

Heisenberg picture tensor network formalism for optical circuits

Dario Cilluffo^{✉,*}, Matthias Kost[✉], Nicola Lorenzoni[✉], and Martin B. Plenio[†]

*Institute of Theoretical Physics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany
and Center for Integrated Quantum Science and Technology (IQST), 89081 Ulm, Germany*



(Received 13 February 2025; revised 3 November 2025; accepted 27 March 2026; published 30 April 2026)

Tensor network formalisms have emerged as powerful tools for simulating quantum state evolution. While widely applied in the study of optical quantum circuits, such as boson sampling, existing tensor network approaches fail to address the complexity mismatch between tensor contractions and the calculation of photon-counting probability amplitudes. Here, we present an alternative tensor network framework that exploits the input-output relations of quantum optical circuits encoded in the unitary interferometer matrix. Our approach bridges the complexity gap by enabling the computation of the permanent—central to boson sampling—with the same computational complexity as the best known classical algorithm based on a graphical representation of the operator-basis matrix product states that we introduce. Furthermore, we exploit the flexibility of tensor networks to extend our formalism to incorporate partial distinguishability and photon loss, two key imperfections in practical interferometry experiments. This work offers a significant step forward in the simulation of large-scale quantum optical systems and the understanding of their computational complexity.

DOI: [10.1103/wtwl-979d](https://doi.org/10.1103/wtwl-979d)

The tensor network formalism for the time evolution of quantum states [1–3] offers numerical speedup primarily due to two key factors: the efficient representation of quantum states and operators with bounded correlations [4,5] through a modular structure and the ability to approximate general quantum states with controlled, quantifiable error. These features lead to significant savings in computational resources [6,7]. They are also the main reason tensor networks have found widespread application in the description of optical quantum circuits, particularly in the context of boson sampling [8–10]. In such applications, the evolution of an optical input state through an interferometer is modeled as the propagation of a bosonic matrix product states (MPS) through a network of matrix product operators (MPOs) representing the phase shifters and beam splitters that define the interferometer. However, when applying these methods to Fock states propagating through such an interferometer—such as in reproducing ideal boson sampling experiments—the computational scaling of tensor network algorithms and the intrinsic complexity of photon-counting tasks scale very differently. In analogy with the fermionic case [11,12], it has been shown [13] that for any defectless passive optical network a product of n bosonic ladder operators evolves as an MPO with a maximum bond dimension of $D = 2^n$, independent of the number of input modes M . However, the compression and manipulation of MPOs involve singular value decompositions with a computa-

tional cost scaling as $\mathcal{O}(D^3)$ [14]. In contrast, photon detection amplitudes can be computed by evaluating the permanent of a special $n \times n$ matrix, which has a complexity of $\mathcal{O}(n^2 2^n)$ [15]. This discrepancy highlights a limitation in the current use of tensor networks for large-scale ideal or nearly ideal systems, despite their success in simulating setups with noise or imperfections [16,17]. In this work, we formulate an alternative tensor-network-based approach, the operator-basis MPS, to bridge this complexity gap. In contrast to the standard approach, where the evolution is represented by MPOs generated by the Hamiltonian in Fock space and subsequently applied to the initial state, our formalism exploits the known input-output relations of the circuit. For an M -mode interferometer, these relations are encoded in the matrix $\mathcal{U} = \{u_k^i\} \in \text{SU}(M)$, with $i, k \in [1, M]$ labeling the input and output ports, respectively. This allows us to directly construct the final state without explicitly performing time propagation, thereby avoiding the primary computational cost of that procedure, which is largely associated with compression. To demonstrate the effectiveness of our formalism, we show that our tensor network description of quantum optical circuits can efficiently compute photodetection amplitudes, which correspond to the permanent of an $n \times n$ matrix,

$$\text{Perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^n A_{i, \sigma(i)}, \quad (1)$$

where σ ranges over all permutations in the symmetric group S_n . Our approach achieves the optimal classical computational complexity, matching that of Ryser's algorithm—the best known method for evaluating matrix permanents. Then, we introduce an alternative graphical representation for our MPS, which not only simplifies the overall language but also provides valuable insights, particularly in understanding the scaling of complexity. This approach also suggests a poten-

*Contact author: dario.cilluffo@uni-ulm.de

†Contact author: martin.plenio@uni-ulm.de

tial strategy for implementation in terms of binary matrices. Finally, we demonstrate that typical imperfections in optical circuits such as partial distinguishability and photon losses can also be treated within our formalism.

I. OPERATOR-BASIS MPS REPRESENTATION OF EVOLVED BOSONIC LADDER OPERATORS

A passive linear optical interferometer represented by the unitary $\mathcal{U} = \{u_k^i\}$ acting on M bosonic modes turns the single ladder operator $\hat{a}_i^\dagger$ acting on input mode i into a linear combination $\sum_{k=1}^M u_k^i \hat{a}_k^\dagger$, which, in terms of augmented vectors (see Appendix A), can be expressed as

$$\mathcal{U}(\hat{a}_i^\dagger) = (u_1^i \hat{a}_1^\dagger \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ u_2^i \hat{a}_2^\dagger & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ u_{M-1}^i \hat{a}_{M-1}^\dagger & \mathbb{1} \end{pmatrix} \times \begin{pmatrix} \mathbb{1} \\ u_M^i \hat{a}_M^\dagger \end{pmatrix}, \quad (2)$$

where, for convenience, we use the same symbol for the channel and its $SU(M)$ representation. Unless otherwise stated, we assume that our evolved operators act on the vacuum state. Considering the local reduced operator basis $\{\hat{a}_k^{n\dagger}\}_{n=0,1}$, the expression above readily maps to the MPO

$$\mathcal{U}(\hat{a}_i^\dagger) = \sum_{\sigma_1, \dots, \sigma_M} A_{i\sigma_1}^{[1]} A_{i\sigma_2}^{[2]} \cdots A_{i\sigma_M}^{[M]} \bigotimes_{k=1}^M (\hat{a}_k^\dagger)^{\sigma_k}, \quad (3)$$

where the tensors $A_i^{[m]\sigma_j}$, with $\sigma_j = 0, 1$, are defined as

$$\begin{aligned} A_{i0}^{[m=1]} &= \begin{pmatrix} 0 & 1 \end{pmatrix}, & A_{i1}^{[m=1]} &= \begin{pmatrix} u_1^i & 0 \end{pmatrix}, \\ A_{i0}^{[m \neq 1, M]} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & A_{i1}^{[m \neq 1, M]} &= \begin{pmatrix} 0 & 0 \\ u_m^i & 0 \end{pmatrix}, \\ A_{i0}^{[m=M]} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & A_{i1}^{[m=M]} &= \begin{pmatrix} 0 \\ u_M^i \end{pmatrix}, \end{aligned} \quad (4)$$

where $1 \leq m \leq M$ and generally $\sigma_1, \dots, \sigma_M \in \{0, 1\}$. However, we note that for $m \neq 1, M$, the $A_{i1}^{[m]}$ -matrices are nilpotent with index 2. As a consequence, the sum in the equation above can be restricted to terms satisfying $\sum_{k=1}^M \sigma_k = 1$.

The bond dimension of the MPO in Eq. (3) is upper bounded by $D = 2$ for any passive linear optics network, which is consistent with the result in Refs. [11–13] for local fermionic operators. When considering a product of n ladder operators $\hat{a}_i^\dagger \equiv \hat{a}_{i_1}^\dagger \hat{a}_{i_2}^\dagger \cdots \hat{a}_{i_n}^\dagger$, evolving as the tensor product of each $\mathcal{U}(\hat{a}_{i_j}^\dagger)$, Eq. (3) generalizes to

$$\mathcal{U}(\hat{a}_i^\dagger) = \sum_{\sum_{m=1}^n n_m = n} \mathcal{A}_i^{[1]n_1} \mathcal{A}_i^{[2]n_2} \cdots \mathcal{A}_i^{[M]n_M} \bigotimes_{k=1}^M (\hat{a}_k^\dagger)^{n_k}, \quad (5)$$

where

$$\sqrt{n_m!} \mathcal{A}_i^{[m]n_m} = \sum_{\sum_{k=1}^n \sigma_k^{(m)} = n_m} A_{i1}^{[m]\sigma_1^{(m)}} \otimes \cdots \otimes A_{i_n}^{[m]\sigma_n^{(m)}}. \quad (6)$$

Probability amplitudes for specific Fock space components can be retrieved by computing the products of the $\mathcal{A}$ -matrices in Eq. (5). We conclude this part by noticing that, while these

objects are formally MPOs, in operator space their tensor structure along the chain of modes, with one physical index per local operator, closely resembles an MPS. Importantly, these operator-basis MPS (OBMPS) are always meant to be applied to the vacuum; in this sense, they encode all the information about the corresponding Fock-state amplitudes. In this structural sense, we define them as operator-basis MPS.

II. GRAPHICAL INTERPRETATION

The expression in Eq. (6) is well suited for a graph-based interpretation, which provides a more straightforward approach to the calculations. We define

$$\square := A_{i0}^{[m]} = \mathbb{1}_2 \quad \blacksquare := m \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (7)$$

where m is a complex number. We also define the two operations

$$\blacksquare \blacksquare := X_i^{[m]} \otimes X_j^{[m]} \quad \blacksquare \square := X_i^{[m]} \times X_i^{[m+1]} \quad (8)$$

where the generic tensor $\blacksquare$ can correspond to either a green or a white square. Note that the nilpotence of the A -matrices translates as $\blacksquare \blacksquare = 0$. Equation (6) translates into

$$\mathcal{A}_i^{[j]n_j} = [\underbrace{\blacksquare \cdots \blacksquare}_{n_j} \underbrace{\square \cdots \square}_{n-n_j}], \quad (9)$$

where the square brackets denote the sum over all $\binom{n}{n_j}$ possible placements of n_j green squares ($\blacksquare$) and $n - n_j$ white squares ($\square$). We will refer to this type of graph and to each arrangement as *lines* and *components*, respectively. Thus, from Eq. (5), the amplitude $\alpha_{\vec{n}}$ corresponding to a component $\hat{a}_j^\dagger |0\rangle = |\vec{n}\rangle$ of the final state is turned into a grid of M lines stacked vertically:

$$\alpha_{\vec{n}} = \langle 0 | \hat{a}_{\vec{j}} \mathcal{U}(\hat{a}_{\vec{i}}^\dagger) | 0 \rangle = \begin{array}{c} [\blacksquare \cdots \blacksquare \square \cdots \square \square \cdots \square \square \cdots \square] \\ [\square \cdots \square \blacksquare \cdots \blacksquare \square \cdots \square \square \cdots \square] \\ \vdots \\ [\square \cdots \square \square \cdots \square \square \cdots \blacksquare \cdots \blacksquare] \end{array} \quad n = |\vec{n}| \quad (10)$$

We will refer to the contraction of two lines, i.e., to the product of two $\mathcal{A}$ -matrices, as *merging*. In terms of graphs,

$$\mathcal{A}_i^{[j]n_j} \mathcal{A}_i^{[j']n_{j'}} = [\underbrace{\blacksquare \cdots \blacksquare}_{n_j + n_{j'}} \underbrace{\square \cdots \square}_{n - n_j - n_{j'}}]. \quad (11)$$

The elementary properties of these operations along with proofs and examples will be presented in detail in Appendix A.

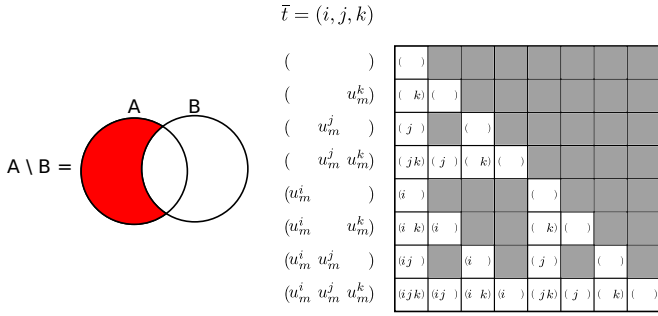


FIG. 1. Left side: The relative complement or set difference between the sets A and B is defined as $A \setminus B = \{x \in A : x \notin B\}$. Right side: Example of generation of the Sierpiński triangle structure for the tensor in Eq. (12). The subsets of $(u_m^i u_m^j u_m^k)$ are arranged in lexicographic order, $()$ represents the empty set, and gray squares correspond to the cases in which $A \cap B = ()$ (see Appendix B for details).

III. INVERSE CHANGE OF BASIS

While we focus on computing probability amplitudes, the formalism is not limited to this. The operator-based construction can in principle be used for general-purpose state manipulation, beyond computing expectation values. For this reason, we complete the set of basic rules of the formalism by defining the transformation of bosonic operator-basis MPSs into standard MPSs. We leverage two key insights: The Kronecker product of multiple A -matrices results in a sparse matrix and the positions of the elements of $\mathcal{A}$ corresponding to different powers of $\hat{a}^\dagger$ follow regular patterns. Regarding the latter, as demonstrated, e.g., in Ref. [18], the multiple tensor products of 2×2 triangular matrices generate a $2^n \times 2^n$ matrix whose elements are distributed according to a fractal structure known as the *Sierpiński triangle*. In our case, we observe the same phenomenon for the repeated tensor products of the elements in our decomposition. This is illustrated below for a three-photon tensor:

$$\begin{pmatrix} 1 & 0 \\ u_m^i \hat{a}^\dagger & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ u_m^j \hat{a}^\dagger & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ u_m^k \hat{a}^\dagger & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_m^k \hat{a}^\dagger & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_m^j \hat{a}^\dagger & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ u_m^{jk} \hat{a}^{\dagger 2} & u_m^j \hat{a}^\dagger & u_m^k \hat{a}^\dagger & 1 & 0 & 0 & 0 & 0 \\ u_m^i \hat{a}^\dagger & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ u_m^{ik} \hat{a}^{\dagger 2} & u_m^i \hat{a}^\dagger & 0 & 0 & u_m^k \hat{a}^\dagger & 1 & 0 & 0 \\ u_m^{ij} \hat{a}^{\dagger 2} & 0 & u_m^i \hat{a}^\dagger & 0 & u_m^j \hat{a}^\dagger & 0 & 1 & 0 \\ u_m^{ijk} \hat{a}^{\dagger 3} & u_m^{ij} \hat{a}^{\dagger 2} & u_m^{ik} \hat{a}^{\dagger 2} & u_m^i \hat{a}^\dagger & u_m^{jk} \hat{a}^{\dagger 2} & u_m^j \hat{a}^\dagger & u_m^k \hat{a}^\dagger & 1 \end{pmatrix}, \quad (12)$$

where we used the shorthand notation $u_m^{ijk} := u_m^i u_m^j u_m^k$. Tensors like the one above can be turned into a standard Fock-basis tensor network by replacing each component with the corresponding $d \times d$ matrix, being d the local dimension cutoff. However, thanks to the regular and sparse structure of the Sierpiński triangle, each and every component can be computed independently of all others. This enables the encoding of components with a specific photon number $n_j < n$ of the single-mode tensor $\mathcal{A}_i^{[j]n_j}$ without generating and storing all possible components. Furthermore, the logical operations required to produce the fractal structure can be executed efficiently and in parallel using binary strings. The Sierpiński triangle corresponding to the matrices $\mathcal{A}_i^{[m]n}$ for a given number of photons n in the target interferometer can

be generated recursively from the set of the lexicographically ordered subsets $[(), u_m^i \hat{a}_m^\dagger, u_m^{i2} \hat{a}_m^{\dagger 2}, u_m^{i1} u_m^{i2} (\hat{a}_m^\dagger)^2, \dots, u_m^i (\hat{a}_m^\dagger)^n]$ with nonempty intersection, as depicted in Fig. 1 for a three-photon scenario (see Appendix B for details). Therefore, we are able to generate exclusively the required components once n_j is fixed, offering significant advantages in memory usage and computational efficiency.

Composition of channels. We consider the representation of the composition of unitary channels $\mathcal{U}$ and $\mathcal{V}$, denoted by $\mathcal{V} \circ \mathcal{U}$, within the bosonic-operator-basis-MPS formalism. We will restrict to single-photon states, as the many-photon case is obtained through tensor products. In this scenario, the input-output matrix that represents the evolution is $\mathcal{U} \cdot \mathcal{V}$. The state of a single photon at the input port i evolved through the composition of unitary channels $\mathcal{V} \circ \mathcal{U}$ reads

$$\begin{aligned} \mathcal{V} \circ \mathcal{U}(\hat{a}_i^\dagger) &= \left(u_1^i \sum_{j=1}^M v_j^1 \hat{a}_j^\dagger \quad 1 \right) \prod_{k=2}^{M-1} \begin{pmatrix} 1 & 0 \\ u_k^i \sum_{j=1}^M v_j^k \hat{a}_j^\dagger & 1 \end{pmatrix} \begin{pmatrix} 1 \\ u_M^i \sum_{j=1}^M v_j^M \hat{a}_j^\dagger \end{pmatrix} \\ &= \left(\sum_{j=1}^M u_j^i v_j^1 \hat{a}_1^\dagger \quad 1 \right) \prod_{k=2}^{M-1} \begin{pmatrix} 1 & 0 \\ \sum_{j=1}^M u_j^i v_j^k \hat{a}_k^\dagger & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \sum_{j=1}^M u_j^i v_j^M \hat{a}_M^\dagger \end{pmatrix}, \end{aligned} \quad (13)$$

where $\mathcal{V} = \{v_j^i\}$ and we used the commutative property (see Appendix A). As expected, the coefficients in the above equation are the sums of the row elements from the input-output matrix of the combined transformation. Notably, for each product of coefficients, the first upper index on the left, i , denotes the input port, while the final lower index on the right represents the output port through which the photon exits the interferometer. Therefore, as a consistency check, if $\mathcal{U} = \mathcal{V}^{-1}$, all coefficients will vanish except for $\sum_{j=1}^n u_j^i v_j^i$, and the state simplifies to the single ladder operator at i . Equation (13) can be readily generalized to scenarios with multiple channels using the following lemma, which directly follows from the commutative property (see Appendix A for details):

Lemma 1. Let $u_j^{(k)i}$ be the elements of a generic k th transformation $\mathcal{U}^{(k)}$ acting on the ladder operator $\hat{a}_i^\dagger$ and $T_{i_2, i_3, \dots, i_m}^{i_1, i_2, \dots, i_n} = u_{i_2}^{(1)i_1} u_{i_3}^{(2)i_2} \dots u_m^{(n)i_n} \hat{a}_m^\dagger$, with $m \in [1, M]$, the non-identity element of the m th matrix in the expansion of $\bigcirc_{k=1}^n \mathcal{U}^{(k)}$. Then, $\forall m, \mathcal{U}^{(n+1)} \circ \bigcirc_{k=1}^n \mathcal{U}^{(k)}$ corresponds to the following transformation on the nonidentity elements:

$$T_{i_2, i_3, \dots, i_m}^{i_1, i_2, i_3, \dots, i_n} \rightarrow T_{i_2, i_3, \dots, i_m, i_{n+1}}^{i_1, i_2, i_3, \dots, i_n, i_{n+1}}. \quad (14)$$

This concludes the basic rules of our formalism. We now move on to benchmark it by calculating the probability amplitude for a specific photon-counting outcome, starting from a given input state. The square modulus of this amplitude corresponds to the coincidence rate in a multiphoton interference experiment.

IV. MULTIPHOTON DETECTION PROBABILITY AMPLITUDES

For the sake of argument, we consider the case $n_i = 1$ for $i \leq n$ and 0 otherwise [19]. This is known to correspond to the permanent of an $n \times n$ complex matrix [8]. The diagrammatic representation of the amplitude $\alpha_{\vec{n}}$ reads

$$\alpha_{\vec{n}} = \left\{ \begin{array}{c} \begin{array}{cccc} \blacksquare & \square & \square & \dots & \square \\ \square & \blacksquare & \square & \dots & \square \\ & & \dots & & \\ \square & \square & \square & \dots & \blacksquare \end{array} \end{array} \right\} n. \quad (15)$$

There is not a unique contraction strategy. The direct computation of all scalar products followed by tensor products results in the same computational scaling as using the definition of permanents directly (see Appendix D). We show here how we can obtain a computational advantage over the standard tensor network approach by working with the diagrams and the given rules. First, we analyze the computational cost of merging the first two lines in Eq. (15):

$$\left\{ \begin{array}{cccc} \blacksquare & \square & \square & \dots & \square \\ \square & \blacksquare & \square & \dots & \square \end{array} \right\} = \left\{ \begin{array}{cccc} \blacksquare & \blacksquare & \square & \dots & \square \end{array} \right\}. \quad (16)$$

Due to the nilpotence of the $A^{[m]}$ -matrices, each of the n component of the first line can be merged with $n - 1$ components of the second one. In Appendix A, we show that the number of operations necessary to perform these operation is $\mathcal{O}(n)$. Thus, the resulting computational cost of merging the first two lines is $n \binom{n}{1} \binom{n-1}{1}$. The same arguments hold for the next

contraction,

$$\left\{ \begin{array}{cccc} \blacksquare & \blacksquare & \square & \dots & \square \\ \square & \square & \blacksquare & \dots & \square \end{array} \right\} = \left\{ \begin{array}{cccc} \blacksquare & \blacksquare & \blacksquare & \dots & \square \end{array} \right\} \quad (17)$$

with the difference that this time the nilpotence of the matrices implies that for a fixed component in the first line we have $\binom{n}{2}$ different components in the subsequent line with nonzero contribution. This gives $n \binom{n}{2} \binom{n-2}{1}$ operations. By proceeding in the same way for the rest of the operations, we find that the computational cost c is given by

$$c = n \cdot \sum_{i=1}^{n-1} \binom{n}{i} \binom{n-i}{1} = n^2 (2^{n-1} - 1) = \mathcal{O}(n^2 2^n). \quad (18)$$

This matches the computational complexity of computing the permanent of an $n \times n$ matrix using Ryser's algorithm. Thus, the bosonic-operator-basis MPS representation provides a computationally optimal description of linear interferometers followed by photon-counting measurements. Furthermore, relaxing the condition $n_i = 1$ reduces the number of merging operations required to compute the amplitude, consistent with the reduction in complexity observed for boson sampling when multiple photons occupy the same input mode [20].

V. PHOTONIC CIRCUITS WITH IMPERFECTIONS

Partial distinguishability of photons introduces important distortions in quantum interference effects and constitutes one of the main sources of imperfection in boson-sampling experiments. In order to introduce this feature, we consider now an internal and, for the sake of clarity, dichotomic degree of freedom for each photon, namely,

$$\hat{a}^\dagger \rightarrow \sqrt{\eta} \hat{a}^\dagger + \sqrt{1-\eta} \hat{a}^{\perp\dagger}, \quad (19)$$

where $\eta \in [0, 1]$ encodes the partial distinguishability. The components generated by the two ladder operators are mutually orthogonal and do not interact within the interferometer. Consequently, the formalism described earlier can be easily extended by using direct sums of orthogonal subspaces. Thus, based on Eq. (4), we introduce the operator basis for distinguishable photons,

$$B_{i0}^{[m]} = A_{i0}^{[m]} \oplus A_{i0}^{[m]}, \quad (20)$$

$$B_{i1}^{[m]} = \sqrt{\eta_i} A_{i1}^{[m]} \oplus \sqrt{1-\eta_i} A_{i1}^{[m]}, \quad (21)$$

where the distinguishability coefficient refers to the photon entering through the i th port. The interferometer will act independently on both the sides of the direct sum. The distributivity of the matrix multiplication over the direct sum ensures that the bond dimension does not change. In the limiting case of many photons being perfectly distinguishable (e.g., each with a unique frequency), the direct sum defined above introduces terms from additional subspaces, effectively summing over contributions from these subspaces. Consequently, the computation of a multiphoton amplitude simplifies to a sum of single-photon contributions, significantly reducing computational complexity. In a less extreme scenario where groups of photons are partially distinguishable, this leads to the sum of smaller permanents, which is consistent with the findings in Ref. [21]. Analogously, photon losses can

be integrated with minimal modifications by considering the enlarged matrix $\tilde{U}$, which includes additional sink modes representing the environment. In the scenario of uniform losses, the linear transformation is $(\sqrt{1-\epsilon} \mathcal{U} \sqrt{\epsilon} \mathbb{1}_M)$. For each input photon, any loss process corresponds to the creation of an excitation in one of these extra modes. As a consequence, Eq. (2) is modified as follows:

$$\tilde{U}(\hat{a}_i^\dagger) = (\tilde{u}_1^i \hat{a}_1^\dagger \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \tilde{u}_2^i \hat{a}_2^\dagger & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} \\ \sqrt{\epsilon} \hat{a}_{M+i}^\dagger \end{pmatrix}, \quad (22)$$

where $\tilde{u}_j^i = \sqrt{1-\epsilon} u_j^i$. The formalism outlined in the previous section can now be applied directly without significant modification. The probability amplitude $\alpha_{\vec{n}}$, given the number of input photons $n \geq n_{\text{out}} = \sum_{i=1}^M n_i$, reads

$$(\mathcal{A}_i^{[1]n_1} \mathcal{A}_i^{[2]n_2} \cdots \mathcal{A}_i^{[M]n_M}) (\mathcal{A}_i^{[M+1]n_{M+1}} \cdots \mathcal{A}_i^{[2M]n_{2M}}), \quad (23)$$

with $\mathcal{A}_i^{[M+k]n_{M+k}}$ expressed as in Eq. (6) with the A -matrices like in Eq. (4) but with all coefficients u replaced by $\sqrt{\epsilon}$. Since the coefficients in the lossy sector are uniform, any contraction involving a fixed number of lost photons contributes the same overall factor, independent of how the losses are distributed among the sink modes. This common prefactor can therefore be computed separately, and in the graphical representation the entire loss contribution reduces to a single effective line that accounts for all lost photons:

$$n_{\text{out}} \left\{ \begin{array}{c} \begin{array}{ccccccc} \blacksquare & \blacksquare & \cdots & \blacksquare & \blacksquare & \cdots & \blacksquare \\ \blacksquare & \blacksquare & \cdots & \blacksquare & \blacksquare & \cdots & \blacksquare \end{array} \\ \vdots \\ \begin{array}{ccccccc} \blacksquare & \blacksquare & \cdots & \blacksquare & \blacksquare & \cdots & \blacksquare \\ \blacksquare & \blacksquare & \cdots & \blacksquare & \blacksquare & \cdots & \blacksquare \end{array} \end{array} \right\} \quad (24)$$

$\underbrace{\hspace{10em}}_{n_\epsilon}$

where $n_\epsilon = \sum_{i=1}^{2M} n_i$. Note that each component resulting from the merging of the first n_{out} lines determines a unique component in the last line. As a consequence, computing the probabilities conditioned to a specific amount n_ϵ of photons lost reduces to the same procedure previously described, with a number of arithmetic operations scaling as $c_L = n^2 \sum_{i=1}^{n-n_{M+1}} \binom{n-1}{i} < n^2 2^n$. In Fig. 2, we report the ratio c_L/c as a function of the number of photon lost. A distinct threshold emerges at 3 dB of losses, which aligns with the best state-of-the-art experiments [10]. We conclude this section by addressing dephasing. In the pure-state framework adopted here, dephasing can be modeled by introducing random phase shifts in the matrix representation of the channel. In photon-counting experiments, this results in expectation values that depend on these phases, which can be averaged *a posteriori*. This process requires an external routine, such as Monte Carlo integration, which is fully parallelizable as long as the phase shifts are independent. However, handling arbitrary defects would be more effectively approached using a density-matrix formalism, which we leave for future exploration.

VI. DISCUSSION

We leverage the known input-output relations of linear optical circuit components to directly construct, through an MPS-like representation, the exact operator expression

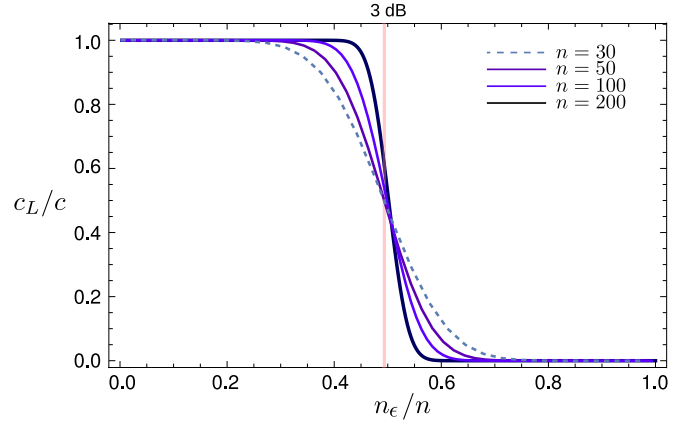


FIG. 2. Ratio between the computational cost of the permanent c [Eq. (18)] and c_L as a function of the fraction of photon lost, n_ϵ/n , for different numbers of input photons n . The drop in complexity corresponds to 50% $\sim$ 3 dB losses.

defining the output of a linear optical network, thereby avoiding explicit time propagation. The graphical representation of our tensors simplifies the identification of optimal contraction strategies and allows for a precise evaluation of the computational scaling involved in calculating probability amplitudes and expectation values of observables expressed as products of bosonic creation and annihilation operators. Specifically, in the context of boson sampling experiments, this approach allows to compute permanents of complex matrices with the same scaling as the best known algorithms. In contrast, until now, the use of tensor networks has been limited to small, idealized setups or larger interferometers plagued by a high degree of defects, which degrade correlations and undermine the claim of quantum advantage [16,17]. An operator-basis MPS can also be efficiently converted into a standard MPO and applied to a state in the Fock basis with minimal overhead. Notably, translating a bosonic operator-basis MPS into a standard MPS via Eq. (2) involves $\mathcal{O}(2^{2n})$ operations, identifying the reason for the complexity gap between our representation and conventional tensor network methods. Moreover, our approach suggests a potential advantage by leveraging the fractal structure of the $\mathcal{A}$ -matrices, offering a more efficient alternative to the standard translation. We have shown how the proposed formalism can be used to describe important sources of defects occurring in quantum optical experiments, namely, partial distinguishability, photon losses, in particular uniform losses, and dephasing. In particular, we can establish thresholds on the level of imperfections an experiment can tolerate while preserving the quantum advantage regime. Provided that the effects of such defects on the dynamical map governing the system's evolution are accurately characterized and represented in the input-output relations of the channel, our approach is well suited to incorporate them. Finally, employing this formalism within the Heisenberg picture offers substantial advantages for simulating observables such as photon-counting moments. This is especially valuable when dealing with highly complex input states (e.g., those with high photon numbers, strong correlations, or non-Gaussian features) or when a transition to the density-matrix formalism becomes necessary. Moreover, the same framework can be

extended to describe coherent states (displaced vacuum) and squeezed states on specific modes by incorporating displacement and squeezing operators as transformations of the local A -matrices.

ACKNOWLEDGMENTS

We acknowledge T. Lacroix, G. Wójtowicz, and G. Magnifico for fruitful discussions. This research was supported by the BMBF project PhoQuant (Grant No. 13N16110).

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: AUGMENTED VECTORS ALGEBRA AND COMBINATION RULES FOR GRAPHS

In this Appendix, we outline the fundamental properties of augmented vectors and establish the combination rules for A -matrices and their associated graphs.

The matrix representation in Eq. (2) corresponds to mapping a sum of operators onto an equivalent product of augmented vectors according to the following:

Definition 1. Vectorized sum:

$$\sum_{i=1}^n \mathbb{x}_i = (\mathbb{x}_1 \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \mathbb{x}_2 & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ \mathbb{x}_{n-1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} \\ \mathbb{x}_n \end{pmatrix}, \quad (\text{A1})$$

where $\mathbb{x}_i$ are elements of a generic vector space over $\mathbb{C}$ and $\mathbb{1}$ is the neutral element of multiplication.

The distributive property of multiplication and commutative property of the sum translate straightforwardly:

Property 1. Distributive property of the product by a scalar:

$$\begin{aligned} \lambda(\mathbb{x}_1 \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \mathbb{x}_2 & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ \mathbb{x}_{n-1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} \\ \mathbb{x}_n \end{pmatrix} \\ = (\lambda\mathbb{x}_1 \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \lambda\mathbb{x}_2 & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ \lambda\mathbb{x}_{n-1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} \\ \lambda\mathbb{x}_n \end{pmatrix}. \end{aligned} \quad (\text{A2})$$

Property 2. Commutative property of the sum:

$$\begin{aligned} (\sum_{i=1}^n \mathbb{x}_i^{(1)} \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \sum_{i=1}^n \mathbb{x}_i^{(2)} & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ \sum_{i=1}^n \mathbb{x}_i^{(M)} & \mathbb{1} \end{pmatrix} \\ = (\sum_{i=1}^M \mathbb{x}_i^{(i)} \quad \mathbb{1}) \begin{pmatrix} \mathbb{1} & 0 \\ \sum_{i=1}^M \mathbb{x}_2^{(i)} & \mathbb{1} \end{pmatrix} \cdots \begin{pmatrix} \mathbb{1} & 0 \\ \sum_{i=1}^M \mathbb{x}_n^{(i)} & \mathbb{1} \end{pmatrix}. \end{aligned} \quad (\text{A3})$$

Property 3. All the components of two lines with at least one green square in the same position return 0 in merging, i.e.,

$$\begin{array}{c} \blacksquare \blacksquare \cdots \blacksquare \\ \blacksquare \blacksquare \cdots \blacksquare \end{array} = 0 \quad (\text{A4})$$

Proof. The distributive property of the matrix product over the tensor product implies that

$$\mathcal{A}_i^{[j]n_j} \mathcal{A}_i^{[j']n_{j'}} = \sum_{\sum_{k=1}^M \sigma_k^{(j)} = n_j} A_{i_1 \sigma_1^{(j)}}^{[j]} A_{i_1 \sigma_1^{(j')}}^{[j']} \otimes \cdots \otimes A_{i_n \sigma_n^{(j)}}^{[j]} A_{i_n \sigma_n^{(j')}}^{[j']}. \quad (\text{A5})$$

Due to the nilpotence of the A -matrices, $A_{i_m \sigma_m^{(j)}}^{[j]} \times A_{i_m \sigma_m^{(j')}}^{[j']}$ returns a null vector if $\sigma_m^{(j)} = \sigma_m^{(j')} = 1$, which corresponds to a graph with two green squares vertically aligned. As a consequence, the product vanishes. $\square$

Property 4. Merging two lines with n and n' green squares, respectively, results in a line with $n + n'$ green squares.

Example A1. To illustrate the property, consider the following component contractions:

$$\begin{array}{cc} \begin{array}{c} \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \\ \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \end{array} & , & \begin{array}{c} \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \\ \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \end{array} \\ & & \begin{array}{c} \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \\ \blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \end{array} \end{array}, \quad (\text{A6})$$

which result in three components with the same shape:

$$\blacksquare \blacksquare \blacksquare \blacksquare \cdots \blacksquare \quad (\text{A7})$$

The sum of these contributions gives

$$\sum_{i=1}^3 u_i v_i w_i \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \mathbb{1}_2 \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \mathcal{I}, \quad (\text{A8})$$

where u_i, v_i, w_i are the nonzero matrix elements for the three components and $\mathcal{I}$ denotes the tensor product of identities over the remaining squares. Therefore, the result of the three contractions can be combined into a single scalar factor multiplying a single tensor product of 2-tensors. This sum can always be expressed as a product $\sum_{i=1}^3 u_i v_i w_i = x_1 x_2 x_3$, leading to

$$\begin{pmatrix} 0 & 0 \\ x_1 & 0 \end{pmatrix} \otimes \mathbb{1}_2 \otimes \begin{pmatrix} 0 & 0 \\ x_2 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ x_3 & 0 \end{pmatrix} \otimes \mathcal{I}. \quad (\text{A9})$$

Thus, in practice, one fixes the pattern of the component resulting from the contraction, collects the corresponding components and coefficients, and then reshapes the result as a single tensor product.

The proof of this property in the general case requires some technical tools. We show first that the product of two $\mathcal{A}_i^{[j]n_j} \mathcal{A}_i^{[j+1]n_{j+1}}$ returns $\binom{n}{n_j+n_{j+1}}$ objects.

Property 5. Let $\mathbf{s}$ be a binary vector of n elements. Then, the nonzero elements of the matrix resulting from the tensor product

$$\bigotimes_{k=1}^n A_{i_k \sigma_k^{(m)}}^{[m]} \Big|_{\{\sigma_k^{(m)}\}=\mathbf{s}} \quad (\text{A10})$$

take the form

$$\prod_{\{k: \sigma_k=1\}} u_m^{i_k}. \quad (\text{A11})$$

The following lemma corresponds to Property 4.

Lemma 2. Let $\mathbf{r}$, $\mathbf{s}^{(m)}$, and $\mathbf{s}^{(m+1)}$ be binary vectors of n elements and norm n_m and n_{m+1} , respectively. Thus, for any $\mathbf{r}$,

$$\sum_{\{\Omega\}} \bigotimes_{k=1}^n A_{i_k \sigma_k^{(m)}}^{[m]} \Big|_{\{\sigma_k^{(m)}\}=\mathbf{s}^{(m)}} \times \bigotimes_{k=1}^n A_{i_k \sigma_k^{(m+1)}}^{[m+1]} \Big|_{\{\sigma_k^{(m+1)}\}=\mathbf{s}^{(m+1)}} \\ = \bigotimes_{k=1}^n \tilde{A}_{i_k \tilde{\sigma}_k^{(m)}}^{[m+1]} \Big|_{\{\tilde{\sigma}_k^{(m)}\}=\mathbf{r}}, \quad (\text{A12})$$

with

$$\Omega = \{(\mathbf{s}^{(m)}, \mathbf{s}^{(m+1)}) : \mathbf{s}^{(m)} + \mathbf{s}^{(m+1)} = \mathbf{r}\} \quad (\text{A13})$$

and

$$\tilde{A}_{i_k 1}^{[m+1]} = \begin{pmatrix} 0 & 0 \\ \tilde{u}_{m+1}^{i_k} & 0 \end{pmatrix}, \quad \tilde{A}_{i_k 0}^{[m+1]} = \mathbb{1}, \quad (\text{A14})$$

$$\tilde{u}_{m+1}^{i_k} : \prod_{k=1}^n \tilde{u}_{m+1}^{i_k} = \sum_{\{\Omega\}} \prod_{\substack{k: \mathbf{s}_k^{(m)}=1 \\ k': \mathbf{s}_{k'}^{(m+1)}=1}} u_m^{i_k} u_{m+1}^{i_{k'}}. \quad (\text{A15})$$

Proof. The right-hand side of Eq. (A12) becomes a sum of tensor products in all of which the nonidentity matrices occupy the same positions. Therefore, given the structure of the A -matrices, these tensor products will return matrices with nonzero elements of the form in Eq. (A11), which depend on the specific vector $\mathbf{s}$, in the exact same positions. Consequently, the result of the sum will also be a matrix with the same structure, where the nonzero elements are the sum of the nonzero elements of the initial matrices. This sum can be equivalently expressed as a product of scalar factors, with one factor for each entry equal to 1 in $\mathbf{r}$. Furthermore, using Property 5 again, we can rewrite the sum of tensor products as a single tensor product, where the factors are matrices with the same structure as the original ones, but whose nonzero elements are determined by the decomposition in Eq. (A10) of the sum expressed in product form. ■

In terms of graphs, this corresponds to the merging of two lines, as it is defined in Property 4. In the computation of the permanent in the main text, we have, by construction, $\sum_{i=0}^n \mathbf{s}_i^{(m)} = m$ and $\sum_{i=0}^n \mathbf{s}_i^{(m+1)} = 1$, which, by definition, implies $\sum_{i=0}^n \mathbf{r}_i = m+1$. Lemma 2 guarantees that the contraction of two $\mathcal{A}$ -matrices yields a sum of tensor products, where each term corresponds to a distinct binary vector $\mathbf{r}$. Given that the number of binary vectors of length n containing $m+1$ ones is $\binom{n}{m+1}$, this is the number of tensor products in the resulting sum. Notably, each step of the decomposition [Eq. (A15)] involves a single division and the increment of a sum by one element; thus, it can be computed with number of operations $\mathcal{O}(n)$.

APPENDIX B: GENERATION OF SIERPIŃSKI TRIANGLES OUT OF $\mathcal{A}$ TENSORS

The single-photon bosonic-operator-basis tensor network can be turned into a standard Fock-basis tensor network through Eq. (2), by expressing the A -matrices in terms of the canonical basis of $\mathcal{M}_{2 \times 2}$ and then by replacing each component with the corresponding $d \times d$ matrix, where d is the local dimension cutoff of bosonic operators. In

the multiphoton case, we can perform the same operation starting from the $\mathcal{A}$ tensors. In both cases, we can isolate the tensor corresponding to a specific photon number n_j by setting to zero all components featuring a power of $\hat{a}_j^\dagger$ different from n_j . While this approach can construct the target tensors, such procedure is highly inefficient as it relies on the generation of all the components followed by postselection, nullifying the advantage of the method compared to ordinary MPS. In this Appendix, we show in detail how we can overcome this problem by constructing the structure of an $\mathcal{A}$ tensor for a generic number of photons as shown in Fig. 1. Given the enumerated collection S of all n coefficients u_m^i involved in constructing the target $\mathcal{A}$ tensor, we generate the corresponding power set. To order the power set, we rely on its isomorphism with the binary representations of integers in the range $\{0, \dots, 2^n - 1\}$. In the case represented in Fig. 1, $S = (u_m^i u_m^j u_m^k)$ and its ordered power set is $\pi(S) = \{(), (u_m^k), (u_m^j), (u_m^i u_m^k), (u_m^i), (u_m^i u_m^k), (u_m^i u_m^j), (u_m^i u_m^j u_m^k)\}$. The set difference or relative complement between the sets A and B is defined as $A \setminus B = \{x \in A : x \notin B\}$. Note that the empty set $()$ is included in $\pi(S)$ and $() \setminus A = \emptyset \forall A \neq ()$ and $A \setminus () = A$. As with many logical functions [22], the truth table of the relative complement over the elements of $\pi(S)$ generates a Sierpiński gasket pattern depicted in Fig. 1, which perfectly reproduces the structure of the tensors in Eq. (12). Furthermore, the regularity of this structure facilitates the identification and construction of both the coefficients and their positions within the tensor for all blocks that share the same power of creation operators, without the need of generating the whole matrix, simplifying the conversion between the bosonic-operator-basis MPS and the standard MPS.

APPENDIX C: EXAMPLE: FOUR-PHOTON INTERFERENCE

Here, we explicitly employ our formalism to compute the probability amplitude of observing a specific output configuration in a four-photon interference experiment in the defectless scenario. This computation of the permanent uses the same operator-basis MPS contractions and tensor network representation described in the main paper, illustrating a concrete application of the formalism.

The input state is $\hat{a}_1^\dagger \hat{a}_2^\dagger \hat{a}_3^\dagger \hat{a}_4^\dagger |\text{vac}\rangle$ evolves through the unitary optical circuit into a superposition of Fock states and we compute the probability of detecting the state $\hat{a}_1^\dagger \hat{a}_3^\dagger \hat{a}_7^\dagger \hat{a}_8^\dagger |\text{vac}\rangle$ through photon-counting measurement on the output state. According to Eq. (4), the target amplitude corresponds to the product

$$\alpha_{\bar{n}} = \mathcal{A}_{\bar{i}}^{[1]1} \mathcal{A}_{\bar{i}}^{[3]1} \mathcal{A}_{\bar{i}}^{[7]1} \mathcal{A}_{\bar{i}}^{[8]1}, \quad (\text{C1})$$

where $\bar{n} = (1, 3, 7, 8)$ and $\bar{i} = (1, 2, 3, 4)$, and

$$\mathcal{A}_{\bar{i}}^{[j]1} = \sum_{\sigma_1^{(j)} + \dots + \sigma_4^{(j)} = 1} A_{1\sigma_1^{(j)}}^{[j]} \otimes A_{2\sigma_2^{(j)}}^{[j]} \otimes A_{3\sigma_3^{(j)}}^{[j]} \otimes A_{4\sigma_4^{(j)}}^{[j]}. \quad (\text{C2})$$

We start by computing the matrix product

$$\begin{aligned} \mathcal{A}_i^{[1]1} \mathcal{A}_i^{[3]1} &= (A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{30}^{[1]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{21}^{[1]} \otimes A_{30}^{[1]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{31}^{[1]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{30}^{[1]} \otimes A_{41}^{[1]}) \\ &\times (A_{11}^{[3]} \otimes A_{20}^{[3]} \otimes A_{30}^{[3]} \otimes A_{40}^{[3]} + A_{11}^{[3]} \otimes A_{21}^{[3]} \otimes A_{30}^{[3]} \otimes A_{40}^{[3]} + A_{11}^{[3]} \otimes A_{20}^{[3]} \otimes A_{31}^{[3]} \otimes A_{40}^{[3]} + A_{11}^{[3]} \otimes A_{20}^{[3]} \otimes A_{30}^{[3]} \otimes A_{41}^{[3]}), \end{aligned} \quad (C3)$$

which, due to the nilpotence of the A -matrices, yields

$$\begin{aligned} \mathcal{A}_i^{[1]1} \mathcal{A}_i^{[3]1} &= (A_{11}^{[1]} \otimes A_{21}^{[3]} \otimes A_{30}^{[1]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{31}^{[3]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{30}^{[1]} \otimes A_{41}^{[3]} + A_{11}^{[3]} \otimes A_{21}^{[1]} \otimes A_{30}^{[1]} \otimes A_{40}^{[1]} \\ &+ A_{11}^{[1]} \otimes A_{21}^{[1]} \otimes A_{31}^{[3]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{21}^{[1]} \otimes A_{30}^{[1]} \otimes A_{41}^{[3]} + A_{11}^{[3]} \otimes A_{20}^{[1]} \otimes A_{31}^{[1]} \otimes A_{40}^{[1]} + A_{11}^{[1]} \otimes A_{21}^{[3]} \otimes A_{31}^{[1]} \otimes A_{40}^{[1]} \\ &+ A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{31}^{[1]} \otimes A_{41}^{[3]} + A_{11}^{[3]} \otimes A_{20}^{[1]} \otimes A_{30}^{[1]} \otimes A_{41}^{[1]} + A_{11}^{[1]} \otimes A_{21}^{[3]} \otimes A_{30}^{[1]} \otimes A_{41}^{[1]} + A_{11}^{[1]} \otimes A_{20}^{[1]} \otimes A_{31}^{[3]} \otimes A_{41}^{[1]}), \end{aligned} \quad (C4)$$

where, with an abuse of notation, we set $A_{i0}^{[k]} \times A_{i1}^{[k']} = A_{i1}^{[k']} = (u_{k'}^i \quad 0)$. Each term in the sum above is a horizontal vector with a single nonzero element given by the product of the u_j^i coefficients corresponding to each $A_i^{[j]}$ -matrix. Analogously, we can compute the product of the last two tensors:

$$\begin{aligned} \mathcal{A}_i^{[7]1} \mathcal{A}_i^{[8]1} &= (A_{11}^{[7]} \otimes A_{21}^{[8]} \otimes A_{30}^{[7]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{20}^{[7]} \otimes A_{31}^{[8]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{20}^{[7]} \otimes A_{30}^{[7]} \otimes A_{41}^{[8]} \\ &+ A_{11}^{[8]} \otimes A_{21}^{[7]} \otimes A_{30}^{[7]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{21}^{[7]} \otimes A_{31}^{[8]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{21}^{[7]} \otimes A_{30}^{[7]} \otimes A_{41}^{[8]} \\ &+ A_{11}^{[8]} \otimes A_{20}^{[7]} \otimes A_{31}^{[7]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{21}^{[8]} \otimes A_{31}^{[7]} \otimes A_{40}^{[7]} + A_{11}^{[7]} \otimes A_{20}^{[7]} \otimes A_{31}^{[7]} \otimes A_{41}^{[8]} \\ &+ A_{11}^{[8]} \otimes A_{20}^{[7]} \otimes A_{30}^{[7]} \otimes A_{41}^{[7]} + A_{11}^{[7]} \otimes A_{21}^{[8]} \otimes A_{30}^{[7]} \otimes A_{41}^{[7]} + A_{11}^{[7]} \otimes A_{20}^{[7]} \otimes A_{31}^{[8]} \otimes A_{41}^{[7]}), \end{aligned} \quad (C5)$$

and with the final contraction we obtain

$$\begin{aligned} \alpha_{\bar{n}} &= u_1^1 u_3^2 u_7^3 u_8^4 + u_1^1 u_3^2 u_8^3 u_7^4 + u_1^1 u_7^2 u_3^3 u_8^4 + u_1^1 u_8^2 u_3^3 u_7^4 + u_1^1 u_7^2 u_8^3 u_3^4 + u_1^1 u_8^2 u_7^3 u_3^4 + u_3^1 u_1^2 u_7^3 u_8^4 + u_3^1 u_1^2 u_8^3 u_7^4 \\ &+ u_1^1 u_7^2 u_3^3 u_8^4 + u_8^1 u_1^2 u_3^3 u_7^4 + u_7^1 u_1^2 u_8^3 u_3^4 + u_8^1 u_1^2 u_7^3 u_3^4 + u_3^1 u_7^2 u_1^3 u_8^4 + u_3^1 u_8^2 u_1^3 u_7^4 + u_7^1 u_3^2 u_1^3 u_8^4 + u_8^1 u_3^2 u_1^3 u_7^4 \\ &+ u_7^1 u_8^2 u_1^3 u_3^4 + u_8^1 u_7^2 u_1^3 u_3^4 + u_3^1 u_7^2 u_8^3 u_1^4 + u_3^1 u_8^2 u_7^3 u_1^4 + u_7^1 u_3^2 u_8^3 u_1^4 + u_8^1 u_3^2 u_7^3 u_1^4 + u_7^1 u_8^2 u_3^3 u_1^4 + u_8^1 u_7^2 u_3^3 u_1^4, \end{aligned} \quad (C6)$$

which, as shown in Ref. [8], is the permanent of the matrix

$$\mathcal{U}_{\bar{i}, \bar{n}} = \begin{pmatrix} u_1^1 & u_1^2 & u_1^3 & u_1^4 \\ u_3^1 & u_3^2 & u_3^3 & u_3^4 \\ u_7^1 & u_7^2 & u_7^3 & u_7^4 \\ u_8^1 & u_8^2 & u_8^3 & u_8^4 \end{pmatrix}. \quad (C7)$$

The same result can be obtained more easily by using the graphical notation presented in the main text. The target probability amplitude is given by the contractions in the following graph according to the rules presented in the main text:

$$\alpha_{\bar{n}} = \begin{matrix} & 1 & 2 & 3 & 4 \\ \begin{matrix} 1 \\ 3 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} \blacksquare & \square & \square & \square \\ \square & \blacksquare & \square & \square \\ \square & \square & \blacksquare & \square \\ \square & \square & \square & \blacksquare \end{bmatrix} \end{matrix} \quad (C8)$$

where we have put $\bar{i}$ and $\bar{n}$ on the top and on the left of the graph, respectively. Merging the first couple of rows generates the following (summed up) configurations, which we report together with the coefficients $\tilde{u}$ corresponding to the

$\tilde{A}$ -matrices computed according to Eq. (A12):

	1	2	3	4	
1, 3					$(1 + u_1^1 u_3^2 / u_1^2 u_3^1, u_1^2 u_3^1)$
1, 3					$(1 + u_1^1 u_3^3 / u_1^3 u_3^1, u_1^3 u_3^1)$
1, 3					$(1 + u_1^1 u_3^4 / u_1^4 u_3^1, u_1^4 u_3^1)$
1, 3					$(1 + u_1^2 u_3^3 / u_1^3 u_3^2, u_1^3 u_3^2)$
1, 3					$(1 + u_1^2 u_3^4 / u_1^4 u_3^2, u_1^4 u_3^2)$
1, 3					$(1 + u_1^3 u_3^4 / u_1^4 u_3^3, u_1^4 u_3^3).$

(C9)

Analogously, for the last couple of rows we obtain

	1	2	3	4	
7, 8					$(1 + u_7^3 u_8^4 / u_7^4 u_8^3, u_7^4 u_8^3)$
7, 8					$(1 + u_7^2 u_8^4 / u_7^4 u_8^2, u_7^4 u_8^2)$
7, 8					$(1 + u_7^2 u_8^3 / u_7^3 u_8^2, u_7^3 u_8^2)$
7, 8					$(1 + u_7^1 u_8^4 / u_7^4 u_8^1, u_7^4 u_8^1)$
7, 8					$(1 + u_7^1 u_8^3 / u_7^3 u_8^1, u_7^3 u_8^1)$
7, 8					$(1 + u_7^1 u_8^2 / u_7^2 u_8^1, u_7^2 u_8^1).$

(C10)

For each row in Eq. (C9), there is exactly one row in Eq. (C10) that can combine with it without resulting in zero. Each

nonzero combination contributes a term to the final result, which is the product of the four $\tilde{u}$ values associated with the two rows. Consequently, the final result can be expressed as follows:

$$\alpha_{\tilde{n}} = (u_1^1 u_3^2 + u_7^2 u_3^1)(u_7^3 u_8^4 + u_7^4 u_8^3) + (u_1^1 u_3^3 + u_7^3 u_3^1)(u_7^2 u_8^4 + u_7^4 u_8^2) + (u_1^1 u_3^4 + u_7^4 u_3^1)(u_7^2 u_8^3 + u_7^3 u_8^2) + (u_7^2 u_3^3 + u_7^3 u_3^2)(u_1^1 u_8^4 + u_7^4 u_8^1) + (u_1^2 u_3^4 + u_7^4 u_3^2)(u_7^1 u_8^3 + u_7^3 u_8^1) + (u_7^3 u_3^4 + u_7^4 u_3^3)(u_1^1 u_8^2 + u_7^2 u_8^1), \quad (\text{C11})$$

which corresponds to the permanent in Eq. (C6). In both the purely algebraic and graphical computation, we observe that the number of operations is $O(n^2 2^n)$.

APPENDIX D: NÄIVE SCALING FOR PERMANENT

For the sake of argument, we consider the case $n_i = 1$ for $i \leq n$, and 0 otherwise. Using the mixed-product property of the Kronecker product over all the matrix products in Eq. (5) yields

$$\mathcal{A}_i^{[1]n_1} \mathcal{A}_i^{[2]n_2} \dots \mathcal{A}_i^{[M]n_M} = \sum_{\sum_{k=1}^n \sigma_k^{(1)} = n_1} \dots \sum_{\sum_{k=1}^n \sigma_k^{(M)} = n_M} \left(A_{i_1 \sigma_1^{(1)}}^{[1]} A_{i_1 \sigma_1^{(2)}}^{[2]} \dots A_{i_1 \sigma_1^{(M)}}^{[M]} \right) \otimes \dots \otimes \left(A_{i_n \sigma_n^{(1)}}^{[1]} A_{i_n \sigma_n^{(2)}}^{[2]} \dots A_{i_n \sigma_n^{(M)}}^{[M]} \right). \quad (\text{D1})$$

Due to the nilpotence of the matrices $A_{i1}^{[m]}$, in each product of matrices in the equation above there can be at most one nonzero (1) σ index. Furthermore, for each valid configuration, for a fixed k , only one $\sigma_j^{(k)}$ is allowed to be 1. Therefore counting the nonzero elements of the sum above corresponds to counting all the possible permutations of the lines of an $n \times n$ identity matrix, that together with the cost of all the matrix multiplications (n) gives the computational cost $n \cdot n!$.

-
- [1] R. Orús, Tensor networks for complex quantum systems, *Nat. Rev. Phys.* **1**, 538 (2019).
 - [2] S. Montangero, *Introduction to Tensor Network Methods* (Springer, Cham, 2018).
 - [3] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011).
 - [4] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner, Entanglement properties of the harmonic chain, *Phys. Rev. A* **66**, 042327 (2002).
 - [5] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, *Rev. Mod. Phys.* **82**, 277 (2010).
 - [6] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
 - [7] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117 (2014).
 - [8] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *STOC '11: Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing, San Jose California USA* (Association for Computing Machinery, New York, NY, 2011), pp. 333–342.
 - [9] R. García-Patrón, J. J. Renema, and V. Shchesnovich, Simulating boson sampling in lossy architectures, *Quantum* **3**, 169 (2019).
 - [10] C. Oh, M. Liu, Y. Alexeev, B. Fefferman, and L. Jiang, Classical algorithm for simulating experimental Gaussian boson sampling, *Nat. Phys.* **20**, 1461 (2024).
 - [11] M. J. Hartmann, J. Prior, S. R. Clark, and M. B. Plenio, Density matrix renormalization group in the Heisenberg picture, *Phys. Rev. Lett.* **102**, 057202 (2009).
 - [12] S. R. Clark, J. Prior, M. J. Hartmann, D. Jaksch, and M. B. Plenio, Exact matrix product solutions in the Heisenberg picture of an open quantum spin chain, *New J. Phys.* **12**, 025005 (2010).
 - [13] D. Cilluffo, N. Lorenzoni, and M. B. Plenio, Simulating Gaussian boson sampling with tensor networks in the Heisenberg picture, [arXiv:2305.11215](https://arxiv.org/abs/2305.11215).
 - [14] C. Oh, K. Noh, B. Fefferman, and L. Jiang, Classical simulation of lossy boson sampling using matrix product operators, *Phys. Rev. A* **104**, 022407 (2021).
 - [15] H. J. Ryser, *Combinatorial Mathematics*, 1st ed. (Mathematical Association of America, Buffalo, 1963), Vol. 14.
 - [16] C. Oh, L. Jiang, and B. Fefferman, On classical simulation algorithms for noisy boson sampling, [arXiv:2301.11532](https://arxiv.org/abs/2301.11532).
 - [17] M. Liu, C. Oh, J. Liu, L. Jiang, and Y. Alexeev, Complexity of Gaussian boson sampling with tensor networks, *Phys. Rev. A* **108**, 052604 (2023).
 - [18] Q. Zhou and Y. Deng, Generating Sierpinski gasket from matrix calculus in Dempster–Shafer theory, *Chaos Solitons Fractals* **166**, 112962 (2023).
 - [19] Note that each A -matrix corresponds to a single element of the matrix $\mathcal{U}$, specifically $A_{i_j \sigma_j^{(k)}}^{[k]} \rightarrow u_k^{i_j}$. Consequently, if we denote with $w(t_i)$ the multiplicity of t_j into $\bar{t}$, for a fixed $\bar{n}$, the set of the A matrices identifies a unique matrix $\mathcal{U}'$ obtained by repeating $w(t_i)$ times its t_i th column and n_i times its i th row. These matrices are precisely those that appear in boson sampling probabilities [8].
 - [20] S. Aaronson and T. Hance, Generalizing and derandomizing Gurvits’s approximation algorithm for the permanent, [arXiv:1212.0025](https://arxiv.org/abs/1212.0025).
 - [21] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley, Efficient classical algorithm for boson sampling with partially distinguishable photons, *Phys. Rev. Lett.* **120**, 220502 (2018).
 - [22] P. S. Denis and P. Grim, Fractal images of formal systems, *J. Philos. Logic* **26**, 181 (1997).