Chordal structure in computer algebra: Permanents

Diego Cifuentes

Laboratory for Information and Decision Systems
Electrical Engineering and Computer Science
Massachusetts Institute of Technology

Joint work with **Pablo A. Parrilo** (MIT)

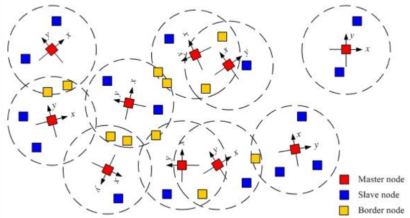
SIAM Conference on Discrete Mathematics - 2016

Algebraic tools in discrete math

Several problems from discrete mathematics can be approached with tools from computer algebra.

Sensor network localization:
Find positions, given a few known fixed anchors and pairwise distances.

$$\|x_i - x_j\|^2 = d_{ij}^2 \quad ij \in \mathcal{A}$$
$$\|x_i - a_k\|^2 = e_{ij}^2 \quad ik \in \mathcal{B}$$



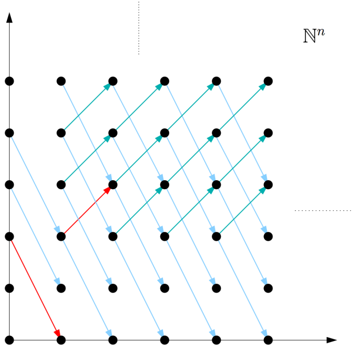
This is a system of quadratic polynomial equations. Can be solved using *Gröbner bases*.

Algebraic tools in discrete math

Several problems from discrete mathematics can be approached with tools from computer algebra.

Lattice walks
Let $\mathcal{A} \subset \mathbb{Z}^n$ consist of integer vectors. Consider a graph with vertex set $\mathbb{N}^n$ in which α, β are adjacent if $\alpha - \beta \in \mathcal{A}$. Describe the connected components.

This can be solved using a *primary decomposition*.



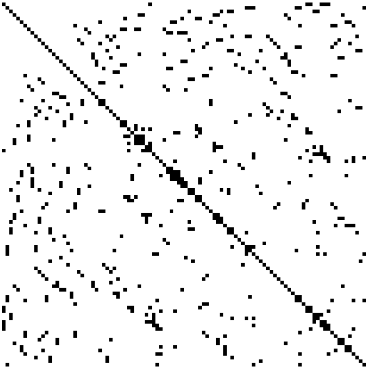
Algebraic tools in discrete math

Several problems from discrete mathematics can be approached with tools from computer algebra.

Permanents (this talk)
Given a $n \times n$ matrix M, compute

$$\mathrm{Perm}(M) := \sum_{\pi} \prod_{i} M_{i,\pi(i)}$$

sum over permutations π.
Important case: sparse matrices.



Algebraic tools in discrete math

- Many more problems can be naturally approached with algebraic tools: graph colorings, independent sets, Hamiltonian cycles, crypto problems, etc.

Algebraic tools in discrete math

- Many more problems can be naturally approached with algebraic tools: graph colorings, independent sets, Hamiltonian cycles, crypto problems, etc.
- However these problems are NP-hard, and these algebraic methods do not work well for relatively small problems.

Algebraic tools in discrete math

- Many more problems can be naturally approached with algebraic tools: graph colorings, independent sets, Hamiltonian cycles, crypto problems, etc.
- However these problems are NP-hard, and these algebraic methods do not work well for relatively small problems.
- Idea: exploit the graphical structure (chordality) of the problem!

Algebraic tools in discrete math

- Many more problems can be naturally approached with algebraic tools: graph colorings, independent sets, Hamiltonian cycles, crypto problems, etc.
- However these problems are NP-hard, and these algebraic methods do not work well for relatively small problems.
- Idea: exploit the graphical structure (chordality) of the problem!
- Complexity aspects? Identify families of tractable instances.

Permanent computation

General facts:

- (Ryser'63) Best general and exact method; complexity $O(n\,2^n)$.
- (Valiant'79) Computing the permanent of a matrix is $\#\mathrm{P}$-hard.

Permanent computation

General facts:

- (Ryser'63) Best general and exact method; complexity $O(n\,2^n)$.
- (Valiant'79) Computing the permanent of a matrix is $\#\mathrm{P}$-hard.

Approximation algorithms:

- (Jerrum,Sinclair,Vigoda'04) FPRAS for nonnegative matrices.

Permanent computation

General facts:

- (Ryser'63) Best general and exact method; complexity $O(n\,2^n)$.
- (Valiant'79) Computing the permanent of a matrix is $\#\mathrm{P}$-hard.

Approximation algorithms:

- (Jerrum,Sinclair,Vigoda'04) FPRAS for nonnegative matrices.

Tractable under special structure:

- (Fisher,Kasteleyn,Temperley'67) Bipartite graph is planar.
- (Barvinok'96) Rank is bounded.
- (Courcelle et al'01) Treewidth is bounded.

Graph abstractions of a matrix

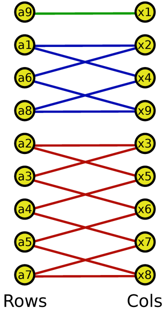
The sparsity structure of a matrix M can be represented with a graph. Let $a_1, \ldots, a_n$ denote the rows and $x_1, \ldots, x_n$ denote the columns. Consider these abstractions:

- G, bipartite adjacency graph: $(a_i, x_j) \in E$ iff $M_{i,j} \neq 0$
- G^s, (symmetrized) adjacency graph: $(i, j) \in E$ iff $|M_{i,j}| + |M_{j,i}| \neq 0$
- G^X, column graph: $(x_j, x_k) \in E$ iff exists a_i such that $M_{i,j} M_{i,k} \neq 0$. Equivalently, the adjacency graph of $M^T M$.

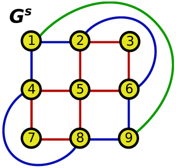
Graph abstractions of a matrix

$$
\begin{array}{c}
 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \end{array} \\
\begin{pmatrix}
0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 0 \\
0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\
0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 1 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

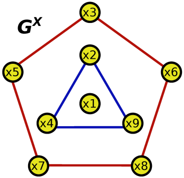
G



G^s



G^x

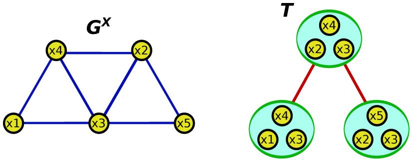


Tree decompositions

Definition: A *tree decomposition* of a graph $G = (X, E)$ is a pair (T, χ), where T is a tree and $\chi : T \to \{0,1\}^X$ assigns some $\chi(t) \subset X$ to each node t of T, that satisfies the following conditions.

i. The union of $\{\chi(t)\}_{t \in T}$ is the whole vertex set X.
ii. For every $(x_i, x_j) \in E$, there is some node t of T with $x_i, x_j \in \chi(t)$.
iii. For every $x_i \in X$ the set $\{t : x_i \in \chi(t)\}$ forms a subtree of T.

The *width* ω of the decomposition is the largest $|\chi(t)|$.



Chordality and treewidth

The *treewidth* of G is the minimum width among all possible tree decompositions. The treewidth $\omega(G)$ of a graph G can be though of as a measure of complexity: the smaller $\omega(G)$, the simpler the graph (closer to a tree).

Meta-theorem: NP-complete problems are "easy" on graphs of small treewidth.

Chordality and treewidth

The *treewidth* of G is the minimum width among all possible tree decompositions. The treewidth $\omega(G)$ of a graph G can be though of as a measure of complexity: the smaller $\omega(G)$, the simpler the graph (closer to a tree).

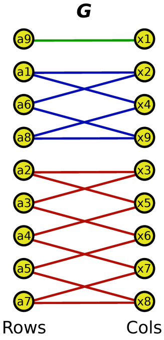
Meta-theorem: NP-complete problems are "easy" on graphs of small treewidth.

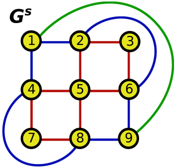
Alternatively, we can define treewidth in terms of chordal completions of the graph (chordal completions are in correspondance with tree decompositions).

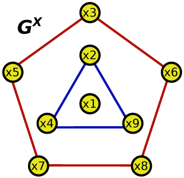
Treewidth of matrix graphs

Simple fact:

- $\operatorname{tw}(G) \leq 2\operatorname{tw}(G^s)$.
- $\operatorname{tw}(G) \leq \operatorname{tw}(G^X) + 1$.
- For a fixed $\operatorname{tw}(G)$ both $\operatorname{tw}(G^s), \operatorname{tw}(G^X)$ are unbounded.



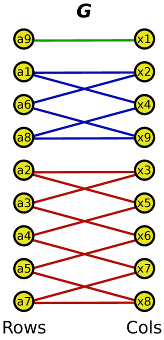


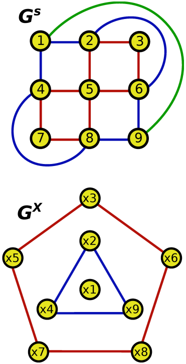


Treewidth of matrix graphs

Simple fact:

- $\mathrm{tw}(G) \leq 2\,\mathrm{tw}(G^s)$.
- $\mathrm{tw}(G) \leq \mathrm{tw}(G^X) + 1$.
- For a fixed $\mathrm{tw}(G)$ both $\mathrm{tw}(G^s), \mathrm{tw}(G^X)$ are unbounded.



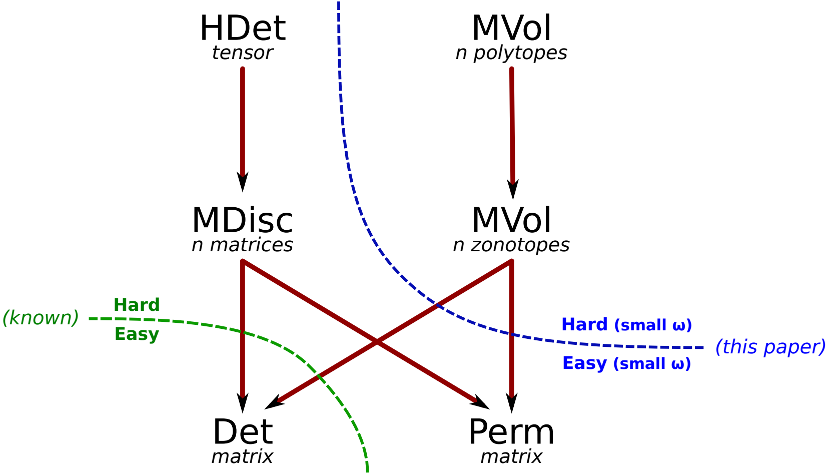


Previous tree decomposition methods are primarily based on the symmetrized graph G^s (Courcelle et al., Flarup et al., Meer), and they do not scale well with the treewidth $O(n\,2^{O(\omega^2)})$.

Our results

- A tree decomposition method to compute permanents, based on the bipartite graph G, with complexity $\widetilde{O}(n\,2^{\omega})$.
- The algorithm naturally extends to higher dimensional problems: mixed discriminants, hyperdeterminants. Complexity $\widetilde{O}(n^2 + n\,3^{\omega})$.
- Hardness results for the case of mixed volumes.

Our results


HDet
tensor
MVol
n polytopes
MDisc
n matrices
MVol
n zonotopes
(known)
Hard
Easy
Hard (small ω)
(this paper)
Easy (small ω)
Det
matrix
Perm
matrix

Basic idea

For block matrices

$$\operatorname{Perm}\begin{pmatrix} A & A' \\ 0 & B \end{pmatrix} = \operatorname{Perm}(A)\operatorname{Perm}(B) = \operatorname{Perm}\begin{pmatrix} 0 & A \\ B & B' \end{pmatrix}$$

Permanent expansion: column graph

Similarly, for the matrix

$$M = \begin{pmatrix} A_{1,1} & 0 & A_{1,3} & A_{1,4} & 0 \\ A_{2,1} & 0 & A_{2,3} & A_{2,4} & 0 \\ 0 & C_{3,2} & C_{3,3} & C_{3,4} & 0 \\ 0 & B_{4,2} & B_{4,3} & 0 & B_{4,5} \\ 0 & B_{5,2} & B_{5,3} & 0 & B_{5,5} \end{pmatrix}$$

we have the following formula

$$\begin{aligned} \mathrm{Perm}(M) = {} & \mathrm{Perm}\begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{2,1} & A_{2,3} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} C_{3,4} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} B_{4,2} & B_{4,5} \\ B_{5,2} & B_{5,5} \end{pmatrix} \\ & + \mathrm{Perm}\begin{pmatrix} A_{1,1} & A_{1,4} \\ A_{2,1} & A_{2,4} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} C_{3,3} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} B_{4,2} & B_{4,5} \\ B_{5,2} & B_{5,5} \end{pmatrix} \\ & + \mathrm{Perm}\begin{pmatrix} A_{1,1} & A_{1,4} \\ A_{2,1} & A_{2,4} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} C_{3,2} \end{pmatrix} \mathrm{Perm}\begin{pmatrix} B_{4,3} & B_{4,5} \\ B_{5,3} & B_{5,5} \end{pmatrix}. \end{aligned}$$

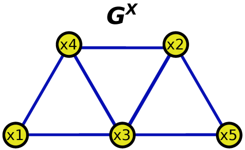
To evaluate we need 14 multiplications: four 2×2 permanetns.

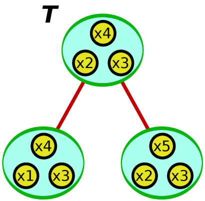
Compare with: $4 \times 5! = 480$ multiplications using definition.

Permanent expansion: column graph

This simple formula exists because its column graph has a simple structure

$$\begin{pmatrix} M_{1,1} & 0 & M_{1,3} & M_{1,4} & 0 \\ M_{2,1} & 0 & M_{2,3} & M_{2,4} & 0 \\ 0 & M_{3,2} & M_{3,3} & M_{3,4} & 0 \\ 0 & M_{4,2} & M_{4,3} & 0 & M_{4,5} \\ 0 & M_{5,2} & M_{5,3} & 0 & M_{5,5} \end{pmatrix}$$





Results

Lemma: Let (T, χ) be a tree decomposition of G^X. Let $c_1, \ldots, c_k$ be the children of a node $t \in T$. Then

$$perm(A_{T_t}, Y) = \sum_{\mathcal{Y}} perm(A_t, Y_t) \prod_{j=1}^{k} perm(A_{T_{c_j}}, Y_{c_j})$$

where the sum is over all $\mathcal{Y} = (Y_t, Y_{c_1}, \ldots, Y_{c_k})$ such that:

$$Y = Y_t \sqcup (Y_{c_1} \sqcup \cdots \sqcup Y_{c_k})$$

$$\chi(T_{c_j}) \setminus \chi(t) \subset Y_{c_j} \subset \chi(T_{c_j}) \qquad\qquad Y_t \subset \chi(t).$$

Lemma: The above formula can be evaluated in $\widetilde{O}(k\, 2^{\omega})$.

Results

Lemma: Let (T, χ) be a tree decomposition of G^X. Let $c_1, \ldots, c_k$ be the children of a node $t \in T$. Then

$$perm(A_{T_t}, Y) = \sum_{\mathcal{Y}} perm(A_t, Y_t) \prod_{j=1}^{k} perm(A_{T_{c_j}}, Y_{c_j})$$

where the sum is over all $\mathcal{Y} = (Y_t, Y_{c_1}, \ldots, Y_{c_k})$ such that:

$$Y = Y_t \sqcup (Y_{c_1} \sqcup \cdots \sqcup Y_{c_k})$$

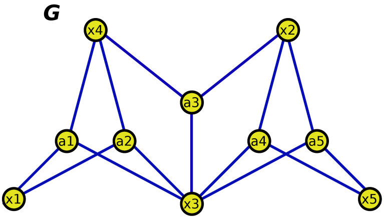
$$\chi(T_{c_j}) \setminus \chi(t) \subset Y_{c_j} \subset \chi(T_{c_j}) \qquad\qquad Y_t \subset \chi(t).$$

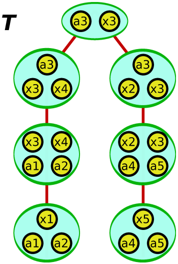
Lemma: The above formula can be evaluated in $\widetilde{O}(k\, 2^{\omega})$.
Proof: Rewrite the equation as

$$P_t(\bar{Y}) = \sum_{\bar{Y}_t \sqcup \bar{Y}_{c_1} \sqcup \cdots \sqcup \bar{Y}_{c_k} = \bar{Y}} Q_t(\bar{Y}_t) \prod_{j=1}^{k} Q_{c_j}(\bar{Y}_{c_j})$$

and use the fast subset convolution (Björklund et al. 07).

Permanent expansion: bipartite graph





$$\begin{aligned}\mathrm{Perm}(M) = {} & perm(\{a_1,a_2\},\{x_1,x_4\})\, perm(\{a_3,a_4,a_5\},\{x_2,x_3,x_5\}) \\ & + perm(\{a_1,a_2,a_3\},\{x_1,x_3,x_4\})\, perm(\{a_4,a_5\},\{x_2,x_5\}) \\ & - M_{3,3}\, perm(\{a_1,a_2\},\{x_1,x_4\})\, perm(\{a_4,a_5\},\{x_2,x_5\})\end{aligned}$$

$$perm(\{a_1,a_2,a_3\},\{x_1,x_3,x_4\}) = M_{3,3}\, perm(\{a_1,a_2\},\{x_1,x_4\}) + M_{3,4}\, perm(\{a_1,a_2\},\{x_1,x_3\}$$

$$perm(\{a_3,a_4,a_5\},\{x_2,x_3,x_5\}) = M_{3,3}\, perm(\{a_4,a_5\},\{x_2,x_5\}) + M_{3,2}\, perm(\{a_4,a_5\},\{x_3,x_5\}$$

To evaluate we need 16 multiplications.

Results

Theorem: Let M be a matrix with associated bipartite graph G. Then we can compute $\mathrm{Perm}(M)$ in $\widetilde{O}(n\,2^{\omega})$ where $\omega = \mathrm{tw}(G)$.

Results

Theorem: Let M be a matrix with associated bipartite graph G. Then we can compute $\mathrm{Perm}(M)$ in $\widetilde{O}(n\,2^{\omega})$ where $\omega = \mathrm{tw}(G)$.

The *mixed discriminant* of n matrices of size $n \times n$ is

$$\mathrm{Disc}(A^1, \ldots, A^n) := \sum_{\pi,\rho} \mathrm{sgn}(\pi)\mathrm{sgn}(\rho) \prod_i A^i_{\pi(i),\rho(i)}$$

where the sum is over pairs of permutations π, ρ.

Theorem: Let M be a list of matrices with associated tripartite graph G. Then we can compute $\mathrm{Disc}(M)$ in $\widetilde{O}(n^2 + n\,3^{\omega})$, where $\omega = \mathrm{tw}(G)$.

Further generalizations

The first Cayley hyperdeterminant is the simplest generalization of the determinant to multidimensional arrays.

Theorem: Let M be a square d-dimensional tensor of length n, with d-partite graph G. We can compute its *hyperdeterminant* in $\widetilde{O}(n^2 + n\,3^{\omega})$.

Further generalizations

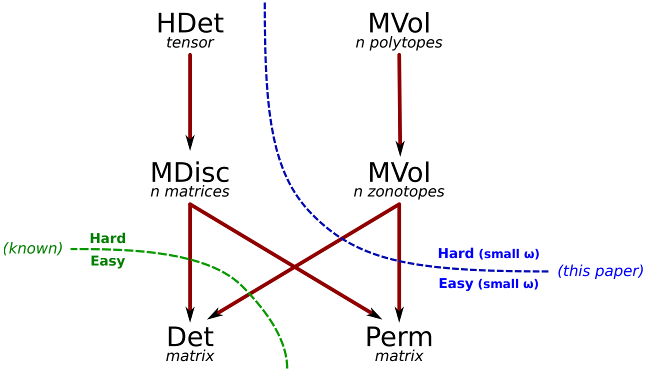
The first Cayley hyperdeterminant is the simplest generalization of the determinant to multidimensional arrays.

Theorem: Let M be a square d-dimensional tensor of length n, with d-partite graph G. We can compute its *hyperdeterminant* in $\widetilde{O}(n^2 + n\,3^\omega)$.

The mixed volume is a geometric generalization of determinants and permanents to arbitrary convex bodies on $\mathbb{R}^n$.

Theorem: Computing mixed volumes of n zonotopes remains $\#\mathrm{P}$-hard when their associated graph has bounded treewidth.

Summary



- D. Cifuentes, P.A. Parrilo, An efficient tree decomposition method for permanents and mixed discriminants. `arXiv:1507.03046`.
- D. Cifuentes, P.A. Parrilo, Exploiting chordal structure in polynomial ideals: a Gröbner basis approach. `arXiv:1411.1745`.
- D. Cifuentes, P.A. Parrilo, Chordal networks of polynomial ideals. `arXiv:1604.02618` .

Thanks for your attention!