An Implicitization Challenge for Binary Factor Analysis

María Angélica Cueto[1]

(Joint work with Enrique Tobis[2] and Josephine Yu[3])

[1] Mathematics Department
Columbia University

[2]Harvard University

[3]GATech

Conference on Tropical Geometry and Computational Biology
Saarbrücken

The Statistical model $\mathcal{F}_{4,2}$

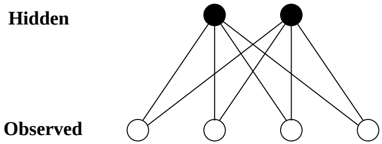


Figure: The undirected graphical model $\mathcal{F}_{4,2}$.

The set of all possible joint probability distributions (X_1, X_2, X_3, X_4) forms an algebraic variety $\mathcal{M}$ inside Δ_{15} with expected codimension one and (multi)homogeneous defining equation f.

Problem (Drton-Sturmfels-Sullivant)

Find the degree and the defining polynomial f / Newton polytope of $\mathcal{M}$.

Geometry of the model $\mathcal{F}_{4,2}$

Parameterization of the model: $p\colon \mathbb{R}^{32} \to \mathbb{R}^{16}$,

$$p_{ijkl} = \sum_{s=0}^{1}\sum_{r=0}^{1} a_{si}b_{sj}c_{sk}d_{sl}e_{ri}f_{rj}g_{rk}h_{rl} \text{ for all } (i,j,k,l) \in \{0,1\}^4.$$

Using homogeneity and the distributive law

$$p\colon (\mathbb{P}^1 \times \mathbb{P}^1)^8 \to \mathbb{P}^{15} \quad p_{ijkl} = (\sum_{s=0}^{1} a_{si}b_{sj}c_{sk}d_{sl}) \cdot (\sum_{r=0}^{1} e_{ri}f_{rj}g_{rk}h_{rl}).$$

So we have a coordinatewise product of two parameterizations of $\mathcal{F}_{4,1}$: the graphical model corresponding to the 4-claw tree with binary nodes.

NICE FACTS: We know a lot about $\mathcal{F}_{4,1}$ and coordinatewise products of projective varieties...

Geometry of the model $\mathcal{F}_{4,2}$

Fact

1. *The binary 4-claw tree model is* $Sec^1(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \subset \mathbb{P}^{15}$.
2. *Coordinatewise products of parameterizations corresponds to Hadamard products of algebraic varieties.*

Definition

$X, Y \subset \mathbb{P}^n$, the **Hadamard product** of X and Y is

$$X \star Y = \overline{\{x \star y := (x_0 y_0 : \ldots : x_n y_n) \,|\, x \in X, y \in Y, x \star y \neq 0\}} \subset \mathbb{P}^n.$$

Geometry of the model $\mathcal{F}_{4,2}$

Corollary

The algebraic variety of the model is $\mathcal{M} = X \,\bullet\, X$ where X is the first secant variety of the Segre embedding $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^{15}$.

Remark

The model is highly symmetric. It is invariant under relabeling of the four observed nodes and changing the role of the two states (0 and 1). Therefore, we have an action of the group $B_4 = \mathbb{S}_4 \ltimes (\mathbb{S}_2)^4$, the group of symmetries of the 4-cube.

Geometry of the model $\mathcal{F}_{4,2}$

Corollary

The algebraic variety of the model is $\mathcal{M} = X \,\mathbf{.}\, X$ where X is the first secant variety of the Segre embedding $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^{15}$.

Remark

The model is highly symmetric. It is invariant under relabeling of the four observed nodes and changing the role of the two states (0 and 1).
Therefore, we have an action of the group $B_4 = \mathbb{S}_4 \ltimes (\mathbb{S}_2)^4$, the group of symmetries of the 4-cube.

Useful facts about X:

1. The ideal $I(X)$ is a well-studied object: it is the 9-dim *irreducible* projective variety of all $2 \times 2 \times 2 \times 2$-tensors of tensor rank ≤ 2.
2. Known set of generators for $I(X)$: 3×3-minors of all three 4×4-flattenings of these tensors $\rightsquigarrow$ 48 polynomials.

Tropicalizing the model

Definition

Given algebraic variety $X \subset \mathbb{C}^n$ with defining ideal $I = I(X) \subset \mathbb{C}[x_1, \ldots, x_n]$, the tropicalization of X or I is defined as:

$$\mathcal{T}X = \mathcal{T}I = \{w \in \mathbb{R}^n \,|\, \mathrm{in}_w(I) \text{ contains no monomial}\}$$

where $\mathrm{in}_w(I) = \langle \mathrm{in}_w(f) : f \in I \rangle$, and $\mathrm{in}_w(f)$ is the sum of all *nonzero* terms of $f = \sum_\alpha c_\alpha x^\alpha$ such that $\alpha \cdot w$ is **maximum**.

1. If X is a hypersurface, $\mathcal{T}X$ is the collection of all codimension one cones in the normal fan of the Newton polytope of X. The multiplicity of a maximal cone is the lattice length of the corresponding edge in the polytope.
2. The lineality space of the fan $\mathcal{T}X$ is the set
$$L = \{w \in \mathcal{T}X : \mathrm{in}_w(I) = I\}.$$
It describes the action of a maximal torus on X (action by $L \cap \mathbb{Z}^{n+1}$).
3. Morphisms can be tropicalized and *monomial maps* have very nice tropicalizations.

Theorem (Sturmfels-Tevelev)

Let $A \in \mathbb{Z}^{d\times n}$, defining a monomial map $\alpha\colon (\mathbb{C}^*)^n \to (\mathbb{C}^*)^d$ and a canonical linear map $A\colon \mathbb{R}^n \to \mathbb{R}^d$. Let $V \subset (\mathbb{C}^*)^n$ be a subvariety. Then

$$\mathcal{T}(\alpha(V)) \;=\; A(\mathcal{T}V).$$

Moreover, if α induces a generically finite morphism on V, we have an explicit formula to push forward the multiplicities of $\mathcal{T}V$ to the multiplicities of $\mathcal{T}(\alpha V)$.

Here, $\mathcal{M} = X \,\mathbf{.}\, X = \alpha(X \times X)$, and A is the matrix $(Id_{16} \mid Id_{16})$.

Theorem (Sturmfels-Tevelev)

Let $A \in \mathbb{Z}^{d\times n}$*, defining a monomial map* $\alpha\colon (\mathbb{C}^*)^n \to (\mathbb{C}^*)^d$ *and a canonical linear map* $A\colon \mathbb{R}^n \to \mathbb{R}^d$*. Let* $V \subset (\mathbb{C}^*)^n$ *be a subvariety. Then*

$$\mathcal{T}(\alpha(V)) \;=\; A(\mathcal{T}V).$$

Moreover, if α *induces a generically finite morphism on* V*, we have an explicit formula to push forward the multiplicities of* $\mathcal{T}V$ *to the multiplicities of* $\mathcal{T}(\alpha V)$*.*

Here, $\mathcal{M} = X \,\mathbf{.}\, X = \alpha(X \times X)$, and A is the matrix $(Id_{16} \mid Id_{16})$.

Theorem (— -Tobis-Yu, Allermann-Rau, ...)

Let $X, Y \subset \mathbb{C}^m$ *be two irreducible varieties. Then*

$$\mathcal{T}(X \times Y) = \mathcal{T}X \times \mathcal{T}Y$$

as weighted polyhedral fans, with $m_{\sigma\times\tau} = m_\sigma m_\tau$ *for maximal cones.*

Corollary: $\mathcal{T}\mathcal{M} = \mathcal{T}(X \,\mathbf{.}\, X) = \mathcal{T}X + \mathcal{T}X$ (as **sets**!).

Computing $\mathcal{T}\mathcal{M}$ from $\mathcal{T}X$

$\mathcal{T}X$ can be computed with `Gfan`. In particular,

- 10-dim. *simplicial* fan in $\mathbb{R}^{16}$,
- 5-dim. lineality space,
- f-vector= $(381, 3\,436, 11\,236, 15\,640, 7\,680)$,
- 13 rays and 49 maximal cones up to B_4-symmetry.

Computing $\mathcal{T}\mathcal{M}$ from $\mathcal{T}X$

$\mathcal{T}X$ can be computed with `Gfan`. In particular,

- 10-dim. *simplicial* fan in $\mathbb{R}^{16}$,
- 5-dim. lineality space,
- f-vector$= (381, 3\,436, 11\,236, 15\,640, 7\,680)$,
- 13 rays and 49 maximal cones up to B_4-symmetry.

Thus we know $\mathcal{T}\mathcal{M} = \mathcal{T}X + \mathcal{T}X$ as a set!

- Dimension = 15 in $\mathbb{C}^{16}$, so $\mathcal{M}$ is a hypersurface!
- Number of maximal cones in $\mathcal{T}X + \mathcal{T}X = 6\,865\,824$.
- $18\,972$ maximal cones up to B_4-symmetry.

Computing $\mathcal{T}\mathcal{M}$ from $\mathcal{T}X$

$\mathcal{T}X$ can be computed with `Gfan`. In particular,

- 10-dim. *simplicial* fan in $\mathbb{R}^{16}$,
- 5-dim. lineality space,
- f-vector$= (381, 3\,436, 11\,236, 15\,640, 7\,680)$,
- 13 rays and 49 maximal cones up to B_4-symmetry.

Thus we know $\mathcal{T}\mathcal{M} = \mathcal{T}X + \mathcal{T}X$ as a set!

- Dimension $=$ 15 in $\mathbb{C}^{16}$, so $\mathcal{M}$ is a hypersurface!
- Number of maximal cones in $\mathcal{T}X + \mathcal{T}X = 6\,865\,824$.
- $18\,972$ maximal cones up to B_4-symmetry.

BUT we want more...

We want to compute multiplicities at *regular points* of $\mathcal{T}\mathcal{M}$.

Our map α is monomial BUT NOT generically finite. However, it is very close to being generically finite. We generalize the [ST] formula to obtain multiplicities in $\mathcal{T}\mathcal{M}$.

Main results

$$\begin{array}{ccc} (\mathbb{C}^*)^n \supseteq V & \xrightarrow{\alpha} & W \subseteq (\mathbb{C}^*)^d \\ \pi \downarrow & & \downarrow \pi \\ V' = V/H & \xrightarrow{\bar{\alpha}} & W/\alpha(H), \end{array}$$

where $H = \Lambda \otimes_{\mathbb{Z}} \mathbb{C}^* \sim (\mathbb{C}^*)^{\dim \Lambda}$.

Theorem (— -Tobis-Yu)

Let $V \subset (\mathbb{C}^*)^n$ be a subvariety with torus action given by a lattice Λ and take the quotient by this action $V' = V/H$.
Assume that $\Lambda' = A(\Lambda)$ is a primitive sublattice of $\mathbb{Z}^d$ and that $\bar{\alpha}$ is generically finite on V' of degree δ. Then:

$$m_w = \frac{1}{\delta} \sum_{\substack{\pi(v) \\ A \cdot v = w}} m_v \cdot index(\mathbb{L}_w \cap \mathbb{Z}^d : A(\mathbb{L}_v \cap \mathbb{Z}^n)).$$

We assume that the number of such $\pi(v)$ is finite, all of them are regular in $\mathcal{T}V$, and $\mathbb{L}_v, \mathbb{L}_w$ are local linear spans of $v \in \mathcal{T}V$ and $w \in A(\mathcal{T}V)$.

The Newton polytope of the implicit equation

KEY: We can recover the *Newton polytope of* f from $\mathcal{T}(f)$ given as a collection of cones *with multiplicities*.

1. $\mathcal{T}(f)$ is the union of the codim 1 cones of the *normal fan of* $\mathrm{NP}(f)$.
2. the multiplicity of a maximal cone is the lattice length of the edge of $\mathrm{NP}(f)$ normal to that cone.

The Newton polytope of the implicit equation

KEY: We can recover the *Newton polytope of* f from $\mathcal{T}(f)$ given as a collection of cones *with multiplicities*.

1. $\mathcal{T}(f)$ is the union of the codim 1 cones of the *normal fan of* $\mathrm{NP}(f)$.
2. the multiplicity of a maximal cone is the lattice length of the edge of $\mathrm{NP}(f)$ normal to that cone.

Theorem (Dickenstein-Feichtner-Sturmfels)

Suppose $w \in \mathbb{R}^n$ *is a generic vector so that the ray* $(w - \mathbb{R}_{>0}\, e_i)$ *intersects* $\mathcal{T}(f)$ *only at regular points of* $\mathcal{T}(f)$*, for all* i*. Let* $\mathcal{P}^w$ *be the vertex of the polytope* $\mathcal{P} = NP(f)$ *that attains the maximum of* $\{w \cdot x : x \in NP(f)\}$*. Then the* i^{th} *coordinate of* $\mathcal{P}^w$ *equals*

$$\mathcal{P}_i^w = \sum_v m_v \cdot |l_{v,i}|,$$

where the sum is taken over all points $v \in \mathcal{T}(f) \cap (w - \mathbb{R}_{>0} e_i)$*,* m_v *is the multiplicity of* v *in* $\mathcal{T}(f)$*, and* $l_{v,i}$ *is the* i^{th} *coordinate of the primitive integral normal vector to* $\mathcal{T}(f)$ *at* v*.*

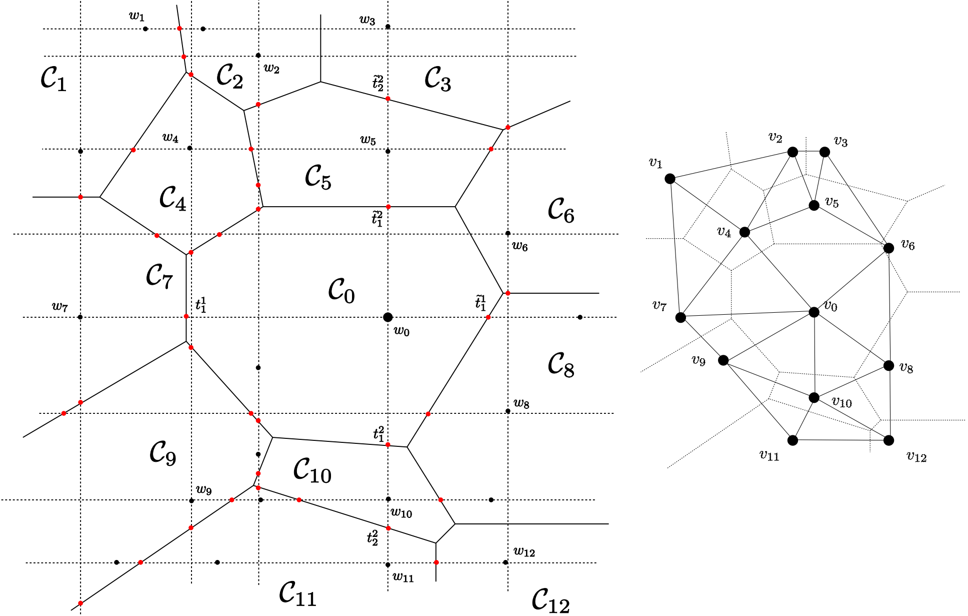


Figure: Ray-shooting and walking algorithms combined. Starting from chamber $\mathcal{C}_0$ we shoot and walk from chamber to chamber, and from vertex to vertex in $\mathrm{NP}(f)$.

The Newton polytope of the implicit equation

Theorem (— -Tobis-Yu)

The hypersurface $\mathcal{M}$ *has multidegree* $(110, 55, 55, 55, 55)$ *with respect to the grading defined by the matrix*

$$\Lambda = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Question: Is there hope of computing $\mathrm{NP(f)}$ by iterating Ray-shooting?
Bottleneck: Going through the list of cones in the **tropical repres.** of $\mathcal{TM}$ ($\sim 7\,000\,000$).

We can do better! $\rightsquigarrow$ Shoot rays and walk from chamber to chamber.

Theorem (— -Tobis-Yu)

The Newton polytope of f *has* $17\,214\,912$ *vertices in* $44\,938$ *orbits and* $70\,646$ *facets in* 246 *orbits under the symmetry group* B_4*.*

Certifying the Newton polytope of the implicit equation

Given $\mathcal{S}$ a (partial) list of vertices of $\mathrm{NP}(f)$, we construct

$$\mathcal{Q} = conv\, hull(\mathcal{S}).$$

FACT: $\mathcal{Q} = \mathrm{NP}(f) \iff$ all facets of $\mathcal{Q}$ are facets of $\mathrm{NP}(f)$.

Certifying the Newton polytope of the implicit equation

Given $\mathcal{S}$ a (partial) list of vertices of $\mathrm{NP}(f)$, we construct

$$\mathcal{Q} = conv\,hull(\mathcal{S}).$$

FACT: $\mathcal{Q} = \mathrm{NP}(f) \Longleftrightarrow$ all facets of $\mathcal{Q}$ are facets of $\mathrm{NP}(f)$.

Lemma

Let $w \in \mathbb{R}^n$ and $\mathcal{T}(f)$ be a tropical representation of the hypersurface. Let d be the dimension of its lineality space. Let $\mathcal{H} = \{\sigma_1, \ldots, \sigma_l\}$ be the list of cones containing w. Let q_i be the normal vector to the cone σ_i for $i = 1, \ldots, l$. TFAE:

- *w is a ray of $\mathcal{T}(f)$,*
- $\dim_{\mathbb{R}} \mathbb{R}\langle q_1, \ldots, q_l\rangle = n - d - 1$,
- *w is a facet direction of $NP(f)$.*

Completing the polytope

Definition

$\mathcal{P} \subset \mathbb{R}^N$ full dim'l and v vertex of $\mathcal{P}$. The tangent cone of $\mathcal{P}$ at v is:

$\mathcal{T}_v^{\mathcal{P}} := v + \mathbb{R}_{\geq 0}\langle w - v : w \in \mathcal{P}\rangle = v + \mathbb{R}_{\geq 0}\langle e : e \text{ edge of } \mathcal{P} \text{ adjacent to } v\rangle$.

Remark

- $\mathcal{T}_v^{\mathcal{P}}$ *is a polyhedron with only ONE vertex* v.
- $\mathcal{P} = \bigcap_{v \text{ vertex of } \mathcal{P}} \mathcal{T}_v^{\mathcal{P}}$.
- *Facet directions of* $\mathcal{P}$ *are facet directions in* $\mathcal{T}_v^{\mathcal{P}}$ *for some vertex* v.
- $\mathcal{T}_v^{\mathcal{Q}} \subseteq \mathcal{T}_v^{\mathcal{P}}$ *and if* $\mathcal{T}_v^{\mathcal{Q}} = \mathcal{T}_v^{\mathcal{P}}$ *then the extremal rays of* $\mathcal{T}_v^{\mathcal{Q}}$ *are edge directions of* $\mathcal{P}$*. We have these edge directions from* $\mathcal{T}(f)$ *(*$\sim 15\,788$*).*

Completing the polytope

Definition

$\mathcal{P} \subset \mathbb{R}^N$ full dim'l and v vertex of $\mathcal{P}$. The tangent cone of $\mathcal{P}$ at v is:

$\mathcal{T}_v^{\mathcal{P}} := v + \mathbb{R}_{\geq 0}\langle w - v : w \in \mathcal{P}\rangle = v + \mathbb{R}_{\geq 0}\langle e : e$ edge of $\mathcal{P}$ adjacent to $v\rangle$.

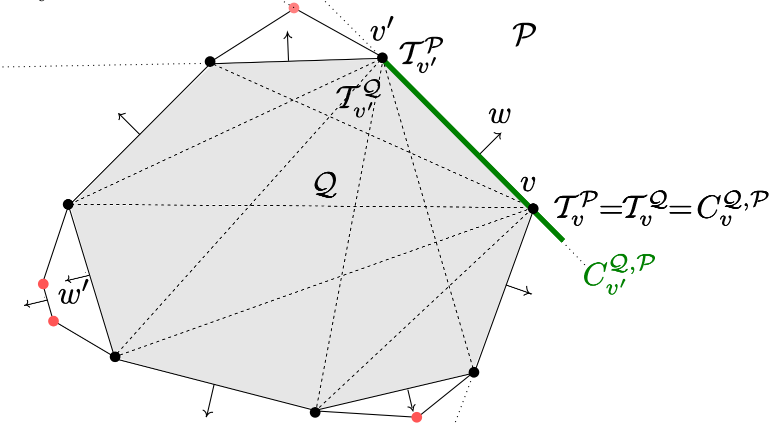
Remark

- *$\mathcal{T}_v^{\mathcal{P}}$ is a polyhedron with only ONE vertex v.*
- *$\mathcal{P} = \bigcap_{v \text{ vertex of } \mathcal{P}} \mathcal{T}_v^{\mathcal{P}}$.*
- *Facet directions of $\mathcal{P}$ are facet directions in $\mathcal{T}_v^{\mathcal{P}}$ for some vertex v.*
- *$\mathcal{T}_v^{\mathcal{Q}} \subseteq \mathcal{T}_v^{\mathcal{P}}$ and if $\mathcal{T}_v^{\mathcal{Q}} = \mathcal{T}_v^{\mathcal{P}}$ then the extremal rays of $\mathcal{T}_v^{\mathcal{Q}}$ are edge directions of $\mathcal{P}$. We have these edge directions from $\mathcal{T}(f)$ ($\sim 15\,788$).*

Definition

$$C_v^{\mathcal{Q},\mathcal{P}} := v + \mathbb{R}_{\geq 0}\langle w - v : w \text{ vertex of } \mathcal{Q}, w - v \sim \text{edge of } \mathcal{P}\rangle \subset \mathcal{T}_v^{\mathcal{Q}}.$$

- In practice: number of generating rays in $C_v^{\mathcal{Q},\mathcal{P}}$ is about 30 (vs. 17 million rays for $\mathcal{T}_v^{\mathcal{Q}}$!).
- Can test $C_v^{\mathcal{Q},\mathcal{P}} \supset \mathcal{T}_v^{\mathcal{Q}}$ by computing facets of $C_v^{\mathcal{Q},\mathcal{P}}$ with `Polymake`.
- If $C_v^{\mathcal{Q},\mathcal{P}} = \mathcal{T}_v^{\mathcal{Q}}$, we can test if its facet directions are facet directions of $\mathcal{T}_v^{\mathcal{P}}$ using our lemma.



- Last: certify that the facet with direction w in $\mathcal{T}_v^{\mathcal{Q}}$ is supported on v. We can do this by using ray-shooting with perturbed w.