

**Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig**

**Reciprocal maximum likelihood degrees
of diagonal linear concentration models**

by

*Christopher Eur, Tara Fife, José Alejandro Samper,
and Tim Seynnaeve*

Preprint no.: 108

2020



RECIPROCAL MAXIMUM LIKELIHOOD DEGREES OF DIAGONAL LINEAR CONCENTRATION MODELS

CHRISTOPHER EUR - TARA FIFE
JOSÉ ALEJANDRO SAMPER - TIM SEYNNAEVE

We show that the reciprocal maximal likelihood degree (rmlD) of a diagonal linear concentration model $\mathcal{L} \subseteq \mathbb{C}^n$ of dimension r is equal to $(-2)^r \chi_M(\frac{1}{2})$, where χ_M is the characteristic polynomial of the matroid M associated to \mathcal{L} . In particular, this establishes the polynomiality of the rmlD for general diagonal linear concentration models, positively answering a question of Sturmfels, Timme, and Zwiernik.

1. Introduction

Let \mathbb{S}^n be the space of (real or complex) $n \times n$ symmetric matrices, and $\mathbb{S}_{>0}^n$ the subset consisting of real positive definite symmetric matrices. For a fixed $S \in \mathbb{S}_{>0}^n$, the **log-likelihood function** $\ell_S : \mathbb{S}_{>0}^n \rightarrow \mathbb{R}$ is defined by

$$\ell_S(K) := \log \det K - \text{trace}(S \cdot K).$$

For a subvariety $\mathcal{L} \subseteq \mathbb{S}^n$, the **maximum likelihood (ML) degree** $\text{mld}(\mathcal{L})$ is the number of invertible complex critical points of ℓ_S on the smooth locus of \mathcal{L} , counted with multiplicity, for a general choice of S . Writing $\mathcal{L}^{-1} \subseteq \mathbb{S}^n$ for the subvariety obtained as the closure of $\{K^{-1} \in \mathbb{S}^n \mid K \in \mathcal{L} \text{ invertible}\}$, one defines the **reciprocal maximum likelihood degree** $\text{rmlD}(\mathcal{L})$ as the number of invertible complex critical points of ℓ_S on the smooth loci of \mathcal{L}^{-1} , counted with multiplicity.

AMS 2020 Subject Classification: 14C17, 05B35, 62R01

Keywords: maximum likelihood degrees, reciprocal spaces, matroids, characteristic polynomials

Computing (reciprocal) ML degrees arises in statistical applications, where $\mathbb{S}_{>0}^n$ is often considered as the set of concentration matrices of multivariate normal distributions [17]. We caution that the terminology here regarding reciprocal vs. non-reciprocal ML degree is the opposite of that in [1, 16], where $\mathbb{S}_{>0}^n$ is considered as the set of *covariance* matrices (inverses of concentration matrices). In particular, our *rml*d is the ML degree of a linear covariance model. Our convention here agrees with [5, 8, 10, 17].

Let $[n] = \{1, \dots, n\}$. A **diagonal linear concentration model** is a linear subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$, where $\mathbb{C}^{[n]}$ is identified with the space of diagonal matrices in \mathbb{S}^n . Let M be the matroid on $[n]$ whose independent subsets are $I \subseteq [n]$ such that the composition $\mathbb{C}^I \hookrightarrow \mathbb{C}^{[n]} \rightarrow \mathcal{L}^\vee$ is injective. Without loss of generality, we always assume that \mathcal{L} is not contained in a coordinate hyperplane, or equivalently, that M is loopless, since in that case $\text{rml}d(\mathcal{L}) = 0$ from the definition. Our main result is the formula for the reciprocal ML degree of \mathcal{L} in terms of M .

Theorem 1.1. *Let $\mathcal{L} \subseteq \mathbb{C}^{[n]}$ be a diagonal linear concentration model of dimension r , and M the associated matroid of rank r on $[n]$. Then we have*

$$\text{rml}d(\mathcal{L}) = (-2)^r \chi_M\left(\frac{1}{2}\right),$$

where χ_M is the characteristic polynomial of M .

In [16, 17], the (non-reciprocal) ML degree of \mathcal{L} was shown to be $\chi_M(0)$. Computing the reciprocal ML degree presents fundamentally new challenges; see Remark 5.1 for a comparison.

From computational experiments, the authors of [16] asked whether the reciprocal ML degree of a *general* diagonal linear concentration model of dimension r in $\mathbb{C}^{[n]}$ is a polynomial in n of degree $r - 1$. Evaluating our Theorem 1.1 at uniform matroids answers their question positively.

Corollary 1.2. *Let $\mathcal{L} \subseteq \mathbb{C}^{[n]}$ be a general linear concentration model of dimension r . Then we have*

$$\text{rml}d(\mathcal{L}) = \sum_{i=1}^r \binom{n-i-1}{r-i} 2^{r-i}.$$

For instance, when $r = 3$ we have $2n^2 - 8n + 7$, and when $r = 4$ we have $4/3n^3 - 10n^2 + 68/3n - 15$, as predicted in [16] from numerical computations.

To prove Theorem 1.1, we use the following alternate description of the reciprocal ML degree, obtained by a standard computation in multivariable calculus. Let \mathcal{L}^\perp denote the orthogonal complement of a subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$ under the standard pairing

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle := \sum_{i=1}^n x_i y_i.$$

Proposition 1.3. [16, Proposition 4.3] *The reciprocal ML degree of a linear subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$ is equal to the number of solutions $(x_1, \dots, x_n) \in (\mathbb{C}^*)^{[n]}$, counted with multiplicity, to the following system of equations, where $s_1, \dots, s_n \in \mathbb{C}$ are generic parameters:*

$$(x_1^{-1}, \dots, x_n^{-1}) \in \mathcal{L} \quad \text{and} \quad (s_1 x_1^2 - x_1, \dots, s_n x_n^2 - x_n) \in \mathcal{L}^\perp.$$

Thus, we prove Theorem 1.1 by establishing the following generalization.

Theorem 1.4. *For an r -dimensional linear subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$, a generic choice of parameters $s_1, \dots, s_n \in \mathbb{C}$, and any integer $d \geq 1$, the number of solutions $(x_1, \dots, x_n) \in (\mathbb{C}^*)^{[n]}$, counted with multiplicity, to the system of equations*

$$(x_1^{-1}, \dots, x_n^{-1}) \in \mathcal{L} \quad \text{and} \quad (s_1 x_1^d - x_1, \dots, s_n x_n^d - x_n) \in \mathcal{L}^\perp \quad (\dagger)$$

is equal to

$$(-d)^r \chi_M\left(\frac{1}{d}\right), \quad \text{or equivalently,} \quad d^r T_M\left(1 - \frac{1}{d}, 0\right),$$

where χ_M is the characteristic polynomial and T_M is the Tutte polynomial of the matroid M associated to \mathcal{L} .

Remark 1.5. For an r -dimensional subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$, let $U(\mathcal{L}) := \mathcal{L} \cap (\mathbb{C}^*)^{[n]}$ be the hyperplane arrangement complement, and M the associated matroid. Then the Poincaré polynomial of $U(\mathcal{L})$

$$P_{U(\mathcal{L})}(q) := \sum_{i \geq 0} (\text{rank } H_i(U(\mathcal{L}); \mathbb{Z})) q^i$$

coincides with polynomial $(-q)^r \chi_M(-\frac{1}{q})$ [11]. In particular, Theorem 1.1 states that $\text{rml}d(\mathcal{L}) = (-1)^r P_{U(\mathcal{L})}(-2)$. This echoes the result of [7], which showed that, for a different log-likelihood function (from discrete statistical models), the ML degree of a smooth very affine variety U is its signed topological Euler characteristic $(-1)^{\dim U} P_U(-1)$. However, ML degrees in our case are not topological invariants of very affine varieties: Observe that $\mathcal{L}^{-1} \cap (\mathbb{C}^*)^{[n]} \simeq \mathcal{L} \cap (\mathbb{C}^*)^{[n]}$ but in general $\text{rml}d(\mathcal{L}^{-1}) = \text{mld}(\mathcal{L}) = \chi_M(0) \neq (-2)^r \chi_M(\frac{1}{2}) = \text{rml}d(\mathcal{L})$. It may still be interesting to find other families of subvarieties $\mathcal{L} \subseteq \mathbb{S}^n$ such that $\text{rml}d(\mathcal{L}) = (-1)^{\dim \mathcal{L}} P_{U(\mathcal{L})}(-2)$, where $U(\mathcal{L}) := \{K \in \mathcal{L} \mid K \text{ invertible}\}$. For example, general pencils of conics form one such family [3, 5].

Outline. In Section 2 we review properties of reciprocal linear spaces \mathcal{L}^{-1} , and introduce score varieties, which together with \mathcal{L}^{-1} encode the system of equations (\dagger) . After establishing two key technical lemmas in Section 3, in Section 4 we compute the number of solutions to the system of equations (\dagger) in $\mathbb{C}^{[n]}$, instead of in $(\mathbb{C}^*)^{[n]}$, in two different ways: One is a Bézout-like computation, and the other is a summation, with each summand corresponding to a set of solutions with specified support (non-zero coordinates). An inclusion-exclusion argument in Section 5 then yields the proof of Theorem 1.4.

Notations

For an affine subvariety $X \subseteq \mathbb{C}^n$, we write $\overline{X} \subseteq \mathbb{P}^n$ for its projective closure. If $X \subseteq \mathbb{C}^n$ is defined by a homogeneous ideal, then we write $\mathbb{P}X \subseteq \mathbb{P}^{n-1}$ for its projectivization. For a point $p \in X \subseteq \mathbb{C}^{[n]}$, we write $TC_p X \subseteq \mathbb{C}^{[n]}$ for the tangent cone of X at p . For $p = (p_1, \dots, p_n) \in \mathbb{C}^{[n]}$, we write $\text{supp}(p) := \{i \in [n] \mid p_i \neq 0\}$ for its support, and for $I \subseteq [n]$, write $p|_I$ for the projection of p onto $\mathbb{C}^I \subseteq \mathbb{C}^{[n]}$.

2. Reciprocal linear spaces and score varieties

We set notations concerning matroids associated to linear subspaces, and review necessary facts about reciprocal linear spaces. We assume familiarity with matroid theory, and refer to [12, 19] as standard references.

Let us fix a linear subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$ of dimension r . Let A be an $r \times n$ matrix whose row-span equals \mathcal{L} . We will often use the fact that the minimal sets among supports of elements in the row-span of A form the cocircuits of M .

For a subset $I \subseteq [n]$, let $\mathcal{L}|_I \subseteq \mathbb{C}^I$ be the image of \mathcal{L} under the coordinate projection $\mathbb{C}^{[n]} \rightarrow \mathbb{C}^I$, and let $\mathcal{L}_{/I}$ be the intersection of \mathcal{L} with the coordinate subspace $\{0\}^I \times \mathbb{C}^{[n] \setminus I}$, considered as a subspace of $\mathbb{C}^{[n] \setminus I}$. The matroid of $\mathcal{L}|_I$ is the restriction $M|I$, whereas the matroid of $\mathcal{L}_{/I}$ is the contraction M/I .

The **reciprocal linear space** \mathcal{L}^{-1} of \mathcal{L} is the Zariski closure in $\mathbb{C}^{[n]}$ of $\{(x_1, \dots, x_n) \in (\mathbb{C}^*)^{[n]} \mid (x_1^{-1}, \dots, x_n^{-1}) \in \mathcal{L}\}$. Note that $\mathcal{L}^{-1} \cap (\mathbb{C}^*)^{[n]}$ is smooth, being isomorphic to $\mathcal{L} \cap (\mathbb{C}^*)^{[n]}$. For $I \subseteq [n]$, we write $\mathcal{L}_{|I}^{-1}$ for $(\mathcal{L}|_I)^{-1}$, and likewise write $\mathcal{L}_{/I}^{-1} = (\mathcal{L}_{/I})^{-1}$. We collect together in the following theorem the known properties of \mathcal{L}^{-1} that we will need.

Theorem 2.1. *Let \mathcal{L}^{-1} be the reciprocal linear space of $\mathcal{L} \subseteq \mathbb{C}^{[n]}$.*

- (a) [18] *The defining ideal of \mathcal{L}^{-1} is homogeneous of degree $|\mu(M)|$, where $\mu(M) := \chi_M(0)$ the Möbius invariant of the matroid M .*
- (b) [13, Proposition 7] *The defining ideal of \mathcal{L}^{-1} is Cohen-Macaulay, with any basis of \mathcal{L}^\perp forming a system of parameters, i.e. $\mathcal{L}^{-1} \cap \mathcal{L}^\perp = \{\mathbf{0}\}$.*
- (c) [13, Proposition 5] *The intersection $\mathcal{L}^{-1} \cap ((\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F})$ is nonempty if and only if $F \subseteq [n]$ is a flat of M , and in that case, one has*

$$\mathcal{L}^{-1} \cap ((\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F}) = (\mathcal{L}_{|F}^{-1} \cap (\mathbb{C}^*)^F) \times \{0\}^{[n] \setminus F}.$$

- (d) [14, Theorem 24] *For a flat $F \subseteq [n]$ and a point $p \in \mathcal{L}^{-1}$ with $\text{supp}(p) = F$, the tangent cone of \mathcal{L}^{-1} at p is the product*

$$TC_p \mathcal{L}^{-1} = TC_{p|_F} \mathcal{L}_{|F}^{-1} \times \mathcal{L}_{/F}^{-1} \simeq \mathcal{L}_{|F} \times \mathcal{L}_{/F}^{-1}.$$

All four statements in Theorem 2.1 can be derived easily from the Gröbner basis for the defining ideal of \mathcal{L}^{-1} computed in [13, Theorem 4]. In [14], the statement of Theorem 2.1.(d) originally reads $TC_p\mathcal{L}^{-1} = p_{|F}^2\mathcal{L}_{|F} \times \mathcal{L}_{|F}^{-1}$, where $p^2\mathcal{L}$ denotes the linear subspace $\{(p_1^2x_1, \dots, p_n^2x_n) \mid (x_1, \dots, x_n) \in \mathcal{L}\}$. It is straightforward to verify that $TC_{p|F}\mathcal{L}_{|F}^{-1} = p_{|F}^2\mathcal{L}_{|F}$.

The reciprocal linear space \mathcal{L}^{-1} encodes the left half of the system of equations in Equation (†). Let us now consider the variety encoding the condition $(s_1x_1^d - x_1, \dots, s_nx_n^d - x_n) \in \mathcal{L}^\perp$. For an integer $d \geq 1$ and a parameter $s = (s_1, \dots, s_n) \in \mathbb{C}^n$, we define the **score variety** as

$$Y(\mathcal{L}, s, d) := \{(x_1, \dots, x_n) \in \mathbb{C}^{[n]} \mid (s_1x_1^d - x_1, \dots, s_nx_n^d - x_n) \in \mathcal{L}^\perp\}.$$

We will simply write Y when we trust that no confusion will arise. We note here that score varieties are smooth for a generic choice of $s \in \mathbb{C}^n$.

Lemma 2.2. *For $d \geq 1$ and a generic choice of $(s_1, \dots, s_n) \in \mathbb{C}^n$, the score variety Y is smooth.*

Proof. If $d = 1$, then Y is linear, so suppose $d \geq 2$. Let A be the $r \times n$ matrix whose row span equals \mathcal{L} , and let g_1, \dots, g_r be the polynomials obtained by multiplying the rows of A with $(s_1x_1^d - x_1, \dots, s_nx_n^d - x_n)^T$. These minimally generate the defining ideal $I_Y \subseteq \mathbb{C}[x_1, \dots, x_n]$ of Y . The Jacobian matrix with respect to these minimal generators is

$$\text{Jac}(x) = A \cdot \text{diag}(ds_1x_1^{d-1} - 1, \dots, ds_nx_n^{d-1} - 1), \quad (1)$$

i.e. matrix A whose i -th column is scaled by $ds_ix_i^{d-1} - 1$ for each $1 \leq i \leq n$. Suppose now that $\text{Jac}(x)$ has rank $< r$ for some $x \in \mathbb{C}^n$, that is, the restriction $M|I$ of the matroid M to the set $I = \{i \in [n] \mid ds_ix_i^{d-1} - 1 \neq 0\}$ has rank $< r$. This happens if and only if I is contained in a hyperplane flat of M , or equivalently, the subset $J := [n] \setminus I$ contains a cocircuit of M . As the minimal supports of the row-space of A constitute the cocircuits of M , let $v = (v_1, \dots, v_n) \in \mathbb{C}^n$ be the element in the row-space of A whose support $C^* = \text{supp}(v) \subseteq [n]$ is a cocircuit of M contained in J . Then we have

$$v \cdot (s_1x_1^d - x_1, \dots, s_nx_n^d - x_n)^T = \sum_{i \in C^*} v_i(s_ix_i^d - x_i) = \left(\frac{1}{d} - 1\right) \sum_{i \in C^*} v_ix_i$$

where last equality follows from $ds_ix_i^{d-1} - 1 = 0$ for $i \in J$. This quantity need be zero if $x \in Y$. We claim that for a general choice of (s_1, \dots, s_n) this quantity can never be zero: Consider the set

$$Z := \{(\zeta_1, \dots, \zeta_n) \in \mathbb{C}^n \mid \zeta_i \text{ is a } (d-1)\text{-th root of } \frac{1}{ds_i} \text{ if } i \in C^*\}.$$

For a general choice of (s_1, \dots, s_n) , no element of Z satisfies $\sum_{i \in C^*} v_ix_i = 0$. \square

3. Two genericity lemmas

We now present the two key technical lemmas for our future intersection multiplicity computations. Both make essential use of the fact that the parameter $s \in \mathbb{C}^n$ can be chosen generically, and that \mathbb{C} has characteristic zero. To state the first lemma, let us define a subscheme of $\mathbb{C}^{[n]}$

$$Y_\infty(\mathcal{L}, s, d) := \{(x_1, \dots, x_n) \in \mathbb{C}^{[n]} \mid (s_1 x_1^d, \dots, s_n x_n^d) \in \mathcal{L}^\perp\}.$$

Lemma 3.1. *For a generic choice of $s \in \mathbb{C}^n$ and any integer $d \geq 1$, one has*

$$\mathcal{L}^{-1} \cap Y_\infty(\mathcal{L}, s, d) = \{\mathbf{0}\}.$$

Proof. Theorem 2.1.(c) and the definition of Y_∞ implies that for $F \subseteq [n]$,

$$\mathcal{L}^{-1} \cap Y_\infty \cap ((\mathbb{C}^*)^F \times \{\mathbf{0}\}^{[n] \setminus F}) = (\mathcal{L}_F^{-1} \cap Y_\infty(\mathcal{L}|_F, s|_F, d) \cap (\mathbb{C}^*)^F) \times \{\mathbf{0}\}^{[n] \setminus F}$$

if F is a flat of M , and empty otherwise. Thus, it suffices to show that $\mathcal{L}^{-1} \cap Y_\infty(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]}$ is empty for a generic choice of $s \in \mathbb{C}^n$. Without loss of generality, we may assume that the $r \times n$ matrix A is of the form $[I_r \mid A']$, where I_r is the $r \times r$ identity matrix. For $1 \leq i \leq r$, let a'_i be the i -th row of A' . The ideal of Y_∞ is by definition minimally generated by

$$\mathfrak{d} = \{s_i x_i^d - a'_i \cdot (s_{r+1} x_{r+1}^d, \dots, s_n x_n^d)^T \mid 1 \leq i \leq r\}.$$

Fixing a generic choice of s_{r+1}, \dots, s_n , and letting s_1, \dots, s_r vary freely, for each $i = 1, \dots, r$ we may consider $s_i x_i^d - a'_i \cdot (s_{r+1} x_{r+1}^d, \dots, s_n x_n^d)^T$ as a pencil \mathfrak{d}_i of hypersurfaces in $\mathbb{C}^{[n]}$. Note that the base locus of the pencil \mathfrak{d}_i is contained in the i -th coordinate hyperplane. We thus obtain a map $\mathcal{L}^{-1} \cap (\mathbb{C}^*)^{[n]} \rightarrow (\mathbb{P}^1)^r$ of smooth varieties. By generic smoothness [6, III.10.7], the fiber over a generic point of $(\mathbb{P}^1)^r$, which is the intersection $\mathcal{L}^{-1} \cap Y_\infty(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]}$ for the corresponding choice of $s \in \mathbb{C}^n$, is either empty or 0-dimensional. The intersection $\mathcal{L}^{-1} \cap Y_\infty(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]}$ cannot be 0-dimensional since the ideals of \mathcal{L}^{-1} and Y_∞ are both homogeneous, and hence must be empty. \square

Remark 3.2. When $s = (1, \dots, 1)$ and $d = 1$, Lemma 3.1 is the second half of Theorem 2.1.(b), which was established by an explicit Gröbner basis computation. For $d \geq 2$ however, the lemma fails in general with $s = (1, \dots, 1)$.

Lemma 3.3. *For a generic choice of $s \in \mathbb{C}^n$ and any integer $d \geq 1$, the intersection $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]}$ is either empty or smooth of dimension 0.*

Proof. The proof is similar to that of Lemma 3.1. Without loss of generality, we again assume that the $r \times n$ matrix A is of the form $[I_r \mid A']$, where I_r is the

$r \times r$ identity matrix. For $1 \leq i \leq r$, let a'_i be the i -th row of A' . The ideal of Y is minimally generated by

$$\mathfrak{d} = \{(s_i x_i^d - x_i) - a'_i \cdot (s_{r+1} x_{r+1}^d - x_{r+1}, \dots, s_n x_n^d - x_n)^T \mid 1 \leq i \leq r\}.$$

Fixing a generic choice of s_{r+1}, \dots, s_n , and letting s_1, \dots, s_r vary freely, for each $i = 1, \dots, r$ we may consider $(s_i x_i^d - x_i) - a'_i \cdot (s_{r+1} x_{r+1}^d - x_{r+1}, \dots, s_n x_n^d - x_n)^T$ as a pencil \mathfrak{d}_i of hypersurfaces in $\mathbb{C}^{[n]}$. Note that the union of the base loci of $\mathfrak{d}_1, \dots, \mathfrak{d}_r$ is contained in the union of the coordinate hyperplanes. Thus, we obtain a map $\mathcal{L}^{-1} \cap (\mathbb{C}^*)^{[n]} \rightarrow (\mathbb{P}^1)^r$ of smooth varieties, and by generic smoothness, the general fiber, which is the intersection $\mathcal{L}^{-1} \cap Y \cap (\mathbb{C}^*)^{[n]}$ for a generic choice of $s \in \mathbb{C}^n$, is either empty or smooth of dimension 0. \square

Let us now denote

$$\mathcal{D}(\mathcal{L}, d) := \text{the degree of the (empty or 0-dimensional) subscheme} \\ \mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]} \subset \mathbb{C}^{[n]}$$

for a generic choice of $s \in \mathbb{C}^n$, which is equal to the number of points in the intersection since it is smooth by Lemma 3.3. Theorem 1.4 is now equivalently stated as $\mathcal{D}(\mathcal{L}, d) = (-d)^r \chi_M(\frac{1}{d})$, and Proposition 1.3 states that $\mathcal{D}(\mathcal{L}, 2) = \text{rml}(\mathcal{L})$.

4. Total intersection multiplicity

We now compute the degree of the intersection $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$ as a subscheme of $\mathbb{C}^{[n]}$ in two different ways. First, we have a Bézout-like computation.

Proposition 4.1. *For a generic choice of $s \in \mathbb{C}^n$, the intersection $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$ is a 0-dimensional scheme of degree $d^r |\mu(M)|$.*

Proof. For $i = 1, \dots, r$, let $\bar{f}_i \in \mathbb{C}[x_0, x_1, \dots, x_n]$ be the homogeneous polynomial obtained as i -th row of A times $(s_1 x_1^d - x_1 x_0^{d-1}, \dots, s_n x_n^d - x_n x_0^{d-1})^T$. We first claim that $(\bar{f}_1, \dots, \bar{f}_r)$ forms a regular sequence on the projective closure $\overline{\mathcal{L}^{-1}} \subset \mathbb{P}^n$. Since the projective variety $\overline{\mathcal{L}^{-1}}$ is (arithmetically) Cohen-Macaulay by Theorem 2.1.(b), it suffices to show that the intersection $\overline{\mathcal{L}^{-1}} \cap V(\bar{f}_1, \dots, \bar{f}_r)$ is 0-dimensional, as every system of parameters in a standard graded Cohen-Macaulay ring is a regular sequence [2, Theorem 2.1.2].

At the hyperplane at infinity, the intersection $V(x_0) \cap \overline{\mathcal{L}^{-1}} \cap V(\bar{f}_1, \dots, \bar{f}_r)$ is isomorphic to $\mathbb{P}\mathcal{L}^{-1} \cap \mathbb{P}Y_\infty$, which is empty for a generic $s \in \mathbb{C}^n$ by Lemma 3.1. On the complement of the hyperplane at infinity, the intersection is equal to $\mathcal{L}^{-1} \cap Y$, since the dehomogenizations of the polynomials $\bar{f}_1, \dots, \bar{f}_r$ give the

defining equations of Y . From Theorem 2.1.(c) and the definition of $Y(\mathcal{L}, s, d)$, it follows that

$$\mathcal{L}^{-1} \cap Y \cap ((\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F}) = (\mathcal{L}_{|F}^{-1} \cap Y(\mathcal{L}_{|F}, s_{|F}, d) \cap (\mathbb{C}^*)^F) \times \{0\}^{[n] \setminus F}$$

if $F \subseteq [n]$ is a flat or empty otherwise, and thus Lemma 3.3 applied to each flat F implies that $\mathcal{L}^{-1} \cap Y$ is 0-dimensional. Thus, the degree d polynomials $(\overline{f}_1, \dots, \overline{f}_r)$ form a regular sequence on $\overline{\mathcal{L}^{-1}}$, and hence the degree of $\mathcal{L}^{-1} \cap V(\overline{f}_1, \dots, \overline{f}_r)$ is $d^r \deg(\overline{\mathcal{L}^{-1}})$. As the ideal of \mathcal{L}^{-1} is homogeneous, the degrees of $\mathbb{P}\mathcal{L}^{-1}$ and $\overline{\mathcal{L}^{-1}}$ are equal, with the value being $|\mu(M)|$ by Theorem 2.1.(a). Lastly, since $\overline{\mathcal{L}^{-1}} \cap V(\overline{f}_1, \dots, \overline{f}_r)$ is empty at the hyperplane at infinity, the degree of the intersection is equal to the degree of $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$. \square

We now compute the degree of $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$ as the sum of contributions from the various strata of \mathcal{L}^{-1} . First, we need the following notation. Note that a 0-dimensional subscheme $X \subset \mathbb{C}^{[n]}$ is a union $\bigcup_{\alpha} X_{\alpha}$ of irreducible (possibly non-reduced) components X_{α} , each of which is topologically a point $(X_{\alpha})_{red}$ in $\mathbb{C}^{[n]}$. For a subset $I \subseteq [n]$, we write X^I to be the subscheme of X defined as the union of components of X whose support is I , i.e.

$$X^F := \bigcup \{X_{\alpha} \mid (X_{\alpha})_{red} \in (\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F}\}.$$

Moreover, recall the notation that $\mathcal{D}(\mathcal{L}, d)$ denotes the degree of the subscheme $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d) \cap (\mathbb{C}^*)^{[n]} \subset \mathbb{C}^{[n]}$ for a generic choice of $s \in \mathbb{C}^n$.

Proposition 4.2. *For a generic choice of $s \in \mathbb{C}^n$, and for a flat $F \subseteq [n]$ of M , the degree of $(\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d))^F \subset \mathbb{C}^{[n]}$ is equal to $\mathcal{D}(\mathcal{L}_{|F}, d) \cdot |\mu(M/F)|$.*

Proof. As topological spaces, the subscheme $(\mathcal{L}^{-1} \cap Y)^F$ is equal to the intersection $\mathcal{L}^{-1} \cap Y \cap ((\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F})$, which is by Theorem 2.1.(c) isomorphic to $\mathcal{L}_{|F}^{-1} \cap Y(\mathcal{L}_{|F}, s_{|F}, d) \cap (\mathbb{C}^*)^F$, which as a scheme is a disjoint union of $\mathcal{D}(\mathcal{L}_{|F}, d)$ many smooth points by Lemma 3.3. It remains only to show that if \tilde{p} is an irreducible component in $(\mathcal{L}^{-1} \cap Y)^F$, then the degree of \tilde{p} is $|\mu(M/F)|$.

For this end, we recall [4, Proposition 1.29]: Suppose two subvarieties X and X' of complementary dimensions in a smooth variety Z intersect dimensionally properly. Then, the multiplicity of the intersection $X \cap X'$ at a point q in the intersection is equal to the product of the degrees of projectivized tangent cones $\mathbb{P}TC_q X$ and $\mathbb{P}TC_q X'$, provided that $\mathbb{P}TC_q X$ and $\mathbb{P}TC_q X'$ are disjoint in $\mathbb{P}T_q Z$.

We apply this to \mathcal{L}^{-1} and Y , which intersect dimensionally properly by Proposition 4.1. Topologically \tilde{p} is a point $p \in (\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F}$. Combining Theorem 2.1.(d) with Theorem 2.1.(a), one has that the degree of $\mathbb{P}TC_p \mathcal{L}^{-1}$ is equal to $|\mu(M/F)|$. The degree of $\mathbb{P}TC_p Y$ is 1 since Y is smooth (Lemma 2.2). Thus, we are done once we show that $\mathbb{P}TC_p \mathcal{L}^{-1}$ and $\mathbb{P}TC_p Y$ are disjoint. This is done in the following lemma. \square

Lemma 4.3. *Let p be a point in $\mathcal{L}^{-1} \cap Y \cap ((\mathbb{C}^*)^F \times \{0\}^{[n] \setminus F})$ for a generic choice of $s \in \mathbb{C}^n$, and for a flat $F \subseteq [n]$. Then we have $\mathbb{P}TC_p \mathcal{L}^{-1} \cap \mathbb{P}TC_p Y = \emptyset$.*

Proof. Let A be an $r \times n$ matrix whose row-span is \mathcal{L} . As computed in the proof of Lemma 2.2 in Equation (1), the tangent cone $TC_p Y$ is equal to

$$\ker \left(A \cdot \text{diag}(ds_1 p_1^{d-1} - 1, \dots, ds_n p_n^{d-1} - 1) \right) \subset \mathbb{C}^{[n]}.$$

Since $p_i = 0$ for $i \in [n] \setminus F$, and since the cocircuits of the matroid M/F are cocircuits of M contained in $[n] \setminus F$, if $x \in TC_p Y$ then $x_{|[n] \setminus F} \in \mathcal{L}_{|F}^\perp \subset \mathbb{C}^{[n] \setminus F}$. On the other hand, by Theorem 2.1.(d) we have $TC_p \mathcal{L}^{-1} = TC_{p|F} \mathcal{L}_{|F}^{-1} \times \mathcal{L}_{|F}^{-1}$. Let us now consider $x \in TC_p \mathcal{L}^{-1} \cap TC_p Y$. We have $x_{|[n] \setminus F} \in \mathcal{L}_{|F}^\perp \cap \mathcal{L}_{|F}^{-1} = \{\mathbf{0}\}$, where the equality follows from Theorem 2.1.(b), and thus $x = x_{|F} \times \mathbf{0}$. That $x \in TC_p Y$ now implies that $x_{|F} \in TC_{p|F} Y(\mathcal{L}_{|F}, s_{|F}, d)$. Thus, we conclude $x_{|F} = \mathbf{0}$, since by Lemma 3.3 the intersection $\mathcal{L}_{|F}^{-1} \cap Y(\mathcal{L}_{|F}, s_{|F}, d)$ is smooth, and in particular transversal at $p_{|F}$, i.e. $TC_{p|F} \mathcal{L}_{|F}^{-1} \cap TC_{p|F} Y(\mathcal{L}_{|F}, s_{|F}, d) = \{\mathbf{0}\}$. \square

Combining Propositions 4.1 and 4.2 yields the following.

Corollary 4.4. *For a generic choice of $s \in \mathbb{C}^n$, we have*

$$d^r |\mu(M)| = \deg(\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)) = \sum_{\substack{F \subseteq [n] \\ \text{a flat of } M}} \mathcal{D}(\mathcal{L}_{|F}, d) |\mu(M/F)|.$$

5. Inclusion-exclusion

We now finish the proof of the main theorem by combining Corollary 4.4 with an inclusion-exclusion argument. For the facts regarding the lattice of flats of a matroid and the Möbius invariant used here, see [21].

Proof of Theorem 1.4. Write $\text{rk} : 2^{[n]} \rightarrow \mathbb{Z}$ for the rank function of the matroid M associated to \mathcal{L} . Let us recall that for a matroid M' of rank r' , one has $(-1)^{r'} \mu(M') = |\mu(M')|$. Then, Corollary 4.4 states that

$$(-d)^{\text{rk}[n]} \mu(M) = \sum_{\substack{F \subseteq [n] \\ \text{a flat of } M}} (-1)^{\text{rk}[n] - \text{rk} F} \mathcal{D}(\mathcal{L}_{|F}, d) \mu(M/F),$$

or more generally, one has, for any flat $F \subseteq [n]$ of M ,

$$(-d)^{\text{rk} F} \mu(M|F) = \sum_{\substack{F' \subseteq F \\ \text{a flat of } M}} (-1)^{\text{rk} F - \text{rk} F'} \mathcal{D}(\mathcal{L}_{|F'}, d) \mu(M|F/F').$$

As $\mu(M|F/F')$ is the value of the Möbius function $\mu(F', F)$ on the lattice of flats of M , applying the Möbius inversion formula [15, Proposition 3.7.1] (with $f(F) = d^{\text{rk}F} \mu(M|F)$ and $g(F) = (-1)^{\text{rk}F} \mathcal{D}(\mathcal{L}|_F, d)$) yields

$$(-1)^{\text{rk}F} \mathcal{D}(\mathcal{L}|_F, d) = \sum_{\substack{F' \subseteq F \\ \text{a flat of } M}} d^{\text{rk}F'} \mu(M|F').$$

Now, letting $F = [n]$ and noting $\text{rk}[n] = r$, we have

$$(-1)^r \mathcal{D}(\mathcal{L}, d) = d^r \sum_{\substack{F' \subseteq [n] \\ \text{a flat of } M}} \left(\frac{1}{d}\right)^{r-\text{rk}F'} \mu(M|F') = d^r \chi_M\left(\frac{1}{d}\right),$$

so that $\mathcal{D}(\mathcal{L}, d) = (-d)^r \chi_M\left(\frac{1}{d}\right) = d^r T_M\left(1 - \frac{1}{d}, 0\right)$ as desired. \square

Remark 5.1. For the non-reciprocal ML degree $\text{mld}(\mathcal{L})$, a standard computation similar to the one that gives Proposition 1.3 (see [16, 17]) yields

$$\text{mld}(\mathcal{L}) = \deg(\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d=0) \cap (\mathbb{C}^*)^{[n]}) = \mathcal{D}(\mathcal{L}, d=0).$$

One can hence recover [17, Corollary 3], which states $\text{mld}(\mathcal{L}) = |\mu(M)|$, by minor modifications of our arguments here. This case is in fact simpler, with no need for the consideration of tangent cones as was done in Proposition 4.2, because the intersection $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$ lies entirely in $(\mathbb{C}^*)^{[n]}$ when $d = 0$. We emphasize that $\mathcal{L}^{-1} \cap Y(\mathcal{L}, s, d)$ *never* lies entirely in $(\mathbb{C}^*)^{[n]}$ when $d \geq 1$.

Remark 5.2. Combining Theorem 1.4 with the “recipe formula” for Tutte-Grothendieck invariants of matroids (see for instance [20, Theorem 2.16]), one has that $\mathcal{D}(\mathcal{L}, d)$ for $d \geq 1$ satisfies the following deletion-contraction relation given an element $e \in [n]$:

$$\mathcal{D}(\mathcal{L}, d) = \begin{cases} 0 & \text{if } e \text{ is a loop,} \\ (d-1) \cdot \mathcal{D}(\mathcal{L}_{/e}, d) & \text{if } e \text{ is a coloop,} \\ \mathcal{D}(\mathcal{L}_{\setminus e}, d) + d \cdot \mathcal{D}(\mathcal{L}_{/e}, d) & \text{if } e \text{ neither loop nor coloop,} \end{cases}$$

where the base cases are $\mathcal{D}(\mathcal{L} = \mathbb{C}^{[1]}, d) = 1$ and $\mathcal{D}(\mathcal{L} = \{0\} \subset \mathbb{C}^{[1]}, d) = 0$. From this relation, it is easy to verify that the following are equivalent:

- (i) $\mathcal{D}(\mathcal{L}, 2) = \text{rml}(\mathcal{L}) = 1$,
- (ii) M is a partition matroid (i.e. every component of M has rank 1),
- (iii) $\text{mld}(\mathcal{L}) = |\mu(M)| = 1$, and
- (iv) \mathcal{L}^{-1} is linear.

It may be interesting to find a proof of Theorem 1.4 that directly reflects the deletion-contraction relation above geometrically.

Proof of Corollary 1.2. For an r -dimensional general subspace $\mathcal{L} \subseteq \mathbb{C}^{[n]}$, the associated matroid is the uniform matroid $U_{r,n}$, for which the Tutte polynomial has the following formula (see for instance [9])

$$T_{U_{r,n}}(x, y) = \sum_{i=1}^r \binom{n-i-1}{r-i} x^i + \sum_{j=1}^{n-r} \binom{n-j-1}{r-1} y^j.$$

Thus, for a general \mathcal{L} , Theorem 1.4 implies that

$$D(\mathcal{L}, d) = d^r T_{U_{r,n}}\left(1 - \frac{1}{d}, 0\right) = d^r \sum_{i=1}^r \binom{n-i-1}{r-i} \left(1 - \frac{1}{d}\right)^i.$$

Evaluating at $d = 2$ gives us

$$\text{rmlD}(\mathcal{L}) = 2^r \sum_{i=1}^r \binom{n-i-1}{r-i} \left(\frac{1}{2}\right)^i = \sum_{i=1}^r \binom{n-i-1}{r-i} 2^{r-i},$$

as desired. □

Remark 5.3. It is interesting to note that, for diagonal linear concentration models, our formula Theorem 1.1 implies that $\text{rmlD}(\mathcal{L})$ is always odd unless it is zero. The same pattern seems to persist for reciprocal ML degrees of general linear concentration models [16, Table 1].

Acknowledgements

We thank Yairon Cid Ruiz and Bernd Sturmfels for helpful conversations. We also thank the organizers of the Linear Spaces of Symmetric Matrices working group at MPI MiS Leipzig. C.E. is partially supported by the US National Science Foundation (DMS-2001854).

REFERENCES

- [1] Tobias Boege, Jane Ivy Coons, Christopher Eur, Aida Maraj, and Frank Röttger. Reciprocal maximum likelihood degrees of brownian motion tree models. *arXiv:2009.04334*, 2020.
- [2] Winfried Bruns and Jürgen Herzog. *Cohen-Macaulay rings*, volume 39 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1993.

- [3] Jane Ivy Coons, Orlando Marigliano, and Michael Ruddy. Maximum likelihood degree of the two-dimensional linear gaussian covariance model. *Algebraic Statistics*, (to appear), 2020.
- [4] David Eisenbud and Joe Harris. *3264 and all that—a second course in algebraic geometry*. Cambridge University Press, Cambridge, 2016.
- [5] Claudia Fevola, Yelena Mandelshtam, and Bernd Sturmfels. Pencils of quadrics: Old and new. *arXiv:2009.04334*, 2020.
- [6] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52.
- [7] June Huh. The maximum likelihood degree of a very affine variety. *Compos. Math.*, 149(8):1245–1266, 2013.
- [8] Laurent Manivel, Mateusz Michałek, Leonid Monin, Tim Seynnaeve, and Martin Vodička. Complete quadrics: Schubert calculus for gaussian models and semidefinite programming. *arXiv:2011.08791*, 2020.
- [9] Criel Merino, Marcelino Ramírez-Ibáñez, and Guadalupe Rodríguez-Sánchez. The Tutte polynomial of some matroids. *Int. J. Comb.*, pages Art. ID 430859, 40, 2012.
- [10] Mateusz Michałek, Leonid Monin, and Jarosław Wiśniewski. Maximum likelihood degree and space of orbits of a C^* action. *arXiv:2004.07735*, 2020.
- [11] Peter Orlik and Louis Solomon. Combinatorics and topology of complements of hyperplanes. *Invent. Math.*, 56(2):167–189, 1980.
- [12] James Oxley. *Matroid theory*, volume 21 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2 edition, 2011.
- [13] Nicholas Proudfoot and David Speyer. A broken circuit ring. *Beiträge Algebra Geom.*, 47(1):161–166, 2006.
- [14] Raman Sanyal, Bernd Sturmfels, and Cynthia Vinzant. The entropic discriminant. *Adv. Math.*, 244:678–707, 2013.
- [15] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
- [16] Bernd Sturmfels, Sascha Timme, and Piotr Zwiernik. Estimating linear covariance models with numerical nonlinear algebra. *Algebraic Statistics*, 11(1):31–52, 2020.
- [17] Bernd Sturmfels and Caroline Uhler. Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry. *Annals of the Institute of Statistical Mathematics*, 62(4):603–638, 2010.
- [18] Hiroaki Terao. Algebras generated by reciprocals of linear forms. *J. Algebra*, 250(2):549–558, 2002.
- [19] Dominic Welsh. *Matroid theory*. Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York, 1976. L. M. S. Monographs, No. 8.
- [20] Dominic Welsh. The Tutte polynomial. volume 15, pages 210–228. 1999. *Statistical physics methods in discrete probability, combinatorics, and theoretical computer science* (Princeton, NJ, 1997).

- [21] Thomas Zaslavsky. The Möbius function and the characteristic polynomial. In *Combinatorial geometries*, volume 29 of *Encyclopedia Math. Appl.*, pages 114–138. Cambridge Univ. Press, Cambridge, 1987.

CHRISTOPHER EUR

Department of Mathematics

Stanford University

e-mail: chriseur@stanford.edu

TARA FIFE

Max Planck Institute

for Mathematics in the Sciences

e-mail: fi.tara@gmail.com

JOSÉ ALEJANDRO SAMPER

Departamento de Matemáticas

Pontificia Universidad Católica de Chile

e-mail: jsamper@mat.uc.cl

TIM SEYNNAEVE

Mathematical Institute

University of Bern

e-mail: tim.seynnaeve@math.unibe.ch