An Implicitization Challenge for Binary Factor Analysis

María Angélica Cueto¹ Enrique Tobis² Josephine Yu³

¹Department of Mathematics University of California, Berkeley

²Department of Mathematics University of Buenos Aires, Argentina

³MSRI

Tropical Grad. Student Seminar - MSRI

- Algebraic Statistics: description of the model.
- Geometry of the model: First Secants of Segre embeddings and Hadamard products.
- S Tropicalization of the model.
- Main results.
- Implicitization Task: build the Newton polytope.

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) form an algebraic variety \mathcal{M} inside Δ_{15} with expected codimension one and (multi)homogeneous defining equation f.

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) form 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 $\mathcal M$

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.

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...

Fact

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

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

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

Corollary

The algebraic variety of the model is $\mathcal{M} = X \cdot X$ 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.

Corollary

The algebraic variety of the model is $\mathcal{M} = X \cdot X$ 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:

• 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 .

• For today: MAX CONVENTION.

Remark

Basic features of $\mathcal{T}(X)$ for $X \subset \mathbb{P}^n$ with homogeneous ideal I = I(X):

- $\mathcal{T}(X)$ is a fan (constant coefficients case).
- **2** The lineality space of the fan $\mathcal{T}(X)$ is the set

$$L = \{ w \in \mathcal{T}(X) : in_w(I) = I \}.$$

It describes action of the maximal torus acting on X (diagonal action by the lattice $L \cap \mathbb{Z}^{n+1}$).

Morphisms can be tropicalized and monomial maps have very nice tropicalizations.

Theorem (Sturmfels-Tevelev-Yu)

Let $A \in \mathbb{Z}^{d \times n}$, defining a monomial map $\alpha : (\mathbb{C}^*)^n \to (\mathbb{C}^*)^d$ and a canonical linear map $A : \mathbb{R}^n \to \mathbb{R}^d$. 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 of degree δ , we have an explicit formula to push forward the multiplicities of $\mathcal{T}(V)$ to multiplicities of $\mathcal{T}(\alpha(V))$. The multiplicity of $\mathcal{T}(\alpha(V))$ at a regular point w equals

$$m_w = \frac{1}{\delta} \cdot \sum_v m_v \cdot \text{ index } (\mathbb{L}_w \cap \mathbb{Z}^d : A(\mathbb{L}_v \cap \mathbb{Z}^n)),$$

where the sum is over all points $v \in \mathcal{T}(V)$ with Av = w. We also assume that the number of such v is finite, all of them are regular in $\mathcal{T}(V)$, and $\mathbb{L}_v, \mathbb{L}_w$ are linear spans of neighborhoods of $v \in \mathcal{T}(V)$ and $w \in A(\mathcal{T}(V))$ respectively.

Main results

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$, and α is the monomial map associated to matrix $(Id_{16} | Id_{16})$.

Main results

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$, and α is the monomial map associated to 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 complexes, with $m_{\sigma \times \tau} = m_{\sigma}m_{\tau}$ for maximal cones $\sigma \subset \mathcal{T}(X), \tau \subset \mathcal{T}(Y)$, and $\sigma \times \tau \subset \mathcal{T}(X \times Y)$.

Theorem (— -Tobis-Yu)

Given $X, Y \subset \mathbb{P}^n$ two projective irreducible varieties none of which is contained in a proper coordinate hyperplane, we can consider the associated projective variety $X \cdot Y \subset \mathbb{P}^n$. Then as sets:

$$\mathcal{T}(X \cdot Y) = \mathcal{T}(X) + \mathcal{T}(Y).$$

 $\mathcal{T}(X)$ can be computed with <code>Gfan.</code> In particular,

- 10-dim. simplicial fan in \mathbb{R}^{16} ,
- 5-dim. lineality space,
- f-vector= (381, 3436, 11236, 15640, 7680),
- 13 rays and 49 maximal cones up to B_4 -symmetry.

 $\mathcal{T}(X)$ can be computed with <code>Gfan.</code> In particular,

- 10-dim. simplicial fan in \mathbb{R}^{16} ,
- 5-dim. lineality space,
- f-vector= (381, 3436, 11236, 15640, 7680),
- 13 rays and 49 maximal cones up to B_4 -symmetry.

Thus we know $\mathcal{T}(\mathcal{M})$ as a set!

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

 $\mathcal{T}(X)$ can be computed with <code>Gfan.</code> In particular,

- 10-dim. simplicial fan in \mathbb{R}^{16} ,
- 5-dim. lineality space,
- f-vector= (381, 3436, 11236, 15640, 7680),
- 13 rays and 49 maximal cones up to B_4 -symmetry.

Thus we know $\mathcal{T}(\mathcal{M})$ as a set!

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

 $\mathcal{T}(X)$ can be computed with <code>Gfan.</code> In particular,

- 10-dim. simplicial fan in \mathbb{R}^{16} ,
- 5-dim. lineality space,
- f-vector= (381, 3436, 11236, 15640, 7680),
- 13 rays and 49 maximal cones up to B_4 -symmetry.

Thus we know $\mathcal{T}(\mathcal{M})$ as a set!

- \bullet Dimension = 15 in $\mathbb{C}^{16},$ so $\mathcal M$ is a hypersurface!
- Number of maximal cones in $T(X) + T(X) = 6\,865\,824$.
- $18\,972$ 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 previous Theorem by [STY] to obtain multiplicities in $\mathcal{T}(\mathcal{M})$.

M.A. Cueto et al. (UC Berkeley)

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. Then,

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

Moreover, if $\Lambda' = A(\Lambda)$ is a primitive sublattice of \mathbb{Z}^d and if $\bar{\alpha}$ induces a generically finite morphism on V' of degree δ , we have an explicit formula to push forward the multiplicities of $\mathcal{T}(V)$ to $\mathcal{T}(\alpha(V))$:

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

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*.

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*.

- T(f) is the union of the codim 1 cones of the normal fan of NP(f).
- Image: multiplicity of a maximal cone is the lattice length of the edge of NP(f) normal to that cone.

The Newton polytope of the implicit equation

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

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

M.A. Cueto et al. (UC Berkeley)

The hypersurface M has multidegree (110, 55, 55, 55, 55) with respect to the grading defined by the matrix



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



Question: Is there hope of computing NP(f) by iterating Ray-shooting? Bottleneck: Going through the list of all maximal cones supporting T(M) (~ 7000000).

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



Question: Is there hope of computing NP(f) by iterating Ray-shooting? Bottleneck: Going through the list of all maximal cones supporting T(M) (~ 7000000).

We can do better!

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

	(1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	$1 \rangle$	
	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	
$\Lambda =$	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 /	

Question: Is there hope of computing NP(f) by iterating Ray-shooting? Bottleneck: Going through the list of all maximal cones supporting T(M) (~ 7000000).

We can do better! \rightsquigarrow **IDEA:** Shoot rays and walk along neighboring chambers.

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

	(1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	$1 \rangle$	
	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	
$\Lambda =$	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 /	

Question: Is there hope of computing NP(f) by iterating Ray-shooting? Bottleneck: Going through the list of all maximal cones supporting T(M) (~ 7000000).

We can do better! \rightsquigarrow **IDEA:** Shoot rays and walk along neighboring chambers.

We obtain $15\,837\,696$ vertices, grouped in $41\,348$ orbits.

M.A. Cueto et al. (UC Berkeley)



Figure: Ray-shooting and walking algorithms combined. Starting from chamber C_0 we shoot and walk from chamber to chamber, and from vertex to vertex in NP(f).

Certifying the Newton polytope of the implicit equation

Given S a (partial) list of vertices of NP(f), we construct Q = conv hull(S).

QUESTION: When do we have Q = NP(f)?

Answer: If all facets of \mathcal{Q} are facets of NP(f).

Certifying the Newton polytope of the implicit equation

Given S a (partial) list of vertices of NP(f), we construct Q = conv hull(S).

QUESTION: When do we have Q = NP(f)? **Answer:** Iff all facets of Q are facets of NP(f).

Lemma

Let $w \in \mathbb{R}^n$ and $\mathcal{T}(f)$ be a tropical hypersurface given by a collection of cones, but with no prescribed fan structure. 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 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 P. The tangent cone of \mathcal{P} at v is:

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

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 $\mathcal{T}(f)$ (15788 in total).

Completing the polytope

Definition

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

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

Remark

•
$$\mathcal{T}_v^{\mathcal{P}}$$
 is a polyhedron with only ONE vertex (v) .

•
$$\mathcal{P} = \bigcap_{v \text{ vertex of } \mathcal{P}} T_v^{\mathcal{P}}.$$

- 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 $\mathcal{T}(f)$ (15788 in total).

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^{Q,P} is about 30 (vs. 15 million rays for T_v^Q!).
- Can test $C_v^{\mathcal{Q},\mathcal{P}} \supset T_v^{\mathcal{Q}}$ by computing facets of $C_v^{\mathcal{Q},\mathcal{P}}$ with Polymake.
- If $C_v^{Q,\mathcal{P}} = \mathcal{T}_v^Q$ can test if facet directions are facet directions of $\mathcal{T}_v^{\mathcal{P}}$ by our Lemma.



• Last: certify that facet with facet direction w in $\mathcal{T}_v^{\mathcal{Q}}$ is supported on v. Can do this by Ray-shooting with perturbed w.

M.A. Cueto et al. (UC Berkeley)

An Implicitization Challenge