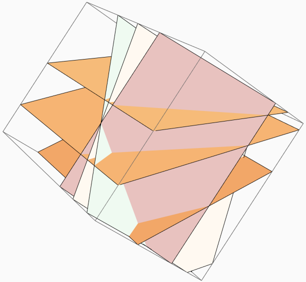
Logarithmic Discriminants of Hyperplane Arrangements

SIAM AG25 – MS59 Discriminants in the Sciences

MAX PLANCK INSTITUTE
FOR MATHEMATICS
IN THE SCIENCES



Leonie Kayser, Andreas Kretschmer & Simon Telen

leokayser.github.io

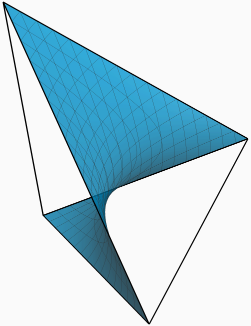
July 9, 2025

Maximum Likelihood Estimation in Algebraic Statistics

- Let $X \subseteq (\mathbb{C}^\times)^{n+1}$ be a d-dimensional smooth variety
- Discrete statistical model $X \cap \Delta_n = \{\, p \in X \cap \mathbb{R}^{n+1} \mid p_i > 0,\ p_0 + \cdots + p_n = 1 \,\}$
- Given data points $u \in \mathbb{N}^{n+1}$, which parameter maximizes the log-likelihood function

$$\mathcal{L}_u(x) = \log x_0^{u_0} \cdots x_n^{u_n}, \qquad x \in X \cap \Delta_n?$$

- Critical equations: $x \in \mathrm{Crit}_X(u) := \{\, x \in X \mid \nabla \mathcal{L}_u(x) = 0 \,\}$
- $\mathrm{Crit}_X(u)$ is a *finite* set of $\mathrm{MLdeg}(X)$ non-degenerate critical points for *general* data $u \in \mathbb{N}^{n+1}$ (or $\mathbb{C}^{n+1}$)
- [Huh13] $\mathrm{MLdeg}(X) = |\chi(X)|$
- Extensively studied for toric models (exponential families), linear models, determinantal varieties, ...



Linear models and scattering amplitudes

- Let $\mathcal{A}$ be an essential arrangement of $n+1$ hyperplanes in $\mathbb{C}^d$

$$\mathcal{A} = \mathbb{V}(\ell_0) \cup \cdots \cup \mathbb{V}(\ell_n) \subseteq \mathbb{C}^d, \qquad (\ell_0(x), \ldots, \ell_n(x))^{\mathsf{T}} = Ax + b, \quad L^{\mathsf{T}} = [b \mid A]$$

- Parametrizes linear model $X := \mathbb{C}^d \setminus \mathcal{A} \overset{\ell}{\hookrightarrow} (\mathbb{C}^\times)^{n+1}$, can assume $\sum_j \ell_j = 1$
- Log-likelihood function or master function given by

$$\mathcal{L}_u(x) = u_0 \log \ell_0(x) + \cdots + u_n \log \ell_n(x), \qquad \nabla \mathcal{L}_u(x) = A^{\mathsf{T}} \operatorname{diag}(1/\ell_0, \ldots, 1/\ell_n) u$$

- Varchenko: For real arrangements, $\mathrm{MLdeg}(X) = \#$bounded chambers of $\mathcal{A} \cap \mathbb{R}^d$
- Critical equations appear as scattering equations in bi-adjoint scalar ϕ^3-theories (Cachazo, He & Yuan [CHY14])

What is “general” data?

- Moving from general to special $u \in \mathbb{P}^n = \mathbb{P}(\mathbb{C}^{n+1})$, what can happen to $\mathrm{Crit}_X(u)$?
  1. Two critical points collide to form a non-reduced/degenerate point
  2. A positive-dimensional component appears
  3. A critical point disappears to infinity
- Outside of 1.-3. the finite set of critical points has constant size $\mathrm{MLdeg}(X)$
- The closure of 1.-3. was called the *data discriminant* by Rodriguez & Tang [RT15]
- 3. was studied by Sattelberger & van der Veer [SvdV23]

Definition (Ad-hoc definition of $\nabla_{\log}(X)$)

The *logarithmic discriminant* of a (smooth) variety $X \hookrightarrow (\mathbb{C}^\times)^{n+1}$ is

$$\nabla_{\log}(X) := \overline{\{\, u \in \mathbb{P}^n \mid \mathrm{Crit}_X(u) \text{ is infinite or non-reduced} \,\}}.$$

⇝ Goal: Understand logarithmic discriminants of hyperplane arrangements!

Three points enter a bar

- Three points on a line $\mathcal{A} = \{0, 1, b\} \subseteq \mathbb{C}^1$
- Model is a line $X = \mathbb{C}^1 \setminus \mathcal{A} \hookrightarrow (\mathbb{C}^\times)^3$ parametrized by $(x, x-1, x-b)$,

$$\mathcal{L}_u(x) = u_0 \log x + u_1 \log(x-1) + u_2 \log(x-b)$$

- Single critical equation in $x \in \mathbb{C}^1 \setminus \mathcal{A}$

$$\frac{u_0}{x} + \frac{u_1}{x-1} + \frac{u_2}{x-b} = 0 \quad \Longleftrightarrow \quad u_0(x-1)(x-b) + u_1 x(x-b) + u_2 x(x-1) = 0$$

- When does this quadric in x have a double root? Highschool discriminant vanishes!

$$\Delta_{\log}(X) = (b-1)^2\, u_0^2 + 2b(b-1)\, u_0 u_1 + b^2\, u_1^2 - 2(b-1)\, u_0 u_2 + 2b\, u_1 u_2 + u_2^2$$

- $\Delta_{\log}(X)$ itself is a smooth quadric in u with discriminant $-4b^2(b-1)^2$

Ramification and its consequences

- Let $f\colon V \to W$ be a dominant map of smooth irreducible varieties of dimension n
- The ramification locus $\mathrm{Ram}(f) \subseteq V$ is the hypersurface

$$\mathrm{Ram}(f) = \left\{\, x \in V \;\middle|\; x \in f^{-1}(f(x)) \text{ is not isolated or reduced} \,\right\} = \mathbb{V}(\det J_f(x))$$

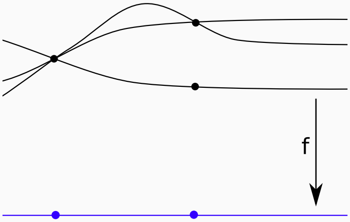
- The branch locus is the image closure $\mathrm{Branch}(f) = \overline{f(\mathrm{Ram}(f))} \subseteq W$
- Apply this to the likelihood correspondence

$$f\colon \mathcal{L}_X^\circ := \{\, (u, x) \in \mathbb{P}^n \times X \mid \nabla \mathcal{L}_u(x) = 0 \,\} \to \mathbb{P}^n$$

Definition (True definition of $\nabla_{\log}(X)$)

The logarithmic discriminant is the branch locus of the projection f. The ramification locus is defined in $\mathbb{P}^n \times X$ by

$$\nabla \mathcal{L}_u(x) = 0, \quad \det \mathrm{Hess}_x(\mathcal{L}_u(x)) = 0.$$



The ramification locus of a linear model

- $X = \mathbb{C}^d \setminus \mathbb{V}(\ell_0 \cdots \ell_n), \quad (\ell_0(x), \ldots, \ell_n(x))^\mathsf{T} = Ax + b$
- Here the equations of the ramification locus have a very concrete form

$$\nabla \mathcal{L}_u(x) = A^\mathsf{T} \cdot \mathrm{diag}(1/\ell_0, \ldots, 1/\ell_n) \cdot u = 0$$

$$h = \det\left(A^\mathsf{T} \cdot \mathrm{diag}\left(\frac{u_0}{\ell_0^2}, \ldots, \frac{u_n}{\ell_n^2}\right) \cdot A\right) = \sum_{\substack{I \subseteq \{0,\ldots,n\} \\ |I|=d}} |A_I|^2 \frac{u^I}{(\ell^I)^2}$$

- Critical equations are linear in the u_j $\rightsquigarrow$ substitute them in h to obtain

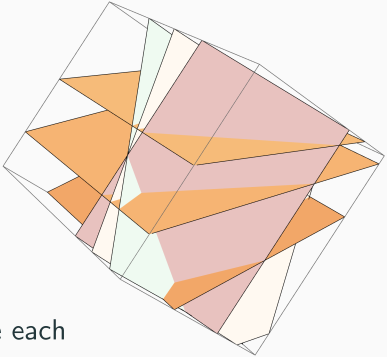
$$\tilde{h} \in \mathbb{C}[u_d, \ldots, u_n; x], \qquad \mathrm{Ram}(f) \cong \mathbb{V}(\tilde{h}) \subseteq \mathbb{P}^{n-d} \times X$$

- $\mathcal{A}$ is bi-uniform if both $\mathrm{M}(A^\mathsf{T})$ and $\mathrm{M}([b|A]^\mathsf{T})$ are uniform matroids

Theorem (Irreducibility of $\mathrm{Ram}(f)$)

If the arrangement contains a subset of $d+2$ hyperplanes which is bi-uniform, then $\mathrm{Ram}(f)$ and hence $\nabla_{\log}(X)$ are irreducible varieties.

A split discriminant!



- Consider the arrangement $\mathcal{A}$ of six planes

$$\ell = (1, x_1, x_2, x_3) \cdot \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 2 & 1 & 0 & 1 \\ 1 & \frac{3}{2} & \frac{3}{2} & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix}$$

- The first and the last three planes intersect in a line each
- The logarithmic discriminant decomposes as

$$\begin{aligned} \nabla_{\log}(X) = {} & \mathbb{V}(144u_0^2 + 120u_0u_1 + 168u_0u_2 + 25u_1^2 - 70u_1u_2 + 49u_2^2) \\ & \cup \mathbb{V}(u_3^2 - 2u_3u_4 + 4u_3u_5 + u_4^2 + 4u_4u_5 + 4u_5^2) \\ & \cup \mathbb{V}(u_0 + u_1 + u_2, u_3 + u_4 + u_5). \end{aligned}$$

A complete answer in $\mathbb{C}^1$

Theorem

Let $\mathcal{A} \subseteq \mathbb{C}^1$ be an arrangement of $n+1 \geq 3$ distinct points.

1. *The ramification locus is a smooth irreducible hypersurface in $\mathcal{L}_X^\circ \subset \mathbb{P}^n \times (\mathbb{C}^1 \setminus \mathcal{A})$.*
2. *Its class in the Chow ring $\mathrm{A}^\bullet(\mathbb{P}^n \times \mathbb{P}^1) = \mathbb{Z}[\alpha, \beta]/\langle \alpha^{n+1}, \beta^2 \rangle$ is $\alpha^2 + 2(n-1)\alpha\beta$.*
3. *The projection $f\colon \mathrm{Ram}(f) \to \nabla_{\log}(X)$ is generically bijective.*
4. *$\nabla_{\log}(X) \subseteq \mathbb{P}^n$ is an irreducible hypersurface of degree $2(n-1)$.*

- Explicit formula for defining polynomial ($\mathcal{A} = \{b_0, \ldots, b_n\}$)

$$\Delta_{\log}(X) = \mathrm{Disc}_x\left(\sum_{i=0}^{n} u_i \prod_{k \neq i} (x - b_k)\right)$$

- For $n+1 = 4$ points $\nabla_{\log}(X) \subseteq \mathbb{P}^3$ is always a singular surface of degree 4

The Hurwitz Discriminant and general arrangements

Theorem

Let $\mathcal{A}$ be a bi-uniform arrangement of $n+1 \geq d+2$ hyperplanes in $\mathbb{C}^d$.

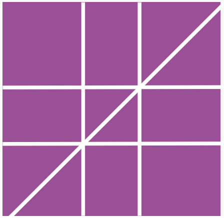
1. $\nabla_{\log}(X)$ *is an irreducible and reduced hypersurface.*
2. $\nabla_{\mathrm{Hu}}(X)$ *is a hypersurface of degree* $2d\binom{n-1}{d}$ *with full Newton polytope*
3. $\nabla_{\log}(X) \subseteq \nabla_{\mathrm{Hu}}(X)$ *coincide as sets, so* $\Delta_{\mathrm{Hu}}(X) = \Delta_{\log}(X)^e$ *for some* $e \geq 1$.
4. *If the arrangement is defined by real affine linear forms, then* $\nabla_{\log} \cap \mathbb{R}_+^{n+1} = \emptyset$.

- Main tool: Hurwitz discriminant $\nabla_{\mathrm{Hu}}(X) \supseteq \nabla_{\log}(X)$
  1. Reciprocal linear space $\mathcal{R} := \overline{\mathrm{Im}(\ell_0^{-1} : \cdots : \ell_n^{-1})} \subseteq \mathbb{P}^n$
  2. Hurwitz form $\mathcal{Z}_1(\mathcal{R}) \subseteq \mathrm{Gr}(n-d, \mathbb{P}^n)$, $\deg \mathcal{Z}_1(\mathcal{R}) = 2(n-d)\binom{n}{d-1}$
  3. $\nabla_{\mathrm{Hu}} := \varphi^{-1}(\mathcal{Z}_1(\mathcal{R}))$, pullback along $\varphi\colon \mathbb{P}^n \dashrightarrow \mathrm{Gr}(n-d, \mathbb{P}^n)$, $u \mapsto \mathrm{Ker}(A^\top \operatorname{diag}(u))$
- Key idea: $x \in \mathrm{Crit}_X(u)$ if and only if $(\ell_0^{-1}(x) : \cdots : \ell_n^{-1}(x)) \in \varphi(u) \cap \mathcal{R}$

The discriminant of $\mathcal{M}_{0,m}$

- $\mathcal{M}_{0,m}$ parametrizes tuples of m distinct points on $\mathbb{P}^1$
- Fixing $(0, 1, x_1, \ldots, x_{m-3}, \infty)$, it can be realized in $\mathbb{C}^{m-3}$ as the complement of the minors of

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & x_1 & x_2 & \cdots & x_{m-3} & 1 \end{bmatrix}$$



- Mandelstam variables s_{ij} corresponding to minor (i, j)
- Discriminant for $m = 5$ has degree $4 < 2 \cdot 2 \cdot \binom{5-2}{2} = 12$

$$\Delta_{\log}(\mathcal{M}_{0,5}) = (s_{13}s_{24} + s_{13}s_{34} + s_{14}s_{34} + s_{14}s_{23} + s_{23}s_{34} + s_{24}s_{34} + s_{34}^2)^2 - 4s_{13}s_{14}s_{23}s_{24}$$

- The Hurwitz discriminant has the extra factors

$$\Delta_{\mathrm{Hu}}(\mathcal{M}_{0,5}) = (s_{13} + s_{23} + s_{34})^2 \cdot (s_{14} + s_{24} + s_{34})^2 \cdot \Delta_{\log}(\mathcal{M}_{0,5})$$

- Conjecturally rich nested structure, degrees of $\nabla_{\log}(\mathcal{M}_{0,m})$ are $4, 30, 208, 1540, \ldots$

Beyond hyperplane arrangements

- Let $f_0, \dots, f_n \in \mathbb{C}[x]$ be polynomials parametrizing a model

$$X \cong \mathbb{C}^d \setminus \mathbb{V}(f_0 \cdots f_n) \hookrightarrow (\mathbb{C}^\times)^{n+1}, \qquad x \mapsto (f_0(x), \dots, f_n(x))$$

- Case $(f, x_1, \dots, x_d)$ closely related to toric models
- $d = 1$: $\nabla_{\log}(X)$ is an irreducible hypersurface of degree $2(\# \mathbb{V}(f_0 \cdots f_n) - 2)$
- Consider a family of conics $X_z \subset \mathbb{A}^2_{\mathbb{C}[z]}$ degenerating to two lines as $z \to 0$

$$f_0 = (x_1 + x_2 + 1)(-x_1 + x_2 - 2) + z, \quad f_1 = x_1, \quad f_2 = x_2$$

- X_0 is a bi-uniform arrangement of 4 lines, hence $\deg \nabla_{\log}(X_0) = 2 \cdot 2 \cdot \binom{4-2}{2} = 4$
- $\nabla_{\log}(X_z)$ has degree 6, $\Delta_{\log}(X_z)|_{z=0} = u_0^2 \cdot \Delta_{\log}(X_0)$
- The discriminant is factor of the tact invariant $\Delta_{\log}(X_z) = \frac{1}{u_0^6} \cdot \mathrm{tact}_x(\mathcal{L}_X^\circ)$

Many open questions - even for linear models

- Missing piece of the puzzle: Is ∇_{Hu} reduced for a bi-uniform arrangement?
- (When) is the projection $\mathrm{Ram}(f) \to \nabla_{\log}$ generically one-to-one? …bijective?
- Is there any arrangement such that $\nabla_{\log}(\mathcal{A})$ is *not* reduced?
- Is there an arrangement of *lines* whose $\nabla_{\log}(\mathcal{A})$ is reducible?
- Is the degree of $\nabla_{\log}(\mathcal{A})$ an invariant of the matroids?
- What is the meaning of the components $\nabla_{\mathrm{Hu}} \setminus \nabla_{\log}$?

Thank you! Questions?

arXiv:2410.11675

References i

Freddy Cachazo, Song He, and Ellis Ye Yuan.
Scattering equations and Kawai-Lewellen-Tye orthogonality.
Physical Review D, 90(6):065001, 2014.

June Huh.
The maximum likelihood degree of a very affine variety.
Compositio Mathematica, 149(8):1245–1266, 2013.

Leonie Kayser, Andreas Kretschmer, and Simon Telen.
Logarithmic discriminants of hyperplane arrangements, 2024.

References ii

Jose Israel Rodriguez and Xiaoxian Tang.
Data-discriminants of likelihood equations.
In *Proceedings of the 2015 ACM on international symposium on symbolic and algebraic computation*, pages 307–314, 2015.

Anna-Laura Sattelberger and Robin van der Veer.
Maximum likelihood estimation from a tropical and a Bernstein–Sato perspective.
International Mathematics Research Notices, 2023(6):5263–5292, 2023.