Double-Cover-Based Analysis of the Bethe Permanent of Non-negative Matrices

Kit Shing NG
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
nks020@ie.cuhk.edu.hk

Pascal O. Vontobel
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
pascal.vontobel@ieee.org

Abstract—The permanent of a non-negative matrix appears naturally in many information processing scenarios. Because of the intractability of the permanent beyond small matrices, various approximation techniques have been developed in the past. In this paper, we study the Bethe approximation of the permanent and add to the body of literature showing that this approximation is very well behaved in many respects. Our main technical tool are topological double covers of the normal factor graph whose partition function equals the permanent of interest, along with a transformation of these double covers.

I. INTRODUCTION

Let n be a positive integer. Recall that the permanent of a matrix $\mathbf{A} = (a_{i,j}) \in \mathbb{R}^{n \times n}$ is defined to be (see, e.g., [1])

$$\operatorname{perm}(\mathbf{A}) \triangleq \sum_{\sigma \in \mathcal{S}_n} \prod_{i \in [n]} a_{i,\sigma(i)}, \tag{1}$$

where $\mathcal{S}_n$ is the symmetric group of degree n, i.e., the group of all the $n!$ permutations of $[n] \triangleq \{1, \ldots, n\}$. It is well known that exactly computing the permanent is suspected to be a hard problem in general (see, e.g., the discussion in [2]).

In this paper we focus on the particularly important special case when $\mathbf{A}$ is a non-negative matrix, i.e., when the entries of the matrix $\mathbf{A}$ take on non-negative values. Various approaches have been proposed to efficiently numerically approximate the permanent of such matrices (see, e.g., the discussion in [2]). One of these approximations is the so-called Bethe permanent $\operatorname{perm}_{\mathrm{B}}(\mathbf{A})$, which is based on the Bethe approximation from statistical physics [3] and is given as the solution of some optimization problem derived from $\mathbf{A}$.

In contrast to the original definition of $\operatorname{perm}_{\mathrm{B}}(\mathbf{A})$, which is given in terms of an optimization problem, one can, using the techniques that were developed in [4], give a combinatorial characterization of $\operatorname{perm}_{\mathrm{B}}(\mathbf{A})$. Namely,

$$\operatorname{perm}_{\mathrm{B}}(\mathbf{A}) = \limsup_{M \to \infty} \operatorname{perm}_{\mathrm{B},M}(\mathbf{A}), \tag{2}$$

$$\operatorname{perm}_{\mathrm{B},M}(\mathbf{A}) \triangleq \sqrt[M]{\left\langle \operatorname{perm}(\mathbf{A}^{\uparrow\tilde{\mathbf{P}}}) \right\rangle_{\tilde{\mathbf{P}} \in \tilde{\Phi}_M}}. \tag{3}$$

Here the expression under the root sign represents the (arithmetic) average of $\operatorname{perm}(\mathbf{A}^{\uparrow\tilde{\mathbf{P}}})$ over all M-covers of $\mathbf{A}$, $M \geq 1$. (See the upcoming sections for the technical details.)

Note that we can write

$$\underbrace{\frac{\operatorname{perm}(\mathbf{A})}{\operatorname{perm}_{\mathrm{B}}(\mathbf{A})}}_{①} = \underbrace{\frac{\operatorname{perm}(\mathbf{A})}{\operatorname{perm}_{\mathrm{B},2}(\mathbf{A})}}_{②} \cdot \underbrace{\frac{\operatorname{perm}_{\mathrm{B},2}(\mathbf{A})}{\operatorname{perm}_{\mathrm{B}}(\mathbf{A})}}_{③}. \tag{4}$$

Numerically computing the ratios in (4) for various choices of matrices $\mathbf{A}$ shows that a significant contribution to the ratio ① comes from the ratio ②. Therefore, understanding the ratio ② can give useful insights to understanding the ratio ①. The central topic of this paper is to make this observation mathematically more precise.

Let $\mathbf{A}$ be an arbitrary non-negative matrix of size $n \times n$. The key technical result of this paper is that

$$\frac{\operatorname{perm}_{\mathrm{B},2}(\mathbf{A})}{\operatorname{perm}(\mathbf{A})} = \sqrt{\sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} p(\sigma_1) \cdot p(\sigma_2) \cdot 2^{-c(\sigma_1,\sigma_2)}}, \tag{5}$$

where $p(\sigma) \triangleq \left(\prod_{i \in [n]} a_{i,\sigma(i)} \right) / \operatorname{perm}(\mathbf{A})$ is the probability mass function on $\mathcal{S}_n$ induced by $\mathbf{A}$ and where $c(\sigma_1, \sigma_2)$ is the number of cycles of length larger than one in the cycle notation expression of the permutation $\sigma_1 \circ \sigma_2^{-1}$.

This result is then leveraged to make the following analytical statements related to (4):

- If $\mathbf{A} = \mathbf{1}_{n \times n}$, i.e., the all-one matrix of size $n \times n$, then[1]

$$\frac{\operatorname{perm}(\mathbf{A})}{\operatorname{perm}_{\mathrm{B}}(\mathbf{A})} \sim \sqrt{\frac{2\pi n}{\mathrm{e}}}, \quad \frac{\operatorname{perm}(\mathbf{A})}{\operatorname{perm}_{\mathrm{B},2}(\mathbf{A})} \sim \sqrt[4]{\frac{\pi n}{\mathrm{e}}}. \tag{6}$$

  Observe that, up to a factor $\sqrt{2}$, the ratios ② and ③ in (4) are the same for this matrix!

- If $\mathbf{A}$ is a random matrix of size $n \times n$ whose entries are i.i.d. according to some distribution with support over the non-negative reals, then

$$\gamma_{\mathrm{B},2}(n) \triangleq \frac{\sqrt{\mathbb{E}\big[\operatorname{perm}(\mathbf{A})^2\big]}}{\sqrt{\mathbb{E}\big[\operatorname{perm}_{\mathrm{B},2}(\mathbf{A})^2\big]}} \sim \gamma_{\mathrm{B},2}'(n) \triangleq \sqrt[4]{\frac{\pi n}{\mathrm{e}}}. \tag{7}$$

  Interestingly, although the numerator and the denominator on the left-hand side of the above expression both *depend* on the chosen distribution, the right-hand side of the above expression is *independent* of this distribution!

---

[1]The notation $a(n) \sim b(n)$ stands for $\lim_{n \to \infty} \frac{a(n)}{b(n)} = 1$.
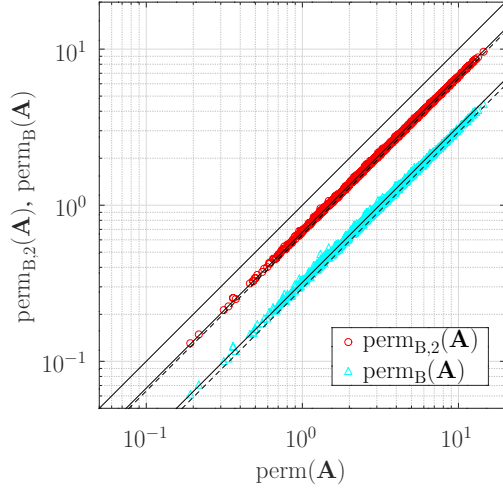
Fig. 1: Simulation results discussed in Section I.

Empirically, it actually appears that much stronger results than in (7) can potentially be proven. Namely, the plot in Fig. 1 shows the following:

- For $n = 5$, we randomly generated 1000 matrices $\mathbf{A}$ of size $n \times n$, where the entries are i.i.d. according to the uniform distribution in the interval $[0, 1]$. For each matrix $\mathbf{A}$, we plotted a red circle at the location $\big(\mathrm{perm}(\mathbf{A}), \mathrm{perm}_{\mathrm{B},2}(\mathbf{A})\big)$ and a cyan triangle at the location $\big(\mathrm{perm}(\mathbf{A}), \mathrm{perm}_{\mathrm{B}}(\mathbf{A})\big)$.
- Let $\gamma_{\mathrm{B},2}(n)$ be the quantity defined in (7). The solid black line going through the cluster of red circles is the location of pairs $\big(\mathrm{perm}(\mathbf{A}), \mathrm{perm}_{\mathrm{B},2}(\mathbf{A})\big)$ for which $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B},2}(\mathbf{A}) = \gamma_{\mathrm{B},2}(n)$.
- Let $\gamma'_{\mathrm{B},2}(n)$ be the quantity defined in (7). The dashed black line going through the cluster of red circles is the location of pairs $\big(\mathrm{perm}(\mathbf{A}), \mathrm{perm}_{\mathrm{B},2}(\mathbf{A})\big)$ for which $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B},2}(\mathbf{A}) = \gamma'_{\mathrm{B},2}(n)$.
- The solid and dashed black lines going through the cluster of cyan triangles are similar to the solid and dashed red lines going through the cluster of red circles, respectively. They are based on our conjecture that the value of the ratio $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B}}(\mathbf{A})$ follows closely the value of the ratio $\mathrm{perm}(\mathbf{1}_{n \times n})/\mathrm{perm}_{\mathrm{B}}(\mathbf{1}_{n \times n})$.

Motivated by Fig. 1, we leave it to future research to make stronger analytical statements than in (7) w.r.t. the distribution of the ratio $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})$, about the distribution of the ratio $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B}}(\mathbf{A})$, etc. Such results will contribute toward rigorously justifying the empirical observation of the usefulness of the Bethe approximation of the permanent (see, e.g., [5]–[12]). Some of these papers replace the (usually intractable) optimization problem $\arg\max_{\mathbf{A} \in \mathcal{A}} \mathrm{perm}(\mathbf{A})$ by the more tractable optimization problem $\arg\max_{\mathbf{A} \in \mathcal{A}} \mathrm{perm}_{\mathrm{B}}(\mathbf{A})$, where $\mathcal{A}$ is a set of matrices of interest. For this approximation to work well, the value of the ratio $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B}}(\mathbf{A})$ is irrelevant as long as it is nearly the same for all matrices $\mathbf{A} \in \mathcal{A}$.

It is worthwhile to mention that various results for the ratio $\mathrm{perm}(\mathbf{A})/\mathrm{perm}_{\mathrm{B}}(\mathbf{A})$ have been developed in the past. In
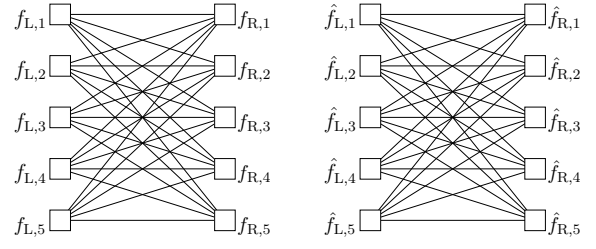

Fig. 2: NFGs used in Sections II and III. (Here, $n = 5$.)

particular, lower and upper bounds on this ratio can be found in [13]–[16] for arbitrary non-negative matrices $\mathbf{A}$ of size $n \times n$ and in [17] for non-negative matrices $\mathbf{A}$ of size $n \times n$ and of a given non-negative matrix rank. While rather non-trivial, these results are not strong enough / not suitable to derive the results mentioned above.

This paper is structured as follows: in Section II we discuss a family of normal factor graphs whose partition function equals the permanent of a non-negative matrix $\mathbf{A}$. Afterwards, we apply a technique from [18] for analyzing $\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})$: in Section III for general non-negative matrices, in Section IV for all-one matrices, and in Section V for random non-negative matrices with i.i.d. entries. We conclude the paper in Section VI. Finally, we have collected many of the proofs in the appendices.

The main tool of this paper are topological graph covers. Note that graph covers have also been used in other contexts toward understanding and quantifying the Bethe approximation of various quantities of interest. For example, graph covers were used to analyze so-called log-supermodular graphical models [18], [19] and weighted homomorphism counting problems over bipartite graphs [20].

II. NORMAL FACTOR GRAPH REPRESENTATION

We assume that the reader is familiar with the basics of factor graphs [21]–[23], in particular with a variant of factor graphs called normal factor graphs (NFGs). Recall that an NFG is a graph consisting of vertices (called function nodes) and edges, where so-called local functions are associated with vertices and where variables are associated with edges. The local functions are such that their arguments are only variables associated with edges incident on the corresponding factor node. The global function is then defined to be the product of the local functions, and the partition function is defined to be the sum of the global function over all possible assignments to the variables associated with the edges.

Let $\mathbf{A} = (a_{i,j})$ be an arbitrary non-negative matrix of size $n \times n$. There are various ways of defining an NFG such that its partition function equals $\mathrm{perm}(\mathbf{A})$. In this paper, we use the NFG $\mathsf{N}(\mathbf{A})$ in Fig. 2 (left), which is the same as in [2]:

- The NFG $\mathsf{N}(\mathbf{A})$ is based on a complete bipartite graph with two times n vertices.
- For every $i, j \in [n]$, let the variable associated with the edge connecting $f_{\mathrm{L},i}$ with $f_{\mathrm{R},j}$ be called $x_{i,j}$ and take value in the set $\mathcal{X} \triangleq \{0, 1\}$.

- For every $i \in [n]$, let

$$f_{\mathrm{L},i}(x_{i,1}, \ldots, x_{i,n}) \triangleq \begin{cases} \sqrt{a_{i,j}} & \begin{aligned} &\exists j \in [n] \text{ s.t. } x_{i,j} = 1; \\ &x_{i,j'} = 0, \ \forall j' \in [n] \setminus \{j\} \end{aligned} \\ \\ 0 & \text{(otherwise)} \end{cases}$$

For every $j \in [n]$, let

$$f_{\mathrm{R},j}(x_{1,j}, \ldots, x_{n,j}) \triangleq \begin{cases} \sqrt{a_{i,j}} & \begin{aligned} &\exists i \in [n] \text{ s.t. } x_{i,j} = 1; \\ &x_{i',j} = 0, \ \forall i' \in [n] \setminus \{i\} \end{aligned} \\ \\ 0 & \text{(otherwise)} \end{cases}$$

- The global function is defined to be

$$g(x_{1,1}, x_{1,2}, \ldots, x_{n,n}) \triangleq \left(\prod_{i \in [n]} f_{\mathrm{L},i}(x_{i,1}, \ldots, x_{i,n}) \right) \\ \cdot \left(\prod_{j \in [n]} f_{\mathrm{R},j}(x_{1,j}, \ldots, x_{n,j}) \right).$$

- The partition function (or partition sum) is defined to be

$$Z(\mathsf{N}) \triangleq \sum_{x_{1,1}, x_{1,2}, \ldots, x_{n,n}} g(x_{1,1}, x_{1,2}, \ldots, x_{n,n}).$$

One can verify the following:

- $g(x_{1,1}, x_{1,2}, \ldots, x_{n,n}) = \prod_{i \in [n]} a_{i,\sigma(i)}$ if there exists a permutation $\sigma \in \mathcal{S}_n$ such that for all $i, j \in [n]$ either $x_{i,j} = 1$ if $j = \sigma(i)$ or $x_{i,j} = 0$ otherwise.
- $g(x_{1,1}, x_{1,2}, \ldots, x_{n,n}) = 0$ if there exists no such permutation $\sigma \in \mathcal{S}_n$.
- $Z(\mathsf{N}) = \mathrm{perm}(\mathbf{A})$.

For NFGs whose local functions take on non-negative values, the paper [3] introduced the Bethe approximation $Z_{\mathrm{B}}(\mathsf{N})$ of the partition function $Z(\mathsf{N})$ as the solution of some optimization problem derived from N. In [4] it was then shown that $Z_{\mathrm{B}}(\mathsf{N})$ has the following combinatorial characterization:

$$Z_{\mathrm{B}}(\mathsf{N}) = \limsup_{M \to \infty} Z_{\mathrm{B},M}(\mathsf{N}), \tag{8}$$

$$Z_{\mathrm{B},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\tilde{\mathsf{N}}) \right\rangle_{\tilde{\mathsf{N}} \in \tilde{\mathcal{N}}_M}}, \tag{9}$$

where the expression under the root sign represents the (arithmetic) average of $Z(\tilde{\mathsf{N}})$ over all M-covers $\tilde{\mathsf{N}}$ of N, $M \geq 1$. For the details of the definition of (topological) graph covers, we refer to [4].

Interestingly, in the context of the NFG $\mathsf{N}(\mathbf{A})$, the expressions in (9) turns into the expression in (3), where for positive integers M we have defined

$$\mathbf{A}^{\uparrow \tilde{\mathbf{P}}} \triangleq \begin{pmatrix} a_{1,1} \tilde{\mathbf{P}}^{(1,1)} & \cdots & a_{1,n} \tilde{\mathbf{P}}^{(1,n)} \\ \vdots & & \vdots \\ a_{n,1} \tilde{\mathbf{P}}^{(n,1)} & \cdots & a_{n,n} \tilde{\mathbf{P}}^{(n,n)} \end{pmatrix},$$

$$\tilde{\Phi}_M \triangleq \left\{ \tilde{\mathbf{P}} = \left\{ \tilde{\mathbf{P}}^{(i,j)} \right\}_{i \in [n], j \in [n]} \ \middle| \ \tilde{\mathbf{P}}^{(i,j)} \in \mathcal{P}_{M \times M} \right\},$$

$$\mathcal{P}_{M \times M} \triangleq (\text{set of permutation matrices of size } M \times M).$$

Note that the matrix $\mathbf{A}^{\uparrow \tilde{\mathbf{P}}}$ has size $(Mn) \times (Mn)$. (For more details, see the discussion in [4, Section VI].)

In the rest of this paper, we will analyze double covers of $\mathsf{N}(\mathbf{A})$, i.e., graph covers of $\mathsf{N}(\mathbf{A})$ for $M = 2$.

III. DOUBLE-COVER-BASED ANALYSIS: GENERAL CASE

The paper [18] introduced a technique for analyzing double covers of an arbitrary NFG N. We refer the interested reader to [18] for all the technical details. Here we just state the main result when applied to the NFG $\mathsf{N}(\mathbf{A})$.

Namely, using [18, Theorem 4], one obtains the following result. (Note that on the right-hand side of (10) only a single NFG appears, which is in contrast to the right-hand side of (9) that averages over multiple NFGs.)

Proposition 1. *Let $\mathbf{A}$ be a non-negative matrix of size $n \times n$. It holds that*

$$\mathrm{perm}_{\mathrm{B},2}(\mathbf{A}) = \sqrt{Z(\hat{\mathsf{N}}(\mathbf{A}))}, \tag{10}$$

where the NFG $\hat{\mathsf{N}}(\mathbf{A})$ is defined as follows (see Fig. 2 (right)):

- *The NFG $\hat{\mathsf{N}}(\mathbf{A})$ is based on a complete bipartite graph with two times n vertices, i.e., the same graph underlying $\mathsf{N}(\mathbf{A})$.*
- *For every $i, j \in [n]$, let the variable associated with the edge connecting $\hat{f}_{\mathrm{L},i}$ with $\hat{f}_{\mathrm{R},j}$ be called $\hat{x}_{i,j}$ and take value in the set*

$$\hat{\mathcal{X}} \triangleq \mathcal{X} \times \mathcal{X} = \big\{ (0,0), \ (0,1), \ (1,0), \ (1,1) \big\}.$$

- *For every $i \in [n]$, let $\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) \triangleq$*

$$\begin{cases} a_{i,j} & \begin{aligned} &\exists j \in [n] \text{ s.t. } \hat{x}_{i,j} = (1,1); \\ &\hat{x}_{i,j'} = (0,0), \ \forall j' \in [n] \setminus \{j\} \end{aligned} \\ \\ \sqrt{a_{i,j} a_{i,j'}} & \begin{aligned} &\exists j, j' \in [n], \ j \neq j', \text{ s.t.} \\ &\hat{x}_{i,j} = (0,1), \ \hat{x}_{i,j'} = (0,1); \\ &\hat{x}_{i,j''} = (0,0), \ \forall j'' \in [n] \setminus \{j, j'\} \end{aligned} \\ \\ 0 & \textit{(otherwise)} \end{cases}$$

- *For every $j \in [n]$, let $\hat{f}_{\mathrm{R},j}(\hat{x}_{1,j}, \ldots, \hat{x}_{n,j}) \triangleq$*

$$\begin{cases} a_{i,j} & \begin{aligned} &\exists i \in [n] \text{ s.t. } \hat{x}_{i,j} = (1,1); \\ &\hat{x}_{i',j} = (0,0), \ \forall i' \in [n] \setminus \{i\} \end{aligned} \\ \\ \sqrt{a_{i,j} a_{i',j}} & \begin{aligned} &\exists i, i' \in [n], \ i \neq i', \text{ s.t.} \\ &\hat{x}_{i,j} = (0,1), \ \hat{x}_{i',j} = (0,1); \\ &\hat{x}_{i'',j} = (0,0), \ \forall i'' \in [n] \setminus \{i, i'\} \end{aligned} \\ \\ 0 & \textit{(otherwise)} \end{cases}$$

- *The global function is defined to be*

$$\hat{g}(\hat{x}_{1,1}, \hat{x}_{1,2}, \ldots, \hat{x}_{n,n}) \triangleq \left(\prod_{i \in [n]} \hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) \right) \\ \cdot \left(\prod_{j \in [n]} \hat{f}_{\mathrm{R},j}(\hat{x}_{1,j}, \ldots, \hat{x}_{n,j}) \right).$$

- *The partition function (or partition sum) is defined to be*

$$Z(\hat{\mathsf{N}}) \triangleq \sum_{\hat{x}_{1,1},\hat{x}_{1,2},\ldots,\hat{x}_{n,n}} \hat{g}(\hat{x}_{1,1}, \hat{x}_{1,2}, \ldots, \hat{x}_{n,n}).$$

Proof. See Appendix A □

Using a mapping from $\mathcal{S}_n \times \mathcal{S}_n$ to the set of valid configurations of $Z(\hat{\mathsf{N}})$, i.e., the set of configurations $(\hat{x}_{1,1}, \hat{x}_{1,2}, \ldots, \hat{x}_{n,n})$ that yield $\hat{g}(\hat{x}_{1,1}, \hat{x}_{1,2}, \ldots, \hat{x}_{n,n}) \neq 0$, allows us to reformulate Proposition 1 as follows.

Proposition 2. *Let $\mathbf{A}$ be a non-negative matrix of size $n \times n$. It holds that*

$$\frac{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}{\mathrm{perm}(\mathbf{A})} = \sqrt{\sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} p(\sigma_1) \cdot p(\sigma_2) \cdot 2^{-c(\sigma_1,\sigma_2)}}, \quad (11)$$

where

$$p(\sigma) \triangleq \frac{\prod\limits_{i \in [n]} a_{i,\sigma(i)}}{\mathrm{perm}(\mathbf{A})} \quad (12)$$

is the probability mass function on $\mathcal{S}_n$ induced by $\mathbf{A}$ and where $c(\sigma_1, \sigma_2)$ is the number of cycles of length larger than one in the cycle notation expression of the permutation $\sigma_1 \circ \sigma_2^{-1}$, or, equivalently, where $c(\sigma_1, \sigma_2)$ is the number of orbits of length larger than one of the permutation $\sigma_1 \circ \sigma_2^{-1}$.

Proof. See Appendix B. □

In the following, we evaluate the expression in Proposition 2 for different setups.

IV. DOUBLE-COVER-BASED ANALYSIS: ALL-ONE MATRIX

We have the following result for the all-one matrix of size $n \times n$. (Note that (13) was already proven in [4, Lemma 48].)

Theorem 1. *Let $\mathbf{A} \triangleq \mathbf{1}_{n \times n}$, i.e., the all-one matrix of size $n \times n$. It holds that*

$$\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B}}(\mathbf{A})} \sim \sqrt{\frac{2\pi n}{\mathrm{e}}}, \quad (13)$$

$$\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})} \sim \sqrt[4]{\frac{\pi n}{\mathrm{e}}}, \quad (14)$$

$$\frac{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B}}(\mathbf{A})} \sim \sqrt{2} \cdot \sqrt[4]{\frac{\pi n}{\mathrm{e}}}. \quad (15)$$

We see that understanding the ratio $\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}$ goes a long way toward understanding the ratio $\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B}}(\mathbf{A})}$.

Proof. A sketch of the proof of this theorem is given below. For all the details of the proof, see Appendix C. □

Let us sketch the proof of Theorem 1. First, thanks to the special structure of the all-one matrix, Proposition 2 simplifies as follows.

Corollary 1 (of Proposition 2). *Let $\mathbf{A} = \mathbf{1}_{n \times n}$. It holds that*

$$\frac{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}{\mathrm{perm}(\mathbf{A})} = \sqrt{\frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} 2^{-c(\sigma)}}, \quad (16)$$

where $c(\sigma)$ is the number of cycles of length larger than one in the cycle notation expression of the permutation σ, or, equivalently, where $c(\sigma)$ is the number of orbits of length larger than one of the permutation σ.

Proof. This follows from $p(\sigma) = 1/n!$ for all $\sigma \in \mathcal{S}_n$ and simplifying the summation. □

Using results for the cycle index of the symmetric group (see, e.g., [24], [25]), one can analyze the expression

$$Z_n \triangleq \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} 2^{-c(\sigma)} = \frac{1}{|\mathcal{S}_n|} \sum_{\sigma \in \mathcal{S}_n} 2^{-c(\sigma)}$$

for $n \in \mathbb{N}$. Indeed, one obtains the following results.

- Let $Z_0 \triangleq 1$. Then for $n \in \mathbb{N}$ it holds that

$$Z_n = \frac{1}{n} \cdot \left(Z_{n-1} + \frac{1}{2} \cdot \sum_{\ell=2}^{n} Z_{n-\ell} \right).$$

- Analyzing the expression $n \cdot Z_n - (n-1) \cdot Z_{n-1}$ leads to the following simplified expression for $n \geq 2$:

$$Z_n = Z_{n-1} - \frac{Z_{n-2}}{2n}.$$

This expression is convenient for numerically evaluating Z_n, however, it does not appear that there are closed-form expressions for Z_n.
- Nevertheless, using mathematical induction, one can obtain the following lower and upper bounds which hold for all positive integers n:

$$\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{n}} \leq Z_n \leq \frac{3}{2\sqrt{2}} \cdot \frac{1}{\sqrt{n}}.$$

- In order to obtain an even more precise characterization of Z_n, one can use

$$Z_n = \frac{1}{n!} \cdot B_n \left(0! \cdot 1, 1! \cdot \frac{1}{2}, \ldots, (n-1)! \cdot \frac{1}{2} \right),$$

where B_n denotes the n-th complete exponential Bell polynomial (see, e.g., [26]). Then, carefully bounding the above expression, one obtains

$$Z_n \sim \sqrt{\frac{\mathrm{e}}{\pi n}}.$$

This expression, via Corollary 1, immediately yields (14).
- Finally, Eq. 15 is then obtained by combining (13) and (14).

V. Double-Cover-Based Analysis: I.I.D. Matrix

In this section we consider matrices satisfying the following assumption.

Assumption 1. *Let $\mathbf{A}$ be a random matrix of size $n \times n$ whose entries are i.i.d. according to some distribution with support over the non-negative reals, first moment μ_1, second moment μ_2, and, consequently, variance $\mu_2 - \mu_1^2$.*

Theorem 2. *Given Assumption 1, it holds that*

$$\frac{\sqrt{\mathbb{E}[\mathrm{perm}(\mathbf{A})^2]}}{\sqrt{\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2]}} \sim \sqrt[4]{\frac{\pi n}{\mathrm{e}}}. \tag{17}$$

Interestingly, although $\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2]$ and $\mathbb{E}[\mathrm{perm}(\mathbf{A})^2]$ both *depend* on the chosen distribution (see the calculations below), the right-hand side of (17) is *independent* of this distribution!

Proof. A sketch of the proof of this theorem is given below. For all the details of the proof, see Appendix D. $\square$

Let us sketch the proof of Theorem 2. We start by evaluating $\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2]$. From Proposition 2 we obtain the following corollary.

Corollary 2 (of Proposition 2)**.** *Given Assumption 1, it holds that*

$$\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2] = n! \cdot \sum_{\sigma \in \mathcal{S}} \mu_2^{c_1(\sigma)} \cdot \left(\prod_{\ell \geq 2} \mu_1^{2 \cdot \ell \cdot c_\ell(\sigma)}\right) \cdot 2^{-c(\sigma)}$$

where for $\ell \geq 1$ we define $c_\ell(\sigma)$ to be the number of cycles of length ℓ in the cycle notation expression of the permutation σ. Note that $\sum_{\ell \geq 1} \ell \cdot c_\ell = n$ and $\sum_{\ell \geq 2} c_\ell(\sigma) = c(\sigma)$.

Proof. From Proposition 2 it follows that

$$\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2]$$
$$= \sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} \mathbb{E}\left[\left(\prod_{i_1 \in [n]} a_{i_1,\sigma_1(i_1)}\right) \cdot \left(\prod_{i_2 \in [n]} a_{i_2,\sigma_2(i_2)}\right)\right] \cdot 2^{-c(\sigma_1,\sigma_2)}$$
$$= \sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} \mu_2^{c_1(\sigma_1,\sigma_2)} \cdot \left(\prod_{\ell \geq 2} \mu_1^{2 \cdot \ell \cdot c_\ell(\sigma_1,\sigma_2)}\right) \cdot 2^{-c(\sigma_1,\sigma_2)}$$
$$= n! \cdot \sum_{\sigma \in \mathcal{S}} \mu_2^{c_1(\sigma)} \cdot \left(\prod_{\ell \geq 2} \mu_1^{2 \cdot \ell \cdot c_\ell(\sigma)}\right) \cdot 2^{-c(\sigma)},$$

where for $\ell \geq 1$ we define $c_\ell(\sigma_1,\sigma_2)$ to be the number of cycles of length ℓ in the cycle notation expression of the permutation $\sigma_1 \circ \sigma_2^{-1}$. $\square$

Using results for the cycle index of the symmetric group, one can analyze the expression

$$Z_n \triangleq \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \mu_2^{c_1(\sigma)} \cdot \left(\prod_{\ell \geq 2} \mu_1^{2 \cdot \ell \cdot c_\ell(\sigma)}\right) \cdot 2^{-c(\sigma)}$$
$$= \frac{1}{|\mathcal{S}_n|} \sum_{\sigma \in \mathcal{S}_n} \mu_2^{c_1(\sigma)} \cdot \left(\prod_{\ell \geq 2} \mu_1^{2 \cdot \ell \cdot c_\ell(\sigma)}\right) \cdot 2^{-c(\sigma)},$$

for positive integers n. In order to evaluate Z_n, one can use

$$Z_n = \frac{1}{n!} \cdot B_n\left(0! \cdot \mu_2, 1! \cdot \frac{1}{2}(\mu_1^2)^2, \ldots, (n-1)! \cdot \frac{1}{2}(\mu_1^2)^n\right),$$

where B_n denotes the n-th complete exponential Bell polynomial. Then, carefully bounding the above expression, one obtains

$$Z_n \sim \mu_1^{2n} \cdot \mathrm{e}^{(\mu_2/\mu_1^2)-1} \cdot \sqrt{\frac{\mathrm{e}}{\pi n}},$$

The combination of the above results yields

$$\mathbb{E}[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2] \sim (n!)^2 \cdot \mu_1^{2n} \cdot \mathrm{e}^{(\mu_2/\mu_1^2)-1} \cdot \sqrt{\frac{\mathrm{e}}{\pi n}}.$$

Turning our attention now to $\mathbb{E}[\mathrm{perm}(\mathbf{A})^2]$, calculations similar to the above calculations yield

$$\mathbb{E}[\mathrm{perm}(\mathbf{A})^2] = n! \cdot \sum_{\sigma \in \mathcal{S}} \mu_2^{c_1(\sigma)} \cdot \left(\prod_{\ell \geq 2} (\mu_1^2)^{\ell c_\ell(\sigma)}\right)$$
$$= \Psi_n(\mu_2, \mu_1^2, 1)$$
$$\sim (n!)^2 \cdot \mu_1^{2n} \cdot \mathrm{e}^{(\mu_2/\mu_1^2)-1}.$$

Finally, combining the above results, we get (17).

VI. Conclusion

We conclude this paper with a few remarks:

- In this paper we have studied the ratios $\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B}}(\mathbf{A})}$ and $\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}$. While it is more desirable to characterize the former, the latter seems to be more tractable. In particular, the latter can be used to obtain qualitative *and* quantitative insights into the former. We leave it to future research to strengthen the results that are presented in this paper w.r.t. these ratios.
- For the all-one matrices studied in Sections IV, some further investigations show that the summation in (5) is dominated by permutations $\sigma_1, \sigma_2 \in \mathcal{S}_n$ for which $c(\sigma_1, \sigma_2) = \Theta(\ln(n))$, ultimately leading to the result $\frac{\mathrm{perm}(\mathbf{A})}{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})} = \Theta(\sqrt[4]{n})$.
  We expect a similar behavior for the matrices of the setup in Section V. More precisely, we conjecture that with high probability the matrix $\mathbf{A}$ is such that the summation in (5) is dominated by permutations $\sigma_1, \sigma_2 \in \mathcal{S}_n$ for which $c(\sigma_1, \sigma_2) = \Theta(\ln(n))$.
- The proofs in Appendices C and D can be generalized and unified. For details, see Appendix E.

Proposition 1 is obtained by applying [18, Theorem 4] to $N(A)$. (Because the following derivations rely heavily on the technique presented in [18], the reader is advised to first read that paper.)

Recall the definition of $\mathcal{X} \triangleq \{0, 1\}$ in Section II and the definition of $\hat{\mathcal{X}} \triangleq \mathcal{X} \times \mathcal{X} = \{(0,0), (0,1), (1,0), (1,1)\}$ in Section III. In this appendix, we will also use the set

$$\tilde{\mathcal{X}} \triangleq \mathcal{X} \times \mathcal{X} = \{(0,0), (0,1), (1,0), (1,1)\}.$$

For all these sets, their entries are considered to be ordered as shown above. Moreover, if M is some matrix, then M^{T} is its transpose. If M_1 and M_2 are two matrices of arbitrary size, then $M_1 \otimes M_2$ is the Kronecker product of M_1 and M_2.

We start by recalling some notations from [18].

- Let $\mathcal{A} \triangleq \{a_1, a_2, \ldots, a_{|\mathcal{A}|}\}$ and $\mathcal{B} \triangleq \{b_1, b_2, \ldots, b_{|\mathcal{B}|}\}$ be some finite sets whose entries are considered to be ordered as shown here. We associate the following matrix of size $|\mathcal{A}| \times |\mathcal{B}|$ with a function $f : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$:

$$\mathbf{T}_f \triangleq \begin{pmatrix} f(a_1, b_1) & f(a_1, b_2), & \cdots & f(a_1, b_{|\mathcal{B}|}) \\ f(a_2, b_1) & f(a_2, b_2), & \cdots & f(a_2, b_{|\mathcal{B}|}) \\ \vdots & \vdots & \ddots & \vdots \\ f(a_{|\mathcal{A}|}, b_1) & f(a_{|\mathcal{A}|}, b_2) & \cdots & f(a_{|\mathcal{A}|}, b_{|\mathcal{B}|}) \end{pmatrix}.$$

Similarly, if $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are some finite sets, then we associate an array $\mathbf{T}_f$ of size $|\mathcal{A}| \times |\mathcal{B}| \times |\mathcal{C}|$ with a function $f : \mathcal{A} \times \mathcal{B} \times \mathcal{C} \to \mathbb{R}$, etc.

- The function[2]

$$\Phi : \tilde{\mathcal{X}} \times \hat{\mathcal{X}} \to \mathbb{R}$$

is defined as in [18, Section III]. Namely, Φ is such that

$$\mathbf{T}_\Phi \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that $\mathbf{T}_\Phi^{\mathsf{T}} = \mathbf{T}_\Phi$ and $\mathbf{T}_\Phi^{-1} = \mathbf{T}_\Phi$.

- As in [18, Section III], we define the matrices

$$\tilde{\mathbf{E}}_{\text{nocross}} \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\tilde{\mathbf{E}}_{\text{cross}} \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix},$$

[2]Note that, although Φ (which is introduced here) and $\tilde{\Phi}_M$ (which appears in Sections I and II) are both used to describe graph covers, they are not directly related.

whose rows and columns are labeled by the elements of $\tilde{\mathcal{X}}$. Based on these matrices, we define the matrices

$$\hat{\mathbf{E}}_{\text{nocross}} \triangleq \mathbf{T}_\Phi^{\mathsf{T}} \cdot \tilde{\mathbf{E}}_{\text{nocross}} \cdot \mathbf{T}_\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\hat{\mathbf{E}}_{\text{cross}} \triangleq \mathbf{T}_\Phi^{\mathsf{T}} \cdot \tilde{\mathbf{E}}_{\text{cross}} \cdot \mathbf{T}_\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix},$$

and

$$\mathbf{T}_{\hat{E}_e} \triangleq \frac{1}{2} \cdot \hat{\mathbf{E}}_{\text{nocross}} + \frac{1}{2} \cdot \hat{\mathbf{E}}_{\text{cross}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{pmatrix}.$$

whose rows and columns are labeled by the elements of $\hat{\mathcal{X}}$.

In order to apply the technique from [18], we need, in a first step, to compute the functions $\tilde{f}_{L,i}$ and $\hat{f}_{L,i}$ derived from $f_{L,i}$, $i \in [n]$, and the functions $\tilde{f}_{R,j}$ and $\hat{f}_{R,j}$ derived from $f_{R,j}$, $j \in [n]$. Here we will only find $\tilde{f}_{L,i}$ and $\hat{f}_{L,i}$, $i \in [n]$, because $\tilde{f}_{R,j}$ and $\hat{f}_{R,j}$, $j \in [n]$, are obtained in an analogous manner.

Fix some positive integer n and some $i \in [n]$. According to [18], $\tilde{f}_{L,i}$ and $\hat{f}_{L,i}$ are defined as, respectively,

$$\tilde{f}_{L,i} : \tilde{\mathcal{X}}^n \to \mathbb{R}$$
$$(\tilde{x}_{i,1}, \ldots, \tilde{x}_{i,n}) \mapsto f(x_{i,1}, \ldots, x_{i,n}) \cdot f(x'_{i,1}, \ldots, x'_{i,n}),$$

$$\hat{f}_{L,i} : \hat{\mathcal{X}}^n \to \mathbb{R}$$
$$(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) \mapsto \sum_{\tilde{x}_1, \ldots, \tilde{x}_n \in \tilde{\mathcal{X}}} \tilde{f}_{L,i}(\tilde{x}_{i,1}, \ldots, \tilde{x}_{i,n})$$
$$\cdot \prod_{j=1}^{n} \Phi(\tilde{x}_{i,j}, \hat{x}_{i,j}),$$

where we have used $\tilde{x}_{i,j} \triangleq (x_{i,j}, x'_{i,j}) \in \tilde{\mathcal{X}}$, $j \in [n]$, in the definition of $\tilde{f}_{L,i}$.

In the following, we want to find the explicit expression for $\hat{f}_{L,i}$. However, before tackling the case of general n, we will first consider the cases $n = 2$ and $n = 3$.

A. Function $\hat{f}_{L,i}$: case $n = 2$

Let $n = 2$ and fix some $i \in [n]$. The matrix associated with the function $f_{L,i}$ turns out to be

$$\mathbf{T}_{f_{L,i}} \triangleq \begin{pmatrix} 0 & \sqrt{a_{i,2}} \\ \sqrt{a_{i,1}} & 0 \end{pmatrix}.$$

Reusing some of the calculations in [18, Section IV], the matrices associated with $\tilde{f}_{L,i}$ and $\hat{f}_{L,i}$ are, respectively,

$$\mathbf{T}_{\tilde{f}_{L,i}} \triangleq \mathbf{T}_{f_{L,i}} \otimes \mathbf{T}_{f_{L,i}} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & a_{i,2} \\ \hline 0 & 0 & \sqrt{a_{i,2}a_{i,1}} & 0 \\ 0 & \sqrt{a_{i,1}a_{i,2}} & 0 & 0 \\ \hline a_{i,1} & 0 & 0 & 0 \end{array}\right),$$

$$\mathbf{T}_{\hat{f}_{L,i}} \triangleq \mathbf{T}_{\Phi} \cdot \mathbf{T}_{\tilde{f}_{L,i}} \cdot \mathbf{T}_{\Phi} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & a_{i,1} \\ \hline 0 & \sqrt{a_{i,1}a_{i,2}} & 0 & 0 \\ 0 & 0 & -\sqrt{a_{i,1}a_{i,2}} & 0 \\ \hline a_{i,2} & 0 & 0 & 0 \end{array}\right),$$

whose rows and columns are labeled by the elements of $\hat{\mathcal{X}}$.

Because all the entries of $\mathbf{T}_{\hat{E}_e}$ in row $(1,0)$ and all entries of $\mathbf{T}_{\hat{E}_e}$ in column $(1,0)$ are equal to zero, it turns out that $Z\big(\hat{N}(\mathbf{A})\big)$ is unchanged if the function $\hat{f}_{L,i}$ is replaced by the function $\hat{f}_{L,i}$ such that

$$\mathbf{T}_{\hat{f}_{L,i}} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & a_{i,1} \\ \hline 0 & \sqrt{a_{i,1}a_{i,2}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline a_{i,2} & 0 & 0 & 0 \end{array}\right).$$

The reader can verify that this matches the function $\hat{f}_{L,i}$ stated in Proposition 1 for the case $n = 2$.

B. Function $\hat{f}_{L,i}$: case $n = 3$

Consider now the case $n = 3$ and fix some $i \in [n]$. (Note that in the following we do not show the arrays associated with $f_{L,i}$ and $\tilde{f}_{L,i}$ and directly show the array associated with $\hat{f}_{L,i}$.)

Reusing some of the calculations in [18, Section IV], the array $\mathbf{T}_{\hat{f}_{L,i}}$ of size $4 \times 4 \times 4$ associated with $\hat{f}_{L,i}$ is[3]

$$\left(\begin{array}{c|ccc} 0 & 0 & 0 & a_{i,2} \\ \hline 0 & \sqrt{a_{i,1}a_{i,2}} & 0 & 0 \\ 0 & 0 & -\sqrt{a_{i,1}a_{i,2}} & 0 \\ \hline a_{i,1} & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} 0 & \sqrt{a_{i,2}a_{i,3}} & 0 & 0 \\ \hline \sqrt{a_{i,1}a_{i,3}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} 0 & 0 & -\sqrt{a_{i,2}a_{i,3}} & 0 \\ \hline 0 & 0 & 0 & 0 \\ -\sqrt{a_{i,1}a_{i,3}} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} a_{i,3} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

where all three dimensions are labeled by the elements of $\hat{\mathcal{X}}$.

Because all the entries of $\mathbf{T}_{\hat{E}_e}$ in row $(1,0)$ and all entries of $\mathbf{T}_{\hat{E}_e}$ in column $(1,0)$ are equal to zero, it turns out that

---

[3]Note that there were some typos in the paragraph before Lemma 6 in [18]. Namely, the second occurrence of $\det(\mathbf{T}_{f|a_1=0})$ should be $\det(\mathbf{T}_{f|a_2=0})$ and the third occurrence of $\det(\mathbf{T}_{f|a_1=0})$ should be $\det(\mathbf{T}_{f|a_2=1})$. Moreover, the last term in the expression for $\hat{f}(\hat{0},\hat{0},\hat{0})$ should be $t_{001}t_{110}$ instead of $t_{000}t_{110}$.

$Z\big(\hat{N}(\mathbf{A})\big)$ is unchanged if the function $\hat{f}_{L,i}$ is replaced by the function $\hat{f}_{L,i}$ such that the array $\mathbf{T}_{\hat{f}_{L,i}}$ of size $4 \times 4 \times 4$ associated with $\hat{f}_{L,i}$ is

$$\left(\begin{array}{c|ccc} 0 & 0 & 0 & a_{i,2} \\ \hline 0 & \sqrt{a_{i,1}a_{i,2}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline a_{i,1} & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} 0 & \sqrt{a_{i,2}a_{i,3}} & 0 & 0 \\ \hline \sqrt{a_{i,1}a_{i,3}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{c|ccc} a_{i,3} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right).$$

The reader can verify that this matches the function $\hat{f}_{L,i}$ stated in Proposition 1 for the case $n = 3$.

C. Function $\hat{f}_{L,i}$: general n

We now consider an arbitrary $n \in \mathbb{N}$ and fix some $i \in [n]$.

The following characterization of the function $\tilde{f}_{L,i}$ follows immediately from the properties of the function $f_{L,i}$. (Recall the definition of $f_{L,i}$ in Section II.)

- If there exists a $j \in [n]$ s.t. $\tilde{x}_{i,j} = (1,1)$ and $\tilde{x}_{i,j'} = (0,0)$ for all $j' \in [n] \setminus \{j\}$, then

$$\tilde{f}_{L,i}(\tilde{x}_{i,1}, \ldots, \tilde{x}_{i,n}) = \sqrt{a_{i,j}} \cdot \sqrt{a_{i,j}} = a_{i,j}. \quad (18)$$

- If there exist $j, j' \in [n]$, $j \neq j'$, s.t. $\tilde{x}_{i,j} = (0,1)$, $\tilde{x}_{i,j'} = (1,0)$ and $\tilde{x}_{i,j''} = (0,0)$ for all $j'' \in [n] \setminus \{j,j'\}$, then

$$\tilde{f}_{L,i}(\tilde{x}_{i,1}, \ldots, \tilde{x}_{i,n}) = \sqrt{a_{i,j} \cdot a_{i,j'}}. \quad (19)$$

- Otherwise,

$$\tilde{f}_{L,i}(\tilde{x}_{i,1}, \ldots, \tilde{x}_{i,n}) = 0. \quad (20)$$

Based on this, the function $\hat{f}_{L,i}$ can be characterized as follows.

- If there exists a $j \in [n]$ s.t. $\hat{x}_{i,j} = (1,1)$ and $\hat{x}_{i,j'} = (0,0)$ for all $j' \in [n] \setminus \{j\}$, then

$$\hat{f}_{L,i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = a_{i,j}.$$

This follows from (18) and

$$\Phi(\tilde{x}_{i,j}, \hat{x}_{i,j}) = \begin{cases} 1 & \text{if } \tilde{x}_{i,j} = (1,1) \\ 0 & \text{otherwise} \end{cases},$$

$$\Phi(\tilde{x}_{i,j'}, \hat{x}_{i,j'}) = \begin{cases} 1 & \text{if } \tilde{x}_{i,j'} = (0,0) \\ 0 & \text{otherwise} \end{cases}.$$

- If there exist $j, j' \in [n]$, $j \neq j'$, s.t. $\hat{x}_{i,j} = (0,1)$, $\hat{x}_{i,j'} = (0,1)$ and $\hat{x}_{i,j''} = (0,0)$ for all $j'' \in [n] \setminus \{j,j'\}$, then

$$\hat{f}_{L,i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = \sqrt{a_{i,j} \cdot a_{i,j'}}.$$

This follows from (19) and

$$\Phi(\tilde{x}_{i,j}, \hat{x}_{i,j}) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j} \in \{(0,1),\ (1,0)\} \\ 0 & \text{otherwise} \end{cases},$$

$$\Phi(\tilde{x}_{i,j'}, \hat{x}_{i,j'}) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j'} \in \{(0,1),\ (1,0)\} \\ 0 & \text{otherwise} \end{cases},$$

$$\Phi(\tilde{x}_{i,j''}, \hat{x}_{i,j''}) = \begin{cases} 1 & \text{if } \tilde{x}_{i,j''} = (0,0) \\ 0 & \text{otherwise} \end{cases}.$$

- If there exist $j, j' \in [n]$, $j \neq j'$, s.t. $\hat{x}_{i,j} = (1,0)$, $\hat{x}_{i,j'} = (1,0)$ and $\hat{x}_{i,j''} = (0,0)$ for all $j'' \in [n] \setminus \{j, j'\}$, then

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = -\sqrt{a_{i,j} \cdot a_{i,j'}}.$$

This follows from (19) and

$$\Phi(\tilde{x}_{i,j}, \hat{x}_{i,j}) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j} = (0,1) \\ -\frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j} = (1,0) \\ 0 & \text{otherwise} \end{cases},$$

$$\Phi(\tilde{x}_{i,j'}, \hat{x}_{i,j'}) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j'} = (0,1) \\ -\frac{1}{\sqrt{2}} & \text{if } \tilde{x}_{i,j'} = (1,0) \\ 0 & \text{otherwise} \end{cases},$$

$$\Phi(\tilde{x}_{i,j''}, \hat{x}_{i,j''}) = \begin{cases} 1 & \text{if } \tilde{x}_{i,j''} = (0,0) \\ 0 & \text{otherwise} \end{cases}.$$

- Otherwise,

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = 0.$$

This follows from (20) and/or from the exact cancellation of terms.

Finally, we can make a similar observation as in Sections A-A and A-B. Namely, because all the entries of $\mathbf{T}_{\hat{E}_e}$ in row $(1,0)$ and all entries of $\mathbf{T}_{\hat{E}_e}$ in column $(1,0)$ are equal to zero, it turns out that $Z\big(\hat{\mathsf{N}}(\mathbf{A})\big)$ is unchanged if the function $\hat{f}_{\mathrm{L},i}$ is replaced by the function $\hat{f}_{\mathrm{L},i}$ where $\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n})$ is set to 0 if there exists a $j \in [n]$ such that $\hat{x}_{i,j} = (1,0)$. We then obtain

- If there exists a $j \in [n]$ s.t. $\hat{x}_{i,j} = (1,1)$ and $\hat{x}_{i,j'} = (0,0)$ for all $j' \in [n] \setminus \{j\}$, then

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = a_{i,j}.$$

- If there exist $j, j' \in [n]$, $j \neq j'$, s.t. $\hat{x}_{i,j} = (0,1)$, $\hat{x}_{i,j'} = (0,1)$ and $\hat{x}_{i,j''} = (0,0)$ for all $j'' \in [n] \setminus \{j, j'\}$, then

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = \sqrt{a_{i,j} \cdot a_{i,j'}}.$$

- Otherwise,

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = 0.$$

The reader can verify that this matches the function $\hat{f}_{\mathrm{L},i}$ stated in Proposition 1 for arbitrary $n \in \mathbb{N}$.

As mentioned earlier, here we only showed how to find $\tilde{f}_{\mathrm{L},i}$ and $\hat{f}_{\mathrm{L},i}$, $i \in [n]$, because $\tilde{f}_{\mathrm{R},j}$ and $\hat{f}_{\mathrm{R},j}$, $j \in [n]$, are obtained in an analogous manner.

APPENDIX B
PROOF OF PROPOSITION 2

In this appendix, we want to prove that

$$\frac{\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})}{\mathrm{perm}(\mathbf{A})} = \sqrt{\sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} p(\sigma_1) \cdot p(\sigma_2) \cdot 2^{-c(\sigma_1, \sigma_2)}},$$

which can be rewritten as

$$\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2$$
$$= \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} 2^{-c(\sigma_1, \sigma_2)} \cdot \left(\prod_{i \in [n]} a_{i,\sigma_1(i)} \right) \cdot \left(\prod_{i \in [n]} a_{i,\sigma_2(i)} \right).$$

However, because of Proposition 1, this is equivalent to proving

$$Z(\hat{\mathsf{N}}(\mathbf{A})) = \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} 2^{-c(\sigma_1, \sigma_2)} \cdot \left(\prod_{i \in [n]} a_{i,\sigma_1(i)} \right) \cdot \left(\prod_{i \in [n]} a_{i,\sigma_2(i)} \right). \tag{21}$$

In order to prove (21), we need to better understand the valid configurations of $\hat{\mathsf{N}}(\mathbf{A})$ and their global function value.

Consider the NFG $\hat{\mathsf{N}}(\mathbf{A})$ as specified in Proposition 1. Recall that its global function is called $\hat{g}(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}} \triangleq (\hat{x}_{1,1}, \hat{x}_{1,2}, \ldots, \hat{x}_{n,n}) \in \hat{\mathcal{X}}^{(n^2)}$ is an arbitrary configuration of $\hat{\mathsf{N}}(\mathbf{A})$. Let $\mathcal{C}(\hat{\mathsf{N}}(\mathbf{A}))$ be the set of all valid configurations of $\hat{\mathsf{N}}(\mathbf{A})$, i.e., the set of all $\hat{\mathbf{x}} \in \hat{\mathcal{X}}^{(n^2)}$ such that $\hat{g}(\hat{\mathbf{x}}) \neq 0.$[4]

Consider Fig. 3, which is a reproduction of Fig. 2 (right). Let $\hat{\mathbf{x}} \in \mathcal{C}(\hat{\mathsf{N}}(\mathbf{A}))$ be a valid configuration. From Proposition 1 it follows that every vertex of $\hat{\mathsf{N}}(\mathbf{A})$ is either

- the endpoint of exactly a $(1,1)$-edge, where a $(1,1)$-edge is defined to be an edge such that $\hat{x}_{i,j} = (1,1)$, or
- a vertex of exactly one $(0,1)$-cycle, where a $(0,1)$-cycle is defined to be a simple cycle such that $\hat{x}_{i,j} = (0,1)$ for all its edges.

Example 1. *The NFG in Fig. 3 shows a possible valid configuration of $\hat{\mathsf{N}}(\mathbf{A})$ for the case $n = 5$. Here, an edge is colored in blue if $\hat{x}_{i,j} = (1,1)$, it is colored in red if $\hat{x}_{i,j} = (0,1)$, and it is colored in black if $\hat{x}_{i,j} = (0,0)$. (Note that $\hat{x}_{i,j} = (1,0)$ cannot occur in a valid configuration.)*

Definition 1. Consider the mapping

$$h : \mathcal{S}_n \times \mathcal{S}_n \to \mathcal{C}(\hat{\mathsf{N}}(\mathbf{A}))$$
$$(\sigma_1, \sigma_2) \mapsto \hat{\mathbf{x}},$$

where $\hat{x}_{i,j}$, $i, j \in [n]$, is defined as follows:

$$\hat{x}_{i,j} \triangleq \begin{cases} (1,1) & \text{if } j = \sigma_1(i) = \sigma_2(i) \\ (0,1) & \text{if } j = \sigma_1(i) \neq \sigma_2(i) \\ (0,1) & \text{if } j = \sigma_2(i) \neq \sigma_1(i) \\ (0,0) & \text{otherwise} \end{cases}.$$

---

[4]In the following, we assume that all entries of $\mathbf{A}$ are positive. The case where some entries of $\mathbf{A}$ are zero can be handled by suitable adaptations, or by considering a sequence of matrices with positive entries where some entries converge to 0 in the limit and using continuity of the relevant expressions.
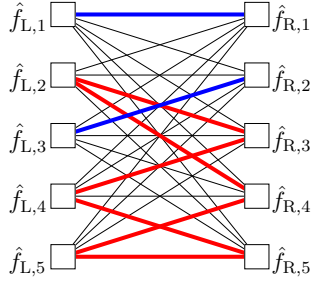
Fig. 3: NFG used in Appendix B. The two $(1,1)$-edges are colored in blue and the edges of the $(0,1)$-cycle are colored in red.
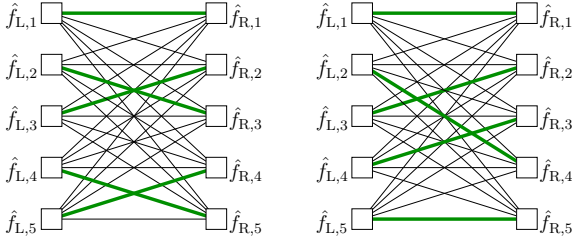

Fig. 4: NFGs used in Appendix B. Visualization of permutations $\sigma \in \mathcal{S}_n$. For $i,j \in [n]$, the edge connecting $\hat{f}_{L,i}$ to $\hat{f}_{R,j}$ is colored in green if and only if $\sigma(i) = j$. (Here, $n = 5$.)

(It can be readily verified that, indeed, $\hat{\mathbf{x}} \in \mathcal{C}(\hat{N}(\mathbf{A}))$.)

Example 2. *One can verify that if $\sigma_1, \sigma_2 \in \mathcal{S}_n$ are such that*

$$\sigma_1(1) = 1, \ \sigma_1(2) = 3, \ \sigma_1(3) = 2, \ \sigma_1(4) = 5, \ \sigma_1(5) = 4,$$
$$\sigma_2(1) = 1, \ \sigma_2(2) = 4, \ \sigma_2(3) = 2, \ \sigma_2(4) = 3, \ \sigma_2(5) = 5,$$

then $h(\sigma_1, \sigma_2)$ equals the valid configuration that is highlighted in the NFG in Fig. 3. Note that σ_1 and σ_2 are visualized in Fig. 4 (left) and Fig. 4 (right), respectively.

Example 3. *Similarly, one can verify that if $\sigma_1, \sigma_2 \in \mathcal{S}_n$ are such that*

$$\sigma_1(1) = 1, \ \sigma_1(2) = 4, \ \sigma_1(3) = 2, \ \sigma_1(4) = 3, \ \sigma_1(5) = 5,$$
$$\sigma_2(1) = 1, \ \sigma_2(2) = 3, \ \sigma_2(3) = 2, \ \sigma_2(4) = 5, \ \sigma_2(5) = 4,$$

then $h(\sigma_1, \sigma_2)$ equals the valid configuration that is highlighted in the NFG in Fig. 3. Note that σ_1 and σ_2 are visualized in Fig. 4 (right) and Fig. 4 (left), respectively.

As can be seen from Examples 2 and 3, a $(0,1)$-cycle in the valid configuration $h(\sigma_1, \sigma_2)$ arises from alternatingly picking an edge selected by σ_1 and an edge selected by σ_2.

Toward proving the upcoming Lemmas 1 and 2, we consider the following examples.

Example 4. *Consider again the selection of $\sigma_1, \sigma_2 \in \mathcal{S}_n$ in Example 2. Let $\sigma \triangleq \sigma_1 \circ \sigma_2^{-1}$, i.e., the permutation obtained by first applying the inverse of σ_2 and then σ_1. Then*

$$\sigma(1) = 1, \ \sigma(2) = 2, \ \sigma(3) = 5, \ \sigma(4) = 3, \ \sigma(5) = 4,$$

which in cycle notation is

$$\sigma = (1)(2)(354).$$

Note that

- *the cycle (1) of σ of length 1 corresponds to the blue-colored edge ending in $\hat{f}_{R,1}$ in Fig. 3,*
- *the cycle (2) of σ of length 1 corresponds to the blue-colored edge ending in $\hat{f}_{R,2}$ in Fig. 3,*
- *the cycle (354) of σ of length 3 corresponds to the blue-colored $(0,1)$-cycle in Fig. 3 of length $2 \cdot 3 = 6$ going through the right-hand side vertices $\hat{f}_{R,3}$, $\hat{f}_{R,5}$, $\hat{f}_{R,4}$.*

Example 5. *Consider again the selection of $\sigma_1, \sigma_2 \in \mathcal{S}_n$ in Example 3. Let $\sigma \triangleq \sigma_1 \circ \sigma_2^{-1}$. Then*

$$\sigma(1) = 1, \ \sigma(2) = 2, \ \sigma(3) = 4, \ \sigma(4) = 5, \ \sigma(5) = 3,$$

which in cycle notation is

$$\sigma = (1)(2)(345).$$

Note that

- *the cycle (1) of σ of length 1 corresponds to the blue-colored edge ending in $\hat{f}_{R,1}$ in Fig. 3,*
- *the cycle (2) of σ of length 1 corresponds to the blue-colored edge ending in $\hat{f}_{R,2}$ in Fig. 3,*
- *the cycle (345) of σ of length 3 corresponds to the blue-colored $(0,1)$-cycle in Fig. 3 of length $2 \cdot 3 = 6$ going through the right-hand side vertices $\hat{f}_{R,3}$, $\hat{f}_{R,4}$, $\hat{f}_{R,5}$.*

For $\sigma_1, \sigma_2 \in \mathcal{S}_n$, recall from Proposition B that $c(\sigma_1, \sigma_2)$ is defined to be the number of cycles of length larger than one in the cycle notation expression of the permutation $\sigma_1 \circ \sigma_2^{-1}$.

Lemma 1. *For any $\sigma_1, \sigma_2 \in \mathcal{S}_n$, the number of $(0,1)$-cycles in the valid configuration $h(\sigma_1, \sigma_2)$ equals $c(\sigma_1, \sigma_2)$.*

Proof. This follows from generalizing Examples 4 and 5. In particular, note that a cycle of length L, $L \geq 2$, of $\sigma_1 \circ \sigma_2^{-1}$ corresponds to a $(0,1)$-cycle of length $2L$ in the valid configuration $h(\sigma_1, \sigma_2)$ □

Lemma 2. *Fix an arbitrary $\hat{\mathbf{x}} \in \mathcal{C}(\hat{N}(\mathbf{A}))$. The number of pre-images of $\hat{\mathbf{x}}$ under the mapping h equals $2^{c(\sigma_1, \sigma_2)}$, i.e.,*

$$\left| \{ (\sigma_1, \sigma_2) \in \mathcal{S}_n \times \mathcal{S}_n \mid h(\sigma_1, \sigma_2) = \hat{\mathbf{x}} \} \right| = 2^{c(\sigma_1, \sigma_2)}.$$

Proof. This follows from generalizing Examples 2–5. In particular, note that for every $(0,1)$-cycle in $\hat{\mathbf{x}}$ there are two walk directions for going around the $(0,1)$-cycle, and with that two different ways of defining the relevant function values of σ_1 and σ_2. Because the walk direction can be chosen independently for every $(0,1)$-cycle in $\hat{\mathbf{x}}$, there are, ultimately, $2^{c(\sigma_1, \sigma_2)}$ ways of selecting $\sigma_1, \sigma_2 \in \mathcal{S}_n$ such that $h(\sigma_1, \sigma_2) = \hat{\mathbf{x}}$. □

Lemma 3. *For any $\sigma_1, \sigma_2 \in \mathcal{S}_n$, it holds that*

$$\hat{g}(h(\sigma_1, \sigma_2)) = \left(\prod_{i \in [n]} a_{i, \sigma_1(i)} \right) \cdot \left(\prod_{i \in [n]} a_{i, \sigma_2(i)} \right).$$

Proof. Fix some $\sigma_1, \sigma_2 \in \mathcal{S}_n$. Let $\hat{\mathbf{x}} \triangleq h(\sigma_1, \sigma_2)$. Then

$$\hat{g}\big(h(\sigma_1, \sigma_2)\big) = \hat{g}(\hat{\mathbf{x}})$$

$$= \left(\prod_{i \in [n]} \hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) \right)$$

$$\cdot \left(\prod_{j \in [n]} \hat{f}_{\mathrm{R},j}(\hat{x}_{1,j}, \ldots, \hat{x}_{n,j}) \right)$$

$$= \left(\prod_{i \in [n]} a_{i,\sigma_1(i)} \right) \cdot \left(\prod_{i \in [n]} a_{i,\sigma_2(i)} \right),$$

where the second equality follows from the specification of $\hat{\mathsf{N}}(\mathbf{A})$ in Proposition 1 and where the third equality is a consequence of the following observations:

- As seen above (see the text before Example 1), every vertex of $\hat{\mathsf{N}}(\mathbf{A})$ is either the endpoint of exactly one $(1,1)$-edge or the vertex of exactly one $(0,1)$-cycle.
- Consider a $(1,1)$-edge of $\hat{\mathbf{x}}$ connecting $\hat{f}_{\mathrm{L},i}$ and $\hat{f}_{\mathrm{R},j}$. Because

$$\hat{f}_{\mathrm{L},i}(\hat{x}_{i,1}, \ldots, \hat{x}_{i,n}) = a_{i,j},$$
$$\hat{f}_{\mathrm{R},j}(\hat{x}_{1,j}, \ldots, \hat{x}_{n,j}) = a_{i,j},$$

we get a contribution of $a_{i,j}^2$ from the two endpoints of this $(1,1)$-edge. Note that

$$a_{i,j}^2 = a_{i,\sigma_1(i)} \cdot a_{i,\sigma_2(i)}.$$

- Consider a $(0,1)$-cycle of length $2L$, $L \geq 2$, of $\hat{\mathbf{x}}$ through $\hat{f}_{\mathrm{L},i_1}, \hat{f}_{\mathrm{R},j_1}, \hat{f}_{\mathrm{L},i_2}, \hat{f}_{\mathrm{R},j_2}, \ldots, \hat{f}_{\mathrm{L},i_L}, \hat{f}_{\mathrm{R},j_L}, \hat{f}_{\mathrm{L},i_1}$, where

$$\sigma_1(i_1) = j_1, \quad \sigma_1(i_2) = j_2, \quad \ldots, \quad \sigma_1(i_L) = j_L,$$
$$\sigma_2(i_1) = j_L, \quad \sigma_2(i_2) = j_1, \quad \ldots, \quad \sigma_2(i_L) = j_{L-1}.$$

Because

$$\hat{f}_{\mathrm{L},i_1}(\hat{x}_{i_1,1}, \ldots, \hat{x}_{i_1,n}) = \sqrt{a_{i_1,j_L} \cdot a_{i_1,j_1}},$$
$$\hat{f}_{\mathrm{R},j_1}(\hat{x}_{1,j_1}, \ldots, \hat{x}_{n,j_1}) = \sqrt{a_{i_1,j_1} \cdot a_{i_2,j_1}},$$
$$\hat{f}_{\mathrm{L},i_2}(\hat{x}_{i_2,1}, \ldots, \hat{x}_{i_2,n}) = \sqrt{a_{i_2,j_1} \cdot a_{i_2,j_2}},$$
$$\hat{f}_{\mathrm{R},j_2}(\hat{x}_{1,j_2}, \ldots, \hat{x}_{n,j_2}) = \sqrt{a_{i_2,j_2} \cdot a_{i_3,j_2}},$$
$$\vdots \qquad\qquad \vdots$$
$$\hat{f}_{\mathrm{L},i_L}(\hat{x}_{i_L,1}, \ldots, \hat{x}_{i_L,n}) = \sqrt{a_{i_L,j_{L-1}} \cdot a_{i_L,j_L}},$$
$$\hat{f}_{\mathrm{R},j_L}(\hat{x}_{1,j_L}, \ldots, \hat{x}_{n,j_L}) = \sqrt{a_{i_L,j_L} \cdot a_{i_1,j_L}},$$

we get a contribution of

$$\left(\prod_{\ell=1}^{L} a_{i_\ell,\sigma_1(i_\ell)} \right) \cdot \left(\prod_{\ell=1}^{L} a_{i_\ell,\sigma_2(i_\ell)} \right)$$

from the $2L$ vertices of this $(0,1)$-cycle.

$\square$

Putting together the above results, we can finally prove Proposition B. Namely, we obtain

$$Z(\hat{\mathsf{N}}(\mathbf{A})) = \sum_{\hat{\mathbf{x}} \in \mathcal{C}(\hat{\mathsf{N}}(\mathbf{A}))} \hat{g}(\hat{\mathbf{x}})$$

$$= \sum_{\hat{\mathbf{x}} \in \mathcal{C}(\hat{\mathsf{N}}(\mathbf{A}))} \sum_{\substack{\sigma_1, \sigma_2 \in \mathcal{S}_n: \\ h(\sigma_1, \sigma_2) = \hat{\mathbf{x}}}} 2^{-c(\sigma_1, \sigma_2)} \cdot \hat{g}(\hat{\mathbf{x}})$$

$$= \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} 2^{-c(\sigma_1, \sigma_2)} \cdot \hat{g}\big(h(\sigma_1, \sigma_2)\big)$$

$$= \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} 2^{-c(\sigma_1, \sigma_2)} \cdot \left(\prod_{i \in [n]} a_{i,\sigma_1(i)} \right) \cdot \left(\prod_{i \in [n]} a_{i,\sigma_2(i)} \right),$$

where the second equality follows from Lemma 2 and where the fourth equality follows from Lemma 3. This proves the validity of (21), and with that the validity of Proposition B. $\square$

APPENDIX C
PROOF OF THEOREM 1

Let $\mathcal{G}$ be a subgroup of $\mathcal{S}_n$. The cycle index of $\mathcal{G}$ is defined to be (see, e.g., [26, Section 6.6])

$$Z(\mathcal{G}) \triangleq \frac{1}{|\mathcal{G}|} \sum_{\sigma \in \mathcal{G}} \prod_{k \in [n]} z_k^{c_k(\sigma)},$$

where $c_k(\sigma)$, $k \in [n]$, denotes the number of cycles of length k in the cycle notation expression of the permutation σ, and where z_k, $k \in [n]$, are indeterminates. If $\mathcal{G} = \mathcal{S}_n$, which is the case of interest here, then

$$Z(\mathcal{S}_n) \triangleq \frac{1}{|\mathcal{S}_n|} \sum_{\sigma \in \mathcal{S}_n} \prod_{k \in [n]} z_k^{c_k(\sigma)}$$

$$= \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{k \in [n]} z_k^{c_k(\sigma)}.$$

The following well-known result gives a convenient recursive expression for $Z(\mathcal{S}_n)$.

Lemma 4. *Define $Z(\mathcal{S}_0) \triangleq 1$. For $n \geq 1$ it holds that*

$$Z(\mathcal{S}_n) = \frac{1}{n} \sum_{\ell \in [n]} z_\ell \cdot Z(\mathcal{S}_{n-\ell}). \tag{22}$$

Proof. For $\ell \in [n]$, let $\mathcal{S}_n(\ell)$ be subset of $\mathcal{S}_n$ that contains all permutations σ such that n is contained in a cycle of σ of length ℓ. Note that $|\mathcal{S}_n(\ell)|/|\mathcal{S}_{n-\ell}| = \binom{n-1}{\ell-1} \cdot (\ell-1)!$. Therefore,

$$Z(\mathcal{S}_n) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{k \in [n]} z_k^{c_k(\sigma)}$$

$$= \frac{1}{n!} \sum_{\ell \in [n]} \sum_{\sigma \in \mathcal{S}_n(\ell)} \prod_{k \in [n]} z_k^{c_k(\sigma)}$$

$$= \frac{1}{n!} \sum_{\ell \in [n]} \binom{n-1}{\ell-1} \cdot (\ell-1)! \cdot z_\ell \cdot \sum_{\sigma \in \mathcal{S}_{n-\ell}} \prod_{k \in [n-\ell]} z_k^{c_k(\sigma)}$$

$$= \frac{1}{n} \sum_{\ell \in [n]} z_\ell \cdot Z(\mathcal{S}_{n-\ell}).$$

□

For $n \in \mathbb{N} \cup \{0\}$, define $Z_n \triangleq Z(\mathcal{S}_n)$, where

$$z_k \triangleq \begin{cases} 1 & \text{if } k = 1 \\ \frac{1}{2} & \text{if } k \geq 2 \end{cases}$$

Rewriting (22) for this choice of z_k, $k \geq 1$, we obtain

$$Z_n = \frac{1}{n} Z_{n-1} + \frac{1}{2n} \sum_{\ell=2}^{n} Z_{n-\ell}. \tag{23}$$

Contemplating Fig. 5, it appears that $Z_n = \Theta(\frac{1}{\sqrt{n}})$. In the following, we will prove this observation. In fact, we will prove

$$\frac{C_l}{\sqrt{n}} \leq Z_n \leq \frac{C_u}{\sqrt{n}}, \ n \geq 1, \qquad \text{and} \qquad Z_n \sim \frac{C}{\sqrt{n}},$$

where

$$C_l \triangleq \frac{1}{\sqrt{2}}, \quad C_u \triangleq \frac{3}{2\sqrt{2}}, \quad C \triangleq \sqrt{\frac{e}{\pi}}. \tag{24}$$

Lemma 5. *For all $n \in \mathbb{N}$ it holds that*

$$Z_n \leq \frac{C_u}{\sqrt{n}},$$

where C_u was defined in (24). Thus $Z_n = \mathrm{O}\big(\frac{1}{\sqrt{n}}\big)$.

Proof. The proof is by strong induction.

Base cases ($n = 1$ and $n = 2$): $Z_1 = 1 \leq \frac{3}{2\sqrt{2}} = \frac{C_u}{\sqrt{1}}$ and $Z_2 = \frac{3}{4} = \frac{3}{2\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{C_u}{\sqrt{2}}$.

Induction step ($n = k+1$ for some $k \in \mathbb{N}$, $k \geq 2$): Assume the claim is true for every $j \in [k]$. Using (23), we obtain

$$\begin{aligned}
Z_{k+1} &= \frac{Z_k}{k+1} + \frac{1}{2(k+1)} \sum_{j=2}^{k+1} Z_{k+1-j} \\
&= \frac{2 + Z_k}{2(k+1)} + \frac{1}{2(k+1)} \sum_{j=2}^{k-1} Z_{k+1-j} \\
&= \frac{2 + Z_k}{2(k+1)} + \frac{1}{2(k+1)} \sum_{j=2}^{k} Z_j \\
&\leq \frac{1}{k+1} \cdot \left(1 + \frac{C_u}{2\sqrt{k}}\right) + \frac{1}{2(k+1)} \sum_{j=2}^{k} \frac{C_u}{\sqrt{j}} \\
&\leq \frac{C_u}{k+1} \cdot \left(\frac{1}{C_u} + \frac{1}{2\sqrt{k}}\right) + \frac{C_u}{2(k+1)} \int_1^k \frac{1}{\sqrt{x}} \, \mathrm{d}x \\
&= \frac{C_u}{k+1} \cdot \left(\frac{1}{C_u} + \frac{1}{2\sqrt{k}} + (\sqrt{k} - 1)\right) \\
&\leq \frac{C_u}{k+1} \cdot \sqrt{k+1} \\
&= \frac{C_u}{\sqrt{k+1}}.
\end{aligned}$$

Notice that the last inequality is valid since the function $f : \mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto \sqrt{x+1} - \sqrt{x} - \frac{1}{2\sqrt{x}} + 1 - \frac{2\sqrt{2}}{3}$ is positive for $x \geq 2$, which can be proven by the observing that $f(2) \approx 0.0215 > 0$ and that f is strictly increasing because
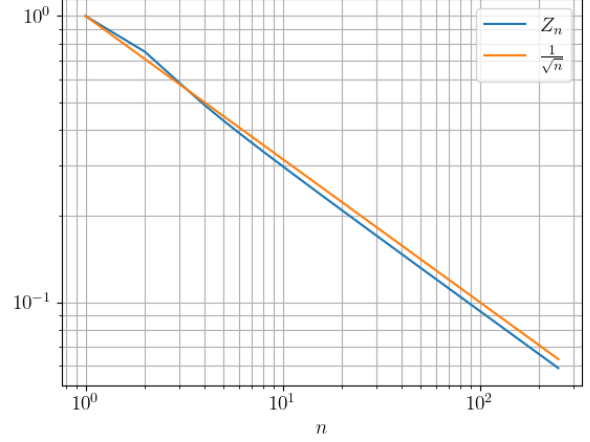

Fig. 5: Graph comparing the functions $n \mapsto Z_n$ with $n \mapsto \frac{1}{\sqrt{n}}$ for $n \in [250]$.
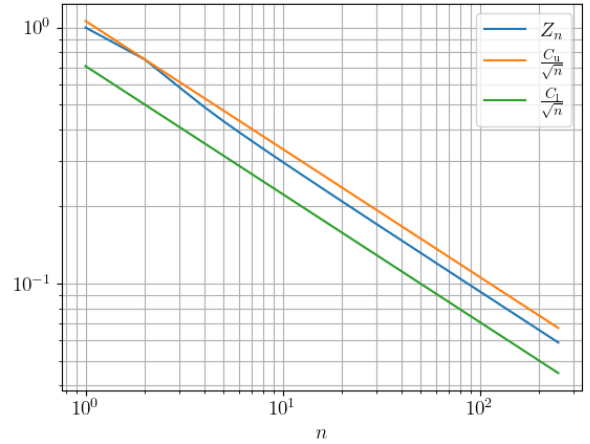

Fig. 6: Graph comparing the functions $n \mapsto Z_n$ with $n \mapsto \frac{C_u}{\sqrt{n}}$ and $n \mapsto \frac{C_l}{\sqrt{n}}$ for $n \in [250]$.

its derivative satisfies $f'(x) = \frac{1}{2\sqrt{x+1}} - \frac{1}{2\sqrt{x}} + \frac{1}{4x^{3/2}} > 0$ for all $x > 0$. □

Lemma 6. *For all $n \in \mathbb{N}$ it holds that*

$$Z_n \geq \frac{C_l}{\sqrt{n}},$$

where C_l was defined in (24). Thus $Z_n = \Omega\big(\frac{1}{\sqrt{n}}\big)$.

Proof. The proof is by strong induction.

Base cases ($n = 1$ and $n = 2$): $Z_1 = 1 \geq \frac{1}{\sqrt{2}} = \frac{C_l}{\sqrt{1}}$ and $Z_2 = \frac{3}{4} = \frac{3}{2} \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \geq \frac{C_l}{\sqrt{2}}$.

Induction step ($n = k+1$ for some $k \in \mathbb{N}$, $k \geq 2$): Assume

the claim is true for every $j \in [k]$. Using (23), we obtain

$$
\begin{aligned}
Z_{k+1} &= \frac{Z_k}{k+1} + \frac{1}{2(k+1)} \sum_{j=2}^{k+1} Z_{k+1-j} \\
&= \frac{2 + Z_k}{2(k+1)} + \frac{1}{2(k+1)} \sum_{j=2}^{k-1} Z_{k+1-j} \\
&\geq \frac{1}{k+1} \cdot \left(1 + \frac{C_1}{2\sqrt{k}}\right) + \frac{1}{2(k+1)} \sum_{j=2}^{k} \frac{C_1}{\sqrt{j}} \\
&\geq \frac{C_1}{k+1} \cdot \left(\frac{1}{C_1} + \frac{1}{2\sqrt{k}}\right) + \frac{C_1}{2(k+1)} \int_2^{k+1} \frac{1}{\sqrt{x}} \, dx \\
&= \frac{C_1}{k+1} \cdot \left(\frac{1}{C_1} + \frac{1}{2\sqrt{k}} + (\sqrt{k+1} - \sqrt{2})\right) \\
&\geq \frac{C_1}{k+1} \cdot \sqrt{k+1} \\
&= \frac{C_1}{\sqrt{k+1}}.
\end{aligned}
$$

Notice that the last inequality is valid since the function $f : \mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto \frac{1}{2\sqrt{x}}$ is positive for $x \geq 2$. $\qquad\square$

Proposition 3. *It holds that*

$$
Z_n = \Theta\left(\frac{1}{\sqrt{n}}\right).
$$

Proof. This follows immediately from Lemmas 5 and 6. $\quad\square$

A comparison of the function $n \mapsto Z_n$ with the upper and lower bounds in Lemmas 5 and 6 is shown in Fig. 6.

Toward proving $Z_n \sim \frac{C}{\sqrt{n}}$, where $C \triangleq \sqrt{\frac{e}{\pi}}$, we use the following ingredients:

- It holds that

$$
Z(\mathcal{S}_n) = \frac{1}{n!} \cdot B_n\big(0! \cdot z_1, 1! \cdot z_2, \ldots, (n-1)! \cdot z_n\big),
$$

where B_n denotes the n-th complete exponential Bell polynomial. (This relationship follows from the fact that a permutation is, by definition, bijective, and from identifying each monomial in $Z(\mathcal{S}_n)$ with a partitioning of $[n]$.) Consequently,

$$
Z_n = \frac{1}{n!} \cdot B_n\left(0! \cdot 1, 1! \cdot \frac{1}{2}, \ldots, (n-1)! \cdot \frac{1}{2}\right).
$$

- It holds that (see, e.g., [25, Theorem 4.38])

$$
Z(\mathcal{S}_n) = \left(\frac{\partial}{\partial t}\right)^n C(t) \bigg|_{t=0},
$$

where

$$
C(t) \triangleq \exp\left(\sum_{k=1}^{\infty} \frac{z_k t^k}{k}\right)
$$

is the generating function of the cycle index of the symmetric group.

- We recall the following well-known power series, which are convergent in a neighborhood around $t = 0$:

$$
\ln(1 - t) = -\sum_{k=1}^{\infty} \frac{t^k}{k},
$$

$$
e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!},
$$

$$
(1 - t)^{-1/2} = \sum_{k=0}^{\infty} \frac{(2k)!}{4^k \cdot (k!)^2} t^k.
$$

- We recall the following bounds on the k-th central binomial coefficient for $k \in \mathbb{N}$:

$$
\frac{4^k}{\sqrt{(k + \frac{1}{2})\pi}} \leq \binom{2k}{k} \leq \frac{4^k}{\sqrt{k\pi}}. \tag{25}
$$

Lemma 7. *It holds that*

$$
\lim_{n \to \infty} Z_n \cdot \sqrt{n} \geq C.
$$

Proof. Substituting

$$
z_k \triangleq \begin{cases} 1 & \text{if } k = 1 \\ \frac{1}{2} & \text{if } k \geq 2 \end{cases}
$$

into $C(t)$, we get

$$
\begin{aligned}
C(t) &= \exp\left(t + \frac{1}{2} \sum_{k=2}^{\infty} \frac{t^k}{k}\right) \\
&= \exp\left(t + \frac{1}{2}\big(-\ln(1 - t) - t\big)\right) \\
&= \exp\left(\frac{1}{2}\big(t - \ln(1 - t)\big)\right) \\
&= e^{t/2}(1 - t)^{-1/2} \\
&= \left(\sum_{k=0}^{\infty} \frac{(\frac{t}{2})^k}{k!}\right) \cdot \left(\sum_{k=0}^{\infty} \frac{(2k)!}{4^k \cdot (k!)^2} t^k\right) \\
&= \sum_{n=0}^{\infty} \sum_{\ell=0}^{n} \frac{1}{2^\ell \cdot \ell!} \cdot \frac{1}{4^{n-\ell}} \cdot \binom{2(n-\ell)}{n-\ell} \cdot t^n.
\end{aligned}
$$

Letting $\tau \triangleq \frac{1}{2}$, we obtain for all $n \in \mathbb{N}$

$$
\begin{aligned}
Z_n &= \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{4^{n-\ell}} \cdot \binom{2(n-\ell)}{n-\ell} \\
&\overset{(a)}{\geq} \frac{\tau^n}{n!} + \sum_{\ell=0}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{4^{n-\ell}} \cdot \frac{4^{n-\ell}}{\sqrt{(n - \ell + \frac{1}{2})\pi}} \\
&\geq \frac{\tau^n}{n!} + \sum_{\ell=0}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{(n + \frac{1}{2})\pi}} \\
&\geq \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{(1 + \frac{1}{2n})\pi}} \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!},
\end{aligned}
$$

where step (a) is due to (25). It follows that

$$\lim_{n \to \infty} Z_n \cdot \sqrt{n} \geq \lim_{n \to \infty} \frac{1}{\sqrt{(1 + \frac{1}{2n})\pi}} \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!} = \frac{\exp(\tau)}{\sqrt{\pi}} = \sqrt{\frac{e}{\pi}}.$$

□

Lemma 8. *It holds that*

$$\lim_{n \to \infty} Z_n \cdot \sqrt{n} \leq C.$$

Proof. The proof is similar to the proof of Lemma 7. In particular, from the proof of Lemma 7 we know that for all $n \in \mathbb{N}$ it holds that

$$Z_n = \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{4^{n-\ell}} \cdot \binom{2(n-\ell)}{n-\ell}.$$

Using (25), we obtain

$$Z_n \leq \frac{\tau^n}{n!} + \sum_{\ell=0}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{4^{n-\ell}} \cdot \frac{4^{n-\ell}}{\sqrt{(n-\ell)\pi}}$$

$$= \frac{\tau^n}{n!} + \frac{1}{\sqrt{\pi}} \sum_{\ell=0}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-\ell}}$$

$$= \frac{\tau^n}{n!} + \frac{1}{\sqrt{\pi}} \left(\sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-\ell}} + \sum_{\ell=h(n)+1}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-\ell}} \right)$$

$$\leq \frac{\tau^n}{n!} + \frac{1}{\sqrt{\pi}} \left(\sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-h(n)}} + \sum_{\ell=h(n)+1}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-\ell}} \right)$$

$$= \frac{\tau^n}{n!} + \frac{(1-\frac{h(n)}{n})^{-1/2}}{\sqrt{n}\cdot\sqrt{\pi}} \sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n-\ell}},$$

where $h : \mathbb{N} \to \mathbb{N}$ is a function satisfying $h(n) = \omega_n(1)$ and $h(n) = o(n)$. In particular, there exists an $N \in \mathbb{N}$ such that $\forall\, n \geq N$, it holds that $1 < h(n) < n$. Then notice that

$$0 \leq \sum_{\ell=h(n)+1}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{\sqrt{n}}{\sqrt{n-\ell}}$$

$$\overset{(a)}{\leq} \sum_{\ell=h(n)+1}^{n-1} \tau^\ell \cdot \frac{\sqrt{n}}{\sqrt{n-\ell}} \cdot \frac{1}{\sqrt{2\pi\ell}} \left(\frac{e}{\ell}\right)^\ell e^{-1/(12\ell+1)}$$

$$\leq \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \frac{\sqrt{n}}{\sqrt{\ell(n-\ell)}} \left(\frac{\tau e}{\ell}\right)^\ell e^{-0}$$

$$\overset{(b)}{\leq} \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \frac{\sqrt{n}}{\sqrt{(n-1)(n-(n-1))}} \left(\frac{\tau e}{\ell}\right)^\ell$$

$$\leq \frac{1}{\sqrt{2\pi}} \left(1-\frac{1}{n}\right)^{-1/2} \sum_{\ell=h(n)+1}^{n-1} \left(\frac{\tau e}{h(n)}\right)^\ell$$

$$= \frac{1}{\sqrt{2\pi}} \left(1-\frac{1}{n}\right)^{-1/2} \frac{\left(\frac{\tau e}{h(n)}\right)^{h(n)+1} - \left(\frac{\tau e}{h(n)}\right)^n}{1 - \frac{\tau e}{h(n)}}, \quad (26)$$

where step (a) is due to Robbins' approximations for factorial function [27], and where step (b) is due to the observation that $x = n-1$ is a global minimum for the polynomial function $x \mapsto x \cdot (n-x)$ on $[n-1]$.

Note that the limit of the right-hand side of (26) equals 0 as $n \to \infty$ since $\frac{\tau e}{h(n)} < 1$ for sufficiently large n, which is guaranteed by $h(n) = \omega_n(1)$. So, we obtain

$$\lim_{n \to \infty} Z_n \cdot \sqrt{n}$$

$$\leq \lim_{n \to \infty} \left(\frac{\sqrt{n} \cdot \tau^n}{n!} + \frac{(1-\frac{h(n)}{n})^{-1/2}}{\sqrt{\pi}} \sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} \right.$$

$$\left. + \sum_{\ell=h(n)+1}^{n} \frac{\tau^\ell}{\ell!} \cdot \frac{\sqrt{n}}{\sqrt{n-\ell}} \right)$$

$$= 0 + \frac{(1-0)^{-1/2}}{\sqrt{\pi}} \cdot \exp\left(\frac{1}{2}\right) + 0 = \sqrt{\frac{e}{\pi}}.$$

□

Proposition 4. *It holds that*

$$Z_n \sim C \cdot \frac{1}{\sqrt{n}},$$

where C was defined in (24).

Proof. This follows directly from Lemmas 7 and 8. □

APPENDIX D
PROOF OF THEOREM 2

Given Assumption 1, we are interested in the expectation value of $\text{perm}(\mathbf{A})$, $\text{perm}(\mathbf{A})^2$, and $\text{perm}_{\text{B},2}(\mathbf{A})^2$.

The expectation value of the permanent of $\mathbf{A}$ is given by[5]

$$\mathbb{E}\big[\text{perm}(\mathbf{A})\big] = \mathbb{E}\left[\sum_{\sigma \in \mathcal{S}_n} \prod_{i \in [n]} a_{i,\sigma(i)} \right]$$

$$\overset{(a)}{=} \sum_{\sigma \in \mathcal{S}_n} \mathbb{E}\left[\prod_{i \in [n]} a_{i,\sigma(i)} \right]$$

$$\overset{(b)}{=} \sum_{\sigma \in \mathcal{S}_n} \prod_{i \in [n]} \mathbb{E}\big[a_{i,\sigma(i)}\big]$$

$$= n! \cdot \mu_1^n,$$

where step (a) is due to the linearity of the expectation value, and where step (b) is due to the independence between the entries of $\mathbf{A}$. Similarly, we obtain

$$\mathbb{E}\big[\text{perm}(\mathbf{A})^2\big] = \sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} \prod_{i \in [n]} \mathbb{E}\big[a_{i,\sigma_1(i)} \cdot a_{i,\sigma_2(i)}\big]$$

$$= \sum_{\sigma_1,\sigma_2 \in \mathcal{S}_n} \prod_{i \in [n]} h_i(\sigma_1, \sigma_2)$$

where

$$h_i(\sigma_1, \sigma_2) \triangleq \mathbb{E}\big[a_{i,\sigma_1(i)} \cdot a_{i,\sigma_2(i)}\big] = \begin{cases} \mu_2 & \text{if } \sigma_1(i) = \sigma_2(i) \\ \mu_1^2 & \text{otherwise} \end{cases}.$$

---

[5]Note that we actually do not need this result in the following.

Because

$$h_i(\sigma_1, \sigma_2) = h_i(\sigma_1 \circ \sigma_2^{-1}, \sigma_2 \circ \sigma_2^{-1}) = h_i(\sigma_1 \circ \sigma_2^{-1}, \mathrm{id}),$$

where id is the identity permutation, we obtain

$$\begin{aligned}
\mathbb{E}\big[\mathrm{perm}(\mathbf{A})^2\big] &= \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} \prod_{i \in [n]} h_i(\sigma_1, \sigma_2) \\
&= \sum_{\sigma_1, \sigma_2 \in \mathcal{S}_n} \prod_{i \in [n]} h_i(\sigma_1 \circ \sigma_2^{-1}, \mathrm{id}) \\
&= n! \cdot \sum_{\sigma \in \mathcal{S}_n} \prod_{i \in [n]} h_i(\sigma, \mathrm{id}),
\end{aligned}$$

where

$$h_i(\sigma, \mathrm{id}) = \begin{cases} \mu_2 & \text{if } \sigma(i) = i \\ \mu_1^2 & \text{otherwise} \end{cases}.$$

Note that every cycle of length 1 of σ contributes a factor of μ_2 to $\prod_{i \in [n]} h_i(\sigma, \mathrm{id})$. Similarly, every cycle of length k, $k \geq 2$, contributes a factor of $(\mu_1^2)^k$ to $\prod_{i \in [n]} h_i(\sigma, \mathrm{id})$. Therefore, using the cycle index of $\mathcal{S}_n$, we can write

$$\mathbb{E}\big[\mathrm{perm}(\mathbf{A})^2\big] = Z(\mathcal{S}_n),$$

where

$$z_k = \begin{cases} \mu_2 & \text{if } k = 1 \\ (\mu_1^2)^k & \text{if } k \geq 2 \end{cases}.$$

Applying the generation function technique as in the proofs of Lemmas 7 and 8, we obtain

$$\begin{aligned}
\mathbb{E}\big[\mathrm{perm}(\mathbf{A})^2\big] &= (n!)^2 \cdot \sum_{\ell=0}^{n} \frac{(\mu_2 - \mu_1^2)^\ell \mu_1^{2(n-\ell)}}{\ell!} \\
&= (n!)^2 \cdot \mu_1^{2n} \cdot \sum_{\ell=0}^{n} \frac{\left(\frac{\mu_2}{\mu_1^2} - 1\right)^\ell}{\ell!} \\
&\sim (n!)^2 \cdot \mu_1^{2n} \cdot \exp\left(\frac{\mu_2}{\mu_1^2} - 1\right). \quad (27)
\end{aligned}$$

Similarly, one obtains

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{perm}_{B,2}(\mathbf{A})^2\big] \\
&= (n!)^2 \cdot \mu_1^{2n} \cdot \sum_{\ell=0}^{n} \frac{\left(\frac{\mu_2}{\mu_1^2} - \frac{1}{2}\right)^\ell}{\ell!} \cdot \frac{(2(n-\ell))!}{4^{n-\ell} \cdot ((n-\ell)!)^2}. \quad (28)
\end{aligned}$$

By carefully reusing some calculations in the proofs of Lemmas 7 and 8, the summation appearing in (28) can then be bounded between

$$\frac{\mu_1^{2n}}{\sqrt{(n + \frac{1}{2})\pi}} \cdot \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!}$$

and

$$\mu_1^{2n}\left(\frac{\tau^n}{n!} + \frac{(1 - \frac{h(n)}{n})^{-1/2}}{\sqrt{n\pi}} \cdot \sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n-1} \frac{\tau^\ell}{\ell!} \cdot \frac{1}{\sqrt{n - \ell}}\right),$$

where $\tau \triangleq \frac{\mu_2}{\mu_1^2} - \frac{1}{2}$. Finally, we obtain

$$\frac{\mathbb{E}\big[\mathrm{perm}_{B,2}(\mathbf{A})^2\big]}{(n!)^2} \sim \frac{\mathrm{e}^\tau}{\sqrt{\pi}} \cdot \frac{\mu_1^{2n}}{\sqrt{n}}. \quad (29)$$

Combining (27) and (29), we get

$$\begin{aligned}
\frac{\mathbb{E}\big[\mathrm{perm}(\mathbf{A})^2\big]}{\mathbb{E}\big[\mathrm{perm}_{B,2}(\mathbf{A})^2\big]} &\sim \frac{(n!)^2 \cdot \mu_1^{2n} \cdot \exp\left(\frac{\mu_2}{\mu_1^2} - 1\right)}{(n!)^2 \cdot \frac{\mathrm{e}^\tau}{\sqrt{\pi}} \cdot \frac{\mu_1^{2n}}{\sqrt{n}}} \\
&= \sqrt{\pi n} \cdot \exp\left(\tau - \frac{1}{2} - \tau\right) \\
&= \sqrt{\frac{\pi n}{\mathrm{e}}},
\end{aligned}$$

which, surprisingly, does not depend on the chosen distribution!

We remark that the results in Appendix C can be regarded as a special case of the results in this section. Indeed, if the entries of $\mathbf{A}$ are chosen i.i.d. according to the (degenerate) distribution that assigns $a_{i,j} = 1$ with probability 1, then $\mu_1 = 1$ and $\mu_2 = 1$, and with that the right-hand side of (29) simplifies to $\sqrt{\frac{\mathrm{e}}{\pi}} \cdot \frac{1}{\sqrt{n}}$.

APPENDIX E
UNIFICATION OF THE RESULTS IN APPENDICES C AND D

In this appendix, we show that the results in Appendices C and D can be obtained as special cases of a more general result.

Recall that for $n \in \mathbb{N}$, we defined $[n] \triangleq \{1, 2, \ldots, n\}$. More generally, for $n_1, n_2 \in \mathbb{Z}$ such that $n_1 \leq n_2$, we define $[\![n_1, n_2]\!] \triangleq \{n_1, n_1 + 1, \ldots, n_2\}$.

Definition 2. Let $n \in \mathbb{N}$. We define the function

$$\begin{aligned}
\Psi_n : \mathbb{R}_{\geq 0}^3 &\to \mathbb{R}_{\geq 0}, \\
(\theta_1, \theta_2, \theta_3) &\mapsto (n!)^2 \cdot \theta_2^n \cdot \sum_{\ell=0}^{n} \binom{n - \ell + \theta_3 - 1}{n - \ell} \cdot \frac{\left(\frac{\theta_1}{\theta_2} - \theta_3\right)^\ell}{\ell!},
\end{aligned}$$

where the binomial coefficient here is the generalized binomial coefficient, which, for $\alpha \in \mathbb{R}$ and $k \in \mathbb{N} \cup \{0\}$, is defined to be

$$\binom{\alpha}{k} \triangleq \frac{\alpha \cdot (\alpha - 1) \cdots (\alpha - k + 1)}{k!}.$$

Lemma 9. For all $n \in \mathbb{N}$ it holds that

1) $\Psi_n(\theta_1, \theta_2, \theta_3) = n! \cdot B_n(0! \cdot \theta_1, 1! \cdot \theta_3 \theta_2^2, \ldots, (n-1)! \cdot \theta_3 \theta_2^n)$,

2) $\Psi_n(\theta_1, \theta_2, \theta_3) \sim (n!)^2 \cdot \theta_2^n \cdot \frac{n^{\theta_3 - 1}}{\Gamma(\theta_3)} \cdot \exp\left(\frac{\theta_1}{\theta_2} - \theta_3\right)$,

where Item 1 holds for $\theta_3 \in \mathbb{R}_{>0}$, Item 2 holds for $\theta_3 \in (0, 1]$ (possibly for $\theta_3 \in \mathbb{R}_{>0}$, but we did not need this generalization), and where $\Gamma(\cdot)$ denotes the Gamma function.

Proof. Item 1 follows from analyzing $C(t)$, the generating function of the cycle index of the symmetric group, with the substitution

$$z_k = \begin{cases} \theta_1 & \text{if } k = 1 \\ \theta_3\theta_2^k & \text{if } k \geq 2 \end{cases}.$$

In the following, we will use the binomial series, i.e.,

$$(1+t)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} t^k,$$

which is convergent for any $\alpha \in \mathbb{R}$ and $t \in (-1, 1)$. (Note that the binomial coefficient here is the generalized binomial coefficient.)

Adapting some calculations in the proof of Lemma 7, we obtain

$$
\begin{aligned}
C(t) &= \exp\left(\theta_1 t + \sum_{k=2}^\infty \frac{\theta_3\theta_2^k}{k} t^k\right) \\
&= \exp(\theta_1 t + \theta_3(-\ln(1 - \theta_2 t) - \theta_2 t)) \\
&= \mathrm{e}^{(\theta_1 - \theta_3\theta_2)t}(1 - \theta_2 t)^{-\theta_3} \\
&= \left(\sum_{k=0}^\infty \frac{((\theta_1 - \theta_3\theta_2)t)^k}{k!}\right) \cdot \left(\sum_{k=0}^\infty \binom{-\theta_3}{k}(-\theta_2 t)^k\right) \\
&= \left(\sum_{k=0}^\infty \frac{(\theta_1 - \theta_3\theta_2)^k}{k!}t^k\right) \cdot \left(\sum_{k=0}^\infty \theta_2^k\binom{k+\theta_3-1}{k}t^k\right) \\
&= \sum_{n=0}^\infty \sum_{\ell=0}^n \frac{\theta_2^{n-\ell}(\theta_1 - \theta_3\theta_2)^\ell}{\ell!} \cdot \binom{n-\ell+\theta_3-1}{n-\ell} \cdot t^n.
\end{aligned}
$$

Then we can verify that

$$
\begin{aligned}
&n! \cdot B_n(0! \cdot \theta_1, 1! \cdot \theta_3\theta_2^2, \ldots, (n-1)! \cdot \theta_3\theta_2^n) \\
&= (n!)^2\theta_2^n \sum_{\ell=0}^n \binom{n-\ell+\theta_3-1}{n-\ell} \cdot \frac{(\frac{\theta_1}{\theta_2} - \theta_3)^\ell}{\ell!} \\
&= \Psi_n(\theta_1, \theta_2, \theta_3).
\end{aligned}
$$

For Item 2, we first notice that for any $n \in \mathbb{N}$, $\ell \in [\![0, n-1]\!]$, $\theta_3 \in (0, 1]$, we have

$$
\begin{aligned}
&\binom{n-\ell+\theta_3-1}{n-\ell} \\
&= \frac{\Gamma((n-\ell+\theta_3-1)+1)}{\Gamma((n-\ell)+1)\Gamma((n-\ell+\theta_3-1)-(n-\ell)+1)} \\
&= \frac{\Gamma(n-\ell+\theta_3)}{\Gamma(n-\ell+1)\Gamma(\theta_3)}.
\end{aligned}
$$

In the following, we will need Wendel's double inequalities for the ratio of two Gamma function values [28]:

$$\left(\frac{x}{x+s}\right)^{1-s} \leq \frac{\Gamma(x+s)}{x^s\Gamma(x)} \leq 1, \tag{30}$$

where $x > 0 \in \mathbb{R}$, $s \in [0, 1]$. This double inequalities can be rewritten as

$$\left(\frac{x}{x+s}\right)^{1-s} \leq \frac{\Gamma(x+s)}{x^{s-1}\Gamma(x+1)} \leq 1, \tag{31}$$

or as

$$(x+s)^{s-1} \leq \frac{\Gamma(x+s)}{\Gamma(x+1)} \leq x^{s-1}. \tag{32}$$

Notice that (31) is due to the property $\Gamma(x+1) = x \cdot \Gamma(x)$ of the Gamma function. Substituting $x = n - \ell$ and $s = c$ into (32), we obtain

$$(n-\ell+\theta_3)^{\theta_3-1} \leq \frac{\Gamma(n-\ell+\theta_3)}{\Gamma(n-\ell+1)} \leq (n-\ell)^{\theta_3-1}.$$

Moreover,

$$\frac{(n-n+\theta_3)^{\theta_3-1}}{\Gamma(\theta_3)} = \frac{\Gamma(1)}{\theta_3^{1-\theta_3}\Gamma(\theta_3)} = \frac{\Gamma(c+(1-\theta_3))}{\theta_3^{1-\theta_3}\Gamma(\theta_3)} \leq 1,$$

where the inequality is obtained from substituting $x = \theta_3$ and $s = 1 - \theta_3$ into (30). Observe that, because Γ is decreasing in $(0, 1]$ and because $\Gamma(1) = 1$, it holds that $\frac{1}{\Gamma(\theta_3)} \leq \frac{1}{\Gamma(1)} = 1$. Then, letting $\tau \triangleq \frac{\theta_1}{\theta_2} - \theta_3$, one obtains the expression in (33) (see top of the next page) and the expression in (34) (see top of the next page), where $h : \mathbb{N} \to \mathbb{N}$ is a function satisfying $h(n) = \omega_n(1)$ and $h(n) = o(n)$. In summary, we obtain the expression in (35) (see top of the next page).

Finally, by using parts of the proof of Lemma 8, we obtain

$$
\begin{aligned}
&\sum_{\ell=h(n)+1}^{n-1} \frac{n^{1-\theta_3}}{(n-\ell)^{1-\theta_3}} \cdot \frac{\tau^\ell}{\ell!} \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \frac{n^{1-\theta_3}}{(n-\ell)^{1-\theta_3}} \cdot \frac{1}{\sqrt{\ell}}\left(\frac{\tau\mathrm{e}}{\ell}\right)^\ell \cdot \mathrm{e}^{-0} \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \left(\frac{n}{\ell(n-\ell)}\right)^{1-\theta_3} \cdot \left(\frac{\tau\mathrm{e}}{h(n)}\right)^\ell \\
&\leq \frac{1}{\sqrt{2\pi}} \cdot \left(1 - \frac{1}{n}\right)^{-(1-\theta_3)} \cdot \frac{\left(\frac{\tau\mathrm{e}}{h(n)}\right)^{h(n)+1} - \left(\frac{\tau\mathrm{e}}{h(n)}\right)^n}{1 - \frac{\tau\mathrm{e}}{h(n)}}
\end{aligned}
$$

for any $\theta_3 \in \left[\frac{1}{2}, 1\right]$ since $\ell^{-1/2} \leq \ell^{-(1-\theta_3)}$ and

$$
\begin{aligned}
&\sum_{\ell=h(n)+1}^{n-1} \frac{n^{1-\theta_3}}{(n-\ell)^{1-\theta_3}} \cdot \frac{\tau^\ell}{\ell!} \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \frac{n^{1-\theta_3}}{(n-\ell)^{1-\theta_3}} \cdot \frac{1}{\sqrt{\ell}}\left(\frac{\tau\mathrm{e}}{\ell}\right)^\ell \\
&= \frac{1}{\sqrt{2\pi}} \sum_{\ell=h(n)+1}^{n-1} \left(\frac{n}{\ell(n-\ell)}\right)^{1-\theta_3} \cdot \frac{\tau\mathrm{e}}{\ell^{1/2+\theta_3}}\left(\frac{\tau\mathrm{e}}{\ell}\right)^{\ell-1} \\
&\leq \frac{\tau\mathrm{e}}{\sqrt{2\pi}} \cdot \left(1 - \frac{1}{n}\right)^{-(1-\theta_3)} \frac{\left(\frac{\tau\mathrm{e}}{h(n)}\right)^{h(n)} - \left(\frac{\tau\mathrm{e}}{h(n)}\right)^{n-1}}{1 - \frac{\tau\mathrm{e}}{h(n)}}
\end{aligned}
$$

for any $\theta_3 \in \left(0, \frac{1}{2}\right)$. The desired result is then obtained by letting $n \to \infty$ and applying the sandwich theorem. $\quad\square$

$$\Psi_n(\theta_1,\theta_2,\theta_3) = (n!)^2 \cdot \theta_2^n \cdot \frac{\tau^n}{n!} + (n!)^2 \cdot \theta_2^n \cdot \sum_{\ell=0}^{n-1} \frac{\Gamma(n-\ell+\theta_3)}{\Gamma(n-\ell+1)\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!}$$

$$\geq (n!)^2 \cdot \theta_2^n \cdot \frac{\tau^n}{n!} \cdot \frac{(n-n+\theta_3)^{\theta_3-1}}{\Gamma(\theta_3)} + (n!)^2 \cdot \theta_2^n \sum_{\ell=0}^{n-1} \frac{(n-\ell+\theta_3)^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!}$$

$$\geq (n!)^2 \cdot \theta_2^n \cdot \frac{(n+\theta_3)^{\theta_3-1}}{\Gamma(\theta_3)} \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!}$$

$$= (n!)^2 \cdot \theta_2^n \cdot \frac{n^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \left(1+\frac{\theta_3}{n}\right)^{\theta_3-1} \cdot \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!}, \tag{33}$$

$$\Psi_n(\theta_1,\theta_2,\theta_3) \leq (n!)^2 \cdot \theta_2^n \cdot \frac{\tau^n}{n!} + (n!)^2 \cdot \theta_2^n \cdot \sum_{\ell=0}^{n-1} \frac{(n-\ell)^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!}$$

$$= (n!)^2 \cdot \theta_2^n \cdot \frac{\tau^n}{n!} + (n!)^2 \cdot \theta_2^n \cdot \left(\sum_{\ell=0}^{h(n)} \frac{(n-\ell)^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n-1} \frac{(n-\ell)^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!} \right)$$

$$\leq (n!)^2 \cdot \theta_2^n \cdot \frac{\tau^n}{n!} + (n!)^2 \cdot \theta_2^n \cdot \left(\sum_{\ell=0}^{h(n)} \frac{(n-h(n))^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n-1} \frac{(n-\ell)^{\theta_3-1}}{\Gamma(1)} \cdot \frac{\tau^\ell}{\ell!} \right)$$

$$= (n!)^2 \cdot \theta_2^n \left[\frac{\tau^n}{n!} + \frac{n^{\theta_3-1}}{\Gamma(\theta_3)} \cdot \left(1-\frac{h(n)}{n}\right)^{c-1} \cdot \sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n-1} \frac{1}{(n-\ell)^{1-\theta_3}} \cdot \frac{\tau^\ell}{\ell!} \right], \tag{34}$$

$$\frac{(1+\frac{\theta_3}{n})^{\theta_3-1}}{\Gamma(\theta_3)} \sum_{\ell=0}^{n} \frac{\tau^\ell}{\ell!} \leq \frac{\Psi_n(\theta_1,\theta_2,\theta_3)}{(n!)^2\theta_2^n \cdot n^{\theta_3-1}}$$

$$\leq \frac{\tau^n \cdot n^{1-\theta_3}}{n!} + \frac{\left(1-\frac{h(n)}{n}\right)^{\theta_3-1}}{\Gamma(\theta_3)} \sum_{\ell=0}^{h(n)} \frac{\tau^\ell}{\ell!} + \sum_{\ell=h(n)+1}^{n-1} \frac{n^{1-\theta_3}}{(n-\ell)^{1-\theta_3}} \cdot \frac{\tau^\ell}{\ell!}. \tag{35}$$

Corollary 3. *Given Assumption 1, it holds that*

$$\mathbb{E}\big[\mathrm{perm}(\mathbf{A})^2\big] = \Psi_n(\mu_2, \mu_1^2, 1),$$

$$\mathbb{E}\big[\mathrm{perm}_{\mathrm{B},2}(\mathbf{A})^2\big] = \Psi_n\left(\mu_2, \mu_1^2, \frac{1}{2}\right),$$

$$\mathrm{perm}(\mathbf{1}_{n\times n})^2 = \Psi_n(1,1,1),$$

$$\mathrm{perm}_{\mathrm{B},2}(\mathbf{1}_{n\times n})^2 = \Psi_n\left(1,1,\frac{1}{2}\right).$$

Proof. These expressions follow from expressing the quantities on the left-hand side in terms of Bell polynomials and using Lemma 9. □

REFERENCES

[1] H. Minc, *Permanents*. Reading, MA: Addison-Wesley, 1978.

[2] P. O. Vontobel, "The Bethe permanent of a nonnegative matrix," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1866–1901, 2013.

[3] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

[4] P. O. Vontobel, "Counting in graph covers: A combinatorial characterization of the Bethe entropy function," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 6018–6048, 2013.

[5] M. Chertkov, L. Kroc, and M. Vergassola, "Belief propagation and beyond for particle tracking," *CoRR, available online under* http://arxiv.org/abs/0806.1199, Jun. 2008.

[6] B. Huang and T. Jebara, "Approximating the permanent with belief propagation," *CoRR, available online under* http://arxiv.org/abs/0908.1769, Aug. 2009.

[7] Y. Watanabe and M. Chertkov, "Belief propagation and loop calculus for the permanent of a non-negative matrix," *J. Phys. A: Math. and Theor.*, vol. 43, p. 242002, 2010.

[8] P. O. Vontobel, "The Bethe approximation of the pattern maximum likelihood distribution," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, 1–6 Jul 2012, pp. 2012–2016.

[9] ——, "The Bethe and Sinkhorn approximations of the pattern maximum likelihood estimate and their connections to the Valiant–Valiant estimate," in *Proc. Inf. Theory and Appl. Workshop*, UC San Diego, La Jolla, CA, USA, 9–14 Feb 2014, pp. 1–10.

[10] J. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2942–2959, Oct. 2014.

[11] M. Chertkov and A. B. Yedidia, "Approximating the permanent with fractional belief propagation," *J. Mach. Learn. Res.*, vol. 14, no. 26, pp. 2029–2066, 2013.

[12] M. Schwartz and P. O. Vontobel, "Improved lower bounds on the size of balls over permutations with the infinity metric," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6227–6239, Oct 2017.

[13] L. Gurvits, "Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation," *Elec. Coll. Comp. Compl.*, Dec. 2011.

[14] L. Gurvits and A. Samorodnitsky, "Bounds on the permanent and some applications," in *Proc. IEEE Symp. Found. Comp. Sci.*, Philadelphia, PA, USA, Oct. 18–21 2014.

[15] N. Anari and A. Rezaei, "A tight analysis of Bethe approximation for permanent," in *Proc. IEEE Ann. Symp. on Found. Comp. Sci.*, 2019, pp.

1434–1445.

[16] D. Straszak and N. K. Vishnoi, "Belief propagation, Bethe approximation, and polynomials," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4353–4363, Jul. 2019.

[17] N. Anari, M. Charikar, K. Shiragur, and A. Sidford, "The Bethe and Sinkhorn permanents of low rank matrices and implications for profile maximum likelihood," in *Proc. 34th Conf. Learn. Theory*, ser. Proc. Mach. Learn. Res., M. Belkin and S. Kpotufe, Eds., vol. 134, 15–19 Aug 2021, pp. 93–158.

[18] P. O. Vontobel, "Analysis of double covers of factor graphs," in *Proc. Int. Conf. Sig. Proc. and Comm.*, Bangalore, India, 12–15 Jun 2016, pp. 1–5.

[19] N. Ruozzi, "The Bethe partition function of log-supermodular graphical models," in *Proc. Neural Inf. Proc. Sys. Conf.*, Lake Tahoe, NV, USA, Dec. 3–6 2012.

[20] P. Csikvári, N. Ruozzi, and S. Shams, "Markov random fields, homomorphism counting, and Sidorenko's conjecture," *IEEE Trans. Inf. Theory*, to appear.

[21] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[22] G. D. Forney, Jr., "Codes on graphs: normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.

[23] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Sig. Proc. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[24] N. L. Biggs, *Discrete Mathematics*, 2nd ed. New York: The Clarendon Press and Oxford University Press, 1989.

[25] H. S. Wilf, *Generatingfunctionology*, 3rd ed. A. K. Peters / CRC Press, 2005.

[26] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Dordrecht: Springer Netherlands, 1974.

[27] H. Robbins, "A remark on Stirling's formula," *The Amer. Math. Monthly*, vol. 62, no. 1, pp. 26–29, 1955. [Online]. Available: http://www.jstor.org/stable/2308012

[28] J. G. Wendel, "Note on the Gamma function," *The Amer. Math. Monthly*, vol. 55, no. 9, pp. 563–564, 1948. [Online]. Available: http://www.jstor.org/stable/2304460