Secant nondefectivity of determinantal moment varieties

Oskar Henriksson, University of Copenhagen
Joint work with Kristian Ranestad, Lisa Seccia and Teresa Yu

Secant nondefectivity of determinantal moment varieties

Oskar Henriksson, University of Copenhagen
Joint work with Kristian Ranestad, Lisa Seccia and Teresa Yu

The method of moments

Let X be a random variable with density depending on unknown parameters $\theta = (\theta_1, \dots, \theta_n)$.
Goal: estimate θ from sample moments $\widehat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^k$ for an iid sample $x_1, \dots, x_N$.

Secant nondefectivity of determinantal moment varieties

Oskar Henriksson, University of Copenhagen
Joint work with Kristian Ranestad, Lisa Seccia and Teresa Yu

The method of moments

Let X be a random variable with density depending on unknown parameters $\theta = (\theta_1, \ldots, \theta_n)$.
Goal: estimate θ from sample moments $\widehat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^k$ for an iid sample $x_1, \ldots, x_N$.

For many distributions, the moments $m_r(\theta) = \mathbb{E}[X^r]$ for $r = 1, 2, 3, \ldots, d$ are polynomial in θ.

Secant nondefectivity of determinantal moment varieties

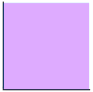
Oskar Henriksson, University of Copenhagen
Joint work with Kristian Ranestad, Lisa Seccia and Teresa Yu

The method of moments

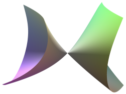
Let X be a random variable with density depending on unknown parameters $\theta = (\theta_1, \dots, \theta_n)$.
Goal: estimate θ from sample moments $\widehat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^k$ for an iid sample $x_1, \dots, x_N$.

For many distributions, the moments $m_r(\theta) = \mathbb{E}[X^r]$ for $r = 1, 2, 3, \dots, d$ are polynomial in θ.

The dth moment variety $\mathcal{M}_d$ is the closure of the image of



$\mathbb{C}^n \xrightarrow{\theta \mapsto (m_0(\theta):m_1(\theta):\cdots:m_d(\theta))}$



$\mathcal{M}_d \subseteq \mathbb{P}^n$

Secant nondefectivity of determinantal moment varieties

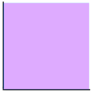
Oskar Henriksson, University of Copenhagen
Joint work with Kristian Ranestad, Lisa Seccia and Teresa Yu

The method of moments

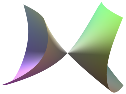
Let X be a random variable with density depending on unknown parameters $\theta = (\theta_1, \dots, \theta_n)$.
Goal: estimate θ from sample moments $\widehat{m}_r = \frac{1}{N}\sum_{i=1}^{N} x_i^k$ for an iid sample $x_1, \dots, x_N$.

For many distributions, the moments $m_r(\theta) = \mathbb{E}[X^r]$ for $r = 1, 2, 3, \dots, d$ are polynomial in θ.

The dth moment variety $\mathcal{M}_d$ is the closure of the image of



$\mathbb{C}^n \xrightarrow{\theta \mapsto (m_0(\theta):m_1(\theta):\cdots:m_d(\theta))}$



$\mathcal{M}_d \subseteq \mathbb{P}^n$

How many moments do we need to have algebraic or rational identifiability?

Moment varieties of mixture distributions

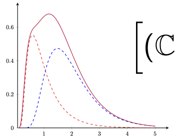
Moment varieties of mixture distributions

Let p_θ be a density with parameters $\theta = (\theta_1, \ldots, \theta_n)$ and moment variety $\mathcal{M}_d \subseteq \mathbb{P}^d$.

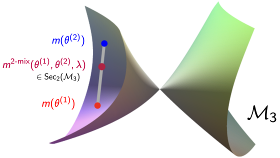
Moment varieties of mixture distributions

Let p_θ be a density with parameters $\theta = (\theta_1, \ldots, \theta_n)$ and moment variety $\mathcal{M}_d \subseteq \mathbb{P}^d$.

The moment variety of a *k*-mixture is the *k*th secant variety $\mathrm{Sec}_k(\mathcal{M}_d)$:



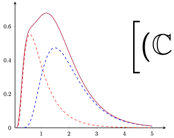
$$\left[(\mathbb{C}^n)^k \times \mathcal{V}\left(1 - \textstyle\sum_{i=1}^k \lambda_i\right)\right] / \mathfrak{S}_k \to \mathrm{Sec}_k(\mathcal{M}_d)$$
$$(\theta^{(1)}, \ldots, \theta^{(k)}, \lambda_1, \ldots, \lambda_k) \mapsto \textstyle\sum_{i=1}^k \lambda_i m(\theta^{(i)})\,.$$



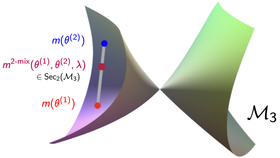
Moment varieties of mixture distributions

Let p_θ be a density with parameters $\theta = (\theta_1, \ldots, \theta_n)$ and moment variety $\mathcal{M}_d \subseteq \mathbb{P}^d$.

The moment variety of a k-mixture is the kth secant variety $\mathrm{Sec}_k(\mathcal{M}_d)$:



$$\left[(\mathbb{C}^n)^k \times \mathcal{V}\left(1 - \textstyle\sum_{i=1}^k \lambda_i\right)\right] / \mathfrak{S}_k \to \mathrm{Sec}_k(\mathcal{M}_d)$$
$$(\theta^{(1)}, \ldots, \theta^{(k)}, \lambda_1, \ldots, \lambda_k) \mapsto \textstyle\sum_{i=1}^k \lambda_i m(\theta^{(i)}) .$$



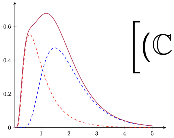
Theorem (Pearson 1894; Améndola–Ranestad–Sturmfels 2016; Améndola–Rodriguez–Lindberg 2025)

For **Gaussian mixtures**, we have algebraic identifiability for $d \geq 3k - 1$, and rational identifiability for $d \geq 3k + 2$.

Moment varieties of mixture distributions

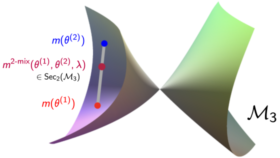
Let p_θ be a density with parameters $\theta = (\theta_1, \ldots, \theta_n)$ and moment variety $\mathcal{M}_d \subseteq \mathbb{P}^d$.

The moment variety of a *k*-mixture is the *k*th secant variety $\mathrm{Sec}_k(\mathcal{M}_d)$:



$$\left[(\mathbb{C}^n)^k \times \mathcal{V}\left(1 - \sum_{i=1}^k \lambda_i\right)\right] / \mathfrak{S}_k \to \mathrm{Sec}_k(\mathcal{M}_d)$$

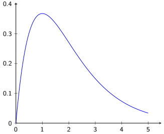
$$(\theta^{(1)}, \ldots, \theta^{(k)}, \lambda_1, \ldots, \lambda_k) \mapsto \sum_{i=1}^k \lambda_i m(\theta^{(i)}) \, .$$



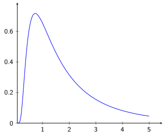
Theorem (Pearson 1894; Améndola–Ranestad–Sturmfels 2016; Améndola–Rodriguez–Lindberg 2025)

For **Gaussian mixtures**, we have algebraic identifiability for $d \geq 3k-1$, and rational identifiability for $d \geq 3k+2$.

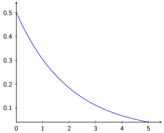
What about mixtures of other distributions?



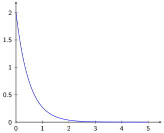
Gamma



Inverse Gaussian



χ^2



Exponential

Theorem (**H.**, Seccia, Yu 2024)

Gamma distribution:

Ideal: $I_3\begin{pmatrix} 0 & m_1 & 2m_2 & 3m_3 & \cdots & (d-1)m_{d-1} \\ m_0 & m_1 & m_2 & m_3 & \cdots & m_{d-1} \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-1}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Theorem (**H.**, Seccia, Yu 2024)

Gamma distribution:

Ideal: $I_3\begin{pmatrix} 0 & m_1 & 2m_2 & 3m_3 & \cdots & (d-1)m_{d-1} \\ m_0 & m_1 & m_2 & m_3 & \cdots & m_{d-1} \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-1}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Invere Gaussian distribution:

Ideal: $I_3\begin{pmatrix} m_0^2 & m_0 & m_1 & m_2 & \cdots & m_{d-2} \\ 0 & m_1 & 3m_2 & 5m_3 & \cdots & (2d-3)m_{d-1} \\ m_1^2 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2+\binom{d-1}{2}t^3\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-2}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Theorem (**H.**, Seccia, Yu 2024)

Gamma distribution:

Ideal: $I_3\begin{pmatrix} 0 & m_1 & 2m_2 & 3m_3 & \cdots & (d-1)m_{d-1} \\ m_0 & m_1 & m_2 & m_3 & \cdots & m_{d-1} \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-1}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Invere Gaussian distribution:

Ideal: $I_3\begin{pmatrix} m_0^2 & m_0 & m_1 & m_2 & \cdots & m_{d-2} \\ 0 & m_1 & 3m_2 & 5m_3 & \cdots & (2d-3)m_{d-1} \\ m_1^2 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2+\binom{d-1}{2}t^3\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-2}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Theorem (**H.**, Ranestad, Seccia, Yu 2025)

For k-mixtures of the inverse Gaussian or gamma distribution, we have algebraic identifiability for $d \geq 3k-1$, and rational identifiability for $d \geq 3k+2$.

Theorem (**H.**, Seccia, Yu 2024)

Gamma distribution:

Ideal: $I_3\begin{pmatrix} 0 & m_1 & 2m_2 & 3m_3 & \cdots & (d-1)m_{d-1} \\ m_0 & m_1 & m_2 & m_3 & \cdots & m_{d-1} \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-1}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Invere Gaussian distribution:

Ideal: $I_3\begin{pmatrix} m_0^2 & m_0 & m_1 & m_2 & \cdots & m_{d-2} \\ 0 & m_1 & 3m_2 & 5m_3 & \cdots & (2d-3)m_{d-1} \\ m_1^2 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2+\binom{d-1}{2}t^3\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-2}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Theorem (**H.**, Ranestad, Seccia, Yu 2025)

For k-mixtures of the inverse Gaussian or gamma distribution, we have algebraic identifiability for $d \geq 3k-1$, and rational identifiability for $d \geq 3k+2$.

Theorem (**H.**, Seccia, Yu 2024)

For k-mixtures of the exp or χ^2 distribution, we have rational identifiability for $d \geq 2k-1$.

Theorem (**H.**, Seccia, Yu 2024)

Gamma distribution:

Ideal: $I_3\begin{pmatrix} 0 & m_1 & 2m_2 & 3m_3 & \cdots & (d-1)m_{d-1} \\ m_0 & m_1 & m_2 & m_3 & \cdots & m_{d-1} \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-1}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Invere Gaussian distribution:

Ideal: $I_3\begin{pmatrix} m_0^2 & m_0 & m_1 & m_2 & \cdots & m_{d-2} \\ 0 & m_1 & 3m_2 & 5m_3 & \cdots & (2d-3)m_{d-1} \\ m_1^2 & m_2 & m_3 & m_4 & \cdots & m_d \end{pmatrix}$

Hilbert series: $\frac{1}{(1-t)^3}\left(1+(d-2)t+\binom{d-1}{2}t^2+\binom{d-1}{2}t^3\right)$.

Singularities: $\mathcal{V}(m_0, m_1, \ldots, m_{d-2}) \cup \mathcal{V}(m_1, m_2, \ldots, m_d)$

Theorem (**H.**, Ranestad, Seccia, Yu 2025)

For k-mixtures of the inverse Gaussian or gamma distribution, we have algebraic identifiability for $d \geq 3k-1$, and rational identifiability for $d \geq 3k+2$.

Theorem (**H.**, Seccia, Yu 2024)

For k-mixtures of the exp or χ^2 distribution, we have rational identifiability for $d \geq 2k-1$.

Future directions: Unifying results, joins, identifiability degrees, real root counts, ED degrees, ...

A New Direction in Algebraic Statistics: Design of Experiments with Application in Cryptography and Cybersecurity

A. J. Saka[1,2] L. A. Akanbi[3] T. G. Jaiyeola[1] Felipe O. Silva[2] R. A. Adetona[1] J. S. S. Bueno[4]

[1]Department of Mathematics
Obafemi Awolowo University, Ile-Ife, Osun state, Nigeria

[2]Department of System Engineering and Automation
Federal University of Lavras, Brazil

[3]Department of Management, Business and Marketing
Birmingham City University Business School Birmingham, United Kingdom.

[4]Department of Statistics
Federal University of Lavras, Brazil.

Introduction

In this study Algebra and Statistics are harmonized to develop a cryptography Algorithm to secure communication in Internet of Things. Here I will like to dowel much on Statistics with bias in Design of Experiments precisely with respect to Construction of Designs; Mutually Orthogonal Latin Square MOLS and Resolvable Balanced Incomplete Block Design RBIBD

Methodology and Result: Automated Data Encryption Generation

Stage 1. (Plaintext): In the name of God, the Entirely Merciful, the Especially Merciful. It is You we worship and You we ask for help. Guide us to the straight path.

Stage 2. (Time Based Vector): Generation of Encrypting as: [9, 3, 8, 4, 7, 3, 2, 0, 2, 1, 6, 8, 5, 1, 4, 2, 0, 3, 7, 3, 3, 5, 3, 0, 2, 1, 6]

Stage 3. (Ciphertex): Ro bkl oank ph Lsh- zie Hqvirgla Reuljnxs- tik Fuuigjgmlb Pgrckfwq. Lc ja Bvv wf cptxlmq god Brw we csm kou qfts? Hujjf wx xs unf swucigjt rftk[

Stage 4. (Matrix Encryption): [[45907, 48623, 58147, 50564, 58386, 31020, 56474, 90152, 47801, 41489, 59161, 55011, 46233, 37397, 61102, 56072, 68895, 85046, 52907, 43669, 58549, 50832, 53385, 47084, 37875, 51607, 49876, 66581, 56724, 48384, 18226, 70224, 57220, 53490, 70702, 56445, 54399, 50622, 38668, 71687, 43651, 60729, 57535, 65357, 58730, 98173, 44893, 52668, 61283, 73150, 67537, 51712, 69239, 53443, 49503, 61178, 54160, 59639, 52020]]

Conclusion

The following are the distinct features and advantages of the ZTM algorithm which distinguishes it from traditional encryption schemes like RSA and AES.

Dynamic Time-Based Key Generation: Unlike static key approaches in RSA and AES, the ZTM algorithm generates time-based keys dynamically, ensuring an additional layer of security for each encryption session.

Resistance to Pattern Recognition: By using a shuffled key list in combination with zig-zag block matrix, the resulting algorithm resists pattern detection that could sometimes appear in block ciphers like AES.

Lightweight and Efficient: The algorithm is optimized for real-time encryption and decryption of data. Its lightweight structure allows it to perform faster than heavier algorithms such as RSA.

Symmetric Timing for Encryption and Decryption: Unlike RSA, which has an intensive decryption process, the ZTM algorithm ensures that decryption times are almost identical to encryption times, regardless of the length of the text.

Low Key Space Attack Vulnerability: The shuffled time-based key list enhances the variability of the encryption key for each session. This makes brute-force attacks more difficult compared to simpler encryption schemes, which rely on smaller fixed key spaces.

On a final note, the proposed hybridized encryption algorithm (ZTM) has a unique feature of dynamic time-based key generation such that it generates distinct ciphertext at different times of encryption for the same data or messages, which distinguishes it from traditional encryption schemes like RSA and AES.

Computing Ideals of Level-*k* Jukes-Cantor Networks

Udani Ranasinghe (University of Hawai'i at Mānoa)

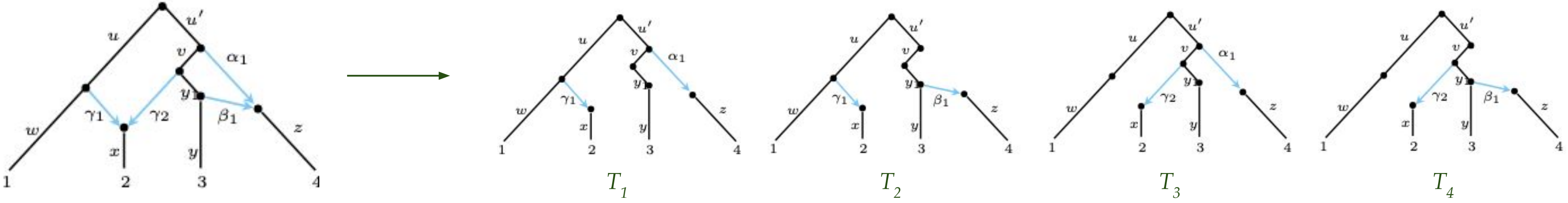
Joint work with Dr. Elizabeth Gross, Dr. Gillian Grindstaff, Dr. Max Hill, Ruiqi Huang, and Nazia Riasat

July 21, 2025

Parameterization of the Jukes-Cantor Model

- Tree parametrization: $q^{T}_{g_1 \dots g_n} = \begin{cases} \prod_{e \in \Sigma(T)} a^{e}_{\sum_{j \in B_e} g_j} & \text{if } \sum_{j=1}^{n} g_j = 0 \\ 0 & \text{otherwise} \end{cases}$

- Network parametrization: $q_{g_1 \dots g_n} = \sum_{T \in \mathcal{N}} \gamma_T q^{T}_{g_1 \dots g_n} = \sum_{T \in \mathcal{N}} q^{T}_{g_1 \dots g_n}$

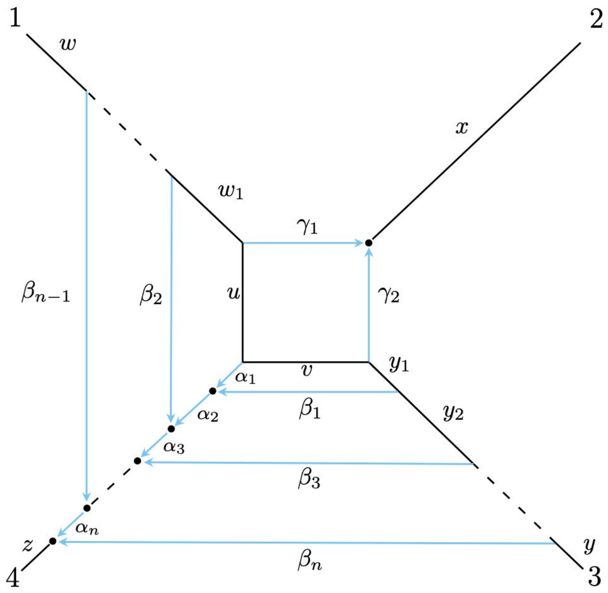


$$q_{g_1 g_2 g_3 g_4} = q^{T_1}_{g_1 g_2 g_3 g_4} + q^{T_2}_{g_1 g_2 g_3 g_4} + q^{T_3}_{g_1 g_2 g_3 g_4} + q^{T_4}_{g_1 g_2 g_3 g_4}$$

- We are interested in studying the ideals corresponding to network parameterizations

Example: Non-Reticulate Half Ziggurat

We implemented a new method of computing the parametrization of a network, and used it analyze a number of families of networks, such as the following:



- To begin our analysis, we calculated dimension for each network numerically - and observed the dimension appeared to stabilize
- Using the Macualay2 package **`MultigradedImplicitization`**, we computed generators of the ideal up to degree 2:

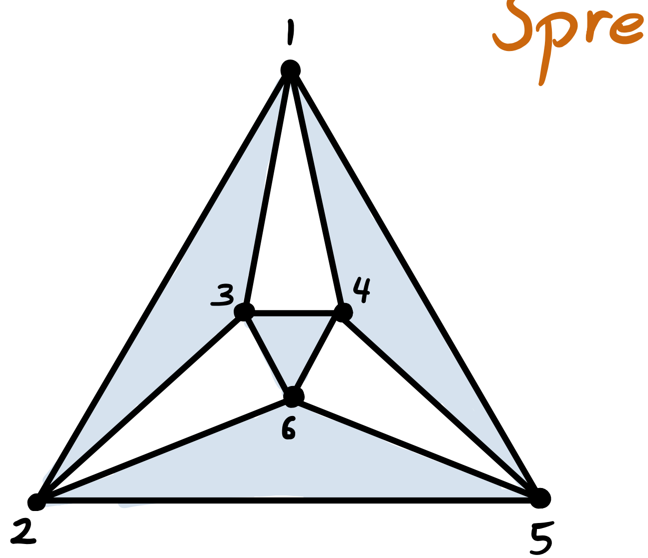
$$f_0 := -q_{1111} - q_{1212} + q_{1122} + q_{1221}$$
$$f_1 := q_{1023}q_{1230} - q_{1010}q_{1212}$$
$$f_2 := q_{0101}q_{1010} - q_{0110}q_{1001} - q_{0011}q_{1100} + q_{1111}q_{0000}$$

- We then proved that as the number of reticulations along leaf 4 grows, the ideal stabilizes to $\langle f_0, f_1, f_2 \rangle$
- We believe that these polynomials are always in the ideal for a larger class of networks

Hypertrees

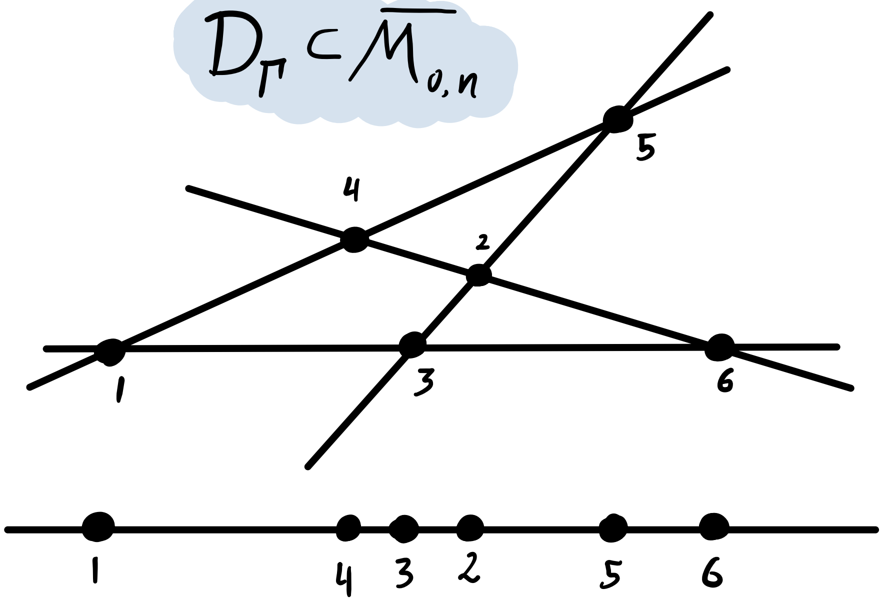
A *hypertree* is a collection of $n-2$ 3-tuples $\Gamma_1, \dots, \Gamma_{n-2}$ on $[n] = \{1, \dots, n\}$ s.t.

1. Each $i \in [n]$ appears in at least 2 tuples
   "Appearance condition"
2. $\left| \bigcup_{i \in S} \Gamma_i \right| \geq |S| + 2$ $\forall$ non-empty subsets $S \subseteq [n-2]$
   "Spreading condition"

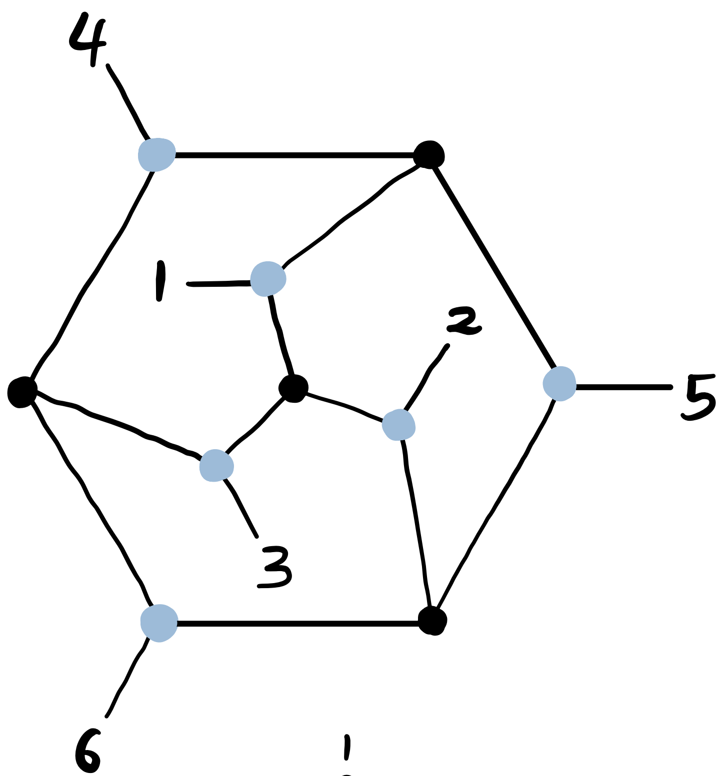


$\{(123), (145), (256), (346)\}$

$D_\Gamma \subset \overline{M}_{0,n}$



Scattering Hypertrees

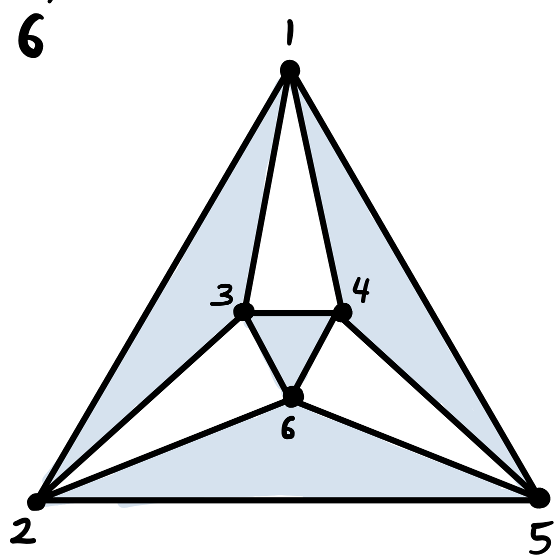


$$L_\Gamma = \sum_{(i,j)\in\Gamma} s_{ij} \log(p_{ij})$$

$$\nabla L_\Gamma = 0$$

- What are the critical points of L_Γ on $\mathcal{M}_{0,n} = Gr(2,n)^\circ / (\mathbb{C}^*)^n$ w/ coordinates given by

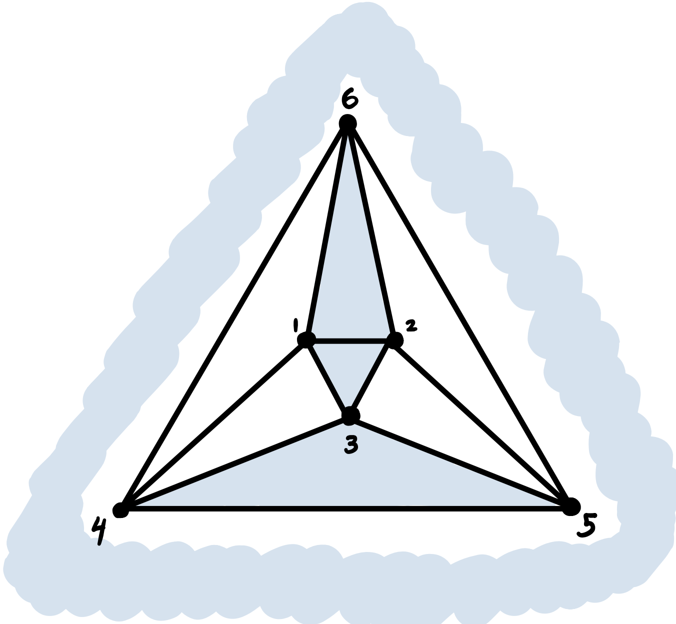
$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & x_1 & x_2 & & x_{n-3} & 1 \end{pmatrix}?$$



- What about the ML degree with coordinates

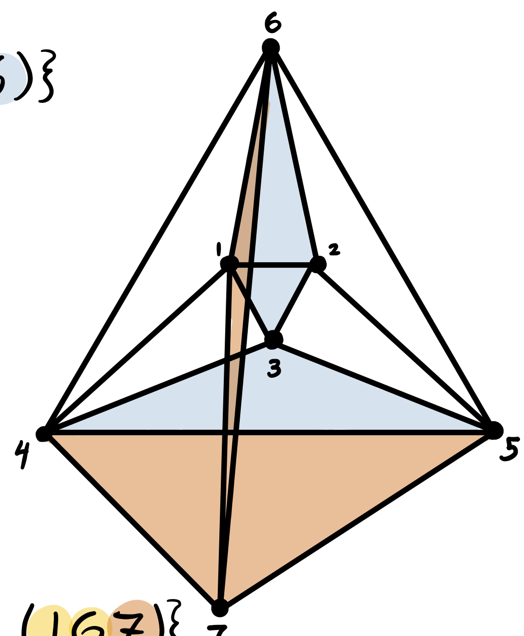
$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & & x_n \end{pmatrix}?$$

An Infinite Family of ML Degree 0

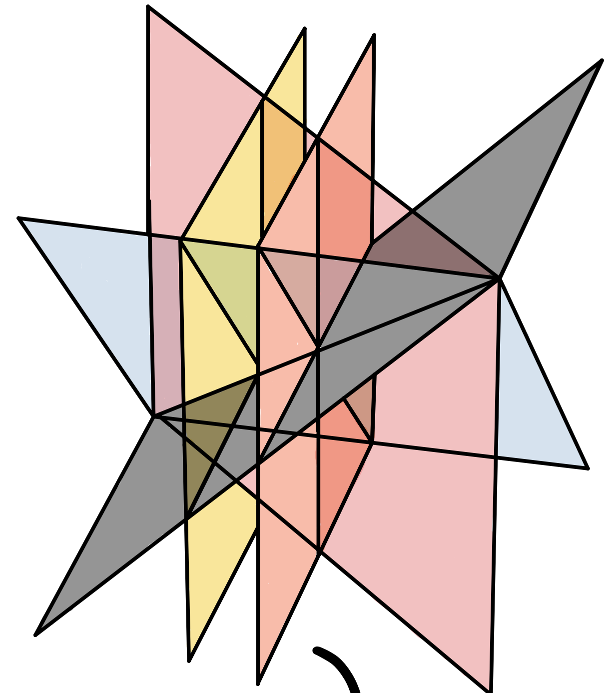


$\{(1\,2\,3), (1\,2\,6), (3\,4\,5), (4\,5\,6)\}$

ADD A VERTEX



$\{(1\,2\,3), (1\,2\,6), (3\,4\,5), (4\,5\,7), (1\,6\,7)\}$



⋮

Higher dimensional arrangements without bounded regions

Geometry of continuous adjoint Newton's system for bivariate quadratics

Francisco Ponce-Carrión

North Carolina State University

July 21, 2025

Definition

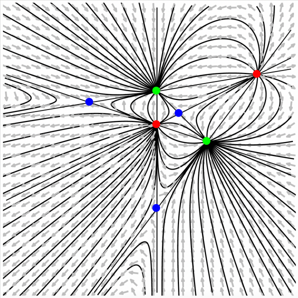
$\frac{dx}{dt} = g(x)$ where $g = |Jf|\ Jf^{-1} f.$

Example

Let f be

$$\begin{bmatrix} -11x_1^2 + 8x_1x_2 - 2x_2^2 + 63x_1 - 112 \\ -8x_1^2 + 5x_1x_2 - 8x_2^2 + 54x_1 + 54x_2 - 196 \end{bmatrix}$$

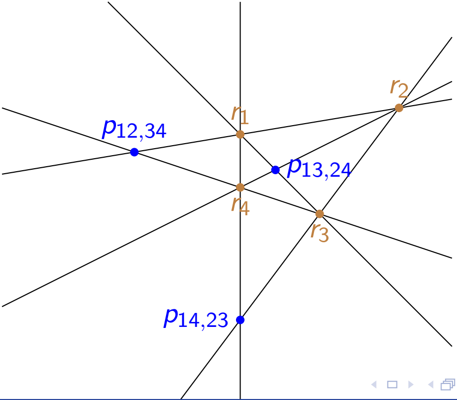
- 7 solutions to $g = 0$.
- Red solutions: source.
- Green solutions: sink.
- Blue solutions: saddle.



Theorem (Hauenstein, Hills, Hong, PC, 2025+)

(1) *Equilibria:*

1. $g = 0$ *has* 7 *solutions:* $\{r_1, r_2, r_3, r_4, p_{12,34}, p_{13,24}, p_{14,23}\}$
2. r_1, r_2, r_3, r_4 *are solutions to* $f = 0$.
3. *Each* $p_{ij,kl}$ *is the intersection point of the lines* L_{ij} *and* L_{kl}.



Theorem (Hauenstein, Hills, Hong, PC, 2025+)

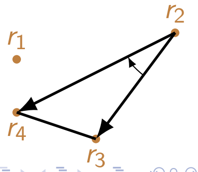
(2) *Eigenvalue / eigenspace at equilibrium e:*

e	eigenvalue of $g'(e)$	eigenspace of $g'(e)$
r_q	$\lambda_q = \frac{(-1)^q}{\Delta_q} C$	$\mathbb{R}^2$
$p_{ij,kl}$	$\lambda_{ij} = \frac{(-1)^{i+j}}{\Delta_{ij,kl}} C$	$\text{span}(r_j - r_i)$
	$\lambda_{kl} = \frac{(-1)^{k+l}}{\Delta_{kl,ij}} C$	$\text{span}(r_l - r_k)$

for some nonzero constant $C \in \mathbb{R}^2$.

Assuming $C < 0$ we can deduce $\lambda_1 < 0$:

$$\lambda_1 = \frac{(-1)^1}{\Delta_1} C, \quad \Delta_1 = \Delta_{23,24} = \frac{1}{2} \begin{vmatrix} r_3 - r_2 & r_4 - r_2 \end{vmatrix}$$



Toric geometry of ReLU neural networks

Definitions

- For any $n \in \mathbb{N}$, the *rectified linear unit* (ReLU) is the map $\varsigma : \mathbb{R}^n \to \mathbb{R}$, $\varsigma(x) = (max\{0, x_1\}, max\{0, x_2\}, \cdots, max\{0, x_n\})$.
- For any number of hidden layers $k \in \mathbb{N}$, a $(k+1)-$layer **feedforward ReLU neural network** (RNN) is defined as follows:

$$f_\theta : \mathbb{R}^{n_0} \xrightarrow{\varsigma \circ A_1} \mathbb{R}^{n_1} \xrightarrow{\varsigma \circ A_2} \cdots \xrightarrow{\varsigma \circ A_k} \mathbb{R}^{n_k} \xrightarrow{A_{k+1}} \mathbb{R}$$

  where $n_0, n_1, \cdots, n_k \in \mathbb{N}$, and $A_i : \mathbb{R}^{n_{i-1}} \to \mathbb{R}^{n_i}$ are all affine-linear maps. The output function f is a piecewise linear function.

Motivation

- Which functions are exactly realized by a given RNN architecture?
- Conversely, given a piecewise linear function, which architectures realize it?

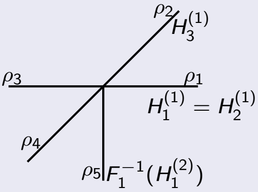
Unbiased RNNs with rational weights

Toric geometry

- Denoted $\Sigma_{f_\theta}^{\text{big}}$ the **ReLU fan** of an unbiased RNN with rational weights is defined to be the **canonical polyhedral complex** associated with f_θ.
- The output function f serves as the support function of a $\mathbb{Q}-$Cartier divisor D_f supported on $\Sigma_{f_\theta}^{\text{big}}$, called **ReLU Cartier divisor**.

Example

Given a 3-layer unbiased feedforward RNN: $f_\theta : \mathbb{R}^2 \xrightarrow{F_1 := L_1 \circ \varsigma} \mathbb{R}^3 \xrightarrow{F_2 := L_2 \circ \varsigma} \mathbb{R} \xrightarrow{L_3} \mathbb{R}$, where $L_1 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 1 & -1 \end{bmatrix}$, $L_2 = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$ and $L_3 = 1$. The ReLU fan $\Sigma_{f_\theta}^{\text{big}}$ is as follows:



. The output piecewise linear function is $f = \max\{0, x, y\}$. The ReLU Cartier divisor associated with f is $D_f = -D_{\rho_1} - D_{\rho_2}$.

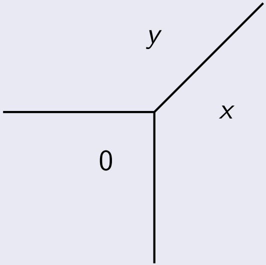
First application of toric geometry framework

Complete classification of functions realizable by unbiased depth 2 RNNs

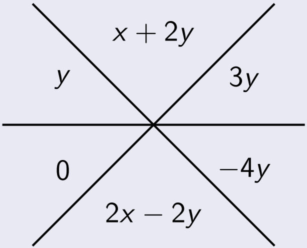
- The classification is obtained with the help of computing **intersection number of divisors and curves**.
- A piecewise linear function f is realizable by an unbiased depth 2 RNN with rational weights iff $D_f \cdot V(\tau_1) = D_f \cdot V(\tau_2)$ for any two walls τ_1, τ_2 coming from the same full hyperplane in the fan.

Counterexamples in $\mathbb{R}^2$

$f = max\{0, x, y\}$:



and



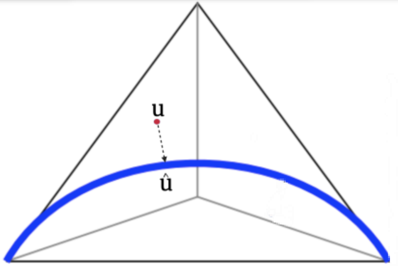
Wasserstein Distance to Small Toric Models

Ikenna Nometa (UH Mānoa)

With: Greg DePaul, Serkan Hoşten, and Nilava Metya

IMSI Conference on New Directions in Algebraic Statistics

(Lightning Talk)



$W_d(\mu, \nu)$ is the optimal value of:

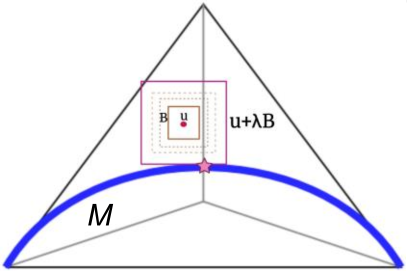
Maximize $\sum_{i=1}^{n} (\mu_i - \nu_i)x_i$ s.t.

$|x_i - x_j| \leq d_{ij}$

for all $i < j \in [n]$

Wasserstein distance between μ and M

$$W_d(\mu, M) = \min_{\nu \in M} W_d(\mu, \nu)$$



Geometrically:

$$D_B(\mu, M) = \min_{\lambda \in \mathbb{R}_{\geq 0}} \{\lambda : (\mu + \lambda B) \cap M \neq \emptyset\}$$

Proposition (Çelik-Jamneshan-Montúfar-Sturmfels-Venturello)[ÇJM+21]

$W_d(\mu, M) = D_B(\mu, M)$; and the # of complex critical points is bounded by

$$\sum_{i=0}^{n-1} \delta_i(M) f_i(B)$$

Polar Degrees of Rational Normal Scrolls

Theorem (DePaul-Hoşten-Metya-N [DHMN24])

Let $X = X_A$ be a rational normal scroll whose toric variety is defined by the A matrix determined by positive integers $n_1, n_2, \ldots, n_d$. Let $N = \sum_{k=1}^{d} n_k$, then X has nonzero i^{th} polar degrees only for $i = 0, 1, 2$. In particular,

$$\delta_i(X) = \begin{cases} N, & \text{if } i = 0, 2 \\ 2(N-1), & \text{if } i = 1 \end{cases}$$

Other Results:

- Polar degrees of graphical models – star, path, and cycle graphs[1]

[1] G. DePaul, S. Hoşten, N. Metya, and I. Nometa. "Degrees of the Wasserstein Distance to Small Toric Models". In: *Algebraic Statistics* 15 (2 2024), pp. 249–267.

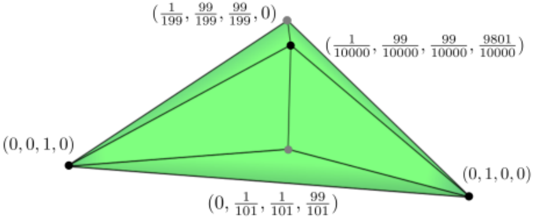
The Correlated Equilibrium Polytope of Zero Sum Games

Danai Deligeorgaki, Max Hill, Bryson Kagy, Miruna-Stefana Sorea

North Carolina State University

July 21, 2025

		Player 2 go	Player 2 stop
Player 1	go	$(-99, -99)$	$(1, 0)$
	stop	$(0, 1)$	$(0, 0)$



Zero Sum Games

Goals

- Characterize the combinatorics of the dimension of the polytope (Phase transitions!)
- Understand how multiple notions of generic interact
- Understand how notions of reducing a game affect the Correlated equilibrium polytope