Permanent estimators via random matrices

Mark Rudelson
joint work with Ofer Zeitouni

Department of Mathematics
University of Michigan

Permanent of a matrix

Let A be an $n \times n$ matrix with $a_{i,j} \geq 0$.

Permanent of A:

$$\mathrm{perm}(A) = \sum_{\pi \in \Pi_n} \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Permanent of a matrix

Let A be an $n \times n$ matrix with $a_{i,j} \geq 0$.

Permanent of A:

$$\text{perm}(A) = \sum_{\pi \in \Pi_n} \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Determinant of A:

$$\det(A) = \sum_{\pi \in \Pi_n} \text{sign}(\pi) \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Permanent of a matrix

Let A be an $n \times n$ matrix with $a_{i,j} \geq 0$.

Permanent of A:

$$\text{perm}(A) = \sum_{\pi \in \Pi_n} \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Determinant of A:

$$\det(A) = \sum_{\pi \in \Pi_n} \text{sign}(\pi) \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Evaluation of determinants is fast: use e.g., triangularization by Gaussian elimination.

Kalfoten–Villard algorithm:
Running time: $O(n^{2.7})$.

Permanent of a matrix

Let A be an $n \times n$ matrix with $a_{i,j} \geq 0$.

Permanent of A:

$$\text{perm}(A) = \sum_{\pi \in \Pi_n} \prod_{j=1}^{n} a_{j,\pi(j)}.$$

Evaluation of permanents is $\#P$-complete (Valiant 1979)
if there exists a polynomial-time algorithm for permanent evaluation, then any $\#P$ problem can be solved in polynomial time.
Fast computation $\Rightarrow$ P=NP.

Determinant of A:

$$\det(A) = \sum_{\pi \in \Pi_n} \text{sign}(\pi) \prod_{j=1}^{n} a_{j,\pi(j)}.$$

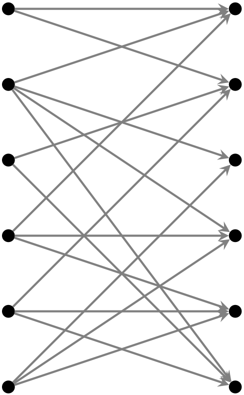
Evaluation of determinants is fast: use e.g., triangularization by Gaussian elimination.

Kalfoten–Villard algorithm:
Running time: $O(n^{2.7})$.

Applications of permanents

Perfect matchings

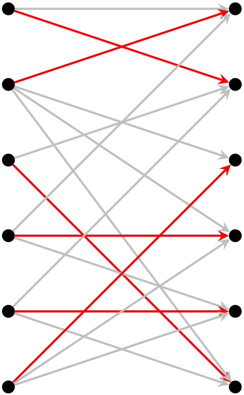
Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.



Applications of permanents

Perfect matchings

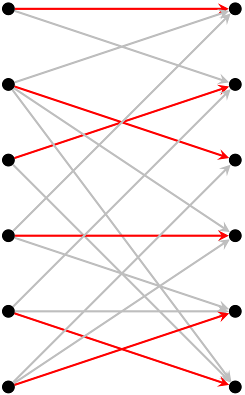
Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A perfect matching is a bijection $\tau : E \to R$ such that $e \to \tau(e)$ for all $e \in E$.



Applications of permanents

Perfect matchings

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A perfect matching is a bijection $\tau : E \to R$ such that $e \to \tau(e)$ for all $e \in E$.



Applications of permanents

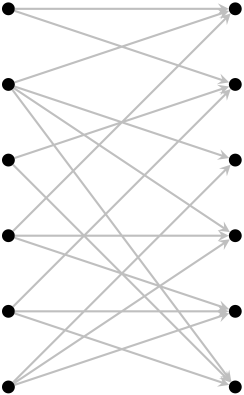
Perfect matchings

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A perfect matching is a bijection $\tau : E \to R$ such that $e \to \tau(e)$ for all $e \in E$.

$$\#(\text{perfect matchings}) = \text{perm}(A),$$

where A is the adjacency matrix of the graph:

$$a_{i,j} = 1 \quad \text{if } i \to j.$$



Deterministic bounds

- Linial–Samorodnitsky–Wigderson algoritm: if $\text{perm}(A) > 0$, then one can find in polynomial time diagonal matrices D, D' such that the renormalized matrix $A' = D'AD$ is almost doubly stochastic:

$$1 - \varepsilon < \sum_{i=1}^{n} a'_{i,j} < 1 + \varepsilon, \qquad \text{for all } j = 1, \dots, n$$

$$1 - \varepsilon < \sum_{j=1}^{n} a'_{i,j} < 1 + \varepsilon, \qquad \text{for all } i = 1, \dots, n$$

Deterministic bounds

- Linial–Samorodnitsky–Wigderson algoritm: if $\text{perm}(A) > 0$, then one can find in polynomial time diagonal matrices D, D' such that the renormalized matrix $A' = D'AD$ is almost doubly stochastic:

$$1 - \varepsilon < \sum_{i=1}^{n} a'_{i,j} < 1 + \varepsilon, \qquad \text{for all } j = 1, \dots, n$$
$$1 - \varepsilon < \sum_{j=1}^{n} a'_{i,j} < 1 + \varepsilon, \qquad \text{for all } i = 1, \dots, n$$

- $\text{perm}(A) = \prod_{i=1}^{n} d_i \cdot \prod_{j=1}^{n} d'_j \cdot \text{perm}(A')$

Deterministic bounds

- Linial–Samorodnitsky–Wigderson algoritm: reduces permanent estimates to almost doubly stochastic matrices
- Van der Waerden conjecture, proved by Falikman and Egorychev: if A is doubly stochastic, then

$$1 \geq \operatorname{perm}(A) \geq \frac{n!}{n^n} \approx e^{-n}$$

- Linial–Samorodnitsky–Wigderson algorithm estimates the permanent with the multiplicative error at most e^n

Deterministic bounds

- Linial–Samorodnitsky–Wigderson algoritm: reduces permanent estimates to almost doubly stochastic matrices
- Van der Waerden conjecture, proved by Falikman and Egorychev: if A is doubly stochastic, then

$$1 \geq \mathrm{perm}(A) \geq \frac{n!}{n^n} \approx e^{-n}$$

- Linial–Samorodnitsky–Wigderson algorithm estimates the permanent with the multiplicative error at most e^n
- Bregman's theorem (1973) implies that if A is doubly stochastic, and $\max a_{i,j} \leq t \cdot \min a_{i,j}$, then

$$\mathrm{perm}(A) \leq e^{-n} \cdot n^{O(t^2)}$$

- Conclusion: if $\max a_{i,j} \leq t \cdot \min a_{i,j}$, then Linial–Samorodnitsky–Wigderson algoritm yields a multiplicative error $n^{O(t^2)}$
- Doesn't cover matrices with zeros.

Probabilistic estimates

- Jerrum–Sinclair–Vigoda algorithm estimates the permanent of any matrix with constant multiplicative error with high probability.

Probabilistic estimates

- Jerrum–Sinclair–Vigoda algorithm estimates the permanent of any matrix with constant multiplicative error with high probability.
- Deficiency: running time is $O(n^{10})$

Probabilistic estimates

- Jerrum–Sinclair–Vigoda algorithm estimates the permanent of any matrix with constant multiplicative error with high probability.
- Deficiency: running time is $O(n^{10})$
- Godsil–Gutman estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let R be an $n \times n$ random matrix with i.i.d. ± 1 entries.
  Form the Hadamard product $R \odot A_{1/2}$: $\quad w_{i,j} = \sqrt{a_{i,j}} \cdot r_{i,j}$.
  Then
  $$\operatorname{perm}(A) = \mathbb{E} \det^2(R \odot A_{1/2}).$$
  Estimator: $\operatorname{perm}(A) \approx \det^2(R \odot A_{1/2})$.

Probabilistic estimates

- Jerrum–Sinclair–Vigoda algorithm estimates the permanent of any matrix with constant multiplicative error with high probability.
- Deficiency: running time is $O(n^{10})$
- Godsil–Gutman estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let R be an $n \times n$ random matrix with i.i.d. ± 1 entries.
  Form the Hadamard product $R \odot A_{1/2}$: $\quad w_{i,j} = \sqrt{a_{i,j}} \cdot r_{i,j}$.
  Then
  $$\operatorname{perm}(A) = \mathbb{E} \det^2(R \odot A_{1/2}).$$
  Estimator: $\operatorname{perm}(A) \approx \det^2(R \odot A_{1/2})$.
- Advantage: Godsil–Gutman estimator is faster than any other algorithm.

Probabilistic estimates

- Jerrum–Sinclair–Vigoda algorithm estimates the permanent of any matrix with constant multiplicative error with high probability.
- Deficiency: running time is $O(n^{10})$
- Godsil–Gutman estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let R be an $n \times n$ random matrix with i.i.d. ± 1 entries.
  Form the Hadamard product $R \odot A_{1/2}$: $\quad w_{i,j} = \sqrt{a_{i,j}} \cdot r_{i,j}$.
  Then
  $$\text{perm}(A) = \mathbb{E} \det{}^2(R \odot A_{1/2}).$$
  Estimator: $\text{perm}(A) \approx \det^2(R \odot A_{1/2})$.
- Advantage: Godsil–Gutman estimator is faster than any other algorithm.
- Deficiency: Godsil–Gutman estimator performs well for "generic" matrices, but fails for large classes of $\{0, 1\}$ matrices, because of arithmetic issues.

Barvinok's estimator

- Godsil–Gutman estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let R be an $n \times n$ random matrix with i.i.d. ± 1 entries.
  Form the Hadamard product $R \odot A_{1/2}$. Then

$$\text{perm}(A) = \mathbb{E} \det^2(R \odot A_{1/2}).$$

  Estimator: $\text{perm}(A) \approx \det^2(R \odot A_{1/2})$.

Barvinok's estimator

- Barvinok's estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let G be an $n \times n$ random matrix with i.i.d. $N(0,1)$ entries.
  Form the Hadamard product $G \odot A_{1/2}$. Then

  $$\text{perm}(A) = \mathbb{E} \det{}^2(G \odot A_{1/2}).$$

  Estimator: $\text{perm}(A) \approx \det{}^2(G \odot A_{1/2})$.
- Barvinok's estimator has no arithmetic issues.

Barvinok's estimator

- Barvinok's estimator Let $A_{1/2}$ be the matrix with entries $a_{i,j}^{1/2}$.
  Let G be an $n \times n$ random matrix with i.i.d. $N(0,1)$ entries.
  Form the Hadamard product $G \odot A_{1/2}$. Then

$$\operatorname{perm}(A) = \mathbb{E} \det^2(G \odot A_{1/2}).$$

  Estimator: $\operatorname{perm}(A) \approx \det^2(G \odot A_{1/2})$.
- Barvinok's estimator has no arithmetic issues.

Theorem (Barvinok)

Let A be any $n \times n$ matrix. Then, with probability $1 - \delta$,

$$((1-\varepsilon) \cdot \theta)^n \operatorname{perm}(A) \le \det^2(G \odot A_{1/2}) \le C \operatorname{perm}(A),$$

where C is an absolute constant and

- $\theta = 0.28$ *for real Gaussian matrices;*
- $\theta = 0.56$ *for complex Gaussian matrices;*
- $\theta = 0.76$ *for quaternionic Gaussian matrices;*

Subexponential bounds for Barvinok's estimator

- Identity matrix: multiplicative error at least $\exp(cn)$ w.h.p.
- Matrix of all ones: multiplicative error at most $\exp(C\sqrt{\log n})$ (Goodman, 1963).
- What happens for other matrices?

Subexponential bounds for Barvinok's estimator

- Identity matrix: multiplicative error at least $\exp(cn)$ w.h.p.
- Matrix of all ones: multiplicative error at most $\exp(C\sqrt{\log n})$ (Goodman, 1963).
- What happens for other matrices?

- Balanced entries (Friedland, Rider, Zeitouni, 2004):
  if $\max a_{i,j} \le t \cdot \min a_{i,j}$, then

$$e^{-o(n)} \le \frac{\det^2(G \odot A_{1/2})}{\mathrm{perm}(A)} \le e^{o(n)}$$

  with probability $1 - o(1)$ as $n \to \infty$.

- The bound is asymptotic.

Balanced entries

Theorem (Costello, Vu, 2009)

If $\max a_{i,j} \leq t \cdot \min a_{i,j}$, *then*

$$\exp\left(-O(n^{2/3}\log n)\right) \leq \frac{\det^2(G \odot A_{1/2})}{\operatorname{perm}(A)} \leq \exp\left(O(n^{2/3}\log n)\right)$$

with probability $1 - n^{-C}$.

Balanced entries

Theorem (Costello, Vu, 2009)

If $\max a_{i,j} \leq t \cdot \min a_{i,j}$, *then*

$$\exp\left(-O(n^{2/3}\log n)\right) \leq \frac{\det^2(G \odot A_{1/2})}{\mathrm{perm}(A)} \leq \exp\left(O(n^{2/3}\log n)\right)$$

with probability $1 - n^{-C}$.

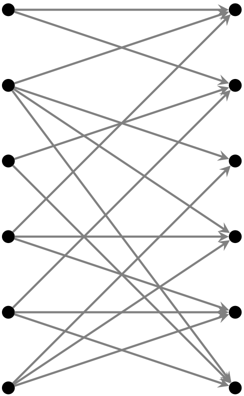
- Not applicable for matrices with zeros.
- Linial–Samorodnitsky–Wigderson algorithm estimates the permanent with polynomial error for balanced matrices.

Question:

for which graphs would Barvinok's estimator yield a small error?

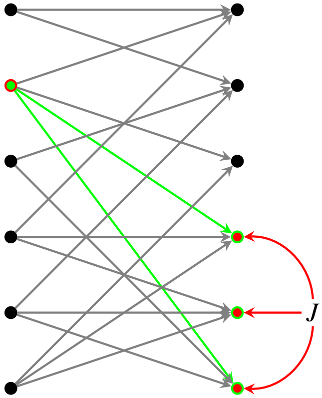
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.



Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected
to a set $J \subset R$ if it is connected
to at least $(\delta/2) \cdot |J|$ vertices of J.



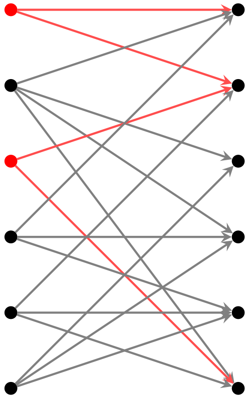
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected
to a set $J \subset R$ if it is connected
to at least $(\delta/2) \cdot |J|$ vertices of J.

Broadly connected graph

Let $\delta, \kappa > 0$, $\delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;



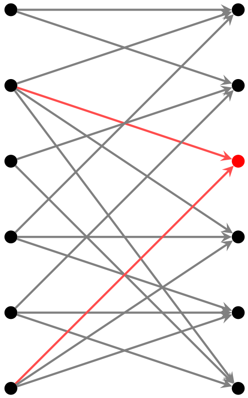
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected
to a set $J \subset R$ if it is connected
to at least $(\delta/2) \cdot |J|$ vertices of J.

Broadly connected graph

Let $\delta, \kappa > 0$, $\delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;
2. Right degree condition: $\deg(j) \geq \delta n$ for all $j \in [n]$;



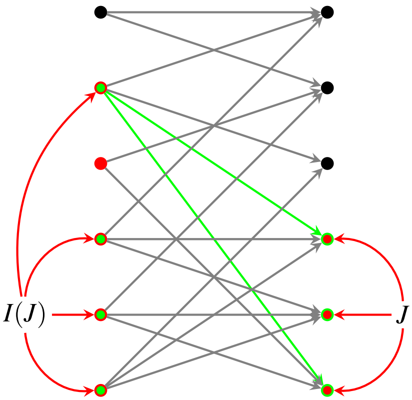
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected to a set $J \subset R$ if it is connected to at least $(\delta/2) \cdot |J|$ vertices of J.

Broadly connected graph

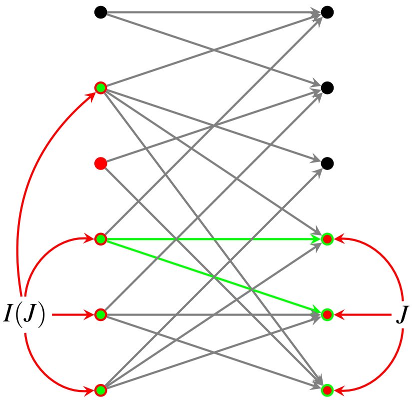
Let $\delta, \kappa > 0$, $\delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;
2. Right degree condition: $\deg(j) \geq \delta n$ for all $j \in [n]$;
3. Strong expansion condition: for any set $J \subset [m]$ the set of its δ-broadly connected neighbors has the cardinality $|I(J)| \geq \min\left((1+\kappa)|J|, n\right)$.



Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected to a set $J \subset R$ if it is connected to at least $(\delta/2) \cdot |J|$ vertices of J.



Broadly connected graph

Let $\delta, \kappa > 0$, $\delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;
2. Right degree condition: $\deg(j) \geq \delta n$ for all $j \in [n]$;
3. Strong expansion condition: for any set $J \subset [m]$ the set of its δ-broadly connected neighbors has the cardinality $|I(J)| \geq \min\big((1+\kappa)|J|, n\big)$.

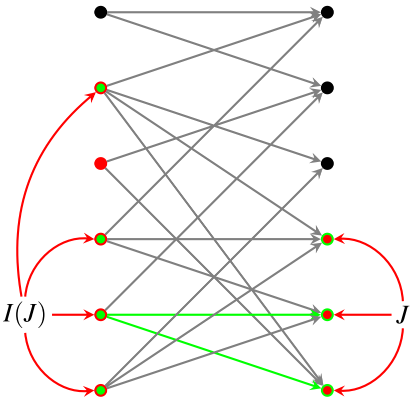
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected to a set $J \subset R$ if it is connected to at least $(\delta/2) \cdot |J|$ vertices of J.

Broadly connected graph

Let $\delta, \kappa > 0$, $\delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;
2. Right degree condition: $\deg(j) \geq \delta n$ for all $j \in [n]$;
3. Strong expansion condition: for any set $J \subset [m]$ the set of its δ-broadly connected neighbors has the cardinality $|I(J)| \geq \min\big((1+\kappa)|J|, n\big)$.



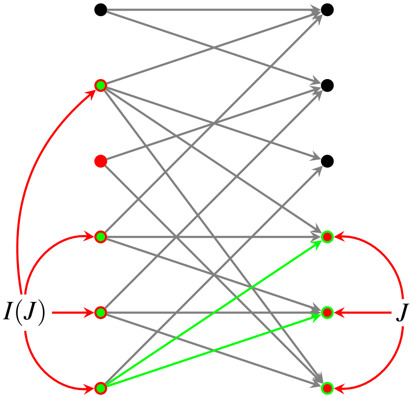
Broadly connected bipartite graphs

Let $\Gamma = (L, R, V)$ be an $n \times n$ bipartite graph.
A vertex $i \in L$ is δ-broadly connected to a set $J \subset R$ if it is connected to at least $(\delta/2) \cdot |J|$ vertices of J.

Broadly connected graph

Let $\delta, \kappa > 0,\ \delta/2 > \kappa$. The graph Γ is (δ, κ)-broadly connected if

1. Left degree condition: $\deg(i) \geq \delta n$ for all $i \in [n]$;
2. Right degree condition: $\deg(j) \geq \delta n$ for all $j \in [n]$;
3. Strong expansion condition: for any set $J \subset [m]$ the set of its δ-broadly connected neighbors has the cardinality $|I(J)| \geq \min\left((1+\kappa)|J|, n\right)$.



Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated:
  the multiplicative error is $O(\exp\left((cn \log n)^{1/3}\right)$ with high probability

Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated: the multiplicative error is $O(\exp\left((cn\log n)^{1/3}\right)$ with high probability
- Bad news: It can be concentrated around a wrong value

Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated: the multiplicative error is $O(\exp\left((cn \log n)^{1/3}\right)$ with high probability
- Bad news: It can be concentrated around a wrong value

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph. Let Γ be a random Gaussian matrix with variance matrix A. Then for any $\tau \geq 1$

$$\mathbb{P}\left[\qquad \leq \frac{\det^2(\Gamma)}{M} \leq \qquad \right]$$

$$\geq 1 - \textit{small}$$

Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated: the multiplicative error is $O(\exp\left((cn\log n)^{1/3}\right)$ with high probability
- Bad news: It can be concentrated around a wrong value

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph.
Let Γ be a random Gaussian matrix with variance matrix A.
Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n\log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n\log n)^{1/3}\right)\right]$$
$$\geq 1 - \textit{small}$$

Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated: the multiplicative error is $O(\exp\left((cn\log n)^{1/3}\right)$ with high probability
- Bad news: It can be concentrated around a wrong value

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph. Let Γ be a random Gaussian matrix with variance matrix A. Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n\log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n\log n)^{1/3}\right)\right]$$
$$\geq 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log n\right)$$

Concentration for Barvinok's estimator: how strong?

- Good news: the Barvinok estimator is strongly concentrated: the multiplicative error is $O(\exp\left((cn\log n)^{1/3}\right)$ with high probability
- Bad news: It can be concentrated around a wrong value

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph.
Let Γ be a random Gaussian matrix with variance matrix A.
Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n\log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n\log n)^{1/3}\right)\right]$$
$$\geq 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log n\right)$$

and

$$\exp\left(-C(n\log n)^{1/2}\right) \leq \frac{M}{\operatorname{perm}(A)} \leq 1.$$

Concentration for Barvinok's estimator: where?

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph.
Let Γ be a random Gaussian matrix with variance matrix A.
Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n \log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n \log n)^{1/3}\right)\right]$$
$$\geq 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log n\right)$$

and

$$\exp\left(-C(n \log n)^{1/2}\right) \leq \frac{M}{\text{perm}(A)} \leq 1.$$

- $\text{perm}(A) = \mathbb{E}\det^2(\Gamma)$;

Concentration for Barvinok's estimator: where?

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph.
Let Γ be a random Gaussian matrix with variance matrix A.
Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n \log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n \log n)^{1/3}\right)\right]$$
$$\geq 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log n\right)$$

and

$$\exp\left(-C(n \log n)^{1/2}\right) \leq \frac{M}{\operatorname{perm}(A)} \leq 1.$$

- $\operatorname{perm}(A) = \mathbb{E}\det^2(\Gamma)$;
- $M = \exp\left(\mathbb{E}\log\det^2(\Gamma)\right)$.

Concentration for Barvinok's estimator: where?

Theorem (R'–Zeitouni, 2013)

Let A be the adjacency matrix A of an $n \times n$ broadly connected bipartite graph.
Let Γ be a random Gaussian matrix with variance matrix A.
Then for any $\tau \geq 1$

$$\mathbb{P}\left[\exp\left(-C(\tau n \log n)^{1/3}\right) \leq \frac{\det^2(\Gamma)}{M} \leq \exp\left(C(\tau n \log n)^{1/3}\right)\right]$$
$$\geq 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log n\right)$$

and

$$\exp\left(-C(n \log n)^{1/2}\right) \leq \frac{M}{\operatorname{perm}(A)} \leq 1.$$

- $\operatorname{perm}(A) = \mathbb{E}\det^2(\Gamma)$;
- $M = \exp\left(\mathbb{E}\log\det^2(\Gamma)\right)$.
- det is a highly non-linear function $\Rightarrow$ $\det(\Gamma)$ has heavy tails.

Results for matrices

Large entries graph

Let $s > 0$ and let B be an $n \times n$ matrix B with non-negative entries.
Define the bipartite graph $\Gamma_B(s)$ connecting the vertices i and j whenever $b_{i,j} \geq s$

Results for matrices

Large entries graph

Let $s > 0$ and let B be an $n \times n$ matrix B with non-negative entries.
Define the bipartite graph $\Gamma_B(s)$ connecting the vertices i and j whenever $b_{i,j} \geq s$

$$B = \begin{pmatrix} 0.7 & 0 & 0.1 & 0.5 \\ 0.1 & 0.6 & 0.8 & 0.2 \\ 0.6 & 0.6 & 0.3 & 0.5 \\ 0.2 & 0.8 & 0.7 & 0.3 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad (s = 0.5)$$

Consider matrices with broadly connected large entries graphs.

Results for matrices

Theorem

Let B be an $n \times n$ matrix such that

$$\sum_{i=1}^{n} b_{i,j} \le 1 \quad \text{for all } j \in [n]; \qquad \text{and} \qquad \sum_{j=1}^{n} b_{i,j} \le 1 \quad \text{for all } i \in [n],$$

and $0 \le b_{i,j} \le b_n/n$, *where* $0 < b_n \le n$.
Assume that the large entries graph $\Gamma_B(1/n)$ *is broadly connected.*
Let Γ *be a random Gaussian matrix with variance matrix* A.

Results for matrices

Theorem

Let B be an $n \times n$ matrix such that

$$\sum_{i=1}^{n} b_{i,j} \le 1 \quad \text{for all } j \in [n]; \qquad \text{and} \qquad \sum_{j=1}^{n} b_{i,j} \le 1 \quad \text{for all } i \in [n],$$

and $0 \le b_{i,j} \le b_n/n$, where $0 < b_n \le n$.
Assume that the large entries graph $\Gamma_B(1/n)$ is broadly connected.
Let Γ be a random Gaussian matrix with variance matrix A. Then for any $\tau \ge 1$

$$\mathbb{P}\left[\exp\left(-C(\tau b_n n)^{1/3}\log^c n\right) \le \frac{\det^2(\Gamma)}{M} \le \exp\left(C(\tau b_n n)^{1/3}\log^c n\right)\right]$$
$$\ge 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log^c n\right)$$

Results for matrices

Theorem

Let B be an $n \times n$ matrix such that

$$\sum_{i=1}^{n} b_{i,j} \le 1 \quad \textit{for all } j \in [n]; \qquad \textit{and} \qquad \sum_{j=1}^{n} b_{i,j} \le 1 \quad \textit{for all } i \in [n],$$

and $0 \le b_{i,j} \le b_n/n$, where $0 < b_n \le n$.
Assume that the large entries graph $\Gamma_B(1/n)$ is broadly connected.
Let Γ be a random Gaussian matrix with variance matrix A. Then for any $\tau \ge 1$

$$\mathbb{P}\left[\exp\left(-C(\tau b_n n)^{1/3} \log^c n\right) \le \frac{\det^2(\Gamma)}{M} \le \exp\left(C(\tau b_n n)^{1/3} \log^c n\right)\right]$$
$$\ge 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log^c n\right)$$

$$\textit{and} \qquad \exp\left(-C(b_n n)^{1/2} \log^c n\right) \le \frac{M}{\operatorname{perm}(A)} \le 1.$$

Results for matrices

Theorem

$$\mathbb{P}\left[\exp\left(-C(\tau b_n n)^{1/3}\log^c n\right) \le \frac{\det^2(\Gamma)}{M} \le \exp\left(C(\tau b_n n)^{1/3}\log^c n\right)\right]$$

$$\ge 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log^c n\right)$$

and $$\exp\left(-C(b_n n)^{1/2}\log^c n\right) \le \frac{M}{\operatorname{perm}(A)} \le 1.$$

- Small maximal entry: $\max b_{i,j} = o(1)$ or $b_n = o(n)$:
  - Barvinok's estimator is well-concentrated about the permanent.

Results for matrices

Theorem

$$\mathbb{P}\left[\exp\left(-C(\tau b_n n)^{1/3}\log^c n\right) \le \frac{\det^2(\Gamma)}{M} \le \exp\left(C(\tau b_n n)^{1/3}\log^c n\right)\right]$$
$$\ge 1 - \exp(-\tau) + \exp\left(-c\sqrt{n}/\log^c n\right)$$

and $$\exp\left(-C(b_n n)^{1/2}\log^c n\right) \le \frac{M}{\text{perm}(A)} \le 1.$$

- Small maximal entry: $\max b_{i,j} = o(1)$ or $b_n = o(n)$:
  - Barvinok's estimator is well-concentrated about the permanent.
- Large maximal entry: $\max b_{i,j} = \Omega(1)$ or $b_n = n \cdot \max b_{i,j} = \Omega(n)$:
  - Barvinok's estimator is well-concentrated: $(\tau b_n n)^{1/3} = O(n^{2/3})$;
  - It may be concentrated exponentially far from the permanent: $\sqrt{b_n n} = \Omega(n)$.
  - Consistent failure is possible.

Example of a consistent failure

Let B be the $n \times n$ matrix with entries

$$b_{i,j} = \begin{cases} \theta & \text{if } i = j \\ \frac{1-\theta}{n-1} & \text{if } i \neq j \end{cases}.$$

- The matrix B is doubly stochastic for $\theta \in (0, 1)$.
- B has no zero entries.
- Γ_B is a complete bipartite graph.

Example of a consistent failure

Let B be the $n \times n$ matrix with entries

$$b_{i,j} = \begin{cases} \theta & \text{if } i = j \\ \frac{1-\theta}{n-1} & \text{if } i \neq j \end{cases}.$$

- The matrix B is doubly stochastic for $\theta \in (0,1)$.
- B has no zero entries.
- Γ_B is a complete bipartite graph.

Theorem

There exists $\theta_0 < 1$ such that for any $\theta \in (\theta_0, 1)$

$$\det^2(B_{1/2} \odot G) < e^{-cn} \operatorname{perm}(B)$$

with high probability.

Approach to concentration

- Aim: $X(G) := \det^2(A_{1/2} \odot G)$ is concentrated.
- $\det^2(A_{1/2} \odot G)$ is highly non-linear $\quad \Rightarrow \quad \log(\det^2(A_{1/2} \odot G))$ is easier to control.
- Modified aim : $Y(G) = \log\det^2(A_{1/2} \odot G)$ is concentrated around its expectation.
  We will have to compare the concentration for $X(G)$ and $Y(G)$ at the end.

Approach to concentration

- Aim: $X(G) := \det^2(A_{1/2} \odot G)$ is concentrated.
- $\det^2(A_{1/2} \odot G)$ is highly non-linear $\quad \Rightarrow \quad \log(\det^2(A_{1/2} \odot G))$ is easier to control.
- Modified aim : $Y(G) = \log \det^2(A_{1/2} \odot G)$ is concentrated around its expectation.
  We will have to compare the concentration for $X(G)$ and $Y(G)$ at the end.
- There exists a subgaussian concentration inequality for Lipschitz functions on $\mathbb{R}^{n\times n}$ with respect to the gaussian measure:

$$\mathbb{P}\left(|F(G) - \mathbb{E}F(G)| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2L^2(F)}\right).$$

Approach to concentration

- Aim: $X(G) := \det^2(A_{1/2} \odot G)$ is concentrated.
- $\det^2(A_{1/2} \odot G)$ is highly non-linear $\Rightarrow$ $\log(\det^2(A_{1/2} \odot G))$ is easier to control.
- Modified aim : $Y(G) = \log \det^2(A_{1/2} \odot G)$ is concentrated around its expectation.
  We will have to compare the concentration for $X(G)$ and $Y(G)$ at the end.
- There exists a subgaussian concentration inequality for Lipschitz functions on $\mathbb{R}^{n \times n}$ with respect to the gaussian measure:

$$\mathbb{P}\left(|F(G) - \mathbb{E}F(G)| \geq t\right) \leq 2 \exp\left(-\frac{t^2}{2L^2(F)}\right).$$

- $\log \det^2(A_{1/2} \odot G)$ is not Lipschitz.

What is non-Lipschitz in log determinant?

- $\log \det^2(A_{1/2} \odot G) = 2 \sum_{j=1}^n \log s_j(A_{1/2} \odot G)$.
- The maping $G \to A_{1/2} \odot G$ is Lipschitz.
- The mapping $M \to \big(s_1(M), \dots, s_n(M)\big)$ is Lipschitz.
- Logarithm is not a Lipschitz function.

What is non-Lipschitz in log determinant?

- $\log \det^2(A_{1/2} \odot G) = 2\sum_{j=1}^{n} \log s_j(A_{1/2} \odot G)$.
- The maping $G \to A_{1/2} \odot G$ is Lipschitz.
- The mapping $M \to \big(s_1(M), \ldots, s_n(M)\big)$ is Lipschitz.
- Truncated logarithm $\log_\varepsilon x = \max(\log x, \log \varepsilon)$ is a Lipschitz function.
- We have to guarantee that $s_n(A_{1/2} \odot G) > \varepsilon$ with high probability.

Estimating the log determinant

1. Use non-asymptotic random matrix theory to show that $\mathbb{P}\left(s_n(A_{1/2} \odot G) > \varepsilon\right) = 1 - o(1)$;
2. Replace

$$\log \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log s_j(A_{1/2} \odot G)$$

by $$\log_{\varepsilon} \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log_{\varepsilon} s_j(A_{1/2} \odot G)$$

(holds with high probability);
3. Establish concentration for $\log_{\varepsilon} \det^2(A_{1/2} \odot G)$.

Estimating the log determinant: first attempt

1. Use non-asymptotic random matrix theory to show that $\mathbb{P}\left(s_n(A_{1/2} \odot G) > \varepsilon\right) = 1 - o(1)$;
2. Replace

$$\log \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log s_j(A_{1/2} \odot G)$$

by

$$\log_{\varepsilon} \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log_{\varepsilon} s_j(A_{1/2} \odot G)$$

(holds with high probability);

3. Establish concentration for $\log_{\varepsilon} \det^2(A_{1/2} \odot G)$. $\leftarrow$ Fails

Estimating the log determinant: first attempt

1. Use non-asymptotic random matrix theory to show that $\mathbb{P}\left(s_n(A_{1/2} \odot G) > \varepsilon\right) = 1 - o(1)$;
2. Replace

$$\log \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log s_j(A_{1/2} \odot G)$$

by
$$\log_\varepsilon \det{}^2(A_{1/2} \odot G) = 2 \sum_{j=1}^{n} \log_\varepsilon s_j(A_{1/2} \odot G)$$

(holds with high probability);
3. Establish concentration for $\log_\varepsilon \det^2(A_{1/2} \odot G)$. $\leftarrow$ Fails

Reasons for the failure

- The least singular value is too small $\Rightarrow$ the Lipschitz constant is too big.
- The least singular value is not concentrated at all.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;
   - Stronger concentrated.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;
   - Stronger concentrated.
3. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;
   - Stronger concentrated.
3. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
4. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_\varepsilon s_j(A_{1/2} \odot G)$ is a $(1/\varepsilon)$-Lipschitz function
   $\Rightarrow$ we can use concentration for the first singular values

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;
   - Stronger concentrated.
3. $\log \det^2 (A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
4. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_\varepsilon s_j(A_{1/2} \odot G)$ is a $(1/\varepsilon)$-Lipschitz function
   $\Rightarrow$ we can use concentration for the first singular values
5. $\left| \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G) \right| \leq (n-k) |\log \alpha_n| \quad \Rightarrow \quad$ contribution of the last singular values is limited.

Estimating the log determinant: second attempt

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. Intermediate singular values:
   - Larger than the last one $\Rightarrow$ stronger concentration for truncated logarithm;
   - Stronger concentrated.
3. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
4. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_{\varepsilon} s_j(A_{1/2} \odot G)$ is a $(1/\varepsilon)$-Lipschitz function $\Rightarrow$ we can use concentration for the first singular values
5. $\left|\sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)\right| \leq (n-k)|\log \alpha_n| \quad \Rightarrow \quad$ contribution of the last singular values is limited.

A single threshold is not enough to get a fine picture.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
3. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_\varepsilon s_j(A_{1/2} \odot G)$ is a $(1/\varepsilon)$-Lipschitz function
   $\Rightarrow$ we can use concentration for the first singular values
4. $\left| \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G) \right| \leq (n-k) |\log \alpha_n|$
   $\Rightarrow$ contribution of the last singular values is limited.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
3. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_{\varepsilon_j} s_j(A_{1/2} \odot G)$ is a (?)-Lipschitz function $\Rightarrow$ balance the concentration
4. $\left| \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G) \right| \leq (n-k) |\log \alpha_n|$
   $\Rightarrow$ contribution of the last singular values is limited.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
3. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_{\varepsilon_j} s_j(A_{1/2} \odot G)$ is a $(?)$-Lipschitz function $\Rightarrow$ balance the concentration
4. $\left| \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G) \right| \leq \sum_{j=n-k+1}^{n} |\log \varepsilon_j|$
   $\Rightarrow$ contribution of the last singular values is limited.

Estimating the log determinant

1. Non-asymptotic random matrix theory $\Rightarrow$
   Bound for intermediate singular values: $s_j(A_{1/2} \odot G) \geq \varepsilon_j$ with high probability;
2. $\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G)$
3. $\sum_{j=1}^{n-k} \log s_j(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_{\varepsilon_j} s_j(A_{1/2} \odot G)$ is a (?)-Lipschitz function $\Rightarrow$
   balance the concentration
4. $\left| \sum_{j=n-k+1}^{n} \log s_j(A_{1/2} \odot G) \right| \leq \sum_{j=n-k+1}^{n} |\log \varepsilon_j|$
   $\Rightarrow$ contribution of the last singular values is limited.
5. How to choose the threshold k?

- Smaller k $\Rightarrow$ smaller error.
- Larger k $\Rightarrow$ stronger concentration.

Choosing the right threshold

$$\log \det^2(A_{1/2} \odot G) = \sum_{j=1}^{n-k} \log_{\varepsilon_j} s_j(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \text{error terms}$$

Choosing the right threshold

$$\log \det^2(A_{1/2} \odot G) = \widetilde{\log} \det^2(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \text{error terms}$$

$\widetilde{\log} \det^2(A_{1/2} \odot G)$ is concentrated about its expectation.

- Smaller k $\Rightarrow$ smaller error.
- Larger k $\Rightarrow$ stronger concentration.

Choosing the right threshold

$$\log\det^2(A_{1/2}\odot G) = \widetilde{\log}\det^2(A_{1/2}\odot G) + \sum_{j=n-k+1}^{n} \text{error terms}$$

$\widetilde{\log}\det^2(A_{1/2}\odot G)$ is concentrated about its expectation.

- Smaller k $\Rightarrow$ smaller error.
  $\log\det^2(A_{1/2}\odot G)$ is close to $\mathbb{E}\log\det^2(A_{1/2}\odot G)$ with high probability.
  This may be far from $\log\operatorname{perm}(A)$.
- Larger k $\Rightarrow$ stronger concentration.

Choosing the right threshold

$$\log \det^2(A_{1/2} \odot G) = \widetilde{\log} \det^2(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \text{error terms}$$

$\widetilde{\log} \det^2(A_{1/2} \odot G)$ is concentrated about its expectation.

- Smaller $k \Rightarrow$ smaller error.
  $\log \det^2(A_{1/2} \odot G)$ is close to $\mathbb{E} \log \det^2(A_{1/2} \odot G)$ with high probability.
  This may be far from $\log \operatorname{perm}(A)$.
- Larger $k \Rightarrow$ stronger concentration.
  Strong concentration $\Rightarrow$

$$\mathbb{E}\, \widetilde{\log} \det^2(A_{1/2} \odot G) \approx \widetilde{\log}\, \mathbb{E} \det^2(A_{1/2} \odot G)$$

Choosing the right threshold

$$\log \det^2(A_{1/2} \odot G) = \widetilde{\log} \det^2(A_{1/2} \odot G) + \sum_{j=n-k+1}^{n} \text{error terms}$$

$\widetilde{\log} \det^2(A_{1/2} \odot G)$ is concentrated about its expectation.

- Smaller k $\Rightarrow$ smaller error.
  $\log \det^2(A_{1/2} \odot G)$ is close to $\mathbb{E} \log \det^2(A_{1/2} \odot G)$ with high probability.
  This may be far from $\log \operatorname{perm}(A)$.
- Larger k $\Rightarrow$ stronger concentration.
  Strong concentration $\Rightarrow$

$$\mathbb{E} \log \det^2(A_{1/2} \odot G) \approx \log \mathbb{E} \det^2(A_{1/2} \odot G) = \log \operatorname{perm}(A)$$

up to the error terms.

Conclusion

- Barvinok's estimator is always strongly concentrated for broadly connected graphs and matrices.

Conclusion

- Barvinok's estimator is always strongly concentrated for broadly connected graphs and matrices.
- It is concentrated around $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ $\Rightarrow$ the main sourse of errors is deterministic.

Conclusion

- Barvinok's estimator is always strongly concentrated for broadly connected graphs and matrices.
- It is concentrated around $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ $\Rightarrow$ the main sourse of errors is deterministic.
- The gap between $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ and the permanent is subexponential for graphs

Conclusion

- Barvinok's estimator is always strongly concentrated for broadly connected graphs and matrices.
- It is concentrated around $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ $\Rightarrow$ the main sourse of errors is deterministic.
- The gap between $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ and the permanent is subexponential for graphs
- and for almost doubly stochastic matrices in the subcritical case $\max a_{i,j} = o(1)$.

Conclusion

- Barvinok's estimator is always strongly concentrated for broadly connected graphs and matrices.
- It is concentrated around $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ $\Rightarrow$ the main sourse of errors is deterministic.
- The gap between $\exp \mathbb{E} \log \det^2(A_{1/2} \odot G)$ and the permanent is subexponential for graphs
- and for almost doubly stochastic matrices in the subcritical case $\max a_{i,j} = o(1)$.
- It may be exponential in the critical case $\max a_{i,j} = \Omega(1) \Rightarrow$ consistent failure.