Universality of generalized bunching and efficient assessment of Boson Sampling

V. S. Shchesnovich
Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP, 09210-170 Brazil

It is found that N identical bosons (fermions) show generalized bunching (antibunching) in $1 \leq \mathcal{K} \leq M-1$ output modes of a random M-mode linear network: The maximum (minimum) probability of detecting all input particles in $\mathcal{K}$ output modes is attained only when the bosons (fermions) are completely indistinguishable. For fermions $\mathcal{K}$ is arbitrary, for bosons it is either (i) arbitrary for only classically correlated particles or (ii) satisfies $\mathcal{K} \geq N$ (or $\mathcal{K} = 1$) for arbitrary input states. The generalized bunching supplies an efficient assessment protocol of Boson Sampling with an *arbitrary network*, which requires only a polynomial number of runs of experimental device and computation of only one matrix permanent of a positive definite Hermitian matrix (with a value close to 1), with an analytic formula for the Scattershot version of Boson Sampling.

PACS numbers: 42.50.St, 03.67.Ac, 42.50.Ar

arXiv:1509.01561v3 [quant-ph] 2 Nov 2015

Introduction.– Optical networks with photons have become a growing research field with potential application in quantum computing [1–3]. The idea of Boson Sampling (BS) [2] as a non-universal but near-future feasible device aiming at the extended Church-Turing thesis, followed by spectacular experiments with optical networks with growing size and number of photons [4–10], can be a way for benchmark demonstration of quantum supremacy. Growing complexity of networks used for BS poses a problem of assessment of device operation. Indeed, on the one hand, the very nature of computational complexity of the matrix permanents [11, 12] prevents verification of the BS in a polynomial time, without explicit reliance on details of experimental setup. On the other hand, unavoidable experimental errors may degrade output of such a device. Indeed, an unwanted distinguishability introduced by a network and photon losses are found to be the dominating factors in the optical setups [4–10], along with the higher photon numbers introduced by currently used photon sources. Not surprisingly, finding ways to assess how close is output distribution of an experimental device to BS [2, 13] has become one of the central issues. Zero-transmission laws in the Fourier network [14], boson clouding [8], coincidence counting probabilities [9, 10], and statistical benchmarking [15] were considered.

Nevertheless, the assessment problem remains open. The test of Ref. [14] checks indistinguishability of N photons in a fixed (and difficult to build), Fourier network, the statistical method of Ref. [15] can only assess the second-order coherence function [47], and the current methods of assessment of a BS device with a random network rely on computation of exponentially small probabilities (responsible for computational complexity of BS) requiring an exponential number of runs of the device for accuracy. Already with small networks of dozen or so modes some experimentalists resort to Bayesian methods of hypothesis testing [9, 10].

Generalized bunching and efficient assessment of BS.– Since the (believed) quantum supremacy of BS is due to

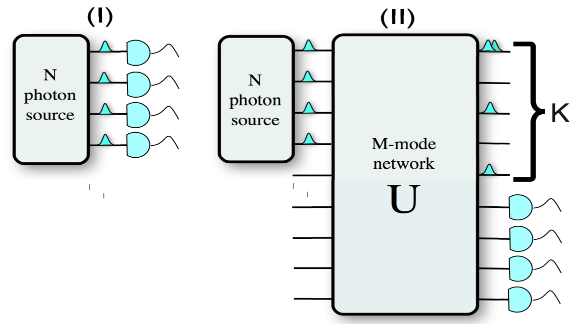

FIG. 1: (Color online) Assessment of a BS device. Step (I): the source is checked for output with only one particle per mode. Step (II), the probability of all N input particles to appear in $\mathcal{K}$ output modes of a M-mode network U is monitored by $M-\mathcal{K}$ bucket detectors: Given U, the maximum probability (given in Eqs. (6) and (8)) is attained *only* by the completely indistinguishable bosons and is *close to* 1 for $N(M-\mathcal{K}) \ll M$.

quantum statistics of indistinguishable bosons, an assessment protocol of a BS device (with a random network) could be based on a network-independent (universal) feature in its output distribution (i) unique for completely indistinguishable bosons and (ii) requiring only a polynomial number of runs of the device. Such a feature is the generalized bunching, i.e., the absolute maximum of probability of detecting *all* N input bosons in $\mathcal{K}$ output modes of an M-mode linear network. The maximum probability is attained *only* by completely indistinguishable bosons (i) over arbitrary input states of particles for $\mathcal{K} \geq N$ (and $\mathcal{K} = 1$) or (ii) for arbitrary $1 \leq \mathcal{K} \leq M-1$ over only classically-correlated bosons. The minimum probability, over all possible input states, is attained by completely indistinguishable fermions (generalized antibunching), with the probability to detect distinguishable particles lying in the middle.

The generalized bunching can be viewed as an N-particle generalization of the Hong-Ou-Madel (HOM) effect [16] to arbitrary M-mode networks. Previously, generalizations of the HOM effect to many-particle multimode setups were reported in coincidence counting [17, 18] and zero transmission laws [17, 19] in Bell multiport networks. In a random M-mode network, boson bunching in a single mode [20] has probability decaying as $\sim e^{-N}$ for $M \geq N^2$ [21]. A similar effect is boson tendency to form clouds in output modes in some networks [8]. In contrast, *the generalized bunching can be observed with a probability close to* 1 *in an arbitrary (nontrivial) quantum network.* Though in the dilute limit, $M \geq N^2$, the generalized bunching/antibunching effect wanes as $N \to \infty$, for any finite N there is a witness of the complete indistinguishability detectable in a polynomial number of runs. The latter enables the assessment protocol of a BS device sketched in Fig. 1.

The generalized bunching/antibunching is found by revealing an equivalence between probability of detecting all input particles in a subset of output modes of a linear network and an eigenvalue problem for positive semidefinite (p.s.d.) Hermitian matrix (Eq. (7) below).

Description of partially distinguishable identical particles in a linear network.– We consider M-mode linear quantum network, Fig. 1(b), with N single identical particles at input modes $k_1, \ldots, k_N$ in arbitrary internal state (ordered products of operators in case of fermions)

$$\rho = \sum_i q_i |\Psi_i\rangle\langle\Psi_i|, \quad |\Psi_i\rangle = \sum_{\mathbf{j}} C^{(i)}_{\mathbf{j}} \prod_{\alpha=1}^{N} a^\dagger_{k_\alpha, j_\alpha}|0\rangle, \quad (1)$$

where $q_i \geq 0$, $\sum_i q_i = 1$, $\mathbf{j} = (j_1, \ldots, j_N)$, $\sum_{\mathbf{j}} |C^{(i)}_{\mathbf{j}}|^2 = 1$, and a mode operator $a^\dagger_{k,j}$ (and below $b^\dagger_{k,j}$) creates a particle in an input (output) mode k and an internal basis state $|j\rangle \in \mathcal{H}$ (e.g., a basis function of spectral shape of a photon). A unitary network with matrix U, Fig. 1(b), relates the input $a_{k,j}$ and output $b_{l,j}$ modes: $a^\dagger_{k,j} = \sum_{l=1}^{M} U_{k,l} b^\dagger_{l,j}$. The internal state of identical particles, defined as

$$\rho^{(int)} = \sum_i q_i |\psi_i\rangle\langle\psi_i|, \quad |\psi_i\rangle \equiv \sum_{\mathbf{j}} C^{(i)}_{\mathbf{j}} \prod_{\alpha=1}^{N}{}^{\otimes} |j_\alpha\rangle, \quad (2)$$

governs their behavior in a linear network. Symmetry properties of $\rho^{(int)}$ under the permutation group $\mathcal{S}_N$ play the key role [22–27], e.g., particles of one species with $\rho^{(int)}$ antisymmetric under permutations emulate behavior of the other species [28, 29]. The probability formula of an output configuration $\mathbf{m} = (m_1, \ldots, m_M)$ [22, 23] (applicable also to fermions [32]) reads

$$\hat{p}(\mathbf{m}) = \frac{1}{\prod_{l=1}^{M} m_l!} \sum_{\tau,\sigma \in \mathcal{S}_N} J(\tau\sigma^{-1}) \prod_{\alpha=1}^{N} U^*_{k_{\tau(\alpha)}, l_\alpha} U_{k_{\sigma(\alpha)}, l_\alpha}, \quad (3)$$

where $l_1, \ldots, l_N$ are output modes with multiplicities $(m_1, \ldots, m_M)$, and a complex-valued function $J(\sigma)$ of $\sigma \in \mathcal{S}_N$ is defined as

$$J(\sigma) = \varepsilon(\sigma) \mathrm{Tr}\big(\rho^{(int)} P_\sigma\big), \; \varepsilon(\sigma) = \begin{cases} 1, & \text{Bosons}, \\ \mathrm{sgn}(\sigma), & \text{Fermions}, \end{cases} \quad (4)$$

where $P_\sigma \prod_{\alpha=1}^{N}{}^{\otimes} |j_\alpha\rangle = \prod_{\alpha=1}^{N}{}^{\otimes} |j_{\sigma^{-1}(\alpha)}\rangle$ is the operator representation of σ in $\mathcal{H}^{\otimes N}$.

Identical particles are called completely indistinguishable if $\rho^{(int)}$ is symmetric under permutations (e.g., particles in the same internal state), i.e., $J^{(id)}(\sigma) = \varepsilon(\sigma)$, whereas particles with orthogonal internal states are distinguishable with $J^{(d)}(\sigma) = \delta_{\sigma,I}$ (see also Refs. [23, 29]). The trace of $\rho^{(int)}$ in the symmetric subspace $S_N \mathcal{H}^{\otimes N}$ with $S_N = (1/N!) \sum_\sigma P_\sigma$, defined as $d(J) \equiv \mathrm{Tr}\{S_N \rho^{(int)}\} = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \varepsilon(\sigma) J(\sigma)$ is a suitable measure of the partial indistinguishability [30] and also in some other context involving identical particles [31].

Note that $J(\sigma)$ of Eq. (4) is a normalized p.s.d. function of $\sigma \in \mathcal{S}_N$, i.e., for any complex-valued function $z(\sigma)$ we have $\sum\limits_{\sigma_1,\sigma_2} z^*(\sigma_1) J(\sigma_1 \sigma_2^{-1}) z(\sigma_2) \geq 0$, while $J(I) = 1$ for the identity permutation I. Therefore, there is a factorization $\theta(\sigma)$ such that [32]

$$J(\sigma) = \sum_{\tau \in \mathcal{S}_N} \theta^*(\tau)\theta(\tau\sigma), \quad \sum_{\sigma \in \mathcal{S}_N} |\theta(\sigma)|^2 = 1. \quad (5)$$

Importantly, any normalized p.s.d. function $J(\sigma)$ corresponds to an input state of single particles [32], i.e., the set of input states of N single particles can be equivalently represented by the convex set of normalized p.s.d. functions $\mathcal{S}_N \to \mathbb{C}$.

Probability to detect all particles in $\mathcal{K}$ output modes.– The probability for all N input particles to gather in $\mathcal{K}$ output modes, say $1, \ldots, \mathcal{K}$ as in Fig. 1(b), reads $p(J) = \sum'_{\mathbf{m}} \hat{p}(\mathbf{m})$ where $\mathbf{m} = (m_1, \ldots, m_{\mathcal{K}}, 0 \ldots, 0)$. Defining an N-dimensional p.s.d. Hermitian matrix H, built from the submatrix of U on the rows $k_1, \ldots, k_N$ and columns $1, \ldots, \mathcal{K}$, and an $(N!)$-dimensional Schur power matrix $\Pi(H)$ indexed by elements of $\mathcal{S}_N$ [48],

$$H_{\alpha,\beta} \equiv \sum_{l=1}^{\mathcal{K}} U_{k_\alpha,l} U^*_{k_\beta,l}, \quad \Pi_{\sigma,\tau}(H) = \prod_{\alpha=1}^{N} H_{\sigma(\alpha),\tau(\alpha)}, \quad (6)$$

we obtain $p(J)$ in the form [32]

$$p(J) = \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \Pi_{I,\sigma} = \sum_{\tau,\tau' \in \mathcal{S}_N} \theta^*(\tau)\theta(\tau')\Pi_{\tau,\tau'}. \quad (7)$$

The probability is therefore a convex combination of eigenvalues of $\Pi(H)$ with functions $\theta(\sigma)$ serving as eigenvectors. For the completely indistinguishable particles $\theta^{(id)}(\sigma) = \varepsilon(\sigma)/\sqrt{N!}$ and Eq. (7) gives

$$p^{(id)}_F = \det(H), \quad p^{(id)}_B = \mathrm{per}(H), \quad (8)$$

for fermions and bosons, respectively. For distinguishable particles $\theta^{(d)}(\sigma) = \delta_{\sigma,I}$ and

$$p_{\mathcal{D}} = \prod_{\alpha=1}^{N} H_{\alpha,\alpha}. \tag{9}$$

The Hadamard-type inequalities [33, 34], i.e., $\det(H) \le \prod_{\alpha=1}^{N} H_{\alpha,\alpha} \le \mathrm{per}(H)$, give $p_F^{(id)} \le p_{\mathcal{D}} \le p_B^{(id)}$ (whose validity extends to an arbitrary input configuration [32]).

The generalized bunching (antibunching) occurs for $\theta(\sigma)$ of Eq. (5) being an eigenvector of the largest (smallest) eigenvalue of $\Pi(H)$ in Eq. (6). A famous result of Schur [35] states that the smallest eigenvalue of $\Pi(H)$ is $\det(H)$, hence, the generalized antibunching is an universal attribute of the completely indistinguishable fermions (a unique minimum for $\mathcal{K} \ge N$). On the other hand, the maximum eigenvalue of $\Pi(H)$ is generally unknown. The permanent-on-top conjecture (POT) [34, 36], stating that universally it is $\mathrm{per}(H)$, proven for $N \le 3$ [37], has turned out false for $N \ge 5$ [38].

Thus, conditions on input states and/or networks are needed to ensure the maximum probability being attained only by the completely indistinguishable bosons. For $\mathcal{K} = 1$ the latter is true for arbitrary network [20, 26, 30], since in this case Eq. (7) gives $p(J) = d(J)N! \prod_{\alpha=1}^{N} |U_{k_\alpha,l}|^2$ with the maximum for $d(J) = 1$, i.e., for $J(\sigma) = 1$. Extensive numerical simulations [49] with random p.s.d. Hermitian matrices reveal that $p_B^{(id)} = \mathrm{per}(H)$ is not the maximum probability *only* for $N \ge 5$ particles in $2 \le \mathcal{K} \le N-1$ output modes (such a state of $N \ge 5$ bosons is necessarily a state of non-classically correlated particles, see below). More importantly, when $\mathcal{K} \ge N$, a *unique eigenvector* $\theta_B^{(id)}(\sigma) = 1/\sqrt{N!}$ corresponds to $\mathrm{per}(H)$, i.e., the maximum probability of N particles to be detected in $\mathcal{K} \ge N$ output modes is attained only by the completely indistinguishable bosons.

If we consider only classically correlated bosons, the completely indistinguishable ones attain the unique maximum of $p(J)$ (7) for all $1 \le \mathcal{K} \le M-1$ (over all classically correlated inputs, including distinguishable particles). Indeed, a classically-correlated internal state can be expressed as a convex combination of pure states, i.e., $\rho^{(int)} = \sum_{\mathbf{j}} \nu_{\mathbf{j}} |\phi_{j_1}\rangle\langle\phi_{j_1}| \otimes \ldots \otimes |\phi_{j_N}\rangle\langle\phi_{j_N}|$ for some arbitrary states $|\phi_{j_\alpha}\rangle \in \mathcal{H}$ and $\nu_{\mathbf{j}} > 0$, $\sum_{\mathbf{j}} \nu_{\mathbf{j}} = 1$. The corresponding J-function reads

$$J^{(cc)}(\sigma) = \sum_{\mathbf{j}} \nu_{\mathbf{j}} \prod_{\alpha=1}^{N} \langle \phi_{j_{\sigma(\alpha)}} | \phi_{j_\alpha} \rangle. \tag{10}$$

Setting $G^{(\mathbf{j})}_{\alpha,\beta} \equiv \langle \phi_{j_\beta} | \phi_{j_\alpha} \rangle$, we get from Eqs. (6), (7), and (10)

$$\begin{aligned} p(J^{(cc)}) &= \sum_{\mathbf{j}} \nu_{\mathbf{j}} \sum_{\sigma \in \mathcal{S}_N} \prod_{\alpha=1}^{N} H_{\alpha,\sigma(\alpha)} G^{(\mathbf{j})}_{\alpha,\sigma(\alpha)} \\ &= \sum_{\mathbf{j}} \nu_{\mathbf{j}} \, \mathrm{per}(H \cdot G^{(\mathbf{j})}), \end{aligned} \tag{11}$$

where the dot stands for the Hadamard (by-element) product. Thus, an old conjecture in linear algebra [39], stating that for two p.s.d. Hermitian matrices H and G (for $G_{\alpha,\alpha} = 1$)

$$\mathrm{per}(H \cdot G) \le \mathrm{per}(H), \tag{12}$$

would imply the claimed result in this case. Eq. (12) requires verification only at matrices H violating the POT conjecture, when $\mathrm{per}(H)$ is not the largest eigenvalue of $\Pi(H)$. It was numerically verified that $\mathrm{per}(H \cdot G) < \mathrm{per}(H)$ for any N random states $|\varphi_1\rangle, \ldots, |\varphi_N\rangle$, with at least two linearly independent.

The assessment protocol of BS device.– Consider the generalized bunching/antibunching effect in a Haar-random network to have an idea of how pronounced it is. By the unitary invariance of the Haar measure, the respective average probability $\langle p_N \rangle$ depends only on the ratio of considered output configurations [32]:

$$\langle p_{B,F}^{(id)} \rangle = \frac{\mathcal{K}(\mathcal{K} \pm 1) \cdot \ldots \cdot (\mathcal{K} \pm N \mp 1)}{M(M \pm 1) \cdot \ldots \cdot (M \pm N \mp 1)}, \tag{13}$$

here (and below) the upper (lower) signs stand for bosons (fermions). For distinguishable particles the probability lies in the middle, for large M it can be shown that [32]

$$\langle p_{\mathcal{D}} \rangle = \left(\frac{\mathcal{K}}{M} \right)^N \left[1 + O\left(\frac{1}{M} \right) \right]. \tag{14}$$

The quantum to classical average probability ratio becomes

$$\frac{\langle p_{B,F}^{(id)} \rangle}{\langle p_{\mathcal{D}} \rangle} = \left[1 + O\left(\frac{1}{M} \right) \right] \prod_{i=1}^{N-1} \frac{1 \pm l/\mathcal{K}}{1 \pm l/M}. \tag{15}$$

For $NL \ll M$, where $L = M - \mathcal{K}$, the detection probability is close to 1: $\langle p_{B,F}^{(id)} \rangle = \prod_{l=0}^{N-1} [1 - L/(M \pm l)] = 1 - O(LN/M)$, whereas the r.h.s. of Eq. (15) is approximated as $\langle p_{B,F}^{(id)} \rangle / \langle p_{\mathcal{D}} \rangle \approx 1 \pm LN(N-1)/(2M^2)$. In this case one needs $R \gg M^4/(N^4 L^2)$ runs for the ratio (15) to surpass the statistical error $O(1/\sqrt{R})$ in experimental data. But the ratio in Eq. (15) is attained *only* by the completely indistinguishable bosons (fermions), hence it is a witness, detectable in polynomial number of runs, of their complete indistinguishability during propagation, i.e., that no decoherence process has contributed to output statistics. Therefore, we have a *network-independent* protocol for assessment of a BS device using only *polynomial number of runs*, Fig. 1. The only known protocol for a BS device with a *Haar-random* network [13], verified in Ref. [8], discriminates BS and uniform distributions only. Our protocol discriminates against all other than BS samplers, which are physically realizable, including the classical ($\mathcal{M}_A$), the fermion ($\mathcal{F}_A$), the random-classical ($\mathcal{B}_A$) samplers of Ref. [13], and the random-phase bosons [14, 15]. Moreover, there is evidence [40] that an efficient classical algorithm may exist to approximate the permanents such as $p_B^{(id)}$ (8) (one per network).

The protocol applies also to the so-called Scattershot BS [41], recently experimentally tested [10], which uses heralded single photons in N different input modes in each run. The probability describing experimental statistics of a Scattershot BS is well approximated by $\langle p_B^{(id)}\rangle$ (13) already for a few hundred runs of the device (see Fig. 2 below), i.e., no classical computation required.

The assessment protocol, depicted in Fig. 1, has two stages. At stage (I), Fig. 1(I), by using photon-number resolving detectors (e.g., by cascading bucket detectors [42]) one checks that sources produce N single photons. At stage (II), Fig. 1(II), employing the universality of generalized bunching, one checks experimental statistics against the probability $p_B^{(id)}$ (8) by using $L = M - \mathcal{K}$ bucket detectors, with typically $L \ll \sqrt{M}$.

Stage (I) would expose attempts to bypass the universality property with inputs having variable number of particles per input mode (not required under certified input). Indeed, an input with any distribution of bosons between M input modes of a unitary network

$$\rho = \sum_{\mathbf{n}} p_{\mathbf{n}} |\mathbf{n}\rangle\langle\mathbf{n}|, \quad p_{\mathbf{n}} \geq 0, \quad \sum_{\mathbf{n}} p_{\mathbf{n}} = 1, \tag{16}$$

where $n_1 + \ldots + n_M = N$, has the Haar-average probability equal to $\langle p_B^{(id)}\rangle$ [32]. Stage (I) in few experimental runs exposes such inputs. For instance, the random-phase bosons of Refs. [14, 15], i.e., an input where each boson is in a coherent superposition of S input modes (and in an internal state $|\phi\rangle$) described by operator $A(\theta) = S^{-\frac{1}{2}} \sum_{j=1}^{S} e^{i\theta_j} a_{k_j,\phi}$ with random phases $\theta_1, \ldots, \theta_S$ [14, 15]. The corresponding density matrix reads

$$\rho_s = \frac{(S-1)!S^N}{(S+N-1)!} \prod_{j=1}^{S} \int_0^{2\pi} \frac{\mathrm{d}\theta_j}{2\pi} \left[A^\dagger(\theta)\right]^N |0\rangle\langle 0| \left[A(\theta)\right]^N$$
$$= \frac{N!(S-1)!}{(S+N-1)!} \sum_{\mathbf{n}} |\mathbf{n}\rangle\langle\mathbf{n}|, \tag{17}$$

where $\mathbf{n} = (n_{k_1}, \ldots, n_{k_S})$, $n_{k_1} + \ldots + n_{k_S} = N$. A source of ρ_s with $S = N$ reveals itself at stage (I) due to a vanishing probability of an input with one particle per each of N input modes, for $N \gg 1$ scaling as $\sim 4^{-N}$. Stage (I) exposes also the sampler $\mathcal{B}_A$ of Ref. [13], a "mockup distribution of BS" physically realized by distributing N uncorrelated distinguishable particles randomly over N input modes, the probability of such a particle to land in an output mode l (assuming $k_\alpha = \alpha$) being $p(l) = \frac{1}{N}\sum_{\alpha=1}^{N} |U_{\alpha,l}|^2$. The probability of an input with one particle per mode reads $N!/N^N \sim e^{-N}$.

Figs. 2 provides numerical tests of the assessment protocol for BS device with fixed (standard BS) and random input modes (Scattershot BS). Here $M = [\frac{1}{2}N^2]$ (integer part), while $L = M - \mathcal{K}$ is optimized to maximize the ratio (15) for a lower bound $\langle p_B^{(id)}\rangle = 0.25$ (M and $\mathcal{K}$ for N in Fig. 2 are given in Table I).

N	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
L	2	2	3	4	5	5	6	7	7	8	9	9	10	11	11	12	13	14
M	5	8	13	18	25	32	41	50	61	72	85	98	113	128	145	162	181	200

TABLE I: The network size M and $L = M - \mathcal{K}$ as functions of N. Here $\mathcal{K}$ is selected by maximizing the ratio of Eq. (15) under the condition that $\langle p_B^{(id)}\rangle \geq 0.25$ (note that $\mathcal{K} \geq N$).

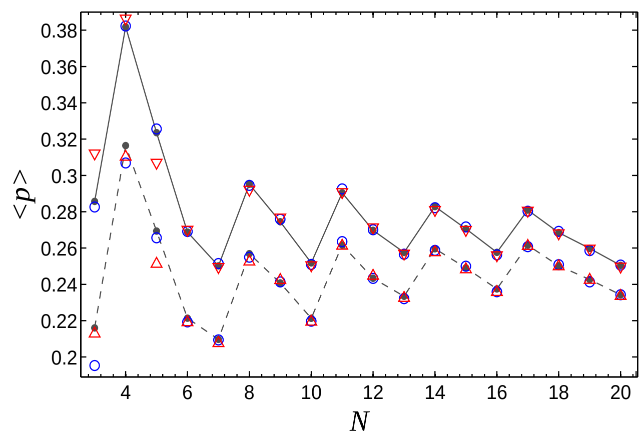

FIG. 2: (Color online) The analytical quantum average probability $\langle p_B^{(id)}\rangle$ Eq. (13) (dots on the solid line) and the approximation in classical case $\langle p_{\mathcal{D}}\rangle$ Eq. (14) (dots on the dashed line) vs. numerical averaging with 1000 Haar-random networks (circles). Quantum and classical probability for Scattershot BS with a random network for each value of N is shown by the triangles.

The trace-distance $\mathcal{D}$ [2] between ideal BS and an experimental device reflects the effect of unavoidable errors in an experimental setup. The protocol allows to estimate $\mathcal{D}$ using the above discussed distinguishability measure $1 - d(J)$ [30]

$$\mathcal{D} \leq 1 - d(J) = 1 - \mathrm{Tr}\{S_N \rho^{(int)}\}, \tag{18}$$

where $1 - d(J)$ is bounded by the experimental value p_{exp} of $p(J)$ (7). Expanding $p(J)$ over the eigenvalues of $\Pi(H)$ (6) gives upper and lower bounds on $1 - d(J)$ [32], by substituting $p_{exp} \to p(J)$ into them we get

$$\frac{\mathrm{per}(H) - p_{exp}}{\mathrm{per}(H) - \det(H)} \leq 1 - d(J) \leq \frac{\mathrm{per}(H) - p_{exp}}{\mathrm{per}(H) - \lambda} \tag{19}$$

with λ being the next largest eigenvalue of $\Pi(H)$ after $\mathrm{per}(H)$. Besides the distinguishability, errors reported in BS experiments [4–10] include the higher-order photon number contribution from the sources, estimated at stage (I) of our protocol, and the photon losses, which can be directly estimated at a network output (see also Ref. [43]). The combined effect of other possible errors can be estimated by the trace distance as in Eqs. (18)-(19). It is expected that a BS device with a fixed-ratio of lost photons is still hard to simulate on a classical computer [2, 45], proved for a fixed number of photons being lost [46]. Since a lossy linear M-mode network is equivalent to an $2M$-mode unitary one, with half of output modes being unaccessible "loss channels" [32], it is clear that our assessment protocol works also for a lossy network, where one should use H built from the proper non-unitary network matrix U, which can be characterized using only classical coherence [44].

Conclusion.– An assessment protocol is discovered for BS, which employs universality of the generalized boson bunching at output of a linear network. It requires only a polynomial number of experimental runs to verify that a device, consisting of a random linear network with single photons at its input, outputs the BS distribution. The protocol requires computation of only one matrix permanent of a p.s.d. Hermitian matrix, of a value close to 1 (conjectured to be easier to approximate than exponentially small probabilities of individual output configurations [40]). Moreover, for assessment of the Scattershot version of BS [41] (having better prospects for scalability) our protocol *does not require classically hard computations*, as there is an analytical result. In general terms, we have found an analog of the famous HOM effect [16] revealing the complete indistinguishability of N photons in an arbitrary (non-trivial) M-mode network, which can find applications where linear networks are used.

This work was supported by the CNPq (Brazil).

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature **409** (2001) 46.
[2] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).
[3] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, Nat. Photonics **9**, 615 (2015).
[4] M. A. Broome *et al*, Science **339**, 794 (2013);
[5] J. B. Spring *et al*, Science, **339**, 798 (2013).
[6] M. Tillmann *et al*, Nature Photonics, **7**, 540 (2013).
[7] A. Crespi *et al*, Nature Photonics, **7**, 545 (2013).
[8] J. Carolan *et al*, Nat. Photon. **8**, 621 (2014).
[9] N. Spagnolo *et al*, Nature Photonics, **8**, 615 (2014).
[10] M. Bentivegna *et al*, Sci. Adv. **1**, 1400255 (2015).
[11] L. G. Valiant, Theoretical Coput. Sci., **8**, 189 (1979).
[12] S. Aaronson, Proc. Roy. Soc. London A, **467**, 3393 (2008).
[13] S. Aaronson and A. Arkhipov, arXiv:1309.7460 [quant-ph].
[14] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Molmer, Phys. Rev. Lett. **113**, 020502 (2014).
[15] M. Walschaers *et al*, arXiv:1410.8547 [quant-ph].
[16] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
[17] Y. L. Lim and A. Beige, New J. Phys. **7**, 155 (2005).
[18] Z. Y. Ou, Phys. Rev. A **77**, 043829 (2008).
[19] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, New Journal of Phys. **14**, 093015 (2012).
[20] N. Spagnolo *et al*, Phys. Rev. Lett. **111**, 130503 (2013).
[21] A. Arkhipov and G.Kuperberg, Geom. Topol. Monogr., **18**, 1 (2012).
[22] V. S. Shchesnovich, Phys. Rev. A **89**, 022333 (2014).
[23] V. S. Shchesnovich, Phys. Rev. A **91**, 013844 (2015).
[24] Z. Y. Ou, Phys. Rev. A **74**, 063808 (2006).
[25] H. de Guise, S.-H. Tan, I. P. Poulin, and B. C. Sanders, Phys. Rev. A **89**, 063819 (2014).
[26] M. C. Tichy, Phys. Rev. A **91**, 022316 (2015).
[27] P. P. Rohde, Phys. Rev. A **91**, 012307 (2015).
[28] J. C. F. Matthews *et al*, Sci. Rep. **3**, 1539 (2013).
[29] V. S. Shchesnovich, Int. J. Quantum Inform. **13**, 1550013 (2015).
[30] V. S. Shchesnovich, Phys. Rev. A **A 91**, 063842 (2015).
[31] M. Miková *et al*, Phys. Rev. A **87**, 042327 (2013).
[32] See the appendix.
[33] E. Fisher, Archiv d. Math. u. Phys. **13**, 32 (1907).
[34] E. H. Lieb, J. Math. and Mech. **16**, 127 (1966).
[35] I. Schur, Math. Z. **1**, 184 (1918).
[36] G. Soules, *Matrix Functions and the Laplace expansion theorem*, PhD Dissertation (University of California, Santa Barbara, 1966).
[37] R. B. Bapat and V. S. Sunder, Linear Algebra Appl. **76**, 153 (1986).
[38] V. S. Shchesnovich, *The permanent-on-top conjecture is false*, accepted to Linear Algebra Appl.
[39] R. B. Bapat and V. S. Sunder, Linear Algebra Appl. **72**, 107 (1985).
[40] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, Phys. Rev. Lett. **114**, 060501 (2015).
[41] A. P. Lund *et al*, Phys. Rev. Lett. **113**, 100502 (2014).
[42] P. P. Rohde, J. Opt. B: Quant. Semiclass. Opt. **7** 82 (2005).
[43] V. S. Shchesnovich, arXiv:1403.4459 [quant-ph].
[44] S. Rahimi-Keshari *et al*, Optics Express **21**, 13450 (2013).
[45] P. P. Rohde and T. C. Ralph, Phys. Rev. A **85**, 022332 (2012).
[46] S. Aaronson and D. J. Brod, arXiv:1510.05245 [quant-ph].
[47] In description of degrees of indistinguishability of N input photons, each in the same mixed internal state ρ, involves $N-1$ parameters, $\mathrm{Tr}(\rho^2), \ldots, \mathrm{Tr}(\rho^N)$ [22, 23], whereas a two-mode operator depends only on $\mathrm{Tr}(\rho^2)$.
[48] For a p.s.d. Hermitian H, as a principal submatrix of the tensor product matrix $H^{\otimes N}$ the Schur power matrix is a p.s.d. Hermitian as well.
[49] Numerical simulations are limited to $N \leq 7$ due to a large dimension of the Schir power matrix.

APPENDIX

Probability of N particles to gather in $\mathcal{K}$ output modes of a linear network

Let us first recall principal steps in derivation of the output probability distribution of N identical bosons/fermions at input of a linear network U (for bosons this result was derived in Refs. [22, 23]). We will consider simultaneously both species (in case of fermions an order creation and annihilation operators is assumed). A more general input is assumed, with $0 \leq n_k \leq N$ particles per input mode k (fermions have linearly independent internal states in each input mode). A general input state of configuration $\mathbf{n} = (n_1, \ldots, n_M)$ reads

$$\rho(\mathbf{n}) = \sum_i q_i |\Psi_i\rangle\langle\Psi_i|, \qquad (20)$$

with $q_i \geq 0$, $\sum_i q_i = 1$, and

$$|\Psi_i\rangle = \frac{1}{\sqrt{\mu(\mathbf{n})}} \sum_{\mathbf{j}} C^{(i)}_{\mathbf{j}} \prod_{\alpha=1}^{N} a^\dagger_{k_\alpha, j_\alpha} |0\rangle, \qquad (21)$$

where $\mathbf{j} = (j_1, \ldots, j_N)$ and $\mu(\mathbf{n}) = \prod_{k=1}^{M} n_k!$. Permutation symmetry (anti-symmetry) of creation operators for bosons (fermions) allows to chose expansion coefficients $C^{(i)}_{\mathbf{j}}$ symmetric (anti-symmetric) with respect to the Young subgroup $\mathcal{S}_{\mathbf{n}} \equiv \mathcal{S}_{n_1} \otimes \ldots \otimes \mathcal{S}_{n_M}$ of the symmetric group $\mathcal{S}_N$, where $\mathcal{S}_{n_k}$ corresponds to permutations of internal states of particles in input mode k between themselves. Such coefficients are normalized by $\sum_{\mathbf{j}} |C^{(i)}_{\mathbf{j}}|^2 = 1$.

The probability of an output configuration $\mathbf{m} = (m_1, \ldots, m_M)$ is given as [22, 23, 29]

$$\hat{p}(\mathbf{m}|\mathbf{n}) = \mathrm{Tr}(\rho(\mathbf{n})\mathcal{D}(\mathbf{m})), \qquad (22)$$

where ρ is the input state Eq. (1) and $\mathcal{D}(\mathbf{m})$ is the detection operator [23, 29] ($|0\rangle$ is Fock vacuum state)

$$\mathcal{D}(\mathbf{m}) = \frac{1}{\mu(\mathbf{m})} \sum_{\mathbf{j}} \left[\prod_{\alpha=1}^{N} b^\dagger_{l_\alpha, j_\alpha}\right] |0\rangle\langle 0| \left[\prod_{\alpha=1}^{N} b_{l_\alpha, j_\alpha}\right]. \qquad (23)$$

One can evaluate the trace in Eq. (22) by first expressing the input mode operators in Eq. (20) through the output ones using $a^\dagger_{k,j} = \sum_{l=1}^{M} U_{k,l} b^\dagger_{l,j}$ and then employing the following identity (see also Refs. [22, 23, 29])

$$\langle 0| \left[\prod_{\alpha=1}^{N} b_{l_\alpha, j_\alpha}\right] \left[\prod_{\alpha=1}^{N} b^\dagger_{l'_\alpha, j'_\alpha}\right] |0\rangle = \sum_{\sigma \in \mathcal{S}_N} \varepsilon(\sigma) \prod_{\alpha=1}^{N} \delta_{l'_\alpha, l_{\sigma(\alpha)}} \delta_{j'_\alpha, j_{\sigma(\alpha)}}. \qquad (24)$$

Substituting Eq. (20) and (23) into Eq. (22) and using Eq. (24) in the two inner products one obtains the probability of an output configuration $\mathbf{m}$ in a linear network U in the form

$$\hat{p}(\mathbf{m}|\mathbf{n}) = \frac{1}{\mu(\mathbf{m})\mu(\mathbf{n})} \sum_{\tau,\sigma} J(\tau\sigma^{-1}) \prod_{\alpha=1}^{N} U^*_{k_{\tau(\alpha)}, l_\alpha} U_{k_{\sigma(\alpha)}, l_\alpha}. \qquad (25)$$

where σ, τ are elements of the symmetric group $\mathcal{S}_N$, $l_1, \ldots, l_N$ are output modes, $1 \leq l_\alpha \leq M$, with multiplicities $(m_1, \ldots, m_M)$, whereas function $J(\sigma)$ and the internal state $\rho^{(int)}$ are given by Eqs. (4) and (2) of the main text.

Note that the permutation symmetry (anti-symmetry) of the input state (20)-(21) for bosons (fermions), i.e., $P_\pi \rho^{(int)} = \rho^{(int)} P_\pi = \varepsilon(\pi)\rho^{(int)}$ for any $\pi \in \mathcal{S}_{\mathbf{n}}$, implies that

$$J(\sigma\pi) = J(\pi\sigma) = J(\sigma), \quad \forall \pi \in \mathcal{S}_{\mathbf{n}}. \qquad (26)$$

The symmetry (26) means that the factorizing function of Eq. (5) of the main text can be chosen to satisfy

$$\theta(\sigma\pi) = \theta(\sigma), \quad \forall \pi \in \mathcal{S}_{\mathbf{n}}. \qquad (27)$$

(In contrast, $\theta(\tau\sigma)$ for all $\tau \in \mathcal{S}_N$ is the same factorization with a different order of terms).

A network input consists of *completely indistinguishable* bosons (fermions) if the corresponding J-function reads $J^{(id)}(\sigma) = \varepsilon(\sigma)$ [23, 29]. This case allows one to completely neglect the internal degrees of freedom. Probabilities at a network output are expressed through the usual matrix permanent and determinant, respectively (in case of fermions $\mu(\mathbf{n}) = 1$). The simplest case of completely indistinguishable particles consists of all particles being in the same internal state $|\phi\rangle$, giving $\rho^{(int)} = (|\phi\rangle\langle\phi|)^{\otimes N}$ and $J(\sigma) = \varepsilon(\sigma)$. The other limiting case, which may be identified as the *classical case*, since the output probabilities are the same as in the case of classical particles, corresponds to a "block-structured" J-function (see also Refs. [22, 23])

$$J^{(d)}(\sigma) = \sum_{\pi \in \mathcal{S}_{\mathbf{n}}} \delta_{\sigma,\pi}. \qquad (28)$$

Function $J(\sigma)$ of Eq. (28) appears when the internal states of identical particles from different input modes become orthogonal: $\mathrm{Tr}\{\rho^{(int)} P_\sigma\} = 0$ for $\sigma \notin \mathcal{S}_{\mathbf{n}}$, whereas $\varepsilon(\sigma)\mathrm{Tr}\{\rho^{(int)} P_\sigma\} = 1$ for $\sigma \in \mathcal{S}_{\mathbf{n}}$ (i.e., distinguishable particles from the same input mode cannot be discriminated by a linear network from the completely indistinguishable bosons). Note that the subgroup $\mathcal{S}_{\mathbf{n}}$ acts as identity on the indices $k_1, \ldots, k_N$ of matrix U in Eq. (25), thus the sum over $\mathcal{S}_{\mathbf{n}}$ in Eq. (28) cancels $\mu(\mathbf{n})$ in the denominator in Eq. (25), resulting in the familiar formula for the probability in the classical case, expressed through the matrix permanent of doubly stochastic matrix with elements $|U_{kl}|^2$. Obviously, for a single-particle input ($n_k \le 1$) we have $J^{(d)}(\sigma) = \delta_{\sigma,I}$.

The total probability of detecting all N input particles at a preselected (and fixed) set of $1 \le \mathcal{K} \le M$ output modes, say, the first $\mathcal{K}$ modes, is a sum of $\hat{p}(\mathbf{m}|\mathbf{n})$ (25) with $m_{\mathcal{K}+1} = \dots m_M = 0$. We have

$$
\begin{aligned}
p(J) &= \sum_{\mathbf{m}}{}' \hat{p}(\mathbf{m}|\mathbf{n}) = \frac{1}{N!}\sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \\
&\times \frac{1}{\mu(\mathbf{n})} \sum_{\tau,\sigma} J(\tau\sigma^{-1}) \prod_{\alpha=1}^{N} U^*_{k_{\tau(\alpha)},l_\alpha} U_{k_{\sigma(\alpha)},l_\alpha} \\
&= \frac{1}{\mu(\mathbf{n})} \sum_{\sigma'} J(\sigma') \prod_{\alpha=1}^{\mathcal{K}} H_{\alpha,\sigma'(\alpha)} = \frac{1}{\mu(\mathbf{n})} \sum_{\sigma\in\mathcal{S}_N} J(\sigma)\Pi_{I,\sigma} \\
&= \frac{1}{\mu(\mathbf{n})} \sum_{\tau,\tau'\in\mathcal{S}_N} \theta^*(\tau)\theta(\tau')\Pi_{\tau,\tau'}. \qquad (29)
\end{aligned}
$$

where we have transformed the sum over output configurations $\mathbf{m}$ into that over output mode indices $l_1,\dots,l_N$ with the combinatorial coefficient $\mu(\mathbf{m})/N!$, used the definition of matrix H (6), reordered the product $\prod_{\alpha=1}^{\mathcal{K}} H_{\sigma(\alpha),\tau(\alpha)} = \prod_{\alpha=1}^{\mathcal{K}} H_{\alpha,\tau\sigma^{-1}(\alpha)}$, and defined $\sigma' = \tau\sigma^{-1}$ (the sum over τ cancels the factor $1/N!$). Here the p.s.d. Hermitian matrix H and the Schur power matrix $\Pi(H)$ are given in Eq. (6) of the main text.

The limit cases

In case of the completely indistinguishable particles $\theta^{(id)}(\sigma) = \frac{\varepsilon(\sigma)}{\sqrt{N!}}$ (since $J^{(id)}(\sigma) = \varepsilon(\sigma)$). Eq. (29) gives for bosons and fermions:

$$p^{(id)}_B = \frac{\mathrm{per}(H)}{\mu(\mathbf{n})}, \quad p^{(id)}_F = \det(H)\delta_{\mu(\mathbf{n}),1}, \qquad (30)$$

where $\mathrm{per}(A) \equiv \sum_\sigma \prod_{\alpha=1}^N A_{\alpha,\sigma(\alpha)}$. For classical particles (e.g., distinguishable bosons or fermions from different input modes) we get from Eq. (28) that

$$\theta^{(d)}(\sigma) = \frac{1}{\sqrt{\mu(\mathbf{n})}} \sum_{\pi\in\mathcal{S}_{in}} \delta_{\sigma,\pi} \qquad (31)$$

which results in the expression of Eq. (9) of the main text.

The following order of the limit-case probabilities can be easily established

$$p^{(id)}_F \le p_D \le p^{(id)}_B, \qquad (32)$$

valid for *arbitrary* number of particles per input mode. Indeed, for a p.s.d. Hermitian matrix H, which we rearrange in a block-matrix form

$$H = \begin{pmatrix} H^{(1,1)} & H^{(1,2)} \\ H^{(2,1)} & H^{(2,2)} \end{pmatrix}. \qquad (33)$$

the following inequality is known for the matrix determinant [33]:

$$\det(H) \le \det(H^{(1,1)})\det(H^{(2,2)}). \qquad (34)$$

Similarly, for the matrix permanent [34]

$$\mathrm{per}(H) \ge \mathrm{per}(H^{(1,1)})\mathrm{per}(H^{(2,2)}). \qquad (35)$$

By repeated application of Eqs. (34) and (35) it is easy to demonstrate that Eq. (32) holds.

Factorization of J-function and its representation through a density matrix

To show that the probability $p^{(id)}_B$ ($p^{(id)}_F$) for the completely indistinguishable bosons (fermions) corresponds to the absolute maximum (minimum) over *arbitrary* input states of particles in a given configuration $\mathbf{n}$, one has to know to what class of functions the physical J-functions, i.e., describing an input of a linear network, belong. Let us show that any normalized by $J(I) = 1$ p.s.d. function $J(\sigma)$ can be represented in the form of Eq. (4) with some state $\rho^{(int)}$ (2). Since $\mathrm{sgn}(\sigma)J(\sigma)$ is also a normalized p.s.d function, it is sufficient to consider bosons, $\varepsilon(\sigma) = 1$.

Consider a linear subspace $\mathcal{L} \subset \mathcal{H}^{\otimes N}$ defined as the linear span of vectors $|\sigma\rangle \equiv P_\sigma|I\rangle$ for $\sigma \in \mathcal{S}_N$, where $|I\rangle \equiv |\phi_1\rangle \otimes \ldots \otimes |\phi_N\rangle$ for some arbitrary orthonormal vectors $|\phi_k\rangle \in \mathcal{H}$, $\langle\phi_k|\phi_l\rangle = \delta_{k,l}$ (note that $\langle\sigma|\tau\rangle = \delta_{\sigma,\tau}$ and $P_\pi = \sum_\tau |\pi\tau\rangle\langle\tau|$ when restricted to $\mathcal{L}$). Using that $P_\sigma|\tau\rangle = |\sigma\tau\rangle$ and $\langle\pi|P_\sigma|\pi\rangle = \delta_{\sigma,I}$, by starting from a trivial identity we get

$$
\begin{aligned}
J(\sigma) &= \frac{1}{N!}\sum_{\pi\in\mathcal{S}_N} \langle\pi|\Big[\sum_\tau J(\tau)P^\dagger_\tau\Big]P_\sigma|\pi\rangle \\
&= \mathrm{Tr}\left\{\rho^{(int)}P_\sigma\right\}, \qquad (36)
\end{aligned}
$$

where the trace is taken in $\mathcal{H}^{\otimes N}$ and we have introduced a p.s.d. Hermitian operator (density matrix) $\rho^{(int)}$ in the Hilbert space $\mathcal{H}^{\otimes N}$

$$\rho^{(int)} \equiv \frac{1}{N!}\sum_{\tau\in\mathcal{S}_N} J(\tau) \sum_{\pi\in\mathcal{S}_N} |\pi\rangle\langle\tau\pi|. \qquad (37)$$

Obviously, $\mathrm{Tr}\{\rho^{(int)}\} = J(I) = 1$. Positivity of $\rho^{(int)}$ follows from the explicit form (now for bosons and fermions)

$$
\begin{aligned}
\rho^{(int)} &= \frac{1}{N!}\sum_{\tau\in\mathcal{S}_N} |\Phi_\tau\rangle\langle\Phi_\tau|, \\
|\Phi_\tau\rangle &\equiv \sum_{\sigma\in\mathcal{S}_N} \xi(\sigma\tau)P_\sigma\{|\phi_1\rangle\otimes\ldots\otimes|\phi_N\rangle\}, \qquad (38)
\end{aligned}
$$

where $\xi(\sigma) \equiv \varepsilon(\sigma)\theta^*(\sigma)$ for the factorizing function $\theta(\sigma)$, see Eq. (5) of the main text.

To show factorization in Eq. (5) of the main text consider an operator $\mathcal{J}$ in $\mathcal{L}$ (note $\mathcal{J} = N!\rho^{(int)}$ of Eq. (37))

$$\mathcal{J} \equiv \sum_{\sigma\in\mathcal{S}_N} J(\sigma) P^\dagger_\sigma, \tag{39}$$

given by following matrix

$$\mathcal{J}_{\nu,\tau} = \langle \nu|\mathcal{J}|\tau\rangle = J(\nu^{-1}\tau), \tag{40}$$

which, by assumption that $J(\sigma)$ is a p.s.d. Hermitian, is a p.s.d. Hermitian matrix. Since any operator $\mathcal{A}$ in $\mathcal{L}$ can be represented as

$$\mathcal{A} = \sum_\sigma A(\sigma)P^\dagger_\sigma, \quad A \in \mathcal{S}_N \to \mathbb{C}, \tag{41}$$

and a product of operators in $\mathcal{L}$ belongs to $\mathcal{L}$, the operator $\mathcal{J}$ (39) is factorized by an operator $\mathcal{B} \in \mathcal{L}$

$$\mathcal{J} = \mathcal{B}^\dagger\mathcal{B}, \quad \mathcal{B} = \sum_\sigma \theta(\sigma)P^\dagger_\sigma. \tag{42}$$

By the group rule $P_\sigma P_\tau = P_{\sigma\tau}$ Eq. (42), in the matrix form, is equivalent to Eq. (5) of the main text.

Derivation of the average probability formulae

Eq. (13) of the main text follows from a simple symmetry argument for a Haar-random unitary U, but Eq. (14) requires a bit more of insight. Here these results are derived by direct evaluation which also demonstrates the validity of the classical formula for $M \gg 1$ (arbitrary $\mathcal{K}$ and N), as observed in numerical simulations. The following identity will be employed

$$\begin{aligned}&\langle \prod_{\alpha=1}^N U_{k_\alpha,l_\alpha} U^*_{k'_\alpha,l'_\alpha}\rangle \\ &= \sum_{\nu,\tau\in\mathcal{S}_N} \mathcal{W}(M,\nu\tau^{-1}) \prod_{\alpha=1}^N \delta_{k'_\alpha,k_{\nu(\alpha)}}\delta_{l'_\alpha,l_{\tau(\alpha)}},\end{aligned} \tag{43}$$

where $\mathcal{W}(M,\sigma)$ is the Weingarten function of the unitary group [1, 2] which depends only on the cycle structure of the relative permutation $\sigma = \nu\tau^{-1}$, i.e., the sequence of numbers $(c_1(\sigma), \ldots, c_N(\sigma))$ of cycles of lengths $(1, \ldots, N)$ in its cycle decomposition (see for more details Ref. [3]). From Eq. (29) we have for the average over Haar-random network

$$\langle p(J)\rangle = \frac{1}{\mu(\mathbf{n})} \sum_{\sigma\in\mathcal{S}_N} J(\sigma) \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} Q_{\mathbf{n},\mathbf{m}}(\sigma), \tag{44}$$

where by application of Eq. (43) (using that the input and output indices have multiplicities $\mathbf{n}$ and $\mathbf{m}$, respectively) we have

$$Q_{\mathbf{n},\mathbf{m}}(\sigma) \equiv \langle \prod_{\alpha=1}^N U_{k_\alpha,l_\alpha} U^*_{k_{\sigma(\alpha)},l_\alpha}\rangle = \sum_{\nu\in\mathcal{S}_\mathbf{n}} \sum_{\tau\in\mathcal{S}_\mathbf{m}} \mathcal{W}(M,\nu\sigma\tau), \tag{45}$$

with summation over permutation invariance subgroups $\mathcal{S}_\mathbf{n}$ and $\mathcal{S}_\mathbf{m}$ of the input and output indices, respectively (the output indices vary between 1 and $\mathcal{K}$). Let us now consider separately bosons, fermions and distinguishable particles for a general input $\mu(\mathbf{n}) \geq 1$.

In case of the completely indistinguishable bosons, $J(\sigma) = 1$, we obtain from Eqs. (44) and (45)

$$\begin{aligned}\langle p^{(id)}_B\rangle &= \frac{1}{\mu(\mathbf{n})} \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \sum_{\sigma\in\mathcal{S}_N} Q_{\mathbf{n},\mathbf{m}}(\sigma) \\ &= \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \mu(\mathbf{m}) \sum_{\sigma'\in\mathcal{S}_N} \mathcal{W}(M,\sigma') \\ &= \frac{(\mathcal{K}+N-1)!}{(\mathcal{K}-1)!} \sum_{\sigma\in\mathcal{S}_N} \mathcal{W}(M,\sigma),\end{aligned} \tag{46}$$

where we have denoted $\sigma' = \nu\sigma\tau$, used that $\sum_{\pi\in\mathcal{S}_\mathbf{n}} 1 = \mu(\mathbf{n})$, and converted the sum over output indices to that over their multiplicities $\mathbf{m}$ with the coefficient $N!/\mu(\mathbf{m})$ (this sum is the number of Fock states of N bosons in $\mathcal{K}$ modes). Observing that for $\mathcal{K} = M$ the probability must be equal to 1, we get the sum of $\mathcal{W}$-functions in Eq. (46)

$$\sum_{\sigma\in\mathcal{S}_N} \mathcal{W}(M,\sigma) = \frac{(M-1)!}{(M+N-1)!}. \tag{47}$$

Eqs. (46) and (47) result in the boson part of Eq. (13) of the main text.

In case of the completely indistinguishable fermions $J(\sigma) = \mathrm{sgn}(\sigma)$ and $\mu(\mathbf{n}) = \mu(\mathbf{m}) = 1$ (no two or more particles per mode). We have from Eqs. (44) and (45)

$$\begin{aligned}\langle p^{(id)}_F\rangle &= \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \sum_{\sigma\in\mathcal{S}_N} \mathrm{sgn}(\sigma)\mathcal{W}(M,\sigma) \\ &= \frac{\mathcal{K}!}{(\mathcal{K}-N)!} \sum_{\sigma\in\mathcal{S}_N} \mathrm{sgn}(\sigma)\mathcal{W}(M,\sigma),\end{aligned} \tag{48}$$

where now $l_\alpha \neq l_\beta$ for $\alpha \neq \beta$ and the first sum is the number of Fock states of N fermions over $\mathcal{K}$ modes with the coefficient $N!$. Setting $\mathcal{K} = M$ in Eq. (48) we obtain

$$\sum_{\sigma\in\mathcal{S}_N} \mathrm{sgn}(\sigma)\mathcal{W}(M,\sigma) = \frac{M!}{(M-N)!}. \tag{49}$$

Eqs. (48) and (49) result in the fermion part of Eq. (13) of the main text.

In case of distinguishable particles we have from Eqs.

(28), (44), and (45)

$$
\begin{aligned}
\langle p_D \rangle &= \frac{1}{\mu(\mathbf{n})} \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \sum_{\sigma \in \mathcal{S}_\mathbf{n}} Q_{\mathbf{n},\mathbf{m}}(\sigma) \\
&= \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \sum_{\nu \in \mathcal{S}_\mathbf{n}} \sum_{\tau \in \mathcal{S}_\mathbf{m}} \mathcal{W}(M, \nu\tau) \\
&= \sum_{\nu \in \mathcal{S}_\mathbf{n}} \sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \nu\sigma) \sum_{l_1=1}^{\mathcal{K}} \cdot \ldots \cdot \sum_{l_N=1}^{\mathcal{K}} \prod_{\alpha=1}^{N} \delta_{l_\alpha, l_{\sigma(\alpha)}} \\
&= \sum_{\nu \in \mathcal{S}_\mathbf{n}} \sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \nu\sigma) \mathcal{K}^{\#\sigma} \qquad (50)
\end{aligned}
$$

where we have set $\#\sigma \equiv c_1(\sigma) + \ldots + c_N(\sigma)$ (the total number of cycles in the cycle decomposition of σ) and used that $\sum_{\pi \in \mathcal{S}_\mathbf{n}} 1 = \mu(\mathbf{n})$ and $\prod_{\alpha=1}^{N} \delta_{l_\alpha, l_{\sigma(\alpha)}} = \sum_{\tau \in \mathcal{S}_\mathbf{m}} \delta_{\sigma,\tau}$. Eq. (50) must coincide with Eq. (46) for all particles in the same input mode, $\mu(\mathbf{n}) = N!$, since the limit of distinguishable particles Eq. (28) is obtained by making identical particles from different modes distinguishable, while particles from the same input mode behave as the completely indistinguishable bosons. On the other hand, for a single-particle input it has a form quite different from that of Eq. (46) for a similar input of the completely indistinguishable bosons. Such an expression must have quite cumbersome form, in general. Consider the special case of single-particle input. Using an asymptotic form of $\mathcal{W}$ [2] for $M \gg 1$

$$
\begin{aligned}
\mathcal{W}(M, \sigma) &= \frac{(-1)^N}{M^{2N}} \prod_{s=1}^{N} (-M g_s)^{c_s(\sigma)} \left(1 + O\left(\frac{1}{M^2}\right)\right) \\
g_s &= \frac{(2s-2)!}{s!(s-1)!}, \qquad (51)
\end{aligned}
$$

we obtain from Eq. (50) the leading term in the average classical probability as a cycle sum

$$
\langle p_D \rangle \approx \frac{(-1)^N}{M^{2N}} \sum_{\sigma \in \mathcal{S}_N} \prod_{s=1}^{N} (-\mathcal{K} M g_s)^{c_s(\sigma)} = \frac{(-1)^N N!}{M^{2N}} Z_N, \qquad (52)
$$

with ($t_s = -\mathcal{K} M g_s$)

$$
Z_N \equiv \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \prod_{s=1}^{N} t_s^{c_s(\sigma)}. \qquad (53)
$$

The cycle sum is evaluated by the generating function method (see, for instance, Ref. [3]) which satisfies the following identity

$$
\mathcal{F}(x) \equiv \sum_{N \geq 1} Z_N x^N = \exp\left\{ \sum_{s \geq 1} t_s \frac{x^s}{s} \right\}. \qquad (54)
$$

In our case (after evaluation of a table sum [4]) we get $\mathcal{F}(x)$ to be

$$
\begin{aligned}
\mathcal{F}(x) &= \exp\left\{ -\mathcal{K} M \sum_{s=1}^{\infty} \frac{(2s-2)!}{(s!)^2} x^s \right\} \qquad (55) \\
&= \left(\frac{2}{1 + \sqrt{1-4x}} \exp\left\{ -[1 - \sqrt{1-4x}] \right\} \right)^{\mathcal{K} M}.
\end{aligned}
$$

For $M \gg 1$, the leading order reads

$$
Z_N = \frac{1}{N!} \frac{d^N \mathcal{F}(0)}{dx^N} \approx \frac{(-\mathcal{K} M)^N}{N!}, \qquad (56)
$$

due to the fact that the derivative of $\mathcal{F}(x)$,

$$
\frac{d\mathcal{F}(x)}{dx} = \frac{-2\mathcal{K} M}{\sqrt{1-4x}(1 + \sqrt{1-4x})} \mathcal{F}(x), \qquad (57)
$$

has as a factor a slowly varying function of $|x| \ll 1$ (in comparison with $\mathcal{F}(x)$). Substituting Eq. (56) into Eq. (52) we obtain the probability for distinguishable particles given by Eq. (14) of the main text.

Estimate on the trace distance to ideal BS

We can expand the factorization $\theta(\sigma)$ of Eq. (5) of the main text over the eigenvectors of the Schur power matrix $\Pi(H)$. For bosons we get

$$
\theta(\sigma) = \sqrt{\frac{d(J)}{N!}} + \sum_j \sqrt{c_j(J)} \theta_j(\sigma), \qquad (58)
$$

where the eigenvector $\theta_B(\sigma) = 1/\sqrt{N!}$ corresponds to the completely indistinguishable particles and we have introduced the expansion coefficients $c_1, \ldots, c_{N!-1}$ satisfying $\sum_j |c_j| = 1 - d$. The probability in Eq. (7) becomes

$$
p(J) = d(J) p_B^{(id)} + \sum_j |c_j(J)| \lambda_j, \qquad (59)
$$

where λ_j, $j = 1, \ldots, N! - 1$, are eigenvalues of the Schur power matrix other than $p_B^{(id)} = \mathrm{per}(H)$, with $\lambda_1 = p_F^{(id)} = \det(H)$. Setting λ_{max} to be the maximum eigenvalue less than $\mathrm{per}(H)$, from Eq. (59) we obtain:

$$
\begin{aligned}
\mathrm{per}(H) - p(J) &\geq (1 - d(J)) \left[\mathrm{per}(H) - \lambda_{max} \right], \qquad (60) \\
\mathrm{per}(H) - p(J) &\leq (1 - d(J)) \left[\mathrm{per}(H) - \det(H) \right]. \qquad (61)
\end{aligned}
$$

These inequalities give

$$
\frac{\mathrm{per}(H) - p(J)}{\mathrm{per}(H) - \det(H)} \leq 1 - d(J) \leq \frac{\mathrm{per}(H) - p(J)}{\mathrm{per}(H) - \lambda_{max}}. \qquad (62)
$$

Equivalent description of a lossy linear network

Below we focus on bosonic particles, for fermions one should replace the commutators below by the anti-commutators. A realistic network U is non-unitary due to (generally path-dependent) losses of particles, its action is described not just by input and output mode operators, $a_1, \ldots, a_M$ and $b_1, \ldots, b_M$, but also by some additional operators $f_1, \ldots, f_M$ accounting for losses:

$$a_k^\dagger = \sum_{l=1}^{M} U_{k,l} b_l^\dagger + f_k^\dagger, \tag{63}$$

where operators f_k and $f_k^\dagger$ commute with the creation and annihilation operators corresponding to network modes: $[f_k, a_l] = [f_k, b_l] = [f_k^\dagger, a_l] = [f_k^\dagger, b_l] = 0$ [5]. Using the latter we obtain

$$[f_k, f_j] = 0, \quad [f_k, f_j^\dagger] = \delta_{k,j} - \sum_{l=1}^{M} U_{k,l}^* U_{j,l}. \tag{64}$$

Eq. (64) can be satisfied if we expand the loss operators $f_1^\dagger, \ldots, f_M^\dagger$ over some additional creation operators

$$f_k^\dagger = \sum_{l=1}^{M} V_{k,l} b_{M+l}^\dagger, \tag{65}$$

where V is a matrix satisfying the following matrix equation (valid both for bosons and fermions)

$$VV^\dagger = I - UU^\dagger, \tag{66}$$

where I is the unit matrix and $U^\dagger$ denotes the Hermitian conjugate to matrix U. Notice that Eq. (66) requires that the singular values of U be bounded by 1, which is the necessary and sufficient condition for an arbitrary complex matrix U to describe a passive linear quantum network. There is a polar decomposition, $U = \sqrt{A}\mathcal{U}$, where $\mathcal{U}$ is a unitary matrix and a Hermitian matrix $A = UU^\dagger$, with the eigenvalues bounded by 1, describing losses in the network.

The expansion in Eq. (65) means that one can imbed an arbitrary (non-unitary) linear M-mode network into a $2M$-mode unitary one [8]. The following embedding unitary network seems to be the simplest one

$$\hat{U} = \begin{pmatrix} U & V \\ -V^\dagger \mathcal{U} & D \end{pmatrix}, \tag{67}$$

here the diagonal matrix $D = \mathrm{diag}(\eta_1, \ldots, \eta_M)$, $0 \le \eta_k \le 1$, is composed of the square-roots of singular values of U (eigenvalues of $\sqrt{A}$, since $A = UU^\dagger$), $V = SQ$, with the unitary matrix S containing the eigenvectors of $\sqrt{A}$, $\sqrt{A} = SDS^\dagger$, $Q = \mathrm{diag}(\sqrt{1-\eta_1^2}, \ldots, \sqrt{1-\eta_M^2})$, and $\mathcal{U}$ is from the polar decomposition $U = \sqrt{A}\mathcal{U}$. Matrix $\hat{U}$ can be also rewritten in a product form

$$\hat{U} = \begin{pmatrix} S & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} D & Q \\ -Q & D \end{pmatrix} \begin{pmatrix} S^\dagger & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathcal{U} & 0 \\ 0 & I \end{pmatrix}, \tag{68}$$

from which it is evident that $\hat{U}$ is unitary. Note that in description of a lossy network U, the unitary matrix $\hat{U}$ has vacuum input in the modes $\{M+1, \ldots, 2M\}$ and output modes $\{M+1, \ldots, 2M\}$ are not accessible "loss channels". The embedding matrix of Eq. (67) reduces to a matrix appeared in Ref. [7] in the special case of path-independent losses, i.e., a diagonal loss matrix A.

[1] D. Weingarten, J. Math. Phys. **19**, 999 (1978).

[2] P. W. Brouwer and C. W. J. Beenakker, J. Math. Phys. **37**, 4904 (1996).

[3] R. P. Stanley, *Enumerative Combinatorics*, 2nd ed., Vol. 1 (Cambridge University Press, 2011).

[4] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integraly i Ryady. Elementarnue funktzii* (in Russian) (Moscow, FizMatLit 1981).

[5] S. M. Barnett, C. R. Gilson, B. Huttner, and N. Imoto, Phys. Rev. Lett. **77**, 1739 (1996).

[6] K. Fan and A.J. Hoffman, Proc. Amer. Math. Soc. **6**,111 (1955).

[7] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A.Fedrizzi, T. C. Ralph, and A. G. White, Opt. Express **21**, 13450 (2013).

[8] Moreover, analysis of necessary conditions on embedding shows that one cannot reduce the size of the embedding network for non-singular matrices U.