1y_2 - x_2x_3y_2 + x_0^2y_3 - x_3^2y_3$.

The conormal variety $\mathcal{N}_{X,Y}$ is a surface in $\mathbb{P}^3_x \times \mathbb{P}^3_y$ with class $4s^3t + 6s^2t^2 + 3st^3$, and hence

$$\text{EDdegree}(X) \;=\; \text{EDdegree}(Y) \;=\; 4+6+3 \;=\; 13.$$

Corollary 6.4 relates this to the number 13 in (3.2). The projective joint ED correspondence $\mathcal{PE}_{X,Y}$ is defined by the above equations together with the four 3×3-minors of the matrix

$$\begin{pmatrix} u \\ x \\ y \end{pmatrix} \;=\; \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \\ x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \end{pmatrix}.$$

For fixed scalars $u_0, u_1, u_2, u_3 \in \mathbb{R}$, this imposes a codimension 2 condition. This cuts out 13 points in $\mathcal{N}_{X,Y} \subset X \times Y \subset \mathbb{P}^3_x \times \mathbb{P}^3_y$. These represent the critical points of d_u on X or Y. $\diamondsuit$

Armed with Theorem 5.4, we can now use the beautiful results in Holme's article [22] to express the ED degree of a smooth projective variety X in terms of its Chern classes.

Theorem 5.8. *Let X be a smooth irreducible subvariety of dimension m in $\mathbb{P}^{n-1}$, and suppose that X is transversal to the isotropic quadric Q. Then*

$$\text{EDdegree}(X) \;=\; \sum_{i=0}^{m} (-1)^i \cdot (2^{m+1-i} - 1) \cdot \deg(c_i(X)). \tag{5.4}$$

Here $c_i(X)$ is the ith Chern class of the cotangent bundle of X. For more information on Chern classes, and alternative formulations of Theorem 5.8, we refer the reader to Section 7.

Proof. By Theorem 5.4 we have $\text{EDdegree}(X) = \sum_{i=0}^{n-2} \delta_i(X)$. We also saw that $\delta_i(X) = 0$ for $i > m$, so we may let i run from 0 to m instead. Substituting the expression

$$\delta_i(X) \;=\; \sum_{j=i}^{m} (-1)^{m-j} \binom{j+1}{i+1} \deg(c_{m-j}(X))$$

from [22, Page 150], and summing over all values of the index i, yields the theorem. $\square$

Corollary 5.9. *Let X be a smooth irreducible curve of degree d and genus g in $\mathbb{P}^{n-1}$, and suppose that X is transversal to Q. Then*

$$\text{EDdegree}(X) \;=\; 3d + 2g - 2. \tag{5.5}$$

Proof. We have from [20, App. A §3] that $\deg(c_0(X)) = d$ and $\deg(c_1(X)) = 2 - 2g$. □

Example 5.10. Consider a 2×3 matrix with entries in $\mathbb{R}[x_1, x_2, x_3, x_4]$ where the first row contains general linear forms, and the second row contains general quadratic forms. The ideal I generated by its three maximal minors defines a smooth irreducible curve in $\mathbb{P}^3$ of degree 7 and genus 5, so Corollary 5.9 gives $\text{EDdegree}(V(I)) = 3 \cdot 7 + 2 \cdot 5 - 2 = 29$. This exceeds the bound of 27 we would get by taking $n = 4, c = 3, d_1 = d_2 = d_3 = 3$ in (2.8). However, while ideal I has $s = 3$ generators, the codimension of its variety $V(I)$ is $c = 2$. Applying Corollary 2.9 to $c = 2, d_1 = d_2 = 3$, we get the correct bound of 45. This is the ED degree for the complete intersection of two cubics in $\mathbb{P}^3$, and it exceeds 29 as desired. ♢

The formula (5.4) is particularly nice for smooth projective *toric* varieties X in $\mathbb{P}^{n-1}$. According to [15], such a toric manifold corresponds to a simple lattice polytope $P \subset \mathbb{R}^m$ with $|P \cap \mathbb{Z}^m| = n$, and $c_{m-j}(X)$ is the sum of classes corresponding to all j-dimensional faces of P. The degree of this class is its *normalized volume*. Therefore, Theorem 5.8 implies

Corollary 5.11. *Let $X \subset \mathbb{P}^{n-1}$ be an m-dimensional smooth projective toric variety, with coordinates such that X is transversal to Q. If V_j denotes the sum of the normalized volumes of all j-dimensional faces of the simple lattice polytope P associated with X, then*

$$\text{EDdegree}(X) \;=\; \sum_{j=0}^{m} (-1)^{m-j} \cdot (2^{j+1} - 1) \cdot V_j.$$

Example 5.12. Consider a rational normal curve X in $\mathbb{P}^n$ in generic coordinates (we denote the ambient space as $\mathbb{P}^n$ instead of $\mathbb{P}^{n-1}$, to compare with Example 4.7). The associated polytope P is a segment of integer length n. The formula above yields

$$\text{EDdegree}(X) \;=\; (2^2 - 1) \cdot V_1 - (2^1 - 1) \cdot V_0 \;=\; 3n - 2.$$

In special coordinates, the ED degree can drop to n; see Corollary 8.7. Interestingly, in those special coordinates, the square root of $3n - 2$ is the average ED degree, by Example 4.7.

All Segre varieties and Veronese varieties are smooth toric varieties, so we can compute their ED degrees (in generic coordinates) using Corollary 5.11. For Veronese varieties, this can be used to verify the $r = 1$ row in the table of Example 5.6. For instance, for $s = 3$, the toric variety X is the Veronese surface in $\mathbb{P}^5$, and the polytope is a regular triangle with sides of lattice length 2. Here, $\text{EDdegree}(X) = 7 \cdot V_2 - 3 \cdot V_1 + V_0 \;=\; 7 \cdot 4 - 3 \cdot 6 + 3 \;=\; 13$. ♢

6 Geometric Operations

Following up on our discussion of duality, this section studies the behavior of the ED degree of a variety under other natural operations. We begin with the dual operations of projecting from a point and intersecting with a hyperplane. Thereafter we discuss homogenizing and dehomogenizing. Geometrically, these correspond to passing from an affine variety to its projective closure and vice versa. We saw in the examples of Section 3 that the ED degree can go up or go down under homogenization. We aim to explain that phenomenon.

Our next two results are corollaries to Theorem 5.4 and results of Piene in [30]. We work in the setting of Section 5, so X is an irreducible projective variety in $\mathbb{P}^{n-1}$ and X^* is its dual, embedded into the same $\mathbb{P}^{n-1}$ by way of the quadratic form $q(x,y) = x_1y_1 + \cdots + x_ny_n$. The polar classes satisfy $\delta_i(X) = \delta_{n-2-i}(X^*)$. These integers are zero for $i \geq \dim(X)$ and $i \leq \mathrm{codim}(X^*) - 2$, and they are strictly positive for all other values of the index i. The first positive $\delta_i(X)$ is the degree of X^*, and the last positive $\delta_i(X)$ is the degree of X. The sum of all $\delta_i(X)$ is the common ED degree of X and X^*. See [22] and our discussion above.

Fix a generic linear map $\pi : \mathbb{C}^n \to \mathbb{C}^{n-1}$. This induces a rational map $\pi : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-2}$, whose base point lies outside X. The image $\pi(X)$ is an irreducible closed subvariety in $\mathbb{P}^{n-2}$. Since the projective space $\mathbb{P}^{n-2}$ comes with a coordinate system $(x_1 : x_2 : \cdots : x_{n-1})$, the ED degree of $\pi(X)$ is well-defined. If $\mathrm{codim}(X) = 1$ then $\pi(X) = \mathbb{P}^{n-2}$ has ED degree 1 for trivial reasons. Otherwise, X maps birationally onto $\pi(X)$, and the ED degree is preserved:

Corollary 6.1. *Let X satisfy the assumptions of Theorem 5.4. If* $\mathrm{codim}(X) \geq 2$ *then*

$$\mathrm{EDdegree}(\pi(X)) \;=\; \mathrm{EDdegree}(X). \tag{6.1}$$

Proof. Piene [30] showed that $\delta_i(\pi(X)) = \delta_i(X)$ for all i. Now use Theorem 5.4. □

Example 6.2. Let I be the prime ideal generated by the 2×2-minors of the symmetric 3×3-matrix whose six entries are generic linear forms in $\mathbb{R}[x_1, x_2, x_3, x_4, x_5, x_6]$. The *elimination ideal* $J = I \cap \mathbb{R}[x_1, x_2, x_3, x_4, x_5]$ is minimally generated by seven cubics. Its variety $\pi(X) = V(J)$ is a random projection of the Veronese surface $X = V(I)$ from $\mathbb{P}^5$ into $\mathbb{P}^4$. Example 5.6 tells us that $\mathrm{EDdegree}(X) = 13$. By plugging $J = I_{\pi(X)}$ into the formula (2.7), and running `Macaulay2` as in Example 2.10, we verify $\mathrm{EDdegree}(\pi(X)) = 13$. ♢

If X is a variety of high codimension, then Corollary 6.1 can be applied repeatedly until the image $\pi(X)$ is a hypersurface. In other words, we can take π to be a generic linear projection $\mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^d$ provided $d > \dim(X)$. Then $\pi(X)$ also satisfies the assumptions of Theorem 5.4, and the formula (6.1) remains valid. This technique is particularly useful when X is a smooth toric variety as in Corollary 5.11. Here, X is parametrized by certain monomials, and $\pi(X)$ is parametrized by generic linear combinations of those monomials.

Example 6.3. Consider a surface in $\mathbb{P}^3$ that is parametrized by four homogeneous polynomials of degree d in three variables. That surface can be represented as $\pi(X)$ where X is the d-fold Veronese embedding of $\mathbb{P}^2$ into $\mathbb{P}^{\binom{d+2}{2}-1}$, and π is a random projection into $\mathbb{P}^3$. By applying Corollary 5.11 to the associated lattice triangle $P = \mathrm{conv}\{(0,0),(0,d),(d,0)\}$, and using Corollary 6.1, we find $\mathrm{EDdegree}(\pi(X)) = \mathrm{EDdegree}(X) = 7d^2 - 9d + 3$. This is to be compared to the number $4d^2 - 4d + 1$, which is the ED degree in (2.5) for the *affine* surface in $\mathbb{C}^3$ parametrized by three inhomogeneous polynomials of degree d in two variables.

A similar distinction arises for Bézier surfaces in 3-space. The ED degree of the affine surface in Example 3.1 is $8d_1d_2 - 2d_1 - 2d_2 + 1$, while $\mathrm{EDdegree}(\pi(X)) = 14d_1d_2 - 6d_1 - 6d_1 + 4$ for the projective surface $\pi(X)$ that is given by four bihomogeneous polynomials ψ_i of degree

(d_1, d_2) in 2+2 parameters. Here, the toric surface is $X = \mathbb{P}^1 \times \mathbb{P}^1$, embedded in $\mathbb{P}^{(d_1+1)(d_2+1)-1}$ by the line bundle $\mathcal{O}(d_1, d_2)$, and the lattice polygon is the square $P = [0, d_1] \times [0, d_2]$. ♢

In the previous example we computed the ED degree of a variety by expressing it as a linear projection from a high-dimensional space with desirable combinatorial properties. This is reminiscent of the technique of *lifting* in optimization theory, where one simplifies a problem instance by optimizing over a higher-dimensional constraint set that projects onto the given constraint set. It would be desirable to develop this connection further, and to find a more direct proof of Corollary 6.1 that works for both projective and affine varieties.

The operation dual to projection is taking linear sections. Let H be a generic hyperplane in $\mathbb{P}^{n-1}$. Then $X \cap H$ is a subvariety of codimension 1 in X. In particular, it lives in the same ambient space $\mathbb{P}^{n-1}$, with the same coordinates $(x_1 : \cdots : x_n)$, and this defines the ED degree of $X \cap H$. By Bertini's Theorem, the variety $X \cap H$ is irreducible provided $\dim(X) \geq 2$.

Corollary 6.4. *Let $X \subset \mathbb{P}^{n-1}$ satisfy the assumptions of Theorem 5.4. Then*

$$\text{EDdegree}(X \cap H) \quad = \quad \begin{cases} \text{EDdegree}(X) - \text{degree}(X^*) & \text{if } \operatorname{codim}(X^*) = 1, \\ \quad \text{EDdegree}(X) & \text{if } \operatorname{codim}(X^*) \geq 2. \end{cases}$$

Proof. Piene [30] showed that $\delta_i(X \cap H) = \delta_{i+1}(X)$ for all $i \geq 0$. By Theorem 5.4, the desired ED degree is the sum of these numbers, so it equals $\text{EDdegree}(X) - \delta_0(X)$. However, we know that $\delta_0(X)$ equals the degree of X^* if X^* is a hypersurface and it is zero otherwise. □

Example 6.5. Let X_r be the projective variety of symmetric 3×3-matrices of rank $\leq r$. We know that $X_r^* = X_{3-r}$ and $\text{EDdegree}(X_2) = \text{EDdegree}(X_1) = 13$. If H is a generic hyperplane in $\mathbb{P}^5$ then $\text{EDdegree}(X_2 \cap H) = 13$ but $\text{EDdegree}(X_1 \cap H) = 13 - 3 = 10$. ♢

If X is a variety of high dimension in $\mathbb{P}^{n-1}$ then Corollary 6.4 can be applied repeatedly until the generic linear section is a curve. This motivates the following definition which parallels its analogue in the multiplicative setting of likelihood geometry [24, §3]. The *sectional ED degree* of the variety X is the following binary form of degree $n-1$ in (x, u):

$$\sum_{i=0}^{\dim(X)-1} \text{EDdegree}(X \cap L_i) \cdot x^i \cdot u^{n-1-i} \tag{6.2}$$

where L_i is a generic linear section of codimension i. Corollary 6.4 implies that, for varieties in generic coordinates as in Theorem 5.4, this equals

$$\sum_{0 \leq i \leq j < \dim(X)} \delta_j(X) \cdot x^i \cdot u^{n-1-i}.$$

It would be interesting to get a better understanding of the sectional ED degree also for varieties in special coordinates. For instance, in light of [24, Conjecture 3.19], we may ask how (6.2) is related to the bidegree of the projective ED correspondence, or to the tridegree of the joint projective ED correspondence. For a concrete application, suppose that X is a determinantal variety, in the special coordinates of the Eckart-Young Theorem (Example 2.3). Minimizing the squared distance function d_u over a linear section $X \cap L_i$ is known as *structured low-rank matrix approximation.* This problem has numerous applications in engineering [7]. A study is under way in collaboration with Pierre-Jean Spaenlehauer.

We now change the topic to homogenization. Geometrically, this is the passage from an affine variety $X \subset \mathbb{C}^n$ to its projective closure $\overline{X} \subset \mathbb{P}^n$. This is a standard operation in algebraic geometry [8, §8.4]. Homogenization often preserves the solution set to a given geometric problem, but the analysis is simpler in $\mathbb{P}^n$ since projective space is compact. Algebraically, we proceed as follows. Given the ideal $I_X = \langle f_1, \ldots, f_s \rangle \subset \mathbb{R}[x_1, \ldots, x_n]$, we introduce a new variable x_0, representing the hyperplane at infinity, $H_\infty = \mathbb{P}^n \backslash \mathbb{C}^n = V(x_0)$. Given a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ of degree d, its *homogenization* $\overline{f} \in \mathbb{R}[x_0, x_1, \ldots, x_n]$ is defined by $\overline{f}(x_0, \ldots, x_n) = x_0^d \cdot f(x_1/x_0, \ldots, x_n/x_0)$. The ideal $I_{\overline{X}}$ of the projective variety $\overline{X}$ is generated by $\{\overline{f} \,:\, f \in I_X\}$. It can be computed (e.g. in `Macaulay2`) by saturation:

$$I_{\overline{X}} \;=\; \langle \overline{f}_1, \ldots, \overline{f}_s \rangle : \langle x_0 \rangle^\infty \quad \subseteq \; \mathbb{R}[x_0, x_1, \ldots, x_n].$$

One might naively hope that $\mathrm{EDdegree}(X) = \mathrm{EDdegree}(\overline{X})$. But this is false in general:

Example 6.6. Let X be the cardioid in Example 1.1. Written in the notation above, its projective closure is the quartic curve $\overline{X} \subset \mathbb{P}^2$ whose defining homogeneous ideal equals

$$I_{\overline{X}} \;=\; \langle\, x_0^2 x_2^2 - 2x_0 x_1^3 - 2x_0 x_1 x_2^2 - x_1^4 - 2x_1^2 x_2^2 - x_2^4 \,\rangle.$$

For this curve we have

$$\mathrm{EDdegree}(X) \;=\; 3 \;<\; 7 \;=\; \mathrm{EDdegree}(\overline{X}).$$

By contrast, consider the affine surface $Y = V(x_1 x_2 - x_3) \subset \mathbb{C}^3$. Its projective closure is the $2{\times}2$-determinant $\overline{Y} = V(x_1 x_2 - x_0 x_3) \subset \mathbb{P}^3$. Here the inequality goes in the other direction:

$$\mathrm{EDdegree}(Y) \;=\; 5 \;>\; 2 \;=\; \mathrm{EDdegree}(\overline{Y}). \tag{6.3}$$

The same phenomenon was seen in our study of Hurwitz determinants in Theorem 3.6. $\diamondsuit$

To explain what is going on here, we recall that $\mathrm{EDdegree}(\overline{X})$ is defined as the ED degree of the affine cone over the projective variety $\overline{X} \subset \mathbb{P}^n$, which we also denote by $\overline{X}$. Explicitly,

$$\overline{X} \;=\; \{\, (t, tx) \mid x \in X, t \in \mathbb{C} \,\} \;\subset\; \mathbb{C}^{n+1}.$$

The ED degree of $\overline{X}$ is for the fixed quadratic form $x_0^2 + x_1^2 + \cdots + x_n^2$ that cuts out the isotropic quadric $Q \subset \mathbb{P}^n$. This is just one of the infinitely many quadratic forms on $\mathbb{C}^{n+1}$ that restrict to the given form $x \cdot x = x_1^2 + \cdots + x_n^2$ on $\mathbb{C}^n$. That is one reason why the ED degrees of X and of $\overline{X}$ are not as closely related as one might hope. Nevertheless, we will now make the relation more explicit. The affine variety X is identified with the intersection of the cone $\overline{X}$ with the hyperplane $\{x_0 = 1\}$. Its part at infinity is denoted $X_\infty := \overline{X} \cap H_\infty$.

The data point $(1, 0) \in \mathbb{C}^{n+1}$ plays a special role, since it is the orthogonal projection of the vertex $(0, 0)$ of the cone $\overline{X}$ onto the affine hyperplane $\{x_0 = 1\}$. The following lemma relates the critical points for $u = 0$ on X to the critical points for $u = (1, 0)$ on $\overline{X}$.

Lemma 6.7. *Assume that all critical points of d_0 on X satisfy $x \cdot x \neq -1$. Then the map*

$$x \mapsto \left(\frac{1}{1+(x\cdot x)}, \frac{1}{1+(x\cdot x)} x \right)$$

is a bijection from the critical points of d_0 on X to the critical points of $d_{(1,0)}$ on $\overline{X} \backslash X_\infty$.

Proof. Let $t \in \mathbb{C}\backslash\{0\}$ and $x \in X\backslash X_{\text{sing}}$. The point $(t, tx) \in \overline{X}$ is critical for $d_{(1,0)}$ if and only if $(1-t, -tx)$ is perpendicular to $T_{(t,tx)}\overline{X}$. That space is spanned by $\{0\} \times T_x X$ and $(1, x)$. Hence $(1-t, -tx)$ is perpendicular to $T_{(t,tx)}\overline{X}$ if and only if $x \perp T_x X$ and $(1-t)-t(x\cdot x) = 0$. The first condition says that x is critical for d_0, and the second gives $t = 1/(1+(x\cdot x))$. □

If, under the assumptions in Lemma 6.7, the number of critical points of d_0 equals the ED degree of X, then we can conclude $\text{EDdegree}(X) \le \text{EDdegree}(\overline{X})$, with equality if none of the critical points of $d_{(1,0)}$ on $\overline{X}$ lies at infinity. To formulate a condition that guarantees equality, we fix the isotropic quadric $Q_\infty = \{x_1^2 + \cdots + x_n^2 = 0\}$ in H_∞. Our condition is:

$$\text{The intersections } X_\infty = \overline{X} \cap H_\infty \text{ and } X_\infty \cap Q_\infty \text{ are both transversal.} \tag{6.4}$$

Lemma 6.8. *If (6.4) holds then none of the critical points of $d_{(1,0)}$ on $\overline{X}$ lies in X_∞.*

Proof. Arguing by contradiction, suppose that $(0, x_\infty) \in X_\infty$ is a critical point of $d_{(1,0)}$ on $\overline{X}$. Then $(1, -x_\infty)$ is perpendicular to $T_{(0,x_\infty)}\overline{X}$, and hence $(0, x_\infty)$ is perpendicular to $H_\infty \cap T_{(0,x_\infty)}\overline{X}$. By transversality of $\overline{X}$ and H_∞, the latter is the tangent space to X_∞ at $(0, x_\infty)$. Hence $T_{(0,x_\infty)}X_\infty$ is contained in $(0, x_\infty)^\perp$, and X_∞ is tangent to Q_∞ at $(0, x_\infty)$. □

Fix $v \in \mathbb{C}^n$ and consider the affine translate $X_v := X - v = \{x - v \mid x \in X\}$. Its projective closure $\overline{X}_v$ is isomorphic to $\overline{X}$ as a projective variety in $\mathbb{P}^n$. However, the metric properties of the corresponding cones in $\mathbb{C}^{n+1}$ are rather different. While $\text{EDdegree}(X_v) = \text{EDdegree}(X)$ holds trivially, it is possible that $\text{EDdegree}(\overline{X}_v) \neq \text{EDdegree}(\overline{X})$. Here is a simple example:

Example 6.9. Consider the unit circle $X = \{x_1^2 + x_2^2 = 1\}$ in the plane. Then $\text{EDdegree}(X) = \text{EDdegree}(\overline{X}) = 2$. For generic $v \in \mathbb{R}^2$, the translated circle X_v has $\text{EDdegree}(\overline{X}_v) = 4$. ♢

Affine translation sheds light on the behavior of the ED degree under homogenization.

Proposition 6.10. *Let X be an irreducible variety in $\mathbb{C}^n$, and let $v \in \mathbb{C}^n$ be a generic vector. Then $\text{EDdegree}(X) \le \text{EDdegree}(\overline{X}_v)$, and equality holds if the hypothesis (6.4) is satisfied.*

The genericity hypothesis (6.4) simply says that X_∞ and $X_\infty \cap Q_\infty$ are smooth. Note that this does not depend on the extension of the quadric Q_∞ to $\mathbb{C}^{n+1}$.

Proof. Since translation of affine varieties preserves ED degree, the inequality follows from Lemma 6.7 provided $x' \cdot x' \neq -1$ for all critical points x' for d_0 on X_v. These are the points $x' = x - v$ with x critical for d_v, i.e., with $(x, v) \in \mathcal{E}_X$. The expression $(x - v)\cdot(x - v)$ is not constant -1 on the irreducible variety $\mathcal{E}_X$, because it is zero on the diagonal $\Delta(X) \subset \mathcal{E}_X$. As a consequence, the variety of pairs $(x, v) \in \mathcal{E}_X$ with $(x-v)\cdot(x-v) = -1$ has dimension $\le n-1$. In particular, it does not project dominantly onto the second factor $\mathbb{C}^n$. Taking v outside that projection, and such that the number of critical points of d_v on X is equal to $\text{EDdegree}(X)$, ensures that we can apply Lemma 6.7. The second statement follows from Lemma 6.8 applied to X_v and the fact that X and X_v have the same behavior at infinity. □

Our main result on homogenization links the discussion above to the polar classes of $\overline{X}$.

Theorem 6.11. *For any irreducible affine variety X in $\mathbb{C}^n$ we have the two inequalities*

$$\mathrm{EDdegree}(X) \;\leq\; \sum_{i=0}^{n-1} \delta_i(\overline{X}) \quad \textit{and} \quad \mathrm{EDdegree}(\overline{X}) \;\leq\; \sum_{i=0}^{n-1} \delta_i(\overline{X}),$$

with equality on the left if (6.4) holds, and equality on the right if the conormal variety $\mathcal{N}_{\overline{X}}$ is disjoint from the diagonal $\Delta(\mathbb{P}^n)$ in $\mathbb{P}^n_x \times \mathbb{P}^n_y$. The equality on the right holds in particular if $X \cap Q$ is smooth and disjoint from X_{sing} (see the statement after Theorem 5.4).

Proof. We claim that for generic $v \in \mathbb{C}^n$ the conormal variety $\mathcal{N}_{\overline{X}_v}$ does not intersect $\Delta(\mathbb{P}^n)$. For this we need to understand how $\mathcal{N}_{\overline{X}_v}$ changes with v. The $(1+n)\times(1+n)$ matrix

$$A_v := \begin{pmatrix} 1 & 0 \\ -v & I_n \end{pmatrix}$$

defines an automorphism $\mathbb{P}^n_x \to \mathbb{P}^n_x$ that maps $\overline{X}$ isomorphically onto $\overline{X}_v$. The second factor $\mathbb{P}^n_y$ is the dual of $\mathbb{P}^n_x$ and hence transforms contragradiently, i.e., by the matrix A_v^{-T}. Hence the pair of matrices (A_v, A_v^{-T}) maps $\mathcal{N}_{\overline{X}}$ isomorphically onto $\mathcal{N}_{\overline{X}_v}$. Consider the variety

$$Z \;:=\; \big\{\, (x, y, v) \in \mathcal{N}_{\overline{X}} \times \mathbb{C}^n \mid A_v x = A_v^{-T} y \big\}.$$

For fixed $(x, y) = ((x_0 : x_\infty), (y_0 : y_\infty)) \in \mathcal{N}_{\overline{X}}$ with $x_0 \neq 0$, the equations defining Z read

$$x_0 = c(y_0 + v^T y_\infty) \quad \text{and} \quad -x_0 v + x_\infty = c y_\infty,$$

for $v \in \mathbb{C}^n$ and a scalar c reflecting that we work in projective space. The second equation expresses v in c, x, y. Substituting that expression into the first equation gives a system for c with at most 2 solutions. This shows that $\dim Z$ is at most $\dim \mathcal{N}_{\overline{X}} = n-1$, so the image of Z in $\mathbb{C}^n$ is contained in a proper subvariety of $\mathbb{C}^n$. For any v outside that subvariety, $\mathcal{N}_{\overline{X}_v} \cap \Delta(\mathbb{P}^n) = \emptyset$. For those v, Theorem 5.4 implies that $\mathrm{EDdegree}(\overline{X}_v)$ is the sum of the polar classes of $\overline{X}_v$, which are also those of $\overline{X}$ since they are projective invariants. Since $\mathrm{EDdegree}(\overline{X}_v)$ can only go down as v approaches a limit point, this yields the second inequality, as well as the sufficient condition for equality there. By applying Proposition 6.10, we establish the first inequality, as well as the sufficient condition (6.4) for equality. □

Example 6.12. Consider the quadric surface $\overline{Y} = V(x_0x_3 - x_1x_2) \subset \mathbb{P}^3$ from Example 6.6. This is the toric variety whose polytope P is the unit square. By Corollary 5.11, the sum of the polar classes equals $7V_2 - 3V_1 + V_0 = 14 - 12 + 4 = 6$. Comparing this with (6.3), we find that neither of the two inequalities in Theorem 6.11 is an equality. This is consistent with the fact that $Y_\infty := \overline{Y} \cap H_\infty = V(x_1x_2)$ is not smooth at the point $(0:0:0:1)$ and the fact that $\overline{Y}$ and Q are tangent at the four points $(1 : a_1 : a_2 : a_1a_2)$ with $a_1, a_2 = \pm i$. ♢

Example 6.13. Consider the threefold $\overline{Z} = V(x_1x_4 - x_2x_3 - x_0^2 - x_0x_1)$ in $\mathbb{P}^4$. Then Z_∞ is isomorphic to $\overline{Y}$ from the previous example and smooth in $\mathbb{P}^3$, but $Z_\infty \cap Q_\infty$ is isomorphic to the $\overline{Y} \cap Q$ from the previous example and hence has four non-reduced points. Here, we have $\mathrm{EDdegree}(Z) = 4 < 8 = \mathrm{EDdegree}(\overline{Z}) = \sum_{i=0}^{3} \delta_i(\overline{Z})$. If we replace x_1x_4 by $2x_1x_4$ in the equation defining $\overline{Z}$, then the four non-reduced points disappear. Now $Z_\infty \cap Q_\infty$ is smooth, we have $\mathrm{EDdegree}(Z) = 8$, and both inequalities in Theorem 6.11 hold with equality. ♢

Example 6.14. Let X be the cardioid from Examples 1.1 and 6.6. This curve violates both conditions for equality in Theorem 6.11. Here $X_\infty = V(x_1^4 + 2x_1^2x_2^2 + x_2^4)$ agrees with $Q_\infty = V(x_1^2 + x_2^2)$ as a subset of $H_\infty \simeq \mathbb{P}^1$, but it has multiplicity two at the two points. $\diamondsuit$

7 ED discriminant and Chern Classes

Catenese and Trifogli [6, 37] studied ED discriminants under their classical name *focal loci*. We present some of their results, including a formula for the ED degree in terms of Chern classes, and we discuss a range of applications. We work in the projective setting, so X is a subvariety of $\mathbb{P}^{n-1}$, equipped with homogeneous coordinates $(x_1 : \ldots : x_n)$ and $\mathcal{PE}_X \subset \mathbb{P}_x^{n-1} \times \mathbb{C}_u^n$ is its projective ED correspondence. By Theorem 4.4, the ED degree is the size of the general fiber of the map $\mathcal{PE}_X \to \mathbb{C}_u^n$. The branch locus of this map is the closure of the set of data points u for which there are fewer than EDdegree(X) complex critical points. Since the variety $\mathcal{PE}_X \subset \mathbb{P}_x^{n-1} \times \mathbb{C}_u^n$ is defined by bihomogeneous equations in x, u, also the branch locus is defined by homogeneous equations and it is a cone in $\mathbb{C}_u^n$. Hence the branch locus defines a projective variety $\Sigma_X \subset \mathbb{P}_u^{n-1}$, which we call the *ED discriminant*. The ED discriminant Σ_X is typically an irreducible hypersurface, by the Nagata-Zariski Purity Theorem, and we are interested in its degree and defining polynomial.

Remark 7.1. In applications, the *uniqueness* of the closest real-valued point $u^* \in X$ to a given data point u is relevant. In many cases, e.g. for symmetric tensors of rank one [12], this closest point is unique for u outside an algebraic hypersurface that strictly contains Σ_X.

Example 7.2. Let $n = 4$ and consider the quadric surface $X = V(x_1x_4 - 2x_2x_3) \subset \mathbb{P}_x^3$. This is the 2×2-determinant in general coordinates, so EDdegree$(X) = 6$. The ED discriminant is a irreducible surface of degree 12 in $\mathbb{P}_u^3$. Its defining polynomial has 119 terms:

$$\begin{aligned}\Sigma_X \;=\; & 65536u_1^{12} + 835584u_1^{10}u_2^2 + 835584u_1^{10}u_3^2 - 835584u_1^{10}u_4^2 + 9707520u_1^9u_2u_3u_4 \\ & +3747840u_1^8u_2^4 - 7294464u_1^8u_2^2u_3^2 + \cdots + 835584u_3^2u_4^{10} + 65536u_4^{12}.\end{aligned}$$

This ED discriminant can be computed using the following `Macaulay2` code:

```
R = QQ[x1,x2,x3,x4,u1,u2,u3,u4];  f = x1*x4-2*x2*x3;
EX = ideal(f) + minors(3,matrix {{u1,u2,u3,u4},{x1,x2,x3,x4},
 {diff(x1,f),diff(x2,f),diff(x3,f),diff(x4,f)} });
g = first first entries gens eliminate({x3,x4},EX);
toString factor discriminant(g,x2)
```

Here `EX` is the ideal of the ED correspondence in $\mathbb{P}_x^3 \times \mathbb{P}_u^3$. The command `eliminate` maps that threefold into $\mathbb{P}^1_{(x_1:x_2)} \times \mathbb{P}_u^3$. We print the discriminant of that hypersurface over $\mathbb{P}^3$. $\diamondsuit$

If X is a general hypersurface of degree d in $\mathbb{P}^{n-1}$ then, by Corollary 2.9,

$$\text{EDdegree}(X) \;=\; d \cdot \frac{(d-1)^{n-1} - 1}{d-2}. \tag{7.1}$$

Trifogli [37] determined the degree of the ED discriminant Σ_X for such a hypersurface X:

Theorem 7.3 (Trifogli)**.** *If X is a general hypersurface of degree d in $\mathbb{P}^{n-1}$ then*

$$\text{degree}(\Sigma_X) \;=\; d(n-2)(d-1)^{n-2} \;+\; 2d(d-1)\frac{(d-1)^{n-2}-1}{d-2}. \tag{7.2}$$

Example 7.4. A general plane curve X has EDdegree$(X) = d^2$ and degree$(\Sigma_X) = 3d(d-1)$. These are the numbers seen for the ellipse ($d = 2$) in Example 4.5. For a plane quartic curve X, we expect EDdegree$(X) = 16$ and degree$(\Sigma_X) = 36$, in contrast to the numbers 3 and 4 for the cardioid in Example 1.1. A general surface in $\mathbb{P}^3$ has EDdegree$(X) = d(d^2 - d + 1)$ and degree$(\Sigma_X) = 2d(d-1)(2d-1)$. For quadrics ($d = 2$) we get 6 and 12, as in Example 7.2. ♢

Example 7.5. The ED discriminant Σ_X of a plane curve X was already studied in the 19th century, under the names *evolute* or *caustic*. Salmon [32, page 96, art. 112] showed that a curve $X \subset \mathbb{P}^2$ of degree d with with δ ordinary nodes and k ordinary cusps has degree$(\Sigma_X) = 3d^2 - 3d - 6\delta - 8k$. For affine $X \subset \mathbb{C}^2$, the same holds provided that $\overline{X} \subseteq \mathbb{P}^2$ is not tangent to the line H_∞ and neither of the two isotropic points on H_∞ is on $\overline{X}$. ♢

We comment on the relation between duality and the ED discriminant Σ_X. Recall that Σ_X is the projectivization of the branch locus of the covering $\mathcal{PE}_X \to \mathbb{C}^n_u$. By the results in Section 5, this is also the branch locus of $\mathcal{PE}_{X,Y} \to \mathbb{C}^n_u$, and hence also of $\mathcal{PE}_Y \to \mathbb{C}^n_u$. This implies that the ED discriminant of a variety X agrees with that of its dual variety $Y = X^*$.

Example 7.6. Let $X \subset \mathbb{P}^2_x$ denote the cubic Fermat curve given by $x_0^3 + x_1^3 + x_2^3 = 0$. Its dual Y is the sextic curve in $\mathbb{P}^2_y$ that is defined by $y_0^6 + y_1^6 + y_2^6 - 2y_0^3y_1^3 - 2y_0^3y_2^3 - 2y_1^3y_2^3$. This pair of curves satisfies EDdegree(X) = EDdegree$(Y) = 9$. The ED discriminant $\Sigma_X = \Sigma_Y$ is an irreducible curve of degree 18 in $\mathbb{P}^2_u$. Its defining polynomial has 184 terms:

$$\Sigma_X = 4u_0^{18} - 204u_0^{16}u_1^2 + 588u_0^{15}u_1^3 - 495u_0^{14}u_1^4 + 2040u_0^{13}u_1^5 - 2254u_0^{12}u_1^6 + 2622u_0^{11}u_1^7 + \cdots + 4u_2^{18}.$$

The computation of the ED discriminant for larger examples in `Macaulay2` is difficult. ♢

The formulas (7.1) and (7.2) are best understood and derived using modern intersection theory; see [16] or [20, Appendix A]. That theory goes far beyond the undergraduate algebraic geometry [8] used in the earlier sections but is indispensable for more general degree formulas, especially for varieties X of codimension ≥ 2. We briefly sketch some of the required vector bundle techniques.

A vector bundle $\mathcal{E} \to X$ on a smooth, m-dimensional projective variety X has a *total Chern class* $c(\mathcal{E}) = c_0(\mathcal{E}) + \ldots + c_m(\mathcal{E})$, which resides in the cohomology ring $H^*(X) = \bigoplus_{i=0}^m H^{2i}(X)$. In particular, the *top Chern class* $c_m(\mathcal{E})$ is an integer scalar multiple of the class of a point, and that integer is commonly denoted $\int c(\mathcal{E})$. If $\mathcal{E}$ has rank equal to $\dim X = m$, and if $s : X \to \mathcal{E}$ is a global section for which $V(s) := \{x \in X \mid s(x) = 0\}$ consists of finitely many simple points, then the cardinality of $V(s)$ equals $\int c(\mathcal{E})$. To apply this to the computation of ED degrees, we shall find $\mathcal{E}$ and s such that the variety $V(s)$ is the set of critical points of d_u, and then compute $\int c(\mathcal{E})$ using vector bundle tools. Among these tools are *Whitney's sum formula* $c(\mathcal{E}) = c(\mathcal{E}')\cdot c(\mathcal{E}'')$ for any exact sequence $0 \to \mathcal{E}' \to \mathcal{E} \to \mathcal{E}'' \to 0$ of vector bundles on X, and the fact that the total Chern class of the pull-back of $\mathcal{E}$ under a morphism $X' \to X$ is the image of $c(\mathcal{E})$ under the ring homomorphism $H^*(X) \to H^*(X')$.

Here is our repertoire of vector bundles on X: the trivial bundle $X \times \mathbb{C}^n$ of rank n; the *tautological line bundle* pulled back from $\mathbb{P}^{n-1}$, which is $\mathcal{R}_X := \{(x,v) \in X \times \mathbb{C}^n \mid v \in x\}$ (also often denoted by $\mathcal{O}_X(-1)$, while the dual $\mathcal{R}_X^*$ is denoted by $\mathcal{O}_X(1)$); the *tangent bundle* TX whose fibers are the tangent spaces T_xX; the *cotangent bundle* T^*X whose fibers are their duals $(T_xX)^*$; and the *normal bundle* N_X whose fibers are the quotient $T_x\mathbb{P}^n/T_xX$. From these building blocks, we can construct new vector bundles using direct sums, tensor products, quotients, duals, and orthogonal complements inside the trivial bundle $X \times \mathbb{C}^n$.

Theorem 7.7. *Let X be a smooth and irreducible projective variety in $\mathbb{P}^{n-1}$ and assume that X intersects the isotropic quadric $Q = V(x_1^2 + x_2^2 + \cdots + x_n^2)$ transversally, i.e. $X \cap Q$ is smooth. Then the EDdegree of X can be computed in $H^*(X)$ by either of the expressions*

$$\mathrm{EDdegree}(X) \;=\; \int \frac{c(\mathcal{R}_X^*)\cdot c(T^*X\otimes\mathcal{R}_X^*)}{c(\mathcal{R}_X)} \;=\; \int \frac{1}{c(\mathcal{R}_X)\cdot c(N_X^*\otimes\mathcal{R}_X^*)}. \tag{7.3}$$

Proof. The first expression is due to Catanese and Trifogli. It is stated after Remark 3 on page 6026 in [6], as a formula for the *inverse* of the total Chern class of what they call *Euclidean normal bundle* (for simplicity we tensor it by $\mathcal{R}_X^*$, differently from [6]). The total space of that bundle, called *normal variety* in [6, 37], is precisely our projective ED correspondence $\mathcal{PE}_X$ from Theorem 4.4.

A generic data point $u \in \mathbb{C}^n$ gives rise to a section $x \mapsto [(x,u)]$ of the quotient bundle $(X \times \mathbb{C}^n)/\mathcal{PE}_X$, whose zero set is exactly the set of critical points of d_u. By Whitney's sum formula, the total Chern class of this quotient is $1/c(\mathcal{PE}_X)$. This explains the inverse and the first formula. The second formula is seen using the identity

$$\frac{1}{c(N_X^*\otimes\mathcal{R}_X^*)} \;=\; \frac{c(T^*X\otimes\mathcal{R}_X^*)}{c(T^*\mathbb{P}^{n-1}\otimes\mathcal{R}_X^*)} \;=\; c(\mathcal{R}_X^*)\cdot c(T^*X\otimes\mathcal{R}_X^*), \tag{7.4}$$

where the second equality follows from the Euler sequence ([20], example II.8.20.1). (We remark that the ED degree of X can also be interpreted as the top Segre class [16] of the Euclidean normal bundle of X.) ☐

We shall now relate this discussion to the earlier formula in Section 5, by offering a

Second proof of Theorem 7.7. If $\mathcal{E}$ is a vector bundle of rank m and $\mathcal{L}$ is a line bundle then

$$c_k(\mathcal{E}\otimes\mathcal{L}) \;=\; \sum_{i=0}^{k} \binom{r-i}{k-i} c_i(\mathcal{E})c_1(\mathcal{L})^{k-i}. \tag{7.5}$$

This formula is [16, Example 3.2.2]. By definition, we have $c_i(X) = (-1)^i c_i(T^*X)$. Setting $c_1(\mathcal{R}_X^*) = h$, the formula (7.5) implies

$$c(T^*X\otimes\mathcal{R}_X^*) \;=\; \sum_{k=0}^{m}\sum_{i=0}^{k}\binom{m-i}{k-i}(-1)^i c_i(X) h^{k-i} \;=\; \sum_{i=0}^{m}(-1)^i c_i(X)\sum_{t=0}^{m-i}\binom{m-i}{t}h^t.$$

We have $c(\mathcal{R}_X^*) = 1 + h$ and $1/c(\mathcal{R}_X) = 1/(1-h) = \sum_{i=0}^{m} h^i$. The equation above implies

$$\frac{c(\mathcal{R}_X^*)\cdot c(T^*X\otimes\mathcal{R}_X^*)}{c(\mathcal{R}_X)} \;=\; \sum_{i=0}^{m}(-1)^i c_i(X)\left(\sum_{t=0}^{m-i}\binom{m-i}{t}h^t\right)\left(1+2\sum_{j=1}^{m}h^j\right).$$

The integral on the left hand side in (7.3) is the coefficient of h^{m-i} in the polynomial in h that is obtained by multiplying the two parenthesized sums. That coefficient equals

$$1+2\sum_{j=0}^{m-i-1}\binom{m-i}{j} \;=\; 2^{m-i+1}-1.$$

We conclude that Theorem 5.8 is in fact equivalent to the first formula in Theorem 7.7. The second formula follows from (7.4), as argued above. □

The Catanese-Trifogli formula in (7.3) is most useful when X has low codimension. In that case, we typically compute the relevant class in the cohomology ring of the ambient projective space $\mathbb{P}^n$, and then pull back to X. On the other hand, if X is a low-dimensional variety, then Theorem 5.8 may be more useful, especially if X is a variety whose cohomology ring we understand well. We illustrate these two scenarios with concrete examples.

Example 7.8. Let X be a generic hypersurface of degree d in $\mathbb{P}^{n-1}$. We compute in $H^*(\mathbb{P}^{n-1}) = \mathbb{Z}[h]/\langle h^n\rangle$. The line bundle $\mathcal{R}_X$ is the pull-back of $\mathcal{R}_{\mathbb{P}^{n-1}}$, whose total Chern class is $1-h$. Since $\operatorname{codim}(X)=1$, the vector bundle N_X is a line bundle. By [20, Example II.8.20.3], we have $N_X = (\mathcal{R}_X^*)^{\otimes d}$, so that $N_X^*\otimes\mathcal{R}_X^* = (\mathcal{R}_X)^{\otimes(d-1)}$. In $H^*(\mathbb{P}^{n-1})$ we have

$$\frac{1}{c(\mathcal{R}_{\mathbb{P}^{n-1}})\cdot c(\mathcal{R}_{\mathbb{P}^{n-1}}^{\otimes d-1})} \;=\; \frac{1}{(1-h)(1-(d-1)h)}.$$

The coefficient of h^{n-2} in this expression equals $\sum_{i=0}^{n-2}(d-1)^i$, and since the image of h^{n-2} in $H^*(X)$ under pull-back equals $d=\operatorname{degree}(X)$ times the class of a point, we find

$$\operatorname{EDdegree}(X) \;=\; \int\frac{1}{c(\mathcal{R}_X)\cdot c(\mathcal{N}_X^*)} \;=\; d\cdot\sum_{i=0}^{n-2}(d-1)^i.$$

A similar reasoning applies when X is a general complete intersection of c hypersurfaces of degrees $d_1,\ldots,d_c$. Again, by working in $H^*(\mathbb{P}^{n-1}) = \mathbb{Z}[h]/\langle h^n\rangle$, we evaluate

$$\operatorname{EDdegree}(X) \;=\; \int\frac{1}{(1-h)\prod_{i=1}^{c}(1-(d_i-1)h)},$$

where $\int$ refers to the coefficient of the point class in the pull-back to X. To compute this, we expand the integrand as a series in h. The coefficient of h^{n-c-1} in that series, multiplied by $\operatorname{degree}(X) = d_1\cdots d_c$, is the formula in (2.8). Proposition 2.5 then follows from Theorem 6.11. Here is the argument. After a transformation (if necessary) of the given equations $f_1,\ldots,f_s$, the variety X' cut out by the first c of them is a complete intersection. Then X is an irreducible component of X'. This implies $\operatorname{EDdegree}(X)\le\operatorname{EDdegree}(X')$. Now, by semicontinuity, $\operatorname{EDdegree}(X')$ is at most the value for a generic complete intersection. ♢

The following computation generalizes Example 5.12 from $X\simeq\mathbb{P}^1$ to higher dimensions.

Proposition 7.9. *After a change of coordinates that creates a transverse intersection with the isotropic quadric Q in $\mathbb{P}^{\binom{m+d}{d}-1}$, the d-th Veronese embedding of $\mathbb{P}^m$ has ED degree*

$$\frac{(2d-1)^{m+1}-(d-1)^{m+1}}{d}. \tag{7.6}$$

Proof. We write $i_d\colon \mathbb{P}^{m-1} \to X$ for the dth-Veronese embedding in question. So, X denotes the image of $\mathbb{P}^{m-1}$ in $\mathbb{P}^{\binom{m+d-1}{d}-1}$ under the map given by a sufficiently general basis for the space of homogeneous polynomials of degree d in m variables. We have $c_i(X) = \binom{m+1}{i}h^i$, so that $\deg c_i(X) = \int (dh)^{m-i} c_i(X) = \binom{m}{i} d^{m-i}$. From Theorem 5.8 we now get

$$\text{EDdegree}(X) \;=\; \sum_{i=0}^{m} (-1)^i (2^{m+1-i}-1) \binom{m+1}{i} d^{m-i}.$$

Using the Binomial Theorem, we see that this alternating sum is equal to (7.6). □

Theorem 7.7 requires X to be smooth. Varieties with favorable desingularizations are also amenable to Chern class computations, but the computations become more technical.

Example 7.10. Let X_r denote the variety of $s \times t$ matrices of rank $\leq r$, in generic coordinates so that X_r intersects Q transversally. Its ED degree can be computed by the desingularization in [38, Proposition 6.1.1.a]. The Chern class formula amounts to a nontrivial computation in the ring of symmetric functions. We implemented this in `Macaulay2` as follows:

```
loadPackage "Schubert2"
ED=(s,t,r)->
(G = flagBundle({r,s-r}); (S,Q) = G.Bundles;
X=projectiveBundle (S^t); (sx,qx)=X.Bundles;
d=dim X; T=tangentBundle X;
sum(d+1,i->(-1)^i*(2^(d+1-i)-1)*integral(chern(i,T)*(chern(1,dual(sx)))^(d-i)))
```

The first values of EDdegree(X_r) are summarized in the following table

(s,t) =	(2,2)	(2,3)	(2,4)	(2,5)	(3,3)	(3,4)	(3,5)	(4,4)	(4,5)	(5,5)
$r=1$	6	10	14	18	39	83	143	284	676	2205
$r=2$					39	83	143	1350	4806	55010
$r=3$								284	676	55010
$r=4$										2205

The $r=1$ row can also be computed with P a product of two simplices in Corollary 5.11. ♢

Using the formalism of Chern classes, Catanese and Trifogli [6, page 6030] derive a general formula for the degree of the ED discriminant Σ_X. Their formula is a complicated expression in terms of the Chow ring of the ED correspondence $\mathcal{PE}_X$. Here are two easier special cases.

Example 7.11. If X is a general smooth curve in $\mathbb{P}^n$ of degree d and genus g then

$$\text{degree}(\Sigma_X) \;=\; 6(d+g-1).$$

For instance, the rational normal curve X in general coordinates in $\mathbb{P}^n$, as discussed in Example 5.12, has $\text{degree}(X) = n$, $\text{EDdegree}(X) = 3n-2$, and $\text{degree}(\Sigma_X) = 6n-6$.

If X is a general smooth surface in $\mathbb{P}^n$ of degree d, with Chern classes $c_1(X), c_2(X)$, then

$$\text{degree}(\Sigma_X) \;=\; 2 \cdot \big(15 \cdot d + c_1(X)^2 + c_2(X) - 9 \cdot \deg c_1(X)\big).$$

The formulas in Example 7.4 can be derived from these expressions, as in [6, page 6034]. ♢

8 Tensors of Rank One

In this section we discuss low rank approximation of multidimensional tensors. We focus on tensors of rank one [12], and we present formulas for their ED degree and average ED degree. The former was computed by Friedland and Ottaviani in [13]. The latter is work in progress of Draisma and Horobeţ [9]. Our analysis includes partially symmetric tensors, and it represents a step towards extending the Eckart-Young theorem from matrices to tensors.

We consider real tensors $x = (x_{i_1 i_2 \cdots i_p})$ of format $m_1 \times m_2 \times \cdots \times m_p$. The space of such tensors is the tensor product $\mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2} \otimes \cdots \otimes \mathbb{R}^{m_p}$, which we identify with $\mathbb{R}^{m_1 m_2 \cdots m_p}$. The corresponding projective space $\mathbb{P}(\mathbb{R}^{m_1} \otimes \cdots \otimes \mathbb{R}^{m_p})$ is likewise identified with $\mathbb{P}^{m_1 m_2 \cdots m_p - 1}$.

A tensor x has *rank one* if $x = t_1 \otimes t_2 \otimes \cdots \otimes t_p$ for some vectors $t_i \in \mathbb{R}^{m_i}$. In coordinates,

$$x_{i_1 i_2 \cdots i_p} \; = \; t_{1 i_1} t_{2 i_2} \cdots t_{p i_p} \qquad \text{for} \quad 1 \leq i_1 \leq m_1, \ldots, 1 \leq i_p \leq m_p. \tag{8.1}$$

The set X of all tensors of rank one is an algebraic variety in $\mathbb{R}^{m_1 m_2 \cdots m_p}$. It is the cone over the *Segre variety* $\mathbb{P}(\mathbb{R}^{m_1}) \times \cdots \times \mathbb{P}(\mathbb{R}^{m_p}) = \mathbb{P}^{m_1 - 1} \times \cdots \times \mathbb{P}^{m_p - 1}$ in its natural embedding in $\mathbb{P}^{m_1 m_2 \cdots m_p - 1}$. By slight abuse of notation, we use the symbol X also for that Segre variety.

Theorem 8.1. ([13, Theorem 4]). *The ED degree of the Segre variety X of rank 1 tensors of format $m_1 \times \cdots \times m_p$ equals the coefficient of the monomial $z_1^{m_1 - 1} \cdots z_p^{m_p - 1}$ in the polynomial*

$$\prod_{i=1}^{p} \frac{(\widehat{z}_i)^{m_i} - z_i^{m_i}}{\widehat{z}_i \, - \, z_i} \qquad \textit{where} \quad \widehat{z}_i \; = \; z_1 + \cdots + z_{i-1} + z_{i+1} + \cdots + z_p.$$

The embedding (8.1) of the Segre variety X into $\mathbb{P}^{m_1 m_2 \cdots m_p - 1}$ is not transversal to the isotropic quadric Q, so our earlier formulas do not apply. However, it is natural in the following sense. The Euclidean distance on each factor $\mathbb{R}^{m_i}$ is preserved under the action by the rotation group $SO(m_i)$. The product group $SO(m_1) \times \cdots \times SO(m_p)$ embeds in the group $SO(m_1 \cdots m_p)$, which acts by rotations on the tensor space $\mathbb{R}^{m_1 m_2 \cdots m_p}$. The Segre map (8.1) from $\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_p}$ to $\mathbb{R}^{m_1 m_2 \cdots m_p}$ is $SO(m_1) \times \cdots \times SO(m_p)$-equivariant. This group invariance becomes crucial when, in a short while, we pass to partially symmetric tensors.

For $p = 2$, when the given tensor u is a matrix, Theorem 8.1 gives the Eckart-Young formula $\text{EDdegree}(X) = \min(m_1, m_2)$. The fact that singular vectors are the eigenvectors of $u^T u$ or $u u^T$, can be interpreted as a characterization of the ED correspondence $\mathcal{E}_X$. The following generalization to arbitrary tensors, due to Lim [26], is the key ingredient used in [13]. Suppose that $u = (u_{i_1 i_2 \cdots i_p})$ is a given tensor, and we seek to find its best rank one approximation $x^* = (x^*_{i_1 i_2 \cdots i_p}) = (t^*_{1 i_1} t^*_{2 i_2} \cdots t^*_{p i_p})$. Then we have the *singular vector equations*

$$u \cdot (t_1^* \otimes \cdots \otimes t_{i-1}^* \otimes t_{i+1}^* \otimes \cdots \otimes t_p^*) \; = \; \lambda t_i^* \tag{8.2}$$

where the scalars λ's are the *singular values* of the tensor u. The dot in (8.2) denotes tensor contraction. In the special case $p = 2$, these are the equations, familiar from linear algebra, that characterize the singular vector pairs of a rectangular matrix [13, (1.1)]. Theorem 8.1 is proved in [13] by counting the number of solutions to (8.2). The arguments used are based on Chern class techniques as described in Section 6.

Consider the ED correspondence $\mathcal{PE}_X$, introduced before Theorem 4.4, but now regarded as a subvariety of $\mathbb{P}^{m_1 \cdots m_p - 1} \times \mathbb{P}^{m_1 \cdots m_p - 1}$. Its equations can be derived as follows. The proportionality conditions of (8.2) are expressed as quadratic equations given by 2×2 minors. This leads to a system of bilinear equations in (x,u). These equations, together with the quadratic binomials in x for the Segre variety X, define the ED correspondence $\mathcal{PE}_X$.

Example 8.2. Let $p=3$, $m_1=m_2=m_3=2$, and abbreviate $a=t_1^*, b=t_2^*, c=t_3^*$, for the Segre embedding of $X=\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1$ into $\mathbb{P}^7$. This toric threefold is defined by the ideal

$$\begin{array}{l}\langle\, x_{101}x_{110}-x_{100}x_{111}\,,\ \ x_{011}x_{110}-x_{010}x_{111}\,,\ \ x_{011}x_{101}-x_{001}x_{111}\\ \ \ x_{010}x_{100}-x_{000}x_{110}\,,\ \ x_{001}x_{100}-x_{000}x_{101}\,,\ \ x_{001}x_{010}-x_{000}x_{011}\\ \ \ x_{010}x_{101}-x_{000}x_{111}\,,\ \ x_{011}x_{100}-x_{000}x_{111}\,,\ \ x_{001}x_{110}-x_{000}x_{111}\,\rangle.\end{array} \tag{8.3}$$

The six singular vector equations (8.2) for the $2\times2\times2$-tensor x reduces to the proportionality between the columns of the following three matrices

$$\begin{pmatrix} u_{000}b_0c_0+u_{001}b_0c_1+u_{010}b_1c_0+u_{011}b_1c_1 & a_0\\ u_{100}b_0c_0+u_{101}b_0c_1+u_{110}b_1c_0+u_{111}b_1c_1 & a_1\end{pmatrix}$$

$$\begin{pmatrix} u_{000}a_0c_0+u_{001}a_0c_1+u_{100}a_1c_0+u_{101}a_1c_1 & b_0\\ u_{010}a_0c_0+u_{011}a_0c_1+u_{110}a_1c_0+u_{111}a_1c_1 & b_1\end{pmatrix}$$

$$\begin{pmatrix} u_{000}a_0b_0+u_{010}a_0b_1+u_{100}a_1b_0+u_{110}a_1b_1 & c_0\\ u_{001}a_0b_0+u_{011}a_0b_1+u_{101}a_1b_0+u_{111}a_1b_1 & c_1\end{pmatrix}$$

We now take the three determinants, by using $a_ib_jc_k=x_{ijk}$, this gives the bilinear equations

$$\begin{array}{rcl} u_{000}x_{100}+u_{001}x_{101}+u_{010}x_{110}+u_{011}x_{111} &=& u_{100}x_{000}+u_{101}x_{001}+u_{110}x_{010}+u_{111}x_{011},\\ u_{000}x_{010}+u_{001}x_{011}+u_{100}x_{110}+u_{101}x_{111} &=& u_{010}x_{000}+u_{011}x_{001}+u_{110}x_{100}+u_{111}x_{101},\\ u_{000}x_{001}+u_{010}x_{011}+u_{100}x_{101}+u_{110}x_{111} &=& u_{001}x_{000}+u_{011}x_{010}+u_{101}x_{100}+u_{111}x_{110}.\end{array} \tag{8.4}$$

The ED correspondence $\mathcal{PE}_X\subset\mathbb{P}^7\times\mathbb{P}^7$ of $X=\mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1$ is defined by (8.3) and (8.4).

By plugging the binomials (8.3) into (2.7), we verify EDdegree$(X)=6$, the number from Theorem 8.1. By contrast, if we scale the x_{ijk} so that X meets the isotropic quadric Q transversally, then EDdegree$(X)=15\cdot6-7\cdot12+3\cdot12-1\cdot8=34$, by Corollary 5.11. $\diamondsuit$

Our duality results in Section 5 have nice consequences for rank one tensor approximation. It is known [17, Chapter XIV] that the dual variety $Y=X^*$ is a hypersurface if and only if

$$2\cdot\max(m_1,m_2,\ldots,m_p)\ \le\ m_1+m_2+\cdots+m_p-p+2. \tag{8.5}$$

In that case, the polynomial defining Y is the *hyperdeterminant* of format $m_1\times m_2\times\cdots\times m_p$. For instance, in Example 8.2, where P is the 3-cube, we get the $2\times2\times2$-*hyperdeterminant*

$$\begin{array}{rcl} Y &=& V\big(\,x_{000}^2x_{111}^2-2x_{000}x_{001}x_{110}x_{111}-2x_{000}x_{010}x_{101}x_{111}-2x_{000}x_{011}x_{100}x_{111}\\ && \quad +4x_{000}x_{011}x_{101}x_{110}+x_{001}^2x_{110}^2+4x_{001}x_{010}x_{100}x_{111}-2x_{001}x_{010}x_{101}x_{110}\\ && \quad -2x_{001}x_{011}x_{100}x_{110}+x_{010}^2x_{101}^2-2x_{010}x_{011}x_{100}x_{101}+x_{011}^2x_{100}^2\,\big).\end{array}$$

The following result was proved for $2\times2\times2$-tensors by Stegeman and Comon [33]. However, it holds for arbitrary $m_1,\ldots,m_p$. The proof is an immediate consequence of Theorem 5.2.

Corollary 8.3. *Let u be a tensor and u^* its best rank one approximation. Then $u - u^*$ is in the dual variety Y. In particular, if (8.5) holds then the hyperdeterminant of $u - u^*$ is zero.*

This result explains the fact, well known in the numerical multilinear algebra community, that tensor decomposition and best rank one approximation are unrelated for $p \geq 3$. The same argument gives the following generalization to arbitrary toric varieties X_A. Following [17], here A is a point configuration, whose convex hull is the polytope P in Corollary 5.11. Fix a projective toric variety $X_A \subset \mathbb{P}^n$ whose dual variety $(X_A)^*$ is a hypersurface. The defining polynomial of that hypersurface is the *A-discriminant* Δ_A. See [17] for details.

Corollary 8.4. *Given a general point $u \in \mathbb{R}^{n+1}$, let x be a point in the cone over X_A which is critical for the squared distance function d_u. The A-discriminant Δ_A vanishes at $u - x$.*

The construction of singular vectors and the ED degree formula in Theorem 8.1 generalizes to partially symmetric tensors. Corollary 8.4 continues to apply in this setting. We denote by $S^a\mathbb{R}^m$ the a-th symmetric power of $\mathbb{R}^m$. Fix positive integers $\omega_1, \ldots, \omega_p$. We consider the embedding of the Segre variety $X = \mathbb{P}(\mathbb{R}^{m_1}) \times \cdots \times \mathbb{P}(\mathbb{R}^{m_p})$ into the space of tensors $\mathbb{P}(S^{\omega_1}\mathbb{R}^{m_1} \otimes \cdots \otimes S^{\omega_p}\mathbb{R}^{m_p})$, sending $(v_1, \ldots, v_p)$ to $v_1^{\omega_1} \otimes \cdots \otimes v_p^{\omega_p}$. The image is called a *Segre-Veronese variety*. When $p = 1$ we get the classical *Veronese variety* whose points are symmetric decomposable tensors in $\mathbb{P}(S^{\omega_1}\mathbb{R}^{m_1})$. A symmetric tensor $x \in S^{\omega_1}\mathbb{R}^{m_1}$ corresponds to a homogeneous polynomial of degree ω_1 in m_1 indeterminates. Such a polynomial sits in the Veronese variety X if it can be expressed as the power of a linear form.

At this point, it is extremely important to note the correct choice of coordinates on the space $S^{\omega_1}\mathbb{R}^{m_1} \otimes \ldots \otimes S^{\omega_p}\mathbb{R}^{m_p}$. We want the group $SO(m_1) \times \cdots \times SO(m_p)$ to act by rotations on that space, and our Euclidean distance must be compatible with that action. In order for this to happen, we must include square roots of appropriate multinomial coefficients in the parametrization of the Segre-Veronese variety. We saw this Example 2.6 for the twisted cubic curve $(p = 1, m_1 = 2, \omega_1 = 3)$ and in Example 3.2 for symmetric matrices $(p = 1, \omega_2 = 3)$. In both examples, the Euclidean distances comes from the ambient space of all tensors.

Example 8.5. Let $p = 2, m_1 = 2, m_2 = 3, \omega_1 = 3, \omega_2 = 2$. The corresponding space $S^3\mathbb{R}^2 \otimes S^2\mathbb{R}^3$ of partially symmetric tensors has dimension 24. We regard this as a subspace in the 72-dimensional space of $2\times2\times2\times3\times3$-tensors. With this, the coordinates on $S^3\mathbb{R}^2 \otimes S^2\mathbb{R}^3$ are x_{ijklm} where $1 \leq i \leq j \leq k \leq 2$ and $1 \leq l \leq m \leq 3$, and the squared distance function is

$$\begin{aligned} d_u(x) \quad = \quad & (u_{11111} - x_{11111})^2 + 2(u_{11112} - x_{11112})^2 + \cdots + (u_{11133} - x_{11133})^2 \\ & + 3(u_{12111} - x_{12111})^2 + 6(u_{12112} - x_{12112})^2 + \cdots + (u_{22233} - x_{22233})^2. \end{aligned}$$

In the corresponding projective space $\mathbb{P}^{23} = \mathbb{P}(S^3\mathbb{R}^2 \otimes S^2\mathbb{R}^3)$, the threefold $X = \mathbb{P}^1 \times \mathbb{P}^2$ is embedded by the line bundle $\mathcal{O}(3,2)$. It is cut out by *scaled* binomial equations such as $3x_{11111}x_{22111} - x_{12111}x_{12111}$. The ED degree of this Segre-Veronese variety X equals 27. $\diamondsuit$

Theorem 8.6. ([13, Theorem 5]). *Let $X \subset \mathbb{P}(S^{\omega_1}\mathbb{C}^{m_1} \otimes \cdots \otimes S^{\omega_p}\mathbb{C}^{m_p})$ be the Segre-Veronese variety of partially symmetric tensors of rank one. In the invariant coordinates described above, the ED degree of X is the coefficient of the monomial $z_1^{m_1-1} \cdots z_p^{m_p-1}$ in the polynomial*

$$\prod_{i=1}^{p} \frac{(\widehat{z}_i)^{m_i} - z_i^{m_i}}{\widehat{z}_i - z_i} \qquad \textit{where} \quad \widehat{z}_i = \left(\textstyle\sum_{j=1}^{p} \omega_j z_j\right) - z_i.$$

The critical points of d_u on X are characterized by the singular vector equations (8.2), obtained by restricting from ordinary tensors to partially symmetric tensors. Of special interest is the case $p = 1$, with $m_1 = m$ and $\omega_1 = \omega$. Here X is the Veronese variety of symmetric $m \times m \times \cdots \times m$ tensors with ω factors that have rank one.

Corollary 8.7. *The Veronese variety $X \subset \mathbb{P}(S^\omega \mathbb{C}^m)$, with $SO(m)$ invariant coordinates, has*

$$\text{EDdegree}(X) \;=\; \frac{(\omega-1)^m - 1}{\omega - 2}.$$

This is the formula in [5] for the number of eigenvalues of a tensor. Indeed, for symmetric tensors, the eigenvector equations of [5] translate into (8.2). This is well-known in the matrix case ($\omega = 2$): computing eigenvalues and computing singular values is essentially equivalent. At present, we do not know how to extend our results to tensors of rank $r \geq 2$.

We now shift gears and examine the average ED degrees of rank one tensors. As above, we write X for the cone over the Segre variety, given by its distinguished embedding (8.1) into $\mathbb{R}^{m_1 m_2 \cdots m_p}$. We fix the standard Gaussian distribution ω centered at the origin in $\mathbb{R}^{m_1 m_2 \cdots m_p}$.

In [9] the average ED degree of X is expressed in terms of the average absolute value of the determinant on a Gaussian-type matrix ensemble constructed as follows. Set $m := \sum_i (m_i - 1)$ and let $A = (a_{k\ell})$ be the symmetric $m \times m$-matrix with $p \times p$-block division into blocks of sizes $m_1 - 1, \ldots, m_p - 1$ whose upper triangular entries $a_{k\ell}$, $1 \leq k \leq \ell \leq m$, are

$$a_{k\ell} = \begin{cases} U_{k\ell} & \text{if } k, \ell \text{ are from distinct blocks,} \\ U_0 & \text{if } k = \ell, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Here U_0 and the $U_{k\ell}$ with $k < \ell$ in distinct blocks are independent normally distributed scalar random variables. For instance, if $p = 3$ and $(m_1, m_2, m_3) = (2, 2, 3)$, then

$$A = \begin{bmatrix} U_0 & U_{12} & U_{13} & U_{14} \\ U_{12} & U_0 & U_{23} & U_{24} \\ U_{13} & U_{23} & U_0 & 0 \\ U_{14} & U_{24} & 0 & U_0 \end{bmatrix}$$

with $U_0, U_{12}, U_{13}, U_{14}, U_{23}, U_{24} \sim N(0,1)$ independent.

Theorem 8.8 ([9])**.** *The average ED degree of the Segre variety X relative to the standard Gaussian distribution on $\mathbb{R}^{m_1 m_2 \cdots m_p}$ equals*

$$\text{aEDdegree}(X) \;=\; \frac{\pi^{p/2}}{2^{m/2} \cdot \prod_{i=1}^{p} \Gamma\left(\frac{m_i}{2}\right)} \cdot \mathbb{E}(|\det(A)|),$$

where $\mathbb{E}(|\det(A)|)$ is the expected absolute determinant of the random matrix A.

The proof of this theorem, which can be seen as a first step in *random tensor theory*, is a computation similar to that in Example 4.7, though technically more difficult. Note the dramatic decrease in dimension: instead of sampling tensors u from an $m_1 \cdots m_p$-dimensional space and computing the critical points of d_u, the theorem allows us to compute the average ED degree by sampling $m \times m$-matrices and computing their determinants. Unlike in Example 4.7, we do not expect that there exists a closed form expression for $\mathbb{E}(|\det(A)|)$, but existing asymptotic results on the expected absolute determinant, e.g. from [35], should still help in comparing aEDdegree(X) with EDdegree(X) for large p. The following table from [9] gives some values for the average ED degree of X and compares them with Theorem 8.1:

Tensor format	aEDdegree	EDdegree
$n \times m$	$\min(n, m)$	$\min(n, m)$
$2^3 = 2 \times 2 \times 2$	4.2891...	6
2^4	11.0647...	24
2^5	31.5661...	120
2^6	98.8529...	720
2^7	333.6288...	5040
2^8	1205.4...	40320
2^9	46130.0...	362880
$2 \times 2 \times n,\ n \geq 3$	5.6038...	8
$2 \times 3 \times 3$	8.8402...	15
$2 \times 3 \times n,\ n \geq 4$	10.3725...	18
$3 \times 3 \times 3$	16.0196...	37
$3 \times 3 \times 4$	21.2651...	55
$3 \times 3 \times n,\ n \geq 5$	23.0552...	61

It is known from [13] that EDdegree(X) stabilizes outside the range (8.5), and we observed the same behavior experimentally for aEDdegree(X). Part of the ongoing work in [9] is to explain this behavior both geometrically and from the formula in Theorem 8.8.

Epilogue

We conclude our investigation of the Euclidean distance degree by loosely paraphrasing Hilbert and Cohn-Vossen in their famous book *Anschauliche Geometrie* [21, Chapter I, §1]:

> *The simplest curves are the planar curves. Among them, the simplest one is the line* (ED degree 1). *The next simplest curve is the circle* (ED degree 2). *After that come the parabola* (ED degree 3), *and, finally, general conics* (ED degree 4).

Acknowledgements.
Jan Draisma was supported by a Vidi grant from the Netherlands Organisation for Scientific Research (NWO), and Emil Horobeţ by the NWO Free Competition grant *Tensors of bounded rank*. Giorgio Ottaviani is member of GNSAGA-INDAM. Bernd Sturmfels was supported by the NSF (DMS-0968882), DARPA (HR0011-12-1-0011), and the Max-Planck Institute für Mathematik in Bonn, Germany. Rekha Thomas was supported by the NSF (DMS-1115293).

References

[1] C. Aholt, B. Sturmfels and R. Thomas: *A Hilbert scheme in computer vision*, Canadian J. Mathematics **65** (2013), no. 5, 961–988.

[2] B. Anderson and U. Helmke: *Counting critical formations on a line*, preprint, 2013.

[3] D. Bates, J. Hauenstein, A. Sommese, and C. Wampler: *Numerically Solving Polynomial Systems with Bertini*, SIAM, 2013.

[4] H.-C.G. von Bothmer and K. Ranestad: *A general formula for the algebraic degree in semidefinite programming*, Bull. London Math. Soc. **41** (2009) 193–197.

[5] D. Cartwright and B. Sturmfels: *The number of eigenvectors of a tensor*, Linear Algebra and its Applications **438** (2013) 942–952.

[6] F. Catanese and C. Trifogli: *Focal loci of algebraic varieties I*, Communications in Algebra **28** (2000) 6017–6057.

[7] M. Chu, R. Funderlic, and R. Plemmons: *Structured low rank approximation*, Linear Algebra Appl. **366** (2003) 157–172.

[8] D. Cox, J. Little, and D. O'Shea: *Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1992.

[9] J. Draisma and E. Horobeţ: *The average number of critical rank-one approximations to a tensor*, manuscript in preparation.

[10] J. Draisma and J. Rodriguez: *Maximum likelihood duality for determinantal varieties*, Int. Math. Res. Not., to appear, `arXiv:1211.3196`.

[11] M. Drton, B. Sturmfels and S. Sullivant: *Lectures on Algebraic Statistics*, Oberwolfach Seminars, Vol 39, Birkhäuser, Basel, 2009.

[12] S. Friedland: *Best rank one approximation of real symmetric tensors can be chosen symmetric*, Front. Math. China **8**(1) (2013) 19–40.

[13] S. Friedland and G. Ottaviani: *The number of singular vector tuples and uniqueness of best rank one approximation of tensors*, `arXiv:1210.8316`.

[14] J.-C. Faugère, M. Safey El Din and P.-J. Spaenlehauer: *Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): algorithms and complexity*, J. Symbolic Comput. **46** (2011) 406–437.

[15] W. Fulton: *Introduction to Toric Varieties*, Princeton University Press, 1993.

[16] W. Fulton: *Intersection Theory*, Springer, Berlin, 1998.

[17] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky: *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[18] D. Grayson and M. Stillman: *Macaulay2, a software system for research in algebraic geometry*, available at `www.math.uiuc.edu/Macaulay2/`.

[19] R. Hartley and P. Sturm: *Triangulation*, Computer Vision and Image Understanding: CIUV (1997) 68(2): 146–157.

[20] R. Hartshorne: *Algebraic Geometry*, Graduate Texts in Mathematics **52**, Springer-Verlag, New York, 1977.

[21] D. Hilbert and S. Cohn-Vossen: *Anschauliche Geometrie*, Springer-Verlag, Berlin, 1932.

[22] A. Holme: *The geometric and numerical properties of duality in projective algebraic geometry*, Manuscripta Math. **61** (1988) 145–162.

[23] S. Hoşten, A. Khetan, and B. Sturmfels: *Solving the likelihood equations*, Foundations of Computational Mathematics **5** (2005) 389–407.

[24] J. Huh and B. Sturmfels: *Likelihood geometry*, `arXiv:1305.7462`.

[25] M. Laurent: *Cuts, matrix completions and graph rigidity*, Mathematical Programming **79** (1997) 255–283.

[26] L.-H. Lim: *Singular values and eigenvalues of tensors: a variational approach,* Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 05), 1 (2005), 129-132.

[27] E. Miller and B. Sturmfels: *Combinatorial Commutative Algebra*, Graduate Texts in Mathematics **227**, Springer, New York, 2004.

[28] The Online Encyclopedia of Integer Sequences, `http://oeis.org/`.

[29] P.A. Parrilo: *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, PhD Thesis, Caltech, Pasadena, CA, May 2000.

[30] R. Piene: *Polar classes of singular varieties*, Ann. Sci. École Norm. Sup. (4) **11** (1978), no. 2, 247–276.

[31] P. Rostalski and B. Sturmfels: *Dualities*, Chapter 5 in G. Blekherman, P. Parrilo and R. Thomas: *Semidefinite Optimization and Convex Algebraic Geometry*, pp. 203–250, MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2013.

[32] G. Salmon: *A Treatise on the Higher Plane Curves*, Dublin, 1879, available on the web at `http://archive.org/details/117724690`.

[33] A. Stegeman and P. Comon: *Subtracting a best rank-*1 *approximation does not necessarily decrease tensor rank*, Linear Algebra Appl. **433** (2010) 1276–1300.

[34] H. Stewénius, F. Schaffalitzky, and D. Nistér: *How hard is 3-view triangulation really?*, Proc. *International Conference on Computer Vision*, Beijing, China (2005) 686–693.

[35] T. Tao and V. Vu: *A central limit theorem for the determinant of a Wigner matrix*, Adv. Math. **231** (2012) 74–101.

[36] J. Thomassen, P. Johansen, and T. Dokken: *Closest points, moving surfaces, and algebraic geometry*, Mathematical methods for curves and surfaces: Tromsø, 2004, 351–362, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2005

[37] C. Trifogli: *Focal loci of algebraic hypersurfaces: a general theory*, Geometriae Dedicata **70** (1998) 1–26.

[38] J. Weyman: *Cohomology of Vector Bundles and Syzygies*, Cambridge Tracts in Mathematics, **14**, Cambridge University Press, Cambridge, 2003.

Authors' addresses:
Jan Draisma, TU Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, `j.draisma@tue.nl`;
and Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands
Emil Horobeţ, TU Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, `e.horobet@tue.nl`
Giorgio Ottaviani, Università di Firenze, viale Morgagni 67A, 50134 Firenze, Italy, `ottavian@math.unifi.it`
Bernd Sturmfels, University of California, Berkeley, CA 94720-3840, USA, `bernd@math.berkeley.edu`
Rekha Thomas, University of Washington, Box 354350, Seattle, WA 98195-4350, USA, `rrthomas@uw.edu`