Quantum-inspired permanent identities

Ulysse Chabaud[1], Abhinav Deshpande[1], and Saeed Mehraban[2]

[1]Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA
[2]Computer Science, Tufts University, Medford, MA 02155, USA

The permanent is pivotal to both complexity theory and combinatorics. In quantum computing, the permanent appears in the expression of output amplitudes of linear optical computations, such as in the Boson Sampling model. Taking advantage of this connection, we give quantum-inspired proofs of many existing as well as new remarkable permanent identities. Most notably, we give a quantum-inspired proof of the MacMahon master theorem as well as proofs for new generalizations of this theorem. Previous proofs of this theorem used completely different ideas. Beyond their purely combinatorial applications, our results demonstrate the classical hardness of exact and approximate sampling of linear optical quantum computations with input cat states.

Contents

uchabaud@caltech.edu
saeed.mehraban@tufts.edu

1 Introduction

The *permanent* of an $m \times m$ matrix $A = (a_{ij})_{1 \leq i,j \leq m}$ is a combinatorial function defined as [1, 2]:

$$\mathrm{Per}(A) = \sum_{\sigma \in \mathcal{S}_m} \prod_{k=1}^{m} a_{k\sigma(k)}, \tag{1}$$

where $\mathcal{S}_m$ is the symmetric group over m symbols. While the closely related determinant can be computed efficiently using many methods such as Gaussian elimination, Valiant famously proved the #P-hardness of computing the permanent [3].

Interestingly, the permanent appears in the expression of the output amplitudes of linear optical quantum computations with noninteracting bosons [4, 5], as in the Boson Sampling model of quantum computation [6]. This connection has lead to several linear optical proofs of existing and new classical complexity results: computation of the permanent is #P-hard [7], (inverse polynomial) multiplicative estimation of the permanent of positive semidefinite matrices is in $\mathsf{BPP}^{\mathsf{NP}}$ [8], multiplicative estimation of the permanent of orthogonal matrices is #P-hard [9], and computation of a class of multidimensional integrals is #P-hard [10]. It has also lead to the introduction of a quantum-inspired classical algorithm for additive estimation of the permanent of positive semidefinite matrices [11]. Moreover, an approach inspired by quantum tomography recently showed that (subexponential) multiplicative estimation of the permanent of positive semidefinite matrices is NP-hard [12].

Beyond its importance for complexity theory, the permanent has numerous applications for solving combinatorial problems [1, 2] and identities for the permanent have been instrumental in these applications. For example, the MacMahon master theorem [13], which relates the permanent to a coefficient of the Taylor series of a determinant, is an invaluable tool for proving combinatorial identities [14–16]. Similarly, Ryser's formula [17] and Glynn's formula [18–20] are routinely used to compute the permanent more efficiently than the naive brute-force approach.

While linear optics has been used as a tool to explore the classical complexity of the permanent, previous work suggests that it can also be a useful way to obtain simple proofs of theorems about the permanent: for instance, [5] shows that the permanent of a unitary matrix U lies in the (closed) complex unit disk by expressing $|\mathrm{Per}(U)|^2$ as an output probability of a linear optical sampling computation, while [6] derives simple permanent identities using various representations of the same linear optical sampling computation. This begs the following questions:

1. Can we use linear optics to prove existing remarkable permanent identities, such as the MacMahon master theorem?
2. Can we use linear optics to derive *new* remarkable permanent identities?

In this work, we show that the answer to both questions is *yes*.

We give quantum-inspired proofs of several important permanent identities in section 5. In particular, we show that the MacMahon master theorem can be understood as two different ways of computing an inner product between two Gaussian quantum states.

We also derive new quantum-inspired identities for the permanent. Our main results are summarized in section 3 and proven in section 6. These include generalizations of the MacMahon master theorem (Theorems 1 and 2), new generating functions for the permanent (Theorem 3) and a formula for the sum of two permanents (Theorem 4).

As a bonus, our findings also have consequences for the classical complexity of exact and approximate sampling of linear optical sampling computations with input cat states

(Theorem 5), which we discuss in section 7. We rigorously prove that the corresponding quantum probability distributions are as hard to sample as the original Boson Sampling distribution [6], for all cat state amplitudes in the exact case and for small enough amplitudes in the approximate case. Until now, a formal proof was available only in the exact case [21].

2 Background

2.1 Notations and preliminary material

$\mathbf{0} = (0, \dots, 0)$	$\boldsymbol{p} \leq \boldsymbol{q} \Leftrightarrow \forall k \in \{1, \dots, m\},\ p_k \leq q_k$
$\mathbf{1} = (1, \dots, 1)$	$-\boldsymbol{z} = (-z_1, \dots, -z_m)$
$\lvert\boldsymbol{p}\rvert = p_1 + \cdots + p_m$	$\boldsymbol{z}^* = (z_1^*, \dots, z_m^*)$
$\boldsymbol{p}! = p_1! \dots p_m!$	$\lVert\boldsymbol{z}\rVert^2 = \lvert z_1\rvert^2 + \cdots + \lvert z_m\rvert^2$
$\lvert\boldsymbol{p}\rangle = \lvert p_1\rangle \otimes \cdots \otimes \lvert p_m\rangle$	$\boldsymbol{z}^{\boldsymbol{p}} = z_1^{p_1} \dots z_m^{p_m}$
$\boldsymbol{p} + \boldsymbol{q} = (p_1 + q_1, \dots, p_m + q_m)$	$\partial_{\boldsymbol{z}}^{\boldsymbol{p}} = \partial_{z_1}^{p_1} \dots \partial_{z_m}^{p_m}$
$\boldsymbol{p} \oplus \boldsymbol{q} = (p_1, \dots, p_m, q_1, \dots, q_m)$	$d^m\boldsymbol{z} d^m\boldsymbol{z}^* = d\operatorname{Re}\{z_1\} d\operatorname{Im}\{z_1\} \dots d\operatorname{Re}\{z_m\} d\operatorname{Im}\{z_m\}$

Table 1: Multi-index notations used in this paper, for $m \in \mathbb{N}^*$, $\boldsymbol{p} = (p_1, \dots, p_m) \in \mathbb{N}^m$, $\boldsymbol{q} = (q_1, \dots, q_m) \in \mathbb{N}^m$ and $\boldsymbol{z} = (z_1, \dots, z_m) \in \mathbb{C}^m$.

We use bold math for multi-index expressions (see Table 1 above). We denote by $\mathbb{T} = \{z \in \mathbb{C}, |z| = 1\}$ the complex unit circle. For all $m \in \mathbb{N}^*$, all $\boldsymbol{z} = (z_1, \dots, z_m) \in \mathbb{C}^m$ and all $\boldsymbol{p} = (p_1, \dots, p_m) \in \mathbb{N}^m$, we use the notation $[\boldsymbol{z}^{\boldsymbol{p}}]$ to denote the coefficient of $\boldsymbol{z}^{\boldsymbol{p}} = z_1^{p_1} \dots z_m^{p_m}$ in an analytic expression. For all $m \in \mathbb{N}^*$, all $\boldsymbol{p} = (p_1, \dots, p_m) \in \mathbb{N}^m$ and all $\boldsymbol{q} = (q_1, \dots, q_m) \in \mathbb{N}^m$, we denote by $A_{\boldsymbol{p},\boldsymbol{q}}$ the matrix obtained from an $m \times m$ matrix A by first repeating its i^{th} row p_i times (deleting the row if $p_i = 0$) for all $i \in \{1, \dots, m\}$ and then repeating its j^{th} column q_j times (deleting the column if $q_j = 0$) for all $j \in \{1, \dots, m\}$. By convention, we set the permanent of non-square matrices to 0 and $\operatorname{Per}(A_{\mathbf{0},\mathbf{0}}) = 1$.

Let $m \in \mathbb{N}^*$ denote the number of modes. We denote (unnormalized) quantum states using the Dirac ket notation. These are vectors in an infinite-dimensional Hilbert space spanned by the orthonormal Fock basis

$$\{\lvert\boldsymbol{p}\rangle = \lvert p_1\rangle \otimes \cdots \otimes \lvert p_m\rangle\}_{\boldsymbol{p}=(p_1,\dots,p_m)\in\mathbb{N}^m}. \tag{2}$$

Hereafter, we label Fock states using p, q, $\boldsymbol{p}$, $\boldsymbol{q}$, and $\boldsymbol{k}$, coherent states using α, β, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$, and two-mode squeezed states using λ, μ, $\boldsymbol{\lambda}$, and $\boldsymbol{\mu}$. In particular, coherent states are defined as

$$\lvert\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{p\geq 0} \frac{\alpha^p}{\sqrt{p!}} \lvert p\rangle\,, \tag{3}$$

for all $\alpha \in \mathbb{C}$, and (unnormalized) two-mode squeezed states as

$$\lvert\lambda\rangle = \sum_{p\geq 0} \lambda^p \lvert pp\rangle\,, \tag{4}$$

for all $\lambda \in \mathbb{C}$ with $|\lambda| < 1$. Moreover, cat states are defined as

$$\lvert\mathrm{cat}_\alpha\rangle = \frac{e^{\frac{1}{2}|\alpha|^2}}{2\sqrt{\sinh(|\alpha|^2)}}(\lvert\alpha\rangle - \lvert-\alpha\rangle), \tag{5}$$

for all $\alpha \in \mathbb{C}$.

We make use of the following inner products involving these states: for all $p \in \mathbb{N}$, $\alpha, \beta \in \mathbb{C}$, and $\lambda \in \mathbb{C}$ with $|\lambda| < 1$,

$$\langle p|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \frac{\alpha^p}{\sqrt{p!}} \tag{6}$$

$$\langle \alpha|\beta\rangle = e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \alpha^*\beta}, \tag{7}$$

$$\langle pp|\lambda\rangle = \lambda^p, \tag{8}$$

$$\langle \lambda| \left(|\alpha\rangle \otimes |\beta\rangle\right) = e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \lambda^*\alpha\beta}, \tag{9}$$

which can be readily checked in the Fock basis. Moreover, for all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ [5, 6],

$$\langle \boldsymbol{p}| \hat{U} |\boldsymbol{q}\rangle = \frac{\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}})}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}}, \tag{10}$$

where $\hat{U}$ is a passive linear operation over m modes whose action on the creation operators of the modes is described by the unitary matrix $U = (u_{ij})_{1 \le i,j \le m}$ as

$$\hat{U} \hat{a}_j^\dagger \hat{U}^\dagger = \sum_{i=1}^m u_{ij} \hat{a}_i^\dagger, \tag{11}$$

for all $j \in \{1, \dots, m\}$. As their name indicates, passive linear operations do not change the total number of photons [22]: for all $n \in \mathbb{N}$,

$$\hat{U} \Pi_n = \Pi_n \hat{U}, \tag{12}$$

where $\Pi_n := \sum_{|\boldsymbol{p}|=n} |\boldsymbol{p}\rangle\langle\boldsymbol{p}|$ is the m-mode projector onto states with total photon number equal to n. Passive linear operations map tensor products of coherent states to tensor products of coherent states [22]: for all $\boldsymbol{\alpha} \in \mathbb{C}^m$,

$$\hat{U} |\boldsymbol{\alpha}\rangle = |U\boldsymbol{\alpha}\rangle \,. \tag{13}$$

Recall that coherent states form an overcomplete basis:

$$\int_{\boldsymbol{\alpha} \in \mathbb{C}^m} |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \frac{d^m\boldsymbol{\alpha} d^m\boldsymbol{\alpha}^*}{\pi^m} = \hat{\mathbb{I}}, \tag{14}$$

where $\hat{\mathbb{I}}$ is the identity operator over m modes. We will also make use of the following Gaussian inner product: for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{C}^m$ with $|\lambda_k| < 1$ and $|\mu_k| < 1$ for all $k \in \{1, \dots, m\}$, and for all passive linear operation $\hat{U}$ over $2m$ modes,

$$\begin{aligned}
\langle \boldsymbol{\lambda}^*| \hat{U} |\boldsymbol{\mu}\rangle &= \int_{\boldsymbol{\beta} \in \mathbb{C}^{2m}} \langle \boldsymbol{\lambda}^*|\boldsymbol{\beta}^*\rangle\langle\boldsymbol{\beta}^*| \hat{U} |\boldsymbol{\mu}\rangle \frac{d^{2m}\boldsymbol{\beta} d^{2m}\boldsymbol{\beta}^*}{\pi^{2m}} \\
&= \int_{\boldsymbol{\beta} \in \mathbb{C}^{2m}} \langle \boldsymbol{\lambda}^*|\boldsymbol{\beta}^*\rangle \left(\langle\boldsymbol{\mu}| U^\dagger \boldsymbol{\beta}^*\rangle\right)^* \frac{d^{2m}\boldsymbol{\beta} d^{2m}\boldsymbol{\beta}^*}{\pi^{2m}} \\
&= \int_{\boldsymbol{\beta} \in \mathbb{C}^{2m}} e^{\sum_{k=1}^m \lambda_k \beta_k^* \beta_{m+k}^*} e^{\sum_{k=1}^m \mu_k (U^T\boldsymbol{\beta})_k (U^T\boldsymbol{\beta})_{m+k}} e^{-\sum_{j=1}^{2m} \beta_j^* \beta_j} \frac{d^{2m}\boldsymbol{\beta} d^{2m}\boldsymbol{\beta}^*}{\pi^{2m}} \\
&= \int_{\boldsymbol{\beta} \in \mathbb{C}^{2m}} \exp\left[-\frac{1}{2} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^* \end{pmatrix}^T V_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^* \end{pmatrix}\right] \frac{d^{2m}\boldsymbol{\beta} d^{2m}\boldsymbol{\beta}^*}{\pi^{2m}} \\
&= \frac{1}{\sqrt{\mathrm{Det}(V_U(\boldsymbol{\lambda}, \boldsymbol{\mu}))}},
\end{aligned} \tag{15}$$

where we have used the overcompleteness of coherent states (14) over $2m$ modes in the first line, the action of passive linear operations on coherent states (13) in the second line, the overlap between two-mode squeezed states and coherent states (9) in the third line, and where we have introduced in the fourth line the $(4m) \times (4m)$ symmetric matrix

$$V_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \begin{pmatrix} -UV_{\boldsymbol{\mu}}U^T & I_{2m} \\ I_{2m} & -V_{\boldsymbol{\lambda}} \end{pmatrix}, \tag{16}$$

where for all $\boldsymbol{w} \in \mathbb{C}^n$,

$$V_{\boldsymbol{w}} := \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{w}) \\ \mathrm{Diag}(\boldsymbol{w}) & 0_m \end{pmatrix}. \tag{17}$$

Note that we associate mode 1 with mode $m+1$, mode 2 with mode $m+2$, and so on, i.e. $|\boldsymbol{\lambda}\rangle = \bigotimes_{k=1}^m |\lambda_k\rangle$, where $|\lambda_k\rangle$ is an unormalized two-mode squeezed state over modes k and $m+k$.

Finally, we note that any $m \times m$ matrix B with $\|B\| \leq 1$, where $\|\cdot\|$ is the spectral norm, may be embedded as the top-left submatrix of a $(2m) \times (2m)$ unitary matrix U (see [6, Lemma 29] for an explicit construction). We will use this fact multiple times throughout the paper to extend identities proven for unitary matrices to the case of generic matrices.

2.2 Boson Sampling

Boson Sampling is a sub-universal model of quantum computation introduced by Aaronson and Arkhipov (AA) in [6], which takes as input a Fock state $|\mathbf{1} \oplus \mathbf{0}\rangle = |1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}$, evolves it according to a passive linear operation $\hat{U}$ over m modes with unitary matrix U, and measures the photon number of each output mode.

With Eq. (10), the probability of detecting $\boldsymbol{p} = (p_1, \ldots, p_m) \in \mathbb{N}^m$ output photons is given by:

$$\begin{aligned} P_{\mathrm{BS}}(\boldsymbol{p}|n) :=& |\langle p_1 \ldots p_m | \hat{U} (|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n})|^2 \\ =& \frac{|\mathrm{Per}(U_{\boldsymbol{p},\mathbf{1}\oplus\mathbf{0}})|^2}{\boldsymbol{p}!}. \end{aligned} \tag{18}$$

This model of quantum computation, while not believed to be universal, is already capable of outperforming its classical counterparts: AA showed that the output probability distribution P_{BS} is hard to sample exactly classically for $m \geq 2n$, or the polynomial hierarchy of complexity classes collapses to its third level [6].

Moreover, under additional plausible conjectures, AA showed that this collapse holds for $m = \Theta(n^5 \log^2 n)$ even if only an efficient classical algorithm for *approximate sampling* exists (i.e. a classical algorithm which samples efficiently from a probability distribution that has a small total variation distance with the ideal Boson Sampling output probability distribution P_{BS}). They further conjectured that this approximate hardness should hold for $m = \Theta(n^2)$.

The approximate hardness proof for Boson Sampling is based on a matrix-hiding argument which requires a specific regime of parameters m and n: given $\delta > 0$, the relation between m and n should be such that the distribution $\mathcal{H}^{n,m}$ of $n \times n$ submatrices of $m \times m$ Haar-random unitary matrices multiplied by $\sqrt{m}$ is $O(\delta)$-close in total variation distance to the distribution $\mathcal{G}^{n \times n}$ of $n \times n$ matrices of i.i.d. Gaussians (see Theorem 35 in [6]). In particular, AA showed that, for any $\delta > 0$, $\|\mathcal{H}^{n,m} - \mathcal{G}^{n \times n}\|_{\mathrm{TV}} = O(\delta)$ when $m \geq \frac{n^5}{\delta} \log^2 \frac{n}{\delta}$.

This result was later refined in [23] with the bound $\|\mathcal{H}^{n,m} - \mathcal{G}^{n \times n}\|_{\mathrm{TV}} \leq \frac{n^3}{2(m-n)}$, which gives a total variation distance $O(\delta)$ for $m \geq \frac{n^3}{\delta}$.

Finally, in the case of orthogonal matrices rather than unitary matrices, it was shown in [24] that (see the equation following Eq. (2.31) in [24], with $n \to m$ and $p = q \to n$):

$$\begin{aligned} D_{\mathrm{KL}}(\mathcal{H}^{n,m}, \mathcal{G}^{n\times n}) &\leq \frac{n^3}{4(m-n)} - \frac{(m-n)-(n+1)}{2}\left[\frac{n^3+O(n^2)}{2(m-n)^2} + \frac{n^4+O(n^3)}{3(m-n)^3}\right] \\ &\leq O\left(\frac{n^2}{m} + \frac{n^4}{m^2} + \frac{n^5}{m^3}\right), \end{aligned} \tag{19}$$

where D_{KL} is the Kullback–Leibler distance and where we obtain the second line under the assumption $n = o(m)$. In particular, for $\delta > 0$, choosing $m \geq \frac{n^2}{\delta^2}$ implies $D_{\mathrm{KL}}(\mathcal{H}^{n,m}, \mathcal{G}^{n\times n}) \leq O(\delta^2)$, and thus $\|\mathcal{H}^{n,m} - \mathcal{G}^{n\times n}\|_{\mathrm{TV}} \leq O(\delta)$ by Pinsker's inequality (see Eq. (1.3) in [24]). It was also argued in [24] that a similar result should hold for unitary matrices.

Summarising these results, approximate hardness of Boson Sampling is proven for passive linear operations described by orthogonal matrices in the regime $m = \Theta(n^2)$ [24] and for unitary matrices in the regime $m = \Theta(n^3)$ [23].

2.3 The MacMahon master theorem

The MacMahon master theorem is an important result in combinatorics which relates the permanent to the determinant:

Theorem (MacMahon master theorem [13])**.** *Let $\boldsymbol{z} = (z_1, \ldots, z_m)$ be formal variables. For any $m \times m$ matrix A,*

$$\sum_{\boldsymbol{p}\in\mathbb{N}^m} \frac{\boldsymbol{z}^{\boldsymbol{p}}}{\boldsymbol{p}!}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \frac{1}{\mathrm{Det}(I - ZA)}, \tag{20}$$

where $Z = \mathrm{Diag}(\boldsymbol{z})$.

This theorem is particularly useful to derive short proofs of combinatorial identities: it expresses the permanent of an $m \times m$ matrix A with rows and columns repeated in the same way as the coefficient

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \boldsymbol{p}![\boldsymbol{z}^{\boldsymbol{p}}]\left(\frac{1}{\mathrm{Det}(I - ZA)}\right), \tag{21}$$

where $Z = \mathrm{Diag}(\boldsymbol{z})$ and $\boldsymbol{p} \in \mathbb{N}^m$, while the same permanent may also be expressed as the coefficient

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \boldsymbol{p}![\boldsymbol{z}^{\boldsymbol{p}}](A\boldsymbol{z})^{\boldsymbol{p}}. \tag{22}$$

Picking a specific matrix A and a pattern $\boldsymbol{p}$ and computing the above expressions yields combinatorial identities, a famous example being the short proof of Dixon's identity [14, 25], $\sum_{k=0}^{2n}(-1)^k\binom{2n}{k}^3 = (-1)^n\binom{3n}{n,n,n}$, by taking

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, \tag{23}$$

and $\boldsymbol{p} = (2n, 2n, 2n)$ for $n \in \mathbb{N}^*$.

Various generalizations of the MacMahon master theorem have been introduced over the years [26–30]. In physics, this theorem plays an important role in the quantum theory of angular momentum [31] and is also oftentimes interpreted as an instance of the boson-fermion correspondence [27].

3 Main results

In this section, we summarize our main findings, which we prove in section 6. We obtain the following generalization of the MacMahon master theorem:

Theorem 1. *Let $\boldsymbol{x} = (x_1, \dots, x_m)$ and $\boldsymbol{y} = (y_1, \dots, y_m)$ be formal variables. For all $m \times m$ matrices A and B,*

$$\sum_{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m} \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) \mathrm{Per}(B_{\boldsymbol{q},\boldsymbol{p}}) = \frac{1}{\mathrm{Det}(I - XAYB)}, \tag{24}$$

where $X = \mathrm{Diag}(\boldsymbol{x})$ and $Y = \mathrm{Diag}(\boldsymbol{y})$.

We further show that this generalization extends to N matrices:

Theorem 2. *Let $N \geq 2$. For all $k \in \{1, \dots, N\}$, let $\boldsymbol{z}_k = (z_{k1}, \dots, z_{km})$ be formal variables. For all $m \times m$ matrices $A^{(1)}, \dots, A^{(N)}$,*

$$\begin{aligned} \sum_{\boldsymbol{p}_1, \dots, \boldsymbol{p}_N \in \mathbb{N}^m} \prod_{k=1}^{N} \frac{\boldsymbol{z}_k^{\boldsymbol{p}_k}}{\boldsymbol{p}_k!} \mathrm{Per}(A^{(1)}_{\boldsymbol{p}_1,\boldsymbol{p}_2}) \mathrm{Per}(A^{(2)}_{\boldsymbol{p}_2,\boldsymbol{p}_3}) \dots \mathrm{Per}(A^{(N)}_{\boldsymbol{p}_N,\boldsymbol{p}_1}) \\ = \frac{1}{\mathrm{Det}(I - Z_1 A^{(1)} \dots Z_N A^{(N)})}, \end{aligned} \tag{25}$$

where $Z_k = \mathrm{Diag}(\boldsymbol{z}_k)$ for all $k \in \{1, \dots, N\}$.

As a corollary of Theorem 1, we obtain:

Corollary 1. *For any $m \times m$ matrix A and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n \in \mathbb{N}$,*

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{\boldsymbol{p}!\boldsymbol{q}!}{n!} [\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}] \left(\boldsymbol{x}^T A \boldsymbol{y}\right)^n. \tag{26}$$

As a consequence of Corollary 1, we obtain the following family of generating functions for the permanent:

Theorem 3. *Let $f(z) = \sum_{n=0}^{+\infty} f_n z^n$ be a series. Let $\boldsymbol{x} = (x_1, \dots, x_m)$ and $\boldsymbol{y} = (y_1, \dots, y_m)$ be formal variables. For any $m \times m$ matrix A,*

$$f(\boldsymbol{x}^T A \boldsymbol{y}) = \sum_{\substack{n \in \mathbb{N}, \boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m \\ |\boldsymbol{p}| = |\boldsymbol{q}| = n}} f_n n! \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}), \tag{27}$$

Equivalently,

$$\partial_{\boldsymbol{x}}^{\boldsymbol{p}} \partial_{\boldsymbol{y}}^{\boldsymbol{q}} f(\boldsymbol{x}^T A \boldsymbol{y})\big|_{\boldsymbol{x}=\boldsymbol{y}=\boldsymbol{0}} = \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}})\, \partial_z^n f(z)\big|_{z=0}, \tag{28}$$

for all $n \in \mathbb{N}$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ such that $|\boldsymbol{p}| = |\boldsymbol{q}| = n$. As a result, when $f_n \neq 0$,

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y} \in \mathbb{T}^m} \left[\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}} \frac{f(\boldsymbol{x}^T A \boldsymbol{y})}{\partial_z^n f(z)\big|_{z=0}} \right], \tag{29}$$

where the average is over random vectors with complex coefficients of modulus 1.

From this theorem we derive various permanent identities, the most notable one being a formula for the sum of two permanents:

Theorem 4. *For all $m \times m$ matrices A and B, all $n \in \mathbb{N}$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ such that $|\boldsymbol{p}| = |\boldsymbol{q}| = n$,*

$$\begin{aligned}&\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) + \mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}}) \\ &\quad= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k}{\binom{n-1}{k}} \sum_{\substack{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{p} \\ \boldsymbol{a}'+\boldsymbol{b}'+\boldsymbol{c}'=\boldsymbol{q} \\ |\boldsymbol{a}|=|\boldsymbol{b}|=|\boldsymbol{a}'|=|\boldsymbol{b}'|=k}} \frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{a}!\boldsymbol{b}!\boldsymbol{c}!\boldsymbol{a}'!\boldsymbol{b}'!\boldsymbol{c}'!} \mathrm{Per}(A_{\boldsymbol{a},\boldsymbol{a}'})\mathrm{Per}(B_{\boldsymbol{b},\boldsymbol{b}'})\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'}). \end{aligned} \tag{30}$$

In particular, when $\boldsymbol{p} = \boldsymbol{q} = \mathbf{1}$,

$$\mathrm{Per}(A) + \mathrm{Per}(B) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k}{\binom{n-1}{k}} \sum_{\substack{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\mathbf{1} \\ \boldsymbol{a}'+\boldsymbol{b}'+\boldsymbol{c}'=\mathbf{1} \\ |\boldsymbol{a}|=|\boldsymbol{b}|=|\boldsymbol{a}'|=|\boldsymbol{b}'|=k}} \mathrm{Per}(A_{\boldsymbol{a},\boldsymbol{a}'})\mathrm{Per}(B_{\boldsymbol{b},\boldsymbol{b}'})\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'}). \tag{31}$$

Section 5 is primarily devoted to quantum-inspired proofs of existing permanent identities, including a new proof of the MacMahon master theorem. Along the way, we obtain the following inner product formula:

Lemma 1. *For all $\alpha \in \mathbb{C}$, all $\boldsymbol{p} \in \mathbb{N}^m$, all $n \leq m$, and any passive linear operation $\hat{U}$ over m modes with unitary matrix U,*

$$\begin{aligned}\langle \boldsymbol{p}|\hat{U}(|\mathrm{cat}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}) &= \frac{\alpha^n}{\sqrt{\sinh^n(|\alpha|^2)}} \langle \boldsymbol{p}|\hat{U}(|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}) \\ &= \frac{\alpha^n}{\sqrt{\sinh^n(|\alpha|^2)}} \frac{\mathrm{Per}(U_{\boldsymbol{p},\mathbf{1}\oplus\mathbf{0}})}{\sqrt{\boldsymbol{p}!}}. \end{aligned} \tag{32}$$

Lemma 1 has the following direct implications for the hardness of Boson Sampling with input cat states [21], which we discuss in section 7:

Theorem 5. *Let $m = \mathrm{poly}\, n \geq 2n$ and $\alpha \neq 0$. Boson Sampling with input $|\mathrm{cat}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}$ with $\alpha \in \mathbb{C}$ is hard to sample exactly classically unless the polynomial hierarchy of complexity classes collapses to its third level.*

Moreover, assuming $|\alpha| = \mathcal{O}(n^{-1/4}\log^{1/4} m)$ Boson Sampling with input cat states is as hard to sample approximately as Boson Sampling with input single-photons, i.e. Boson Sampling with input cat states is hard to sample approximately classically unless the polynomial hierarchy of complexity classes collapses to its third level, in the same regime as Boson Sampling, modulo the complexity conjectures introduced by AA in [6].

This result extends the arguments of [21]—which gave formal proof of classical hardness in the exact sampling case—to the approximate sampling case.

4 Discussion

In the next sections, we introduce quantum-inspired proofs of permanent identities. This approach allows us to give quantum-operational interpretations of seminal results, such as the MacMahon master theorem [13]. In particular, we show that this theorem can be seen as two facets of the same bosonic Gaussian amplitude.

This quantum-inspired approach also yields a breadth of new permanent identities. We give some examples of combinatorial applications of these identities in section 6 and we anticipate that they have many more. Beyond these purely combinatorial applications, it

would be interesting to investigate whether our new identities may be used to obtain more efficient classical algorithms for computing or estimating the permanent. In particular, our Theorem 3 provides new estimators for the permanent which could be of interest, by minimizing the variance of these estimators over the choice of the analytic function f.

We use the formalism of linear optics with noninteracting bosons, but our approach can be applied more generally to linear optics with other types of particles. For instance, we expect our proof techniques to lead to determinant identities in the fermionic case [32], immanant identities in the case of partially distinguishable particles [33, 34], and additional permanent identities in the case of generalized bosons [35]. Moreover, graphical languages are currently being developed for linear optical quantum computations [36, 37], which could lead to graphical proofs of remarkable identities in combination with our approach.

Finally, our Theorem 5 gives solid complexity-theoretic foundations for the hardness of Boson Sampling with input cat states. Generation of such states has progressed tremendously in the recent years, thanks to circuit QED in particular [38–40]. We hope that these foundations will motivate an experimental demonstration of quantum speedup based on Schrödinger's cat states.

5 Quantum-inspired proofs of permanent identities

In this section, we derive quantum-inspired proofs of existing permanent identities. Along the way, we obtain a generalization of Glynn's formula [20] for the permanent of matrices with repeated rows and columns in Eq. (43), as well as a generalization of the Glynn–Kan formula [41] for the permanent of matrices with repeated rows and columns in Eq. (54).

5.1 Glynn's formula

Glynn's formula for the permanent of an $m \times m$ matrix $A = (a_{ij})_{1\leq i,j\leq m}$ is [20]:

$$\mathrm{Per}(A) = \frac{1}{2^{m-1}} \sum_{\substack{\boldsymbol{x}\in\{-1,1\}^m \\ x_1=1}} x_1 \dots x_m \prod_{i=1}^{m} \left(\sum_{j=1}^{m} a_{ij} x_j \right). \tag{33}$$

By symmetry, it is equivalent to the identity

$$\mathrm{Per}(A) = \frac{1}{2^{m}} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \prod_{i=1}^{m} \left(\sum_{j=1}^{m} a_{ij} x_j \right). \tag{34}$$

Proof of Glynn's formula. To prove this identity using quantum-mechanical tools, let us introduce the unnormalized cat state $|\tilde{\mathrm{cat}}_\alpha\rangle := \frac{1}{2\alpha}(|\alpha\rangle - |-\alpha\rangle)$, for $\alpha \in \mathbb{C}$. Using the Fock basis expansion of coherent states (6), we have $\lim_{\alpha\to 0} |\tilde{\mathrm{cat}}_\alpha\rangle = |1\rangle$ in trace distance. Hence, using the Fock basis expansion of $\hat{U}$ (10),

$$\mathrm{Per}(U) = \lim_{\alpha\to 0} \langle 1 \dots 1 | \hat{U} | \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha \rangle, \tag{35}$$

where $\hat{U}$ is a passive linear operation over m modes with unitary matrix $U = (u_{ij})_{1\leq i,j\leq m}$. We compute the right hand side of this equation:

$$\begin{aligned} \langle 1 \dots 1 | \hat{U} | \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha \rangle &= \frac{1}{(2\alpha)^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \langle 1 \dots 1 | \hat{U} | x_1\alpha \dots x_m\alpha \rangle \\ &= \frac{1}{(2\alpha)^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \langle 1 \dots 1 | (\alpha U\boldsymbol{x})_1 \dots (\alpha U\boldsymbol{x})_m \rangle \end{aligned} \tag{36}$$

$$= \frac{1}{(2\alpha)^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m e^{-\frac{1}{2}|\alpha|^2\|U\boldsymbol{x}\|^2} \prod_{i=1}^m (\alpha U\boldsymbol{x})_i$$
$$= \frac{e^{-\frac{m}{2}|\alpha|^2}}{2^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \prod_{i=1}^m \left(\sum_{j=1}^m u_{ij}x_j\right),$$

where we used the action of $\hat{U}$ on coherent states (13) in the second line, the Fock basis expansion of coherent states (6) in the third line, and the fact that U is unitary in the last line. Taking the limit when $\alpha \to 0$ yields Glynn's formula for unitary matrices. □

With the same calculations, we may obtain a more general version of Eq. (36): for all $\alpha \in \mathbb{C}$ and $\boldsymbol{p} \in \mathbb{N}^m$,

$$\begin{aligned}\langle p_1 \dots p_m|\hat{U}|\tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha\rangle &= \frac{1}{(2\alpha)^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m e^{-\frac{1}{2}|\alpha|^2\|U\boldsymbol{x}\|^2} \prod_{i=1}^m \frac{[(\alpha U\boldsymbol{x})_i]^{p_i}}{\sqrt{p_i!}} \\ &= \frac{\alpha^{|\boldsymbol{p}|-m} e^{-\frac{m}{2}|\alpha|^2}}{2^m\sqrt{\boldsymbol{p}!}} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \prod_{i=1}^m \left(\sum_{j=1}^m u_{ij}x_j\right)^{p_i}.\end{aligned} \tag{37}$$

Note that when $|\boldsymbol{p}| < m$ all products in the sum have at least one x_i missing. Hence, by symmetry, $\langle p_1 \dots p_m|\hat{U}|\tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha\rangle = 0$ when $|\boldsymbol{p}| < m$. With the Fock basis expansion of $\hat{U}$ (10), taking the limit when $\alpha \to 0$ yields a version of Glynn's formula for unitary matrices with repeated rows:

$$\mathrm{Per}(U_{\boldsymbol{p},\mathbf{1}}) = \frac{\delta_{|\boldsymbol{p}|,m}}{2^m} \sum_{\boldsymbol{x}\in\{-1,1\}^m} x_1 \dots x_m \prod_{i=1}^m \left(\sum_{j=1}^m u_{ij}x_j\right)^{p_i}, \tag{38}$$

where δ is the Kronecker symbol.

So far, all identities are derived for unitary matrices U. In order to retrieve the same identity for a generic matrix $A = (a_{ij})_{1\le i,j\le n}$ of size n, we can embed $\frac{1}{\|A\|}A$ (or A directly if $A = 0$) as a submatrix of a unitary matrix U of size $m = 2n$ [6, Lemma 29] and compute:

$$\frac{\mathrm{Per}(A_{\boldsymbol{q},\mathbf{1}})}{\sqrt{\boldsymbol{q}!}} = \|A\|^{|\boldsymbol{q}|} \lim_{\alpha\to 0} (\langle q_1 \dots q_n| \otimes \langle 0|^{\otimes n})\hat{U}(|\tilde{\mathrm{cat}}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes n}). \tag{39}$$

To do so, we compute a slightly more general inner product: for $\alpha \in \mathbb{C}$, $\boldsymbol{p} \in \mathbb{N}^m$ and $n \le m$,

$$\begin{aligned}&\langle p_1 \dots p_m|\hat{U}(|\tilde{\mathrm{cat}}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}) \\ &\quad= \frac{1}{(2\alpha)^n} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1 \dots x_n \langle p_1 \dots p_m|\hat{U}|x_1\alpha \dots x_n\alpha\, 0 \dots 0\rangle \\ &\quad= \frac{1}{(2\alpha)^n} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1 \dots x_n e^{-\frac{1}{2}|\alpha|^2\|U(\boldsymbol{x}\oplus\mathbf{0})\|^2} \prod_{i=1}^m \frac{[(\alpha U(\boldsymbol{x}\oplus 0))_i]^{p_i}}{\sqrt{p_i!}} \\ &\quad= \frac{\alpha^{|\boldsymbol{p}|-n} e^{-\frac{n}{2}|\alpha|^2}}{2^n\sqrt{\boldsymbol{p}!}} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1 \dots x_n \prod_{i=1}^m \left(\sum_{j=1}^n u_{ij}x_j\right)^{p_i}.\end{aligned} \tag{40}$$

In particular, setting $m = 2n$ and $\boldsymbol{p} = \boldsymbol{q} \oplus \mathbf{0}$ for $\boldsymbol{q} \in \mathbb{N}^n$,

$$(\langle q_1 \dots q_n| \otimes \langle 0|^{\otimes n})\hat{U}(|\tilde{\mathrm{cat}}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes n}) = \frac{\alpha^{|\boldsymbol{q}|-n} e^{-\frac{n}{2}|\alpha|^2}}{2^n\sqrt{\boldsymbol{q}!}} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1 \dots x_n \prod_{i=1}^n \left(\sum_{j=1}^n u_{ij}x_j\right)^{q_i}. \tag{41}$$

Choosing $u_{ij} = \frac{1}{\|A\|}a_{ij}$ for $1 \leq i,j \leq n$ and letting α go to 0 proves the claim: the left hand side converges to $\frac{1}{\sqrt{\boldsymbol{q}!}}\mathrm{Per}(U_{\boldsymbol{q}\oplus\boldsymbol{0},\boldsymbol{1}\oplus\boldsymbol{0}}) = \frac{1}{\|A\|^{|\boldsymbol{q}|}\sqrt{\boldsymbol{q}!}}\mathrm{Per}(A_{\boldsymbol{q},\boldsymbol{1}})$, while the right hand side converges to $\frac{\delta_{|\boldsymbol{q}|,n}}{\|A\|^{|\boldsymbol{q}|}2^n\sqrt{\boldsymbol{q}!}}\sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1\dots x_n\prod_{i=1}^n\left(\sum_{j=1}^n a_{ij}x_j\right)^{q_i}$. Hence,

$$\mathrm{Per}(A_{\boldsymbol{q},\boldsymbol{1}}) = \frac{\delta_{|\boldsymbol{q}|,n}}{2^n}\sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1\dots x_n\prod_{i=1}^n\left(\sum_{j=1}^n a_{ij}x_j\right)^{q_i}. \tag{42}$$

A similar generalization of Glynn's formula for matrices with repeated rows (or columns) based on roots of unity has previously appeared in [42].

Moreover, for any $m\times m$ matrix $A = (a_{i,j})_{1\leq i,j\leq m}$ and all $\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n$, we have the formula

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{\boldsymbol{q}!}{n^m}\sum_{\boldsymbol{x}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}^m}\boldsymbol{x}^{-\boldsymbol{q}}(A\boldsymbol{x})^{\boldsymbol{p}}, \tag{43}$$

which is a version of Glynn's formula for matrices with repeated rows *and* columns. This formula, which has previously appeared in [43–45], can be easily proven by expanding the product $(A\boldsymbol{x})^{\boldsymbol{p}} = \prod_{i=1}^m\left(\sum_{j=1}^m a_{ij}x_j\right)^{p_i}$ and using properties of roots of unity. We give a quantum-inspired proof in what follows.

For all $n\geq 1$, all $q\leq n$, and all $\alpha\in\mathbb{C}$, we define the following unnormalised superposition of coherent states:

$$|\sigma_\alpha^{q,n}\rangle := \frac{\sqrt{q!}}{n\alpha^q}\sum_{k=0}^{n-1} e^{-\frac{2ikq\pi}{n}}\left|e^{\frac{2ik\pi}{n}}\alpha\right\rangle. \tag{44}$$

Using the definition of coherent states (3), we obtain

$$\begin{aligned}|\sigma_\alpha^{q,n}\rangle &= \frac{\sqrt{q!}e^{-\frac{1}{2}|\alpha|^2}}{n\alpha^q}\sum_{p\geq 0}\frac{\alpha^p}{\sqrt{p!}}\left(\sum_{k=0}^{n-1}e^{\frac{2ik(p-q)\pi}{n}}\right)|p\rangle\\ &= \sqrt{q!}e^{-\frac{1}{2}|\alpha|^2}\sum_{l\geq 0}\frac{\alpha^{ln}}{\sqrt{(ln+q)!}}|ln+q\rangle.\end{aligned} \tag{45}$$

where we used $\sum_{k=0}^{n-1}e^{\frac{2ik(p-q)\pi}{n}} = n$ if $p = q \mod n$ and 0 otherwise in the second line. Hence, $\lim_{\alpha\to 0}|\sigma_\alpha^{q,n}\rangle = |q\rangle$ in trace distance. Using the Fock basis expansion of $\hat{U}$ (10), we thus have, for all $\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n\geq 1$,

$$\frac{\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}})}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} = \lim_{\alpha\to 0}\langle p_1\dots p_m|\hat{U}\bigotimes_{j=1}^m\left|\sigma_\alpha^{q_j,n}\right\rangle, \tag{46}$$

where $\hat{U}$ is a passive linear operation over m modes with unitary matrix $U = (u_{ij})_{1\leq i,j\leq m}$. We compute the right hand side of this equation:

$$\begin{aligned}\langle p_1\dots p_m|\hat{U}\bigotimes_{j=1}^m\left|\sigma_\alpha^{q_j,n}\right\rangle &= \frac{\sqrt{\boldsymbol{q}!}}{n^m\alpha^n}\sum_{k_1,\dots,k_m=0}^{n-1}\prod_{j=1}^m e^{-\frac{2ik_jq_j\pi}{n}}\langle p_1\dots p_m|\hat{U}\bigotimes_{j=1}^m\left|e^{\frac{2ik_j\pi}{n}}\alpha\right\rangle\\ &= \frac{\sqrt{\boldsymbol{q}!}}{n^m\alpha^n}\sum_{\boldsymbol{x}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}^m}\boldsymbol{x}^{-\boldsymbol{q}}\langle p_1\dots p_m|\hat{U}|\alpha x_1\dots\alpha x_m\rangle\\ &= \frac{\sqrt{\boldsymbol{q}!}e^{-\frac{m}{2}|\alpha|^2}}{\sqrt{\boldsymbol{p}!}n^m}\sum_{\boldsymbol{x}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}^m}\boldsymbol{x}^{-\boldsymbol{q}}(U\boldsymbol{x})^{\boldsymbol{p}}.\end{aligned} \tag{47}$$

where we used the definition of $|\sigma_\alpha^{q,n}\rangle$ (44) in the first line and where the derivation for the last line is identical to that of Eqs. (36-37). Letting α go to 0 proves Eq. (43) for unitary matrices. Finally, the case of a generic nonzero $m \times m$ matrix A is obtained as in Eq. (40) by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U of size $2m$ and computing

$$\left(\langle p_1 \dots p_m| \otimes \langle 0|^{\otimes m}\right) \hat{U} \left(\bigotimes_{j=1}^{m} \left|\sigma_\alpha^{q_j,n}\right\rangle \otimes |0\rangle^{\otimes m}\right). \tag{48}$$

5.2 The Glynn–Kan formula

A symmetrized version of Glynn's formula for the permanent of an $m \times m$ matrix A has been recently derived [41] under the name Glynn–Kan formula:

$$\mathrm{Per}(A) = \frac{1}{4^m m!} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m (\boldsymbol{x}^T A \boldsymbol{y})^m. \tag{49}$$

While the Glynn–Kan formula gives a slower classical algorithm than the Glynn formula for computing the permanent, it is motivated by a connection with quantum algorithms for estimating the permanent [41].

Proof of the Glynn–Kan formula. To prove this identity using quantum-mechanical tools, we again use the unnormalized cat state $|\tilde{\mathrm{cat}}_\alpha\rangle := \frac{1}{2\alpha}(|\alpha\rangle - |-\alpha\rangle)$, as in the previous section. Using the Fock basis expansion of coherent states (6), we have $\lim_{\alpha\to 0} |\tilde{\mathrm{cat}}_\alpha\rangle = |1\rangle$ in trace distance. Hence, using the Fock basis expansion of $\hat{U}$ (10),

$$\mathrm{Per}(U) = \lim_{\alpha\to 0} \langle \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha | \hat{U} | \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha \rangle, \tag{50}$$

where $\hat{U}$ is a passive linear operation over m modes with unitary matrix U. We compute the right hand side of this equation:

$$\begin{aligned}
&\langle \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha | \hat{U} | \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha \rangle \\
&\quad = \frac{1}{4^m |\alpha|^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m \langle x_1\alpha \dots x_m\alpha | \hat{U} | y_1\alpha \dots y_m\alpha \rangle \\
&\quad = \frac{1}{4^m |\alpha|^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m \langle x_1\alpha \dots x_m\alpha | (\alpha U\boldsymbol{y})_1 \dots (\alpha U\boldsymbol{y})_m \rangle \\
&\quad = \frac{1}{4^m |\alpha|^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m e^{-\frac{1}{2}|\alpha|^2 \|\boldsymbol{x}\|^2} e^{-\frac{1}{2}|\alpha|^2 \|U\boldsymbol{y}\|^2} \prod_{i=1}^{m} e^{x_i \alpha^* (\alpha U\boldsymbol{y})_i} \\
&\quad = \frac{e^{-m|\alpha|^2}}{4^m |\alpha|^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m e^{|\alpha|^2 \boldsymbol{x}^T U \boldsymbol{y}} \\
&\quad = \frac{e^{-m|\alpha|^2}}{4^m} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m \sum_{k=0}^{+\infty} \frac{|\alpha|^{2k-2m}}{k!} (\boldsymbol{x}^T U \boldsymbol{y})^k,
\end{aligned} \tag{51}$$

where we used the action of $\hat{U}$ on coherent states (13) in the third line, the overlap between coherent states (7) in the fourth line, and the fact that U is unitary in the fifth line. For $k < m$, all products in the expansion of $(\boldsymbol{x}^T U \boldsymbol{y})^k$ have at least one x_i missing. Hence, by symmetry, the terms for $k < m$ vanish and we obtain

$$\langle \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha | \hat{U} | \tilde{\mathrm{cat}}_\alpha \dots \tilde{\mathrm{cat}}_\alpha \rangle = \frac{e^{-m|\alpha|^2}}{4^m} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m \sum_{l=0}^{\infty} \frac{|\alpha|^{2l}}{(m+l)!} (\boldsymbol{x}^T U \boldsymbol{y})^{m+l}. \tag{52}$$

Finally, taking the limit when α goes to 0 yields

$$\mathrm{Per}(U) = \frac{1}{4^m m!} \sum_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m (\boldsymbol{x}^T U \boldsymbol{y})^m. \tag{53}$$

This proves the Glynn–Kan formula for a unitary matrix U. Once again, the proof extends straightforwardly to any matrix A of size n by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U of size $2n$ and computing $\lim_{\alpha\to 0}(\langle \tilde{\mathrm{cat}}_\alpha|^{\otimes n} \otimes \langle 0|^{\otimes n})\hat{U}(|\tilde{\mathrm{cat}}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes n})$. □

Similar to Eq. (43), we obtain a version of the Glynn–Kan formula for matrices with repeated rows and columns using roots of unity as

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{\boldsymbol{p}!\boldsymbol{q}!}{n^{2m} n!} \sum_{\boldsymbol{x},\boldsymbol{y}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}^m} \boldsymbol{x}^{-\boldsymbol{p}} \boldsymbol{y}^{-\boldsymbol{q}} (\boldsymbol{x}^T A \boldsymbol{y})^n, \tag{54}$$

for any $m \times m$ matrix $A = (a_{ij})_{1\le i,j\le m}$ and all $\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n$. This formula, which appears to be new, can be easily proven using our Corollary 1 by expanding the product $(\boldsymbol{x}^T A \boldsymbol{y})^n$ and using properties of roots of unity. We give a quantum-inspired proof in what follows.

For all $n \ge 1$, all $q \le n$, and all $\alpha \in \mathbb{C}$, we again use the unnormalised superposition of coherent states as in Eq. (44): $|\sigma_\alpha^{q,n}\rangle := \frac{\sqrt{q!}}{n\alpha^q} \sum_{k=0}^{n-1} e^{-\frac{2ikq\pi}{n}} \left|e^{\frac{2ik\pi}{n}}\alpha\right\rangle$. Recall that $\lim_{\alpha\to 0} |\sigma_\alpha^{q,n}\rangle = |q\rangle$ in trace distance. Using the Fock basis expansion of $\hat{U}$ (10), we thus have, for all $\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n \ge 1$,

$$\frac{\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}})}{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}} = \lim_{\alpha\to 0} \bigotimes_{j=1}^m \left\langle \sigma_\alpha^{p_j,n}\right| \hat{U} \bigotimes_{j=1}^m \left|\sigma_\alpha^{q_j,n}\right\rangle, \tag{55}$$

where $\hat{U}$ is a passive linear operation over m modes with unitary matrix $U = (u_{ij})_{1\le i,j\le m}$. We compute the right hand side of this equation:

$$\begin{aligned}
\bigotimes_{j=1}^m \left\langle \sigma_\alpha^{p_j,n}\right| \hat{U} \bigotimes_{j=1}^m \left|\sigma_\alpha^{q_j,n}\right\rangle &= \frac{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}}{n^{2m}|\alpha|^{2n}} \sum_{\substack{k_1,\dots,k_m=0\\ l_1,\dots l_m=0}}^{n-1} \prod_{j=1}^m e^{-\frac{2ik_j p_j \pi}{n}} e^{-\frac{2il_j q_j \pi}{n}} \bigotimes_{j=1}^m \left\langle e^{\frac{2ik_j\pi}{n}}\alpha\right| \hat{U} \bigotimes_{j=1}^m \left|e^{\frac{2il_j\pi}{n}}\alpha\right\rangle \\
&= \frac{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}}{n^{2m}|\alpha|^{2n}} \sum_{\boldsymbol{x},\boldsymbol{y}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}^m} \boldsymbol{x}^{-\boldsymbol{p}} \boldsymbol{y}^{-\boldsymbol{q}} \langle x_1\alpha \dots x_m\alpha|\hat{U}|y_1\alpha \dots y_m\alpha\rangle \\
&= \frac{\sqrt{\boldsymbol{p}!\boldsymbol{q}!}e^{-m|\alpha|^2}}{n^{2m}} \sum_{\boldsymbol{x}\in\{1,e^{\frac{2i\pi}{n}},\dots,e^{\frac{2i(n-1)\pi}{n}}\}} \boldsymbol{x}^{-\boldsymbol{p}} \boldsymbol{y}^{-\boldsymbol{q}} \sum_{l=0}^\infty \frac{|\alpha|^{2l}}{(n+l)!} (\boldsymbol{x}^T U \boldsymbol{y})^{n+l}.
\end{aligned} \tag{56}$$

where we used the definition of $|\sigma_\alpha^{q,n}\rangle$ (44) in the first line and where the derivation for the last line is identical to that of Eqs. (51-52). Letting α go to 0 proves Eq. (54) for unitary matrices. Once again, the proof extends straightforwardly to any nonzero matrix A of size m by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U of size $2m$ and computing

$$\left(\bigotimes_{j=1}^m \left\langle \sigma_\alpha^{p_j,n}\right| \otimes \langle 0|^{\otimes m}\right) \hat{U} \left(\bigotimes_{j=1}^m \left|\sigma_\alpha^{q_j,n}\right\rangle \otimes |0\rangle^{\otimes m}\right). \tag{57}$$

5.3 The Cauchy–Binet theorem

The Cauchy–Binet theorem for the permanent expresses the permanent of the product of two $m \times m$ matrices A and B as a sum of products involving permanents of submatrices of these two matrices [1, 2]: for all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$,

$$\mathrm{Per}((AB)_{\boldsymbol{p},\boldsymbol{q}}) = \sum_{\boldsymbol{k}\in\mathbb{N}^m} \frac{1}{\boldsymbol{k}!}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{k}})\mathrm{Per}(B_{\boldsymbol{k},\boldsymbol{q}}). \tag{58}$$

Proof of the Cauchy–Binet theorem. Using the Fock basis expansion of passive linear operations (10) gives a quick quantum-inspired proof of this identity: for U and V two $m\times m$ unitary matrices and for all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$,

$$\begin{aligned}
\mathrm{Per}((UV)_{\boldsymbol{p},\boldsymbol{q}}) &= \sqrt{\boldsymbol{p}!\boldsymbol{q}!}\,\langle \boldsymbol{p}|\,\widehat{UV}\,|\boldsymbol{q}\rangle \\
&= \sqrt{\boldsymbol{p}!\boldsymbol{q}!}\,\langle \boldsymbol{p}|\,\hat{U}\hat{V}\,|\boldsymbol{q}\rangle \\
&= \sqrt{\boldsymbol{p}!\boldsymbol{q}!}\,\langle \boldsymbol{p}|\,\hat{U}\left(\sum_{\boldsymbol{k}\in\mathbb{N}^m} |\boldsymbol{k}\rangle\langle\boldsymbol{k}|\right)\hat{V}\,|\boldsymbol{q}\rangle \\
&= \sum_{\boldsymbol{k}\in\mathbb{N}^m} \sqrt{\boldsymbol{p}!\boldsymbol{q}!}\,\langle \boldsymbol{p}|\,\hat{U}\,|\boldsymbol{k}\rangle\langle\boldsymbol{k}|\,\hat{V}\,|\boldsymbol{q}\rangle \\
&= \sum_{\boldsymbol{k}\in\mathbb{N}^m} \frac{1}{\boldsymbol{k}!}\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{k}})\mathrm{Per}(V_{\boldsymbol{k},\boldsymbol{q}}),
\end{aligned} \tag{59}$$

where we used the fact that Fock states form a basis in the third line and where we used Eq. (10) once in the first line and twice in the last line.

In order to retrieve the formula for generic matrices A and B, we can embed $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U and $\frac{1}{\|B\|}B$ as a submatrix of a unitary matrix V and compute $\mathrm{Per}((UV)_{\boldsymbol{p}\oplus\boldsymbol{0},\boldsymbol{q}\oplus\boldsymbol{0}})$. □

5.4 Generating functions

The permanent may be seen as a monomial coefficient in the Taylor expansion of various functions [1, 2]. For example, the permanent of an $m \times m$ matrix $A = (a_{ij})_{1\leq i,j\leq m}$ with rows repeated according to $\boldsymbol{p} \in \mathbb{N}^m$ and columns repeated according to $\boldsymbol{q} \in \mathbb{N}^m$ is given by the coefficient of $\boldsymbol{z}^{\boldsymbol{q}} = z_1^{q_1}\ldots z_m^{q_m}$ in

$$\boldsymbol{q}!(A\boldsymbol{z})^{\boldsymbol{p}} = \prod_{i=1}^{m} q_i!\left(\sum_{j=1}^{m} a_{ij}z_j\right)^{p_i}. \tag{60}$$

This may be thought of as the 'monomial version' of Glynn's formula in Eq. (42). Formally:

$$\sum_{\boldsymbol{q}\in\mathbb{N}^m} \frac{\boldsymbol{z}^{\boldsymbol{q}}}{\boldsymbol{q}!}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = (A\boldsymbol{z})^{\boldsymbol{p}}, \tag{61}$$

for formal variables $\boldsymbol{z} = (z_1, \ldots, z_m)$.

Proof of 'Glynn's monomial formula'. To prove this relation with quantum mechanical tools, we fix $\boldsymbol{p} \in \mathbb{N}^m$, $\boldsymbol{\alpha} \in \mathbb{C}^m$, a unitary matrix U of size m, and we compute:

$$\sum_{\boldsymbol{q}\in\mathbb{N}^m} \frac{\boldsymbol{\alpha}^{\boldsymbol{q}}}{\boldsymbol{q}!}\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}}) = \sqrt{\boldsymbol{p}!}e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2}\,\langle \boldsymbol{p}|\,\hat{U}\left(\sum_{\boldsymbol{q}\in\mathbb{N}^m} |\boldsymbol{q}\rangle\langle\boldsymbol{q}|\right)|\boldsymbol{\alpha}\rangle$$

$$
\begin{aligned}
&= \sqrt{\boldsymbol{p}!} e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2} \langle \boldsymbol{p}| \hat{U} |\boldsymbol{\alpha}\rangle \\
&= \sqrt{\boldsymbol{p}!} e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2} \langle \boldsymbol{p}| (U\boldsymbol{\alpha})\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2 - \frac{1}{2}\|U\boldsymbol{\alpha}\|^2} (U\boldsymbol{\alpha})^{\boldsymbol{p}} \\
&= (U\boldsymbol{\alpha})^{\boldsymbol{p}},
\end{aligned}
\tag{62}
$$

where we used the Fock basis expansions of coherent states (6) and of $\hat{U}$ (10) in the first line, the fact that Fock states form a basis in the second line, the action of $\hat{U}$ on coherent states (13) in the third line, Eq. (6) again in the fourth line, and the fact that U is unitary in the last line. Once again, the relation for a generic nonzero matrix A is obtained by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U. □

Another generating function for the permanent is due to Jackson [46]: for formal variables $\boldsymbol{x} = (x_1, \ldots, x_m), \boldsymbol{y} = (y_1, \ldots, y_m)$,

$$
\sum_{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m} \frac{\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = e^{\boldsymbol{x}^T A \boldsymbol{y}}. \tag{63}
$$

Note that this relation implies Corollary 1, by expanding the Taylor series of the exponential and considering the $\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}$ coefficient when $|\boldsymbol{p}| = |\boldsymbol{q}| = n \in \mathbb{N}^*$ (an alternative proof of this result is given in the next section):

$$
\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{\boldsymbol{p}!\boldsymbol{q}!}{n!} [\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}](\boldsymbol{x}^T A \boldsymbol{y})^n. \tag{64}
$$

Proof of Jackson's formula. Jackson's formula may be derived using quantum mechanical tools in a similar way: for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}^m$ and a unitary matrix U,

$$
\begin{aligned}
\sum_{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m} \frac{\boldsymbol{\alpha}^{\boldsymbol{p}}\boldsymbol{\beta}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!} \mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}}) &= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2 + \frac{1}{2}\|\boldsymbol{\beta}\|^2} \langle \boldsymbol{\alpha}^*| \left(\sum_{\boldsymbol{p}\in\mathbb{N}^m} |\boldsymbol{p}\rangle\langle\boldsymbol{p}| \right) \hat{U} \left(\sum_{\boldsymbol{q}\in\mathbb{N}^m} |\boldsymbol{q}\rangle\langle\boldsymbol{q}| \right) |\boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2 + \frac{1}{2}\|\boldsymbol{\beta}\|^2} \langle \boldsymbol{\alpha}^*| \hat{U} |\boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2 + \frac{1}{2}\|\boldsymbol{\beta}\|^2} \langle \boldsymbol{\alpha}^*| U\boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\beta}\|^2 - \frac{1}{2}\|U\boldsymbol{\beta}\|^2} e^{\boldsymbol{\alpha}^T U \boldsymbol{\beta}} \\
&= e^{\boldsymbol{\alpha}^T U \boldsymbol{\beta}},
\end{aligned}
\tag{65}
$$

where we used the Fock basis expansions of coherent states (6) and of $\hat{U}$ (10) in the first line, the fact that Fock states form a basis in the second line, the action of $\hat{U}$ on coherent states (13) in the third line, the overlap between coherent states (7) in the fourth line, and the fact that U is unitary in the last line. Once again, the relation for a generic nonzero matrix A is obtained by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U. □

We conclude this section with arguably one of the most remarkable permanent identities, the MacMahon master theorem [13], which relates the permanent and the determinant through a generating function (see section 2.3):

$$
\sum_{\boldsymbol{p}\in\mathbb{N}^m} \frac{\boldsymbol{z}^{\boldsymbol{p}}}{\boldsymbol{p}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \frac{1}{\mathrm{Det}(I - ZA)}, \tag{66}
$$

where $Z = \mathrm{Diag}(\boldsymbol{z})$, with $\boldsymbol{z} = (z_1, \ldots, z_m)$ formal variables.

Proof of the MacMahon master theorem. To prove this relation with quantum mechanical tools, we show that the MacMahon master theorem describes two different ways of computing an inner product between two Gaussian states.

For $\lambda \in \mathbb{C}$ and $|\lambda| < 1$, we make use of (unnormalized) two-mode squeezed states of the form $|\lambda\rangle = \sum_{p\geq 0} \lambda^p |pp\rangle$. We write $|\boldsymbol{\lambda}\rangle$ with $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{C}^m$ a tensor product of m two-mode squeezed states (note that we are associating mode 1 with mode $m+1$, mode 2 with mode $m+2$, and so on). For U a unitary matrix and $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{C}^m$, with $|\lambda_k| < 1$ and $|\mu_k| < 1$ for all $k \in \{1, \ldots, m\}$, we compute, using the Fock basis expansion of $\hat{U}$ (10):

$$\sum_{\boldsymbol{p}\in\mathbb{N}^m} \frac{\boldsymbol{\lambda}^{\boldsymbol{p}}\boldsymbol{\mu}^{\boldsymbol{p}}}{\boldsymbol{p}!} \mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{p}}) = \sum_{\boldsymbol{p}\in\mathbb{N}^m} \boldsymbol{\lambda}^{\boldsymbol{p}}\boldsymbol{\mu}^{\boldsymbol{p}} \langle \boldsymbol{p}| \hat{U} |\boldsymbol{p}\rangle . \tag{67}$$

Using the Fock basis expansion of two-mode squeezed states (8), we have

$$\boldsymbol{\lambda}^{\boldsymbol{p}} = \langle \boldsymbol{\lambda}^*| \boldsymbol{pp}\rangle. \tag{68}$$

Moreover, a quick computation in Fock basis shows that

$$\boldsymbol{\mu}^{\boldsymbol{p}} \langle \boldsymbol{p}| \hat{U} |\boldsymbol{p}\rangle = \langle \boldsymbol{pp}| \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle . \tag{69}$$

Hence, Eq. (67) rewrites

$$\begin{aligned} \sum_{\boldsymbol{p}\in\mathbb{N}^m} \frac{\boldsymbol{\lambda}^{\boldsymbol{p}}\boldsymbol{\mu}^{\boldsymbol{p}}}{\boldsymbol{p}!} \mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{p}}) &= \sum_{\boldsymbol{p}\in\mathbb{N}^m} \langle \boldsymbol{\lambda}^*| \boldsymbol{pp}\rangle\langle \boldsymbol{pp}| \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle \\ &= \langle \boldsymbol{\lambda}^*| \left(\sum_{\boldsymbol{p}\in\mathbb{N}^m} |\boldsymbol{pp}\rangle\langle \boldsymbol{pp}| \right) \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle \\ &= \langle \boldsymbol{\lambda}^*| \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle , \end{aligned} \tag{70}$$

where in the last line we used $\langle \boldsymbol{\lambda}^*| \left(\sum_{\boldsymbol{p}\in\mathbb{N}^m} |\boldsymbol{pp}\rangle\langle \boldsymbol{pp}| \right) = \langle \boldsymbol{\lambda}^*|$, which can be checked in Fock basis.

We now compute the Gaussian inner product $\langle \boldsymbol{\lambda}^*| \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle$. From the action of passive linear operations on creation operators (11), the unitary matrix associated to the passive linear operation $\hat{I} \otimes \hat{U}$ is $I \oplus U$. With the Gaussian overlap derived in Eqs. (15-17), we thus obtain:

$$\langle \boldsymbol{\lambda}^*| \hat{I} \otimes \hat{U} |\boldsymbol{\mu}\rangle = \frac{1}{\sqrt{\mathrm{Det}(V_{I\oplus U}(\boldsymbol{\lambda}, \boldsymbol{\mu}))}}, \tag{71}$$

where

$$V_{I\oplus U}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{pmatrix} -(I \oplus U) V_{\boldsymbol{\mu}} (I \oplus U)^T & I_{2m} \\ I_{2m} & -V_{\boldsymbol{\lambda}} \end{pmatrix}, \tag{72}$$

where for all $\boldsymbol{w} \in \mathbb{C}^m$,

$$V_{\boldsymbol{w}} = \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{w}) \\ \mathrm{Diag}(\boldsymbol{w}) & 0_m \end{pmatrix}. \tag{73}$$

We have

$$\begin{aligned} \mathrm{Det}(V_{I\oplus U}(\boldsymbol{\lambda}, \boldsymbol{\mu})) &= \mathrm{Det}(I - (I \oplus U) V_{\boldsymbol{\mu}} (I \oplus U)^T V_{\boldsymbol{\lambda}}) \\ &= \mathrm{Det}\left(I - (I \oplus U) \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{\mu}) \\ \mathrm{Diag}(\boldsymbol{\mu}) & 0_m \end{pmatrix} (I \oplus U)^T \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{\lambda}) \\ \mathrm{Diag}(\boldsymbol{\lambda}) & 0_m \end{pmatrix} \right) \\ &= \mathrm{Det}\begin{pmatrix} I - \mathrm{Diag}(\boldsymbol{\mu}) U^T \mathrm{Diag}(\boldsymbol{\lambda}) & 0_m \\ 0_m & I - U \mathrm{Diag}(\boldsymbol{\mu}) \mathrm{Diag}(\boldsymbol{\lambda}) \end{pmatrix} \end{aligned} \tag{74}$$

$$= \mathrm{Det}(I - \mathrm{Diag}(\boldsymbol{\mu})U^T\mathrm{Diag}(\boldsymbol{\lambda}))\,\mathrm{Det}(I - U\mathrm{Diag}(\boldsymbol{\mu})\mathrm{Diag}(\boldsymbol{\lambda}))$$
$$= \mathrm{Det}(I - \mathrm{Diag}(\boldsymbol{\lambda})\mathrm{Diag}(\boldsymbol{\mu})U)^2,$$

where the last line is obtained by using Sylvester's determinant theorem $\mathrm{Det}(I+MN) = \mathrm{Det}(I+NM)$, together with the transpose for the first determinant and $\mathrm{Diag}(\boldsymbol{\mu})\mathrm{Diag}(\boldsymbol{\lambda}) = \mathrm{Diag}(\boldsymbol{\lambda})\mathrm{Diag}(\boldsymbol{\mu})$ for the second determinant. Combining Eqs. (70), (71) and (74), we obtain, $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{C}^m$, with $|\lambda_k| < 1$ and $|\mu_k| < 1$ for all $k \in \{1, \dots, m\}$,

$$\sum_{\boldsymbol{p}\in\mathbb{N}^m} \frac{\boldsymbol{\lambda}^{\boldsymbol{p}}\boldsymbol{\mu}^{\boldsymbol{p}}}{\boldsymbol{p}!}\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{p}}) = \langle\boldsymbol{\lambda}^*|\,\hat{I}\otimes\hat{U}\,|\boldsymbol{\mu}\rangle = \frac{1}{\mathrm{Det}(I - \mathrm{Diag}(\boldsymbol{\lambda})\mathrm{Diag}(\boldsymbol{\mu})U)}, \tag{75}$$

where we used the value at $\boldsymbol{\lambda} = \boldsymbol{\mu} = 0$ and the fact that the left hand side is a continuous function of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ to determine the sign of the square root. Replacing $(\lambda_1\mu_1, \dots, \lambda_m\mu_m)$ by the formal variables $\boldsymbol{z} = (z_1, \dots, z_m)$ concludes the proof. The relation for a generic nonzero matrix A of size n is obtained by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U of size $2n$ and taking $\boldsymbol{z} = (z_1, \dots, z_n, 0, \dots, 0)$. □

6 New quantum-inspired permanent identities

In this section, we derive new quantum-inspired identities for the permanent and we give some combinatorial applications of these identities.

6.1 Generalizations of the MacMahon master theorem

In this section, we introduce new generalizations of the MacMahon master theorem (see section 2.3).

We first derive new quantum-inspired identities involving the permanent of even-sized matrices: for M a $(2m) \times (2m)$ matrix,

$$\mathrm{Per}(M) = [z^m]\left(\frac{1}{4^m}\sum_{\boldsymbol{x},\boldsymbol{y}\in\{-1,1\}^m}\frac{x_1\dots x_m y_1\dots y_m}{\sqrt{\mathrm{Det}(I - zV_{\boldsymbol{x}}MV_{\boldsymbol{y}}M^T)}}\right), \tag{76}$$

and, for formal variables $\boldsymbol{x} = (x_1, \dots, x_m), \boldsymbol{y} = (y_1, \dots, y_m)$,

$$\sum_{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m}\frac{\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!}\mathrm{Per}(M_{\boldsymbol{p}\oplus\boldsymbol{p},\boldsymbol{q}\oplus\boldsymbol{q}}) = \frac{1}{\sqrt{\mathrm{Det}(I - V_{\boldsymbol{x}}MV_{\boldsymbol{y}}M^T)}}, \tag{77}$$

where for all $\boldsymbol{w} = (w_1, \dots, w_m)$,

$$V_{\boldsymbol{w}} = \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{w}) \\ \mathrm{Diag}(\boldsymbol{w}) & 0_m \end{pmatrix}. \tag{78}$$

As a direct consequence, when $M = A \oplus B$, with A and B two $m \times m$ matrices, we obtain Theorem 1:

$$\sum_{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m}\frac{\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}})\mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}}) = \frac{1}{\mathrm{Det}(I - XAYB^T)}, \tag{79}$$

where $X = \mathrm{Diag}(x_1, \dots, x_m)$ and $Y = \mathrm{Diag}(y_1, \dots, y_m)$. In particular, $\mathrm{Per}(A)\mathrm{Per}(B)$ is given by the $x_1 \dots x_m y_1 \dots y_m$ coefficient of $1/\mathrm{Det}(I - XAYB^T)$. As a consequence, we obtain Corollary 1: for all $n \in \mathbb{N}^*$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n$,

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{\boldsymbol{p}!\boldsymbol{q}!}{n!}[\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}]\left(\boldsymbol{x}^T A\boldsymbol{y}\right)^n. \tag{80}$$

We also generalize Eq. (79) to the case of N matrices to obtain Theorem 2: for all $n \times n$ matrices $A^{(1)}, \dots, A^{(N)}$,

$$\begin{aligned} &\sum_{\boldsymbol{p}_1,\dots,\boldsymbol{p}_N \in \mathbb{N}^m} \prod_{k=1}^{N} \frac{\boldsymbol{z}_k^{\boldsymbol{p}_k}}{\boldsymbol{p}_k!} \mathrm{Per}(A^{(1)}_{\boldsymbol{p}_1,\boldsymbol{p}_2}) \mathrm{Per}(A^{(2)}_{\boldsymbol{p}_2,\boldsymbol{p}_3}) \dots \mathrm{Per}(A^{(N)}_{\boldsymbol{p}_N,\boldsymbol{p}_1}) \\ &= \frac{1}{\mathrm{Det}(I - Z_1 A^{(1)} \dots Z_N A^{(N)})}, \end{aligned} \tag{81}$$

where $\boldsymbol{z}_k = (z_{k1}, \dots, z_{km})$ are formal variables and $Z_k = \mathrm{Diag}(\boldsymbol{z}_k)$ for all $k \in \{1, \dots, N\}$.

Proofs of Theorem 1, Corollary 1, and Theorem 2. In section 5.1, we have obtained a linear optical proof of Glynn's formula for the permanent using the fact that a cat state $|\mathrm{cat}_\alpha\rangle$ of small amplitude α approximates a single-photon Fock state $|1\rangle$. Hereafter, we consider another approximation of Fock states using superpositions of two-mode squeezed states. Such superpositions were recently studied in [47] in the context of quantum sensing.

To prove the identity in Eq. (76), we introduce $|\lambda^-\rangle := \frac{1}{2\lambda}(|\lambda\rangle - |-\lambda\rangle)$, for $|\lambda| < 1$, where $|\lambda\rangle = \sum_{p \geq 0} \lambda^p |pp\rangle$ is an unnormalized two-mode squeezed state. We have $\lim_{\lambda \to 0} |\lambda^-\rangle = |11\rangle$ in trace distance. Hence, for U a $(2m) \times (2m)$ unitary matrix we have with the Fock basis expansion of $\hat{U}$ (10):

$$\mathrm{Per}(U) = \lim_{\lambda \to 0} {}^{\otimes m}\langle \lambda^- | \hat{U} | \lambda^- \rangle^{\otimes m}. \tag{82}$$

For all $\lambda \in \mathbb{R}$, with $|\lambda| < 1$, let us compute

$${}^{\otimes m}\langle \lambda^- | \hat{U} | \lambda^- \rangle^{\otimes m} = \frac{1}{4^m \lambda^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m g_U(\lambda \boldsymbol{x}, \lambda \boldsymbol{y}), \tag{83}$$

where we have defined for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{C}^m$ (note that we associate mode 1 with mode $m+1$, mode 2 with mode $m+2$, and so on)

$$g_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \langle \boldsymbol{\lambda}^* | \hat{U} | \boldsymbol{\mu} \rangle. \tag{84}$$

With the Gaussian overlap from Eq. (15), we have

$$g_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{1}{\sqrt{\mathrm{Det}(V_U(\boldsymbol{\lambda}, \boldsymbol{\mu}))}}, \tag{85}$$

where

$$V_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{pmatrix} -U V_{\boldsymbol{\mu}} U^T & I_{2m} \\ I_{2m} & -V_{\boldsymbol{\lambda}} \end{pmatrix}, \tag{86}$$

where for all $\boldsymbol{w} \in \mathbb{C}^m$,

$$V_{\boldsymbol{w}} = \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{w}) \\ \mathrm{Diag}(\boldsymbol{w}) & 0_m \end{pmatrix}. \tag{87}$$

Moreover,

$$\mathrm{Det}(V_U(\boldsymbol{\lambda}, \boldsymbol{\mu})) = \mathrm{Det}(I - V_{\boldsymbol{\lambda}} U V_{\boldsymbol{\mu}} U^T), \tag{88}$$

so with Eq. (85)

$$g_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{1}{\sqrt{\mathrm{Det}(I - V_{\boldsymbol{\lambda}} U V_{\boldsymbol{\mu}} U^T)}}. \tag{89}$$

Setting $\boldsymbol{\lambda} = \boldsymbol{\mu} = \lambda \mathbf{1}$ for $\lambda \in \mathbb{R}$ and letting λ go to 0 in Eq. (83), we obtain that $\mathrm{Per}(U)$ is given by

$$\frac{1}{4^m} \sum_{\boldsymbol{x},\boldsymbol{y} \in \{-1,1\}^m} x_1 \dots x_m y_1 \dots y_m [\lambda^{2m}] g_U(\lambda \boldsymbol{x}, \lambda \boldsymbol{y}), \tag{90}$$

where $[\lambda^{2m}] g_U(\lambda \boldsymbol{x}, \lambda \boldsymbol{y})$ is the λ^{2m} coefficient in the Taylor expansion (in λ) of

$$g_U(\lambda \boldsymbol{x}, \lambda \boldsymbol{y}) = \frac{1}{\sqrt{\mathrm{Det}(I - \lambda^2 V_{\boldsymbol{x}} U V_{\boldsymbol{y}} U^T)}}. \tag{91}$$

This concludes the proof of Eq. (76). Once again, the relation for a generic nonzero matrix A is obtained by embedding $\frac{1}{\|A\|} A$ as a submatrix of a unitary matrix U.

Now with Eq. (84), for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{C}^n$,

$$\begin{aligned} g_U(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \sum_{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m} \boldsymbol{\lambda}^{\boldsymbol{p}} \boldsymbol{\mu}^{\boldsymbol{q}} \langle \boldsymbol{p}\boldsymbol{p} | \hat{U} | \boldsymbol{q}\boldsymbol{q} \rangle \\ &= \sum_{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m} \frac{\boldsymbol{\lambda}^{\boldsymbol{p}} \boldsymbol{\mu}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(U_{\boldsymbol{p} \oplus \boldsymbol{p}, \boldsymbol{q} \oplus \boldsymbol{q}}), \end{aligned} \tag{92}$$

where we used the Fock basis expansion of $\hat{U}$ (10) in the second line. Combining Eqs. (84), (89) and (92) we obtain

$$\sum_{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m} \frac{\boldsymbol{\lambda}^{\boldsymbol{p}} \boldsymbol{\mu}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(U_{\boldsymbol{p} \oplus \boldsymbol{p}, \boldsymbol{q} \oplus \boldsymbol{q}}) = \langle \boldsymbol{\lambda}^* | \hat{U} | \boldsymbol{\mu} \rangle = \frac{1}{\sqrt{\mathrm{Det}(I - V_{\boldsymbol{\lambda}} U V_{\boldsymbol{\mu}} U^T)}}, \tag{93}$$

where for all $\boldsymbol{w} \in \mathbb{C}^n$,

$$V_{\boldsymbol{w}} = \begin{pmatrix} 0_m & \mathrm{Diag}(\boldsymbol{w}) \\ \mathrm{Diag}(\boldsymbol{w}) & 0_m \end{pmatrix}. \tag{94}$$

This concludes the proof of Eq. (77). Once again, the relation for a generic nonzero matrix A is obtained by embedding $\frac{1}{\|A\|} A$ as a submatrix of a unitary matrix U.

When $U = A \oplus B$, with A and B two $m \times m$ matrices, we have $U_{\boldsymbol{p} \oplus \boldsymbol{p}, \boldsymbol{q} \oplus \boldsymbol{q}} = A_{\boldsymbol{p},\boldsymbol{q}} \oplus B_{\boldsymbol{p},\boldsymbol{q}}$ and the permanent of a block-diagonal matrix is the product of the permanents of the blocks, so Eq. (93) gives

$$\sum_{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m} \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) \mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}}) = \frac{1}{\sqrt{\mathrm{Det}(I - V_{\boldsymbol{x}} (A \oplus B) V_{\boldsymbol{y}} (A^T \oplus B^T))}}. \tag{95}$$

Writing $X = \mathrm{Diag}(x_1, \dots, x_m)$ and $Y = \mathrm{Diag}(y_1, \dots, y_m)$, we have

$$\begin{aligned} \mathrm{Det}(I - V_{\boldsymbol{x}} (A \oplus B) V_{\boldsymbol{y}} (A^T \oplus B^T)) &= \mathrm{Det}\left(I - \begin{pmatrix} 0_m & X \\ X & 0_m \end{pmatrix} \begin{pmatrix} A & 0_m \\ 0_m & B \end{pmatrix} \begin{pmatrix} 0_m & Y \\ Y & 0_m \end{pmatrix} \begin{pmatrix} A^T & 0_m \\ 0_m & B^T \end{pmatrix} \right) \\ &= \mathrm{Det}\begin{pmatrix} I - XBYA^T & 0_m \\ 0_m & I - XAYB^T \end{pmatrix} \\ &= \mathrm{Det}(I - XBYA^T)\, \mathrm{Det}(I - XAYB^T) \\ &= \mathrm{Det}(I - XAYB^T)^2, \end{aligned} \tag{96}$$

where the last line is obtained by using Sylvester's determinant theorem $\mathrm{Det}(I + MN) = \mathrm{Det}(I + NM)$ together with the transpose for the first determinant. Combining Eqs. (95) and (96) concludes the proof of Theorem 1.

This theorem reduces to the MacMahon master theorem when $A = I$ or $B = I$, since $\mathrm{Per}(I_{\boldsymbol{p},\boldsymbol{q}}) = \boldsymbol{p}!\delta_{\boldsymbol{p},\boldsymbol{q}}$. Another particular case of interest is when $B = J_m$, where J_m is the all-1 matrix of size $m \times m$, which satisfies $\mathrm{Per}(J_m) = m!$. In this case, we obtain, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^m$,

$$\begin{aligned}\sum_{\substack{n\in\mathbb{N},\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m\\|\boldsymbol{p}|=|\boldsymbol{q}|=n}} \frac{n!\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) &= \frac{1}{\mathrm{Det}(I - AYJ_mX)}\\ &= \frac{1}{\mathrm{Det}(I - A\boldsymbol{y}\boldsymbol{x}^T)}\\ &= \frac{1}{1 - \boldsymbol{x}^T A\boldsymbol{y}},\end{aligned} \tag{97}$$

where we used the fact that $A\boldsymbol{y}\boldsymbol{x}^T$ is a rank-one matrix to compute the determinant. This implies Corollary 1: for all $n \in \mathbb{N}^*$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n$,

$$\begin{aligned}\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) &= \frac{\boldsymbol{p}!\boldsymbol{q}!}{n!}[\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}]\left(\frac{1}{1 - \boldsymbol{x}^T A\boldsymbol{y}}\right)\\ &= \frac{\boldsymbol{p}!\boldsymbol{q}!}{n!}[\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}]\left(\boldsymbol{x}^T A\boldsymbol{y}\right)^n,\end{aligned} \tag{98}$$

where we used the Taylor series $\frac{1}{1-z} = \sum_{k=0}^{+\infty} z^k$.

We note that Theorem 1—which is a generalization of the MacMahon master theorem to two matrices—may also be obtained by combining the MacMahon master theorem in Eq. (66) with the Cauchy–Binet theorem in Eq. (58), applied to the matrix AYB^T. As it turns out, we may apply the same proof technique inductively in order to generalize the MacMahon master theorem to N matrices $A^{(1)}, \ldots, A^{(N)}$, for $N \geq 1$ and obtain Theorem 2: assuming that Eq. (81) holds for some $N \geq 2$, we have, for all $m \times m$ matrices $B^{(1)}, \ldots, B^{(N)}$,

$$\begin{aligned}\sum_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N\in\mathbb{N}^m} \prod_{k=1}^{N} \frac{\boldsymbol{z}_k^{\boldsymbol{p}_k}}{\boldsymbol{p}_k!}\mathrm{Per}(B^{(1)}_{\boldsymbol{p}_1,\boldsymbol{p}_2})\mathrm{Per}(B^{(2)}_{\boldsymbol{p}_2,\boldsymbol{p}_3})\ldots\mathrm{Per}(B^{(N)}_{\boldsymbol{p}_N,\boldsymbol{p}_1})\\ = \frac{1}{\mathrm{Det}(I - Z_1B^{(1)}\ldots Z_NB^{(N)})},\end{aligned} \tag{99}$$

where $\boldsymbol{z}_k = (z_{k1}, \ldots, z_{km})$ are formal variables and $Z_k = \mathrm{Diag}(\boldsymbol{z}_k)$ for all $k \in \{1, \ldots, N\}$.

Let $A^{(1)}, \ldots, A^{(N+1)}$ be $m \times m$ matrices, $\boldsymbol{z}_{N+1} = (z_{N+1,1}, \ldots, z_{N+1,m})$ be formal variables and $Z_{N+1} = \mathrm{Diag}(\boldsymbol{z}_{N+1})$. Setting $B^{(N)} = A^{(N)}Z_{N+1}A^{(N+1)}$ and $B^{(k)} = A^{(k)}$ for all $k \in \{1, \ldots, N\}$, we obtain with Eq. (99):

$$\begin{aligned}&\frac{1}{\mathrm{Det}(I - Z_1A^{(1)}\ldots Z_NA^{(N)}Z_{N+1}A^{(N+1)})}\\ &= \sum_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N\in\mathbb{N}^m} \prod_{k=1}^{N} \frac{\boldsymbol{z}_k^{\boldsymbol{p}_k}}{\boldsymbol{p}_k!}\mathrm{Per}(A^{(1)}_{\boldsymbol{p}_1,\boldsymbol{p}_2})\ldots\mathrm{Per}(A^{(N-1)}_{\boldsymbol{p}_{N-1},\boldsymbol{p}_N})\mathrm{Per}((A^{(N)}Z_{N+1}A^{(N+1)})_{\boldsymbol{p}_N,\boldsymbol{p}_1}).\end{aligned} \tag{100}$$

On the other hand, with the Cauchy–Binet theorem from Eq. (58),

$$\begin{aligned}\mathrm{Per}((A^{(N)}Z_{N+1}A^{(N+1)})_{\boldsymbol{p}_N,\boldsymbol{p}_1}) &= \sum_{\boldsymbol{p}_{N+1}\in\mathbb{N}^m} \frac{1}{\boldsymbol{p}_{N+1}!}\mathrm{Per}(A^{(N)}_{\boldsymbol{p}_N,\boldsymbol{p}_{N+1}})\mathrm{Per}((Z_{N+1}A^{(N+1)})_{\boldsymbol{p}_{N+1},\boldsymbol{p}_1})\\ &= \sum_{\boldsymbol{p}_{N+1}\in\mathbb{N}^m} \frac{\boldsymbol{z}_{N+1}^{\boldsymbol{p}_{N+1}}}{\boldsymbol{p}_{N+1}!}\mathrm{Per}(A^{(N)}_{\boldsymbol{p}_N,\boldsymbol{p}_{N+1}})\mathrm{Per}(A^{(N+1)}_{\boldsymbol{p}_{N+1},\boldsymbol{p}_1}),\end{aligned} \tag{101}$$

where we used the fact that multiplying any single row of M by a variable z changes $\mathrm{Per}(M)$ to $z\mathrm{Per}(M)$. Combining Eqs. (100) and (101) completes the induction step and the proof of Theorem 2. □

Applications. Recall that the MacMahon master theorem is a useful tool for deriving combinatorial identities (see section 2.3): it expresses the permanent of an $m \times m$ matrix A with rows and columns repeated in the same way as the coefficient

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \boldsymbol{p}![\boldsymbol{z}^{\boldsymbol{p}}] \left(\frac{1}{\mathrm{Det}(I - ZA)} \right), \tag{102}$$

where $Z = \mathrm{Diag}(\boldsymbol{z})$ and $\boldsymbol{p} \in \mathbb{N}^m$, while the same permanent may also be expressed as the coefficient

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{p}}) = \boldsymbol{p}![\boldsymbol{z}^{\boldsymbol{p}}](A\boldsymbol{z})^{\boldsymbol{p}}. \tag{103}$$

Our generalizations of the MacMahon master theorem may be used in a similar fashion to obtain simple proofs of combinatorial identities. Let us illustrate this with an example: for $n \in \mathbb{N}$ and $a, b \in \mathbb{C}$, setting

$$S_n(a,b) := \sum_{k=0}^{n} \binom{n}{k}^2 a^k b^{n-k}, \tag{104}$$

we aim to prove the relation

$$S_n(a,b)^2 = \sum_{l=0}^{n} \binom{2l}{l} \binom{n+l}{2l} (-1)^{n-l} (a-b)^{2n-2l} S_l(a^2, b^2), \tag{105}$$

which for $a = b = 1$ directly implies the well-known $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$.

Proof. By Theorem 1 we have

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}})\mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}}) = \boldsymbol{p}!\boldsymbol{q}![\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}] \left(\frac{1}{\mathrm{Det}(I - XAYB^T)} \right), \tag{106}$$

where $X = \mathrm{Diag}(x_1, \ldots, x_m)$ and $Y = \mathrm{Diag}(y_1, \ldots, y_m)$. Setting

$$A = B = \begin{pmatrix} 1 & a \\ 1 & b \end{pmatrix}, \tag{107}$$

for $a, b \in \mathbb{C}$, $n \in \mathbb{N}^*$, and $\boldsymbol{p} = \boldsymbol{q} = (n, n)$, we obtain with Eq. (103):

$$\begin{aligned} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) &= (n!)^2 [z_1^n z_2^n] (z_1 + az_2)^n (z_1 + bz_2)^n \\ &= (n!)^2 \sum_{k+l=n} \binom{n}{k} \binom{n}{l} a^k b^l \\ &= (n!)^2 S_n(a,b). \end{aligned} \tag{108}$$

On the other hand, with Eq. (106) we have

$$\begin{aligned} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}})^2 &= (n!)^4 [x_1^n x_2^n y_1^n y_2^n] \left(\frac{1}{\mathrm{Det}(I - \mathrm{Diag}(x_1, x_2) A \mathrm{Diag}(y_1, y_2) A^T)} \right) \\ &= (n!)^4 [x_1^n x_2^n y_1^n y_2^n] \left(\frac{1}{1 - x_1 y_1 - a^2 x_1 y_2 - x_2 y_1 - b^2 x_2 y_2 + (a-b)^2 x_1 x_2 y_1 y_2} \right) \\ &= (n!)^4 \sum_{k=n}^{2n} [x_1^n x_2^n y_1^n y_2^n] \left(x_1 y_1 + a^2 x_1 y_2 + x_2 y_1 + b^2 x_2 y_2 - (a-b)^2 x_1 x_2 y_1 y_2 \right)^k \end{aligned} \tag{109}$$

$$= (n!)^4 \sum_{k=n}^{2n} \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} (a-b)^{2k-2j} \sum_{n_{11}+n_{12}+n_{21}+n_{22}=j} \frac{j! a^{2n_{12}} b^{2n_{22}}}{n_{11}! n_{12}! n_{21}! n_{22}!}$$
$$\times [x_1^n x_2^n y_1^n y_2^n] \left((x_1 x_2 y_1 y_2)^{k-j} x_1^{n_{11}+n_{12}} x_2^{n_{21}+n_{22}} y_1^{n_{11}+n_{21}} y_2^{n_{12}+n_{22}} \right),$$

where we used the Taylor expansion of $\frac{1}{1-z}$ in the third line and the multinomial theorem in the last line. The indices in the above expression must satisfy $n_{11} = n_{22}$, $n_{12} = n_{21}$, which implies $j = 2n_{11} + 2n_{12}$, i.e. j is even. Moreover, $n = k - j + n_{11} + n_{12} = k - j/2$. Relabeling $j = 2l$ and $n_{11} = p$ we have $k = l + n$ and we obtain

$$\begin{aligned} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}})^2 &= (n!)^4 \sum_{l=0}^{n} \binom{n+l}{2l} (-1)^{n-l} (a-b)^{2n-2l} \sum_{p=0}^{l} \frac{(2l)! a^{2l-2p} b^{2p}}{p!^2 (l-p)!^2} \\ &= (n!)^4 \sum_{l=0}^{n} \binom{2l}{l} \binom{n+l}{2l} (-1)^{n-l} (a-b)^{2n-2l} S_l(a^2, b^2). \end{aligned} \tag{110}$$

With Eq. (108) this proves

$$S_n(a,b)^2 = \sum_{l=0}^{n} \binom{2l}{l} \binom{n+l}{2l} (-1)^{n-l} (a-b)^{2n-2l} S_l(a^2, b^2). \tag{111}$$

□

6.2 New generating functions

We have encountered several generating functions for the permanent of an $m \times m$ matrix A with differently repeated rows and columns. Jackson's formula gives [46]:

$$e^{\boldsymbol{x}^T A \boldsymbol{y}} = \sum_{\substack{n \in \mathbb{N}, \boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}), \tag{112}$$

for formal variables $\boldsymbol{x} = (x_1, \ldots, x_m), \boldsymbol{y} = (y_1, \ldots, y_m)$. Moreover, we have obtained in Eq. (97), as a corollary of our generalization of the MacMahon master theorem:

$$\frac{1}{1 - \boldsymbol{x}^T A \boldsymbol{y}} = \sum_{\substack{n \in \mathbb{N}, \boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} n! \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}). \tag{113}$$

Furthermore, Corollary 1 gives:

$$(\boldsymbol{x}^T A \boldsymbol{y})^n = \sum_{\substack{\boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} n! \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}). \tag{114}$$

In this section we prove Theorem 3, which generalizes all three statements: for any series $f(z) = \sum_{n=0}^{+\infty} f_n z^n$,

$$f(\boldsymbol{x}^T A \boldsymbol{y}) = \sum_{\substack{n \in \mathbb{N}, \boldsymbol{p},\boldsymbol{q} \in \mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} f_n n! \frac{\boldsymbol{x}^{\boldsymbol{p}} \boldsymbol{y}^{\boldsymbol{q}}}{\boldsymbol{p}! \boldsymbol{q}!} \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}). \tag{115}$$

One may prove this statement by linearity, using Eq. (114). In what follows, we give a direct quantum-inspired proof which is similar to that of Jackson's formula from section 5.4.

Proof of Theorem 3. For $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}^m$ and a unitary matrix U,

$$\begin{aligned}
\sum_{\substack{n\in\mathbb{N},\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} f_n n! \frac{\boldsymbol{\alpha}^{\boldsymbol{p}}\boldsymbol{\beta}^{\boldsymbol{q}}}{\boldsymbol{p}!\boldsymbol{q}!} \mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{q}}) &= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \sum_{n=0}^{+\infty} f_n n! \sum_{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{N}^m \\ |\boldsymbol{p}|=|\boldsymbol{q}|=n}} \langle \boldsymbol{\alpha}^* | \boldsymbol{p}\rangle\langle \boldsymbol{p}| \hat{U} |\boldsymbol{q}\rangle\langle \boldsymbol{q}| \boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \sum_{n=0}^{+\infty} f_n n! \, \langle \boldsymbol{\alpha}^*| \Pi_n \hat{U} \Pi_n |\boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \sum_{n=0}^{+\infty} f_n n! \, \langle \boldsymbol{\alpha}^*| \Pi_n \hat{U} |\boldsymbol{\beta}\rangle \\
&= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \sum_{n=0}^{+\infty} f_n n! \, \langle \boldsymbol{\alpha}^*| \Pi_n |U\boldsymbol{\beta}\rangle \,,
\end{aligned} \tag{116}$$

where we used the Fock basis expansions of coherent states (6) and of $\hat{U}$ (10) in the first line, the definition of the projector $\Pi_n = \sum_{|\boldsymbol{p}|=n} |\boldsymbol{p}\rangle\langle \boldsymbol{p}|$ in the second line, the fact that passive linear operations conserve the total number of photons (12) in the third line, and the action of $\hat{U}$ on coherent states (13) in the fourth line. Now,

$$\begin{aligned}
e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \langle \boldsymbol{\alpha}^*| \Pi_n |U\boldsymbol{\beta}\rangle &= e^{\frac{1}{2}\|\boldsymbol{\alpha}\|^2+\frac{1}{2}\|\boldsymbol{\beta}\|^2} \sum_{|\boldsymbol{p}|=n} \langle \boldsymbol{\alpha}^* | \boldsymbol{p}\rangle\langle \boldsymbol{p}| U\boldsymbol{\beta}\rangle \\
&= \sum_{|\boldsymbol{p}|=n} e^{\frac{1}{2}\|\boldsymbol{\beta}\|^2-\frac{1}{2}\|U\boldsymbol{\beta}\|^2} \frac{\boldsymbol{\alpha}^{\boldsymbol{p}}(U\boldsymbol{\beta})^{\boldsymbol{p}}}{\boldsymbol{p}!} \\
&= \sum_{|\boldsymbol{p}|=n} \frac{\boldsymbol{\alpha}^{\boldsymbol{p}}(U\boldsymbol{\beta})^{\boldsymbol{p}}}{\boldsymbol{p}!} \\
&= \frac{1}{n!}(\boldsymbol{\alpha}^T U \boldsymbol{\beta})^n,
\end{aligned} \tag{117}$$

where we used the Fock basis expansion of coherent states (6) in the second line, the fact that U is unitary in the third line, and the multinomial theorem in the last line. Combining Eqs. (116) and (117) completes the proof for unitary matrices. Once again, the relation for a generic nonzero matrix A is obtained by embedding $\frac{1}{\|A\|}A$ as a submatrix of a unitary matrix U.

As a result, when $f_n \neq 0$ we have, for any $m \times m$ matrix A, all $n \in \mathbb{N}$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ such that $|\boldsymbol{p}| = |\boldsymbol{q}| = n$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{T}^m} \left[\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}} \frac{f(\boldsymbol{x}^T A \boldsymbol{y})}{\partial_z^n f(z)\big|_{z=0}} \right] &= \frac{\boldsymbol{p}!\boldsymbol{q}!}{\partial_z^n f(z)\big|_{z=0}} \sum_{\substack{k\in\mathbb{N},\boldsymbol{s},\boldsymbol{t}\in\mathbb{N}^m \\ |\boldsymbol{s}|=|\boldsymbol{t}|=k}} f_k k! \frac{\mathrm{Per}(A_{\boldsymbol{s},\boldsymbol{t}})}{\boldsymbol{s}!\boldsymbol{t}!} \\
&\quad \times \prod_{j=1}^m \left(\int_{\theta_j,\varphi_j\in[0,2\pi]} e^{i\theta_j(s_j-p_j)} e^{i\varphi_j(t_j-q_j)} \frac{d\theta_j d\varphi_j}{4\pi^2} \right) \\
&= \frac{1}{\partial_z^n f(z)\big|_{z=0}} f_n n! \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) \\
&= \mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}),
\end{aligned} \tag{118}$$

where we have set $\boldsymbol{x} = (e^{i\theta_1}, \ldots, e^{i\theta_m}) \in \mathbb{T}^m$ and $\boldsymbol{y} = (e^{i\varphi_1}, \ldots, e^{i\varphi_m}) \in \mathbb{T}^m$ in the first line, and where we have used $\int_0^{2\pi} e^{i(k-l)\theta} \frac{d\theta}{2\pi} = \delta_{kl}$ in the third line. □

Applications. These generating functions may be used to derive remarkable identities for the permanent: for instance, for A and B two $m \times m$ matrices, taking $f(z) = e^z$ and equating the $\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}$ coefficients of $e^{\boldsymbol{x}^T(A+B)\boldsymbol{y}}$ and $e^{\boldsymbol{x}^T A\boldsymbol{y}}e^{\boldsymbol{x}^T B\boldsymbol{y}}$ yields the sum formula [1, 2]:

$$\mathrm{Per}((A+B)_{\boldsymbol{p},\boldsymbol{q}}) = \sum_{\substack{\boldsymbol{s}+\boldsymbol{t}=\boldsymbol{p}\\ \boldsymbol{u}+\boldsymbol{v}=\boldsymbol{q}}} \frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!\boldsymbol{u}!\boldsymbol{v}!}\mathrm{Per}(A_{\boldsymbol{s},\boldsymbol{u}})\mathrm{Per}(B_{\boldsymbol{t},\boldsymbol{v}}), \tag{119}$$

for all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$.

Similarly, for A an $m \times m$ matrix, taking $f(z) = z^p$ for $p = k, l, k+l$ and equating the $\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}$ coefficients of $(\boldsymbol{x}^T A\boldsymbol{y})^{k+l}$ and $(\boldsymbol{x}^T A\boldsymbol{y})^k(\boldsymbol{x}^T A\boldsymbol{y})^l$ yields the Laplace expansion formula [1, 2]:

$$\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) = \frac{k!l!}{(k+l)!} \sum_{\substack{\boldsymbol{s}+\boldsymbol{t}=\boldsymbol{p}\\ \boldsymbol{u}+\boldsymbol{v}=\boldsymbol{q}\\ |\boldsymbol{s}|=|\boldsymbol{u}|=k\\ |\boldsymbol{t}|=|\boldsymbol{v}|=l}} \frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!\boldsymbol{u}!\boldsymbol{v}!}\mathrm{Per}(A_{\boldsymbol{s},\boldsymbol{u}})\mathrm{Per}(A_{\boldsymbol{t},\boldsymbol{v}}), \tag{120}$$

for all $k, l \in \mathbb{N}$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}^m$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = k + l$.

Finally, for A and B two $m \times m$ matrices, taking $f(z) = -\log(1-z)$ and equating the $\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}$ coefficients of $-\log\left(1-\boldsymbol{x}^T A\boldsymbol{y}\right) - \log\left(1-\boldsymbol{x}^T B\boldsymbol{y}\right)$ and $-\log[(1-\boldsymbol{x}^T B\boldsymbol{y})(1-\boldsymbol{x}^T A\boldsymbol{y})]$ yields Theorem 4:

$$\begin{aligned}&\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) + \mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}})\\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^k}{\binom{n-1}{k}} \sum_{\substack{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{p}\\ \boldsymbol{a}'+\boldsymbol{b}'+\boldsymbol{c}'=\boldsymbol{q}\\ |\boldsymbol{a}|=|\boldsymbol{b}|=|\boldsymbol{a}'|=|\boldsymbol{b}'|=k}} \frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{a}!\boldsymbol{b}!\boldsymbol{c}!\boldsymbol{a}'!\boldsymbol{b}'!\boldsymbol{c}'!}\mathrm{Per}(A_{\boldsymbol{a},\boldsymbol{a}'})\mathrm{Per}(B_{\boldsymbol{b},\boldsymbol{b}'})\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'}),\end{aligned} \tag{121}$$

after a derivation which we detail below.

Proof of Theorem 4. Let us write

$$-\log[(1-\boldsymbol{x}^T B\boldsymbol{y})(1-\boldsymbol{x}^T A\boldsymbol{y})] = -\log[1-\boldsymbol{x}^T(A+B-B\boldsymbol{y}\boldsymbol{x}^T A)\boldsymbol{y})]. \tag{122}$$

Applying Theorem 3 for $f(z) = -\log(1-z)$ with $z = \boldsymbol{x}^T B\boldsymbol{y}, \boldsymbol{x}^T A\boldsymbol{y}, \boldsymbol{x}^T(A+B-B\boldsymbol{y}\boldsymbol{x}^T A)\boldsymbol{y}$ and considering the $\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}$ coefficient we obtain for all $n \in \mathbb{N}^*$ and all $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{N}$ with $|\boldsymbol{p}| = |\boldsymbol{q}| = n$:

$$\begin{aligned}&\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) + \mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}})\\ &= \frac{\boldsymbol{p}!\boldsymbol{q}!}{(n-1)!}[\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}]\left(\sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l}} (l-1)!\frac{\boldsymbol{x}^{\boldsymbol{s}}\boldsymbol{y}^{\boldsymbol{t}}}{\boldsymbol{s}!\boldsymbol{t}!}\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}})\right)\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l}} \frac{(l-1)!}{(n-1)!}\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!}[\boldsymbol{x}^{\boldsymbol{p}}\boldsymbol{y}^{\boldsymbol{q}}]\left(\boldsymbol{x}^{\boldsymbol{s}}\boldsymbol{y}^{\boldsymbol{t}}\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}})\right)\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}\le\boldsymbol{p},\boldsymbol{t}\le\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!}[\boldsymbol{x}^{\boldsymbol{p}-\boldsymbol{s}}\boldsymbol{y}^{\boldsymbol{q}-\boldsymbol{t}}]\left(\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}})\right)\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}+\boldsymbol{u}=\boldsymbol{p}\\ \boldsymbol{t}+\boldsymbol{v}=\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!}[\boldsymbol{x}^{\boldsymbol{u}}\boldsymbol{y}^{\boldsymbol{v}}]\left(\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}})\right).\end{aligned} \tag{123}$$

Applying the sum formula from Eq. (119) to the matrices $(A+B)$ and $(-B\boldsymbol{y}\boldsymbol{x}^T A)$ we obtain, for all $l \in \mathbb{N}^*$ and all $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{N}$ with $|\boldsymbol{s}| = |\boldsymbol{t}| = l$,

$$\begin{aligned}\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}}) &= \sum_{\substack{\boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|,|\boldsymbol{k}|=|\boldsymbol{k}'|}} \frac{\boldsymbol{s}!\boldsymbol{t}!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})\mathrm{Per}((-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{k},\boldsymbol{k}'})\\ &= \sum_{\substack{\boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|,|\boldsymbol{k}|=|\boldsymbol{k}'|}} \frac{(-1)^{|\boldsymbol{k}|}\boldsymbol{s}!\boldsymbol{t}!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})\mathrm{Per}((B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{k},\boldsymbol{k}'}).\end{aligned} \tag{124}$$

The permanent of the outer product $\boldsymbol{v}\boldsymbol{w}^T$ of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ of the same size k is easily computed as

$$\begin{aligned}\mathrm{Per}(\boldsymbol{v}\boldsymbol{w}^T) &= \sum_{\sigma\in\mathcal{S}_k}\prod_{j=1}^{k} v_j w_{\sigma(j)}\\ &= \sum_{\sigma\in\mathcal{S}_k} v_1\dots v_k w_1\dots w_k\\ &= k!\, v_1\dots v_k w_1\dots w_k.\end{aligned} \tag{125}$$

The matrix $(B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{k},\boldsymbol{k}'}$ is the outer product between the vector $(B\boldsymbol{y})_{\boldsymbol{k}}$ (obtained from the vector $B\boldsymbol{y}$ by repeating k_i times its i^{th} entry) and the vector $(A^T\boldsymbol{x})_{\boldsymbol{k}'}$ (obtained from the vector $A^T\boldsymbol{x}$ by repeating k'_i times its i^{th} entry) of size $|\boldsymbol{k}|$, so its permanent is given by

$$\mathrm{Per}((B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{k},\boldsymbol{k}'}) = (|\boldsymbol{k}|)!(B\boldsymbol{y})^{\boldsymbol{k}}(A^T\boldsymbol{x})^{\boldsymbol{k}'}. \tag{126}$$

With Eq. (124) we obtain

$$\mathrm{Per}((A+B-B\boldsymbol{y}\boldsymbol{x}^T A)_{\boldsymbol{s},\boldsymbol{t}} = \sum_{\substack{\boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|,|\boldsymbol{k}|=|\boldsymbol{k}'|}} \frac{(-1)^{|\boldsymbol{k}|}(|\boldsymbol{k}|)!\boldsymbol{s}!\boldsymbol{t}!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})(B\boldsymbol{y})^{\boldsymbol{k}}(A^T\boldsymbol{x})^{\boldsymbol{k}'}. \tag{127}$$

Plugging this expression into Eq. (123) yields

$$\begin{aligned}&\mathrm{Per}(A_{\boldsymbol{p},\boldsymbol{q}}) + \mathrm{Per}(B_{\boldsymbol{p},\boldsymbol{q}})\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}+\boldsymbol{u}=\boldsymbol{p}\\ \boldsymbol{t}+\boldsymbol{v}=\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{s}!\boldsymbol{t}!}\sum_{\substack{\boldsymbol{c},\boldsymbol{k},\boldsymbol{c}',\boldsymbol{k}'\in\mathbb{N}^m\\ \boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|,|\boldsymbol{k}|=|\boldsymbol{k}'|}} \frac{(-1)^{|\boldsymbol{k}|}(|\boldsymbol{k}|)!\boldsymbol{s}!\boldsymbol{t}!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})[\boldsymbol{x}^{\boldsymbol{u}}\boldsymbol{y}^{\boldsymbol{v}}](B\boldsymbol{y})^{\boldsymbol{k}}(A^T\boldsymbol{x})^{\boldsymbol{k}'}\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}+\boldsymbol{u}=\boldsymbol{p}\\ \boldsymbol{t}+\boldsymbol{v}=\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\boldsymbol{p}!\boldsymbol{q}!\sum_{\substack{\boldsymbol{c},\boldsymbol{k},\boldsymbol{c}',\boldsymbol{k}'\in\mathbb{N}^m\\ \boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|,|\boldsymbol{k}|=|\boldsymbol{k}'|}} \frac{(-1)^{|\boldsymbol{k}|}(|\boldsymbol{k}|)!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})[\boldsymbol{x}^{\boldsymbol{u}}](A^T\boldsymbol{x})^{\boldsymbol{k}'}[\boldsymbol{y}^{\boldsymbol{v}}](B\boldsymbol{y})^{\boldsymbol{k}}\\ &= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}+\boldsymbol{u}=\boldsymbol{p}\\ \boldsymbol{t}+\boldsymbol{v}=\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\boldsymbol{p}!\boldsymbol{q}!\sum_{\substack{\boldsymbol{c},\boldsymbol{k},\boldsymbol{c}',\boldsymbol{k}'\in\mathbb{N}^m\\ \boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|\\ |\boldsymbol{k}|=|\boldsymbol{k}'|=|\boldsymbol{u}|=|\boldsymbol{v}|}} \frac{(-1)^{|\boldsymbol{k}|}(|\boldsymbol{k}|)!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})\frac{\mathrm{Per}(A^T_{\boldsymbol{k}',\boldsymbol{u}})}{\boldsymbol{u}!}\frac{\mathrm{Per}(B_{\boldsymbol{k},\boldsymbol{v}})}{\boldsymbol{v}!}\end{aligned} \tag{128}$$

$$= \sum_{\substack{l\in\mathbb{N}^*,\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v}\in\mathbb{N}^m\\ |\boldsymbol{s}|=|\boldsymbol{t}|=l\\ \boldsymbol{s}+\boldsymbol{u}=\boldsymbol{p}\\ \boldsymbol{t}+\boldsymbol{v}=\boldsymbol{q}}} \frac{(l-1)!}{(n-1)!}\frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{u}!\boldsymbol{v}!} \sum_{\substack{\boldsymbol{c},\boldsymbol{k},\boldsymbol{c}',\boldsymbol{k}'\in\mathbb{N}^m\\ \boldsymbol{c}+\boldsymbol{k}=\boldsymbol{s}\\ \boldsymbol{c}'+\boldsymbol{k}'=\boldsymbol{t}\\ |\boldsymbol{c}|=|\boldsymbol{c}'|\\ |\boldsymbol{k}|=|\boldsymbol{k}'|=|\boldsymbol{u}|=|\boldsymbol{v}|}} \frac{(-1)^{|\boldsymbol{k}|}(|\boldsymbol{k}|)!}{\boldsymbol{c}!\boldsymbol{k}!\boldsymbol{c}'!\boldsymbol{k}'!}\mathrm{Per}(A_{\boldsymbol{u},\boldsymbol{k}'})\mathrm{Per}(B_{\boldsymbol{k},\boldsymbol{v}})\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'})$$

$$= \sum_{k=0}^{\lfloor\frac{n}{2}\rfloor} \frac{(-1)^k k!(n-k-1)!}{(n-1)!} \sum_{\substack{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}'\in\mathbb{N}^m\\ \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{p}\\ \boldsymbol{a}'+\boldsymbol{b}'+\boldsymbol{c}'=\boldsymbol{q}\\ |\boldsymbol{a}|=|\boldsymbol{a}'|=|\boldsymbol{b}|=|\boldsymbol{b}'|=k}} \frac{\boldsymbol{p}!\boldsymbol{q}!}{\boldsymbol{a}!\boldsymbol{b}!\boldsymbol{c}!\boldsymbol{a}'!\boldsymbol{b}'!\boldsymbol{c}'!}\mathrm{Per}(A_{\boldsymbol{a},\boldsymbol{a}'})\mathrm{Per}(B_{\boldsymbol{b},\boldsymbol{b}'})\mathrm{Per}((A+B)_{\boldsymbol{c},\boldsymbol{c}'}),$$

where we used Eq. (61) in the third line, $\mathrm{Per}(M^T) = \mathrm{Per}(M)$ in the fourth line, and where we relabeled $\boldsymbol{u}$ as $\boldsymbol{a}$, $\boldsymbol{k}'$ as $\boldsymbol{a}'$, $\boldsymbol{k}$ as $\boldsymbol{b}$, $\boldsymbol{v}$ as $\boldsymbol{b}'$, and $|\boldsymbol{k}| = |\boldsymbol{k}'| = |\boldsymbol{u}| = |\boldsymbol{v}|$ as $k = n - l$ in the last line. The cutoff of the sum at $\lfloor\frac{n}{2}\rfloor$ comes from the fact that $n = |\boldsymbol{p}| = |\boldsymbol{a}| + |\boldsymbol{b}| + |\boldsymbol{c}| = 2k + |\boldsymbol{c}| \geq 2k$. Writing $\frac{(-1)^k k!(n-k-1)!}{(n-1)!} = \frac{(-1)^k}{\binom{n-1}{k}}$ completes the proof of Theorem 4. □

7 Boson Sampling with input cat states

In this section, we discuss the classical complexity of sampling from the output distribution of linear optical computations with input cat states, which we refer to as Boson Sampling with input cat states.

Since the introduction of Boson Sampling by Aaronson and Arkhipov [6] for the demonstration of quantum computational advantage using noninteracting bosons, several variants of this model have been analyzed [48–55]. These variants were introduced to address two different challenges: on the one hand, to reduce the experimental burden associated with the demonstration of quantum advantage; on the other hand, to understand the resources for this quantum advantage and which computational models are able to reproduce it. Boson Sampling with input cat states is another such variant and has been first considered in [21], where its classical hardness was shown for exact sampling and argued for approximate sampling. We strengthen these results in the following.

We first give a proof of Lemma 1, which provides a closed form expression for output amplitudes of Boson Sampling with input cat states.

Proof of Lemma 1. In the quantum-inspired proof of Glynn's formula in section 5.1 we have obtained the following identity:

$$\langle p_1 \dots p_m|\hat{U}(|\tilde{\mathrm{cat}}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}) = \frac{\alpha^{|\boldsymbol{p}|-n}e^{-\frac{n}{2}|\alpha|^2}}{2^n\sqrt{\boldsymbol{p}!}} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1\dots x_n \prod_{i=1}^m \left(\sum_{j=1}^n u_{ij}x_j\right)^{p_i}, \tag{129}$$

for all $\alpha \in \mathbb{C}$, all $\boldsymbol{p} = (p_1, \dots, p_m) \in \mathbb{N}^m$, all $n \leq m$ and any passive linear operation $\hat{U}$ over m modes with unitary matrix $U = (u_{ij})_{1\leq i,j\leq m}$, where $|\tilde{\mathrm{cat}}_\alpha\rangle = \frac{1}{2\alpha}(|\alpha\rangle - |-\alpha\rangle)$ is an unnormalized cat state. By definition of the cat states (5), these states are related to normalized cat states $|\mathrm{cat}_\alpha\rangle$ by

$$|\mathrm{cat}_\alpha\rangle = \frac{\alpha e^{\frac{1}{2}|\alpha|^2}}{\sqrt{\sinh(|\alpha|^2)}}\,|\tilde{\mathrm{cat}}_\alpha\rangle\,. \tag{130}$$

Setting $|\boldsymbol{p}| = n$ and replacing unnormalized cat states by normalized ones in Eq. (129), we obtain Lemma 1:

$$\begin{aligned}\langle p_1 \dots p_m|\hat{U}(|\mathrm{cat}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}) &= \frac{\alpha^n}{\sqrt{\sinh^n(|\alpha|^2)}2^n\sqrt{\boldsymbol{p}!}} \sum_{\boldsymbol{x}\in\{-1,1\}^n} x_1 \dots x_n \prod_{i=1}^{m}\left(\sum_{j=1}^{n} u_{ij}x_j\right)^{p_i} \\ &= \frac{\alpha^n}{\sqrt{\sinh^n(|\alpha|^2)}} \frac{\mathrm{Per}(U_{\boldsymbol{p},\boldsymbol{1}\oplus\boldsymbol{0}})}{\sqrt{\boldsymbol{p}!}} \\ &= \frac{\alpha^n}{\sqrt{\sinh^n(|\alpha|^2)}} \langle p_1 \dots p_m|\hat{U}(|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}),\end{aligned} \tag{131}$$

where we used Eq. (42) in the second line and the Fock basis expansion of $\hat{U}$ (10) in the last line. □

Interestingly, Lemma 1 implies that, up to a global factor, n cat states of amplitude α reproduce exactly the Boson Sampling statistics of n single photons (however, detection events $\boldsymbol{p}$ with $|\boldsymbol{p}| > n$ can occur for input cat states, contrary to single-photons).

This result has two consequences, summarized by Theorem 5. Firstly, it implies that Boson Sampling with input cat states is hard to sample exactly for $m \geq 2n$ and for all choices of nonzero cat state amplitude unless the polynomial hierarchy collapses to its third level, since some of its outcome probabilities are #P-hard to estimate multiplicatively [6]. Alternative proofs of this statement based either on universality under post-selection or on rejection sampling can be found in [21]. Secondly, as we show hereafter, it implies that Boson Sampling with input cat states of small enough (nonzero) amplitudes is also hard to sample approximately in the same regime as Boson Sampling.

Proof of Theorem 5. Our proof uses arguments similar to the ones used in [21], extended to the case of approximate sampling.

At a high level, we show how to convert any classical algorithm for approximate sampling from the output probability distribution of an m-mode Boson Sampling computation with n input cat states to a classical algorithm for approximate sampling from the output probability distribution of an m-mode Boson Sampling computation with n input Fock states (Lemma 2). We do so by performing rejection sampling: we run the first algorithm and only keep the samples with total photon number equal to n. In particular, we show that if the fraction of samples with correct photon number n is large enough (namely, at least inverse polynomial in m), the rejection sampling subroutine is efficient and the final classical probability distribution is close to the ideal distribution of Boson Sampling with Fock state input. Then, we determine the regime of input cat state amplitude such that the fraction of samples with correct photon number n is at least inverse polynomial in m (Lemma 3).

Let $P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha) := |\langle p_1 \dots p_m|\hat{U}(|\mathrm{cat}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n})|^2$ be the probability of detecting $\boldsymbol{p} = (p_1,\dots,p_m)$ output photons in a Boson Sampling experiment with interferometer $\hat{U}$ and input cat states $|\mathrm{cat}_\alpha\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}$, and let $P_{\mathrm{BS}}(\boldsymbol{p}|n) := |\langle p_1 \dots p_m|\hat{U}(|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n})|^2$ be the probability of detecting $\boldsymbol{p} = (p_1,\dots,p_m)$ output photons in a Boson Sampling experiment with interferometer $\hat{U}$ and input single photons $|1\rangle^{\otimes n} \otimes |0\rangle^{\otimes m-n}$.

We show the following reduction: as long as the fraction of samples with photon number n is large enough, then any efficient classical algorithm for approximate sampling from P_{cat} can be converted to an efficient classical algorithm for approximate sampling from P_{BS}.

Lemma 2. *Suppose there exists an efficient classical algorithm C for approximate sampling from P_{cat}, i.e., that takes as input the description of the m-mode boson sampler with n input cat states with output probability distribution P_{cat} and an error bound ϵ, and that samples from a distribution P_C such that $\|P_C - P_{cat}\|_{TV} \leq \epsilon$ in poly $(m, \frac{1}{\epsilon})$ time, where $\|\cdot\|_{TV}$ is the total variation distance.*

Assume further that the fraction of samples from P_{cat} with photon number equal to n is at least inverse polynomial in m, i.e. $\sum_{|\boldsymbol{p}|=n} P_{cat}(\boldsymbol{p}|n,\alpha) \geq \frac{1}{poly\ m}$.

Then, there exists an efficient classical algorithm $\tilde{C}$ for approximate sampling from P_{BS}, i.e., a classical algorithm that takes as input the description of the m-mode boson sampler with n input Fock states with output probability distribution P_{BS} and an error bound ϵ, and that samples from a distribution P_C^{rej} such that $\|P_C^{rej} - P_{BS}\|_{TV} \leq \epsilon$ in poly $(m, \frac{1}{\epsilon})$ time.

Proof of Lemma 2. Given a sampling algorithm from a probability distribution P over $\mathbb{N}^m$, we define the following rejection sampling subroutine: sample $\boldsymbol{p}$ from P and compute $|\boldsymbol{p}|$; discard the sample if $|\boldsymbol{p}| \neq n$; otherwise, output $\boldsymbol{p}$. We denote by P^{rej} the corresponding output probability distribution. For all $\boldsymbol{p} \in \mathbb{N}^m$, we have

$$P^{\text{rej}}(\boldsymbol{p}) = \frac{\delta_{|\boldsymbol{p}|,n} P(\boldsymbol{p})}{\sum_{|\boldsymbol{q}|=n} P(\boldsymbol{q})}, \tag{132}$$

where δ is the Kronecker symbol. Applying this subroutine to P_{cat} gives a new probability distribution

$$P_{\text{cat}}^{\text{rej}}(\boldsymbol{p}|n,\alpha) = \frac{\delta_{|\boldsymbol{p}|,n} P_{\text{cat}}(\boldsymbol{p}|n,\alpha)}{\sum_{|\boldsymbol{q}|=n} P_{\text{cat}}(\boldsymbol{q}|n,\alpha)}. \tag{133}$$

By Lemma 1 (see Eq. (131)) we have, for all $\boldsymbol{p}$ such that $|\boldsymbol{p}| = n$,

$$P_{\text{cat}}(\boldsymbol{p}|n,\alpha) = \frac{|\alpha|^{2n}}{\sinh^n(|\alpha|^2)} P_{\text{BS}}(\boldsymbol{p}|n). \tag{134}$$

Hence

$$\begin{aligned} \sum_{|\boldsymbol{p}|=n} P_{\text{cat}}(\boldsymbol{p}|n,\alpha) &= \frac{|\alpha|^{2n}}{\sinh^n(|\alpha|^2)} \sum_{|\boldsymbol{p}|=n} P_{\text{BS}}(\boldsymbol{p}|n) \\ &= \frac{|\alpha|^{2n}}{\sinh^n(|\alpha|^2)}, \end{aligned} \tag{135}$$

where we used $\sum_{|\boldsymbol{p}|=n} P_{\text{BS}}(\boldsymbol{p}|n) = 1$ in the second line since $\hat{U}$ conserves the total number of photons (12). Combining Eqs. (133-135), we thus obtain

$$P_{\text{BS}}(\boldsymbol{p}|n) = P_{\text{cat}}^{\text{rej}}(\boldsymbol{p}|n,\alpha). \tag{136}$$

This reproduces the argument for exact sampling hardness from [21].

To show approximate hardness, we define the classical approximate sampling algorithm $\tilde{C}$ by running the algorithm C with input error bound $\frac{\epsilon}{2} \sum_{|\boldsymbol{p}|=n} P_{\text{cat}}(\boldsymbol{p}|n,\alpha)$ and applying the rejection sampling subroutine. By assumption, running the algorithm C can be done in time

$$\text{poly}\left(m, \frac{2}{\epsilon \sum_{|\boldsymbol{p}|=n} P_{\text{cat}}(\boldsymbol{p}|n,\alpha)}\right) = \text{poly}\left(m, \frac{1}{\epsilon}\right), \tag{137}$$

since $\sum_{|\boldsymbol{p}|=n} P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha) \geq \frac{1}{\mathrm{poly}\ m}$, also by assumption. On the other hand, the rejection sampling subroutine induces a computational overhead scaling with the inverse of the fraction of kept samples $\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q})$. We have

$$\begin{aligned}
\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha) - \sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q}) &\leq \left| \sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q}) - \sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha) \right| \\
&\leq \sum_{|\boldsymbol{q}|=n} |P_C(\boldsymbol{q}) - P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)| \\
&\leq \sum_{\boldsymbol{q}\in\mathbb{N}^m} |P_C(\boldsymbol{q}) - P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)| \\
&= 2\|P_C - P_{\mathrm{cat}}\|_{\mathrm{TV}} \\
&\leq \epsilon \sum_{|\boldsymbol{p}|=n} P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha),
\end{aligned} \tag{138}$$

where we used the triangle inequality in the second line, the definition of the total variation distance in the fourth line, and the definition of C in the last line. Hence,

$$\begin{aligned}
\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q}) &\geq (1-\epsilon) \sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha) \\
&\geq \frac{1-\epsilon}{\mathrm{poly}\ m},
\end{aligned} \tag{139}$$

so that the computational overhead induced by the rejection sampling subroutine scales as $\mathrm{poly}(m, \frac{1}{\epsilon})$.

Overall, the classical algorithm $\tilde{C}$ outputs samples from the probability distribution P_C^{rej} in $\mathrm{poly}(m, \frac{1}{\epsilon})$ time (with probability exponentially close to 1). Moreover,

$$\begin{aligned}
\|P_C^{\mathrm{rej}} - P_{\mathrm{BS}}\|_{\mathrm{TV}} &= \frac{1}{2} \sum_{\boldsymbol{p}\in\mathbb{N}^m} |P_C^{\mathrm{rej}}(\boldsymbol{p}) - P_{\mathrm{BS}}(\boldsymbol{p}|n)| \\
&= \frac{1}{2} \sum_{|\boldsymbol{p}|=n} |P_C^{\mathrm{rej}}(\boldsymbol{p}) - P_{\mathrm{BS}}(\boldsymbol{p}|n)| \\
&= \frac{1}{2} \sum_{|\boldsymbol{p}|=n} \left| P_C^{\mathrm{rej}}(\boldsymbol{p}) - P_{\mathrm{cat}}^{\mathrm{rej}}(\boldsymbol{p}|n,\alpha) \right| \\
&= \frac{1}{2} \sum_{|\boldsymbol{p}|=n} \left| \frac{P_C(\boldsymbol{p})}{\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q})} - \frac{P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha)}{\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \right| \\
&\leq \frac{1}{2} \sum_{|\boldsymbol{p}|=n} \left[\frac{|P_C(\boldsymbol{p}) - P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha)|}{\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \right. \\
&\qquad \left. + P_C(\boldsymbol{p}) \left| \frac{1}{\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q})} - \frac{1}{\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \right| \right] \\
&\leq \frac{1}{2\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \sum_{|\boldsymbol{p}|=n} \left[|P_C(\boldsymbol{p}) - P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha)| \right. \\
&\qquad \left. + \frac{P_C(\boldsymbol{p})}{\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q})} \left| \sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha) - \sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q}) \right| \right]
\end{aligned} \tag{140}$$

$$
\begin{aligned}
&\leq \frac{1}{\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \Bigg[\frac{1}{2} \sum_{\boldsymbol{p}\in\mathbb{N}^m} |P_C(\boldsymbol{p}) - P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha)| \\
&\qquad\qquad + \sum_{|\boldsymbol{p}|=n} \frac{P_C(\boldsymbol{p})}{\sum_{|\boldsymbol{q}|=n} P_C(\boldsymbol{q})} \frac{1}{2} \sum_{\boldsymbol{q}\in\mathbb{N}^m} |P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha) - P_C(\boldsymbol{q})| \Bigg] \\
&= \frac{2\|P_C - P_{\mathrm{cat}}\|_{\mathrm{TV}}}{\sum_{|\boldsymbol{q}|=n} P_{\mathrm{cat}}(\boldsymbol{q}|n,\alpha)} \\
&\leq \epsilon.
\end{aligned}
$$

where we used the definition of the total variation distance in the first and eighth lines, the fact that P_C^{rej} and P_{BS} are supported on samples with total photon number n in the second line, Eq. (136) in the third line, Eq. (132) in the fourth line and the triangle inequality in the fifth line.

Hence, $\tilde{C}$ is an efficient classical algorithm for approximate sampling from P_{BS}. □

We are left with identifying the regime of input cat state amplitude $\alpha \in \mathbb{C}$ such that the assumption in Lemma 2 is satisfied, i.e. $\sum_{|\boldsymbol{p}|=n} P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha) \geq \frac{1}{\mathrm{poly}\ m}$.

Lemma 3. *For $0 < |\alpha| = \mathcal{O}(n^{-1/4}\log^{1/4} m)$ and $m = poly\ n$, the detection events satisfying $|\boldsymbol{p}| = n$ represent at least an inverse-polynomial fraction of the outcomes for Boson Sampling with input cat states of amplitude α, i.e. $\sum_{|\boldsymbol{p}|=n} P_{cat}(\boldsymbol{p}|n,\alpha) \geq \frac{1}{poly\ m}$.*

Proof of Lemma 3. Let $\alpha \in \mathbb{C}$ such that $0 < |\alpha| = \mathcal{O}(n^{-1/4}\log^{1/4} m)$. The fraction of outcomes $\boldsymbol{p} \in \mathbb{N}^m$ for m-mode Boson Sampling with n input cat states of amplitude α is given by

$$
\begin{aligned}
\sum_{|\boldsymbol{p}|=n} P_{\mathrm{cat}}(\boldsymbol{p}|n,\alpha) &= \frac{|\alpha|^{2n}}{\sinh^n(|\alpha|^2)} \\
&= e^{n\log\left(|\alpha|^2\right) - n\log\left(\sinh|\alpha|^2\right)} \\
&= e^{n\log\left(|\alpha|^2\right) - n\log\left(|\alpha|^2 + \frac{1}{6}|\alpha|^6 + O(|\alpha|^{10})\right)} \\
&= e^{-n\log\left(1 + \frac{1}{6}|\alpha|^4 + O(|\alpha|^8)\right)} \\
&= e^{-\frac{1}{6}n|\alpha|^4 + O(n|\alpha|^8)} \\
&= e^{-O(\log m) + O(n^{-2}\log^2 m)} \\
&\geq \frac{1}{\mathrm{poly}\ m},
\end{aligned}
\tag{141}
$$

where we used Eq. (135) in the first line, the scaling of $|\alpha|$ in the sixth line, and $O(n^{-2}\log^2 m) = O(1)$ in the last line. □

Combining Lemma 2 and Lemma 3, using any efficient classical algorithm for approximate sampling from the output probability distribution of a Boson Sampling instance with input cat states with $0 < |\alpha| = \mathcal{O}(n^{-1/4}\log^{1/4} n)$, one may also efficiently sample approximately from the output probability distribution of the corresponding Boson Sampling instance with input single photons efficiently, by keeping only the samples $\boldsymbol{p}$ satisfying $|\boldsymbol{p}| = n$.

This proves that Boson Sampling with input cat states with $0 < |\alpha| = \mathcal{O}(n^{-1/4}\log^{1/4} n)$ is hard to sample approximately in the same regime as Boson Sampling with input Fock states, assuming the same conjectures as in [6]. □

Acknowledgments

U. C. thanks Atul Singh Arora and Pierre-Emmanuel Emeriau for interesting discussions. U. C, A. D, and S. M. acknowledge funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant PHY-1733907). A. D. also acknowledges support from the National Science Foundation RAISE-TAQS 1839204 and Amazon Web Services, AWS Quantum Program.

References

[1] H. Minc, "Permanents,", vol. 6. Cambridge University Press, 1984.

[2] J. K. Percus, "Combinatorial methods,", vol. 4. Springer Science & Business Media, 2012.

[3] L. G. Valiant, "The complexity of computing the permanent," *Theoretical computer science* **8**, 189–201 (1979).

[4] E. R. Caianiello, "On quantum field theory—I: explicit solution of Dyson's equation in electrodynamics without use of Feynman graphs," *Il Nuovo Cimento (1943-1954)* **10**, 1634–1652 (1953).

[5] S. Scheel, "Permanents in linear optical networks," quant-ph/0406127.

[6] S. Aaronson and A. Arkhipov, "The computational Complexity of Linear Optics," *Theory of Computing* **9**, 143 (2013), arXiv:1011.3245.

[7] S. Aaronson, "A linear-optical proof that the permanent is# P-hard," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **467**, 3393–3405 (2011).

[8] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, "What can quantum optics say about computational complexity theory?," *Physical review letters* **114**, 060501 (2015).

[9] D. Grier and L. Schaeffer, "New hardness results for the permanent using linear optics," arXiv:1610.04670.

[10] P. P. Rohde, D. W. Berry, K. R. Motes, and J. P. Dowling, "A Quantum Optics Argument for the #P-hardness of a Class of Multidimensional Integrals," arXiv:1607.04960.

[11] L. Chakhmakhchyan, N. J. Cerf, and R. Garcia-Patron, "Quantum-inspired algorithm for estimating the permanent of positive semidefinite matrices," *Physical Review A* **96**, 022329 (2017).

[12] A. Meiburg, "Inapproximability of Positive Semidefinite Permanents and Quantum State Tomography," arXiv:2111.03142.

[13] P. A. MacMahon, "Combinatory Analysis, Volumes I and II,", vol. 137. American Mathematical Soc., 2001.

[14] I. Good, "Proofs of some 'binomial'identities by means of MacMahon's 'Master Theorem'," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, pp. 161–162, Cambridge University Press. 1962.

[15] L. Carlitz, "An application of MacMahon's master theorem," *SIAM Journal on Applied Mathematics* **26**, 431–436 (1974).

[16] L. Carlitz, "Some expansions and convolution formulas related to MacMahon's master theorem," *SIAM Journal on Mathematical Analysis* **8**, 320–336 (1977).

[17] H. J. Ryser, "Combinatorial mathematics,", vol. 14. American Mathematical Soc., 1963.

[18] K. Balasubramanian, *Combinatorics and diagonals of matrices*. PhD thesis, Indian Statistical Institute-Kolkata, 1980.

[19] E. T. Bax, *Finite-difference algorithms for counting problems*. PhD thesis, California Institute of Technology, 1998.
[20] D. G. Glynn, "The permanent of a square matrix," *European Journal of Combinatorics* **31**, 1887–1891 (2010).
[21] P. P. Rohde, K. R. Motes, P. A. Knott, J. Fitzsimons, W. J. Munro, and J. P. Dowling, "Evidence for the conjecture that sampling generalized cat states with linear optics is hard," *Physical Review A* **91**, 012342 (2015).
[22] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Reviews of Modern Physics* **84**, 621 (2012).
[23] A. Leverrier, "$SU(p, q)$ coherent states and a Gaussian de Finetti theorem," *Journal of Mathematical Physics* **59**, 042202 (2018).
[24] T. Jiang and Y. Ma, "Distances between random orthogonal matrices and independent normals," arXiv:1704.05205.
[25] A. C. Dixon, "On the sum of the cubes of the coefficients in a certain expansion by the binomial theorem," *Messenger of mathematics* **20**, 79–80 (1891).
[26] I. Good, "A short proof of MacMahon's 'Master Theorem'," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, pp. 160–160, Cambridge University Press. 1962.
[27] S. Garoufalidis, T. T. Lê, and D. Zeilberger, "The quantum MacMahon master theorem," *Proceedings of the National Academy of Sciences* **103**, 13928–13931 (2006).
[28] M. Konvalinka and I. Pak, "Non-commutative extensions of the MacMahon Master Theorem," *Advances in Mathematics* **216**, 29–61 (2007).
[29] M. P. Tuite, "Some generalizations of the MacMahon Master Theorem," *Journal of Combinatorial Theory, Series A* **120**, 92–101 (2013).
[30] V. V. Kocharovsky, V. V. Kocharovsky, and S. V. Tarasov, "The Hafnian Master Theorem," *Linear Algebra and its Applications* 144–161 (2022).
[31] W. Y. Chen, H. Galbraith, and J. Louck, "Angular momentum theory, umbral calculus, and combinatorics," *Computers & Mathematics with Applications* **41**, 1199–1214 (2001).
[32] B. M. Terhal and D. P. DiVincenzo, "Classical simulation of noninteracting-fermion quantum circuits," *Physical Review A* **65**, 032325 (2002).
[33] V. Shchesnovich, "Partial indistinguishability theory for multiphoton experiments in multiport devices," *Physical Review A* **91**, 013844 (2015).
[34] D. Spivak, M. Y. Niu, B. C. Sanders, and H. de Guise, "Generalized interference of fermions and bosons," *Physical Review Research* **4**, 023013 (2022).
[35] E.-J. Kuo, Y. Xu, D. Hangleiter, A. Grankin, and M. Hafezi, "Boson Sampling for Generalized Bosons," arXiv:2204.08389.
[36] A. Clément, N. Heurtel, S. Mansfield, S. Perdrix, and B. Valiron, "$\mathrm{LO_v}$-Calculus: A Graphical Language for Linear Optical Quantum Circuits," arXiv:2204.11787.
[37] G. De Felice and B. Coecke, "Quantum Linear Optics via String Diagrams," arXiv:2204.12985.
[38] B. Peropadre, G. G. Guerreschi, J. Huh, and A. Aspuru-Guzik, "Proposal for microwave boson sampling," *Physical review letters* **117**, 140505 (2016).
[39] S. Girvin, "Schrödinger cat states in circuit qed," arXiv:1710.03179.
[40] X. Gu, A. F. Kockum, A. Miranowicz, Y.-x. Liu, and F. Nori, "Microwave photonics with superconducting quantum circuits," *Physics Reports* **718**, 1–102 (2017).
[41] J. Huh, "A fast quantum algorithm for computing matrix permanent," arXiv:2205.01328.

[42] S. Aaronson and T. Hance, "Generalizing and Derandomizing Gurvits's Approximation Algorithm for the Permanent," *Quantum Info. Comput.* **14**, 541–559 (2014).
[43] S. Chin and J. Huh, "Generalized concurrence in boson sampling," *Scientific reports* **8**, 1–9 (2018).
[44] M.-H. Yung, X. Gao, and J. Huh, "Universal bound on sampling bosons in linear optics and its computational implications," *National science review* **6**, 719–729 (2019).
[45] V. S. Shchesnovich, "On the classical complexity of sampling from quantum interference of indistinguishable bosons," *International Journal of Quantum Information* **18**, 2050044 (2020).
[46] D. M. Jackson, "The unification of certain enumeration problems for sequences," *Journal of Combinatorial Theory, Series A* **22**, 92–96 (1977).
[47] F. R. Cardoso, D. Z. Rossatto, G. P. Fernandes, G. Higgins, and C. J. Villas-Boas, "Superposition of two-mode squeezed states for quantum information processing and quantum sensing," *Physical Review A* **103**, 062405 (2021).
[48] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, "Boson sampling from a Gaussian state," *Physical review letters* **113**, 100502 (2014).
[49] J. P. Olson, K. P. Seshadreesan, K. R. Motes, P. P. Rohde, and J. P. Dowling, "Sampling arbitrary photon-added or photon-subtracted squeezed states is in the same complexity class as boson sampling," *Physical Review A* **91**, 022317 (2015).
[50] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, "Gaussian boson sampling," *Physical review letters* **119**, 170501 (2017).
[51] A. Lund, S. Rahimi-Keshari, and T. Ralph, "Exact boson sampling using Gaussian continuous-variable measurements," *Physical Review A* **96**, 022301 (2017).
[52] L. Chakhmakhchyan and N. J. Cerf, "Boson sampling with Gaussian measurements," *Physical Review A* **96**, 032326 (2017).
[53] U. Chabaud, T. Douce, D. Markham, P. van Loock, E. Kashefi, and G. Ferrini, "Continuous-variable sampling from photon-added or photon-subtracted squeezed states," *Physical Review A* **96**, 062307 (2017).
[54] N. Quesada, J. M. Arrazola, and N. Killoran, "Gaussian boson sampling using threshold detectors," *Physical Review A* **98**, 062322 (2018).
[55] A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, J. Eisert, *et al.*, "Quantum computational advantage via high-dimensional Gaussian boson sampling," *Science advances* **8**, 7894 (2022).