An Implicitization Challenge for Binary Factor Analysis

María Angélica Cueto\textsuperscript{1} \quad Josephine Yu\textsuperscript{2}

\textsuperscript{1}Department of Mathematics\newline University of California, Berkeley

\textsuperscript{2}Department of Mathematics\newline Massachusetts Institute of Technology

MEGA’09
Outline

1. Algebraic Statistics: description of the model.
2. Geometry of the model: First Secants of Segre embeddings and Hadamard products.
3. Tropicalization of the model.
4. Main results.
5. Implicitization Task: build the Newton polytope.
The Statistical 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 $\Delta_{15}$ with expected codimension one and (multi)homogeneous defining equation $f$.

Figure: The undirected graphical model $\mathcal{F}_{4,2}$. 
The Statistical 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 $\Delta_{15}$ with expected codimension one and (multi)homogeneous defining equation $f$.

**Problem**

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

Parameterization of the model:  \( p : \mathbb{R}^{32} \rightarrow \mathbb{R}^{16}, \)

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

Using homogeneity and the distributive law

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

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.
Geometry of the model

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

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

Using homogeneity and the distributive law

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

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. But...
Fact

1. The binary 4-claw tree model is $\text{Sec}^1(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \subset \mathbb{P}^{15}$.
2. 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

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

Proposition

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. Invariant under relabeling of observed nodes and by changing role of two states (0 or 1). Therefore, we have an action of the group $B_4 = S_4 \ltimes (S_2)^4$, the group of symmetries of the 4-cube.
Geometry of the model

Proposition

The algebraic variety of the model is \( M = X \cdot X \) 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. Invariant under relabeling of observed nodes and by changing role of two states (0 or 1). Therefore, we have an action of the group \( B_4 = S_4 \rtimes (S_2)^4 \), 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 subvariety of all \( 2 \times 2 \times 2 \times 2 \)-tensors of tensor rank at most 2.
2. Known set of generators for \( I(X) \): \( 3 \times 3 \)-minors of all three \( 4 \times 4 \)-flattenings of these tensors.
Tropicalizing the model

Definition

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

$$\mathcal{T}(X) = \mathcal{T}(I) = \{ w \in \mathbb{R}^{n+1} | \text{in}_w(I) \text{ contains no monomial} \}$$

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

**Definition**

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

\[
T(X) = T(I) = \{ w \in \mathbb{R}^{n+1} \mid \text{in}_w(I) \text{ contains no monomial} \}
\]

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

**Example**

\( L = (x + y + 1 = 0) \subset \mathbb{C}^2 \)

gives the well-known picture:
Remark

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

1. It is a rational polyhedral subfan of the Gröbner fan of $I$.
2. If $I$ is prime, then $T(X)$ is pure of the same dimension as $X$ (Bieri-Groves Thm) and it is connected in codimension one.
3. Maximal cones have canonical multiplicities attached to them. With these multiplicities, $T(X)$ satisfies the balancing condition.
4. The lineality space of the fan $T(X)$ is the set
   
   $$ L = \{ w \in T(X) : \text{in}_w(I) = I \}. $$

   It describes action of the maximal torus acting on $X$ (diagonal action by the lattice $L \cap \mathbb{Z}^{n+1}$.)
5. Morphisms can be tropicalized and monomial maps have very nice tropicalizations.
**Theorem (S-T-Y)**

Let $A \in \mathbb{Z}^{d \times n}$, defining a monomial map $\alpha : (\mathbb{C}^*)^n \rightarrow (\mathbb{C}^*)^d$ and a canonical linear map $A : \mathbb{R}^n \rightarrow \mathbb{R}^d$.

Let $V \subset (\mathbb{C}^*)^n$ be a subvariety. Then

$$T(\alpha(V)) = A(T(V)).$$

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

In our case \( \mathcal{M} = X \cdot X = \alpha(X \times X) \) where \( \alpha \) is the monomial map associated to matrix \((\text{Id}_{16} \mid \text{Id}_{16})\).
Main results

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$ where $\alpha$ is the monomial map associated to matrix $(\text{Id}_{16} | \text{Id}_{16})$. In general...
Main results

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$ where $\alpha$ is the monomial map associated to matrix $(\text{Id}_{16} \mid \text{Id}_{16})$. In general...

**Theorem (—but, 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:

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

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$ where $\alpha$ is the monomial map associated to matrix $(\text{Id}_{16} \mid \text{Id}_{16})$. In general...

**Theorem (—, 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 Gfan, so we know $\mathcal{T}(\mathcal{M})$ as a set! BUT we want more...
Main results

In our case \( \mathcal{M} = X \cdot X = \alpha(X \times X) \) where \( \alpha \) is the monomial map associated to matrix \((\text{Id}_{16} | \text{Id}_{16})\). In general...

**Theorem (―, 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:

\[
T(X \cdot Y) = T(X) + T(Y).
\]

\( T(X) \) can be computed with \texttt{Gfan}, so we know \( T(\mathcal{M}) \) as a set! BUT we want more...

We want to compute multiplicities at regular points of \( T(\mathcal{M}) \).
Main results

In our case $\mathcal{M} = X \cdot X = \alpha(X \times X)$ where $\alpha$ is the **monomial map** associated to matrix $(Id_{16} \mid Id_{16})$. In general...

---

**Theorem (—, 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**:

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

$\mathcal{I}(X)$ can be computed with Gfan, so we know $\mathcal{I}(\mathcal{M})$ as a **set**! BUT we want more...

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

Our map $\alpha$ is monomial BUT NOT generically finite but very close to being gen. finite. We generalize the previous theorem to obtain multiplicities in $\mathcal{I}(\mathcal{M})$...
Let $V \subset (\mathbb{C}^*)^n$ be a subvariety with torus action given by a lattice $L$ and take the quotient by this action $V' = V/H$. Then,

$$\mathcal{T}(\bar{\alpha}(V')) = A'(\mathcal{T}(V')).$$

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

\[ \mathbb{T}^n \supset V \xrightarrow{\alpha} W \subseteq \mathbb{T}^d \]
\[ \pi \downarrow \quad \pi \]
\[ V' = V/H \xrightarrow{\bar{\alpha}} W/\alpha(H). \]
Theorem (—, Yu)

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

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

as weighted polyhedral complexes, with $m_{\sigma \times \tau} = m_\sigma m_\tau$ for maximal cones $\sigma \subset \mathcal{I}(X), \tau \subset \mathcal{I}(Y)$, and $\sigma \times \tau \subset \mathcal{I}(X \times Y)$. 

If \( I = (f) \), we can recover the Newton polytope of \( f \) from \( \mathcal{T}(I) \).
If \( I = (f) \), we can recover the *Newton polytope of f* from \( \mathcal{T}(I) \). Why?
The Newton polytope of the implicit equation

If $I = (f)$, we can recover the Newton polytope of $f$ from $\mathcal{T}(I)$. Why?

1. $\mathcal{T}(I)$ is the union of the codim 1 cones of the normal fan of $NP(f)$.
2. 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

If $I = (f)$, we can recover the Newton polytope of $f$ from $\mathcal{T}(I)$. Why?

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

**Theorem (D-F-S)**

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 $P^w$ be the vertex of the polytope $P = NP(f)$ that attains the maximum of $\{w \cdot x : x \in NP(f)\}$. Then the $i^{th}$ coordinate of $P^w$ equals

$$P^w_i = \sum_v m_v \cdot |l_{v,i}|,$$

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 $v$. 
The Newton polytope of the implicit equation

**Theorem**

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

$$L = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}.$$
The Newton polytope of the implicit equation

**Theorem**

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

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

**Bottleneck:** Going through the list of all maximal cones supporting $\mathcal{T}(\mathcal{M})$ ($\sim 7\,000\,000$.)

M.A. Cueto and J. Yu (UC Berkeley, MIT)  An Implicitization Challenge  June 15th 2009  14 / 15
The Newton polytope of the implicit equation

**Theorem**

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

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

**Bottleneck:** Going through the list of all maximal cones supporting $\mathcal{T}(\mathcal{M})$ ($\sim 7 \, 000 \, 000$.)

*We can do better!*
The Newton polytope of the implicit equation

**Theorem**

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

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

**Bottleneck:** Going through the list of all maximal cones supporting $T(M)$ ($\sim 7000000$.)

**We can do better!**

**IDEA:** Shoot rays and walk along neighboring chambers.
Theorem

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

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

Bottleneck: Going through the list of all maximal cones supporting $\mathcal{T}(\mathcal{M})$ ($\sim 7\,000\,000$.)

We can do better!

IDEA: Shoot rays and walk along neighboring chambers.

Up to now, we have computed 1,155,072 vertices of $NP(f)$ (3,030 orbits.)
Thank you!!!