Breaking e^n barrier for deterministic poly-time approximation of the permanent and settling Friedland's conjecture on the Monomer-Dimer Entropy

Leonid Gurvits and Alex Samorodnitsky(Hebrew University)

The City College, Dept. of Computer Science.
e-mail: l.n.gurvits@gmail.com

The Determinant, the Permanent and Subpermanents

The Determinant:

$$det(A) = \sum_{\sigma \in S_n} (-1)^{sign(\sigma)} \prod_{i=1}^{n} A(i, \sigma(i))$$

The Permanent:

$$per(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A(i, \sigma(i))$$

The sum of Subpermanents:

$$per_m(A) =: \sum_{|S|=|T|=m} per(A_{S,T}); per_0(A) = 1.$$

The Matching Polynomial(the bipartite case):

$$M_A(x) =: (-1)^n \sum_{0 \leq m \leq n} per_m(A)(-x)^{n-m}$$

The P-characteristic Polynomial:

$$C_A(t) =: per(tI + A)$$

Why the permanent(and its relatives) is interesting, important etc.

1. **Theoretical Computer Science**: **Permanent vs Determinant** i.e. VP vs VNP; it is Sharp-P-Complete; was and is a testing ground for many (randomized,deterministic) algorithms and many more...
2. **Combinatorics**: number of perfect matchings, number of matchings, number of matching of a given size, number of permutations with restricted positions, Latin Squares, Representation Theory...
3. The Monomer-Dimer Problem: Statistical Physics, Chemistry, Yang-Baxter Theory.....
4. **Computational Algebra** : the Mixed Volume(a generalization, will define later) counts(upper bounds) the number of isolated solutions of systems of polynomial equations,...

5. Approximating the permanent(or just getting the sign) of certain integer matrices is equivalent to **Quantum Computing**.

6. If U is a complex unitary matrix, i.e $UU^* = I$, then the polynomial

$p(x_1, ..., x_n) = per(U diag(x_1, ..., x_n)U^*)$ is a generating function of Linear Quantum Optics distributions. I.e. the coefficients of this homogeneous polynomial are non-negative and sum to one:

$$p(x_1, ..., x_n) = \sum_{\omega_1+...+\omega_n=n} a_{\omega_1,...,\omega_n} x_1^{\omega_1} ... x_n^{\omega_n},$$

$a_{\omega_1,...,\omega_n} \geq 0$ and $\Sigma_{\omega_1+...+\omega_n=n} a_{\omega_1,...,\omega_n} = 1$.

This polynomial happened to be doubly-stochastic:

$$\frac{\partial}{\partial x_i} p(1, 1, ..., 1) = 1, 1 \leq i \leq n.$$

$$det(AB) = det(A)det(B),$$

what about the permanent?

Associate with a square matrix A the following product polynomial: $\Sigma_{\omega_1+\ldots\omega_n=n} c^A_{\omega_1,\ldots,\omega_n} x_1^{\omega_1} \ldots x_n^{\omega_n} =$

$$= Prod_A(x_1, \ldots, x_n) =: \prod_{1 \leq i \leq n} \sum_{1 \leq j \leq n} A(i,j)x_j.$$

Then [Augustin-Louis Cauchy, 1812]:

$$per(AB^*) =< Prod_A, Prod_B >_F =$$

$$=: \sum_{\omega_1+\ldots+\omega_n=n} (\prod_{i=1}^{n} \omega_i!) c^A_{\omega_1,\ldots,\omega_n} \overline{c^B_{\omega_1,\ldots,\omega_n}}.$$

In other words the Fock Hilbert Space(thus the connection to Linear Quantum Optics) appeared in 1812! because of the permanent.

Fairly Inclusive Generalization of the Permanent (of non-negative matrices):

Given a family $A_1, ..., A_k$ of $n \times n$ complex matrices, define the following non-negative number

$$QP(A_1, ..., A_k) =: || \det(\sum_{1 \leq i \leq k} z_i A_i) ||_F^2$$

The permanent of non-negatives matrices: $A_{i,j} = a(i,j) e_i e_j^*$;

The mixed discriminant: $A_{i,j} = e_i v_{i,j}^*, v_{i,j} \in C^n$;

The permanent of PSD matrices: $A_{i,j}$ are diagonal.

Generalized Matching Polynomial:

Given a homogeneous polynomial $p \in Hom_+(m,n)$

$$p(x_1, ..., x_m) = \sum_{\omega_1 + ... \omega_m = n} a_{\omega_1,...,\omega_m} x_1^{\omega_1} ... x_m^{\omega_m}$$

such that $p(1,1,....,1) = 1$. Define the following univariate polynomial

$$M_p(x) = \sum_{\omega_1 + ... \omega_m = n} a_{\omega_1,...,\omega_m} x^{n-|\omega|_+} \prod_{\omega_i \neq 0} (x - \omega_i),$$

where $|\omega|_+$ is the number of positive coordinates in the vector $(\omega_1, ..., \omega_m)$.

Generalized Van Der Waerden-Schrijver like inequalities(and conjectures):
the permanent is the mixed derivative(aka polarization) of the product polynomial:

$$per(A) = \frac{\partial^n}{\partial x_1 ... \partial x_n} Prod_A(0),$$

where $Prod_A(x_1, ..., x_n) = \Pi_{1 \leq i \leq n} \Sigma_{1 \leq j \leq n} A(i,j)x_j$. The mixed discriminant is the mixed derivative(aka polarization) of the determinantal polynomial:

$$D(Q_1, ..., Q_n) =: \frac{\partial^n}{\partial x_1 ... \partial x_n} \det(\sum_{1 \leq j \leq n} x_j Q_j).$$

And the mixed volume is the mixed derivative of the Minkowski(volume) polynomial.

Those observations had led to the following problem: Given an evaluation oracle for a polynomial $p \in Hom_+(n,n)$, to compute/approximate within relative error its mixed derivative $\frac{\partial^n}{\partial x_1 ... \partial x_n} p(0)$... in poly$(n, cmpl(p))$ time, deterministic or randomized.

Define the following two numbers:

$$C - Cap(p) = \inf_{RE(z_i)>0} \frac{|p(z_1, ..., z_n)|}{\Pi_{1 \le i \le n} RE(z_i)},$$
$$Cap(p) = \inf_{x_i > 0} \frac{|p(x_1, ..., x_n)|}{\Pi_{1 \le i \le n} x_i}$$

Define also a sequence of polynomials(derivatives), kind of "polynomial hierarchy":

$$g_n = p, q_{n-1} = [\frac{p}{x_n}], ..., q_i = [\frac{p}{x_n ... x_{i+1}}], ...$$

Here the polynomial $[\frac{p}{x_n ... x_{i+1}}] = \frac{\partial^{n-i}}{\partial x_{i+1} ... \partial x_n} p(x_1, ..., x_i, 0, ..., 0)$.

Always:

$$Cap(q_n) \ge Cap(q_{n-1}) \ge .... \ge Cap(q_i) \ge ... \ge [\frac{p}{x_n ... x_{i+1} ... x_1}]$$

Note that $per(A) = [\frac{Prod_A}{x_n ... x_{i+1} ... x_1}]$.

$C-Cap(p)$ is "hard", $Cap(p)$ is "easy" $(p \in Hom_+(n,n))$.

But if $C - Cap(p) > 0$(**H-Stable polynomials**) then $C - Cap(p) = Cap(p)$.

The main result in this direction:

Suppose that $p \in Hom(n,n)$ is **H-Stable**,

i.e. $p(z_1, ..., z_n) \neq 0$ if $RE(z_i) > 0, 1 \leq i \leq n$, then $Cap(q_{i-1}) \geq G(deg_{q_i}(i))Cap(q_i)$, where

$$G(k) = \left(\frac{k-1}{k}\right)^{k-1}$$

Here $deg_p(i)$ is the degree of ith variable in the polynomial p.

It is easy to see that

$$deg_{q_i}(i) \leq \min(i, deg_{q_n}(\{n, ..., i\}) - n + i) \leq \min(i, deg_{q_n}(i)).$$

Moreover in the **H-Stable** case , given an evaluation oracle for $p = q_n$, the degrees $deg_{q_i}(i)$ can be computed in poly-time via submodular minimization.

This theory is powerful, cool, easy to understand and to teach!!! The simplicity(and the power) is achieved by giving up the matrix structure. The same happened in the spectacular proof of Kadison-Singer...

Yet, when applied to the permanent it gives a poly-time deterministic algorithm with the factor $e^{n-k\log(n)}$ for any fixed k (the same factor for the mixed discriminant and the mixed volume).
Essentially, the algorithm computes via the convex minimization $Cap(q_{n-k\log(n)})$.

To break e^n barrier we needed to get back to the matrices/graphs. Not clear at all whether e^n barrier can be broken for, say, the mixed discriminant.

A few Notations

$\Omega_n =$ The set of $n \times n$ Doubly Stochastic matrices, i.e. of $n \times n$ non-negative matrices with row and column sums all equal 1.

$\Lambda(k, n)$ - the set of $n \times n$ matrices with non-negative integer entries and row and column sums equal to k. Note that $k^{-1}\Lambda(k, n) \subset \Omega_n$.

$Bool(k, n) \subset \Lambda(k, n)$ - the set of $n \times n$ boolean matrices with non-negative integer entries and row and column sums all equal to k.

Matrices from $\Lambda(k, n)$ are adjacency matrices of regular bipartite graphs with multiple edges.

Note that $\frac{1}{k}Bool(k, n) \subset \frac{1}{k}\Lambda(k, n) \subset \Omega_n$.

Monomer-Dimer Problem, I will be talking about

Definition 1:

1. $\alpha(m,n,k) =: \min_{A\in\Lambda(k,n)} per_m(A)$.
2. Consider an integer sequence $m(n)$ such that $m(n) \le n$ and $\lim_{n\to\infty} \frac{m(n)}{n} = p \in [0,1]$:
$$\beta(p,k) =: \lim_{n\to\infty} \frac{\log(\alpha(m(n),n,k))}{n}.$$

∎

The problem is to compute (exactly) this thing $\beta(p,k)$. And the same problem for the boolean matrices.

The function $\beta(p,k)$ is concave on $[0,1]$, moreover $\beta(p,k) + p\log(p)$ is concave.

If $k = 1$ things are simple:

$$A = I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

$per_m(I) = \binom{n}{m}$ and if $\lim_{n\to\infty} \frac{m}{n} = p \in [0,1]$ then

$\lim_{n\to\infty} \frac{\log(\binom{n}{m})}{n} = -(p\log(p) + (1-p)\log(1-p)).$

A, sort of, Concentration Conjecture

The idea, which goes back to Wilf[1966](credits Mark(Marek) Kac) and Schrijver-Valiant[1981]: replace the minimum by the average over some natural distribution on $\Lambda(k,n)$. Herbert Wilf's distribution: $P_1 + ... + P_k$, P_i are IID uniformly distributed permutation matrices.

$$AVP(m,n,k) =: E_{\mu(k,n)} per_m(A); A \in \Lambda(k,n).$$

Schrijver-Valiant distribution: fix some equi-partition $\cup_{1 \leq i \leq n} S_i = [1, kn]$, where $|S_i| = k, 1 \leq i \leq n$. Take a random permutation $\sigma \in S_{kn}$ and construct the intersection matrix $\{A(i,j) = |S_i \cup \sigma(S_j)| : 1 \leq i, j \leq n\}$. Define, assuming that $\frac{m(n)}{n} \to p \in [0,1]$,

$$AAVP(p,k) =: \lim_{n \to \infty} \frac{\log(AVP(m(n),n,k))}{n}, p \in [0,1].$$

Can be computed exactly:

$$AAVP(p,k) = p \log \binom{k}{p} - 2(1-p)\log(1-p) + (k-p)\log(1-\frac{p}{k})$$

Conjecture 2: [Friedland$\leq$ 2002-2005, most likely had been stated somewhere a while ago.] The asymptotic exponential growth of the minimum is the same as of the average:

$$\beta(p,k) = AAVP(p,k).$$

▮

Conjecture 3: [Lu-Mohr-Szekely,2011] Let $A \in \Omega_n$. Then the following (positive correlation) ineq. holds?

$$per(A) \geq S(A)$$

$$S(A) =: \prod_{1\leq i\leq n} \sum_{1\leq j\leq n} A(i,j) \prod_{k\neq i} (1 - A(k,j))$$

▮

Theorem 4 : *Let $A \in \Omega_n$. Then the following (new) inequalities hold:*

$$per(A) \geq F(A) =: \prod_{1 \leq i,j \leq n} (1 - A(i,j))^{1-A(i,j)}; \quad (1)$$

(Note that $S(A) \geq F(A), A \in \Omega_n$.)

$$per(A) \leq C^n \prod_{1 \leq i,j \leq n} (1 - A(i,j))^{1-A(i,j)}, \quad (2)$$

where $C = \max_{0 \leq x \leq 1} \frac{x}{f^{-1}(x)} \frac{e^{1-x}}{(1-x)^{1-x}}$,
$f(x) = 1 - (1-x)a^x$, and $1 < a < e$ is the unique solution of the equation $-\log(\frac{a}{e}) = \frac{a}{e}$.

We prove that $C \leq 2$, but numerics give that $C \approx 1.9022$. Note the similarity of the definition of C with the approximation constant in Goemans-Williamson algorithm.

This gives deterministic poly-time algorithm to approximate the permanent of **general** $n \times n$ nonnegative matrices within the relative factor 1.9022^n:

Step 1. Preprocessing, get so called indecomposable $n_i \times n_i$ matrices B_i, $\Sigma_i n_i = n$, such that

$$per(A) = \prod_i per(B_i).$$

Step 2. Scale B_i to doubly-stochastic matrices

$$B_i \approx Diag_{1,i} D_i Diag_{2,i}, D_i \in \Omega_{n_i}.$$

This step is based on STOC-98 paper by Linial, Samorodnitsky and Wigderson.

Step 3. Output

$$Est(per(A)) = \prod_i \det(Diag_{1,i} Diag_{2,i}) \prod_i F(D_i).$$

Previous Conjectures or Questions

1. $\min_{A \in \Omega_n} per(A) = per(\frac{1}{n} J_{n,n}) = \frac{n!}{n^n}$

Van der Waerden (1926): Bang-Friedland(1977-78)-$\geq e^{-n}$, Falikman(1979-1981),

Egorychev(1980-1981).

2. The case $p = 1$:

$$\min_{A \in \Lambda(k,n)} per(A) =?$$

Voorhove, 1980, $k = 3$; $\geq 6 \left(\frac{4}{3}\right)^{n-3}$;

Schrijver,1998: $\geq \left(k \left(\frac{k-1}{k}\right)^{k-1}\right)^n$

L.G,2005: $\geq k^n \frac{k!}{k^k} \left(\frac{k-1}{k}\right)^{(k-1)(n-k)}$ "hyperbolic polynomials approach", actually a vast generalization, goes far beyond permanents.

3. $\min_{A \in Bool(k,n)} per(A) =?$

Exact values are known only for $k = 2$ or

$k = n, n-1, n-2$.

4. Friedland' Conjecture holds for $p = \frac{k}{k+s}, s = 0, 1, ...$:

Friedland-Gurvits [2006-2008], the second application of my "hyperbolic polynomials approach".

A story: in the beginning there was the Scaling.

Let A be $n \times n$ non-negative matrix.

r_i is the sum of ith row, c_j is the sum of jth column.

$R(A) = Diag(\frac{1}{r_1}, ..., \frac{1}{r_n})A; C(A) = ADiag(\frac{1}{c_1}, ..., \frac{1}{c_n})$.

The Scaling(many names for it) Algorithm :

$...RCR(A)$

When does it converge and to what limit?

$KLD(Q,A) = \Sigma_{1 \leq i,j \leq n} Q(i,j) \log\left(\frac{Q(i,j)}{A(i,j)}\right)$

The algorithm converges iff there exists a doubly stochastic Q such that $KLD(Q,A)$ is defined:

$Supp(Q) \subset Supp(A)$. Which is equivalent to the positivity of the permanent: $per(A) > 0$. In this case the limit is a doubly stochastic P:

$$KLD(P,A) = \min_{Q \in \Omega_n} KLD(Q,A).$$

Moreover the minimum is unique, it has the largest support,

and if $Supp(Q) = Supp(A)$ then

$$A = Diag(a_1, ..., a_n) P Diag(b_1, ..., b_n);$$

$$-KLD(P,A) = \sum_{1 \leq i \leq n} \log(a_i) + \log(b_i).$$

Van Der Waerden Conjecture implies that

$\frac{n!}{n^n} = \min_{Q \in \Omega_n} per(Q) \leq \frac{per(A)}{\Pi_{1 \leq i \leq n} a_i b_i} \leq \max_{Q \in \Omega_n} per(Q).$

$$\max_{Q \in \Omega_n} per(Q) = 1$$

Essentially: convex relaxations of $\log(per(A))$**.**

The Sinkhorn's (matrix structure is respected!):

$$per(A) \leq \log(\prod_i a_i \cdot \prod_j b_j) = \max_{B \in \Omega_n} \sum_{1 \leq i,j \leq i,j} b_{i,j} \log(\frac{a_{i,j}}{b_{i,j}}) \tag{3}$$

The "polynomial hierarchy" (matrix structure is lost!):

$$\log(\prod_i a_i \cdot \prod_j b_j) = \log(Cap(Prod_A)), \tag{4}$$

$$Cap(p) = \inf_{x_i > 0} \frac{p(x_1, ..., x_n)}{x_1 ... x_n}.$$

Define a sequence of polynomials

$$g_n = p, q_{n-1} = [\frac{p}{x_n}], ..., q_i = [\frac{p}{x_n ... x_{i+1}}], ...$$

Here the polynomial $[\frac{p}{x_n ... x_{i+1}}] = \frac{\partial^{n-i}}{\partial x_{i+1} ... \partial x_n} p(x_1, ..., x_i, 0, ..., 0)$.

Always:

$$Cap(q_n) \geq Cap(q_{n-1} \geq .... \geq Cap(q_i) \geq ...[\frac{p}{x_n ... x_{i+1} ... x_1}].$$

Note that $per(A) = [\frac{Prod_A}{x_n ... x_{i+1} ... x_1}]$.

The main result in this direction:

Suppose that $p \in Hom(n,n)$ is **H-Stable**,

i.e. $p(z_1, ..., z_n) \neq 0$ if $RE(z_i) > 0, 1 \leq i \leq n$, then

$Cap(q_{i-1}) \geq G(deg_{q_i}(i))Cap(q_i)$, where

$G(k) = \left(\frac{k-1}{k}\right)^{k-1}$.

Here $deg_p(i)$ is the degree of ith variable in the polynomial p.

It is easy to see that

$$deg_{q_i}(i) \leq \min(i, deg_{q_n}(\{n, ..., i\}) - n + i) \leq \min(i, deg_{q_n}(i)).$$

Moreover in the **H-Stable** case , given an evaluation oracle for $p = q_n$, the degrees $deg_{q_i}(i)$ can be computed in poly-time via submodular minimization.

Bethe Approximation.

Define for a pair (A, Q) of non-negative matrices the following functional $CW(Q, A)$ as

$$\sum_{1 \le i,j \le n} \log(1-Q(i,j))(1-Q(i,j)) + \sum_{1 \le i,j \le n} Q(i,j) \log\left(\frac{A(i,j)}{Q(i,j)}\right).$$

Theorem 5:

$$\log(per(A)) \geq \max_{Q \in \Omega_n} CW(Q, A) (\geq CW(P, A) \forall P \in \Omega_n) \tag{5}$$

Corollary 6: *If* $A \in \Omega_n$ *then*

$$per(A) \geq F(A) =: \prod_{1 \le i,j \le n} (1 - A(i,j))^{1-A(i,j)}. \tag{6}$$

Theorem (5) has been conjectured(stated without a proof) by Pascal Vontobel in 2010. Was motivated by numerical evidences (Chertkov, Watanabe, Huang, Jebara, ...) and some simple examples.

Theorem (5) is (essentially) equivalent to the following Schrijver's inequality(1998):

$$per(\widetilde{A}) \geq \prod_{1 \leq i,j \leq n} (1 - A(i,j)); A \in \Omega_n; \qquad (7)$$

where $\widetilde{A}(i,j) =: A(i,j)(1 - A(i,j)), 1 \leq i,j \leq n$.

Very HARD PROOF!

Most of reseachers in the area know the famous corollary of (7):

$$per(A) \geq k^n G(k)^n, \qquad (8)$$

where $A \in \Lambda(k,n)$ and $G(k) = \left(\frac{k-1}{k}\right)^{k-1}$.

My proofs of improved, generalized(non-regural, mixed discriminants, etc.) versions of (8) are just simple, period.

Very fortunate that they have appeared after Schrijver's 1998 paper!

Conjecture 7: If $A \in \Omega_n$ then

$$\frac{F(A)}{per(A)} \leq (\sqrt{2})^n.$$

∎

If true it would provide a deterministic poly-time algorithm to approximate the permanent with the factor $(\sqrt{2})^n$.

There are several versions of this conjecture, it is true, for instance, for regular graphs, i.e. for $A \in \frac{1}{k}\Lambda(k, n)$.

Expressing per_m and the matching polynomial as a Single Permanent

Define the following $2n - m \times 2n - m$ matrix

$$B = \begin{pmatrix} A & J_{n,n-m} \\ (J_{n,n-m})^T & 0 \end{pmatrix}.$$

Then

$$per_m(A) = ((n-m)!)^{-2} per(B).$$

So, there is $FPRAS$ for $per_m(A)$, $A \geq 0$.

Yet, there are situations (planar graphs, for example) when $per(A)$ is easy, but $per_m(A)$ is hard.

The matching polynomial $M_A(x) = (n!)^{-1} per(C)$, where

$$C = (-1)^n \begin{pmatrix} -xA & I \\ J_{n,n} & J_{n,n} \end{pmatrix}.$$

MSS-2013: Matching polynomials and the existence of Ramunujan bipartite graphs of an arbitrary degrees

Another expression for the matching polynomial, A is a boolean matrix:

$M_A(x) = E_{\pm}(Det(xI - A_{\pm}A_{\pm}^T)$, where

$A_{\pm}(i,j) = A(i,j)b(i,j)$, $b(i,j)$ are IID with

$$Prob(b(i,j) = 1) = Prob(b(i,j) = -1) = \frac{1}{2}.$$

Heilmann-Lieb, **Theory of monomer-dimer systems**,1972 :

the roots of the matching polynomial $M_A(x)$ are real;

If $A \in Bool(d,n)$ then the roots of $M_A(x^2)$ are in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$...Alon-Boppana.....

Kadison-Singer Problem: given independent random rank one matrices $X_i X_i^*, X_i \in C^n$ such that

$\Sigma_i\, E\,(X_i X_i^*) = I$ and $\max_i tr\,(E(X_i X_i^*)) = \epsilon.$

Define the mixed characteristic polynomial(a generalization of the bipartite matching polynomial):

$$p(x) = E\left(det(xI - \sum_i X_i X_i^*)\right).$$

What is needed is a generalization of Heilmann-Lieb bound.

MSS-2013: the roots are bounded by $(1 + \sqrt{\epsilon})^2$.

In the Ramunujan graphs context: $X_i = (Ae_i)_{\pm}$–ith column of A with random signs.

A Proof of Friedland's Conjecture

Let $A \in \Lambda(k, n)$. Consider the following $2n - m \times 2n - m$ matrix:

$$K = \begin{pmatrix} aA & bJ_{n,n-m} \\ (bJ_{n,n-m})^T & 0 \end{pmatrix}, \tag{9}$$

where $a = \frac{p}{k}, p = \frac{m}{n}$, $b = \frac{1}{n}$, and $J_{n,n-m}$ is $n \times (n - m)$ matrix of all ones.

It is easy to check that this matrix K is doubly-stochastic.

Importantly, the following identity holds:

$$per_m(A) = \frac{per(K)}{a^m b^{2(n-m)}((n-m)!)^2}. \tag{10}$$

Now, just apply the new lower bound

$$per(K) \geq \prod_{1 \leq i,j \leq 2n-m} (1 - K(i,j))^{1-K(i,j)}, \tag{11}$$

and play a bit with the Stirling Formula. Still can't believe it worked out!

A Disproof of [Lu-Mohr-Szekely] Conjecture

The initial observation: Consider $A \in Bool(k,n)$ and construct the doubly-stochastic matrix K. For this matrix K our (proved!) lower bound

$$F(K) = \prod_{1 \le i,j \le n} (1 - K(i,j))^{1-K(i,j)}$$

and [L-M-S] conjectured lower bound $S(K)$ can be computed explicitely and they are constant for fixed k, n, m. The direct inspection shows that $F(K)$ is asymptotically "close" to the average of $per(K)$ over random $A \in Bool(k,n)$(actually required a bit of extra care: the average was known only for random $A \in \Lambda(k,n)$).

On the other hand $S(K)$ happened to be far above the average for large enough n. This gave the existence of a doubly-stochastic K such that $per(K) < S(K)$ by a standard probabilistic argument. And it works for all densities $\lim_{n \to \infty} \frac{m}{n} = 0 < p < 1$. But not for $p = 1$ or $p = 0$!

The simplification(initially overlooked): the idea works even for $k = 1$. In this case there is no need for the averaging for there is, up to permutations of columns and rows, only one 1-regular matrix. I have choosen $p = .5$, and it resulted in 135×135 counter-example:

$$K = \begin{pmatrix} \frac{1}{2} I_{90} & \frac{1}{90} J_{90,45} \\ (\frac{1}{90} J_{90,45})^T & 0 \end{pmatrix}$$

A Probabilistic Interpretation of the permanent and the product polynomial

Let A be $n \times n$ stochastic matrix, i.e. the rows of A are probabilistic distributions on $\{1, ..., n\}$; $(e_1, ..., e_n)$ is the standard basis in R^n: $e_1 = (1, 0, ..., 0), ...$ Let $\mathbf{V} =: (V_1, ..., V_n)$ be a n-tuple of independent random vectors: $Prob(V_i = e_k) = A(i, k); 1 \le i, k \le n$. The distribution of the sum $V_1 + ... + V_n$ coincides with the vector of the coefficients of the product polynomial

$$Prod_A(x_1, ..., x_n) = \prod_{1 \le i \le n} \sum_{1 \le j \le n} A(i, j) x_j,$$

i.e. the probability $Prob(V_1 + ... + V_n = (\omega_1, ..., \omega_n))$ is the coefficient $a_{\omega_1, \ldots, \omega_n}$ of the monomial $\Pi_{1 \le i \le n} x_i^{\omega_i}$ in the polynomial $Prod_A$. In particular, considering the monomial $x_1 \ldots x_n$

$$per(A) = \frac{\partial^n}{\partial x_1 ... \partial x_n} Prod_A(0) = Prob(V_1 + ... + V_n = e), \tag{12}$$

where $e = (1, 1, ..., 1)$ is the vector of all ones.

Note that the expected value $E(V_1 + ... + V_n) = (c_1, ..., c_n)$, where c_j is the sum of the jth column of A. Thus in the doubly-stochastic case:
the expectation $E(W) = (1, 1, ..., 1), W =: V_1 + ... + V_n$
and the covariance matrix

$$E(WW^T) - E(W)E(W)^T = I - A^T A.$$

$$per(A) = Prob(V_1 + ... + V_n = E(V_1 + ... + V_n)) =$$

$$= Prob(||V_1 + ... + V_n - E(V_1 + ... + V_n)||_2^2 < 2),$$

and the lower bounds on the permanent of doubly-stochastic matrices can be viewed as concentration inequalities for sums of independent random vectors. This interpretation raises a number of natural questions...CLT...

Define the following n events:

$$NE_i = \{(V_1, ..., V_n) : V_i \notin \{V_j, j \neq i\}\}; 1 \leq i \leq n.$$

Equivalent representation of the permanent of doubly-stochastic A:

$$per(A) = Prob(\cap_{1 \leq i \leq n} NE_i). \tag{13}$$

[Lu, Mohr, Szekely; 2011] noticed that $Prob(NE_i) = \Sigma_{1 \leq j \leq n} A(i,j) \Pi_{k \neq i}(1 - A(k,j)$ and conjectured the following beautiful positive correlation inequality for doubly-stochastic matrices $A \in \Omega_n$:

$$per(A) \geq S(A) =: \prod_{1 \leq i \leq n} Prob(NE_i) =$$

$$= \prod_{1 \leq i \leq n} \sum_{1 \leq j \leq n} A(i,j) \prod_{k \neq i} (1 - A(k,j))?$$

Another, Sidak-like, correlation conjecture:

$per(A) = Prob(\cap_{1 \leq j \leq n} \{|(V_1 + ... + V_n)_j - 1| < 1\}) \geq SID(A) =: \Pi_{1 \leq j \leq n} Prob(\{|(V_1 + ... + V_n)_j - 1| < 1\})?$

Why I called the Sidak-like: if $(\xi_1, ..., \xi_n)$ is a Gaussian Vector with zero mean, i.e. its density $p(x_1, ..., x_n) = (2\pi)^{-n/2}\sqrt{\det(Q^{-1})}exp\left(-\frac{1}{2} < Q^{-1}X, X >\right)$ then $Prob(\cap_{1\leq j\leq n}\{|\xi_j| \leq a_i\}) \geq \Pi_{1\leq j\leq n} Prob(\{|\xi_j| \leq a_i\})$.

Why this new lower bound is so amazing?

If $A \in \Omega_n$ then $per(A) \geq F(A) \geq G(n)^n > e^{-n}$, where $G(k) = \left(\frac{k-1}{k}\right)^{k-1}$:
just a bit worse than the Falikman-Egorychev exact lower bound $\frac{n!}{n^n}$.
Assume that A has at most k nonzero entries in each column. Then:
$per(A) \geq F(A) \geq G(k)^n$. Implies the following inequality for matrices $B \in \Lambda(k, n)$:

$$per(B) = k^n per(\frac{1}{k}B) \geq (kG(k))^n.$$

This is the celebrated Schrijver's result [1998].

Let $|col(j)|$ be the number of non-zero entries in jth column.

Theorem 8: *[2005 arxiv; 2006 STOC, 2008 EJC]. If A is a doubly-stochastic $n \times n$ matrix then*

$$Per(A) \geq \prod_{2 \leq j \leq n} G(\min(|col(j)|, j)) \geq \prod_{2 \leq i \leq n} G(j) = \frac{n!}{n^n}.$$

My "hyperbolic polynomials based" lower bound is stronger, much more general, easy to prove, can be even applied to the Mixed Volume.

Schrijver's proof is very, very hard; "frightening technicalities", very narrow specialized to the permanent. Yet, he proved much more and fortunatelly! clearly stated that:

$$per(\widetilde{A}) \geq \prod_{1 \leq i,j \leq n} (1 - A(i,j));$$

where $\widetilde{A}(i,j) =: A(i,j)(1 - A(i,j)), 1 \leq i,j \leq n$ and A is doubly-stochastic.

A way to fix Sidak-like corellation inequality

Let $p \in Hom_+(n,n)$ be **H-Stable** and doubly-stochastic, i.e. $p(z_1,...,z_n) \neq 0$ if $RE(z_i) > 0, 1 \leq i \leq n$ and $\frac{\partial}{\partial x_i} p(1,1,...,1) = 1, 1 \leq i \leq n$.
Define $p_i \in Hom_+(n-1,n-1)$,

$$p_i = \frac{\partial}{\partial x_i} p(x_i = 0; x_j, j \neq i); 1 \leq i \leq n.$$

Also define $Cap(p) = \inf_{x_i > 0} \frac{p(x_1,...,x_n)}{\Pi_{1 \leq i \leq n} x_i}$.

Conjecture 9:

$$\frac{\partial^n}{\partial x_1 ... \partial x_n} p(0) \geq \prod_{1 \leq i \leq n} Cap(p_i)?$$

▮

The (generally wrong) Sidak-like inequality:

$$\frac{\partial^n}{\partial x_1 ... \partial x_n} p(0) \geq \prod_{1 \leq i \leq n} p_i(e_{n-1}),$$

note that

$p_i(e_{n-1}) = p_i(1,1,...,1) \geq Cap(p_i) = \inf_{x_j > 0, j \neq i} \frac{p(x_j > 0, j \neq i)}{\Pi_{j \neq i} x_j}$.

Another conjecture: the same **H-Stable** and doubly-stochastic polynomial $p \in Hom_+(n,n)$.

Define $rt(i,j)$ as the jth root of the unvariate polynomial $p(te - e_i)$, where $e = e_1 + ... + e_n$ is the vector of all ones.

Conjecture 10:

$$\frac{\partial^n}{\partial x_1 ... \partial x_n} p(0) \geq \prod_{1 \leq i,j \leq n} (1 - rt(i,j))^{1-rt(i,j)}$$

∎

Why this conjecture (9) is interesting?

Let $A \in \Omega_n$ and define the following **H-Stable** and doubly-stochastic polynomial:

$Prod_A(x_1, ..., x_n) = \Pi_{1 \le i \le n} \Sigma_{1 \le j \le n} A(i,j) x_j$.

Then the following (elementary log-concavity reasoning) inequality holds:

$$Cap((Prod_A)_j) \ge \prod_{1 \le i \le n} (1 - A(i,j))^{1 - A(i,j)}.$$

So, it will give a hyperbolic proof and generalization of the main new lower bound $per(A) \ge F(A)$;

will give better algorithms for the mixed discriminant and, possibly, for the mixed volume.