The Linear Programming Degree

Madeleine Goertz

UC Berkeley

Seminar on Nonlinear Algebra
Max Planck Institute for Mathematics in the Sciences
July 22, 2026

(Sectional) LO Degrees

Fix $X \subset \mathbb{C}^n$ an irreducible affine variety with $\dim(X) = d$.

The linear optimization (LO) degree of X, $\mathrm{LOdeg}(X)$, is the number of critical points of a generic linear functional on $\mathbb{C}^n$ restricted to X_{sm}.

For $i = 0, \ldots, d$, the i-th sectional LO degree of X is

$$\sigma_i(X) := \mathrm{LOdeg}(X \cap H_1 \cap \cdots \cap H_i),$$

where $H_1, \ldots, H_i$ are generic affine hyperplanes in $\mathbb{C}^n$.

$\sigma_0(X) = \mathrm{LOdeg}(X)$, $\sigma_d(X) = \deg(X)$.

Application: Polyhedral Norm Optimization

Let $u \in \mathbb{C}^n$ be a point. Find the closest point on X to u in a *polyhedral norm*, a norm whose unit ball in the Euclidean metric is a polytope P.

One algorithm involves solving a linear optimization problem for each of the faces of P.

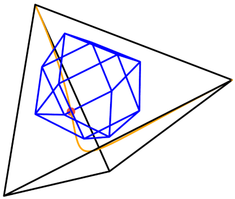

Figure: Wasserstein ball for discrete metric touching the twisted cubic at an edge.[1]

[1] [DeP+24, Figure 1]

Application: Polyhedral Norm Optimization

Çelik, Jamneshan, Montúfar, Sturmfels, Venturello: the algebraic degree of each facial linear optimization problem is bounded by a polar degree of X.[2]

For X an independence model, this bound is experimentally not tight.

Sectional LO degrees improve these bounds.

[2] [Çel+21]

Polar Degrees

$Y \subset \mathbb{P}^N$ irreducible projective variety. The conormal variety of Y:

$$N_Y = \overline{\{(x, H) \in \mathbb{P}^N \times (\mathbb{P}^N)^\vee \mid x \in Y_{sm} \text{ and } H \text{ is tangent to } Y \text{ at } x.\}}$$

$$\mathbb{P}^N \supset Y \xleftarrow{\mathrm{pr}_1} N_Y \xrightarrow{\mathrm{pr}_2} Y^\vee \subset (\mathbb{P}^N)^\vee$$

The polar degrees of Y, $\delta_i(Y)$, defined as bidegrees of N_Y:

$$[N_Y] = \sum_{i=0}^{\dim(Y)} \delta_i(Y)[\mathbb{P}^i \times \mathbb{P}^{N-1-i}] \in A_*(\mathbb{P}^N \times (\mathbb{P}^N)^\vee).$$

$L_{N-i} \subset \mathbb{P}^N$, $L_{i+1} \subset (\mathbb{P}^N)^\vee$ generic linear subspaces, $\dim(L_{N-i}) = N - i$, $\dim(L_{i+1}) = i + 1$. $L_{N-i} \times L_{i+1}$ intersects N_Y transversely at $\delta_i(Y)$ points.

$\delta_{\dim(Y)}(Y) = \deg(Y)$, $\delta_e(Y) = \deg(Y^\vee)$, $\delta_i(Y) = 0 \ \forall i < e$,
$e = \operatorname{codim}(Y^\vee) - 1$.

LO degree & Discriminants

Coordinates on $\mathbb{C}^n$: $x_1, \ldots, x_n$. $\overline{X} \subset \mathbb{P}^n$ the projective closure of X, obtained by homogenizing equations of X by x_0. Dual coordinates $a_0, \ldots, a_n$ on $(\mathbb{P}^n)^\vee$.

Theorem 1 (G. '26+)

Suppose $\overline{X}^\vee$ is a hypersurface with defining equation $\Delta_{\overline{X}}$, a homogeneous irreducible polynomial in the variables $a_0, \ldots, a_n$ of total degree $\delta_0(\overline{X})$. Then the following are equal:

1. *The linear optimization degree* $\mathrm{LOdeg}(X)$.
2. *The degree of $\Delta_{\overline{X}}$ in the variable a_0, denoted* $\deg_{a_0}(\Delta_{\overline{X}})$.

Rostalski & Sturmfels assume X defined over $\mathbb{R}$ and $X(\mathbb{R}) \subset \mathbb{R}^n$ irreducible, smooth, compact affine variety.[3]

[3] [RS10, Theorem 3.2]

Sectional LO degrees as bidegrees

The affine conormal variety of X:

$$T_X^* = \overline{\left\{(x, H) \in T^*\mathbb{C}^n \cong \mathbb{C}_x^n \times \mathbb{C}_u^n \mid x \in X_{sm} \text{ and } u\Big|_{T_xX} = 0\right\}}.$$

The LO bidegrees of Y, $\beta_i(X)$, defined as bidegrees of $\overline{T^*X}$:

$$[\overline{T_X^*}] = \sum_{i=0}^{d} \beta_i(X)[\mathbb{P}^i \times \mathbb{P}^{n-i}] \in A_*(\mathbb{P}^n \times \mathbb{P}^n).$$

Theorem 2 (Maxim-Rodriguez-Wang-Wu '24[4])

The sectional LO degrees and LO bidegrees coincide:
$\sigma_i(X) = \beta_i(X)$ *for all* $i = 0, \ldots, d$.

[4] [Max+24, Theorem 1.4]

The LO gaps

$$\mathbb{P}^n \supset \overline{X} \xleftarrow{\mathrm{pr}_1} N_{\overline{X}} \xrightarrow{\mathrm{pr}_2} \overline{X}^\vee \subset (\mathbb{P}^n)^\vee$$

$H_\infty = \{x_0 = 0\} \in (\mathbb{P}^n)^\vee$ is the hyperplane at infinity.

Set $B = \mathbb{P}^n \times \{H_\infty\}$.

$F = \mathrm{pr}_2^{-1}(H_\infty) \subset B \subset N_{\overline{X}}$ is the fiber of pr_2 over H_∞.

The *i-th LO gap* $\mu_i(X)$ are the coefficients of the Chow class of $F \subset B \cong \mathbb{P}^n$:

$$[F] = \sum_{i=0}^{d} \mu_i(X)[\mathbb{P}^i] \in A_*(B).$$

The LO threshold & LO gap

Theorem 3 (G. '26+)

1. $\sigma_i(X) + \mu_i(X) = \delta_i(\overline{X})$ *for all* $i = 0, \ldots, d$.
2. $H_\infty \notin \overline{X}^\vee$ *if and only if* $\mu_i(X) = 0$ *for all* $i = 0, \ldots, d$.
3. *Let* $d_F = \dim(F)$.
   - $\mu_i(X) = 0$ *for all* $i > d_F$, *and* $\mu_i(X) > 0$ *for all* $i \leq d_F$.
   - *If* $H_\infty \in \overline{X}^\vee$ *then* $\operatorname{codim}(\overline{X}^\vee) \leq d_F \leq d$.

Maxim-Rodriguez-Wang-Wu prove (2).[5]

The LO threshold of X is $\mathrm{LOthresh}(X) = d_F + 1$.

[5] [Max+24, Proposition 6.2]

Connection to Gröbner Degeneration

Let $I = I(N_{\overline{X}})$. If $f \in I$, $\mathrm{Lt}_w(I)$ is sum of terms of f where a_0 has degree $\deg_{a_0}(f)$. Set $\mathrm{Lt}_w(I) = \langle \mathrm{Lt}_f(I) \mid f \in I \rangle$.

$Y = V(\mathrm{Lt}_w(I))$, $Y_\infty = Y \cap B$, $Y_{\mathrm{aff}} = \overline{Y \setminus Y_\infty}$.

Corollary 4 (G. '26+)

1. $[Y] = [Y_\infty] + [Y_{\mathrm{aff}}]$.
2. $[Y] = [N_{\overline{X}}] = \sum_{i=0}^{d} \delta_i(X)[\mathbb{P}^i \times \mathbb{P}^{n-1-i}] \in A_*(\mathbb{P}^n \times (\mathbb{P}^n)^\vee)$.
3. $[Y_\infty] = [F] = \sum_{i=0}^{d} \mu_i(X)[\mathbb{P}^i] \in A_*(B)$.
4. $[Y_{\mathrm{aff}}] = \sum_{i=0}^{d} \sigma_i(X)[\mathbb{P}^i \times \mathbb{P}^{n-i}] \in A_*(\mathbb{P}^n \times (\mathbb{P}^n)^\vee)$.

Computing Sectional LO Degrees of Parametric Varieties

Suppose X has a polynomial parametrization:

$$\varphi = (\varphi_1, \ldots, \varphi_n) \colon \mathbb{C}^d \to \mathbb{C}^n,$$

where each $\varphi_i(u) \colon \mathbb{C}[u_1, \ldots, u_d]$ is a polynomial such that X is the Zariski closure of the image of φ.

Consider this optimization problem:

$$\min_{u \in \mathbb{C}^d} f_0(\varphi(u)) \quad \text{subject to} \quad f_1(\varphi(u)) = \cdots = f_i(\varphi(u)) = 0,$$

where $f_0(x)$ is a generic linear functional on $\mathbb{C}^n$, and $f_1(x), \ldots, f_i(x)$ are the equations of generic affine hyperplanes

The Lagrangian $\mathcal{L}(u,\lambda) = f_0(\varphi(u)) - \sum_{j=1}^{i} \lambda_j f_j(\varphi(u))$ determines a Lagrange system:

$$L = (f_1(\varphi(u)), \ldots, f_i(\varphi(u)), \ell_1, \ldots, \ell_d), \quad \ell_j = \frac{\partial \mathcal{L}}{\partial u_j}.$$

Theorem 5 (Lindberg-Monin-Rose '24[6])

If the system L is a generic sparse system of polynomials with strongly admissible support, then all critical points of the optimization problem are smooth and lie in the algebraic torus $(\mathbb{C}^*)^n$. The number of critical points is equal to the mixed volume $MV(\mathrm{Newt}(f_1(\varphi(u))), \ldots, \mathrm{Newt}(f_i(\varphi(u))), \mathrm{Newt}(\ell_1), \ldots, \mathrm{Newt}(\ell_d))$

Corollary 6

Under the assumptions of this theorem, if φ is generically p-to-1, then this mixed volume equals $p \cdot \sigma_i(X)$.

[6] [LMR24, Theorem B]

Independence Models

Let $m, r \in \mathbb{N}^k$ with each $m_i \geq 2$. Let $(m_i)_{r_i}$ denote a multinomial distribution with m_i possible outcomes and r_i trials. Consider the joint probability distributions of k independent multinomial distributions $(m_1)_{r_1}, \ldots, (m_k)_{r_k}$ as points in $\mathbb{C}^n$, where $n = \prod_{j=1}^{k} \binom{m_j + r_j - 1}{r_j}$.

The independence model $\mathcal{M}_{m,r} \subset \mathbb{C}^n$ is the Zariski closure of these distributions. $\dim(\mathcal{M}_{m,r}) = |m| - k$.

The Segre-Veronese variety $\mathcal{V}_{m,r} \subset \mathbb{P}^{n-1}$ is the image of the embedding

$$\mathbb{P}^{m_1 - 1} \times \cdots \times \mathbb{P}^{m_k - 1} \hookrightarrow \mathbb{P}^{n-1}$$

given by the complete linear system $|\mathcal{O}(r_1, \ldots, r_k)|$.

Lemma 7

$\delta_i(\overline{\mathcal{M}_{m,r}}) = \delta_i(\mathcal{V}_{m,r})$ for all $i = 0, \ldots, d$. When defined, $\Delta_{\overline{\mathcal{M}_{m,r}}}$ and $\Delta_{\mathcal{V}_{m,r}}$ differ by a specific linear change of variables.

Example: $(2, 2)$

Let $k = 2$, $m = (2, 2)$, $r = (1, 1)$, so $d = 2$ and $n = 4$. The joint probability distributions are

$$(pq, p(1-q), (1-p)q, (1-p)(1-q)) \in \mathbb{C}^4 \quad \forall\, p, q \in [0, 1].$$

The ideal of $\mathcal{M}_{m,r} \subset \mathbb{C}^n$ is generated by

$$x_1x_4 - x_2x_3, x_1 + x_2 + x_3 + x_4 - 1.$$

The ideal of $\mathcal{V}_{m,r} \subset \mathbb{P}^{n-1}$ is generated by

$$x_1x_4 - x_2x_3.$$

$\mathcal{V}_{m,r}$ is the Segre embedding $\mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^3$.

Example: (2_3)

Let $k = 1$, $m = (2)$, $r = (3)$, so $d = 1$ and $n = 4$. The joint probability distributions are

$$(p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3) \in \mathbb{C}^4 \quad \forall\ p \in [0,1].$$

The ideal of $\mathcal{M}_{m,r} \subset \mathbb{C}^n$ is generated by

$$3x_1x_3 - x_2^2, 9x_1x_4 - x_2x_3, 3x_2x_4 - x_3^2, x_1 + x_2 + x_3 + x_4 - 1.$$

The ideal of $\mathcal{V}_{m,r} \subset \mathbb{P}^{n-1}$ is generated by

$$x_1x_3 - x_2^2, x_1x_4 - x_2x_3, x_2x_4 - x_3^2.$$

$\mathcal{V}_{m,r}$ is the twisted cubic, the degree 3 Veronese embedding $\mathbb{P}^1 \hookrightarrow \mathbb{P}^3$.

For the independence model $\mathcal{M}_{m,r}$, write $l = (m_1 - 1, \ldots, m_k - 1)$.
Let $i = 0, \ldots, |l| = \sum_{i=1}^{k} l_i = \dim(\mathcal{M}_{m,r})$.

Theorem 8 (Sodomaco '20[7])

The i-th polar degree of the Segre-Veronese variety $\mathcal{V}_{m,r}$ is

$$\delta_i(\mathcal{V}_{m,r}) = \sum_{j=0}^{|l|-i} (-1)^j \binom{|l|+1-j}{i+1} \sum_{\substack{\gamma=(\gamma_s) \\ 0\le\gamma_s\le l_s \\ |\gamma|=j}} \binom{|l|-j}{l_1-\gamma_1, \ldots, l_k-\gamma_k} \prod_{s=1}^{k} \binom{l_s+1}{\gamma_s} r_s^{l_s-\gamma_s}.$$

Theorem 9 (G. '26+)

The i-th sectional LO degree of the independence model $\mathcal{M}_{m,r}$ is

$$\sigma_i(\mathcal{M}_{m,r}) = \sum_{j=0}^{|l|-i} (-1)^j \binom{|l|-j}{i} \sum_{\substack{\gamma=(\gamma_s) \\ 0\le\gamma_s\le l_s \\ |\gamma|=j}} \binom{|l|-j}{l_1-\gamma_1, \ldots, l_k-\gamma_k} \prod_{s=1}^{k} \binom{l_s}{\gamma_s} r_s^{l_s-\gamma_s}.$$

[7] Published in [Koz+25, Proposition 6.7].

Example: $(2,2)$

For $m = (2,2)$, $r = (1,1)$, $\mathcal{M}_{m,r} \subset \mathbb{C}^n$ for $n = 4$ has dimension $d = 2$.

Table: The sectional LO degrees σ_i, LO gaps μ_i, and polar degrees δ_i of the independence model $(2,2)$.

i	$\sigma_i(\mathcal{M}_{m,r})$	$\mu_i(\mathcal{M}_{m,r})$	$\delta_i(\overline{\mathcal{M}_{m,r}})$
0	1	1	2
1	2	0	2
2	2	0	2

$\deg(\mathcal{M}_{m,r}) = \sigma_d = \delta_d = 2$, $\operatorname{codim}(\overline{\mathcal{M}_{m,r}}^{\vee}) = 1$, $\operatorname{LOdeg}(\mathcal{M}_{m,r}) = \sigma_0 = 1$, and $\operatorname{LOthresh}(\mathcal{M}_{m,r}) = 1$.

Example: $(2,2)$

$\Delta_{\overline{\mathcal{M}_{m,r}}}$ is a polynomial in $a_0, \ldots, a_4$ with 6 terms of total degree 2 and its degree in a_0 is 1:

$$(a_1 - a_2 - a_3 + a_4) \cdot a_0 + a_1 a_4 - a_2 a_3$$

The LO degree of $\mathcal{M}_{m,r}$ is $\sigma_0(\mathcal{M}_{m,r}) = 1$.

$\mu_0(\mathcal{M}_{m,r}) = 1$. $\delta_0(\overline{\mathcal{M}_{m,r}}) = 2$.

Example: (2_3)

For $m = (2)$, $r = (3)$, $\mathcal{M}_{m,r} \subset \mathbb{C}^n$ for $n = 4$ has dimension $d = 1$.

Table: The sectional LO degrees σ_i, LO gaps μ_i, and polar degrees δ_i of the independence model (2_3).

i	$\sigma_i(\mathcal{M}_{m,r})$	$\mu_i(\mathcal{M}_{m,r})$	$\delta_i(\overline{\mathcal{M}_{m,r}})$
0	2	2	4
1	3	0	3

$\deg(\mathcal{M}_{m,r}) = \sigma_d = \delta_d = 3$, $\mathrm{codim}(\overline{\mathcal{M}_{m,r}}^{\vee}) = 1$,
$\mathrm{LOdeg}(\mathcal{M}_{m,r}) = \sigma_0 = 2$, and $\mathrm{LOthresh}(\mathcal{M}_{m,r}) = 1$.

Example: (2_3)

$\Delta_{\overline{\mathcal{M}_{m,r}}}$ is a polynomial in $a_0, \ldots, a_4$ with 27 terms of total degree 4 and its degree in a_0 is 2:

$$
\begin{aligned}
&(a_1^2 - 6a_1a_2 + 9a_2^2 + 6a_1a_3 - 18a_2a_3 \\
&\qquad + 9a_3^2 - 2a_1a_4 + 6a_2a_4 - 6a_3a_4 + a_4^2) \cdot a_0^2 \\
&+ (4a_2^3 - 6a_1a_2a_3 - 6a_2^2a_3 + 12a_1a_3^2 - 6a_2a_3^2 + 4a_3^3 + 2a_1^2a_4 \\
&\quad - 6a_1a_2a_4 + 12a_2^2a_4 - 6a_1a_3a_4 - 6a_2a_3a_4 + 2a_1a_4^2) \cdot a_0 \\
&\qquad - 3a_2^2a_3^2 + 4a_1a_3^3 + 4a_2^3a_4 - 6a_1a_2a_3a_4 + a_1^2a_4^2
\end{aligned}
$$

The LO degree of $\mathcal{M}_{m,r}$ is $\sigma_0(\mathcal{M}_{m,r}) = 2$.

$\mu_0(\mathcal{M}_{m,r}) = 2$. $\delta_0(\overline{\mathcal{M}_{m,r}}) = 4$.

The $2 \times \cdots \times 2$ Hyperdeterminant & Derangements

For $m = (2, \ldots, 2), r = (1, \ldots, 1) \in \mathbb{Z}^k$, $\mathcal{M}_{m,r}$ is the k-bit independence model. The discriminant $\Delta_k = \Delta_{\mathcal{V}_{m,r}}$ is the *hyperdeterminant of format* $2 \times \cdots \times 2$.[8] The total degree of Δ_k is

$$\delta_0(\mathcal{V}_{m,r}) = \sum_{j=0}^{k} (-1)^j (k+1-j) \binom{k}{j} (k-j)! 2^j.$$

$0, 2, 4, 24, 128, 880, 6816, 60032$ for $k = 1, \ldots, 8$.

[8] See [GKZ94, Chapter 14].

Corollary 10

The degree in the variable a_0 of Δ_k is

$$\mathrm{LOdeg}(\mathcal{M}_{m,r}) = \sigma_0(\mathcal{M}_{m,r}) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)!.$$

$0, 1, 2, 9, 44, 265, 1854$ for $k = 1, \ldots, 8$. $\mathrm{LOdeg}(\mathcal{M})$ is the number of derangements of size k.

Counting Nash Equilibria

Corollary 11

Let $m = (m_1, \ldots, m_k)$*,* $r = (1, \ldots, 1) \in \mathbb{Z}^k$*. The maximum number of totally mixed Nash equilibria in a* k*-player game where player* i *has exactly* m_i *pure strategies is* $\mathrm{LOdeg}(\mathcal{M}_{m,r})$*.*

Next Steps

- Find a combinatorial formula for the leading coefficient term $\mathrm{Lt}_w(I)$, where I is the ideal of the conormal variety $N_{\overline{X}}$.
- Find explicit formulas for the sectional LO degrees of other classes of parametric varieties.
- Characterize when Wasserstein degrees are less than sectional LO degrees.
- ...

Thank you!

[Çel+21] Türkü Özlüm Çelik et al. "Wasserstein Distance to Independence Models". In: *J. Symb. Comput.* 104 (2021), pp. 855–873. ISSN: 0747-7171,1095-855X. DOI: 10.1016/j.jsc.2020.10.005.

[DeP+24] Greg DePaul et al. "Degrees of the Wasserstein Distance to Small Toric Models". In: *Algebr. Stat.* 15.2 (Dec. 2024), pp. 249–267. ISSN: 2693-3004. DOI: 10.2140/astat.2024.15.249. (Visited on 06/10/2026).

[GKZ94] I. M. Gel'fand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants.* Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1994. ISBN: 978-0-8176-3660-9. DOI: 10.1007/978-0-8176-4771-1.

[Koz+25] Khazhgali Kozhasov et al. "On the Minimal Algebraic Complexity of the Rank-One Approximation Problem for General Inner Products". In: *Math. Comput.* (Dec. 2025). ISSN: 1088-6842, 0025-5718. DOI: 10.1090/mcom/4176. (Visited on 06/02/2026).

References II

[LMR24] Julia Lindberg, Leonid Monin, and Kemal Rose. *The Algebraic Degree of Sparse Polynomial Optimization*. 2024. arXiv: 2308.07765.

[Max+24] Laurentiu G. Maxim et al. "Linear Optimization on Varieties and Chern-Mather Classes". In: *Adv. Math.* 437 (2024), Paper No. 109443, 22. ISSN: 0001-8708,1090-2082. DOI: 10.1016/j.aim.2023.109443.

[RS10] Philipp Rostalski and Bernd Sturmfels. "Dualities in Convex Algebraic Geometry". In: *Rendiconti Mat. E Delle Sue Appl. Ser. VII* 30.3-4 (2010), pp. 285–327. ISSN: 1120-7183,2532-3350.

Strongly Admissible Support

$\mathcal{A} = (\mathcal{A}_0, \ldots, \mathcal{A}_i)$ where $\mathcal{A}_j \in \mathbb{N}^n$ and $P_j := \mathrm{conv}(\mathcal{A}_j)$ is admissible if for every $j = 0, \ldots, i$:

1. $\mathbb{R}^n_{\geq 0}$ is a cone in a normal fan of the polytope P_j, and
2. The polytope P_j meets every coordinate hyperplane of $\mathbb{R}^n$.

Suppose $\mathcal{A}$ is admissible. The proper normal toric variety X with polyhedral fan Σ is appropriate for $\mathcal{A}$[9] if

1. The normal fan $\Sigma(P_j)$ is refined by Σ for each $j = 0, \ldots, i$;
2. The nonnegative orthant $\mathbb{R}^n_{\geq 0}$ is a cone in the fan Σ; and
3. For generic functions f_j with monomial support $\mathcal{A}_j$, $\overline{V(f_1, \ldots, f_i)} \cap X_{\mathrm{sing}} = \varnothing$.

A point configuration $\mathcal{A}$ is strongly admissible if $\mathcal{A}$ is admissible and if the toric variety X given by the normal fan $\Sigma(P_0 + \cdots + P_i)$ is appropriate for $\mathcal{A}$.[10]

[9] Guarantees smoothness

[10] [LMR24, Definitions 3.1, 3.3, 3.4]