Bounding the Permanent of a Non-negative Matrix via its Degree-M Bethe and Sinkhorn Permanents

Yuwen Huang and Pascal O. Vontobel

Department of Information Engineering
The Chinese University of Hong Kong
yuwen.huang@ieee.org, pascal.vontobel@ieee.org

ISIT 2023

The journal version on Arxiv: 2306.02280

Outline

Outline

- **Overview of the main results**

An S-FG representation of the permanent

Analyzing the permanent and its degree-M Bethe permanent

Bounding the permanent via its approximations

Conclusion

Overview of the main results

Definition

- $[n] \triangleq \{1, 2, \ldots, n\}$.
- $\boldsymbol{\theta} \triangleq \big(\theta(i,j)\big)_{i,j\in[n]} \in \mathbb{R}_{\geq 0}^{n\times n}$: **a non-negative real-valued matrix**.
- $\mathcal{S}_{[n]}$ is the set of all $n!$ **permutations** in $[n]$.
- The determinant:

$$\det(\boldsymbol{\theta}) \triangleq \sum_{\sigma\in\mathcal{S}_{[n]}} \mathrm{sgn}(\sigma) \cdot \prod_{i\in[n]} \theta\big(i, \sigma(i)\big).$$

  The complexity of computing the determinant is $O(n^3)$.
- The permanent:

$$\mathrm{perm}(\boldsymbol{\theta}) \triangleq \sum_{\sigma\in\mathcal{S}_{[n]}} \prod_{i\in[n]} \theta\big(i, \sigma(i)\big).$$

  Computing the permanent is in the complexity class **#P** (a counting problem in the class **NP**).

Overview of the main results

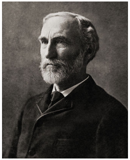
The Bethe permanent $\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})$ is a **graphical-model-based** method for approximating the permanent of a **non-negative matrix**.

$$1 \leq \frac{\mathrm{perm}(\boldsymbol{\theta})}{\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})} \leq 2^{n/2}.$$

- The first inequality was **proven** by Gurvits [Gurvits, 2011] with the help of an **inequality by Schrijver** [Schrijver, 1998].
- The second inequality was **conjectured** by Gurvits [Gurvits, 2011] and **proven** by Anari and Rezaei [Anari and Rezaei, 2019].

The **sum-product algorithm (SPA)** finds $\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})$ **efficiently**.

Overview of the main results

	 Josiah W. Gibbs	 Hans Bethe
	Permanent	Bethe permanent
Combinatorial	$\mathrm{perm}(\boldsymbol{\theta}) = \sum_{\sigma \in \mathcal{S}_{[n]}} \prod_{i \in [n]} \theta\big(i, \sigma(i)\big)$ (the sum of weighted configurations)	???
Analytical	$\mathrm{perm}(\boldsymbol{\theta}) = \exp\left(- \min_{\gamma \in \Gamma_n} F'_{\mathrm{G},\boldsymbol{\theta}}(\gamma) \right)$	$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \exp\left(- \min_{\gamma \in \Gamma_n} F_{\mathrm{B},\boldsymbol{\theta}}(\gamma) \right)$

Overview of the main results

Main idea: Bound $\mathrm{perm}(\boldsymbol{\theta})$ via $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$.

Definition [Vontobel, 2013a] Let $M \in \mathbb{Z}_{\geq 1}$.

The **degree-M Bethe permanent** is defined to be

$$\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}) \triangleq \sqrt[M]{\left\langle \mathrm{perm}(\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}) \right\rangle_{\boldsymbol{P}_M \in \tilde{\Psi}_M}},$$

where

$$\left\langle \mathrm{perm}(\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}) \right\rangle_{\boldsymbol{P}_M \in \tilde{\Psi}_M} \triangleq \frac{1}{|\tilde{\Psi}_M|} \sum_{\boldsymbol{P}_M \in \tilde{\Psi}_M} \mathrm{perm}(\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}),$$

and $\tilde{\Psi}_M$ is **the set of all possible P_M-lifting** of $\boldsymbol{\theta}$.

Overview of the main results

$$\boldsymbol{\theta} = \begin{pmatrix} \theta(1,1) & \cdots & \theta(1,n) \\ \vdots & \ddots & \vdots \\ \theta(n,1) & \cdots & \theta(n,n) \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

The **P_M-lifting of θ** is defined to be

$$\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M} \triangleq \begin{pmatrix} \theta(1,1) \cdot \boldsymbol{P}^{(1,1)} & \cdots & \theta(1,n) \cdot \boldsymbol{P}^{(1,n)} \\ \vdots & \ddots & \vdots \\ \theta(n,1) \cdot \boldsymbol{P}^{(n,1)} & \cdots & \theta(n,n) \cdot \boldsymbol{P}^{(n,n)} \end{pmatrix} \in \mathbb{R}_{\geq 0}^{Mn \times Mn}, \qquad \boldsymbol{P}_M \in \tilde{\Psi}_M.$$

Overview of the main results

Consider

$$\boldsymbol{\theta} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

For $M = 2$, a possible $\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}$ is given by

$$\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M} = \left(\begin{array}{c|c} a \cdot \boldsymbol{P}^{(1,1)} & b \cdot \boldsymbol{P}^{(1,2)} \\ \hline c \cdot \boldsymbol{P}^{(2,1)} & d \cdot \boldsymbol{P}^{(2,2)} \end{array}\right)$$

$$= \left(\begin{array}{cc|cc} a & 0 & b & 0 \\ 0 & a & 0 & b \\ \hline c & 0 & d & 0 \\ 0 & c & 0 & d \end{array}\right),$$

where

$$\boldsymbol{P}^{(1,1)} = \boldsymbol{P}^{(1,2)} = \boldsymbol{P}^{(2,1)} = \boldsymbol{P}^{(2,2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Overview of the main results

Consider

$$\boldsymbol{\theta} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

For $M = 2$, a possible $\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}$ is given by

$$\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M} = \left(\begin{array}{c|c} a \cdot \boldsymbol{P}^{(1,1)} & b \cdot \boldsymbol{P}^{(1,2)} \\ \hline c \cdot \boldsymbol{P}^{(2,1)} & d \cdot \boldsymbol{P}^{(2,2)} \end{array}\right)$$

$$= \left(\begin{array}{cc|cc} a & 0 & b & 0 \\ 0 & a & 0 & b \\ \hline c & 0 & 0 & d \\ 0 & c & d & 0 \end{array}\right),$$

where

$$\boldsymbol{P}^{(1,1)} = \boldsymbol{P}^{(1,2)} = \boldsymbol{P}^{(2,1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{P}^{(2,2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

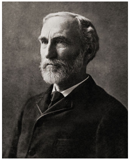
Overview of the main results

Theorem [Vontobel, 2013b]

A **combinatorial characterization** for $\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})$:

$$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \limsup_{M \to \infty} \mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}).$$

Overview of the main results

	 Josiah W. Gibbs	 Hans Bethe
	Permanent	Bethe permanent
Combinatorial	$\mathrm{perm}(\boldsymbol{\theta}) = \sum_{\sigma \in \mathcal{S}_{[n]}} \prod_{i \in [n]} \theta\big(i, \sigma(i)\big)$	$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \limsup_{M \to \infty} \mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}).$
Analytical	$\mathrm{perm}(\boldsymbol{\theta}) = \exp\left(-\min_{\gamma \in \Gamma_n} F'_{\mathrm{G},\boldsymbol{\theta}}(\gamma)\right)$	$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \exp\left(-\min_{\gamma \in \Gamma_n} F_{\mathrm{B},\boldsymbol{\theta}}(\gamma)\right)$

Overview of the main results

Definition Define

$$\boldsymbol{\theta}^{M\cdot\boldsymbol{\gamma}} \triangleq \prod_{i,j\in[n]} \big(\theta(i,j)\big)^{M\cdot\gamma(i,j)}, \qquad \boldsymbol{\gamma} \in \Gamma_{M,n}, \quad M\cdot\gamma(i,j)\in\mathbb{Z}_{\geq 0},$$

where $\Gamma_{M,n}$ is the set of **doubly stochastic matrices** of size $n \times n$ with all entries being **integer multiples** of $1/M$.

The **first main result**:

Lemma There are collections of **non-negative real** numbers

$$\big\{C_{M,n}(\boldsymbol{\gamma})\big\}_{\boldsymbol{\gamma}\in\Gamma_{M,n}}, \quad \big\{C_{\mathrm{B},M,n}(\boldsymbol{\gamma})\big\}_{\boldsymbol{\gamma}\in\Gamma_{M,n}},$$

such that

$$\big(\mathrm{perm}(\boldsymbol{\theta})\big)^M = \sum_{\boldsymbol{\gamma}\in\Gamma_{M,n}} \boldsymbol{\theta}^{M\cdot\boldsymbol{\gamma}} \cdot C_{M,n}(\boldsymbol{\gamma}),$$

$$\big(\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})\big)^M = \sum_{\boldsymbol{\gamma}\in\Gamma_{M,n}} \boldsymbol{\theta}^{M\cdot\boldsymbol{\gamma}} \cdot C_{\mathrm{B},M,n}(\boldsymbol{\gamma}).$$

Overview of the main results

The **second main result**:

Theorem

For every $\gamma \in \Gamma_{M,n}$, the coefficients $C_{M,n}(\gamma)$ and $C_{\mathrm{B},M,n}(\gamma)$ satisfy

$$1 \leq \frac{C_{M,n}(\gamma)}{C_{\mathrm{B},M,n}(\gamma)} \leq (2^{n/2})^{M-1}.$$

Then we **bound** $\mathrm{perm}(\boldsymbol{\theta})$ via $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$:

$$1 \leq \frac{\mathrm{perm}(\boldsymbol{\theta})}{\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})} < \left(2^{n/2}\right)^{\frac{M-1}{M}}.$$

This theorem **proves some of the conjectures** in [Vontobel, 2013a].

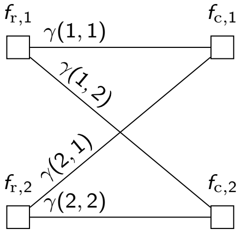
Outline

An S-FG representation of the permanent

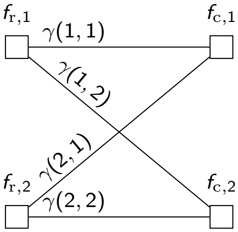
The standard factor graph (S-FG) N for $\boldsymbol{\theta}$ consists of

$$\boldsymbol{\theta} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1. edges: $(1,1), (1,2), (2,1), (2,2)$;
2. variables associated with edges: $\gamma(1,1), \gamma(1,2), \gamma(2,1), \gamma(2,2)$;
3. **binary alphabets**: $\mathcal{X}_{1,1}, \mathcal{X}_{1,2}, \mathcal{X}_{2,1}, \mathcal{X}_{2,2} = \{0,1\}$ for variables $\gamma(1,1), \gamma(1,2), \gamma(2,1), \gamma(2,2)$, respectively;
4. $\boldsymbol{\gamma} \triangleq \begin{pmatrix} \gamma(1,1) & \gamma(1,2) \\ \gamma(2,1) & \gamma(2,2) \end{pmatrix} \in \{0,1\}^{n \times n}$.
5. **nonnegative-valued** local functions $f_{\mathrm{r},1}$, $f_{\mathrm{r},2}$, and $f_{\mathrm{c},1}$, $f_{\mathrm{c},2}$;



An S-FG representation of the permanent



The details of the standard factor graph (S-FG) N for $\boldsymbol{\theta}$ are as follows:

- the global function:

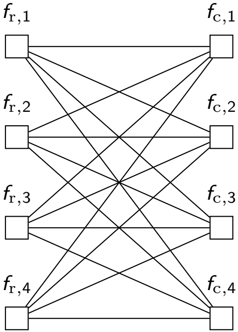
$$g(\gamma) \triangleq f_{\mathrm{r},1}\big(\gamma(1,:)\big) \cdot f_{\mathrm{r},2}\big(\gamma(2,:)\big) \cdot f_{\mathrm{c},1}\big(\gamma(:,1)\big) \cdot f_{\mathrm{c},2}\big(\gamma(:,2)\big);$$

- the partition function:

$$Z(\boldsymbol{\theta}) = \sum_{\boldsymbol{\gamma} \in \{0,1\}^{2\times 2}} g(\gamma) = \mathrm{perm}(\boldsymbol{\theta}).$$

An S-FG representation of the permanent

$$\boldsymbol{\theta} = \begin{pmatrix} \theta(1,1) & \cdots & \theta(1,4) \\ \vdots & \ddots & \vdots \\ \theta(4,1) & \cdots & \theta(4,4) \end{pmatrix} \in \mathbb{R}_{\geq 0}^{4\times 4}.$$



An S-FG representation of the permanent

	Permanent	Bethe permanent
Combinatorial	$\mathrm{perm}(\boldsymbol{\theta}) = \sum_{\sigma \in \mathcal{S}_{[n]}} \prod_{i \in [n]} \theta\big(i, \sigma(i)\big)$	$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \limsup_{M \to \infty} \mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$.
Analytical	$\mathrm{perm}(\boldsymbol{\theta}) = \exp\left(-\min_{\gamma \in \Gamma_n} F'_{\mathrm{G},\boldsymbol{\theta}}(\gamma)\right)$	$\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta}) = \exp\left(-\min_{\gamma \in \Gamma_n} F_{\mathrm{B},\boldsymbol{\theta}}(\gamma)\right)$
Complexity	**#P complete**	**Running the sum-product algorithm on the associated S-FG finds $\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})$ efficiently.**

$$1 \leq \frac{\mathrm{perm}(\boldsymbol{\theta})}{\mathrm{perm}_{\mathrm{B}}(\boldsymbol{\theta})} \leq 2^{n/2}.$$

Main Question

Can we bound $\mathrm{perm}(\theta)$ **via** $\mathrm{perm}_{\mathrm{B},M}(\theta)$**?**

This is indeed the case.

Outline

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example ($n = 2$ and $M = 2$)

$$\boldsymbol{\theta} \triangleq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{\geq 0}^{n \times n}.$$

---

1. We define Γ_n be the set of all **doubly stochastic matrices of size** $n \times n$.

2. We define $\Gamma_{M,n}$ to be **the subset of Γ_n** that contains all matrices where the entries are **multiples of $1/M$**.

3. $\boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}} \triangleq \prod_{i,j \in [n]} \big(\theta(i,j)\big)^{M \cdot \gamma(i,j)}$, for $\boldsymbol{\gamma} \in \Gamma_{M,n}$.

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example continued ($n = 2$ and $M = 2$)

For, $k_1, k_2 \in \mathbb{Z}_{\geq 0}$, define

$$\gamma^{(k_1,k_2)} \triangleq \frac{1}{k_1 + k_2} \cdot \begin{pmatrix} k_1 & k_2 \\ k_2 & k_1 \end{pmatrix} \in \Gamma_{k_1+k_2,2} \,.$$

The permanent and the degree-M Bethe permanent satisfy

$$\begin{aligned}
\mathrm{perm}(\boldsymbol{\theta}) &= a \cdot d + b \cdot c, \\
(\mathrm{perm}(\boldsymbol{\theta}))^2 &= \big(a \cdot d\big)^2 + 2 \cdot a \cdot b \cdot c \cdot d + \big(c \cdot b\big)^2 \\
&= 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,0)}} + 2 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,1)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(0,1)}}, \\
(\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}))^2 &= \big\langle \mathrm{perm}(\boldsymbol{\theta}^{\uparrow \boldsymbol{P}_M}) \big\rangle_{\boldsymbol{P}_M \in \tilde{\Psi}_M} \\
&= \big(a \cdot d\big)^2 + a \cdot b \cdot c \cdot d + \big(c \cdot b\big)^2 \\
&= 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,0)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,1)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(0,1)}}.
\end{aligned}$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example continued ($n = 2$ and $M = 2$)

$$2 \cdot (\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}))^2 = 2 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,0)}} + 2 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,1)}} + 2 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(0,1)}},$$

$$(\mathrm{perm}(\boldsymbol{\theta}))^2 = 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,0)}} + 2 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,1)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(0,1)}},$$

$$(\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta}))^2 = 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,0)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(1,1)}} + 1 \cdot \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}^{(0,1)}}.$$

Analyzing $\operatorname{perm}(\boldsymbol{\theta})$ and $\operatorname{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example continued ($n = 2$ and $M = 2$)

There are collections of coefficients

$$\{C_{M,n}(\boldsymbol{\gamma})\}_{\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^{(0,1)},\boldsymbol{\gamma}^{(1,0)},\boldsymbol{\gamma}^{(1,1)}\}}, \qquad \{C_{\mathrm{B},M,n}(\boldsymbol{\gamma})\}_{\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^{(0,1)},\boldsymbol{\gamma}^{(1,0)},\boldsymbol{\gamma}^{(1,1)}\}},$$

such that

$$(\operatorname{perm}(\boldsymbol{\theta}))^2 = \sum_{\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^{(0,1)},\boldsymbol{\gamma}^{(1,0)},\boldsymbol{\gamma}^{(1,1)}\}} C_{M,n}(\boldsymbol{\gamma}) \cdot \boldsymbol{\theta}^{M\cdot\boldsymbol{\gamma}},$$

$$(\operatorname{perm}_{\mathrm{B},M}(\boldsymbol{\theta}))^2 = \sum_{\boldsymbol{\gamma}\in\{\boldsymbol{\gamma}^{(0,1)},\boldsymbol{\gamma}^{(1,0)},\boldsymbol{\gamma}^{(1,1)}\}} C_{\mathrm{B},M,n}(\boldsymbol{\gamma}) \cdot \boldsymbol{\theta}^{M\cdot\boldsymbol{\gamma}}.$$

The following bounds hold

$$1 \leq \frac{C_{M,n}(\boldsymbol{\gamma})}{C_{\mathrm{B},M,n}(\boldsymbol{\gamma})} \leq 2, \qquad 1 \leq \frac{(\operatorname{perm}(\boldsymbol{\theta}))^2}{(\operatorname{perm}_{\mathrm{B},M}(\boldsymbol{\theta}))^2} < 2.$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example ($n = 2$ and arbitrary $M \in \mathbb{Z}_{\geq 1}$)

$$
\begin{aligned}
\big(\mathrm{perm}(\boldsymbol{\theta})\big)^{M+1} &= (a \cdot d + b \cdot c)^{M+1} \\
&= (a \cdot d + b \cdot c)^{M} \cdot (a \cdot d + b \cdot c) \\
&= \left(\sum_{k=0}^{M} \binom{M}{k} \cdot a^k \cdot d^k \cdot b^{M-k} \cdot c^{M-k} \right) \cdot (a \cdot d + b \cdot c) \\
&= \sum_{k=0}^{M+1} \underbrace{\left(\binom{M}{k-1} + \binom{M}{k} \right)}_{\binom{M+1}{k}} \cdot a^k \cdot d^k \cdot b^{M+1-k} \cdot c^{M+1-k} \\
&= \sum_{k=0}^{M+1} \binom{M+1}{k} \cdot a^k \cdot d^k \cdot b^{M-k} \cdot c^{M-k}.
\end{aligned}
$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example continued ($n = 2$ and arbitrary $M \in \mathbb{Z}_{\geq 1}$)

For the above special setup, it holds that

$$C_{M,n}\big(\gamma^{(k,M-k)}\big) = \binom{M}{k}.$$

The recursion

$$\binom{M+1}{k} = \binom{M}{k-1} + \binom{M}{k},$$

is equivalent to

$$C_{M+1,n}\big(\gamma^{(k,M+1-k)}\big) = C_{M,n}\big(\gamma^{(k-1,M+1-k)}\big) + C_{M,n}\big(\gamma^{(k,M-k)}\big).$$

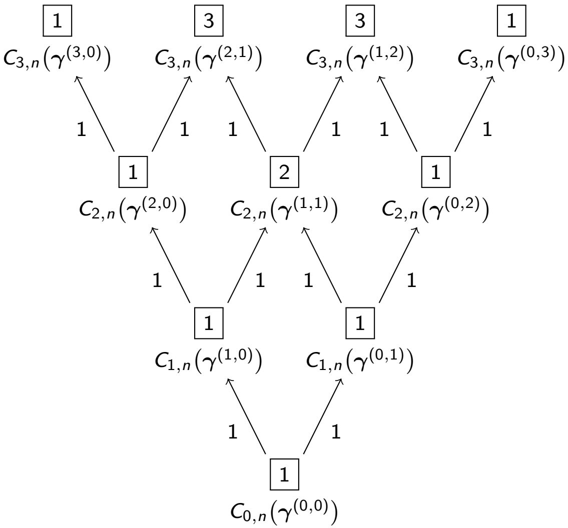


Figure: Pascal's triangle visualizing the recursion for $C_{M,n}$.

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example continued ($n = 2$ and arbitrary $M \in \mathbb{Z}_{\geq 1}$)

For the above special setup, it holds that

$$C_{\mathrm{B},M,n}\big(\gamma^{(k,M-k)}\big) = 1.$$

We have the recursion

$$\begin{aligned} & C_{\mathrm{B},M+1,n}\big(\gamma^{(k,M+1-k)}\big) \\ &= \begin{cases} C_{\mathrm{B},M,n}\big(\gamma^{(k,M-k)}\big) & k = 0 \\ C_{\mathrm{B},M,n}\big(\gamma^{(k-1,M+1-k)}\big) & k = M+1 \\ \frac{1}{2} \cdot C_{\mathrm{B},M,n}\big(\gamma^{(k-1,M+1-k)}\big) + \frac{1}{2} \cdot C_{\mathrm{B},M,n}\big(\gamma^{(k,M-k)}\big) & 1 \leq k \leq M \end{cases} \end{aligned}$$

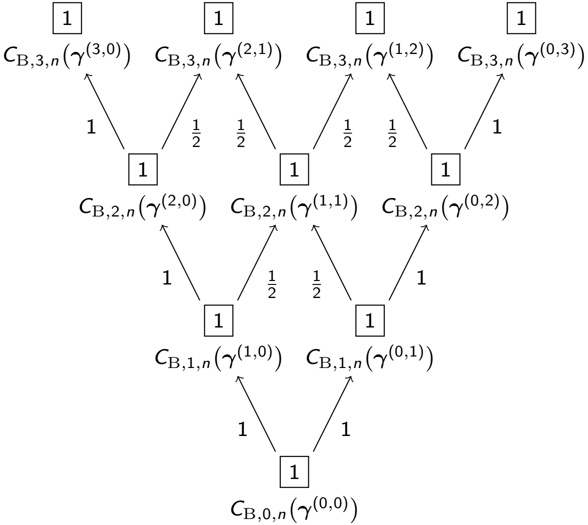


Figure: Generalization of Pascal's triangle visualizing the recursion $C_{\mathrm{B},M,n}$.

Analyzing $\operatorname{perm}(\boldsymbol{\theta})$ and $\operatorname{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

General Case (Arbitrary $n, M \in \mathbb{Z}_{\geq 1}$)

Lemma Consider collections of **non-negative real numbers**

$$\left\{ C_{M,n}(\boldsymbol{\gamma}) \right\}_{\boldsymbol{\gamma} \in \Gamma_{M,n}}, \quad \left\{ C_{\mathrm{B},M,n}(\boldsymbol{\gamma}) \right\}_{\boldsymbol{\gamma} \in \Gamma_{M,n}}.$$

The permanent and its degree-M Bethe permanent satisfy

$$\left(\operatorname{perm}(\boldsymbol{\theta})\right)^M = \sum_{\boldsymbol{\gamma} \in \Gamma_{M,n}} \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}} \cdot C_{M,n}(\boldsymbol{\gamma}),$$

$$\left(\operatorname{perm}_{\mathrm{B},M}(\boldsymbol{\theta})\right)^M = \sum_{\boldsymbol{\gamma} \in \Gamma_{M,n}} \boldsymbol{\theta}^{M \cdot \boldsymbol{\gamma}} \cdot C_{\mathrm{B},M,n}(\boldsymbol{\gamma}).$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

General Case (Arbitrary $n, M \in \mathbb{Z}_{\geq 1}$)

Lemma Let $M \in \mathbb{Z}_{\geq 2}$ and $\gamma \in \Gamma_{M,n}$. The following recursions hold

$$C_{M,n}(\gamma) = \sum_{\sigma_1 \in \mathcal{S}_{[n]}(\gamma)} C_{M-1,n}\big(\gamma_{\sigma_1}\big),$$

$$C_{\mathrm{B},M,n}(\gamma) = \frac{1}{\mathrm{perm}(\hat{\gamma}_{\mathcal{R},\mathcal{C}})} \cdot \sum_{\sigma_1 \in \mathcal{S}_{[n]}(\gamma)} C_{\mathrm{B},M-1,n}\big(\gamma_{\sigma_1}\big).$$

- The main idea is to **bound** $C_{M,n}(\gamma)$ via $C_{\mathrm{B},M,n}(\gamma)$ using **bounds** on $\mathrm{perm}(\hat{\gamma}_{\mathcal{R},\mathcal{C}})$.
- The details of $\mathrm{perm}(\hat{\gamma}_{\mathcal{R},\mathcal{C}})$ and γ_{σ_1} are **omitted** here.

Outline

Bounding the permanent via its approximations

Lemma We **bound** $C_{M,n}$ via $C_{\mathrm{B},M,n}$:

$$1 \leq \frac{C_{M,n}(\gamma)}{C_{\mathrm{B},M,n}(\gamma)} \leq (2^{n/2})^{M-1}.$$

Theorem Based on

$$\big(\mathrm{perm}(\boldsymbol{\theta})\big)^M = \sum_{\gamma \in \Gamma_{M,n}} \boldsymbol{\theta}^{M \cdot \gamma} \cdot C_{M,n}(\gamma),$$

$$\big(\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})\big)^M = \sum_{\gamma \in \Gamma_{M,n}} \boldsymbol{\theta}^{M \cdot \gamma} \cdot C_{\mathrm{B},M,n}(\gamma),$$

we **bound** the permanent $\mathrm{perm}(\boldsymbol{\theta})$ via its degree-M Bethe permanent:

$$1 \leq \frac{\mathrm{perm}(\boldsymbol{\theta})}{\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})} < \left(2^{n/2}\right)^{\frac{M-1}{M}}.$$

Bounding the permanent via its approximations

Another **well-known approximation** to $\mathrm{perm}(\boldsymbol{\theta})$ is the **scaled Sinkhorn permanent** [Anari et al., 2021].

Theorem

We **bound** $\mathrm{perm}(\boldsymbol{\theta})$ via its degree-M scaled Sinkhorn permanent:

$$\left(\mathrm{perm}_{\mathrm{scS},M}(\boldsymbol{\theta})\right)^M = \sum_{\gamma \in \Gamma_{M,n}} \boldsymbol{\theta}^{M \cdot \gamma} \cdot C_{\mathrm{scS},M,n}(\gamma),$$

$$\left(\frac{M^M}{M!}\right)^n \cdot \left(\frac{n!}{n^n}\right)^{M-1} \leq \frac{C_{M,n}(\gamma)}{C_{\mathrm{scS},M,n}(\gamma)} \leq \left(\frac{M^M}{M!}\right)^n,$$

$$\frac{M^n}{(M!)^{n/M}} \cdot \left(\frac{n!}{n^n}\right)^{\frac{M-1}{M}} < \frac{\mathrm{perm}(\boldsymbol{\theta})}{\mathrm{perm}_{\mathrm{scS},M}(\boldsymbol{\theta})} \leq \frac{M^n}{(M!)^{n/M}}.$$

Outline

Overview of the main results

Analyzing the permanent and its degree-M Bethe permanent

Bounding the permanent via its approximations

- **Conclusion**

Conclusion

- It is possible to **bound the permanent** of a non-negative matrix by its **degree-M Bethe and scaled Sinkhorn permanents**.
- Our main results **prove conjectures** in [Vontobel, 2013a].
- Our proofs used some **rather strong results** from [Schrijver, 1998, Gurvits, 2011, Anari and Rezaei, 2019, Egorychev, 1981, Falikman, 1981].
- We leave it as an open problem to find **"more basic" proofs** for some of the inequalities that were established in this paper.

Selected References I

Anari, N., Charikar, M., Shiragur, K., and Sidford, A. (2021).
The Bethe and Sinkhorn permanents of low rank matrices and implications for profile maximum likelihood.
In *Proc. 34th Conf. Learn. Theory*, pages 93–158, Boulder, CO, USA.

Anari, N. and Rezaei, A. (2019).
A tight analysis of Bethe approximation for permanent.
In *Proc. IEEE Ann. Symp. on Found. Comp. Sci.*, pages 1434–1445.

Egorychev, G. P. (1981).
The solution of an van der Waerden's problem for permanents.
Adv. Math., 42:299–305.

Selected References II

Falikman, D. I. (1981).
Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix.
Mathematical Notes of the Academy of Sciences of the USSR, 29(6):475–479.

Gurvits, L. (2011).
Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation.
Elec. Coll. Comp. Compl.

Schrijver, A. (1998).
Counting 1-factors in regular bipartite graphs.
J. Comb. Theory, Ser. B, 72(1):122–135.

Selected References III

Vontobel, P. O. (2013a).
The Bethe permanent of a nonnegative matrix.
IEEE Trans. Inf. Theory, 59(3):1866–1901.

Vontobel, P. O. (2013b).
Counting in graph covers: A combinatorial characterization of the Bethe entropy function.
IEEE Trans. Inf. Theory, 59(9):6018–6048.

Thank you!

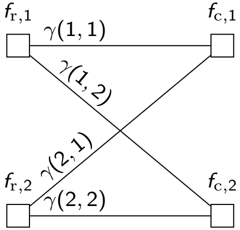
The journal version on Arxiv: 2306.02280

An S-FG representation of the permanent

The details of the standard factor graph (S-FG) N for $\boldsymbol{\theta}$ are as follows.

- local functions:

$$f_{\mathrm{r},1}(\gamma(1,:)) \triangleq \begin{cases} \sqrt{a} & \gamma(1,:) = \begin{pmatrix} 1 & 0 \end{pmatrix} \\ \sqrt{b} & \gamma(1,:) = \begin{pmatrix} 0 & 1 \end{pmatrix} \\ 0 & \text{Otherwise} \end{cases}$$



$$f_{\mathrm{r},2}(\gamma(2,:)) \triangleq \begin{cases} \sqrt{c} & \gamma(2,:) = \begin{pmatrix} 1 & 0 \end{pmatrix} \\ \sqrt{d} & \gamma(2,:) = \begin{pmatrix} 0 & 1 \end{pmatrix} \\ 0 & \text{Otherwise} \end{cases}$$

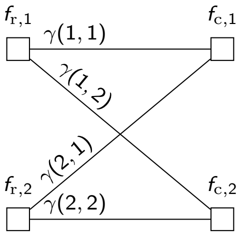
An S-FG representation of the permanent

The details of the standard factor graph (S-FG) N for $\boldsymbol{\theta}$ are as follows.

- local functions:

$$f_{\mathrm{c},1}(\gamma(:,1)) \triangleq \begin{cases} \sqrt{a} & \gamma(:,1) = \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathsf{T}} \\ \sqrt{c} & \gamma(:,1) = \begin{pmatrix} 0 & 1 \end{pmatrix}^{\mathsf{T}} \\ 0 & \text{Otherwise} \end{cases}$$

$$f_{\mathrm{c},2}(\gamma(:,2)) \triangleq \begin{cases} \sqrt{b} & \gamma(:,2) = \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathsf{T}} \\ \sqrt{d} & \gamma(:,2) = \begin{pmatrix} 0 & 1 \end{pmatrix}^{\mathsf{T}} \\ 0 & \text{Otherwise} \end{cases}$$



Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Example Let $M = 2$. Consider

$$\boldsymbol{\gamma}^{(1,1)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \Gamma_{2,2}.$$

If $\sigma_1 \in \mathcal{S}_{[n]}(\boldsymbol{\gamma})$ is chosen to be

$$\sigma_1(1) = 1, \quad \sigma_1(2) = 2,$$

then

$$\boldsymbol{\gamma}_{\sigma_1} = \frac{1}{2-1} \cdot \left(2 \cdot \boldsymbol{\gamma}^{(1,0)} - \boldsymbol{P}_{\sigma_1}\right) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \boldsymbol{\gamma}^{(0,1)} \in \Gamma_{1,2}.$$

It holds that

$$C_{2,n}(\boldsymbol{\gamma}^{(1,1)}) = 2, \qquad C_{1,n}(\boldsymbol{\gamma}^{(0,1)}) = 1.$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Let $M = 2$. Consider

$$\gamma^{(1,1)} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \Gamma_{2,2}.$$

If $\sigma_1 \in \mathcal{S}_{[n]}(\gamma)$ is chosen to be

$$\sigma_1(1) = 2, \quad \sigma_1(2) = 1,$$

then

$$\gamma_{\sigma_1} = \frac{1}{2-1} \cdot \left(2 \cdot \gamma^{(1,1)} - \boldsymbol{P}_{\sigma_1}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \gamma^{(1,0)} \in \Gamma_{1,2}.$$

It holds that

$$C_{2,n}(\gamma^{(1,1)}) = 2, \qquad C_{1,n}(\gamma^{(1,0)}) = 1.$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Definition Consider $M, N \in \mathbb{Z}_{\geq 1}$ and $\gamma \in \Gamma_{M,N}$.

1. The coefficient $C_{M,N}(\gamma)$ is defined to be **the number of** $\boldsymbol{\sigma}_{[M]} = (\sigma_1, \ldots, \sigma_M)$ in $\mathcal{S}_{[N]}^M$ such that $\boldsymbol{\sigma}_{[M]}$ **decomposes** γ, *i.e.*,
$$C_{M,N}(\gamma) = \sum_{\boldsymbol{\sigma}_{[M]} \in \mathcal{S}_{[N]}^M} \left[\gamma = \left\langle \boldsymbol{P}_{\sigma_m} \right\rangle_{m \in [M]} \right].$$
where $[S] \triangleq 1$ if the statement S is **true** and $[S] \triangleq 0$ if the statement is **false** and
$$\left\langle \boldsymbol{P}_{\sigma_m} \right\rangle_{m \in [M]} \triangleq \frac{1}{M} \cdot \sum_{m \in [M]} \boldsymbol{P}_{\sigma_m}.$$
2. $$C_{\mathrm{B},M,N}(\gamma) = (M!)^{2N - N^2} \cdot \prod_{i,j} \frac{(M - M \cdot \gamma(i,j))!}{(M \cdot \gamma(i,j))!}.$$

Analyzing $\mathrm{perm}(\boldsymbol{\theta})$ and $\mathrm{perm}_{\mathrm{B},M}(\boldsymbol{\theta})$

Definition Consider

$$M \in \mathbb{Z}_{\geq 2}, \quad \boldsymbol{\gamma} \in \Gamma_{M,n}, \quad \sigma_1 \in \mathcal{S}_{[n]}(\boldsymbol{\gamma}),$$
$$\mathcal{S}_{[n]}(\boldsymbol{\gamma}) \triangleq \left\{ \sigma \in \mathcal{S}_{[n]} \;\middle|\; \gamma\big(i, \sigma(i)\big) > 0, \ \forall i \in [n] \right\}.$$

We define

$$\boldsymbol{\gamma}_{\sigma_1} \triangleq \frac{1}{M-1} \cdot \left(M \cdot \boldsymbol{\gamma} - \boldsymbol{P}_{\sigma_1} \right) \in \Gamma_{M-1,n},$$

where

$$P_{\sigma_1}(i,j) \triangleq \begin{cases} 1 & j = \sigma_1(i) \\ 0 & \text{Otherwise} \end{cases}, \qquad i,j \in [n].$$

The matrix $\boldsymbol{\gamma}_{\sigma_1}$ is obtained by **"peeling off"** $\boldsymbol{P}_{\sigma_1}$ from $\boldsymbol{\gamma}$.