THE EUCLIDEAN DISTANCE DEGREE OF ORTHOGONALLY INVARIANT MATRIX VARIETIES

DMITRIY DRUSVYATSKIY, HON-LEUNG LEE, GIORGIO OTTAVIANI, AND REKHA R. THOMAS

Abstract. The Euclidean distance degree of a real variety is an important invariant arising in distance minimization problems. We show that the Euclidean distance degree of an orthogonally invariant matrix variety equals the Euclidean distance degree of its restriction to diagonal matrices. We illustrate how this result can greatly simplify calculations in concrete circumstances.

1. Introduction

The problem of minimizing the Euclidean distance (ED) of an observed data point $y \in \mathbb{R}^n$ to a real algebraic variety $\mathcal{V} \subseteq \mathbb{R}^n$ arises frequently in applications, and amounts to solving the polynomial optimization problem

$$\text{minimize} \quad \sum_{i=1}^{n}(y_i - x_i)^2 \quad \text{subject to} \quad x \in \mathcal{V}.$$

The algebraic complexity of this problem is closely related to the number of complex regular critical points of y on the Zariski closure $\mathcal{V}_{\mathbb{C}}$ of $\mathcal{V}$. We will call such points the *ED critical points* of y with respect to $\mathcal{V}$; see Definition 4.1. The authors of [11] showed that the number of ED critical points of a general data point $y \in \mathbb{C}^n$ is a constant, and hence is an invariant of $\mathcal{V}$. This number is called the *Euclidean distance degree* (ED degree) of $\mathcal{V}$. As noted in [11], the computation of EDdegree($\mathcal{V}$) can be subtle, since it may change considerably under a linear transformation of $\mathcal{V}$.

In this work, we explore the ED degree of *orthogonally invariant* matrix varieties $\mathcal{M} \subseteq \mathbb{R}^{n\times t}$, meaning those varieties $\mathcal{M}$ satisfying

$$U\mathcal{M}V^\top = \mathcal{M} \quad \text{for all real orthogonal matrices} \quad U \in O(n), V \in O(t).$$

Without loss of generality, suppose $n \leq t$. Clearly, membership of a matrix M in such a variety $\mathcal{M}$ is fully determined by its vector of singular values $\sigma(M) = (\sigma_1(M), \ldots, \sigma_n(M))$, where we use the convention $\sigma_{i-1}(M) \geq \sigma_i(M)$ for each i. Indeed, we may associate with any orthogonally invariant matrix variety $\mathcal{M}$ its diagonal restriction $S = \{x : \text{Diag}\,(x) \in \mathcal{M}\}$. The variety S thus defined is *absolutely symmetric* (invariant under signed permutations) and satisfies the key relation $\mathcal{M} = \sigma^{-1}(S)$. Conversely, any absolutely symmetric set $S \subseteq \mathbb{R}^n$ yields the orthogonally invariant matrix variety $\sigma^{-1}(S)$; see e.g. [14, Theorem 3.4] and [7, Proposition 1.1].

Date: March 3, 2016.

Drusvyatskiy was partially supported by the AFOSR YIP award FA9550-15-1-0237. Lee and Thomas were partially supported by the NSF grant DMS-1418728.

In this paper, we prove the elegant formula

$$\boxed{\text{EDdegree}(\mathcal{M}) = \text{EDdegree}(S).} \tag{$\star$}$$

In most interesting situations, the diagonal restriction $S \subseteq \mathbb{R}^n$ has simple geometry, as opposed to the matrix variety $\mathcal{M}$, and hence our main result ($\star$) provides elementary and transparent means to compute the ED degree of $\mathcal{M}$ by working with the simpler object S. Interesting consequences flow from there. For example, consider the *r-th rank variety*

$$\mathbb{R}^{n\times t}_r := \{X \in \mathbb{R}^{n\times t} : \operatorname{rank} X \leq r\},$$

and the *essential variety*

$$\mathcal{E} := \{X \in \mathbb{R}^{3\times 3} : \sigma_1(X) = \sigma_2(X),\, \sigma_3(X) = 0\}$$

from computer vision [1, 17, 21]; both $\mathbb{R}^{n\times t}_r$ and $\mathcal{E}$ are orthogonally invariant. The diagonal restrictions of $\mathbb{R}^{n\times t}$ and $\mathcal{E}$ are finite unions of linear subspaces and their ED degrees are trivial to compute. Moreover, our results readily imply that all ED critical points of a general real data matrix Y on $\mathbb{R}^{n\times t}_r$ and on $\mathcal{E}$ are real. This result has been previously shown for $\mathbb{R}^{n\times t}_r$ in [11] – a generalization of the Eckart-Young theorem – and is entirely new for the essential variety $\mathcal{E}$. A related further investigation of the essential variety appears in [15].

Our investigation of orthogonally invariant matrix varieties fits in a broader scope. The idea of studying orthogonally invariant matrix sets $\mathcal{M}$ via their diagonal restrictions S – the theme of our paper – is not new, and goes back at least to von Neumann's theorem on unitarily invariant matrix norms [28]. In recent years, it has become clear that various analytic properties of $\mathcal{M}$ and S are in one-to-one correspondence, and this philosophy is sometimes called the "transfer principle"; see for instance, [7]. For example, $\mathcal{M}$ is C^p-smooth around a matrix X if and only if S is C^p-smooth around $\sigma(X)$ [5, 12, 20, 25, 27]. Other properties, such as convexity [8], positive reach [6], partial smoothness [5], and Whitney conditions [13] follow the same paradigm. In this sense, our paper explores the transfer principle for the ED degree of algebraic varieties. To the best of our knowledge, this is the first result in this body of work that is rooted in algebraic geometry.

Though our main result ($\star$) is easy to state, the proof is subtle; moreover, the result itself is surprising in light of the discussion in [14, Section 5]. The outline of the paper is as follows. In Section 2 we investigate invariance properties of the Zariski closure $\mathcal{M}_{\mathbb{C}} \subseteq \mathbb{C}^{n\times t}$ of an orthogonally invariant matrix variety $\mathcal{M} \subseteq \mathbb{R}^{n\times t}$, as well as the correspondence between irreducible components of S and those of $\mathcal{M}$. In Section 3, we discuss "algebraic singular value decompositions" for general matrices $Y \in \mathbb{C}^{n\times t}$, leading to Section 4 containing our main results. When S is a subspace arrangement, our results yield particularly nice consequences generalizing several classical facts in matrix theory – the content of Section 5.

2. Zariski closure, irreducibility, and dimension of matrix varieties

Setting the stage, we begin with some standard notation. For the fields $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, the symbol $\mathbb{F}[x] = \mathbb{F}[x_1, \ldots, x_n]$ will denote the ring of polynomials in $x_1, \ldots, x_n$ with coefficients in $\mathbb{F}$. Given polynomials $f_1, \ldots, f_s \in \mathbb{F}[x]$ the set $\mathcal{V} := \{x \in \mathbb{F}^n \ : \ f_1(x) = \cdots = f_s(x) = 0\}$ is called an (algebraic) *variety* over $\mathbb{F}$. The *Zariski closure* of an arbitrary set T in $\mathbb{C}^n$, denoted $\overline{T}$, is the smallest variety over $\mathbb{C}$ containing T. Unless otherwise specified, the topology on $\mathbb{C}^n$ is fixed to be

the *Zariski topology*, obtained by defining the closed sets to be the varieties over $\mathbb{C}$. The topology on any subset of $\mathbb{C}^n$ will then always be the one induced by the Zariski topology.

Consider a real algebraic variety $\mathcal{V} \subseteq \mathbb{R}^n$. The *vanishing ideal* of $\mathcal{V}$ is defined to be $I(\mathcal{V}) := \{f \in \mathbb{R}[x] : f(x) = 0 \text{ for all } x \in \mathcal{V}\}$. Viewing $\mathcal{V}$ as a subset of $\mathbb{C}^n$, the Zariski closure of $\mathcal{V}$, denoted $\mathcal{V}_{\mathbb{C}}$, can be written as $\{x \in \mathbb{C}^n : f(x) = 0 \text{ for all } f \in I(\mathcal{V})\}$; see e.g. [29]. Note this notation is slightly redundant since by definition we have $\overline{\mathcal{V}} = \mathcal{V}_{\mathbb{C}}$. Nonetheless, we prefer to keep both symbols to ease notation when appropriate.

2.1. **Invariance under closure.** Consider a group G acting linearly on $\mathbb{C}^n$. Then G acts on the right on $\mathbb{C}[x]$ via

$$g \cdot f(x) = (x \mapsto f(g \cdot x)) \qquad \text{for any } g \in G,\ f \in \mathbb{C}[x].$$

A subset $T \subseteq \mathbb{C}^n$ is *G-invariant* if $g \cdot x$ lies in T for any $g \in G$ and $x \in T$. A polynomial $f \in \mathbb{C}[x]$ is *G-invariant* provided $g \cdot f = f$ for all $g \in G$. We begin with the following elementary result.

Lemma 2.1. *If a set $T \subseteq \mathbb{C}^n$ is G-invariant, then its closure $\overline{T}$ is also G-invariant.*

Proof. Fixing $g \in G$, the map $\mu_g : \mathbb{C}^n \to \mathbb{C}^n$ given by $\mu_g(x) = g \cdot x$ is a linear isomorphism. Hence, assuming T is G-invariant, we deduce $\mu_g(\overline{T}) = \overline{\mu_g(T)} = \overline{T}$, as claimed. □

We now specialize the discussion to the main setting of the paper. For a positive integer s, the symbol $O(s)$ will denote the set of all $s \times s$ real orthogonal matrices. This is both a group and a real variety and its Zariski closure $O_{\mathbb{C}}(s)$ is the set of all $s \times s$ complex orthogonal matrices — those satisfying $Q^\top Q = QQ^\top = I$. Henceforth, we fix two positive integers n and t with $n \leq t$, and consider the groups $O(n) \times O(t)$ and $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$, along with the group $\Pi_n^{\pm}$ of all signed permutations of $\{1, \ldots, n\}$. Recall that we always consider the action of $O(n) \times O(t)$ on $\mathbb{R}^{n \times t}$ and the action of $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$ on $\mathbb{C}^{n \times t}$ by conjugation $(U, V) \cdot X = UXV^\top$.

Now suppose $\mathcal{M} \subseteq \mathbb{R}^{n \times t}$ is a $O(n) \times O(t)$-invariant (orthogonally invariant) matrix variety. Then Lemma 2.1 shows that $\mathcal{M}_{\mathbb{C}}$ is $O(n) \times O(t)$-invariant. We now prove the stronger statement: $\mathcal{M}_{\mathbb{C}}$ is invariant under the larger group $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$.

Proposition 2.2 (Closure invariance)**.** *A matrix variety $\mathcal{M} \subseteq \mathbb{R}^{n \times t}$ is $O(n) \times O(t)$-invariant if and only if $\mathcal{M}_{\mathbb{C}}$ is $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$-invariant. Similarly, a variety $S \subseteq \mathbb{R}^n$ is $\Pi_n^{\pm}$-invariant if and only if $S_{\mathbb{C}}$ is $\Pi_n^{\pm}$-invariant.*

Proof. Since the proofs are similar, we only prove the first claim. Suppose first $\mathcal{M}_{\mathbb{C}}$ is $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$-invariant. Then for every $X \in \mathcal{M}$ and $(U, V) \in O(n) \times O(t)$, the matrix UXV^T lies in $\mathcal{M}_{\mathbb{C}} \cap \mathbb{R}^{n \times t} = \mathcal{M}$. Therefore $\mathcal{M}$ is $O(n) \times O(t)$-invariant, as claimed. Suppose conversely that $\mathcal{M}$ is $O(n) \times O(t)$-invariant. Let $X \in \mathcal{M}_{\mathbb{C}}$ be fixed. Then the map $\gamma_X : O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t) \to \mathbb{C}^{n \times t}$ defined by $\gamma(g) = g \cdot X$ is continuous. Lemma 2.1 yields the inclusion $O(n) \times O(t) \subseteq \gamma_X^{-1}(\mathcal{M}_{\mathbb{C}})$. Since $\gamma_X^{-1}(\mathcal{M}_{\mathbb{C}})$ is closed by continuity, we conclude $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t) \subseteq \gamma_X^{-1}(\mathcal{M}_{\mathbb{C}})$. This completes the proof. □

2.2. **Irreducible components of orthogonally invariant varieties.** For the rest of the section, fix a $\Pi_n^{\pm}$-invariant (absolutely symmetric) variety S in $\mathbb{R}^n$. Then the $O(n) \times O(t)$-invariant matrix set $\mathcal{M} := \sigma^{-1}(S)$ is a real variety in $\mathbb{R}^{n \times t}$; see [14, Theorem 3.4] or [7, Proposition 1.1]. Moreover, the diagonal restriction $\{x \in \mathbb{R}^n : \mathrm{Diag}\,(x) \in \mathcal{M}\}$ coincides with S. Here, we call a $n \times t$ matrix D diagonal if and only if $D_{ij} = 0$ whenever $i \neq j$, and for any vector $x \in \mathbb{R}^n$ the symbol $\mathrm{Diag}\,(x)$ denotes the diagonal matrix with $D_{ii} = x_i$ for each $i = 1, \ldots, n$.

In this section, we highlight the correspondence between the irreducible components of S and those of $\mathcal{M}$. Recall that a real or complex variety $\mathcal{V}$ is *irreducible* if it cannot be written as a union $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ of two proper subvarieties $\mathcal{V}_1$ and $\mathcal{V}_2$. Any variety $\mathcal{V}$ can be written as a union of finitely many irreducible subvarieties $\mathcal{V}_i$ satisfying $\mathcal{V}_i \nsubseteq \mathcal{V}_j$ for distinct indices i and j. The varieties $\mathcal{V}_i$ are called the *irreducible components* of $\mathcal{V}$, and are uniquely defined up to indexing.

Coming back to the aim of this section, let $\{S_i\}_{i=1}^k$ be the irreducible components of S. The varieties S_i are typically not absolutely symmetric. Hence we define their symmetrizations $S_i^{\pi} := \bigcup_{\pi \in \Pi_n^{\pm}} \pi S_i$ and the real varieties $\mathcal{M}_i := \sigma^{-1}(S_i^{\pi})$. It is standard that a signed permutation maps an irreducible component of S to another irreducible component of S.

We record the following elementary observation for ease of reference.

Lemma 2.3. *For any pair of indices i, j, the following implications hold:*

$$\begin{aligned} S_i^{\pi} \subseteq S_j^{\pi} \quad &\Longrightarrow \quad S_i^{\pi} = S_j^{\pi}, \text{ and} \\ \mathcal{M}_i \subseteq \mathcal{M}_j \quad &\Longrightarrow \quad \mathcal{M}_i = \mathcal{M}_j \end{aligned}$$

Proof. If $S_i^{\pi} \subseteq S_j^{\pi}$, then we deduce that $S_i = \bigcup_{\pi \in \Pi_n^{\pm}} (S_i \cap \pi S_j)$. Hence for some $\pi \in \Pi_n^{\pm}$, the inclusion $S_i \subseteq \pi S_j$ holds. Since both S_i and πS_j are irreducible components of S, it must be that $S_i = \pi S_j$ and hence, $S_i^{\pi} = S_j^{\pi}$, as claimed. The second implication follows immediately. □

For any $U \in O(n)$ and $V \in O(t)$, the map $X \mapsto UXV^{\top}$ is an automorphism of $\mathcal{M}_i$, and therefore maps an irreducible component of $\mathcal{M}_i$ to another irreducible component of $\mathcal{M}_i$. We now show that this action is transitive, just as the action of $\Pi_n^{\pm}$ on the components of S_i^{π}.

Lemma 2.4. *For any index i, the group $O(n) \times O(t)$ acts transitively on the irreducible components of $\mathcal{M}_i$. Consequently, the real variety $\mathcal{M}_i$ is equidimensional.*

Proof. Let $\mathcal{H}$ be an irreducible component of $\mathcal{M}_i$. Note that the set $\Gamma := \bigcup \{U\mathcal{H}V^{\top} : U \in O(n), V \in O(t)\}$ is a union of a nonempty collection of irreducible components of $\mathcal{M}_i$. Let Z be the union of the irreducible components of $\mathcal{M}_i$ not contained in Γ (if any). Observe that Z is an orthogonally invariant variety. Hence the two absolutely symmetric varieties $\{x : \mathrm{Diag}\,(x) \in \Gamma\}$ and $\{x : \mathrm{Diag}\,(x) \in Z\}$ cover S_i^{π}. Since S_i is irreducible, either Γ or Z coincides with all of $\mathcal{M}_i$. Since the latter is impossible by construction, we conclude that $\Gamma = \mathcal{M}_i$, as claimed. □

We end with the following theorem, which will play a key role in the proof of Proposition 4.10, leading to the main result of the paper.

Proposition 2.5. *Each variety $\mathcal{M}_i$ is a union of some irreducible components of $\mathcal{M}$.*

Proof. Let $\{C_l\}$ be the set of irreducible components of $\mathcal{M}$. Then for any index l, $C_l = \bigcup_j (C_l \cap \mathcal{M}_j)$ which implies that C_l is contained in $\mathcal{M}_j$ for some index $j(l)$.

Fix an $\mathcal{M}_i$ and let $\mathcal{H}$ be an irreducible component of $\mathcal{M}_i$. From the equality $\mathcal{H} = \bigcup_l (C_l \cap \mathcal{H})$ we conclude that the inclusion $\mathcal{H} \subseteq C_l$ holds for some index l. This implies that $\mathcal{H} \subseteq C_l \subseteq \mathcal{M}_{j(l)}$, and hence by Lemma 2.4, $\mathcal{M}_i \subseteq \mathcal{M}_{j(l)}$. Lemma 2.3 then implies the equality $\mathcal{M}_i = \mathcal{M}_{j(l)}$, yielding $\mathcal{H} \subseteq C_l \subseteq \mathcal{M}_i$. Taking the union of this inclusion over all the irreducible components $\mathcal{H}$ of $\mathcal{M}_i$ and the corresponding C_l, we deduce that $\mathcal{M}_i$ is a union of some irreducible components of $\mathcal{M}$. □

Since closures of irreducible components of a real variety $\mathcal{V}$ are the irreducible components of $\mathcal{V}_{\mathbb{C}}$, Proposition 2.5 immediately implies that $\overline{\mathcal{M}_i}$ is a union of some irreducible components of $\mathcal{M}_{\mathbb{C}}$.

2.3. **Dimension of orthogonally invariant varieties.** We next show how to read off the dimension of $\mathcal{M}_{\mathbb{C}}$ from the absolutely symmetric variety $S \subseteq \mathbb{R}^n$. To this end, note first that since the equality $\dim(\mathcal{M}_{\mathbb{C}}) = \dim(\mathcal{M})$ holds (see [29, Lemma 8]), it suffices to compute the dimension of the real variety $\mathcal{M}$ from S. We will assume that $\Pi_n^{\pm}$ acts transitively on the irreducible components of S, that is in the notation of Section 2.2 we have $S_i^{\pi} = S$ for all indices i. If this is not the case, we can treat each set S_i^{π} separately. With this simplification, both varieties S and $\mathcal{M}$ are equidimensional (Lemma 2.4).

The following recipe follows that in [5, Section 2.3] and [7] and hence we skip some of the explanations. The basic idea is to understand the dimension of the fiber $\sigma^{-1}(x^*)$ where $x^* \in S$ is (carefully) chosen so that the sum of the dimension of the fiber and the dimension of S equals $\dim(\mathcal{M})$.

Fixing notation, consider the convex cone

$$\mathbb{R}^n_{+,\geq} := \{x \in \mathbb{R}^n : x_1 \geq x_2 \geq \ldots \geq x_n \geq 0\}.$$

Observe that $\mathbb{R}^n_{+,\geq}$ is exactly the range of σ on $\mathbb{R}^{n\times t}$. Along with a point $x \in \mathbb{R}^n_{+,\geq}$, we associate the partition $\mathcal{P}_x = \{P_1, \ldots, P_{\rho_x}, P_0\}$ of the index set $\{1, \ldots, n\}$ so that $x_i = x_j$ if and only if $i, j \in P_l$, and $x_i > x_j$ for any $i \in P_q$ and $j \in P_r$ with $q > r$. We assume that P_0 contains the indices of the zero coordinates in x, and we define $p_l := |P_l|$. It could be that $p_0 = 0$ for a given x. On the other hand, we have $p_l > 0$ for all $l = 1, \ldots, \rho_x$. Recall the equality

$$\sigma^{-1}(x) = \{U \operatorname{Diag}(x) V^\top \,:\, U \in O(n), V \in O(t)\}.$$

Let

$$(O(n) \times O(t))_x := \{(U, V) \in O(n) \times O(t) \,:\, \operatorname{Diag}(x) = U \operatorname{Diag}(x) V^\top\}$$

denote the stabilizer of $\operatorname{Diag}(x)$, under the action of $O(n)\times O(t)$. Then one can check that (U, V) lies in the stabilizer $(O(n)\times O(t))_x$ if and only if U is block diagonal with blocks $U_i \in O(p_i)$ for $i = 0, \ldots, \rho_x$ and V is block diagonal with blocks $V_i \in O(p_i)$ for $i = 1, \ldots, \rho_x$, and a block $V_0 \in O(p_0 + (t - n))$. Further, $U_i V_i^\top = I$ for all $i = 1, \ldots, \rho_x$ which means that the U_i's determine the corresponding V_i's for all i except $i = 0$. This implies that the dimension of $(O(n) \times O(t))_x$ is

$$\dim((O(n) \times O(t))_x) = \sum_{l=0}^{\rho_x} \frac{p_l(p_l - 1)}{2} + \frac{(p_0 + t - n)(p_0 + t - n - 1)}{2}$$

yielding

$$\begin{aligned}\dim(\sigma^{-1}(x)) &= \dim(O(n)\times O(t)) - \dim((O(n)\times O(t))_x)\\ &= \frac{n(n-1)+t(t-1)}{2} - \sum_{l=0}^{\rho_x}\frac{p_l(p_l-1)}{2} - \frac{(p_0+t-n)(p_0+t-n-1)}{2}\\ &= \sum_{0\le i<j\le \rho_x} p_ip_j + \frac{t(t-1)}{2} - \frac{(p_0+t-n)(p_0+t-n-1)}{2}. \qquad (2.1)\end{aligned}$$

Here we used the observation

$$\frac{n(n-1)}{2} - \sum_{l=0}^{\rho_x}\frac{p_l(p_l-1)}{2} = \binom{\sum_{l=0}^{\rho_x} p_l}{2} - \sum_{l=0}^{\rho_x}\binom{p_l}{2} = \sum_{0\le i<j\le\rho_x} p_ip_j.$$

For a partition $\mathcal{P}$ of $[n]$, define the set $\Delta_{\mathcal{P}} := \{x \in \mathbb{R}^n_{+,\geq} : \mathcal{P}_x = \mathcal{P}\}$. The set of all such Δ's defines an affine stratification of $\mathbb{R}^n_{+,\geq}$. Let $\mathcal{P}_*$ correspond to a stratum Δ_* in this stratification satisfying $S\cap\Delta_* \neq \emptyset$ and having maximal dimension among all strata that have a nonempty intersection with S. Then for any point $x^* \in S\cap\Delta_*$, we can choose a sufficiently small $\delta > 0$ satisfying $S \cap B_\delta(x^*) \subseteq \Delta_*$. Hence the fibers $\sigma^{-1}(x)$ have the same dimension for all $x \in S\cap B_\delta(x^*)$ and the preimage $\sigma^{-1}(S\cap B_\delta(x^*))$ is an open (in the Euclidean topology) subset of $\mathcal{M}$. Taking into account that both S and $\mathcal{M}$ are equidimensional, we deduce

$$\dim(\sigma^{-1}(S)) = \dim(S) + \dim(\sigma^{-1}(x^*)).$$

Appealing to (2.1), we arrive at the formula

(2.2)

$$\dim(\mathcal{M}) = \dim(S) + \left(\sum_{0\le i<j\le\rho^*} p_i^*p_j^*\right) + \frac{t(t-1)}{2} - \frac{(p_0^*+t-n)(p_0^*+t-n-1)}{2}.$$

Example 2.6 (Rank variety)**.** Recall the rank variety $\mathbb{R}^{n\times t}_r$ of matrices of rank at most r. In this case, S is the union of all coordinate planes in $\mathbb{R}^n$ of dimension r and $S_{\mathbb{C}}$ is the set of all r-dimensional coordinate planes in $\mathbb{C}^n$. Also, $\mathcal{M}_{\mathbb{C}} = \mathbb{C}^{n\times t}_r$, the set of all matrices in $\mathbb{C}^{n\times t}$ of rank at most r.

Note that S is equidimensional. Then along with a point x^* we have $p_0^* = n-r$ and $p_i^* = 1$ for all $i = 1,\dots,r$. Applying (2.2) we get that the dimension of $\mathbb{C}^{n\times t}_r$ is

$$r + \left(\binom{r}{2} + r(n-r)\right) + \frac{t(t-1)}{2} - \frac{(t-r)(t-r-1)}{2} = r(t+n-r).$$

Example 2.7 (Essential variety)**.** The essential variety is $\mathcal{E} = \{E \in \mathbb{R}^{3\times3} : \sigma_1(E) = \sigma_2(E),\ \sigma_3(E) = 0\}$. Its Zariski closure $\mathcal{E}_{\mathbb{C}} \subseteq \mathbb{C}^{3\times3}$ is known to be irreducible and of dimension six [9]. In this case, $S \subseteq \mathbb{R}^3$ consists of the six lines defined by $x_1 = \pm x_2$, $x_1 = \pm x_3$ and $x_2 = \pm x_3$ with the remaining coordinate set to zero in each case.

We can verify $\dim(\mathcal{E}_{\mathbb{C}}) = 6$ using (2.2). Indeed, picking a general point x^* on the line $x_1 = x_2$ in $\mathbb{R}^3_+$, we see that $\mathcal{P}_{x^*}$ has $p_0^* = 1$ and $p_1^* = 2$. Now applying the formula (2.2) we get $\dim(\mathcal{E}_{\mathbb{C}}) = 1 + 1\cdot2 + 3 - 0 = 6$.

3. Algebraic Singular Value Decompositions and GIT quotients

In this section we fix a $\Pi_n^{\pm}$-invariant variety $S \subseteq \mathbb{R}^n$ and the induced $O(n)\times O(t)$-invariant matrix variety $\mathcal{M} := \sigma^{-1}(S)$. The description of $\mathcal{M}$ as the preimage

$\sigma^{-1}(S)$ is not convenient when seeking to understand the algebraic geometric correspondences between $\mathcal{M}$ and S, since σ is not a polynomial map. Instead, we may equivalently write

$$\mathcal{M} = \{U \operatorname{Diag}(x) V^\top \ : \ U \in O(n),\ V \in O(t),\ x \in S\}. \tag{3.1}$$

In this notation, it is clear that $\mathcal{M}$ is obtained from S by an algebraic group action – a description that is more amenable to an algebraic analysis. Naturally then to understand geometric correspondences between the closures $\mathcal{M}_\mathbb{C}$ and $S_\mathbb{C}$, we search for a description analogous to (3.1), with $\mathcal{M}$, S, $O(n)$, and $O(t)$ replaced by their Zariski closures $\mathcal{M}_\mathbb{C}$, $S_\mathbb{C}$, $O_\mathbb{C}(n)$, and $O_\mathbb{C}(t)$. The difficulty is that an exact equality analogous to (3.1) usually fails to hold; instead, equality holds only in a certain generic sense that is sufficient for our purposes. We now make this precise.

3.1. **Algebraic SVD.** Our strategy revolves around an "algebraic singular value decomposition", a notion to be made precise shortly. Note that the common extension of a singular value decomposition (SVD) from real to complex matrices using unitary matrices, their conjugates, and the Hermitian metric does not fit well in the algebraic setting because unitary matrices form a real (but not a complex) variety and conjugation is not an algebraic operation. In particular, it is not suitable for studying the EDdegree of a matrix variety. Hence we will need an algebraic analog of SVD that uses complex orthogonal matrices. For a recent geometric treatment of SVD rooted in algebraic geometry see the survey [23].

Definition 3.1 (Algebraic SVD)**.** We say that a matrix $A \in \mathbb{C}^{n\times t}$ admits an *algebraic SVD* if it can be factored as $A = UDV^\top$ for some orthogonal matrices $U \in O_\mathbb{C}(n)$ and $V \in O_\mathbb{C}(t)$, and a complex diagonal matrix $D \in \mathbb{C}^{n\times t}$.

Not all matrices admit an algebraic SVD; indeed, this is the main obstruction to an equality analogous to (3.1) in which the varieties $\mathcal{M}$, S, $O(n)$, and $O(t)$ are replaced by their closures. A simple example is the matrix $A = \left(\begin{smallmatrix} 1 & i \\ 0 & 0 \end{smallmatrix}\right)$, with $i = \sqrt{-1}$. Indeed, in light of the equality $AA^\top = 0$, if it were possible to write $A = UDV^\top$ for some $U, V \in O_\mathbb{C}(2)$ and a diagonal matrix D, then we would deduce that $UDD^\top U^\top = 0$ which implies that $A = 0$, a contradiction. Fortunately, the existence question has been completely answered by Choudury and Horn [4, Theorem 2 & Corollary 3].

Theorem 3.2 (Existence of an algebraic SVD)**.** *A matrix $A \in \mathbb{C}^{n\times t}$ admits an algebraic SVD, if and only if, $AA^\top$ is diagonalizable and* $\operatorname{rank}(A) = \operatorname{rank}(AA^\top)$.

Suppose A admits an algebraic SVD $A = U \operatorname{Diag}(d) V^\top$ for some orthogonal matrices $U \in O_\mathbb{C}(n)$ and $V \in O_\mathbb{C}(t)$, and a vector $d \in \mathbb{C}^n$. Then the numbers d_i^2 are eigenvalues of $A^\top A$ and $AA^\top$, and the columns of U are eigenvectors of $AA^\top$ and the columns of V are eigenvectors of $A^\top A$, arranged in the same order as d_i. We call the complex numbers d_i the *algebraic singular values* of A. They are determined up to sign.

We record the following immediate consequence of Theorem 3.2 for ease of reference.

Corollary 3.3. *Consider a matrix $A \in \mathbb{C}^{n\times t}$. If the eigenvalues of $AA^\top$ are nonzero and distinct, then A has an algebraic SVD.*

Suppose $\mathcal{V}$ is a variety over $\mathbb{R}$ or $\mathbb{C}$. We say that a property holds for a *general point* $x \in \mathcal{V}$ if the set of points $x \in \mathcal{V}$ for which the property holds contains an open

dense subset of $\mathcal{V}$ (in Zariski topology). In this terminology, Theorem 3.2 implies that general complex matrices $A \in \mathbb{C}^{n\times t}$ do admit an algebraic SVD.

We can now prove the main result of this section (cf. equation (3.1)).

Theorem 3.4 (Generic description)**.** *Suppose that a set $Q \subseteq S_{\mathbb{C}}$ contains an open dense subset of $S_{\mathbb{C}}$. Consider the set*

$$\mathcal{N}_Q := \{U \operatorname{Diag}(x) V^\top \ : \ U \in O_{\mathbb{C}}(n),\ V \in O_{\mathbb{C}}(t),\ x \in Q\}.$$

Then $\mathcal{N}_Q$ is a dense subset of $\mathcal{M}_{\mathbb{C}}$, and $\mathcal{N}_Q$ contains an open dense subset of $\mathcal{M}_{\mathbb{C}}$.

Proof. After we show $\mathcal{N}_Q$ is a dense subset of $\mathcal{M}_{\mathbb{C}}$, Chevalley's theorem [16, Theorem 3.16] will immediately imply that $\mathcal{N}_Q$ contains an open dense subset of $\mathcal{M}_{\mathbb{C}}$, as claimed.

We first argue the inclusion $\mathcal{N}_{S_{\mathbb{C}}} \subseteq \mathcal{M}_{\mathbb{C}}$ (and hence $\mathcal{N}_Q \subseteq \mathcal{M}_{\mathbb{C}}$). To this end, for any $f \in I(\mathcal{M}_{\mathbb{C}})$, note that the polynomial $q(x) := f(\operatorname{Diag}(x))$ vanishes on S and therefore on $S_{\mathbb{C}}$. Hence the inclusion $\{\operatorname{Diag}(x) \ : \ x \in S_{\mathbb{C}}\} \subseteq \mathcal{M}_{\mathbb{C}}$ holds. Since $\mathcal{M}_{\mathbb{C}}$ is $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$-invariant (Proposition 2.2), we conclude that $\mathcal{N}_{S_{\mathbb{C}}} \subseteq \mathcal{M}_{\mathbb{C}}$, as claimed. Moreover, clearly $\mathcal{M}$ is a subset of $\mathcal{N}_{S_{\mathbb{C}}}$, and hence the inclusion $\mathcal{M}_{\mathbb{C}} \subseteq \overline{\mathcal{N}_{S_{\mathbb{C}}}}$ holds. We conclude the equality $\mathcal{M}_{\mathbb{C}} = \overline{\mathcal{N}_{S_{\mathbb{C}}}}$.

Now suppose that Q contains an open dense subset of $S_{\mathbb{C}}$ and consider the continuous polynomial map $P : O_{\mathbb{C}}(n) \times S_{\mathbb{C}} \times O_{\mathbb{C}}(t) \to \mathcal{M}_{\mathbb{C}}$ given by

$$P(U, x, V) := U \operatorname{Diag}(x) V^\top.$$

Noting the equations $\overline{Q} = S_{\mathbb{C}}$ and $\mathcal{N}_Q = P(O_{\mathbb{C}}(n) \times Q \times O_{\mathbb{C}}(t))$, we obtain

$$\begin{aligned}\overline{\mathcal{N}_Q} &= \overline{P(O_{\mathbb{C}}(n) \times Q \times O_{\mathbb{C}}(t))} = \overline{P(\overline{O_{\mathbb{C}}(n) \times Q \times O_{\mathbb{C}}(t)})} \\ &= \overline{P(O_{\mathbb{C}}(n) \times \overline{Q} \times O_{\mathbb{C}}(t))} = \overline{\mathcal{N}_{S_{\mathbb{C}}}} = \mathcal{M}_{\mathbb{C}}.\end{aligned}$$

Hence $\mathcal{N}_Q$ is a dense subset of $\mathcal{M}_{\mathbb{C}}$, as claimed. □

Remark 3.5. The variety $\mathcal{M}_{\mathbb{C}}$ may contain matrices that do not admit an algebraic SVD and hence the closure operation in Theorem 3.4 is not superfluous. For example the Zariski closure of $\mathbb{R}^{2\times 2}_1$ contains the matrix $\left(\begin{smallmatrix} 1 & i \\ 0 & 0 \end{smallmatrix}\right)$, which we saw earlier does not have an algebraic SVD.

Though in the notation of Theorem 3.4, the set $\mathcal{N}_{S_{\mathbb{C}}}$ coincides with $\mathcal{M}_{\mathbb{C}}$ only up to closure, we next show that equality does hold unconditionally when restricted to diagonal matrices. For any matrix $B \in \mathbb{C}^{n\times n}$ we define $e_1(B), \ldots, e_n(B)$ to be the n coefficients of the characteristic polynomial of B, that is $e_1(B), \ldots, e_n(B)$ satisfy

$$\det(\lambda I - B) = \lambda^n - e_1(B)\lambda^{n-1} + \cdots + (-1)^n e_n(B).$$

For any point $b \in \mathbb{C}^n$, we define $e_i(b) = e_i(\operatorname{Diag}(b))$ for every $i = 1, \ldots, n$. In other words, $e_1(b), \ldots, e_n(b)$ are the elementary symmetric polynomials in $b_1, \ldots, b_n$.

Theorem 3.6. *The equality, $S_{\mathbb{C}} = \{x \in \mathbb{C}^n : \operatorname{Diag}(x) \in \mathcal{M}_{\mathbb{C}}\}$, holds.*

Proof. The inclusion $\subseteq$ follows immediately from the inclusion $\mathcal{N}_{S_{\mathbb{C}}} \subseteq \mathcal{M}_{\mathbb{C}}$ established in Theorem 3.4. For the reverse inclusion, define the set

$$\Omega := \{y \in \mathbb{C}^n \ : \ y_i = e_i(x_1^2, \ldots, x_n^2) \quad \text{for some } x \in S_{\mathbb{C}}\}.$$

We first claim that Ω is a variety. To see this, by [26, Proposition 2.6.4], the variety $S_{\mathbb{C}}$ admits some $\Pi_n^{\pm}$-invariant defining polynomials $f_1, \ldots, f_k \in \mathbb{C}[x]$. Since f_j are invariant under coordinate sign changes, they are in fact symmetric polynomials in

the squares $x_1^2, \ldots, x_n^2$. Then by the fundamental theorem of symmetric polynomials, we may write each f_j as some polynomial q_j in the quantities $e_i(x_1^2, \ldots, x_n^2)$. We claim that Ω is precisely the zero set of $\{q_1, \ldots, q_k\}$. By construction q_j vanish on Ω. Conversely, suppose $q_j(y) = 0$ for each j. Letting $x_1^2, \ldots, x_n^2$ be the roots of the polynomial $\lambda^n - y_1\lambda^{n-1} + \cdots + (-1)^n y_n$, we obtain a point $x \in \mathbb{C}^n$ satisfying $y_i = e_i(x_1^2, \ldots, x_n^2)$ for each i. We deduce then that x lies in $S_{\mathbb{C}}$ and hence y lies in Ω as claimed. We conclude that Ω is closed.

Observe the mapping $\pi \colon \mathcal{M}_{\mathbb{C}} \to \mathbb{C}^n$ defined by $\pi(X) = (e_1(XX^\top), \ldots, e_n(XX^\top))$ satisfies $\pi(\mathcal{N}_{S_{\mathbb{C}}}) \subseteq \Omega$, and so we deduce $\pi(\mathcal{M}_{\mathbb{C}}) = \pi(\overline{\mathcal{N}_{S_{\mathbb{C}}}}) \subseteq \overline{\pi(\mathcal{N}_{S_{\mathbb{C}}})} \subseteq \Omega$. Hence for any $y \in \mathbb{C}^n$ satisfying $\operatorname{Diag}(y) \in \mathcal{M}_{\mathbb{C}}$, there exists $x \in S_{\mathbb{C}}$ satisfying $e_i(x_1^2, \ldots, x_n^2) = e_i(y_1^2, \ldots, y_n^2)$ for each index $i = 1 \ldots, n$. We deduce that $x_1^2, \ldots, x_n^2$ and $y_1^2, \ldots, y_n^2$ are all roots of the same characteristic polynomial of degree n. Taking into account that $S_{\mathbb{C}}$ is $\Pi_n^{\pm}$-invariant, we conclude that y lies in $S_{\mathbb{C}}$. The result follows. □

We conclude with the following two enlightening corollaries, which in particular characterize matrices in $\mathcal{M}_{\mathbb{C}}$ admitting an algebraic SVD.

Corollary 3.7 (SVD in the closure)**.** *A matrix $X \in \mathcal{M}_{\mathbb{C}}$ admits an algebraic SVD if and only if $XX^\top$ is diagonalizable,* $\operatorname{rank}(X) = \operatorname{rank}(XX^\top)$*, and the vector of algebraic singular values of X lies in $S_{\mathbb{C}}$.*

Proof. This follows immediately from Proposition 2.2, and Theorems 3.2 and 3.6. □

Corollary 3.8 (Eigenvalues in the closure)**.** *If X is a matrix in $\mathcal{M}_{\mathbb{C}}$, then the vector of the square roots of the eigenvalues of $XX^\top$ lies in $S_{\mathbb{C}}$.*

Proof. Recall that, if $\mathcal{U}$ is an open dense subset of a variety $\mathcal{V}$, then $\mathcal{U}$ has nonempty intersection with any irreducible component of $\mathcal{V}$; see [3, 1.2 Proposition]. Hence the intersection of $\mathcal{U}$ with any irreducible component of $\mathcal{V}$ is open dense in that component, and is Euclidean dense in that component as well; see [22, page 60, Corollary 1]. Consequently, $\mathcal{U}$ is Euclidean dense in $\mathcal{V}$.

From Theorem 3.4 we know $\mathcal{N}_{S_{\mathbb{C}}}$ contains an open dense subset of $\mathcal{M}_{\mathbb{C}}$. It follows from the above discussion that $\mathcal{N}_{S_{\mathbb{C}}}$ is Euclidean dense in $\mathcal{M}_{\mathbb{C}}$. Given $X \in \mathcal{M}_{\mathbb{C}}$, we let x be the vector of the square roots of the eigenvalues of $XX^\top$, which is defined up to sign and order. We know there is a sequence $X_k := U_k \operatorname{Diag}(x^k) V_k^\top$, where $U_k \in O_{\mathbb{C}}(n)$, $V_k \in O_{\mathbb{C}}(t)$, and $x^k \in S_{\mathbb{C}}$ such that $X_k \to X$ as $k \to \infty$. Hence

$$(e_1(X_k X_k^\top), \ldots, e_n(X_k X_k^\top)) \to (e_1(XX^\top), \ldots, e_n(XX^\top)).$$

Since roots of polynomials are continuous with respect to the coefficients [30, Theorem 1], we deduce that the roots of the characteristic polynomial $\det(\lambda I - X_k X_k^\top)$, namely $((x_1^k)^2, \ldots, (x_n^k)^2)$, converge to $(x_1^2, \ldots, x_n^2)$ up to a coordinate reordering of x^k's and x. Passing to a subsequence, we deduce that x_k converge to x up to a signed permutation. Since $S_{\mathbb{C}}$ is closed, we conclude that x lies in $S_{\mathbb{C}}$, as claimed. □

3.2. **GIT perspective of algebraic SVD.** The algebraic SVD can be viewed from the perspective of Geometric Invariant Theory (GIT) [10, Chapter 2]. Let G be the group $O_{\mathbb{C}}(n) \times O_{\mathbb{C}}(t)$ acting on $\mathbb{C}^{n \times t}$ via $(U, V) \cdot A = UAV^\top$. For any variety $\mathcal{V}$ over $\mathbb{C}$, let $\mathbb{C}[\mathcal{V}]$ be the ring of polynomial maps $\mathcal{V} \to \mathbb{C}$. Fix the G-invariant variety $\mathcal{M}_{\mathbb{C}}$ and define the *invariant ring*

$$\mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G := \{f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}] \ : f \text{ is } G\text{-invariant}\}$$

as a subring of $\mathbb{C}[\mathcal{M}_{\mathbb{C}}]$. Consider a function $f \in \mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G$. Since the map $q(x) := f \circ \mathrm{Diag}\,(x)$ lies in $\mathbb{C}[S_{\mathbb{C}}]$ and is $\Pi_n^{\pm}$-invariant, we may write q as a polynomial map in the values $e_i(x_1^2, \ldots, x_n^2)$. Hence by passing to the limit, f itself can be expressed as a polynomial over $\mathbb{C}$ in the ordered sequence of coefficients $e_1(XX^\top), \ldots, e_n(XX^\top)$. In other words, the following equality holds:

$$\mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G = \mathbb{C}[e_1(XX^\top), \ldots, e_n(XX^\top)]$$

Observe that $\mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G$ is a finitely generated reduced $\mathbb{C}$-algebra, and as such, there is a variety over $\mathbb{C}$ denoted by $\mathcal{M}_{\mathbb{C}} /\!\!/ G$, such that $\mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G$ is isomorphic to $\mathbb{C}[\mathcal{M}_{\mathbb{C}} /\!\!/ G]$. This variety (up to isomorphism) is the categorical G-quotient constructed in the affine setting of GIT, and is denoted by $\mathcal{M}_{\mathbb{C}} /\!\!/ G$. Concretely, we may write $\mathcal{M}_{\mathbb{C}} /\!\!/ G$ as the variety corresponding to the ideal

$$\{f \in \mathbb{C}[x] : f(e_1(XX^\top), \ldots, e_n(XX^\top)) = 0 \quad \text{for all } X \in \mathcal{M}_{\mathbb{C}}\}.$$

A bit of thought shows that in our case, we may equivalently write

$$\mathcal{M}_{\mathbb{C}} /\!\!/ G = \{y \in \mathbb{C}^n \ : \ y_i = e_i(x_1^2, \ldots, x_n^2) \quad \text{for some } x \in S_{\mathbb{C}}\}.$$

This was already implicitly shown in the proof of Theorem 3.6.

The *quotient map* $\pi : \mathcal{M}_{\mathbb{C}} \to \mathcal{M}_{\mathbb{C}} /\!\!/ G$ is the surjective polynomial map associated to the inclusion $\mathbb{C}[\mathcal{M}_{\mathbb{C}}]^G \hookrightarrow \mathbb{C}[\mathcal{M}_{\mathbb{C}}]$. To be precise, in our case we have

$$\pi(X) = (e_1(XX^\top), \ldots, e_n(XX^\top))$$

Intuitively $\mathcal{M}_{\mathbb{C}} /\!\!/ G$ can be "identified" with the space of closed orbits for the action of G on $\mathcal{M}_{\mathbb{C}}$, but not the orbit space. It can be proved that a G-orbit in $\mathcal{M}_{\mathbb{C}}$ is closed if and only if it is the orbit of a diagonal matrix. In other words, the orbit of a matrix X is closed if and only if X admits an algebraic SVD. By contrast, all $O(n) \times O(t)$-orbits in $\mathcal{M}$ are closed (compare these facts with [24, §16]).

4. ED critical points of an orthogonally invariant variety

We are now ready to prove our main results characterizing ED critical points of a data point $Y \in \mathbb{C}^{n \times t}$ with respect to an orthogonally invariant matrix variety $\mathcal{M} \subseteq \mathbb{R}^{n \times t}$. We first give the precise definition of an ED critical point; see [11, §2]. For any variety $\mathcal{V}$ over $\mathbb{R}$ or $\mathbb{C}$, we let $\mathcal{V}^{\mathrm{reg}}$ be the open dense subset of regular points in $\mathcal{V}$. Recall that if $\mathcal{V}$ is a union of irreducible varieties $\mathcal{V}_i$, then $\mathcal{V}^{\mathrm{reg}}$ is the union of $\mathcal{V}_i^{\mathrm{reg}}$ minus the points in the intersection of any two irreducible components. In what follows, for any two vectors $v, w \in \mathbb{C}^n$, the symbol $v \perp w$ means $v^\top w = 0$, and for any set $Q \subseteq \mathbb{C}^n$ we define $Q^\perp := \{v \in \mathbb{C}^n : v^\top w = 0 \text{ for all } w \in Q\}$.

Definition 4.1 (ED critical point, ED degree)**.** Let $\mathcal{V}$ be a real variety in $\mathbb{R}^n$ and consider a data point $y \in \mathbb{C}^n$. An *ED critical point* of y with respect to $\mathcal{V}$ is a point $x \in \mathcal{V}_{\mathbb{C}}^{\mathrm{reg}}$ such that $y - x \in \mathcal{T}_{\mathcal{V}_{\mathbb{C}}}(x)^\perp$, where $\mathcal{T}_{\mathcal{V}_{\mathbb{C}}}(x)$ is the tangent space of $\mathcal{V}_{\mathbb{C}}$ at x.

For any general point y in $\mathbb{C}^n$, the number of ED critical points of y with respect to $\mathcal{V}$ is a constant; see [11] called the *ED degree* of $\mathcal{V}$ and denoted by EDdegree($\mathcal{V}$).

Here is a basic fact that will be needed later.

Lemma 4.2. *Let $\mathcal{V} \subseteq \mathbb{R}^n$ be a variety and let $\mathcal{W}$ be an open dense subset of $\mathcal{V}_{\mathbb{C}}$. Then all ED critical points of a general $y \in \mathbb{C}^n$ with respect to $\mathcal{V}$ lie in $\mathcal{W}$.*

Proof. The proof is a dimensional argument explained in [11]. Without loss of generality assume that $\mathcal{V}_{\mathbb{C}}$ is irreducible. Consider the ED correspondence $\mathcal{E}_{\mathcal{V}_{\mathbb{C}}}$, as defined in [11, §4], with its two projections π_1 on $\mathcal{V}_{\mathbb{C}}$ and π_2 on $\mathbb{C}^n$. Since π_1 is an affine vector bundle over $\mathcal{V}_{\mathbb{C}}^{reg}$ of rank $n - \dim\mathcal{V}_{\mathbb{C}}$, it follows that $\pi_2\left(\pi_1^{-1}(\mathcal{V}_{\mathbb{C}} \setminus \mathcal{W})\right)$ has dimension smaller than n. □

Remark 4.3. We mention in passing, that the ED degree of a variety $\mathcal{V}$, as defined above equals the sum of the ED degrees of its irreducible components $\mathcal{V}_i$, which coincides with the original definition of ED degree in [11]. This follows from Lemma 4.2 by noting that the set $\mathcal{V}_{\mathbb{C}}^{\text{reg}} \cap (\mathcal{V}_i)_{\mathbb{C}}$ is an open dense subset of $(\mathcal{V}_i)_{\mathbb{C}}$ for each i.

We say that two matrices X and Y admit a *simultaneous algebraic SVD* if there exist orthogonal matrices $U \in O_{\mathbb{C}}(n)$ and $V \in O_{\mathbb{C}}(t)$ so that both $U^\top XV$ and $U^\top YV$ are diagonal matrices. Our first main result is that every ED critical point X of a general matrix $Y \in \mathbb{C}^{n\times t}$ with respect to an orthogonally invariant variety $\mathcal{M}$ admits a simultaneous algebraic SVD with Y.

Theorem 4.4 (Simultaneous SVD)**.** *Fix an $O(n) \times O(t)$-invariant matrix variety $\mathcal{M} \subseteq \mathbb{R}^{n\times t}$. Consider a matrix $Y \in \mathbb{C}^{n\times t}$ so that the eigenvalues of $YY^\top$ are nonzero and distinct. Then any ED critical point X of Y with respect to $\mathcal{M}$ admits a simultaneous algebraic SVD with Y.*

The proof of this theorem relies on the following three lemmas.

Lemma 4.5. *The tangent space of $O_{\mathbb{C}}(n)$ at a point $U \in O_{\mathbb{C}}(n)$ is*

$$\begin{aligned}\mathcal{T}_{O_{\mathbb{C}}(n)}(U) &= \{ZU \ : \ Z \in \mathbb{C}^{n\times n} \text{ is skew-symmetric}\}\\ &= \{UZ \ : \ Z \in \mathbb{C}^{n\times n} \text{ is skew-symmetric}\}.\end{aligned}$$

Proof. Recall $O_{\mathbb{C}}(n) = \{W \in \mathbb{C}^{n\times n} \ : \ WW^\top = I\}$. Consider the map $F : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ given by $W \mapsto WW^\top$. Note that for any $W, B \in \mathbb{C}^{n\times n}$ and $t \in \mathbb{R}$, one has

$$(W + tB)(W + tB)^\top = WW^\top + t(WB^\top + BW^\top) + t^2BB^\top.$$

Hence given $U \in O_{\mathbb{C}}(n)$, we have $[\nabla F(U)](B) = UB^\top + BU^\top$. The tangent space $\mathcal{T}_{O_{\mathbb{C}}(n)}(U)$ is the kernel of the linear map $\nabla F(U)$. Consider the matrix $Z := BU^\top$. Then $[\nabla F(U)](B) = 0$ if and only if $Z^\top + Z = 0$ which means Z is skew-symmetric. This proves the first description of $\mathcal{T}_{O_{\mathbb{C}}(n)}(U)$. The second description follows by considering the map $W \mapsto W^\top W$ instead of F. □

Lemma 4.6. *A matrix $A \in \mathbb{C}^{n\times n}$ is symmetric if and only if* $\operatorname{trace}(AZ) = 0$ *for any skew-symmetric matrix $Z \in \mathbb{C}^{n\times n}$.*

Proof. The "if" part follows because $A_{ij} - A_{ji} = \operatorname{trace}(A(E^{ij} - E^{ji}))$ where E^{ij} denotes the $n \times n$ matrix whose (i, j)-entry is one and all other entries are zero. The "only if" part follows by the same reasoning since $\{E^{ij} - E^{ji}\}$ is a basis for the space of skew-symmetric matrices. □

Lemma 4.7. *Consider a matrix $A \in \mathbb{C}^{n\times t}$ and a diagonal matrix $D \in \mathbb{C}^{n\times t}$ with nonzero diagonal entries d_i such that the squares d_i^2 are distinct. Then if $AD^\top$ and $D^\top A$ are both symmetric, the matrix A must be diagonal.*

Proof. The symmetry of $AD^\top$ means $A_{ij}d_j = A_{ji}d_i$ for any $i, j = 1, \ldots, n$. In addition, the symmetry of $D^\top A$ implies $A_{ij}d_i = A_{ji}d_j$ for all $i, j = 1, \ldots, n$ and $A_{ij}d_i = 0$ for any $i = 1, \ldots, n$ and $j > n$. Therefore for any i, j, one has

$$A_{ij}d_id_j = A_{ji}d_i^2 \qquad \text{and} \qquad A_{ij}d_id_j = A_{ji}d_j^2.$$

Since $d_i^2 \neq d_j^2$ for all $i \neq j$, we get $A_{ij} = 0$ for all $i \neq j$, $i, j = 1, \ldots, n$. Since the d_i's are all nonzero and $A_{ij}d_i = 0$ for any $i = 1, \ldots, n$ and $j > n$, we have $A_{ij} = 0$ for any $i = 1, \ldots, n$ and $j > n$. Thus A is diagonal. □

Remark 4.8. The assumption $d_i^2 \neq d_j^2$ for $i \neq j$ is necessary in Lemma 4.7. For example consider $D = I$ and the symmetric matrices

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix} \in O(2), \ \ \theta \in \mathbb{R}$$

for which $AD^\top$ and $D^\top A$ are both symmetric. However, A is diagonal only when $\theta = k\pi$ with $k \in \mathbb{Z}$.

Proof of Theorem 4.4. By Corollary 3.3, we may write $Y = UDV^\top$ for some $U \in O_{\mathbb{C}}(n)$, $V \in O_{\mathbb{C}}(t)$, and a diagonal matrix $D \in \mathbb{C}^{n\times t}$. Let X be an ED critical point of Y with respect to $\mathcal{M}$. Then $A := U^\top XV$ lies in $\mathcal{M}_{\mathbb{C}}$ (Proposition 2.2). To prove the theorem, we need to show that $A \in \mathbb{C}^{n\times t}$ is diagonal.

Consider the map $F : O_{\mathbb{C}}(n) \to \mathcal{M}_{\mathbb{C}}$ given by $W \mapsto WAV^\top$. Then

$$[\nabla F(U)](B) = BAV^\top \in \mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X),$$

for any $B \in \mathcal{T}_{O_{\mathbb{C}}(n)}(U)$. By Lemma 4.5, we may write $B = UZ$ for a skew-symmetric Z, yielding $UZAV^\top \in \mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X)$. Varying B, we see that the tangent space of $\mathcal{M}_{\mathbb{C}}$ at X contains $\{UZAV^\top \ : \ Z^\top = -Z\}$. Then, by the definition of ED critical point we have $\operatorname{trace}((Y - X)(UZAV^\top)^\top) = 0$ for any skew-symmetric matrix Z, and hence

$$0 = \operatorname{trace}(U(D - A)V^\top VA^\top Z^\top U^\top) = \operatorname{trace}((D - A)A^\top Z^\top).$$

By Lemma 4.6, this means $(D - A)A^\top$ is symmetric. Since $AA^\top$ is symmetric, we have that $DA^\top$ is symmetric; therefore the transpose $AD^\top$ is symmetric.

By considering $F : O_{\mathbb{C}}(t) \to \mathcal{M}_{\mathbb{C}}$ given by $W \mapsto UAW^\top$, we get as above, that $\{UAZ^\top V^\top \ : \ Z^\top = -Z\} \subseteq \mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X)$. It follows that

$$0 = \operatorname{trace}((U(D - A)V^\top)^\top UAZ^\top V^\top) = \operatorname{trace}((D - A)^\top AZ^\top)$$

for any skew-symmetric matrix Z, and by Lemma 4.6, $(D - A)^\top A$ is symmetric. Again, since $A^\top A$ is symmetric, we get that $D^\top A$ is symmetric. Since $AD^\top$ and $D^\top A$ are both symmetric, we conclude A is diagonal by Lemma 4.7, as claimed. □

The next ingredient in our development is a version of Sard's Theorem in algebraic geometry (often called "generic smoothness" in textbooks); see [18, III, Corollary 10.7].

Theorem 4.9 (Generic smoothness on the target)**.** *Let $\mathcal{V}$ and $\mathcal{W}$ be varieties over $\mathbb{C}$. Consider a dominant polynomial map $f : \mathcal{V} \to \mathcal{W}$. Then there is an open dense subset $\mathcal{W}'$ of $\mathcal{W}^{\mathrm{reg}}$ (and hence of $\mathcal{W}$) such that for any $w \in \mathcal{W}'$ and any point $v \in \mathcal{V}^{\mathrm{reg}} \cap f^{-1}(w)$, the linear map $\nabla f(v) \colon \mathcal{T}_{\mathcal{V}}(v) \to \mathcal{T}_{\mathcal{W}}(w)$ is surjective.*

We now establish a key technical result: a representation of the tangent space of $\mathcal{M}_{\mathbb{C}}$ at a general matrix $X \in \mathcal{M}_{\mathbb{C}}$ in terms of the tangent space of $S_{\mathbb{C}}$ at the vector of algebraic singular values of X.

Proposition 4.10 (Transfer of tangent spaces)**.** *Consider a $\Pi_n^{\pm}$-invariant variety $S \subseteq \mathbb{R}^n$ and the induced real variety $\mathcal{M} := \sigma^{-1}(S)$. Then the following statements hold.*

(a) A general point $X \in \mathcal{M}_{\mathbb{C}}$ lies in $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$, admits an algebraic SVD, and its vector of algebraic singular values lies in $S_{\mathbb{C}}^{\mathrm{reg}}$. Moreover, the tangent space $\mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X)$ admits the representation

$$\mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X) = \left\{ \begin{array}{c} U Z_1 \operatorname{Diag}(x) V^\top + U \operatorname{Diag}(x) Z_2^\top V^\top + U \operatorname{Diag}(a) V^\top \ : \\ a \in \mathcal{T}_{S_{\mathbb{C}}}(x),\ Z_1,\ Z_2 \text{ are skew-symmetric} \end{array} \right\}, \tag{4.1}$$

for any $U \in O_{\mathbb{C}}(n)$, $V \in O_{\mathbb{C}}(t)$, and $x \in S_{\mathbb{C}}^{\mathrm{reg}}$ satisfying $X = U \operatorname{Diag}(x) V^\top$.

(b) A general point $x \in S_{\mathbb{C}}$ lies in $S_{\mathbb{C}}^{\mathrm{reg}}$. Moreover, for any $U \in O_{\mathbb{C}}(n), V \in O_{\mathbb{C}}(t)$, the point $X = U \operatorname{Diag}(x) V^\top$ lies in $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$, and satisfies (4.1).

Proof. We begin by proving claim (a). By Theorem 3.4 with $Q = S_{\mathbb{C}}^{\mathrm{reg}}$, a general point $X \in \mathcal{M}_{\mathbb{C}}$ admits an algebraic SVD: $X = U' \operatorname{Diag}(x') V'^\top$ for some $x' \in S_{\mathbb{C}}^{\mathrm{reg}}$. As $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$ is an open dense subset of $\mathcal{M}_{\mathbb{C}}$, we can assume that X lies in $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$. Consider the dominant polynomial map $P : O_{\mathbb{C}}(n) \times S_{\mathbb{C}} \times O_{\mathbb{C}}(t) \to \mathcal{M}_{\mathbb{C}}$ given by

$$P(\widetilde{U}, \widetilde{x}, \widetilde{V}) := \widetilde{U} \operatorname{Diag}(\widetilde{x}) \widetilde{V}^\top.$$

By Theorem 4.9 we can assume that $\nabla P(U, x, V)$ is surjective whenever we can write $X = U \operatorname{Diag}(x) V^\top$ for some $U \in O_{\mathbb{C}}(n)$, $V \in O_{\mathbb{C}}(t)$ and $x \in S_{\mathbb{C}}^{\mathrm{reg}}$. Therefore the description of tangent space in (4.1) follows from Leibniz rule on P and Lemma 4.5. Hence claim (a) is proved.

Next, we argue claim (b). To this end, let Θ be the dense open subset of $\mathcal{M}_{\mathbb{C}}$ guaranteed to exist by (a). We claim that we can assume that Θ is in addition orthogonally invariant. To see this, observe that all the claimed properties in (a) continue to hold on the dense, orthogonally invariant subset $\Gamma := \bigcup\{U\Theta V^T : U \in O_{\mathbb{C}}(n), V \in O_{\mathbb{C}}(t)\}$ of $\mathcal{M}_{\mathbb{C}}$. By Lemma 2.1, the set $\overline{\mathcal{M}_{\mathbb{C}} \setminus \Gamma}$ is an orthogonally invariant variety. Note now the inclusions $\Theta \subseteq \mathcal{M}_{\mathbb{C}} \setminus (\overline{\mathcal{M}_{\mathbb{C}} \setminus \Gamma}) \subseteq \Gamma$. It follows that $\mathcal{M}_{\mathbb{C}} \setminus (\overline{\mathcal{M}_{\mathbb{C}} \setminus \Gamma})$ is an orthogonally invariant, open, dense variety in $\mathcal{M}_{\mathbb{C}}$ on which all the properties in (a) hold. Replacing Θ with $\mathcal{M}_{\mathbb{C}} \setminus (\overline{\mathcal{M}_{\mathbb{C}} \setminus \Gamma})$, we may assume that Θ is indeed orthogonally invariant in the first place.

Next, we appeal to some results of Section 2.2. Let $\{S_i\}_{i=1}^k$ be the irreducible components of S and define the symmetrizations $S_i^\pi := \bigcup_{\pi \in \Pi_n^{\pm}} \pi S_i$ and the varieties $\mathcal{M}_i := \sigma^{-1}(S_i^\pi)$. Observe that $\overline{S_i}$ are the irreducible components of $S_{\mathbb{C}}$ and we have $\overline{S_i^\pi} = \bigcup_{\pi \in \Pi_n^{\pm}} \pi \overline{S_i}$. Note also that $\mathcal{M}_{\mathbb{C}}$ is the union of the varieties $\overline{\mathcal{M}_i}$.

By Proposition 2.5, each variety $\overline{\mathcal{M}_i}$ is a union of some irreducible components of $\mathcal{M}_{\mathbb{C}}$. Since the intersection of $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$ with any irreducible component of $\mathcal{M}_{\mathbb{C}}$ is open and dense in that component, we deduce that the intersection $\overline{\mathcal{M}_i} \cap \mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$ is an open dense subset of $\overline{\mathcal{M}_i}$ for each index i. Similarly, the intersection $\Theta \cap \overline{\mathcal{M}_i}$ is open and dense in each variety $\overline{\mathcal{M}_i}$. Then clearly Θ intersects $\mathcal{N}_{\overline{S_i^\pi}}$ for each index i, since by Theorem 3.4 the set $\mathcal{N}_{\overline{S_i^\pi}}$ contains an open dense subset of $\overline{\mathcal{M}_i}$. Therefore for each index i, the set Θ contains $\operatorname{Diag}(x_i)$ for some $x_i \in \overline{S_i^\pi}$.

We deduce that the diagonal restriction of Θ, namely the set

$$W := \{x \in \mathbb{C}^n : \operatorname{Diag}(x) \in \Theta\},$$

is an absolutely symmetric, open subset of $S_{\mathbb{C}}$ and it intersects each variety $\overline{S_i^\pi}$. In particular, W intersects each irreducible component $\overline{S_i}$. Since nonempty open

subsets of irreducible varieties are dense, we deduce that W is dense in $S_{\mathbb{C}}$. Moreover, since for any point $x \in W$, the matrix $\operatorname{Diag}(x)$ lies in Θ, we conclude

- $\operatorname{Diag}(x)$ lies in $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$ (and hence by orthogonal invariance so do all matrices $U \operatorname{Diag}(x) V^\top$ with $U \in O_{\mathbb{C}}(n), V \in O_{\mathbb{C}}(t)$) and x lies in $S_{\mathbb{C}}^{reg}$,
- equation (4.1) holds for $X = \operatorname{Diag}(x)$, and hence by orthogonal invariance of $\mathcal{M}_{\mathbb{C}}$ and of the description (4.1), the equation continues to hold for $X = U \operatorname{Diag}(x) V^\top$, where U and V arbitrary orthogonal matrices.

Thus all the desired conclusions hold for any x in the open dense subset W of $S_{\mathbb{C}}$. The result follows. □

We are now ready to prove the main result of this paper, equation ($\star$) from the introduction. As a byproduct, we will establish an explicit bijection between the ED critical points of a general matrix $Y = U \operatorname{Diag}(y) V^\top \in \mathbb{C}^{n\times t}$ on $\mathcal{M}$ and the ED critical points of y on $S_{\mathbb{C}}$.

Theorem 4.11 (ED degree)**.** *Consider a $\Pi_n^{\pm}$-invariant variety $S \subseteq \mathbb{R}^n$ and the induced real variety $\mathcal{M} := \sigma^{-1}(S)$. Then a general matrix $Y \in \mathbb{C}^{n\times t}$ admits a decomposition $Y = U \operatorname{Diag}(y) V^\top$, for some matrices $U \in O_{\mathbb{C}}(n)$, $V \in O_{\mathbb{C}}(t)$, and $y \in \mathbb{C}^n$. Moreover, then the set of ED critical points of Y with respect to $\mathcal{M}$ is*

$$\{U \operatorname{Diag}(x) V^\top \ : \ x \text{ is an ED critical point of } y \text{ with respect to } S\},$$

In particular, equality $\operatorname{EDdegree}(\mathcal{M}) = \operatorname{EDdegree}(S)$ *holds.*

Proof. For general $Y \in \mathbb{C}^{n\times t}$, the eigenvalues of $YY^\top$ are nonzero and distinct. Then by Corollary 3.3, we can be sure that Y admits an algebraic SVD. We fix such a decomposition $Y = U \operatorname{Diag}(y) V^\top$, for some $U \in O_{\mathbb{C}}(n)$, $V \in O_{\mathbb{C}}(t)$, and $y \in \mathbb{C}^n$.

Let X be an ED critical point of Y with respect to $\mathcal{M}$. By Theorem 4.4, we can assume that X and Y admit a simultaneous SVD, that is both $U'^\top X V'^\top$ and $U'^\top Y V'^\top$ are diagonal for some matrices $U' \in O_{\mathbb{C}}(n)$, $V' \in O_{\mathbb{C}}(t)$. Notice that the columns of U and U' are equal up to a sign change and a permutation. Similarly the first n columns of V and V' are equal up to a sign change and a permutation. Hence we may assume that X can be written as $X = U \operatorname{Diag}(x) V^\top$ for some $x \in S_{\mathbb{C}}$. By Lemma 4.2 and Proposition 4.10, we can further assume that X lies in $\mathcal{M}_{\mathbb{C}}^{\mathrm{reg}}$ and x lies in $S_{\mathbb{C}}^{\mathrm{reg}}$, and moreover the tangent space $\mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X)$ at $X = U \operatorname{Diag}(x) V^\top$ is given in (4.1).

We will now show that x is an ED critical point of y with respect to S. To see this, observe the inclusion

$$\{U \operatorname{Diag}(a) V^\top \ : \ a \in \mathcal{T}_{S_{\mathbb{C}}}(x)\} \subseteq \mathcal{T}_{\mathcal{M}_{\mathbb{C}}}(X).$$

and hence

$$0 = \operatorname{trace}(U \operatorname{Diag}(y - x) V^\top (U \operatorname{Diag}(a) V^\top)^\top) \quad \text{for any } a \in \mathcal{T}_{S_{\mathbb{C}}}(x).$$

Simplifying, we immediately conclude $(y - x)^\top a = 0$ for any $a \in \mathcal{T}_{S_{\mathbb{C}}}(x)$, and hence x is an ED critical point of y with respect to S.

Conversely, suppose $x \in S_{\mathbb{C}}^{reg}$ is an ED critical point of y with respect to S. Applying Theorem 3.4, we deduce that if a set $Q \subseteq \mathbb{C}^n$ contains an open dense set in $\mathbb{C}^n$, then

$$\{\widehat{U} \operatorname{Diag}(z) \widehat{V}^\top : z \in Q,\ \widehat{U} \in O_{\mathbb{C}}(n),\ \widehat{V} \in O_{\mathbb{C}}(t)\}$$

contains an open dense subset of $\mathbb{C}^{n\times t}$. Define now the matrix $X := U\,\mathrm{Diag}\,(x)\,V^\top$. Then by Lemma 4.2 and Proposition 4.10, we may assume that X is regular and the tangent space of $\mathcal{M}_\mathbb{C}$ at X is generated by all matrices of the form

i) $UZ\,\mathrm{Diag}\,(x)\,V^\top$ with Z skew-symmetric,
ii) $U\,\mathrm{Diag}\,(a)\,V^\top$ where a belongs to the tangent space of $S_\mathbb{C}$ at x,
iii) $U\,\mathrm{Diag}\,(x)\,Z^\top V^\top$ with Z skew-symmetric.

We will show

$$Y - X \perp \mathcal{T}_{\mathcal{M}_\mathbb{C}}(X) \tag{4.2}$$

by dividing the proof according to the three cases *i*),*ii*), *iii*) above. For *i*), observe

$$\mathrm{trace}\left((X-Y)(UZ\,\mathrm{Diag}\,(x)\,V^\top)^\top\right) = \mathrm{trace}\left(\mathrm{Diag}\,(x-y)\,\mathrm{Diag}\,(x)^\top Z^\top\right) = 0,$$

where the last equality follows from Lemma 4.6. The computation for *iii*) is entirely analogous. For *ii*), we obtain

$$\mathrm{trace}\Big((X-Y)(U\,\mathrm{Diag}\,(a)\,V^\top)^\top\Big) = \mathrm{trace}\Big((\mathrm{Diag}\,(x-y)\mathrm{Diag}\,(a)^\top\Big) = 0,$$

where the last equation follows from the hypothesis that x is an ED critical point of y on $S_\mathbb{C}$. We conclude that X is an ED critical point of Y relative to $\mathcal{M}_\mathbb{C}$, as claimed. The equality, $\mathrm{EDdegree}(\mathcal{M}) = \mathrm{EDdegree}(S)$, quickly follows. □

Example 4.12. To illustrate Theorem 4.11, we now derive the ED degree of some notable orthogonally invariant varieties summarized in the following table. The pairs $(\mathcal{M}, S)$ in all these examples were also discussed in [14, Section 4]. The dimension of $\mathcal{M}$ or $\mathcal{M}_\mathbb{C}$ can be computed using (2.2).

orthogonally invariant variety $\mathcal{M}$	dimension	absolutely symmetric variety S	dimension	EDdegree
$\mathbb{R}^{n\times t}_r$	$r(n+t-r)$	$\mathbb{R}^n_r$	r	$\binom{n}{r}$
$\mathcal{E}$	6	$E_{3,2}$	1	6
$O(n)$	$\binom{n}{2}$	$\{(\pm1,\ldots,\pm1)\}$	0	2^n
$SL^\pm_n$	n^2-1	H_n	$n-1$	$n2^n$
$\mathcal{F}_{n,t,d}$ (d even)	$nt-1$	$F_{n,d}$	$n-1$	[19, Cor. 2.12]

In the first three examples, the set S is a subspace arrangement and hence its ED degree is the number of distinct maximal subspaces in the arrangement. We will elaborate on this situation in Section 5.

The matrix variety $SL^\pm_n$ consists of all matrices $A \in \mathbb{C}^{n\times n}$ satisfying $\det(A) = \pm1$. The ED degree of $\mathrm{SL}^\pm_n$ was explicitly computed in [2]. We show below how our main theorem provides a simple alternate proof of their result.

The absolutely symmetric variety S in this case is $H_n := \{x \in \mathbb{R}^n \,:\, x_1x_2\cdots x_n = \pm1\}$. To compute the ED degree of H_n, we add up the ED degrees of its two irreducible components

$$H^+_n := \{x \in \mathbb{R}^n \,:\, x_1x_2\cdots x_n = 1\}$$

and

$$H^-_n := \{x \in \mathbb{R}^n \,:\, x_1x_2\cdots x_n = -1\}.$$

To compute the ED degree of H^+_n, we begin with a point $y \in \mathbb{C}^n$. Then by a straightforward computation, x is an ED critical point of y with respect to H^+_n if

and only if x solves the system

$$\begin{cases} x_i(x_i - y_i) = x_n(x_n - y_n) & \text{for all } i = 1, \dots, n-1 \\ x_1 \cdots x_n = 1. \end{cases} \tag{4.3}$$

By Bézout's Theorem, we know $\text{EDdegree}(H_n^+) \leq n2^{n-1}$. We now argue that the data point $y = 0$ has $n2^{n-1}$ ED critical points with respect to H_n^+ which proves that $\text{EDdegree}(H_n^+) = n2^{n-1}$.

When $y = 0$, the system (4.3) is equivalent to

$$\begin{cases} x_1^2 = \cdots = x_n^2 \\ x_1 \cdots x_n = 1. \end{cases}$$

which has $n2^{n-1}$ solutions in $(H_n^+)_{\mathbb{C}}$. Indeed, choose x_1 such that $x_1^n = \pm 1$ ($2n$ choices); then choose x_i for $i = 2, \dots, n-1$ such that $x_i^2 = x_1^2$ (2 choices for each i); finally set $x_n = \frac{1}{x_1 \cdots x_{n-1}}$. Hence $\text{EDdegree}(H_n^+) = n2^{n-1}$. Similarly, $\text{EDdegree}(H_n^-) = n2^{n-1}$, and therefore we conclude $\text{EDdegree}(H_n) = n2^n$.

The variety $\mathcal{F}_{n,t,d} = \{X \in \mathbb{R}^{n \times t} : \|X\|_d = 1\}$ is the unit ball of the Schatten d-norm $\|X\|_d := \left[\sum_{i=1}^n \sigma_i(X)^d\right]^{\frac{1}{d}}$. When d is even, the corresponding absolute symmetric variety is the *affine Fermat hypersurface*

$$F_{n,d} := \left\{ x \in \mathbb{R}^n \ : \ \sum_{i=1}^n x_i^d = 1 \right\}.$$

The ED degree of a Fermat hypersurface was computed in [19].

5. Orthogonally invariant varieties from subspace arrangements

In this section, we augment the results of the previous section in the special (and important) case when S is a subspace arrangement. Many important matrix varieties, such as the rank varieties $\mathbb{R}_r^{n \times t}$ and the essential variety $\mathcal{E}$, fall in this category. Recall that S is a *subspace arrangement* if S can be written as a union of finitely many affine subspaces $\{S_i\}_{i=1}^k$ of $\mathbb{R}^n$. Assuming that the representation of S is chosen in such a way that S_i is not contained in S_j for any distinct i, j, we call S_i the *affine components* of S. The following result follows directly from Theorem 4.11.

Corollary 5.1 (Affine arrangements)**.** *Consider a $\Pi_n^{\pm}$-invariant subspace arrangement $S \subseteq \mathbb{R}^n$ with affine components $\{S_i\}_{i=1}^k$, and define the induced real variety $\mathcal{M} := \sigma^{-1}(S)$. Then the equality,* $\text{EDdegree}(\mathcal{M}) = k$*, holds.*

Moreover, a general data point Y in $\mathbb{R}^{n \times t}$ has exactly k ED critical points with respect to $\mathcal{M}$: for any decomposition $Y = U \operatorname{Diag}(\sigma(Y)) V^\top$ with orthogonal matrices $U \in O(n)$ and $V \in O(t)$, the set of ED critical points is precisely

$$\{U \operatorname{Diag}(x) V^\top : x \text{ is the orthogonal projection of } \sigma(Y) \text{ onto } S_i\}.$$

In particular, all ED critical points of Y with respect to $\mathcal{M}$ are real.

Proof. Let Θ be the dense open subset of $\mathbb{C}^{n \times t}$ guaranteed to exist by Theorem 4.11. Clearly we can also assume that each matrix $Y \in \Theta$ has $\text{EDdegree}(\mathcal{M})$ many ED critical points with respect to $\mathcal{M}$. A standard argument shows that the set $\Theta_{\mathbb{R}} := \{Y \in \mathbb{R}^{n \times t} : Y \in \Theta\}$ is a dense open subset of $\mathbb{R}^{n \times t}$. Fix a matrix $Y \in \Theta_{\mathbb{R}}$ and consider a singular value decomposition $Y = U \operatorname{Diag}(\sigma(Y)) V^\top$ with orthogonal

matrices $U \in O(n)$ and $V \in O(t)$. By Theorem 4.11, the set of ED critical points of Y with respect to $\mathcal{M}$ is given by

$$\{U \operatorname{Diag}(x) V^\top \ : \ x \text{ is an ED critical point of } \sigma(Y) \text{ with respect to } S\}.$$

Since $\sigma(Y)$ is a real vector, the ED critical points of $\sigma(Y)$ with respect to S are precisely the orthogonal projections of $\sigma(Y)$ on each component S_i. Therefore we deduce $k = \text{EDdegree}(S) = \text{EDdegree}(\mathcal{M})$. □

The first three examples in Example 4.12 illustrate Corollary 5.1. Typically, as the data point $y \in \mathbb{R}^n$ varies, the number of real ED critical points of y with respect to a variety $\mathcal{V} \subseteq \mathbb{R}^n$ varies. Corollary 5.1 shows that when S is a subspace arrangement, all ED critical points of a real data point with respect to $\mathcal{M} = \sigma^{-1}(S)$ are again real and their number is constant. This unusual feature is easy to see using Theorem 4.11 that creates a bijection between the ED critical points of $\mathcal{M}$ and S, but is not at all obvious if S is not in the picture.

In common examples, outside of the subspace arrangement case, many ED critical points of a real data point may be non-real and the number of real critical points typically varies as the data point moves around. For instance, the hyperbola H_n in Example 4.12 can have complex ED critical points for a general $y \in \mathbb{R}^n$. The same is therefore true for $\mathrm{SL}_n^{\pm}$.

In a sense, Corollary 5.1 generalizes the fact that the pairs of singular vectors of a real matrix are real. Indeed, the pairs of singular vectors of a real matrix Y correspond to the ED critical points of Y with respect to the orthogonally invariant variety of rank one matrices; the corresponding absolutely symmetric variety is the union of all coordinate axes.

Remark 5.2. Results analogous to those in this paper hold for symmetric matrices under the action of the orthogonal group $U \cdot A = UAU^\top$. More precisely, consider the space of real $n \times n$ symmetric matrices $\mathcal{S}^n$. A set $\mathcal{M} \subseteq \mathcal{S}^n$ is *orthogonally invariant* provided $U\mathcal{M}U^\top = \mathcal{M}$ for all matrices $U \in O(n)$. Such a set $\mathcal{M}$ can be written as $\lambda^{-1}(S)$ where $\lambda \colon \mathcal{S}^n \to \mathbb{R}^n$ assigns to each matrix X the vector of its eigenvalues in a nonincreasing order and S is the diagonal restriction $S = \{x \in \mathbb{R}^n : \operatorname{Diag}(x) \in \mathcal{M}\}$. Conversely any permutation invariant set $S \subseteq \mathbb{R}^n$ gives rise to the orthogonally invariant set $\lambda^{-1}(S)$. Similar techniques to the ones developed here can then be used to study the correspondence between ED critical points of algebraic varieties S and $\lambda^{-1}(S)$. This research direction deserves further investigation.

References

[1] S. Agarwal, H.-L. Lee, B. Sturmfels, and R.R. Thomas. On the existence of epipolar matrices. *arXiv:1510.01401*.

[2] J.A. Baaijens and J. Draisma. Euclidean distance degrees of real algebraic groups. *Linear Algebra Appl.*, 467:174–187, 2015.

[3] A. Borel. *Linear Algebraic Groups*, volume 126. Springer Science & Business Media, 2012.

[4] D. Choudhury and R.A. Horn. An analog of the singular value decomposition for complex orthogonal equivalence. *Linear and Multilinear Algebra*, 21(2):149–162, 1987.

[5] A. Daniilidis, D. Drusvyatskiy, and A.S. Lewis. Orthogonal invariance and identifiability. *SIAM J. Matrix Anal. Appl.*, 35(2):580–598, 2014.

[6] A. Daniilidis, A.S. Lewis, J. Malick, and H. Sendov. Prox-regularity of spectral functions and spectral sets. *J. Convex Anal.*, 15(3):547–560, 2008.

[7] A. Daniilidis, J. Malick, and H.S. Sendov. Locally symmetric submanifolds lift to spectral manifolds. *Preprint U.A.B.* **23***/2009, 43 p., arXiv:1212.3936 [math.OC].*

[8] C. Davis. All convex invariant functions of Hermitian matrices. *Arch. Math.*, 8:276–278, 1957.

[9] M. Demazure. Sur deux problemes de reconstruction. Technical Report 992, INRIA, 1988.
[10] H. Derksen and G. Kemper. *Computational Invariant Theory*, volume 130. Springer Science & Business Media, 2013.
[11] J. Draisma, E. Horobeţ, G. Ottaviani, B. Sturmfels, and R.R. Thomas. The Euclidean distance degree of an algebraic variety. *Foundations of Computational Mathematics*, pages 1–51, 2015.
[12] D. Drusvyatskiy and C. Kempton. Variational analysis of spectral functions simplified. *Preprint arXiv:1506.05170*, 2015.
[13] D. Drusvyatskiy and M. Larsson. Approximating functions on stratified sets. *Trans. Amer. Math. Soc.*, 367(1):725–749, 2015.
[14] D. Drusvyatskiy, H.-L. Lee, and R.R. Thomas. Counting real critical points of the distance to orthogonally invariant matrix sets. *SIAM J. Matrix Anal. Appl.*, 36(3):1360–1380, 2015.
[15] G. Fløystad, J. Kileel, and G. Ottaviani. The Chow variety of the essential variety in computer vision. *In preparation.*
[16] J. Harris. *Algebraic Geometry: A First Course*, volume 133. Springer-Verlag, 1992.
[17] R. Hartley and A. Zisserman. *Multiview Geometry in Computer Vision.* Cambridge University Press, second edition, 2003.
[18] R. Hartshorne. *Algebraic Geometry*, volume 52. Springer, 1977.
[19] H. Lee. The Euclidean distance degree of Fermat hypersurfaces. *arXiv:1409.0684*, 2014.
[20] A.S. Lewis. Derivatives of spectral functions. *Math. Oper. Res.*, 21(3):576–588, 1996.
[21] S. Maybank. *Theory of Reconstruction from Image Motion*, volume 28. Springer Science & Business Media, 1993.
[22] D. Mumford. *The Red Book of Varieties and Schemes*, volume 1358 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, expanded edition, 1999.
[23] G. Ottaviani and R. Paoletti. A geometric perspective on the singular value decomposition. *Rend. Istit. Mat. Trieste*, 47:1–20, 2015.
[24] C. Procesi. The invariant theory of $n \times n$ matrices. *Advances in Math.*, 19(3):306–381, 1976.
[25] M. Šilhavý. Differentiability properties of isotropic functions. *Duke Math. J.*, 104(3):367–373, 2000.
[26] B. Sturmfels. *Algorithms in Invariant Theory.* Springer Science & Business Media, 2008.
[27] J. Sylvester. On the differentiability of $\mathrm{O}(n)$ invariant functions of symmetric matrices. *Duke Math. J.*, 52(2):475–483, 1985.
[28] J. von Neumann. Some matrix inequalities and metrization of matrix-space. *Tomck. Univ. Rev.*, 1:286–300, 1937.
[29] H. Whitney. Elementary structure of real algebraic varieties. *Annals of Mathematics*, 66(3):545–556, 1957.
[30] M. Zedek. Continuity and location of zeros of linear combinations of polynomials. *Proceedings of the American Mathematical Society*, 16(1):78–84, 1965.

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350

E-mail address: [ddrusv, hllee, rrthomas]@uw.edu

*Università di Firenze, viale Morgagni 67A, 50134 Firenze, Italy

E-mail address: ottavian@math.unifi.it