

Discriminating distinguishabilityStasja Stanisic^{1,2} and Peter S. Turner¹¹*Quantum Engineering Technology Laboratories, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol, United Kingdom*²*Quantum Engineering Centre for Doctoral Training, University of Bristol, Bristol, United Kingdom*

(Received 29 June 2018; published 19 October 2018)

Particle distinguishability is a significant challenge for quantum technologies, in particular for photonics, where the Hong-Ou-Mandel (HOM) effect clearly demonstrates it is detrimental to quantum interference. We take a representation theoretic approach in first quantization, separating particles' Hilbert spaces into degrees of freedom that we control and those we do not, yielding a quantum-information-inspired bipartite model where distinguishability can arise as correlation with an environment carried by the particles themselves. This makes it clear that the HOM experiment is an instance of a (mixed) state discrimination protocol, which can be generalized to interferometers that discriminate unambiguously between ideal indistinguishable states and interesting distinguishable states, leading to bounds on the success probability of an arbitrary HOM generalization for multiple particles and modes. After setting out the first quantized formalism in detail, we consider several scenarios and provide a combination of analytical and numerical results for up to nine photons in nine modes. Although the quantum Fourier transform features prominently, we see that it is suboptimal for discriminating completely distinguishable states.

DOI: [10.1103/PhysRevA.98.043839](https://doi.org/10.1103/PhysRevA.98.043839)**I. INTRODUCTION**

Interference lies at the heart of quantum mechanics, and thus it holds the promise of fundamental advantages over nonquantum technologies, with far-reaching ramifications in communication, metrology, simulation, and computation. The nemesis of quantum interference is distinguishability, with the Hong-Ou-Mandel (HOM) effect [1] being a prototypical example. Recent advances in scaling linear optics for universal quantum computation [2–4], and the race to demonstrate quantum computational supremacy via analog computations that sample the scattering amplitudes of multipartite states [5–13], highlight the need for a thorough understanding of distinguishability in multimode quantum interference [14–23].

Rather than the usual second quantized approach, we can gain insight by bringing quantum information concepts to bear in first quantization [24–27]. Distinguishability can then be modeled, for example, as entanglement between controlled and uncontrolled degrees of freedom of individual particles, with loss of interference being caused by the decoherence that results when the uncontrolled Hilbert space is marginalized. This can be formalized by observing that bosonic (and fermionic) Fock states of two (sets of) degrees of freedom can have natural Schmidt decompositions, corresponding to so called unitary-unitary duality in many-body physics [28].

An example of a pertinent idea from quantum information is state discrimination [29–31]; we start by showing how this reproduces the well-known HOM distinguishability test for two particles. In principle, the formalism accommodates any number of particles and modes, and we show how this generalizes for multimode quantum interference, taking a representation theoretic approach (Secs. II and III); this com-

plements a number of generalizations in the literature [32–40]. We set up the state discrimination problem in the linear optical framework, assuming we have access to passive transformations (networks of phase shifters and beam splitters) and projective measurements via photon number counting detectors (Sec. IV A). This restriction on the allowed measurements yields a highly nontrivial constraint on the mixed-state discrimination scenario—this problem is what we study here. In particular, the optimization problem that results is nonlinear, as is usually the case in multiphoton interferometry [41], necessitating numerical techniques described in Sec. IV C.

The results are as follows: In Sec. V A, we present two general upper bounds valid for any photon number when discriminating (i) a state with a single distinguishable photon from the completely indistinguishable state and (ii) the completely distinguishable from the completely indistinguishable state; in Sec. V B, we show why the HOM test is the only test of distinguishability for arbitrary states of two photons and demonstrate the generality of the formalism by considering three photons in two modes; in Sec. V C, we use a mix of analytical and numerical techniques to argue the optimality of a balanced three-mode network (tritter) as a discriminator for both completely distinguishable and singly distinguishable states; in Sec. V D 1, we look at discrimination of singly distinguishable states with higher photon numbers up to $N = 9$ and show that the quantum Fourier transform (QFT) saturates the bound, suggesting it is the optimal interferometer for all N ; and finally, in Sec. V D 2, we look at the discrimination of completely distinguishable states with higher photon numbers and give examples of the best-known interferometers up to $N = 8$, found by observing a pattern emerging from the optimizations. Most of these results are summarized in Table I. Although it is not surprising that the QFT features heavily,

the results show that it is not optimal for discriminating completely distinguishable states, thus motivating the search for optimal discriminating networks for other states of interest.

II. MOTIVATION

A. Hong-Ou-Mandel interference

We will use the HOM scenario as an example that sets out the main features of our distinguishability model and its relationship to state discrimination. Each HOM photon has two pertinent degrees of freedom: one is spatial, namely the interferometer arms, and the other is temporal, namely the arrival times. We are usually interested in the case where it is the spatial degree of freedom over which we have control (via interferometry) and so we call this the system degree of freedom. We interpret the temporal degree of freedom as a label—in general, this would include all the particles' degrees of freedom which we do not control. Since we assume complete control of the system (including the possibility of making the system states identical), it is the label that determines the particles' distinguishability, via correlations between the system and label degrees of freedom. In a real HOM experiment, we are interested in preparing situations with varying distinguishability, so we do in fact manipulate the temporal label degree of freedom as well, but for applications we usually think of the system-label correlations as having been determined by means beyond our control.

The HOM scenario has two spatial system modes which we will call “top” and “bottom” ($s = \uparrow, \downarrow$), and two photons, requiring two temporal label modes that we will call “early” and “late” ($l = \leftarrow, \rightarrow$). (Note that these symbols will need to be ordered—we have avoided the obvious choice of s and $l = 1, 2$ to reduce confusion with other indices in this section; in Sec. III A we will revert to integers for the general case.) Photon creators are written as \hat{a}_{sl}^\dagger [42], giving rise to Fock states, which we can write as arrays where rows correspond to system modes and columns to label modes. An example of a completely distinguishable two-photon state is

$$|\psi_d\rangle = \hat{a}_{\uparrow\leftarrow}^\dagger \hat{a}_{\downarrow\rightarrow}^\dagger |\text{vac}\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (1)$$

with an early photon in the top arm and a late one in the bottom, while

$$|\psi_i\rangle = \hat{a}_{\uparrow\leftarrow}^\dagger \hat{a}_{\downarrow\leftarrow}^\dagger |\text{vac}\rangle = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad (2)$$

corresponds to an indistinguishable state where both photons are early.

Ideally, an interferometer acts only upon the system, corresponding to a unitary transformation on the two spatial modes

$$\hat{a}_{sl}^\dagger \mapsto \sum_t \hat{a}_{tl}^\dagger U_{ts}. \quad (3)$$

Here U is a 2×2 unitary matrix corresponding to the two-port interferometer, sometimes called the transfer matrix. We assume that the interferometer acts trivially upon the label modes (the photons remain early or late), corresponding to the 2×2 identity transfer matrix \mathbb{I} . For a suitable choice of

ordering of the four possible creators, the full 4×4 transfer matrix acting on all four modes ($\uparrow\leftarrow, \uparrow\rightarrow, \downarrow\leftarrow, \downarrow\rightarrow$) is given by

$$U \otimes \mathbb{I}. \quad (4)$$

It is tempting to interpret the tensor product in Eq. (4) as that between the system and the label. A quantum information theoretic approach to distinguishability would then ignore (trace out) the label, arriving at reduced states on the system where all the nontrivial transformations and measurements occur. However, this matrix acts on the space of operators, not on the state space which is a tensor product of four system-label harmonic oscillators in the second quantized model. In order to trace out the label, we will use a first quantized description.

Second quantized Fock states can be related to first quantized single-particle states as follows. Viewing each excitation of the four-mode aggregate as a particle with four available states ($\uparrow\leftarrow, \uparrow\rightarrow, \downarrow\leftarrow, \downarrow\rightarrow$), and recognizing that as bosons the total state must be symmetric under particle exchange, we have a one-to-one relationship between the Fock states of two bosons in four modes and symmetric states of two four-dimensional particles (qudits, here with $d = 4$). Applying this procedure to the indistinguishable state of Eq. (2), we have

$$|\psi_i\rangle = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad (5)$$

$$= \text{Sym}(|\uparrow\leftarrow\rangle_1 |\downarrow\leftarrow\rangle_2) \quad (6)$$

$$= \frac{1}{\sqrt{2}} (|\uparrow\leftarrow\rangle_1 |\downarrow\leftarrow\rangle_2 + |\downarrow\leftarrow\rangle_1 |\uparrow\leftarrow\rangle_2) \quad (7)$$

$$= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle_s + |\downarrow\uparrow\rangle_s) |\leftarrow\leftarrow\rangle_L, \quad (8)$$

where the subscripts 1 and 2 have been used as (fictitious) particle labels that get permuted, and we have rearranged the tensor product structure in the last line to arrive at a state in the S(ystem) \otimes L(abel) basis. Similarly, one finds

$$|\psi_d\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (9)$$

$$= \text{Sym}(|\uparrow\leftarrow\rangle_1 |\downarrow\rightarrow\rangle_2) \quad (10)$$

$$= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle_s |\leftarrow\rightarrow\rangle_L + \frac{1}{\sqrt{2}} |\downarrow\uparrow\rangle_s |\rightarrow\leftarrow\rangle_L). \quad (11)$$

We see that Eq. (8) is in a product state (Schmidt rank 1) of system and label [43], so the label states are uncorrelated to the system states; learning the label does not allow one to learn anything about the system, as expected for indistinguishable particles. Equation (11) is entangled (Schmidt rank 2), with the system states perfectly correlated to the labels (\uparrow to \leftarrow and \downarrow to \rightarrow), making the photons completely distinguishable.

It will be useful to rewrite states of both the system and label according to their permutation symmetry. Schur-Weyl duality [28,44] ensures that this basis can be chosen to also have good quantum numbers for the unitary group action of the interferometer, in this case $U(2)$ [45]. The irreducible representations (irreps) of $U(2)$ are well known, and for only two particles Young diagrams provide a compact notation for the basis states that carry these irreps; they are (for arbitrary,

ordered single-particle quantum numbers $x < y$) the symmetric triplet

$$|\boxed{x}\boxed{x}\rangle = |xx\rangle, \quad (12)$$

$$\sqrt{2}|\boxed{x}\boxed{y}\rangle = |xy\rangle + |yx\rangle, \quad (13)$$

$$|\boxed{y}\boxed{y}\rangle = |yy\rangle, \quad (14)$$

and the antisymmetric singlet

$$\sqrt{2}|\boxed{x}\boxed{y}\rangle = |xy\rangle - |yx\rangle. \quad (15)$$

We can now rewrite Eqs. (8) and (11) as

$$|\psi_i\rangle = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = |\boxed{\uparrow\downarrow}\boxed{s}\boxed{\leftarrow\leftarrow}\rangle_L, \quad (16)$$

$$|\psi_d\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}}|\boxed{\uparrow\downarrow}\boxed{s}\boxed{\leftarrow\rightarrow}\rangle_L + \frac{1}{\sqrt{2}}|\boxed{\uparrow\downarrow}\boxed{s}\boxed{\rightarrow\leftarrow}\rangle_L. \quad (17)$$

Note that total exchange symmetry is preserved because the system and label states in the second term of Eq. (17) are both antisymmetric. We can now see clearly that in this case the Schur-Weyl bases provide a Schmidt decomposition of the Fock arrays and that the completely distinguishable state has nonzero amplitude outside the totally symmetric irrep; we will discuss the generalization of these features in Sec. III B.

Tracing out the label degree of freedom, we arrive at the reduced density matrices that describe the state of the system. Another feature of the Schur-Weyl basis is that these states will be block diagonal, each block corresponding to an irrep. Thus, ordering our triplet-singlet basis as $\{|\boxed{\uparrow\uparrow}\rangle, |\boxed{\uparrow\downarrow}\rangle, |\boxed{\downarrow\downarrow}\rangle, |\boxed{\uparrow}\rangle\}$, we have

$$\rho_i = \text{Tr}_L[|\psi_i\rangle\langle\psi_i|] = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 0 \end{bmatrix}, \quad (18)$$

$$\rho_d = \text{Tr}_L[|\psi_d\rangle\langle\psi_d|] = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 1 \end{bmatrix}. \quad (19)$$

A coincidence count occurs when both the top and bottom modes are occupied, defining the coincidence subspace spanned by $\{|\boxed{\uparrow\downarrow}\rangle, |\boxed{\downarrow\uparrow}\rangle\}$. The projector onto this subspace has matrix representation

$$M_{(1,1)} = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & 1 \end{bmatrix}, \quad (20)$$

where we have used an occupation (one excitation in each of the two system modes) in the subscript.

The unitary evolution of these input states due to the interferometer is given by the two-photon representation of the transfer matrix. Again, in the Schur-Weyl basis this is block diagonal, specifically a direct sum of the triplet and singlet matrix representations of $U(2)$. The matrix elements in the coincidence subspace for an arbitrary two-mode

interferometer with transfer matrix U are

$$U^{\otimes 2} \cong U^{\square} \oplus U^{\square} = \begin{bmatrix} * & * & * & \\ * & \text{per}U & * & \\ * & * & * & \\ & & & \det U \end{bmatrix}, \quad (21)$$

where per and det are the matrix permanent and determinant functions, * are matrix elements for events outside the coincidence subspace, and we use \cong to denote the fact that $U \otimes U$ only equals $U^{\square} \oplus U^{\square}$ after the basis change. This can be confirmed by direct calculation from Eq. (3), or equivalently by using the Schur-Weyl transformation, which for $U(2)$ is the familiar Clebsch-Gordan transformation of angular momentum theory.

The probability of a coincidence count is given by the Born rule, which from Eqs. (18)–(21) is given by

$$P_{(1,1)} = \text{Tr}[(U^{\square} \oplus U^{\square})\rho(U^{\square} \oplus U^{\square})^{\dagger}M_{(1,1)}] \quad (22)$$

$$= \text{Tr}[(U^{\square}\rho U^{\square\dagger} + U^{\square}\rho U^{\square\dagger})M_{(1,1)}] \quad (23)$$

$$= \begin{cases} |\text{per}U|^2 & \text{if } \rho = \rho_i \\ \frac{1}{2}|\text{per}U|^2 + \frac{1}{2}|\det U|^2 = \text{per}|U|^2 & \text{if } \rho = \rho_d \end{cases} \quad (24)$$

where we have written $|U|^2$ for the elementwise absolute value squared of a matrix U .

It follows that in order to see no coincidences for an indistinguishable state, which has only a triplet component, we need an interferometer whose transfer matrix permanent vanishes. By parametrizing an arbitrary $U \in U(2)$, one can confirm that only a balanced beam splitter has this property (see Sec. V B). We also see that the distinguishable state has a singlet component that scatters through any U according to the determinant, and since any element of $U(2)$ has $|\det U| = 1$, this component will always give rise to coincidences. Thus, in a HOM experiment one uses a balanced beam splitter to see a “dip” in coincidence counts in the system as one manipulates the label degree of freedom from distinguishable to indistinguishable and back again.

B. State discrimination

By choosing to measure a coincidence count as well as U to be a balanced beam splitter, the HOM situation described above ensures that $P_{(1,1)} = 0$ when the input is ρ_i , while $P_{(1,1)}$ happens to be maximized when the input is ρ_d (see Sec. V B). This is reminiscent of what is known as unambiguous mixed-state discrimination [46].

A general state discrimination protocol [30,31] consists of two parties, a source (Alice) and a detector (Bob), who agree on an ensemble of states $\{p_k, \rho_k\}$ to be discriminated. The source draws a random sample from this ensemble according to the distribution $\{p_k\}$ and sends it to the detector, whose task is to identify which state was sent as best as possible. This is accomplished by finding a measurement, given by a set of POVM elements $\{E_k\}$ that maximize the expected probability of success: $\sum_k p_k \text{Tr}[\rho_k E_k]$, known as minimum error discrimination. For unambiguous discrimination (UD), we have the further constraint that no mistakes are allowed to be made, that is, $\text{Tr}[\rho_k E_j] = 0$ for all $k \neq j$, at the price of

having to add an outcome E_7 to the POVM that corresponds to failing to identify the state.

By rearranging Eq. (22) and defining

$$M_{(1,1)}(U) = (U^{\square} \oplus U^{\square})^\dagger M_{(1,1)}(U^{\square} \oplus U^{\square}), \quad (25)$$

the HOM measurement scenario described above can now be summarized by

$$\text{find } U \text{ maximizing } \text{Tr}[\rho_d M_{(1,1)}(U)] \quad (26)$$

$$\text{subject to } \text{Tr}[\rho_i M_{(1,1)}(U)] = 0. \quad (27)$$

That is, find an interferometer that maximizes the probability of seeing a coincidence for a distinguishable input state, subject to the constraint that it never gives coincidences for an indistinguishable input state. It is now clear this is an instance of an UD problem, with the solution being a balanced beam splitter in the HOM case.

This gives a direction in which to generalize the HOM scenario to any number of particles in any number of modes as a UD problem. A key distinction from general UD is the restricted form of the available POVM elements, which must be projective measurements defined by the interferometer U and the N -photon occupation \underline{n} being detected. In particular, we expect that known optimal measurements for two-state discrimination will not be available in linear optics. When speaking generally about measurements, we will use the notation E for POVM elements, while, as above, $M_{\underline{n}}(U)$ is reserved for photon counts. Because $M_{\underline{n}}(U)$ is a polynomial of degree N in the variables U and U^\dagger , this measurement restriction makes the UD optimization problem nonlinear.

III. BACKGROUND: MANY PARTICLES AND MODES

From the HOM example [e.g., Eq. (17)], we see that the symmetry of the system and label states and the correlations between them play a key role in the distinguishability of the particles. We therefore proceed with an analysis for any number of particles and modes using Schur-Weyl duality in Sec. III A; this applies equally well to any degree of freedom, in particular it applies to the system and label independently. In Sec. III B, we then consider bosonic particles with two degrees of freedom specifically using unitary-unitary duality [28].

A. Schur-Weyl duality in first quantization

In the first quantized picture of the HOM example above, each photon was considered as a d -dimensional quantum system, with d the number of modes available. Schur-Weyl duality states that the Hilbert space of N qudits can be decomposed as

$$(\mathbb{C}^d)^{\otimes N} \cong \bigoplus_{\lambda} \mathbb{C}^{[\lambda]} \otimes \mathbb{C}^{(\lambda)}, \quad (28)$$

where $\mathbb{C}^{[\lambda]}$ carries irrep λ of the group of unitary transformations on a qudit, $U(d)$, $\mathbb{C}^{(\lambda)}$ carries irrep λ of the group of permutations of qudits, S_N , and \cong signifies that the left- and right-hand sides are related by a change of basis (a Schur-Weyl transform). Following Ref. [44], a Schur-Weyl

basis which realizes this decomposition is denoted $|\lambda qp\rangle$ where λ labels the irrep of both the unitary and the symmetric groups simultaneously [47], $q = 1, 2, \dots, d_{[\lambda]}$ indexes a basis of the unitary irrep, and $p = 1, 2, \dots, d_{(\lambda)}$ indexes a basis of the symmetric irrep. These dimensions can be computed, for example, by the Weyl character and hook length formulas respectively [48]. There is an implied dependence of q and p on λ , the set of which in turn depends on the number of particles N and the number of modes d . The irrep $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)$ can be specified using Young diagrams, where λ_j is the number of boxes in row j , $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, and $\sum_j \lambda_j = N$. The indices q and p correspond to the different ways of filling boxes with the numbers $\{1, \dots, d\}$ and $\{1, \dots, N\}$ to make semistandard (with repetition) and standard (without repetition) Young tableaux respectively, where numbers cannot decrease as you move right in a tableau and must increase as you go down.

We can further refine this notation by observing that the basis can be chosen such that the representation theoretic *weight* of a state corresponds to the occupation \underline{n} , which has also been called a *type* in this context [44]. Subspaces of states with the same occupation are then invariant under the Schur-Weyl transform in this basis, and the unitary index q can be uniquely specified by an occupation \underline{n} and an inner multiplicity r (the number of which is also known as a Kostka number), which accounts for the fact that there can be more than one orthogonal state with the same weight in a unitary irrep λ . As we are focusing on the action of the unitary group, p will be referred to as an outer multiplicity, accounting for the fact that the same unitary irrep λ can occur more than once. We can therefore write Schur-Weyl basis states in the form $|\lambda p \underline{n} r\rangle$, where the irrep dependence of p , \underline{n} , and r has again been suppressed to prevent clutter. We will often shorten the notation such that $|\lambda p \underline{n}\rangle := |\lambda, p, \underline{n}, r = 1\rangle$, $|\lambda \underline{n} r\rangle := |\lambda, p = 1, \underline{n}, r\rangle$, and $|\lambda \underline{n}\rangle := |\lambda, p = 1, \underline{n}, r = 1\rangle$, reducing clutter when the multiplicity is trivial; since λ and \underline{n} are vectors while p and r are scalars there should be no ambiguity. Coincident input or output will be denoted with occupation number $\underline{1} = (1, 1, \dots, 1)$, with exactly one particle in each mode (corresponding in first quantization to one qudit in each basis state).

For small N and d , writing states in terms of Young tableaux can be more compact, as in the HOM discussion of the previous section. The shape of a tableau is specified by λ , which is filled with mode indices specified by \underline{n} , following the rules for semistandard tableaux. The inner multiplicity r corresponds to different semistandard fillings of the same λ and \underline{n} , while the outer multiplicity p (corresponding to different standard fillings with fictitious particle indices) will be labeled with a subscript. For example, for three photons ($N = 3$) in three modes ($d = 3$), the coincident $\underline{n} = \underline{1} = (1, 1, 1)$ subspace for irrep $\lambda = (2, 1) = \square$ is spanned by four states given by $p, r \in \{1, 2\}$. If we index the modes 1, 2, and 3, the two notations are related as

$$|\lambda = (2, 1), p = 1, \underline{n} = (1, 1, 1), r = 1\rangle = \left| \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 1 \\ \hline \end{array} \right\rangle_1 \quad (29)$$

$$|\lambda = (2, 1), p = 1, \underline{n} = (1, 1, 1), r = 2\rangle = \left| \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & 1 \\ \hline \end{array} \right\rangle_1 \quad (30)$$

$$|\lambda = (2, 1), p = 2, \underline{n} = (1, 1, 1), r = 1\rangle = \left[\begin{array}{c|c} 1 & 2 \\ \hline 3 & 2 \end{array} \right]_2 \quad (31)$$

$$|\lambda = (2, 1), p = 2, \underline{n} = (1, 1, 1), r = 2\rangle = \left[\begin{array}{c|c} 1 & 3 \\ \hline 2 & 2 \end{array} \right]_2, \quad (32)$$

while, e.g., the $\underline{n} = (2, 1, 0)$ subspace for irrep $\lambda = (2, 1)$ is spanned by the two states

$$|\lambda = (2, 1), p = 1, \underline{n} = (2, 1, 0), r = 1\rangle = \left[\begin{array}{c|c} 1 & 1 \\ \hline 2 & 1 \end{array} \right]_1 \quad (33)$$

$$|\lambda = (2, 1), p = 2, \underline{n} = (2, 1, 0), r = 1\rangle = \left[\begin{array}{c|c} 1 & 1 \\ \hline 2 & 2 \end{array} \right]_2, \quad (34)$$

because the Young tableau $\left[\begin{array}{c|c} 1 & 2 \\ \hline 1 & 1 \end{array} \right]$ is not semistandard and therefore such states do not exist.

1. Implementation of the Schur-Weyl transform

An example of a Schur-Weyl transformation is the triplet-singlet basis change given in Eqs. (12)–(15), where (when $d = 2$) it is the same as the well-known Clebsch-Gordan transformation. There are several ways to implement this basis change more generally [49,50]; we use the method described in Ref. [28], which we will briefly outline here.

Every irrep $\{\lambda\}$ of $U(d)$ can be assigned a highest weight state, which is annihilated by an appropriate set of raising operators that are realized in terms of the bosonic creators and annihilators. Given as a Young tableau, this state can be expressed in terms of single-particle (qudit) states using Slater determinants, with the single-particle basis being indexed from 1 to d . In much the same way as is done for $U(2)$ in angular momentum theory, we then use corresponding lowering operators to find a set of states that span the irrep. The size of this set is known, namely $d_{(\lambda)}$. A Gram-Schmidt procedure is then used to orthonormalize the set (note that there is freedom in choosing how to do so when there are multiplicities; see, e.g., Sec. III C 2). Outer multiplicities are handled by utilizing the dual S_N action to permute a highest weight state in order to find corresponding highest weights for the multiple copies of irrep $\{\lambda\}$. Again, the number of such linearly independent highest weight states is known, namely $d_{(\lambda)}$, and orthonormalization is required. The lowering procedure is then repeated until a complete set of λ states are found. Iterating through all λ then gives a complete set of states $\{|\lambda qp\rangle\}$, from which we can determine the required basis transformation. Transformations for different N and d can be computed once and stored for later use.

B. Unitary-unitary duality

In the HOM example, we saw that each photon had two degrees of freedom, the system and the label, and that, as bosons, first quantized multiphoton states had to be totally symmetric under particle permutations. Independently decomposing both the system and label Hilbert spaces according to Schur-Weyl, one is then led to ask what states of the form

$$\sum_{\substack{\lambda qp \\ \lambda' q' p'}} \psi_{\lambda' q' p'}^{\lambda qp} |\lambda qp\rangle_S |\lambda' q' p'\rangle_L \quad (35)$$

are totally symmetric? This can be viewed as a coupling problem for irreps of the symmetric group—we wish to construct composite states of permutational momentum zero. The answer turns out much like it does in angular momentum theory: that λ, p must equal λ', p' , respectively, and that the coupling coefficients are all equal and independent of p [25,51]. Thus, totally symmetric pure system-label states are of the form

$$\sum_{\lambda qp} \psi_{\lambda qp} |\lambda qp\rangle_{SL}, \quad (36)$$

where we have defined

$$|\lambda qp\rangle_{SL} := \frac{1}{\sqrt{d_{(\lambda)}}} \sum_{p=1}^{d_{(\lambda)}} |\lambda qp\rangle_S |\lambda q' p\rangle_L. \quad (37)$$

These states carry the symmetric irrep of the global unitary group, $U(d_S d_L)$, acting on the $d_S d_L$ modes of the combined system and label. As discussed above, we can replace q with pairs \underline{n}, r in all of these expressions.

Equations (36) and (37) imply a decomposition of the totally symmetric irrep of $U(d_S d_L)$ into irreps of its unitary subgroups $U(d_S)$ and $U(d_L)$ that act on the system and label independently. These irreps are labeled simultaneously by λ , hence “unitary-unitary duality”:

$$\text{Sym}((\mathbb{C}^{d_S} \otimes \mathbb{C}^{d_L})^{\otimes N}) \cong \bigoplus_{\lambda} \mathbb{C}^{(\lambda)_S} \otimes \mathbb{C}^{(\lambda)_L}, \quad (38)$$

where we include subscripts on the right-hand side to remind us which unitary subgroups the irreps belong to Ref. [52]. An interferometer U is given by an element of the system unitary subgroup $U(d_S)$, and thus it acts on states in irrep λ according to the irreducible matrix representation U^λ ,

$$U : |\lambda qp\rangle_{SL} \mapsto \sum_{q''} |\lambda q'' q'\rangle_{SL} U_{q'' q}^\lambda. \quad (39)$$

Just as with a single degree of freedom, the space of second quantized $d_S \times d_L$ Fock arrays can be put into one-to-one correspondence with first quantized totally symmetric states by the procedure exemplified in Eqs. (5)–(8). Thus, we can write an arbitrary partially distinguishable state, which is an element of the totally symmetric subspace of $(\mathbb{C}^{d_S} \otimes \mathbb{C}^{d_L})^{\otimes N}$, in a basis of first quantized states given by Eq. (37). We may now trace out the label to arrive at mixed states describing any partially distinguishable state of N photons in d_S modes. We can order the basis so that the reduced system state and the action of any system interferometer will both be block diagonal according to irreps λ , a potentially significant simplification.

C. States of interest

We will focus our attention on three types of N -photon states: completely indistinguishable, singly distinguishable, and completely distinguishable, described below (the general case will be discussed in Sec. VI). We are not considering loss (where entire qudits would be traced out), so N will be fixed throughout. Situations with mixed system-label states and/or partial distinguishability can be written in terms of the basis of Eq. (36) [53], or equivalently that of Fock arrays. We give

examples of this generality with partial distinguishability for two photons in two modes in Sec. VB 1 and of mixed system-label states for three photons in three modes in Sec. VC 2. Otherwise, we will restrict ourselves to the case where the total system-label state is pure, corresponding to a source that produces states that are always (in)distinguishable in exactly the same way; generalization is, in principle, straightforward.

In practice, the space of label states available to a photon is as large as that of the uncontrolled degrees of freedom; in general, it is infinite dimensional many times over. However, in any given N particle experiment with N fixed, in order to model the distinguishability we need only consider the subspace spanned by the label states, which can be at most N dimensional. In other words, the most distinguishable N photons can be is for each of the N labels to be in an orthogonal state, and so we consider only $d_L \leq N$.

In order to set d_S , consider first two photons that are in the same system mode. In first quantization, this means each photon is in the same system state, implying the two-photon system state is symmetric, and so in order to maintain total symmetry—or by unitary-unitary duality—the state of the label must also be symmetric; cf. Eq. (8). This means that states in the antisymmetric irrep are not available and so restricts the combined state to a subspace of those allowed in Eq. (36). This argument extends to any number of photons, where the situation corresponds to any Fock array that is not full rank N . Thus in order to consider the full set of states available and arbitrary distinguishability, we must have input states that have a single photon in each system mode, i.e., $d_S \geq N$. Unless indicated otherwise, we will consider the case with $d_S = d_L = N$; following common terminology, we will say the photons are (system) coincident, and we will call the space of such system-label states the coincident subspace. The reader may wish to refer ahead to Sec. VC for concrete examples of the following.

1. Completely indistinguishable states

A completely indistinguishable state is one in which every photon’s label state is the same. As mentioned above, such a state lies in the symmetric label subspace with $\lambda = (N)$. Since the symmetric irrep of S_N is one dimensional, $d_{(N)} = 1$, and Schur-Weyl duality tells us that the corresponding unitary irrep is always outer multiplicity free. Moreover, (N) is also inner multiplicity free (there is only one way to symmetrize a product of single-particle states), so we can replace q with the system occupation $\underline{1}$ and q' with the label occupation $(N, \underline{0})$ (ordering our label modes such that the occupied one is first and with the understanding that the list of zeros is as long as it needs to be, in this case $N - 1$). The total state in Eq. (37) therefore becomes

$$\hat{a}_{11}^\dagger \hat{a}_{21}^\dagger \cdots \hat{a}_{N1}^\dagger |\text{vac}\rangle = \left| \begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 1 & 0 & \cdots & 0 \end{array} \right\rangle \quad (40)$$

$$= \text{Sym}(|11\rangle|21\rangle \cdots |N1\rangle) \quad (41)$$

$$= |(N), \underline{1}\rangle_S |(N), (N, \underline{0})\rangle_L, \quad (42)$$

where we have included $N - 1$ redundant zero columns in the Fock array so we can easily compare with the other states in this section. In the second line, we have written the state in the single-particle basis, cf. Eqs. (6) and (10), and we have suppressed trivial multiplicities in the last line. We see that this is always a product state, with no correlation between the system and label, as expected for completely indistinguishable particles. The reduced system state is

$$\rho_i = \text{Tr}_L[|(N), \underline{1}\rangle|(N), (N, \underline{0})\rangle\langle(N), \underline{1}|\langle(N), (N, \underline{0})|] \quad (43)$$

$$= |(N), \underline{1}\rangle\langle(N), \underline{1}|, \quad (44)$$

supported on the one-dimensional intersection of the symmetric system subspace given by (N) with the coincident subspace defined by the system occupation number $\underline{1}$.

2. Singly distinguishable states

The next state we consider is one where a single photon has become distinguishable from the rest; assuming all efforts are being made to produce the completely indistinguishable state, this should be the most likely error to occur. Ordering our modes so that the “bad” photon is in system mode N and label mode 2, we have

$$\hat{a}_{11}^\dagger \hat{a}_{21}^\dagger \cdots \hat{a}_{N2}^\dagger |\text{vac}\rangle = \left| \begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 1 & \cdots & 0 \end{array} \right\rangle \quad (45)$$

$$= \text{Sym}(|11\rangle|21\rangle \cdots |N2\rangle), \quad (46)$$

where in the last line we have not yet performed the Schur-Weyl transform. Considering this symmetrization, one observes that although all $N!$ permutations of the N distinct system indices will occur, since only two distinct label modes are involved, there are only N single-particle states available to the label degree of freedom, namely those with the j th photon in label mode 2 and the rest in label mode 1; denote these states $|2_j\rangle_L$. Such a label state will be perfectly correlated to all system states with the j th photon in mode N ; for each j we can factor these $(N - 1)!$ system states off and denote the resulting normalized state $|N_j\rangle_S$. Thus, in the system-label basis, the singly distinguishable state can be written as

$$\text{Sym}(|11\rangle|21\rangle \cdots |N2\rangle) = \frac{1}{\sqrt{N}} \sum_{j=1}^N |N_j\rangle_S |2_j\rangle_L, \quad (47)$$

e.g., Eq. (11). These sets of states are orthonormal, and we recognize this as an entangled state with Schmidt coefficients $1/\sqrt{N}$.

Now consider Schur-Weyl transforming this state into the form of Eq. (36). Because there are only two distinct label modes involved, the only label irreps that can occur are those whose Young diagrams have two or fewer rows. Moreover, because only a single photon is “bad,” the only two-rowed diagram allowed is that with a single box in the second row. Thus, the label state is supported only by irreps $\lambda = (N)$ and $(N - 1, 1)$. By unitary-unitary duality, the system is therefore also supported only on these two irreps. The

totally symmetric irrep (N) is always both inner and outer multiplicity free; for irrep $(N - 1, 1)$, the outer multiplicity is $d_{((N-1,1))} = N - 1$. It remains only to work out the inner multiplicities for irrep $(N - 1, 1)$. The system and label have occupations $\underline{1}$ and $(N - 1, 1, \underline{0})$ respectively (the marginals of the Fock array). There is only one Young tableau of shape $(N - 1, 1)$ consistent with occupation $(N - 1, 1, \underline{0})$ (that with the 2 in the second row box), so the label states are inner multiplicity free. The system occupation $\underline{1}$ is consistent with $N - 1$ Young tableau of shape $(N - 1, 1)$ (all those without a 1 in the second row box), and so the system inner multiplicity is $N - 1$. When we insert these observations into Eq. (36), the Schur-Weyl transformed state is

$$\begin{aligned} &\psi_{(N),\underline{1},1,(N-1,1,\underline{0}),1}|(N), 1, \underline{1}, 1\rangle_S |(N), 1, (N - 1, 1, \underline{0}), 1\rangle_L \\ &+ \sum_{r=1}^{N-1} \frac{\psi_{(N-1,1),\underline{1},r,(N-1,1,\underline{0}),1}}{\sqrt{N-1}} \sum_{p=1}^{N-1} |(N - 1, 1), p, \underline{1}, r\rangle_S \\ &\times |(N - 1, 1), p, (N - 1, 1, \underline{0}), 1\rangle_L. \end{aligned} \quad (48)$$

We can factor the second term and redefine coefficients to yield another Schmidt decomposition:

$$\begin{aligned} &\psi_{(N)}|(N), 1, \underline{1}, 1\rangle_S |(N), 1, (N - 1, 1, \underline{0}), 1\rangle_L \\ &+ \frac{\psi_{(N-1,1)}}{\sqrt{N-1}} \sum_{p=1}^{N-1} \left(\sum_{r=1}^{N-1} \phi_r |(N - 1, 1), p, \underline{1}, r\rangle_S \right) \\ &\times |(N - 1, 1), p, (N - 1, 1, \underline{0}), 1\rangle_L. \end{aligned} \quad (49)$$

Because the Schur-Weyl transformations yielding Eq. (37) are performed independently, the system-label entanglement cannot be changed. From Eq. (47), we know that the Schmidt coefficients are all $1/\sqrt{N}$, so we must have $\psi_{(N)} = 1/\sqrt{N}$ and $\psi_{(N-1,1)} = \sqrt{(N-1)/N}$. The amplitudes ϕ_r do not affect this entanglement at all—they depend on how one chooses to orthonormalize multiplicities in the Schur-Weyl transform and encode the fact that we chose the “bad” photon to be in system mode N . We can always choose $r = 1$ to correspond to this specific situation, and then use the subgroup of $U(d_S)$ that permutes system modes to find the states corresponding to the “bad” photon being in any other mode.

Making this choice and tracing out the label in Eq. (49) yields the singly distinguishable reduced state (now suppressing trivial multiplicities)

$$\begin{aligned} \rho_s &= \frac{1}{N} |(N), \underline{1}\rangle \langle (N), \underline{1}| \\ &+ \frac{1}{N} \sum_{p=1}^{N-1} |(N - 1, 1), p, \underline{1}, 1\rangle \langle (N - 1, 1), p, \underline{1}, 1|. \end{aligned} \quad (50)$$

We see that this is mixed over N dimensions of the coincident subspace, overlapping the symmetric and almost symmetric $(N - 1, 1)$ irreps.

3. Completely distinguishable states

A completely distinguishable state has each particle in a distinct label mode, paired with a unique system mode. We

can choose to order the modes such that the corresponding Fock array is diagonal; cf. Eq. (1). Generalizing the symmetrization procedure of Eqs. (5)–(8) to N particles, one finds that all $N!$ possible terms will occur in the single-particle picture, and they will each occur once. The unique pairing of system and label modes manifests as maximal entanglement between the system and label single-particle states in the coincident subspace. As above, because the Schur-Weyl transformations yielding Eq. (37) are performed independently, the system-label entanglement is preserved. This means that the transformed state must also be maximally entangled with the same Schmidt rank. Thus,

$$\begin{aligned} \hat{a}_{11}^\dagger \hat{a}_{22}^\dagger \cdots \hat{a}_{NN}^\dagger |\text{vac}\rangle &= \left| \begin{array}{cccc} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{array} \right\rangle \quad (51) \\ &= \text{Sym}(|11\rangle|22\rangle \cdots |NN\rangle) \quad (52) \\ &= \frac{1}{\sqrt{N!}} \sum_{\lambda pr} |\lambda, p, \underline{1}, r\rangle_S |\lambda, p, \underline{1}, r\rangle_L, \end{aligned} \quad (53)$$

with the sum running over all allowed values of irrep, outer, and inner multiplicities. The completely distinguishable reduced system state is therefore

$$\begin{aligned} \rho_d &= \frac{1}{N!} \sum_{\lambda pr} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r| \quad (54) \\ &= \frac{1}{N!} |(N), \underline{1}\rangle \langle (N), \underline{1}| \\ &+ \frac{1}{N!} \sum_{\lambda \neq (N), p, r} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r|, \end{aligned} \quad (55)$$

which is completely mixed over the $N!$ -dimensional coincident subspace.

D. Unitary parametrization

The unitary subgroup $U(d_S)$ corresponds to the set of interferometers that act on the system modes. We can parametrize these unitaries with what is known as a Reck scheme in optics [54,55], decomposing an arbitrary U into a sequence of single-mode unitaries (phase shifters) and unitaries that act on neighboring modes (beam splitters). As shown in Fig. 1, such a scheme can be viewed as $d_s - 1$ layers, indexed by k , each with k phase shifters and beam splitters, followed by a final phase shift on each mode. Because we are only interested in number state inputs and number counting measurements, only the phase shifters between beam splitters play a role. Hereafter, when we refer to U , we will therefore be referring to this smaller interferometer, without the initial and final sets of phase shifters.

E. Measurements

We will assume that we have access to photon number resolving detectors for the system (see Sec. VI for a discussion of a relaxation). The measurement POVM elements are projections on all states with photon occupation \underline{n} ,

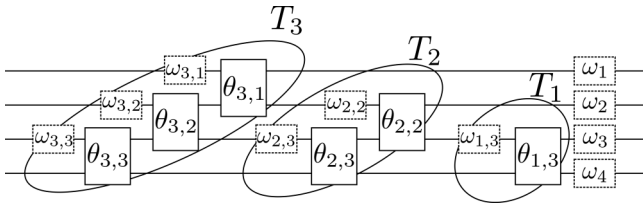


FIG. 1. Example of a Reck scheme parametrizing an arbitrary unitary transformation on four modes ($d_S = 4$), grouped into layers T_k . Each one-mode (phase shifter) and two-mode (beam splitter) subtransformation contributes one real parameter. Only the phase shifters situated between beam splitters ($\omega_{1,3}$, $\omega_{2,2}$, $\omega_{2,3}$) contribute to our problem.

$M_{\underline{n}} = \sum_{\lambda, \rho} |\lambda, \rho\rangle \langle \lambda, \rho|$. Note that this includes projections onto system states that are not symmetric; as shown in Eq. (17), distinguishable states can contain nonsymmetric system components that still give rise to clicks. Comparing with Eq. (54), we see that $M_{\underline{1}} = N! \rho_d$ —that is, up to normalization, a coincidence count is a projection onto the completely distinguishable state. As discussed above, we will usually include the interferometer in our definition of a measurement, yielding parametrized POVM elements

$$M_{\underline{n}}(U) = (\oplus_{\lambda} U^{\lambda} \otimes \mathbb{1}^{\lambda})^{\dagger} M_{\underline{n}} (\oplus_{\lambda'} U^{\lambda'} \otimes \mathbb{1}^{\lambda'}), \quad (56)$$

where $\mathbb{1}^{\lambda}$ corresponds to the irrep of the identity permutation in accordance with Eq. (28), [note that we omit this when it is only one dimensional, e.g., Eq. (21)].

IV. DISCRIMINATION OF DISTINGUISHABLE STATES

We will be interested in two problems: Discriminating the completely indistinguishable state, ρ_i , from the distinguishable states (i) ρ_s and (ii) ρ_d . From Eqs. (44), (50), and (55), we observe that each of these states is of the form

$$\rho = \alpha \rho_i + (1 - \alpha) \rho_{\bar{i}}, \quad \alpha \neq 0, \quad (57)$$

where ρ_i is pure and $\rho_{\bar{i}}$ is diagonal in the Schur-Weyl basis with support outside the symmetric subspace $\lambda = (N)$. From well-known results for the discrimination of two mixed states [56], the fact that ρ_i lies within the support of the mixed state to be discriminated means that the optimal measurement is essentially the same for either minimum error or unambiguous discrimination; one wishes to project onto the support of $\rho_{\bar{i}}$. In particular for UD, the error-free constraint means that we are forced to set $E_i = 0$, and thus the prior probabilities do not affect the optimal choice of measurement operators. This reflects the fact that there is no way to unambiguously discriminate the indistinguishable state ρ_i —we can either conclude that the state was distinguishable by observing an output that is completely suppressed by quantum interference or fail to conclude anything at all. Our task is therefore to minimize the probability of failure $E_{\gamma} = \mathbb{1} - E_{s,d}$, equivalently maximizing the probability of unambiguously detecting a singly or completely distinguishable state, respectively.

If our measurements are unrestricted, the best choice of POVM is to project onto the nonsymmetric subspace. This

choice is suitable for not only the states $\rho_{s,d}$ but by extension any state to be discriminated from ρ_i . However, as mentioned in Sec. III E, in practice we only have access to number counting measurements—we will therefore want to approximate this projection as best possible. The approximation will be sensitive to the state we are discriminating: for example, Eqs. (50) and (55) show that ρ_s can be optimally discriminated by projecting onto only the $(N-1, 1)$ irrep, while for ρ_d one wants to project on to all of the nonsymmetric irreps. As we will see, this can lead to different interferometers being optimal for discriminating different distinguishable states.

A. Restriction to linear optical measurements

In order to discriminate distinguishability in linear optics, we wish to find the best we can do with the measurements we have, namely those in Eq. (56). In the HOM case, the UD problem described by Eqs. (26) and (27) involves only a single occupation POVM element, the coincidence count $M_{\underline{n}}(U)$ with $\underline{n} = (1, 1)$. There are many ways we can approach the generalization of the HOM case. One way would be, given a specific occupation \underline{n} , to find U maximizing $\text{Tr}[\rho M_{\underline{n}}(U)]$ subject to $\text{Tr}[\rho_i M_{\underline{n}}(U)] = 0$. Notice that any \underline{n} that can be made to satisfy $\text{Tr}[\rho_i M_{\underline{n}}(U)] = 0$ for a suitable U is an unambiguous discriminator, but a single \underline{n} is not necessarily the optimal choice. In general, it is possible for multiple occupations to satisfy the UD constraint simultaneously, each contributing to the probability of success.

We therefore consider a different optimization, where we wish to find the subset of all discriminating occupations, call it D , that optimizes the success probability simultaneously, for the same choice of U :

$$\text{find } U \text{ and } D \text{ maximizing } \sum_{\underline{n} \in D} \text{Tr}[\rho M_{\underline{n}}(U)] \quad (58)$$

$$\text{subject to, for all } \underline{n} \in D, \quad \text{Tr}[\rho_i M_{\underline{n}}(U)] = 0. \quad (59)$$

Note that the quantity we are maximizing gives us the total probability of successful discrimination, which is the sum over all the unambiguously discriminating events in the set of occupations D .

While the former choice of optimization focuses on giving an optimal interferometer for discrimination given a specific measurement occupation, the latter optimization focuses on the highest probability of discrimination across all measurement patterns. In general, we find that these two problems give different optimal interferometers; here we will focus on the latter optimization over both U and D ; see Sec. VI for a discussion of a variation of the problem.

B. Scattering probabilities

Let us look at what the probability of a specific measurement pattern \underline{n} being detected at the output of an arbitrary interferometer U is for the states of interest, starting with the

completely distinguishable state. From Eqs. (55) and (56),

$$\begin{aligned}
\text{Tr}[\rho_d M_{\underline{n}}(U)] &= \text{Tr} \left[\left(\frac{1}{N!} \sum_{\lambda, p, \underline{1}, r} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r| \right) \left(\oplus_{\mu} U^{\mu} \otimes \mathbb{I}^{\mu} \right)^{\dagger} \left(\sum_{\lambda', p', \underline{n}, r'} |\lambda', p', \underline{n}, r'\rangle \langle \lambda', p', \underline{n}, r'| \right) \left(\oplus_{\mu'} U^{\mu'} \otimes \mathbb{I}^{\mu'} \right) \right] \\
&= \frac{1}{N!} \sum_{\lambda, p, \underline{n}, r, r'} \text{Tr} [|\lambda, p, \underline{n}, r\rangle \langle \lambda, p, \underline{n}, r'| (U^{\lambda} \otimes \mathbb{I}^{\lambda}) |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r| (U^{\lambda} \otimes \mathbb{I}^{\lambda})^{\dagger}] \\
&= \frac{1}{N!} \sum_{\lambda, p, \underline{n}, r, r'} |\langle \lambda, p, \underline{n}, r | U^{\lambda} \otimes \mathbb{I}^{\lambda} | \lambda, p, \underline{1}, r' \rangle|^2 \\
&= \frac{1}{N!} \sum_{\lambda, r, r'} d_{(\lambda)} |\langle \lambda, \underline{n}, r | U^{\lambda} | \lambda, \underline{1}, r' \rangle|^2, \tag{60}
\end{aligned}$$

where in the last line we have used the fact that outer multiplicities p give rise to identical copies of unitary irreps to write the probability in terms of irreducible unitary matrix elements. When $r = r' = 1$, these matrix elements are immanants [57] of a matrix $U_{\underline{1}}^{\underline{n}}$ whose rows and columns are determined by the input and output occupations of the interferometer given by U [58,59]. Moreover, the completely distinguishable case can be interpreted as independent classical particles evolving stochastically [60], leading to the remarkable fact that the sum in Eq. (60) can always be written in terms of the permanent of the matrix given by the elementwise square amplitudes of $U_{\underline{1}}^{\underline{n}}$, cf. Eq. (24) and note that $U_{\underline{1}}^{\underline{1}} = U$.

The calculation for the singly distinguishable and completely indistinguishable state is the same as Eq. (60), only with fewer irreps occurring. If we recall from Sec. III C 2 that $d_{((N-1,1))} = N - 1$, Eq. (50) gives

$$\begin{aligned}
\text{Tr}[\rho_s M_{\underline{n}}(U)] &= \frac{1}{N} | \langle (N), \underline{n} | U^{(N)} | (N), \underline{1} \rangle |^2 \\
&+ \frac{N-1}{N} \sum_r | \langle (N-1, 1), \underline{n}, r | \\
&\times U^{(N-1,1)} | (N-1, 1), \underline{1}, 1 \rangle |^2, \tag{61}
\end{aligned}$$

where the sum is over all r consistent with \underline{n} , and Eq. (44) gives

$$\text{Tr}[\rho_i M_{\underline{n}}(U)] = | \langle (N), \underline{n} | U^{(N)} | (N), \underline{1} \rangle |^2, \tag{62}$$

where as mentioned above these matrix elements are expressible in terms of per $U_{\underline{1}}^{\underline{n}}$ [61].

We observe that not all occupations are useful for unambiguous discrimination. Measurements where all the photons are bunched into a single mode only occur in the symmetric irrep; that is, if $\underline{n} = (0, \dots, 0, N, 0, \dots, 0)$, then $M_{\underline{n}} = | \langle (N), 1, \underline{n}, 1 \rangle \langle (N), 1, \underline{n}, 1 |$. In this case, Eqs. (60) and (61) are proportional to Eq. (62), and since Eq. (59) has to be satisfied, they will always give zero. Completely bunched events can therefore never help unambiguously discriminate the indistinguishable state, and we will exclude such events from our searches.

C. Numerical optimization approach

In the Results section, there is a mixture of analytical and numerical results. To construct the cost function for our

numerical work, we took into consideration the following criteria: The measurement operator $M_{\underline{n}}$ can only be included in the optimization if Eq. (59) is satisfied; when this is the case, it is added to a sum being optimized as per Eq. (58). The cost function chosen was

$$C(U) = - \sum_{\underline{n}} \exp(-\xi \text{Tr}[\rho_i M_{\underline{n}}(U)]) \text{Tr}[\rho_s M_{\underline{n}}(U)], \tag{63}$$

where ξ is adjusted (usually depending on the choice of N , and ranging from 2 to 60) to penalize results where $M_{\underline{n}}$ might be added to Eq. (58) and optimized without satisfying Eq. (59). A high penalty ξ guarantees that the value of $\text{Tr}[\rho_i M_{\underline{n}}(U)]$ is close to zero before $\text{Tr}[\rho_s M_{\underline{n}}(U)]$ is optimized and added to the sum. Combining this with the Eqs. (60) and (61), we have

$$\begin{aligned}
C_d(U) &= \frac{-1}{N!} \sum_{\lambda \neq (N)} d_{(\lambda)} \sum_{\underline{n}} e^{-\xi | \langle (N), \underline{n} | U^{(N)} | (N), \underline{1} \rangle |^2} \\
&\times \sum_{r, r'} |\langle \lambda, \underline{n}, r | U^{\lambda} | \lambda, \underline{1}, r' \rangle|^2, \tag{64}
\end{aligned}$$

$$\begin{aligned}
C_s(U) &= \frac{1-N}{N} \sum_{\underline{n}, r} e^{-\xi | \langle (N), \underline{n} | U^{(N)} | (N), \underline{1} \rangle |^2} | \langle (N-1, 1), \underline{n}, r | \\
&\times U^{(N-1,1)} | (N-1, 1), \underline{1} \rangle |^2. \tag{65}
\end{aligned}$$

PYTHON was used to optimize these functions with the Scipy library function BASINHOPPING using Broyden-Fletcher-Goldfarb-Shanno (BFGS) as the optimization algorithm. The seeds were generated using Numpy random number generation. Though this optimization function will help us explore the space, it neither guarantees that minimum is global nor exactly solves the original optimization problem. This will be problematic with minima that are close together, as, for example, $\exp(-\xi \text{Tr}[\rho_i M_{\underline{n}}(U)])$ gets closer to 1 for values of $\text{Tr}[\rho_i M_{\underline{n}}(U)]$ that are close to 0. In some situations, this value can be quite high combined with a high value of $\text{Tr}[\rho_s M_{\underline{n}}(U)]$, skewing the results toward a possible nonoptimal solution for the original problem. We could avoid this by choosing an appropriately high ξ as a function of the number of occupations $\binom{N+d_s-1}{N}$; however, if too high, $\exp(-\xi \text{Tr}[\rho_i M_{\underline{n}}(U)])$ will behave like a step function, which does not reward transitional values enough. Therefore, we do not make any strong claims

of optimality for the interferometers found numerically when they do not saturate the general bounds presented in Sec. V A.

V. RESULTS

A. General bounds

Recall from Sec. IV A the best possible unrestricted discrimination measurement is to project onto the nonsymmetric subspace, $E_{(N)} = \sum_{\lambda \neq (N), p, n, r} |\lambda p n r\rangle \langle \lambda p n r|$. Such a POVM element would be equally good for both singly and completely distinguishable states, and indeed any distinguishable state of the form in Eq. (57). The success probability of such a measurement is given by

$$\begin{aligned} \text{Tr}[\rho(\oplus_{\lambda} U^{\lambda} \otimes \mathbb{I}^{\lambda})^{\dagger} E_{(N)} (\oplus_{\lambda} U^{\lambda} \otimes \mathbb{I}^{\lambda})] \\ = \text{Tr}[(\alpha \rho_i + (1 - \alpha) \rho_d) E_{(N)}] \\ = 1 - \alpha \\ = \begin{cases} 1 - \frac{1}{N} & \text{if } \rho = \rho_s \\ 1 - \frac{1}{N!} & \text{if } \rho = \rho_d \end{cases}, \end{aligned} \quad (66)$$

where we have used the fact that any projector onto irreps is unitarily invariant. These then are universal upper bounds on the success probability for singly and completely distinguishable states, respectively. However, since we are restricted to photon number counting measurements, we will see that while the first bound is achievable, the second is not, in general. We will go through various examples in detail in the following sections.

B. Two modes

1. Two photons in two modes

In the case of two photons in two modes, the states to be discriminated are, from Eqs. (44), (50), and (55),

$$\rho_i = |112\rangle \langle 112|, \quad \text{and} \quad (67)$$

$$\rho_s = \rho_d = \frac{1}{2} |112\rangle \langle 112| + \frac{1}{2} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|. \quad (68)$$

Observing that there is only one available state which is not symmetric, it is easy to write down an arbitrary *partially* distinguishable system state in this case, since there is but one parameter:

$$\rho = \alpha |112\rangle \langle 112| + (1 - \alpha) \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|. \quad (69)$$

As discussed in Sec. IV B, only occupations that do not have all the photons bunched in the same mode can be used for meaningful discrimination, in this case leaving only one choice of projector, the coincidence $M_{(1,1)} = |112\rangle \langle 112| + \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|$.

In our discussion in Sec. II, we claimed that the optimal discriminator is given by a coincidence count and a balanced beam splitter; we can now prove this assertion. First, note that since there is only one antisymmetric state, the antisymmetric irreducible representation of any U has but one matrix element and so the action of any interferometer on this state is trivial [in Eq. (21) given by its determinant]. Thus, the only contribution to the nonsymmetric part of Eq. (60)

is $\left| \left(\frac{1}{2} \right) U \left(\frac{1}{2} \right) \right| = 1$, and there is nothing to maximize in Eq. (58). All that is left is to satisfy the constraint, Eq. (59). Parametrizing U as

$$\begin{bmatrix} e^{i\phi} \cos \theta & e^{i\varphi} \sin \theta \\ -e^{-i\varphi} \sin \theta & e^{-i\phi} \cos \theta \end{bmatrix}, \quad (70)$$

one finds that the constraint is then $\text{per} U = \cos^2 \theta - \sin^2 \theta = \cos 2\theta = 0$, with the family of solutions $\{(\phi, \varphi, \pi/4) | 0 \leq \phi \leq \pi, 0 \leq \varphi \leq \pi\}$. The solutions do not depend on the phases ϕ or φ , as we would expect from the discussion in Sec. III D, but only on the choice of the beam-splitter reflectivity, which is balanced as claimed.

We see that not only does unambiguous discrimination return the HOM measurement as was discussed in Sec. II A, it is optimal for an arbitrary partially distinguishable two-photon state.

2. Three photons in two modes

As an example of the utility of the formalism, in this subsection we will make a slight digression and consider the simplest nontrivial case with $N (= 3) > d_S (= 2)$. As mentioned in Sec. III C, this restricts the kinds of distinguishable states that can occur; we consider situations with two photons in one system mode and the third in the other. The indistinguishable state is $\hat{a}_{11}^{\dagger} \hat{a}_{11}^{\dagger} \hat{a}_{21}^{\dagger} |\text{vac}\rangle = |2_1^2\rangle$, with reduced state

$$\rho_i = |112\rangle \langle 112|. \quad (71)$$

There are essentially two types of distinguishable state in this situation. The first is $\hat{a}_{11}^{\dagger} \hat{a}_{11}^{\dagger} \hat{a}_{22}^{\dagger} |\text{vac}\rangle = |2_0^0\rangle$, and the second is $\hat{a}_{11}^{\dagger} \hat{a}_{12}^{\dagger} \hat{a}_{21}^{\dagger} |\text{vac}\rangle = |1_1^1\rangle$. Other states are equivalent to the above for the reasons discussed in Sec. III C 2. Further, the (now incompletely) distinguishable state $\hat{a}_{11}^{\dagger} \hat{a}_{12}^{\dagger} \hat{a}_{23}^{\dagger} |\text{vac}\rangle = |1_0^1 1_0^1\rangle$ has a reduced state that is the same as Eq. (72), and will therefore have the same discrimination measurement and success probability. The reduced state for the first case is

$$\begin{aligned} \rho_{s_1} = \frac{1}{3} |112\rangle \langle 112| \\ + \frac{1}{3} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|_1 + \frac{1}{3} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|_2, \end{aligned} \quad (72)$$

while that for the second case is

$$\begin{aligned} \rho_{s_2} = \frac{4}{6} |112\rangle \langle 112| \\ + \frac{1}{6} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|_1 + \frac{1}{6} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right|_2. \end{aligned} \quad (73)$$

Note that Eq. (58) does not depend on the amplitude of the symmetric part of the state—its contribution has to be zero by Eq. (59). It only depends on the nonsymmetric components, and since ρ_{s_1} and ρ_{s_2} are equally weighted across the available nonsymmetric states, the optimal discriminator will be the same. However ρ_{s_2} does have half of the amplitude of ρ_{s_1} in this subspace, which will halve the success probability.

There are four possible occupations to measure; however, as mentioned in Sec. IV B, the bunched ones can be disregarded and the optimization carried out on $M_{(2,1)}$

and $M_{(1,2)}$. We parametrize U again as in Eq. (70). For $M_{(2,1)}$, Eq. (59) reduces to $|\langle \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} | U | \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \rangle| = |(\cos \theta + 3 \cos 3\theta)/4| = 0$. Since $0 \leq \theta \leq \pi$, this equation is true for $\theta \in \{\pi/2, \arccos(\sqrt{2/3}), \arccos(-\sqrt{2/3})\}$. On the other hand, Eq. (59) for $M_{(1,2)}$ is $|\langle \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} | U | \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \rangle| = |(\sin \theta - 3 \cos 3\theta)/4|$. This equation cannot be zero for the above choice of angles that ensure $|\langle \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} | U | \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \rangle| = 0$. Thus, only one of the outcomes can be used to discriminate these states; without loss of generality, we choose to optimize for $M_{(2,1)}$. In this case, we want to maximize $\text{Tr}[\rho_{s_1} M_{(2,1)}(U)] = 2 \left| \left\langle \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \middle| U^\dagger \begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix} \right\rangle \right|^2 = 2 \cos^2 \theta/3$. When $\theta = \pi/2$, we get success probability of 0. When $\theta = \pm \arccos(\sqrt{2/3})$, we get success probability of 4/9. Thus, an optimal discriminating interferometer is $U = \begin{bmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{bmatrix} / \sqrt{3}$, with success probabilities 4/9 for ρ_{s_1} and 2/9 for ρ_{s_2} .

C. Three modes

From now on, we will only consider coincident input with $N = d_S$. For three photons in three system modes, the completely indistinguishable reduced state is from Eq. (44),

$$\rho_i = |(3), \underline{1}\rangle \langle (3), \underline{1}| = \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix}. \quad (74)$$

There are now three different singly distinguishable states, depending on which system mode the bad photon is in. In the Schur-Weyl basis (see Sec. III A 1), their full system-label states, as per the discussion in Sec. III C 2, are

$$\begin{aligned} \sqrt{3} \hat{a}_{11}^\dagger \hat{a}_{21}^\dagger \hat{a}_{32}^\dagger |\text{vac}\rangle &= \sqrt{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad + \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad + \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix}, \end{aligned} \quad (75)$$

$$\begin{aligned} \sqrt{3} \hat{a}_{11}^\dagger \hat{a}_{22}^\dagger \hat{a}_{31}^\dagger |\text{vac}\rangle &= \sqrt{3} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad - \frac{1}{2} \left(\begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \sqrt{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \right) \\ &\quad - \frac{1}{2} \left(\begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \sqrt{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \right), \end{aligned} \quad (76)$$

$$\begin{aligned} \sqrt{3} \hat{a}_{12}^\dagger \hat{a}_{21}^\dagger \hat{a}_{31}^\dagger |\text{vac}\rangle &= \sqrt{3} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad - \frac{1}{2} \left(\begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} - \sqrt{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \right) \\ &\quad - \frac{1}{2} \left(\begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} - \sqrt{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \right). \end{aligned} \quad (77)$$

While for completely distinguishable states permuting system modes has no effect on the reduced state, here the reduced states will not be invariant. However, because permutations of system modes lie inside the set of allowed operations [that is, $S_{d_S} \subset U(d_S)$], if we optimize for one of these states, the

resulting interferometer will be easily related to the others by including some mode swapping. Therefore, we can focus on one of these states and the success probabilities that we find will be the same for the other two; Eq. (75) has the reduced state [cf. Eq. (50)]

$$\begin{aligned} \rho_s &= \frac{1}{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad + \frac{1}{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \frac{1}{3} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix}. \end{aligned} \quad (78)$$

It is also natural to ask about discrimination of mixtures of the three states in Eqs. (75), (76), and (77); we will discuss this in Sec. V C 2.

The completely distinguishable state corresponding to $\hat{a}_{11}^\dagger \hat{a}_{22}^\dagger \hat{a}_{33}^\dagger |\text{vac}\rangle = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ per Eq. (55) is

$$\begin{aligned} \rho_d &= \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad + \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \\ &\quad + \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} + \frac{1}{6} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix}. \end{aligned} \quad (79)$$

For the following, let us define two sets of measurement operators: Those with two photons in one mode, $M_{(2,1,0)} = \sum_{\lambda \neq \square, p} |\lambda, p, (2, 1, 0)\rangle \langle \lambda, p, (2, 1, 0)|$, $M_{(2,0,1)}$, $M_{(1,0,2)}$, $M_{(1,2,0)}$, $M_{(0,1,2)}$, and $M_{(0,2,1)}$, which we denote \mathcal{M}_2 , and those with each photon in a different mode, that is, $\mathcal{M}_1 \ni M_{(1,1,1)} = \sum_{\lambda, p, r} |\lambda, p, \underline{1}, r\rangle \langle \lambda, p, \underline{1}, r|$. As discussed in Sec. IV B, the measurements $M_{(3,0,0)} = \begin{bmatrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{bmatrix} \begin{bmatrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{bmatrix}$, $M_{(0,3,0)}$, and $M_{(0,0,3)}$ will not be helpful for discrimination.

1. Discriminating singly distinguishable states

Let ρ^λ denote the (un-normalized) part of a state supported on the subspace of irrep λ . Notice that ρ_s has no support in the antisymmetric subspace, so that $\sum_{\underline{n}} \text{Tr}[\rho_s^\square M_{\underline{n}}(U)] = 2/3$ and $\sum_{\underline{n}} \text{Tr}[\rho_s^\square M_{\underline{n}}(U)] = 0$. This means that for any subset of occupations D and any U for which Eq. (59) holds, the success probability will be bounded by 2/3. It is well known how to saturate this; use a balanced tritter, $U = QFT_3$, and all the occupations from \mathcal{M}_2 , where QFT_N is defined as

$$QFT_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega^1 & \dots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \dots & \omega^{(N-1)(N-1)} \end{bmatrix} \quad (80)$$

and $\omega = \exp \frac{2\pi i}{N}$. A parametrization that realizes a balanced tritter is given in Fig. 2.

2. Discriminating mixed singly distinguishable states

A short digression regarding mixed system-label states: If we were (uniformly) ignorant about which mode the bad photon was in, we would have an equal mixture of Eqs. (75),

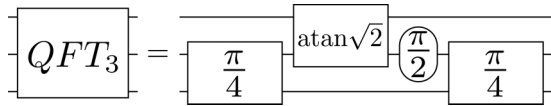


FIG. 2. The best-known interferometer for discriminating completely indistinguishable from distinguishable states of three photons in three modes is QFT_3 , with a success probability of $2/3$. Up to phases, it consists of two balanced beam splitters, one 2 : 1 beamsplitter, and one $\pi/2$ phase shifter.

(76), and (77). The resulting mixed state is

$$\begin{aligned} \rho_{\text{sm}} := & \frac{1}{3} |\underline{123}\rangle \langle \underline{123}| \\ & + \frac{1}{6} |\underline{12}\rangle_1 \langle \underline{12}\rangle_1 + \frac{1}{6} |\underline{12}\rangle_2 \langle \underline{12}\rangle_2 \\ & + \frac{1}{6} |\underline{13}\rangle_1 \langle \underline{13}\rangle_1 + \frac{1}{6} |\underline{13}\rangle_2 \langle \underline{13}\rangle_2. \end{aligned} \quad (81)$$

The overlap $\sum_{\underline{n}} \text{Tr}[\rho_{\text{sm}}^{\square} M_{\underline{n}}(U)] = 2/3$ for \mathcal{M}_2 again saturates the bound, and a balanced tritter remains the best choice of interferometer. This can be seen from the symmetry of the QFT which treats a bad photon in any mode essentially the same way and so should be true for analogous singly distinguishable mixed states for all N ; however, we will not discuss mixed system-label states further here.

3. Discriminating completely distinguishable states

Using the cost function from Eq. (64) and a range of penalties $\xi \in \{2, 4, 6, 8, 10\}$, we find that the highest success probability in the completely distinguishable case is $2/3$. The measurement operators are always the full set \mathcal{M}_2 with a balanced tritter as a solution, just as in the previous section. However, this does not saturate the bound in Sec. V A, which is $5/6$ in the case of three photons. To investigate this further, we try to understand the structure of the state a bit better and use numerical evidence to show that a balanced tritter is likely to be optimal.

From Eq. (79), we have $\sum_{\underline{n}} \text{Tr}(\rho_{\text{d}}^{\square} M_{\underline{n}}(U)) = 2/3$ and $\sum_{\underline{n}} \text{Tr}(\rho_{\text{d}}^{\square} M_{\underline{n}}(U)) = 1/6$, so that $\sum_{\underline{n}} \text{Tr}(\rho_{\text{d}}^{\square} M_{\underline{n}}(U)) = 5/6$, which is the discrimination bound. Notice that operators from \mathcal{M}_2 do not have support on the antisymmetric subspace. Therefore, if we only pick operators from \mathcal{M}_2 as the discriminating operators and assume they can simultaneously satisfy Eq. (59), then $\sum_{\underline{n} \in \mathcal{M}_2} \text{Tr}(\rho_{\text{d}} M_{\underline{n}}(U)) = \sum_{\underline{n} \in \mathcal{M}_2} \text{Tr}(\rho_{\text{d}}^{\square} M_{\underline{n}}(U)) \leq 2/3$. This is exactly what happens for the interferometers from our optimization.

This tells us that if we want the success probability to be larger than $2/3$, the only operator left, $M_{(1,1,1)}$, would have to be included. Our numerical results show that, on the contrary, it is unlikely for any D that includes $M_{(1,1,1)}$ to give a success probability over $1/2$. We do this with a new cost function, much like Eq. (64) but modified to force $M_{(1,1,1)}$ to be included:

$$C_{\text{d},111}(U) = \eta \text{Tr}(\rho_i M_{(1,1,1)}(U)) + C_{\text{d}}(U), \quad (82)$$

where η is a penalty to ensure Eq. (59) for $M_{(1,1,1)}$ has to be satisfied and $C_{\text{d}}(U)$ is as defined in Eq. (63). This penalty is

set to $\eta = 10$, making the first term an order of magnitude higher than the second term of Eq. (82), where we took $\xi = 6$. As we learned in Sec. III D, we can ignore the outside phase shifters of the standard Reck parametrization; therefore, we are only optimizing over four parameters, $\theta_{2,1}, \theta_{2,2}, \theta_{1,2}$, and $\omega_{1,2}$. The lowest value of the cost function found by the optimization techniques in Sec. IV C is -0.500426 . This corresponds to a success probability of 0.5 in discriminating between the two states, which is lower than the $2/3$ achievable when $M_{(1,1,1)}$ is not included.

While this does not give us definitive proof that no scheme that includes a threefold coincidence can give success probability higher than $2/3$, it does strongly indicate that this should be true. Moreover, with the same optimization functions we investigated how many of the other operators alongside $M_{(1,1,1)}$ we can pick at the same time, and it seems that the best we can do is to have four from \mathcal{M}_2 satisfy Eq. (59) simultaneously. However, in all the situations when this occurs, some of the terms in Eq. (58) are zero; thus, the success probability remains at $1/2$, which can be achieved using just $M_{(1,1,1)}$ and a balanced beam splitter.

The balanced tritter uses all the measurement operators from \mathcal{M}_2 , with each contributing $1/9$ to achieve the success probability $2/3$. To draw attention to the difference between optimizing a single operator and multiple operators at once, mentioned in Sec. IV A, we notice that optimizing for one operator from the set \mathcal{M}_2 yields a success probability higher than $1/9$ (for some other choice of U). Taking this further, we can search numerically for the single best outcome, with a cost function similar to that of Eq. (64), except we now focus only on a single \underline{n} , that is $C(\underline{n}, U) = -2 \sum_{\lambda, r, r'} \exp(-\xi | \langle (3), \underline{n} | U | (3), \underline{1} \rangle |^2) | \langle \lambda, \underline{n}, r' | U | \lambda, \underline{1}, r \rangle |^2$. We find $M_{(1,1,1)}$ is a clear winner with a total success probability of $1/2$, achievable by a balanced beam splitter as mentioned above. All of the other operators by themselves only ever give an optimized success probability of $1/8$. Notice that $6 \times 1/8 = 3/4 > 2/3$, showing that the strategy that gives us the best success chance with a single operator from \mathcal{M}_2 can not be achieved simultaneously by all six of them.

D. Four and more modes



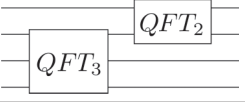

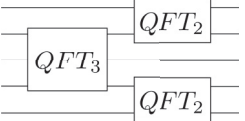

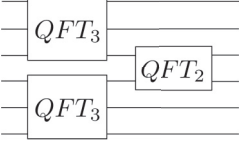
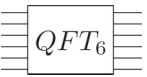
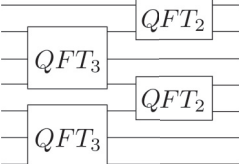
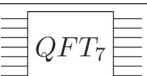
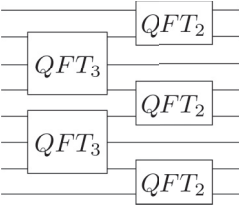

1. Discriminating singly distinguishable states

Using the numerical optimization described in Sec. IV C, we also examined the discrimination of singly distinguishable states for $N = 4$ and 5 photons. Together with the results for $N = 2$ and 3, we see the optimization return interferometers equivalent to QFT_N , each giving a success probability $1 - 1/N$, saturating the bound in Sec. V A. We have confirmed this behavior by direct calculation up to $N = 9$.

2. Discriminating completely distinguishable states

Numerical optimization for the $N = 4$ and 5 photon completely distinguishable states yields success probabilities of $19/24$ and $31/36$, respectively. Both of these are less than the general bounds of Sec. V A ($23/24$ and $35/36$, respectively), and so we cannot conclude they are optimal. We observe that they do both exceed the singly distinguishable bound of $1 - 1/N$, consistent with the intuition that it is easier to

TABLE I. The best known interferometers for discrimination of the singly and completely distinguishable states of $N = 2$ to 8 photons in N modes. For $N = 2$ and 3 the quantum Fourier transform (QFT_N) is optimal for both ρ_s and ρ_d , but for $N \geq 4$ the interferometers for each are different; we include all probabilities of success for comparison. For singly distinguishable states, the quantum Fourier transform saturates the bound and so is optimal for each N ; due to the QFT 's symmetry it does not matter which port the 'bad' photon (see Sec. III C 2) is in, however this is not true of the ρ_d interferometers and so we include best, worst and average success probabilities assuming each port is equally likely to be 'bad'. The completely distinguishable state is essentially unique, so there is only one success probability to report; an asterisk * indicates extensive numerical optimization leads us to believe the $N = 3, 4, 5$ cases are optimal despite being far from the bound; it is remarkable that the ρ_d interferometers have constant optical depth (made up of QFT_3 s followed by QFT_2 s) for each N . Interestingly, the two success probabilities for the QFT are always equal except for $N = 6$, the only case in the table that is not a power of a prime (see Discussion). The measurement outcomes that lead to these probabilities are specified in Table II.

N	U	Singly distinguishable, ρ_s				Completely distinguishable, ρ_d	
		Bound ($1 - 1/N$)	Best	Worst	Avg	Bound ($1 - 1/N!$)	Success probability
2		$\frac{1}{2} = 0.5000$	$\frac{1}{2} = 0.5000$			$\frac{1}{2} = 0.5000$	$\frac{1}{2} = 0.5000$
3		$\frac{2}{3} \approx 0.6666$	$\frac{2}{3} \approx 0.6666$			$\frac{5}{6} \approx 0.8333$	$\frac{2}{3} \approx 0.6666^*$
4		$\frac{3}{4} = 0.7500$	$\frac{25}{36} \approx 0.6944$	$\frac{1}{4} = 0.2500$	$\frac{7}{12} \approx 0.5833$	$\frac{23}{24} \approx 0.9583$	$\frac{19}{24} \approx 0.7916^*$
			$\frac{3}{4} = 0.7500$				$\frac{3}{4} = 0.7500$
5		$\frac{4}{5} = 0.8000$	$\frac{8}{15} \approx 0.7222$	$\frac{1}{4} = 0.2500$	$\frac{13}{18} \approx 0.5333$	$\frac{119}{120} \approx 0.9917$	$\frac{31}{36} \approx 0.8611^*$
			$\frac{4}{5} = 0.8000$				$\frac{4}{5} = 0.8000$
6		$\frac{5}{6} \approx 0.8333$	$\frac{167}{243} \approx 0.6872$	$\frac{167}{243} \approx 0.6872$	$\frac{167}{243} \approx 0.6872$	$\frac{719}{720} \approx 0.9986$	$\frac{671}{729} \approx 0.9204$
			$\frac{5}{6} \approx 0.8333$				$\frac{65}{72} \approx 0.9028$
7		$\frac{6}{7} \approx 0.8571$	$\frac{695}{972} \approx 0.7150$	$\frac{1}{4} = 0.2500$	$\frac{361}{567} \approx 0.6367$	$\frac{5039}{5040} \approx 0.9998$	$\frac{2765}{2916} \approx 0.9482$
			$\frac{6}{7} \approx 0.8571$				$\frac{6}{7} \approx 0.8571$
8		$\frac{7}{8} = 0.8750$	$\frac{695}{972} \approx 0.7150$	$\frac{1}{4} = 0.2500$	$\frac{97}{162} \approx 0.5988$	$\frac{40319}{40320} \approx 1.0000$	$\frac{45095}{46656} \approx 0.9665$
			$\frac{7}{8} = 0.8750$				$\frac{7}{8} = 0.8750$

discriminate a completely distinguishable state than one that is less distinguishable.

The numerics are sensitive to the penalties used in Eq. (64), due to the existence of interferometers with very similar performance. For $N = 4$, a penalty $\xi = 10$ returns an interferometer with success probability $25/32$ that minimizes the cost function with a value of -0.839477 , while a better interferometer with success probability $19/24$ exists but gives a higher value of -0.836287 . Increasing the penalty to 50 yields costs -0.78455 and -0.79277 for these two interferometers respectively, showing that the latter is now the minimum. However, increasing the penalty makes optimization more difficult, because the landscape flattens and gradients go to zero. For this reason, penalties of 10, 13, 15, 17, 20, 25, 35, and 50 were used for $N = 4$ and 10, 12, 14, 15, 16, 18, 20, 35, and 60 for $N = 5$.

While the complexity of the calculations precluded any further optimization for $N > 5$, we notice that the best interferometers for $N = 2, 3, 4, 5$ can be composed out of QFT_3 followed by QFT_2 s. This suggests a recursive structure for the best discriminating interferometers; for $N = 6, 7, 8$ we tried combinations of QFT_N , QFT_{N-1} , and so on, and found that discriminators composed of QFT_3 s followed by QFT_2 s performed the best. This is remarkable as these are of constant optical depth (the maximum number of beam splitters and phase shifters that each photon encounters), independent of N . Indeed, increasing the optical depth beyond this seems to decrease the success probability, which allowed us to limit our search to a manageable number of configurations. These are educated guesses, however, and do not rule out the existence of better interferometers that might be found.

Table I contains a summary of these results. We report the probabilities for the best interferometers found to successfully discriminate ρ_s and ρ_d from ρ_i up to $N = 8$. The measurement outcomes that achieve these probabilities up to $N = 5$ are specified in Table II, where in the interest of saving space we give the occupations that fail (i.e., correspond to the *ambiguous* POVM element $E_?$) instead of the successful discrimina-

tors, because the latter far exceed the former. For comparison, for each interferometer we include success probabilities for both states of interest to be discriminated from the completely indistinguishable state. Note that as discussed above for $N = 3$, although the QFT_N interferometer is optimal for ρ_s no matter which system mode the bad photon is in, this will not be true for interferometers that lack the symmetry of QFT_N . Indeed, the best ρ_d discriminator does not treat each system mode the same way, and so when using such an interferometer to discriminate ρ_s we report best, worst, and average success probabilities, assuming each system mode is equally likely to contain the bad photon.

VI. DISCUSSION AND FURTHER WORK

Although we have focused on single and complete distinguishability, as shown in Sec. III B the formalism admits arbitrary states. Consider, for example, Fock arrays with a single excitation in each system mode and an arbitrary label occupation; call it \underline{n}_L . By applying the Schur-Weyl transform and focusing on the symmetric irrep (N), where the support is one dimensional, we see that the reduced system state will be of the form

$$\frac{\underline{n}_L!}{N!} |(N), \underline{1}\rangle \langle (N), \underline{1}| + \left(1 - \frac{\underline{n}_L!}{N!}\right) \rho_{\bar{1}}, \quad (83)$$

where $\underline{n}_L! = n_1!n_2!\cdots$. This gives a bound of $1 - \underline{n}_L!/N!$ on the probability for successfully discriminating such a state from the completely indistinguishable one and includes the singly and completely distinguishable cases above. The exact form of such states could be found by reasoning as in Sec. III C.

We can use the formalism to compute the number of parameters that describe an arbitrary partially distinguishable collection of N particles in N (or more generally d) modes. Because of the maximal entanglement over p in Eq. (37), when we trace out the label of an arbitrary totally symmetric state in Eq. (36), the resulting mixed state has identical blocks

TABLE II. Measurement occupations corresponding to the ambiguous POVM element $E_?$ that do *not* discriminate the two states of interest for the numerically optimized interferometers in Table I ($N = 2, 3, 4, 5$)—these are, in general, far fewer than the number of successful discriminating occupations and so easier to list. Recall that for ρ_s , the optimal choice of QFT_N does not depend on the mode in which the single distinguishable photon is present, and neither do the occupations. Note that although all of the occupations not listed here satisfy Eq. (59), some might have zero probability of occurring and therefore not contribute to discrimination.

N	ρ_s	ρ_d
2	20,02	20,02
3	300,030,003 111	300,030,003 111
4	4000,0400,0040,0004 2020,0202 2101,1210,1012,0121	4000,0400,0040,0004 3100,1300,1030,1003,0130,0103
5	50000,05000,00500,00050,00005 31001,30110,13100,11030,10301,10013,03011,01310,01103,00131 22010,21200,20102,20021,12002,10220,02201,02120,01022,00212 11111	50000,05000,00500,00050,00005 40010,40001,10040,10004,04010,04001,01040,01004 31010,31001,13010,13001,10310,10301,10031,10013,01310,01301,01031,01013 20120,20102,02120,02102

for each copy of λ (the number of identical blocks being equal to the outer multiplicity). Thus, the most general mixed state is described by a single (Hermitian) block for each irrep. Recalling that the number of real parameters in a d -dimensional Hermitian matrix is d^2 , we have for an arbitrary partially distinguishable mixed state of N bosons in N modes (subtracting one for normalization) $\sum_{\lambda} d_{(\lambda)}^2 - 1 = \binom{N^2+N-1}{N} - 1$ real parameters. If we restrict to coincident input, the number of states is given by the number of *standard* Young tableaux, $d_{(\lambda)}$. This is because coincidence implies each single particle state is different, and so semistandard tableaux become standard; in this case, we have $\sum_{\lambda} d_{(\lambda)}^2 - 1 = N! - 1$ real parameters. This number decreases significantly if pure label states are assumed. A pure state in d dimensions has $2(d-1)$ real parameters, and every pure state added to a set can add at most one parameter beyond those required to describe its projection onto the up to $d-1$ dimensional space spanned by the states that came before it (namely the angle it makes with this subspace). Thus, there are $\sum_{d=2}^N (2(d-2) + 1) = (N-1)^2$ real parameters in this case, which agrees with previous analyses [16,62] but is far fewer than the general case. Note that all of these quantities are of course larger than $\binom{N}{2}$, the number of pairwise distinguishabilities classical intuition might lead one to believe are necessary to measure [25].

There are many other state discrimination scenarios we could consider. For example, we could try to unambiguously discriminate ρ_d from ρ_s , or two entirely different states, or more than two states. Note that due to the nested structure of our three states of interest [cf. Eq. (57)], attempting to find a UD POVM $\{E_i, E_d, E_s, E_\gamma\}$ reduces to only being able to discriminate ρ_d from the rest. Another version of discrimination to consider is using bucket (yes-no) instead of number-resolving detectors, which are simpler to engineer. While our focus has been on optimizing over all the possible measurement patterns to obtain the highest possible success probability, as mentioned in Sec. III E another type of optimization that can be carried out is choosing a fixed set of patterns and optimizing the interferometer U only. The difference would be that in Eqs. (58) and (59) D would now be fixed, simplifying the problem. As an example, during the preparation of this paper, a closely related paper was released [40], where the authors study a single reference photon input into a QFT_{N-1} , followed by QFT_2 HOM tests on the $N-1$ outputs with the rest of the $N-1$ photons (for a total of N photons in

$2N-2$ modes). This is equivalent to a UD procedure where D is fixed as the set of N -fold coincidences. The approach is different and so it is not surprising that it is suboptimal for discrimination; however, this interferometer's behavior is clear for all N .

We can also link the idea of unambiguous discrimination of distinguishable states with prior work done on suppression or zero-transmission laws [36,63]. Suppression laws for an interferometer U identify output occupations whose probabilities are zero for states that are quantum (completely indistinguishable), when compared to states that are classical (completely distinguishable). The focus has been on finding formulas that predict suppression of events for either specific unitaries or families of unitaries. The UD procedure studied here can be viewed as a search over the space of all interferometers, where the optimality of U is measured by the total probability of completely suppressed occupations. Although none of the suppression results so far offer higher total suppression probability than that of the QFT for the choices of N given in Table I, our work shows that the QFT is not the optimal choice of interferometer when it comes to total suppression probability. Furthermore, we have checked the success probabilities of the QFT up to $N=12$, and only the cases $N=6, 10$, and 12 have higher success probabilities than $1-1/N$; these non-power-of-a-prime cases are exactly those for which it is known that the suppression law does not account for all the suppressed patterns.

Finally, we have no doubt that proofs for many of the results here, such as the optimality of QFT_N for discriminating singly distinguishable states, should be possible, but they are left as further work.

The data associated with this paper are available for download at the University of Bristol data repository, data.bris [64].

ACKNOWLEDGMENTS

The authors are pleased to acknowledge helpful conversations with S. Bartlett, S. Croke, A. Doherty, H. de Guise, and J. Silverstone, and especially P. Birchall, S. Pallister, T. Rudolph, C. Sparrow, and T. Sugiyama, who undertook early investigations, as well as valuable comments from the referee. S.S. was supported by the Bristol Quantum Engineering Centre for Doctoral Training, EPSRC Grant No. EP/L015730/1. P.S.T. was supported in part by EPSRC First Grant No. EP/N014812/1.

-
- [1] C.-K. Hong, Z.-Y. Ou, and L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).
 - [2] E. Knill, R. Laflamme, and G. J. Milburn, *Nature (London)* **409**, 46 (2001).
 - [3] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh *et al.*, *Science* **349**, 711 (2015).
 - [4] T. Rudolph, *APL Photon.* **2**, 030901 (2017).
 - [5] S. Aaronson and A. Arkhipov, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (ACM Press, New York, 2011), pp. 333–342.
 - [6] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys *et al.*, *Science* **339**, 798 (2013).
 - [7] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, *Science* **339**, 794 (2013).
 - [8] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, *Nat. Photon.* **7**, 545 (2013).
 - [9] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Nat. Photon.* **7**, 540 (2013).

- [10] M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Viggianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvão, A. Crespi *et al.*, *Sci. Adv.* **1**, e1400255 (2015).
- [11] H. Wang, Y. He, Y.-H. Li, Z.-E. Su, B. Li, H.-L. Huang, X. Ding, M.-C. Chen, C. Liu, J. Qin *et al.*, *Nat. Photon.* **11**, 361 (2017).
- [12] A. Neville, C. Sparrow, R. Clifford, E. Johnston, P. M. Birchall, A. Montanaro, and A. Laing, *Nat. Phys.* **13**, 1153 (2017).
- [13] N. Viggianiello, F. Flamini, M. Bentivegna, N. Spagnolo, A. Crespi, D. J. Brod, R. Osellame, and F. Sciarrino, [arXiv:1710.03578](https://arxiv.org/abs/1710.03578).
- [14] M. C. Tichy, *J. Phys. B* **47**, 103001 (2014).
- [15] M. C. Tichy, *Phys. Rev. A* **91**, 022316 (2015).
- [16] A. J. Menssen, A. E. Jones, B. J. Metcalf, M. C. Tichy, S. Barz, W. S. Kolthammer, and I. A. Walmsley, *Phys. Rev. Lett.* **118**, 153603 (2017).
- [17] V. S. Shchesnovich, *Phys. Rev. A* **91**, 013844 (2015).
- [18] V. Shchesnovich, [arXiv:1712.03191](https://arxiv.org/abs/1712.03191).
- [19] M. Tillmann, S.-H. Tan, S. E. Stoeckl, B. C. Sanders, H. de Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Phys. Rev. X* **5**, 041015 (2015).
- [20] V. Tamma and S. Laibacher, *Phys. Rev. Lett.* **114**, 243601 (2015).
- [21] P. P. Rohde, *Phys. Rev. A* **91**, 012307 (2015).
- [22] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley, *Phys. Rev. Lett.* **120**, 220502 (2018).
- [23] T. Brünner, G. Dufour, A. Rodriguez, and A. Buchleitner, *Phys. Rev. Lett.* **120**, 210401 (2018).
- [24] S. Popescu, *Phys. Rev. Lett.* **99**, 250501 (2007).
- [25] R. B. A. Adamson, P. S. Turner, M. W. Mitchell, and A. M. Steinberg, *Phys. Rev. A* **78**, 033832 (2008).
- [26] M. Oszmaniec and D. J. Brod, [arXiv:1801.06166](https://arxiv.org/abs/1801.06166).
- [27] A. E. Moylett and P. S. Turner, *Phys. Rev. A* **97**, 062329 (2018).
- [28] D. Rowe, M. Carvalho, and J. Repka, *Rev. Mod. Phys.* **84**, 711 (2012).
- [29] A. Chefles, *Phys. Lett. A* **239**, 339 (1998).
- [30] J. A. Bergou, U. Herzog, and M. Hillery, in *Quantum State Estimation* (Springer, Berlin, 2004), pp. 417–465.
- [31] S. M. Barnett and S. Croke, *Adv. Opt. Photon.* **1**, 238 (2009).
- [32] R. A. Campos, *Phys. Rev. A* **62**, 013809 (2000).
- [33] S. P. Walborn, A. N. de Oliveira, S. Pádua, and C. H. Monken, *Phys. Rev. Lett.* **90**, 143601 (2003).
- [34] Y. L. Lim and A. Beige, *New J. Phys.* **7**, 155 (2005).
- [35] B.-G. Englert, D. Kaszlikowski, L. C. Kwek, and W. H. Chee, *Int. J. Quantum. Inform.* **6**, 129 (2008).
- [36] M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, and A. Buchleitner, *Phys. Rev. Lett.* **104**, 220405 (2010).
- [37] J. C. Garcia-Escartin and P. Chamorro-Posada, *Phys. Rev. A* **87**, 052330 (2013).
- [38] A. Crespi, *Phys. Rev. A* **91**, 013811 (2015).
- [39] L. Rigovacca, C. Di Franco, B. J. Metcalf, I. A. Walmsley, and M. S. Kim, *Phys. Rev. Lett.* **117**, 213602 (2016).
- [40] D. J. Brod, E. F. Galvão, N. Viggianiello, F. Flamini, N. Spagnolo, and F. Sciarrino, [arXiv:1804.01334](https://arxiv.org/abs/1804.01334).
- [41] N. M. VanMeter, P. Lougovski, D. B. Uskov, K. Kieling, J. Eisert, and J. P. Dowling, *Phys. Rev. A* **76**, 063808 (2007).
- [42] For example, $\hat{a}_{sl}^\dagger = \int d\omega f_l(\omega)\hat{a}_s^\dagger(\omega)$, where f_l is a spectral envelope function indexed by l .
- [43] The system state itself is entangled with respect to the impractical particle tensor product—this has been referred to as “free” entanglement by Aaronson [5].
- [44] A. Harrow, Applications of coherent classical communication and the Schur transform to quantum information theory, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2005 (unpublished).
- [45] We need not be concerned with the inconsequential difference between unitary groups $U(d)$ and special unitary groups $SU(d)$.
- [46] T. Rudolph, R. W. Spekkens, and P. S. Turner, *Phys. Rev. A* **68**, 010301 (2003).
- [47] This one-to-one correspondence between irreps of different groups is the duality.
- [48] W. Fulton and J. Harris, *Representation Theory: A First Course*, Graduate Texts in Mathematics (Springer, Berlin, 1991), Vol. 129.
- [49] I. M. Gelfand and M. L. Zetlin, *Dokl. Akad. Nauk Ser. Fiz.* **71**, 825 (1950).
- [50] M. Moshinsky, *J. Math. Phys.* **4**, 1128 (1963).
- [51] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, New York, 1962).
- [52] An analogous result holds for fermions and the antisymmetric subspace, where one couples irreps with their transpose Young diagrams.
- [53] P. P. Rohde, *Phys. Rev. A* **86**, 052321 (2012).
- [54] A. Hurwitz, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1897**, 71 (1897).
- [55] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Phys. Rev. Lett.* **73**, 58 (1994).
- [56] U. Herzog and J. A. Bergou, *Phys. Rev. A* **70**, 022302 (2004).
- [57] D. E. Littlewood and A. R. Richardson, *Philos. Trans. Royal Soc. London, Ser. A* **233**, 99 (1934).
- [58] B. Kostant, *J. Am. Math. Soc.* **8**, 181 (1995).
- [59] H. de Guise, D. Spivak, J. Kulp, and I. Dhand, *J. Phys. A: Math. Theor.* **49**, 09LT01 (2016).
- [60] S. Aaronson and A. Arkhipov, [arXiv:1309.7460](https://arxiv.org/abs/1309.7460).
- [61] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics (Springer-Verlag, Berlin, 1997).
- [62] J. Wu, H. de Guise, and B. C. Sanders, *Phys. Rev. A* **98**, 013817 (2018).
- [63] C. Dittel, G. Dufour, M. Walschaers, G. Weihs, A. Buchleitner, and R. Keil, *Phys. Rev. Lett.* **120**, 240404 (2018).
- [64] S. Stanisic and P. S. Turner, Data from distinguishability, <https://doi.org/10.5523/bris.3bj7o4rqo2kxd2f8nnfdx5gw6p>.