Approximating the Permanent of Positive Semidefinite Matrices

Nima Anari

Joint work with

Leonid Gurvits

Shayan Oveis Gharan

Amin Saberi
Determinant

\[ \text{det}(M) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) M_{1,\sigma(1)} \cdots M_{n,\sigma(n)} \]

Permanent

\[ \text{per}(M) = \sum_{\sigma \in S_n} M_{1,\sigma(1)} \cdots M_{n,\sigma(n)} \]
Determinant

\[
det(M) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) M_{1,\sigma(1)} \cdots M_{n,\sigma(n)}
\]

Permanent

\[
\text{per}(M) = \sum_{\sigma \in S_n} M_{1,\sigma(1)} \cdots M_{n,\sigma(n)}
\]

2 \times 2 Example

\[
M = \begin{bmatrix}
a & b \\
c & d \\
\end{bmatrix}
\]

\[
det(M) = ad - bc
\]

\[
\text{per}(M) = ad + bc
\]
Complexity of Permanent

- \#P-hard to compute $\text{per}(M)$ for 0/1 matrices [Valiant’79].
Complexity of Permanent

- #P-hard to compute $\text{per}(M)$ for 0/1 matrices [Valiant’79].

- #P-hard to compute sign of $\text{per}(M)$ [Aaronson’11].
Complexity of Permanent

- \( \#P \)-hard to compute \( \text{per}(M) \) for 0/1 matrices [Valiant’79].

- \( \#P \)-hard to compute sign of \( \text{per}(M) \) [Aaronson’11].
- \( \#P \)-hard to compute \( \text{per}(M) \) for \( M \succeq 0 \) [Grier-Schaeffer’16].
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:

\[ \text{per}(M) \geq 0. \]
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].

Positive Matrices ($M \succeq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]

- Randomized $(1 + \epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda’04].
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]

  - Randomized $(1 + \epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda’04].
  - Deterministic $2^n$-approximation [Gurvits-Samorodnitsky’14].

PSD Matrices ($M \succeq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]

- Randomized $(1 + \epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda’04].

- Deterministic $2^n$-approximation [Gurvits-Samorodnitsky’14].

PSD Matrices ($M \succeq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]

- Deterministic $n!$-approximation [Marcus’63]: $M_{1,1} \ldots M_{n,n}$.

Improved to $n!^k \frac{n}{k}$-approximation in time $2^{O(k + \log(n))}$ [Lieb’66].
Approximating the Permanent

Additive $\pm \epsilon |M|^n$ approximation [Gurvits’05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]
- Randomized $(1 + \epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda’04].
- Deterministic $2^n$-approximation [Gurvits-Samorodnitsky’14].

PSD Matrices ($M \succeq 0$)

- Permanent is always nonnegative:
  \[ \text{per}(M) \geq 0. \]
- Deterministic $n!$-approximation [Marcus’63]: $M_{1,1} \ldots M_{n,n}$.
- Improved to $\frac{n!}{k!n^k}$-approximation in time $2^{O(k + \log(n))}$ [Lieb’66].
Theorem [A-Gurvits-Oveis Gharan-Saberi’17]

The permanent of PSD matrices $M \in \mathbb{C}^{n \times n}$ can be approximated, in deterministic polynomial time, within

$$(e^{\gamma+1})^n \approx 4.84^n.$$
Complex Gaussians

$z \sim \mathbb{CN}(0, 1)$

$\mathbb{P}[z] = \frac{1}{\pi} e^{-|z|^2}$
Complex Gaussians

Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathcal{CN}(0, 1)$.

$z \sim \mathcal{CN}(0, 1)$

$\mathbb{P}[z] = \frac{1}{\pi}e^{-|z|^2}$
Complex Gaussians

- Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathcal{CN}(0, 1)$.
- General (circularly-symmetric) complex Gaussian:

\[
g = Cz, \\
g \sim \mathcal{CN}(0, CC^\dagger).
\]
Complex Gaussians

- Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathcal{CN}(0, 1)$.
- General (circularly-symmetric) complex Gaussian:
  \[
g = Cz, \quad g \sim \mathcal{CN}(0, CC^\dagger).
\]

Wick’s Formula

\[
\mathbb{E} \left[ |g_1|^2 \cdots |g_n|^2 \right] = \text{per}(CC^\dagger).
\]
Schur Power

The Schur power of an $n \times n$ matrix $M$ is

$$\left\{ \begin{array}{c}
\vdots \\
M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} \\
\vdots \\
\vdots
\end{array} \right\}_{n!} \left\{ \begin{array}{c}
\vdots \\
\vdots \\
\vdots
\end{array} \right\}_{n!}$$
Schur Power

The Schur power of an $n \times n$ matrix $M$ is

$$
\left\{ \begin{array}{c}
\vdots \\
\vdots \\
M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} \\
\vdots \\
\vdots \\
\end{array} \right\}
$$

$n!$

$n!$

- The Schur power is a minor of $M^\otimes n$.

$$
M \succeq 0 \implies \text{schur}(M) \succeq 0
$$
The Schur power of an $n \times n$ matrix $M$ is

$$\begin{bmatrix}
\vdots & \vdots & \vdots \\
\vdots & M_{\sigma(1),\tau(1)} & \vdots \\
\vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots \\
\end{bmatrix}$$

$n!$

$n!$

- The Schur power is a minor of $M^\otimes n$.

\[
M \succeq 0 \implies \text{schur}(M) \succeq 0
\]

- The permanent is an eigenvalue:

\[
\text{schur}(M)1 = \text{per}(M)1.
\]

\[
M \succeq 0 \implies \text{per}(M) \geq 0
\]
Schur Power

The Schur power of an $n \times n$ matrix $M$ is

$$
\begin{bmatrix}
\vdots & \vdots & \vdots \\
M_{\sigma(1),\tau(1)} & \cdots & M_{\sigma(n),\tau(n)} \\
\vdots & \vdots & \vdots \\
\end{bmatrix}
$$

- The Schur power is a minor of $M^{\otimes n}$.

- The permanent is an eigenvalue:

$$\text{schur}(M)1 = \text{per}(M)1.$$  

- Permanent is monotone w.r.t. $\succeq$:

$$M \succeq 0 \implies \text{per}(M) \geq 0$$

$$M_1 \succeq M_2 \succeq 0 \implies \text{per}(M_1) \geq \text{per}(M_2) \geq 0$$
Approximation using Monotonicity

- Permanent is monotone w.r.t. $\succeq$:

$$D \succeq M \succeq vv^\top \implies \per(D) \geq \per(M) \geq \per(vv^\top).$$
Approximation using Monotonicity

- Permanent is monotone w.r.t. \(\succeq\):

\[
D \succeq M \succeq vv^\top \implies \text{per}(D) \geq \text{per}(M) \geq \text{per}(vv^\top).
\]

**Theorem [A-Gurvits-Oveis Gharan-Saberi'17]**

For any \(M \succeq 0\) there exist diagonal matrix \(D\) and rank-1 matrix \(vv^\top\) such that

\[
D \succeq M \succeq vv^\top,
\]

and \(\text{per}(D) \leq 4.85^n \text{per}(vv^\top)\).
Computing the Approximation

Solve and output the following

$$\inf_D \per(D),$$
subject to $$D \succeq M.$$
Computing the Approximation

- Solve and output the following

\[ \inf_D \ \text{per}(D), \]
\[ \text{subject to } D \succeq M. \]

- Equivalently solve the convex program

\[ \inf_{D^{-1}} \ \log(\text{per}((D^{-1})^{-1})), \]
\[ \text{subject to } M^{-1} \succeq D^{-1} \succeq 0. \]

- No such convex program for the best rank-1 matrix.
Sketch of Proof

- Renormalize rows and columns to assume $D = I$. 

By duality, there is $B \succeq 0$ with $\text{diag}(B) = 1$ such that $\left( I \cdot M \right) B = 0$.

$B$ is called a correlation matrix.

Let $P = \text{proj}_{\text{imag} \left( B \right)}$. Then $M \succeq P$ because $x^2_{\text{imag} \left( B \right)} = x = Mx = MB = By = x = Px$.

Prove the "PSD Vander Waerden" PSD Van der Waerden [A-Gurvits-Oveis Gharan-Saberi'17]

If $B$ is a correlation matrix and $P$ the orthogonal projection onto the image of $B$, then $\text{per} \left( P \right) \leq 4^{\frac{10}{14}}$. 
Sketch of Proof

- Renormalize rows and columns to assume $D = I$.
- By duality, there is $B \succeq 0$ with $\text{diag}(B) = 1$ such that $(I - M)B = 0$:

  $$B = MB.$$

  B is called a correlation matrix.
Sketch of Proof

- Renormalize rows and columns to assume \( D = I \).
- By duality, there is \( B \succeq 0 \) with \( \text{diag}(B) = 1 \) such that \((I - M)B = 0\):
  \[
  B = MB.
  \]
  B is called a correlation matrix.
- Let \( P = \text{proj}_{\text{imag}(B)} \). Then \( M \succeq P \) because
  \[
  x \in \text{imag}(B) \implies x = By \implies Mx = MBy = By = x = Px.
  \]
Sketch of Proof

- Renormalize rows and columns to assume $D = I$.
- By duality, there is $B \succeq 0$ with $\text{diag}(B) = 1$ such that $(I - M)B = 0$:
  
  \[ B = MB. \]

$B$ is called a correlation matrix.

- Let $P = \text{proj}_{\text{imag}(B)}$. Then $M \succeq P$ because
  \[ x \in \text{imag}(B) \implies x = By \implies Mx = MBy = By = x = Px. \]

- Prove the “PSD Van der Waerden”

**PSD Van der Waerden [A-Gurvits-Oveis Gharan-Saberi’17]**

If $B$ is a correlation matrix and $P$ the orthogonal projection onto the image of $B$, then

\[ \text{per}(P) \geq 4.85^{-n}. \]
PSD Van der Waerden

Given correlation matrix $B$ (i.e. $B \succeq 0$ and $\text{diag}(B) = 1$), want to show

$$\text{per}(\text{proj}_{\text{imag}}(B)) \geq 4.85^{-n}.$$
PSD Van der Waerden

- Given correlation matrix $B$ (i.e. $B \succeq 0$ and $\text{diag}(B) = 1$), want to show

$$\text{per}(\text{proj}_{\text{imag}(B)}) \geq 4.85^{-n}.$$ 

- Show for some unit vector $v \in \text{imag}(B)$

$$\text{per}(vv^\dagger) \geq 4.85^{-n}.$$
PSD Van der Waerden

- Given correlation matrix $B$ (i.e. $B \succeq 0$ and $\text{diag}(B) = 1$), want to show

$$\text{per}(\text{proj}_{\text{imag}(B)}) \geq 4.85^{-n}.$$ 

- Show for some unit vector $v \in \text{imag}(B)$

$$\text{per}(vv^\dagger) \geq 4.85^{-n}.$$ 

- Let $B$ be the Gram matrix of unit vectors $u_1, \ldots, u_n$. Generate $v$ by normalizing the projection vector of $u_1, \ldots, u_n$ onto some direction $g$

$$v = \frac{[g^\dagger u_1 \ldots g^\dagger u_n]}{|[g^\dagger u_1 \ldots g^\dagger u_n]|}.$$
GM-AM Ratio

Let \( u \) be a random vector (e.g., uniformly sampled from \( u_1, \ldots, u_n \)). Define the GM-AM ratio as:

\[
\frac{e^{E[\log(|u|^2)]}}{E[|u|^2]}
\]
GM-AM Ratio

- Let $u$ be a random vector (e.g., uniformly sampled from $u_1, \ldots, u_n$). Define the GM-AM ratio as:

$$\frac{e^{\mathbb{E}[\log(|u|^2)]}}{\mathbb{E}[|u|^2]}$$

- The GM-AM ratio is always $\leq 1$. Equality happens when $|u| = 1$. 

Lemma [A-Gurvits-Oveis Gharan-Saberi'17] If $u$ is a random unit vector, there exists $g$ such that the GM-AM ratio of $gy$ is at least $e^{12/14}$. 


GM-AM Ratio

- Let $u$ be a random vector (e.g., uniformly sampled from $u_1, \ldots, u_n$). Define the GM-AM ratio as:

$$\frac{e^{E[\log(|u|^2)]}}{E[|u|^2]}$$

- The GM-AM ratio is always $\leq 1$. Equality happens when $|u| = 1$.

**Lemma [A-Gurvits-Oveis Gharan-Saberi’17]**

If $u$ is a random unit vector, there exists $g$ such that the GM-AM ratio of $g^\dagger u$ is at least $e^{-\gamma}$.
Complex Gaussians Come Back

Let $g$ be a standard complex Gaussian. Then with positive probability we have:

$$GM-AM(g^\dagger u) \geq \frac{E\left[e^{E[\log(|g^\dagger u|^2)]}\right]}{E[|g^\dagger u|^2]} \geq \frac{e^{E[\log(|g^\dagger u|^2)]}}{E[|g^\dagger u|^2]}$$
Complex Gaussians Come Back

Let $g$ be a standard complex Gaussian. Then with positive probability we have:

$$\text{GM-AM}(g^\dagger u) \geq \frac{\mathbb{E} \left[ e^{\mathbb{E} \left[ \log(|g^\dagger u|^2) \right]} \right]}{\mathbb{E} \left[ |g^\dagger u|^2 \right]} \geq \frac{\mathbb{E} \left[ e^{\mathbb{E} \left[ \log(|g^\dagger u|^2) \right]} \right]}{\mathbb{E} \left[ |g^\dagger u|^2 \right]}$$

But

$$\mathbb{E} \left[ \log(|g^\dagger u|^2) \right] = -\gamma,$$

and

$$\mathbb{E} \left[ |g^\dagger u|^2 \right] = 1.$$
Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?
Conclusion and Open Questions

- \((e^{\gamma+1})^n\)-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?
- Markov chains to get \((1 + \epsilon)\)-approximation?
Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?
- Markov chains to get $(1 + \epsilon)$-approximation?

thank you!