

Approximating Permanent of Random Matrices with Vanishing Mean: Made Better and Simpler

Zhengfeng Ji¹ **Zhihan Jin**² Pinyan Lu³

¹University of Technology Sydney

²ETH

³Shanghai University of Finance and Economics

January 11, 2021

Definition of permanent

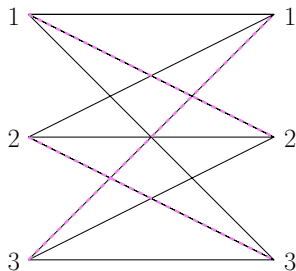
- Given a matrix $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$, the permanent is given by

$$\text{Per}(A) \triangleq \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Definition of permanent

- Given a matrix $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$, the permanent is given by

$$\text{Per}(A) \triangleq \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$



$$\begin{aligned} \text{Per}(A) &= a_{1,1} a_{2,2} a_{3,3} + a_{1,1} a_{2,3} a_{3,2} \\ &\quad + a_{1,2} a_{2,1} a_{3,3} + \mathbf{a_{1,2} a_{2,3} a_{3,1}} \\ &\quad + a_{1,3} a_{2,1} a_{3,2} + a_{1,3} a_{2,2} a_{3,1}. \end{aligned}$$

- Computing permanent exactly is $\#P$ -hard. (Valiant 79')

- Computing permanent exactly is $\#P$ -hard. (Valiant 79')
- There is a FPRAS to approximate (ϵ -approximation) permanent with non-negative entries. (Jerrum, Sinclair and Vigoda 01')

Worst-case results

- Computing permanent exactly is $\#P$ -hard. (Valiant 79')
- There is a FPRAS to approximate (ϵ -approximation) permanent with non-negative entries. (Jerrum, Sinclair and Vigoda 01')
- Deciding the sign of the permanent is also $\#P$ -hard. (Aaronson 11')

Worst-case results

- Computing permanent exactly is $\#P$ -hard. (Valiant 79')
- There is a FPRAS to approximate (ϵ -approximation) permanent with non-negative entries. (Jerrum, Sinclair and Vigoda 01')
- Deciding the sign of the permanent is also $\#P$ -hard. (Aaronson 11')
- Seemingly impossible to derive a multiplicative approximation.

- Exact computing the permanent is $\#P$ -hard even for a constant portion of matrices with iid standard complex Gaussian entries. (Aaronson and Arkhipov 10')

Average-case computational complexity

- Exact computing the permanent is $\#P$ -hard even for a constant portion of matrices with iid standard complex Gaussian entries. (Aaronson and Arkhipov 10')
- What about average-case approximations?

Average-case computational complexity

- Exact computing the permanent is $\#P$ -hard even for a constant portion of matrices with iid standard complex Gaussian entries. (Aaronson and Arkhipov 10')
- What about average-case approximations?
- Permanent-of-Gaussians Conjecture: approximating the permanent of Gaussian matrices is $\#P$ -hard. (Aaronson and Arkhipov 10')

Average-case computational complexity

- Exact computing the permanent is $\#P$ -hard even for a constant portion of matrices with iid standard complex Gaussian entries. (Aaronson and Arkhipov 10')
- What about average-case approximations?
- Permanent-of-Gaussians Conjecture: approximating the permanent of Gaussian matrices is $\#P$ -hard. (Aaronson and Arkhipov 10')
- Easy for quantum computers but maybe hard for classical ones.

- Relaxation: entries are iid sampled from a complex Gaussian with mean μ and unit variance.

Average-case approximation

- Relaxation: entries are iid sampled from a complex Gaussian with mean μ and unit variance.
- $\mu \rightarrow \infty$ is easy while $\mu = 0$ meets the Permanent-of-Gaussians Conjecture.

Average-case approximation

- Relaxation: entries are iid sampled from a complex Gaussian with mean μ and unit variance.
- $\mu \rightarrow \infty$ is easy while $\mu = 0$ meets the Permanent-of-Gaussians Conjecture.
- To (dis-)prove the Permanent-of-Gaussians Conjecture, it is equivalent to consider $\mu = n^{-c}$ for some small constant $c > 0$. (Eldar and Mehraban 18')

- There is a quasi-polynomial scheme to approximate random permanents when $\mu = 1/\text{poly log log } n$. (Eldar and Mehraban 18')

Average-case approximation

- There is a quasi-polynomial scheme to approximate random permanents when $\mu = 1/\text{poly log log } n$. (Eldar and Mehraban 18')
- Can we improve the guarantee for μ ?

Average-case approximation

- There is a quasi-polynomial scheme to approximate random permanents when $\mu = 1/\text{poly log log } n$. (Eldar and Mehraban 18')
- Can we improve the guarantee for μ ?
- **Our result:** there is a quasi-polynomial scheme to approximate random permanents when $\mu = 1/\text{poly log } n$.

- mean μ and variance 1 \leftrightarrow mean 1 and variance μ^{-2} .

- mean μ and variance 1 \leftrightarrow mean 1 and variance μ^{-2} .
- WLOG, consider matrix $M = J + zA$ where J the all-ones matrix and each entry of A is iid sampled from \mathcal{D} :

$$\mathbb{E}_{x \sim \mathcal{D}}[x] = 0, \quad \text{Var}_{x \sim \mathcal{D}}[x] = 1, \quad \mathbb{E}_{x \sim \mathcal{D}} |x - \mu|^3 = \rho < \infty.$$

- mean μ and variance 1 \leftrightarrow mean 1 and variance μ^{-2} .
- WLOG, consider matrix $M = J + zA$ where J the all-ones matrix and each entry of A is iid sampled from \mathcal{D} :

$$\mathbb{E}_{x \sim \mathcal{D}}[x] = 0, \quad \text{Var}_{x \sim \mathcal{D}}[x] = 1, \quad \mathbb{E}_{x \sim \mathcal{D}} |x - \mu|^3 = \rho < \infty.$$

- $z = 1/\mu = \text{poly log } n$.

- mean μ and variance 1 \leftrightarrow mean 1 and variance μ^{-2} .
- WLOG, consider matrix $M = J + zA$ where J the all-ones matrix and each entry of A is iid sampled from \mathcal{D} :

$$\mathbb{E}_{x \sim \mathcal{D}}[x] = 0, \quad \text{Var}_{x \sim \mathcal{D}}[x] = 1, \quad \mathbb{E}_{x \sim \mathcal{D}} |x - \mu|^3 = \rho < \infty.$$

- $z = 1/\mu = \text{poly log } n$.
- ρ is a constant when
 - μ -biased standard (complex) Gaussian: unit variance and mean μ ;
 - biased Bernoulli: $-1 + \mu$ w.p. $1/2$ while $1 + \mu$ w.p. $1/2$.

Expansion of permanent

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$. $P_{n,n} = S_n$.

Expansion of permanent

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$. $P_{n,n} = S_n$.

$$\begin{aligned} \frac{\text{Per}(J + zA)}{n!} &= \frac{1}{n!} \sum_{\sigma \in P_{n,n}} \prod_{t=1}^n (1 + za_{t,\sigma(t)}) \\ &= \frac{1}{n!} \sum_{\sigma \in P_{n,n}} \sum_{k=0}^n \sum_{\{i_1, \dots, i_k\} \in C_{n,k}} \prod_{t=1}^k za_{i_t, \sigma(i_t)} \\ &= \frac{1}{n!} \sum_{k=0}^n z^k \sum_{\{i_1, \dots, i_k\} \in C_{n,k}} \sum_{\sigma \in P_{n,n}} \prod_{t=1}^k a_{i_t, \sigma(i_t)} \\ &= \frac{1}{n!} \sum_{k=0}^n z^k \sum_{\{i_1, \dots, i_k\} \in C_{n,k}} \sum_{(j_1, \dots, j_k) \in P_{n,k}} (n-k)! \prod_{t=1}^k a_{i_t, j_t} \end{aligned}$$

Expansion of permanent

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$. $P_{n,n} = S_n$.

$$\begin{aligned} \frac{\text{Per}(J + zA)}{n!} &= \frac{1}{n!} \sum_{k=0}^n z^k \sum_{\{i_1, \dots, i_k\} \in C_{n,k}} \sum_{(j_1, \dots, j_k) \in P_{n,k}} (n-k)! \prod_{t=1}^k a_{i_t, j_t} \\ &= \sum_{k=0}^n z^k \left(\frac{1}{n^k} \sum_{\{i_1, \dots, i_k\} \in C_{n,k}} \sum_{(j_1, \dots, j_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t} \right) \\ &= \sum_{k=0}^n \left(\frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{(i_1, \dots, i_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t} \right) z^k \\ &=: \sum_{k=0}^n a_k z^k. \end{aligned}$$

- Use $\sum_{k=0}^t a_k z^k$ to approximate $\frac{1}{n!} \text{Per}(J + zA)$. Here $t = \ln \frac{n}{\epsilon}$ and the running time is simply $\mathcal{O}(n^{2t})$.

- Use $\sum_{k=0}^t a_k z^k$ to approximate $\frac{1}{n!} \text{Per}(J + zA)$. Here $t = \ln \frac{n}{\epsilon}$ and the running time is simply $\mathcal{O}(n^{2t})$.
- It is sufficient to prove that w.p. $1 - o(1)$,

$$\left| \sum_{k=t+1}^n a_k z^k \right| < n^{-\gamma} \epsilon, \quad \left| \sum_{k=0}^t a_k z^k \right| > n^{-\gamma}.$$

A second moment bound

- $a_k = \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{(i_1, \dots, i_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t}$.
- Each a_k is a sum of degree- k multilinear monomials whose means are 0.

A second moment bound

- $a_k = \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{(i_1, \dots, i_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t}$.
- Each a_k is a sum of degree- k multilinear monomials whose means are 0.
- Any two monomials in a_k are uncorrelated.

A second moment bound

- $a_k = \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{(i_1, \dots, i_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t}$.
- Each a_k is a sum of degree- k multilinear monomials whose means are 0.
- Any two monomials in a_k are uncorrelated.
- Any monomial in a_k is uncorrelated with any monomial in a_l .

A second moment bound

- $a_k = \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{(i_1, \dots, i_k) \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t}$.
- Each a_k is a sum of degree- k multilinear monomials whose means are 0.
- Any two monomials in a_k are uncorrelated.
- Any monomial in a_k is uncorrelated with any monomial in a_l .
-

$$\mathbb{E}[a_k \bar{a}_l] = \mathbb{1}_{k=l} \cdot \frac{1}{k!}.$$

A second moment bound

- $\mathbb{E}[a_k \bar{a}_l] = \mathbb{1}_{k=l} \cdot \frac{1}{k!}$, $z = \text{poly log } n$, $t = \log \frac{n}{\epsilon}$.

$$\text{Var} \left[\sum_{k=t+1}^n a_k z^k \right] = \mathbb{E} \left[\sum_{k,l=t+1}^n a_k \bar{a}_l z^k \bar{z}^l \right] = \sum_{k=t+1}^n \frac{|z|^{2k}}{k!} = \left(\frac{\epsilon}{n} \right)^{\omega(1)}.$$

- Chebyshev's inequality.

Approximation of a_k 's

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$.

Approximation of a_k 's

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$.

$$\begin{aligned} a_k &= \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{\{i_1, \dots, i_k\} \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t} \\ &\approx \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{i_1, \dots, i_k \in [n]} \prod_{t=1}^k a_{i_t, j_t} \\ &= \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \left(\sum_{i=1}^n a_{i, j_1} \right) \cdots \left(\sum_{i=1}^n a_{i, j_k} \right) =: V_k. \end{aligned}$$

Approximation of a_k 's

- $C_{n,k} \triangleq \{\{i_1, \dots, i_k\} \subseteq [n]\}$;
- $P_{n,k} \triangleq \{(j_1, \dots, j_k) \in [n]^k : j_\alpha \neq j_\beta \text{ for } \alpha \neq \beta\}$.

$$\begin{aligned} a_k &= \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{\{i_1, \dots, i_k\} \in P_{n,k}} \prod_{t=1}^k a_{i_t, j_t} \\ &\approx \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \sum_{i_1, \dots, i_k \in [n]} \prod_{t=1}^k a_{i_t, j_t} \\ &= \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \left(\sum_{i=1}^n a_{i, j_1} \right) \cdots \left(\sum_{i=1}^n a_{i, j_k} \right) =: V_k. \end{aligned}$$

- $\left| \sum_{k=0}^t a_k z^k - \sum_{k=0}^t V_k z^k \right| < n^{-0.1}$ with high probability.

Approximation of a_k 's

- Let $C_j \triangleq \frac{a_{1,j} + a_{2,j} + \dots + a_{n,j}}{\sqrt{n}}$ to be the normalized sum of the j th column.

Approximation of a_k 's

- Let $C_j \triangleq \frac{a_{1,j} + a_{2,j} + \dots + a_{n,j}}{\sqrt{n}}$ to be the normalized sum of the j th column.
- C_j is “more concentrated” than $a_{i,j}$'s.

Approximation of a_k 's

- Let $C_j \triangleq \frac{a_{1,j} + a_{2,j} + \dots + a_{n,j}}{\sqrt{n}}$ to be the normalized sum of the j th column.
- C_j is “more concentrated” than $a_{i,j}$'s.
- Rewrite V_k with respect to C_j 's:

$$\begin{aligned} V_k &= \frac{1}{n^k} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} \left(\sum_{i=1}^n a_{i,j_1} \right) \left(\sum_{i=1}^n a_{i,j_2} \right) \cdots \left(\sum_{i=1}^n a_{i,j_k} \right) \\ &= \frac{1}{n^{k/2}} \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} C_{j_1} \cdots C_{j_k} \approx a_k. \end{aligned}$$

Properties of V_k 's

- $n^{k/2} V_k = \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} C_{j_1} \cdots C_{j_k}$'s are the elementary symmetric polynomials.

Properties of V_k 's

- $n^{k/2} V_k = \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} C_{j_1} \cdots C_{j_k}$'s are the elementary symmetric polynomials.
- Let $D_k \triangleq n^{-k/2} \sum_{j=1}^n C_j^k$.

Properties of V_k 's

- $n^{k/2} V_k = \sum_{\{j_1, \dots, j_k\} \in C_{n,k}} C_{j_1} \cdots C_{j_k}$'s are the elementary symmetric polynomials.
- Let $D_k \triangleq n^{-k/2} \sum_{j=1}^n C_j^k$.
- Newton's identities:

$$V_k = \frac{V_{k-1} V_1 - V_{k-2} D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}.$$

Approximation of V_k 's

- $V_k = \frac{V_{k-1}V_1 - V_{k-2}D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}$, where
 $D_k \triangleq n^{-\frac{k}{2}} \sum_j C_j^k$.

Approximation of V_k 's

- $V_k = \frac{V_{k-1} V_1 - V_{k-2} D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}$, where $D_k \triangleq n^{-\frac{k}{2}} \sum_j C_j^k$.
- D_2 concentrates to $\mathbb{E}[D_2] = \xi =: \mathbb{E}_{x \sim \mathcal{D}}[x^2]$.

Approximation of V_k 's

- $V_k = \frac{V_{k-1} V_1 - V_{k-2} D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}$, where $D_k \triangleq n^{-\frac{k}{2}} \sum_j C_j^k$.
- D_2 concentrates to $\mathbb{E}[D_2] = \xi =: \mathbb{E}_{x \sim \mathcal{D}}[x^2]$.
- $|D_k| \leq n^{-0.1k}$ for all $k \geq 3$ with high probability.

Approximation of V_k 's

- $V_k = \frac{V_{k-1} V_1 - V_{k-2} D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}$, where $D_k \triangleq n^{-\frac{k}{2}} \sum_j C_j^k$.
- D_2 concentrates to $\mathbb{E}[D_2] = \xi =: \mathbb{E}_{x \sim \mathcal{D}}[x^2]$.
- $|D_k| \leq n^{-0.1k}$ for all $k \geq 3$ with high probability.
- A good approximation of V_k :

$$V'_k = \begin{cases} 1, & k = 0, \\ V_1, & k = 1, \\ \frac{V'_{k-1} V_1 - V'_{k-2} \xi}{k}, & k \geq 2. \end{cases}$$

Approximation of V_k 's

- $V_k = \frac{V_{k-1} V_1 - V_{k-2} D_2 + \sum_{i=2}^{k-1} (-1)^i V_{k-1-i} D_{i+1}}{k}$, where $D_k \triangleq n^{-\frac{k}{2}} \sum_j C_j^k$.
- D_2 concentrates to $\mathbb{E}[D_2] = \xi =: \mathbb{E}_{x \sim \mathcal{D}}[x^2]$.
- $|D_k| \leq n^{-0.1k}$ for all $k \geq 3$ with high probability.
- A good approximation of V_k :

$$V'_k = \begin{cases} 1, & k = 0, \\ V_1, & k = 1, \\ \frac{V_{k-1} V_1 - V_{k-2} \xi}{k}, & k \geq 2. \end{cases}$$

- $\left| \sum_{k=0}^t V_k z^k - \sum_{k=0}^t V'_k z^k \right| < n^{-0.08}$ with high probability.

Approximation of V_k 's

- Represent V_k 's with “probabilists’ Hermite polynomials”:

$$V'_k = V'_k(V_1) = \begin{cases} 1, & k = 0 \\ V_1, & k = 1 \\ \frac{V'_{k-1} V_1 - V'_{k-2} \xi}{k}, & k \geq 2 \end{cases} = \frac{1}{k!} \xi^{\frac{k}{2}} H_{e_k} \left(V_1 / \sqrt{\xi} \right).$$

Approximation of V_k 's

- Represent V_k 's with “probabilists’ Hermite polynomials”:

$$V'_k = V'_k(V_1) = \begin{cases} 1, & k = 0 \\ V_1, & k = 1 \\ \frac{V'_{k-1} V_1 - V'_{k-2} \xi}{k}, & k \geq 2 \end{cases} = \frac{1}{k!} \xi^{\frac{k}{2}} H_{e_k} \left(V_1 / \sqrt{\xi} \right).$$

-

$$\sum_{k=0}^t V'_k z^k \approx \sum_{k=0}^{\infty} V'_k z^k = e^{V_1 z - \frac{\xi z^2}{2}} = (*).$$

$|z| \leq (\ln n)^{0.05} \rightarrow (*)$ is small only if $|V_1|$ is relatively large, which is with probability $o(1)$.

- $\frac{1}{n!} \text{Per}(J + zA) = \sum_{k=0}^n a_k z^k.$

Summary

- $\frac{1}{n!} \text{Per}(J + zA) = \sum_{k=0}^n a_k z^k$.
- $\left| \sum_{k=t+1}^n a_k z^k \right| < n^{-1} \epsilon$ with high probability.

Summary

- $\frac{1}{n!} \text{Per}(J + zA) = \sum_{k=0}^n a_k z^k$.
- $|\sum_{k=t+1}^n a_k z^k| < n^{-1}\epsilon$ with high probability.
-

$$\sum_{k=0}^t a_k z^k \approx \sum_{k=0}^t V_k z^k \approx \sum_{k=0}^t V'_k z^k \approx \sum_{k=0}^{\infty} V'_k z^k = e^{V_1 z - \frac{\xi z^2}{2}}$$

is large ($> \frac{1}{2}n^{-0.01}$) with high probability.

Drawbacks and future questions

- From a_k to V_k , we need $\text{Var}[z^k(a_k - V_k)]$ to be small, which does not hold for $\mu^{-1} = z = n^{0.0001}$.

Drawbacks and future questions

- From a_k to V_k , we need $\text{Var}[z^k(a_k - V_k)]$ to be small, which does not hold for $\mu^{-1} = z = n^{0.0001}$.
- Is there a (quasi-)polynomial algorithm for $\mu = 1/\text{poly}(n)$?

Drawbacks and future questions

- From a_k to V_k , we need $\text{Var}[z^k(a_k - V_k)]$ to be small, which does not hold for $\mu^{-1} = z = n^{0.0001}$.
- Is there a (quasi-)polynomial algorithm for $\mu = 1/\text{poly}(n)$?
- Is there a method to directly derive (approximately) the distribution of random permanents when $\mu = 1/\text{poly}(n)$?

Drawbacks and future questions

- From a_k to V_k , we need $\text{Var}[z^k(a_k - V_k)]$ to be small, which does not hold for $\mu^{-1} = z = n^{0.0001}$.
- Is there a (quasi-)polynomial algorithm for $\mu = 1/\text{poly}(n)$?
- Is there a method to directly derive (approximately) the distribution of random permanents when $\mu = 1/\text{poly}(n)$?
- Does the distribution of random permanents concentrated or rather anti-concentrated over \mathbb{C} ?

Thank you.