

Generating entanglement with linear optics

Stasja Stanisc,^{1,2} Noah Linden,³ Ashley Montanaro,³ and Peter S. Turner¹

¹*Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and
Department of Electrical & Electronic Engineering, University of Bristol, UK*

²*Quantum Engineering Centre for Doctoral Training, University of Bristol, UK*

³*School of Mathematics, University of Bristol, UK*

(Dated: September 26, 2017)

Entanglement is the basic building block of linear optical quantum computation, and as such understanding how to generate it in detail is of great importance for optical architectures. We prove that Bell states cannot be generated using only 3 photons in the dual-rail encoding, and give strong numerical evidence for the optimality of the existing 4 photon schemes. In a setup with a single photon in each input mode, we find a fundamental limit on the possible entanglement between a single mode Alice and arbitrary Bob. We investigate and compare other setups aimed at characterizing entanglement in settings more general than dual-rail encoding. The results draw attention to the trade-off between the entanglement a state has and the probability of postselecting that state, which can give surprising constant bounds on entanglement even with increasing numbers of photons.

I. INTRODUCTION

Research into quantum technologies has gained significant momentum in the last several years, with applications ranging across metrology, communications, security, simulation and computation [1–4]. One of the important resources lying behind many of these advances is quantum entanglement [5, 6]. Long before it was a potential technological resource, entanglement was studied as one of the phenomena lying at the foundations of quantum mechanics [7–9]. That there exist non-classical correlations between physical systems is now well established, while how best to generate, verify and quantify such entangled states in practice is an ongoing field of activity. What is practical in any given situation depends on the physical platform under consideration; here we will be interested in the generation of entanglement using linear optics and postselection.

In linear optics we study collections of optical modes, modelled as harmonic oscillators whose excitations correspond to photons. Interactions are restricted to Hamiltonians that leave the total number of photons fixed, giving rise to unitary transformations on modes (interferometers), as well as possible measurement and postselection of quantum states (heralding). This realization introduces an interesting set of constraints on the entanglement problem. Most work to date focuses on either single- or dual-rail encoding of photons into two-dimensional qubits, and then applying the usual approaches to quantum computation such as the circuit model or measurement-based schemes. Gates are carried out via ancilla modes and photon detection measurements [4]. The dual-rail encoding, where qubits are realized as single photons in pairs of spatial or polarization modes, is the commonly accepted standard for quantum computation with linear optics, and allows us to discuss entanglement in terms of standard concepts such as Bell and GHZ states [4, 10, 11]. However, the requirement of postselection means generation of such states is non-

deterministic, and the probability of success is often low; for example, the best known Bell state generation scheme has success probability of $1/4$ [12] and if the postprocessing technique known as procrustean distillation is not allowed, then the probability drops to 0.1875 [13]. When we consider the number of Bell states needed to construct two-dimensional cluster states [11], the requirements can be quite daunting, though promising proposals exist [14].

This helps to motivate the study of entanglement generation in linear optics more generally; in particular, it is natural to consider entanglement between two subsets of modes, foregoing encoding altogether. While this is currently not the preferred way of generating entanglement, any bounds that can be found present fundamental limits on linear optical architectures, as well as for other quantum information processing tasks such as boson sampling [15]. A different perspective on this issue, which considers bosonic entanglement in terms of observables, can be found in (e.g.) [16, 17].

In this paper we will consider two main themes regarding bipartite entanglement in linear optics; that where the parts are encoded qubits, and that where they are collections of modes. Section II introduces the background and notation used throughout. Section III examines qubit entanglement within the standard linear optical dual-rail encoding. When we speak of dual-rail encoding, we mean qubit states that are post-selected such that there is exactly one photon in each pair of modes. First we prove that one cannot generate a Bell state using only 3 photons, and then we give strong numerical evidence for the known 4 photon Bell state generator (with a success probability of 0.1875) being optimal. In Section IV we compare qubit and mode entanglement, including an investigation of the expected average entanglement over uniformly (Haar) distributed interferometers. In Section V we shift our focus to mode entanglement, considering bipartite systems made from two sets of optical modes, Alice and Bob, with a fixed total number of photons. We see two types of behaviour. In the case of bunched photon

Bound (ebits)	Parameters	Input state	Section
$O(\log n)$	$M_A = 1, M_B = 1, M_H = 0$	Bunched	V A 1
2	$M_A = 1, M_B \geq 1, M_H = 0$	Unbunched	V A 2
$\log 3$	$M_A = M_B = 1, M_H \geq 1$	Unbunched	V A 3
$\log \left(2^{\binom{M_A + \frac{n-1}{2}}{M_A}} \right)$	$M_A = M_B, n$ odd	Any	V B 1
$\log \left(2^{\frac{n+M_A}{n} \binom{M_A + \frac{n}{2}-1}{M_A}} \right)$	$M_A = M_B, n$ even	Any	V B 1
n	$M_A = M_B$	Fock state	V B 2

TABLE I. Entanglement bounds proven in this paper. The notation is as defined in Section II (see Figure 1).

input and single mode Alice, we find the entanglement can grow as $\log n$ where n is the number of photons. On the other hand, looking at the case of at most a single photon per input mode (as in, for example, boson sampling [15]), a single mode Alice and no measurement, the entanglement is bounded above by 2 ebits regardless of how many photons are present. If we also restrict Bob to a single mode and furnish the remaining modes with number resolving detectors, the expected entanglement is bounded by $\log 3$ ebits. We then find provable universal bounds on the mode entanglement stemming from the dimensionality of the bipartite Fock states involved, and from the linearity of the optical transformation. Finally, we conjecture a third bound due to unitarity which extends the previously mentioned constant bound in the case of Alice having a single mode to multi-mode Alice, and we provide numerical evidence for this conjecture. The maximum mode entanglement is summarized in Table I.

II. BACKGROUND

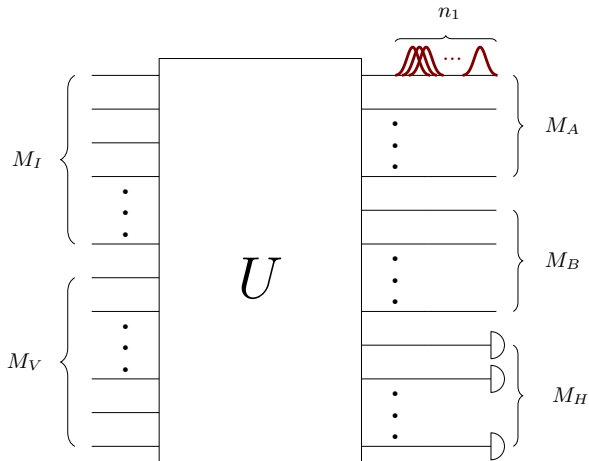


FIG. 1. The generic setup used throughout this paper; see text for an explanation of the notation.

Figure 1 introduces the generic linear optical setup and notation used throughout the paper. The interferometer has M input modes and M output modes. The mode

transformation describing this (photon number preserving) interferometer is an $M \times M$ unitary matrix $U \in U(M)$. The top M_I input modes contain n input photons, while the bottom M_V modes are ancilla vacua. The representation of U carried by the n photon, M mode Hilbert space in the number state (Fock) basis is denoted $\mathcal{U}^{(n)}$. The top M_A output modes belong to one party, Alice, the middle M_B modes belong to Bob, and the bottom M_H modes – Harold – get measured using photon counting detectors. Harold’s detection pattern is labelled $\underline{h} = (n_{M_S+1}, \dots, n_M)$ where n_i gives the photon number of output mode i , and $M_S = M_A + M_B$ is the number of modes in the “system”, i.e. modes that do not belong to Harold and are therefore unmeasured. If $n_H = \sum_{k=M_S+1}^M n_k = \|\underline{h}\|_1$ total photons have been detected, the number of photons left in the system is $n_S = n - n_H = n_A + n_B$. The Hilbert space of subsystem X (a subset of modes), given that it contains exactly n_X photons, is denoted $\mathcal{H}_X^{n_X}$.

Let the input to the interferometer be a Fock state

$$|\psi_{\text{in}}\rangle = |n_1, n_2, \dots, n_{M_I}, \underbrace{0, \dots, 0}_{M_V}\rangle \quad (1)$$

$$= \prod_{k=1}^{M_I} \frac{(\hat{a}_k^\dagger)^{n_k}}{\sqrt{n_k!}} |\text{vac}\rangle, \quad (2)$$

where $|\text{vac}\rangle = |0\rangle^{\otimes N}$. The input transforms according to

$$\mathcal{U} |\psi_{\text{in}}\rangle = \prod_{k=1}^{M_I} \frac{1}{\sqrt{n_k!}} \left(\mathcal{U} \hat{a}_k^\dagger \mathcal{U}^\dagger \right)^{n_k} \mathcal{U} |\text{vac}\rangle \quad (3)$$

$$= \prod_{k=1}^{M_I} \frac{1}{\sqrt{n_k!}} \left(\sum_{j=1}^M \hat{a}_j^\dagger U_{jk} \right)^{n_k} |\text{vac}\rangle, \quad (4)$$

where U_{jk} are the matrix elements of the mode transformation U , \mathcal{U} is the representation of U on the multimode Fock space, and we have used the fact that the vacuum is invariant under all such transformations (as is customary, we have suppressed identity operators on ancillary modes). We will usually be interested in the case of single photon Fock inputs, where $n_k = 1$ or vacuum for all input modes k , a situation we will refer to as unbunched. If all the photons are found in one mode and the rest contain vacuum, we will refer to the state as completely bunched.

When $M_H > 0$ the ideal number resolving detectors will register a detection pattern $\underline{h} = (n_{M_S+1}, \dots, n_M)$ of n_H photons. The output will be the post-measurement state consisting of $n_S = n - n_H$ photons remaining in the system modes $1, \dots, M_S$, given by

$$|\psi_S(\underline{h}, U)\rangle = \frac{\langle \underline{h} | \mathcal{U} | \psi_{\text{in}} \rangle}{\| \langle \underline{h} | \mathcal{U} | \psi_{\text{in}} \rangle \|}. \quad (5)$$

Note that this is a pure state on the system $S = AB$, because $|\underline{h}\rangle$ only has support on subsystem H . We will denote the unnormalized output by $|\tilde{\psi}_S(\underline{h}, U)\rangle = \langle \underline{h} | \mathcal{U} | \psi_{\text{in}} \rangle$. The Hilbert space of the system is

$$\mathcal{H}_S^{n_S} = \bigoplus_{n_A=0}^{n_S} \mathcal{H}_A^{n_A} \otimes \mathcal{H}_B^{n_B}, \quad (6)$$

where $n_B = n_S - n_A$ is the number of photons in Bob's subsystem. We are interested in entanglement with respect to this tensor product structure. The dimension of the Hilbert space of n photons in M modes is $\binom{M+n-1}{n}$, and so

$$\dim \mathcal{H}_S^{n_S} = \sum_{n_A=0}^{n_S} \binom{M_A + n_A - 1}{n_A} \binom{M_B + n_B - 1}{n_B} \quad (7)$$

$$= \binom{M_S + n_S - 1}{n_S} \quad (8)$$

as $M_S = M_A + M_B$ and $n_S = n_A + n_B$. The totality of states available to Alice can be thought of as the Hilbert space $\bigoplus_{n_A=0}^{n_S} \mathcal{H}_A^{n_A}$, and we may index its Fock basis as $\{|\underline{a}\rangle_A : \underline{a} = (n_1, n_2, \dots, n_{M_A}), \|\underline{a}\|_1 = n_A\}$. Similarly for Bob. Expanding the output in this basis, we have

$$|\tilde{\psi}_S(\underline{h}, U)\rangle = \sum_{\underline{a}, \underline{b}} \tilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U) |\underline{a}\rangle_A \otimes |\underline{b}\rangle_B. \quad (9)$$

The coefficients \tilde{C} are related to permanents of the matrix U [15, 18]. More specifically, consider an input Fock state $|\psi\rangle = |n_1 \dots n_M\rangle$ and an output Fock state $|\phi\rangle = |n'_1 \dots n'_M\rangle$ both with a total number of n photons. Construct a new matrix $U_{\psi\phi}$ from U in two steps. First, define the matrix U_ψ consisting of n_j copies of the j -th column of U for all $j \in \{1, \dots, M\}$. Next, construct the matrix $U_{\psi\phi}$ by using n'_j copies of the j -th row of U_ψ for all $j \in \{1, \dots, M\}$. Then

$$\langle \phi | \mathcal{U} | \psi \rangle = \frac{\text{perm}(U_{\psi\phi})}{\sqrt{n_1! \dots n_M! n'_1! \dots n'_M!}}. \quad (10)$$

In our notation, $|\psi\rangle = |\psi_{\text{in}}\rangle$ and $|\phi\rangle = |\underline{abh}\rangle$, we therefore have $\tilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U) = \langle \underline{abh} | \mathcal{U} | \psi_{\text{in}} \rangle$. The probability of detecting pattern \underline{h} is $P(\underline{h}, U) = \sum_{\underline{a}, \underline{b}} |\tilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U)|^2$, and defining $C_{\underline{a}, \underline{b}} = \tilde{C}_{\underline{a}, \underline{b}} / \sqrt{P(\underline{h}, U)}$, the normalized state can be written as $|\psi_S(\underline{h}, U)\rangle = \sum_{\underline{a}, \underline{b}} C_{\underline{a}, \underline{b}}(\underline{h}, U) |\underline{a}\rangle_A |\underline{b}\rangle_B$.

For future convenience we define coefficients of the output states in particle notation, where the Fock state

$|n_1 \dots n_M\rangle$ is written as $|\underbrace{1 \dots 1}_{n_1} \dots \underbrace{M \dots M}_{n_M}\rangle$. We denote relevant coefficients in particle notation by γ , which are related to the above mentioned permanent as

$$\gamma_{1 \dots 1 \dots M \dots M}(\underline{h}, U) = \frac{\tilde{C}_{\underline{a}, \underline{b}}(\underline{h}, U)}{\sqrt{n_1! \dots n_{M_S}!}}. \quad (11)$$

These are the coefficients of the output states as expressed in terms of the creation operators assuming unbunched input to the interferometer, see Eq.(12).

Equation (9) provides a decomposition we can use to bound the entanglement. However, the fact that the total number of photons in the system, n_S , is preserved implies that not all conceivable bipartite basis states $|\underline{a}\rangle_A \otimes |\underline{b}\rangle_B$ are available, so the system should not simply be viewed as the tensor product of two qudits i.e. Eq.(8) is not simply the product of $\dim \mathcal{H}_A$ and $\dim \mathcal{H}_B$. In particular, this means that states that are maximally entangled in the usual sense do not exist. For example, Alice can have many states with n_S photons, but there is only one possible Bob state to which they can be correlated, namely the vacuum (see Section V B 1).

The entanglement measure that will be used is the von Neumann entropy; given a pure state $|\psi_S(\underline{h}, U)\rangle$, its density matrix is defined $\rho_{AB}(\underline{h}, U) = |\psi_S(\underline{h}, U)\rangle \langle \psi_S(\underline{h}, U)|$, and its reduced density matrices on subsystems are the marginals $\rho_A(\underline{h}, U) = \text{Tr}_B[\rho_{AB}(\underline{h}, U)]$ and $\rho_B(\underline{h}, U) = \text{Tr}_A[\rho_{AB}(\underline{h}, U)]$. The von Neumann entropy is then $S(\rho_A(\underline{h}, U)) = -\text{Tr}[\rho_A(\underline{h}, U) \cdot \log \rho_A(\underline{h}, U)] = -\sum_a \lambda_a \cdot \log \lambda_a$ where $\{\lambda_a\}_a$ are the non-zero eigenvalues of the reduced state. Unless stated otherwise, logarithms will be assumed to be base 2. Finally, we will use ebits as the unit of bipartite entanglement where 1 ebit corresponds to the von Neumann entropy of a Bell state.

III. QUBIT ENTANGLEMENT

In this section we will be considering the dual-rail encoding of two qubits. This means that $M_A = M_B = 2$ and states are postselected so that subsystems A and B have exactly one photon each, $n_A = n_B = 1$; all the other states are discarded. (In general, the k -th qubit consists of the modes $2k - 1$ and $2k$ via the mapping $|10\rangle_{2k-1, 2k} \rightarrow |0\rangle_k$, $|01\rangle_{2k-1, 2k} \rightarrow |1\rangle_k$.) Despite the full Hilbert space of the system being of dimension 10 (see Eq. (13)), these constraints limit the space of permissible states to $\dim \mathcal{H}_A = \dim \mathcal{H}_B = 2$, encoding two qubits. To entangle photons in this encoding using only passive linear optics, the use of ancillas and postselection is necessary [4], so $M_H > 0$.

A. Generating Bell states with three photons is impossible

It is known that generating a Bell state in dual-rail encoding with just two photons is impossible [4, 19]. Kieling

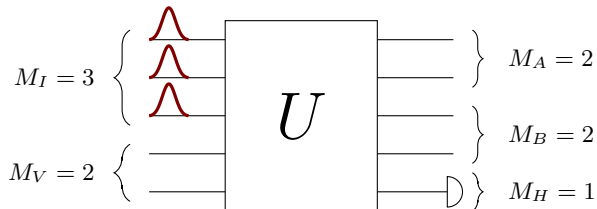


FIG. 2. The setup used in Section III A, with $M_I = n = 3$, $M_V = 2$, $M_A = M_B = 2$, and $M_H = 1$. We show that no such setup can create an entangled state in dual-rail qubit encoding with any non-zero probability. On the other hand, with 4 input photons it is possible to create a Bell state with probability of $1/4$ [13].

observed it is also impossible with three photons, using an algebraic geometry approach to the problem [19]. Here we offer an explicit proof that not only is it impossible with three photons, it is only possible to create product states.

Proposition III.1. *In a passive linear optical setup using dual-rail encoding, ancillas and postselection, it is not possible to create an entangled state using 3 photon input.*

Proof. First, let us consider the case where there are five modes ($M = 5$); four system modes ($M_A + M_B = 4$) and one ancilla ($M_H = 1$), as illustrated in Figure 2. Let the input be three unbunched photons ($n = M_I = 3$). Dual-rail encoding has a total of two photons in a valid qubit state output ($n_S = 2$), implying here that one photon is detected ($n_H = 1$). As there is only one measurement ancilla, the only possible measurement pattern is $\underline{h} = (1)$ (one photon in the fifth mode).

As discussed in Sec. II, the amplitudes are related to the permanents of the matrix U :

$$\gamma_{kj}((1), U) = \begin{cases} \frac{1}{2} \sum_{\sigma \in S_3} U_{k,\sigma(1)} U_{k,\sigma(2)} U_{5,\sigma(3)}, & k = j \\ \sum_{\sigma \in S_3} U_{k,\sigma(1)} U_{j,\sigma(2)} U_{5,\sigma(3)} & k \neq j \end{cases} \quad (12)$$

defined $\forall k, j \in \{1, 2, 3, 4\}$. The unnormalized state following detection is

$$\begin{aligned} |\tilde{\psi}((1), U)\rangle &= \sqrt{2}\gamma_{11} |2000\rangle + \sqrt{2}\gamma_{22} |0200\rangle \\ &+ \sqrt{2}\gamma_{33} |0020\rangle + \sqrt{2}\gamma_{44} |0002\rangle \\ &+ \gamma_{13} |1010\rangle + \gamma_{24} |0101\rangle \\ &+ \gamma_{12} |1100\rangle + \gamma_{34} |0011\rangle \\ &+ \gamma_{14} |1001\rangle + \gamma_{23} |0110\rangle, \end{aligned} \quad (13)$$

occurring with probability $P((1), U) = 2 \sum_{k=1}^4 |\gamma_{kk}|^2 + \sum_{\substack{k,j=1 \\ k \neq j}}^4 |\gamma_{kj}|^2$.

In dual-rail encoding it is possible to do any local unitary deterministically by adding beamsplitters and phase shifters to each of the qubits [4]. Thus it suffices to show that it is not possible to create any state of the form $\alpha |0\rangle_A |0\rangle_B + \beta |1\rangle_A |1\rangle_B$ where $|\alpha|^2 + |\beta|^2 = 1$ and $\alpha \neq 0$,

$\beta \neq 0$, because any entangled pure state can be transformed into one of this form by local unitary operations. The coefficients must therefore satisfy

$$\gamma_{11} = \gamma_{22} = \gamma_{33} = \gamma_{44} = 0, \quad (14)$$

$$\gamma_{12} = \gamma_{14} = \gamma_{23} = \gamma_{34} = 0 \quad \text{and} \quad (15)$$

$$|\gamma_{13}| = \alpha\sqrt{p}, |\gamma_{24}| = \beta\sqrt{p}, \quad (16)$$

where $p = P((1), U)$, the probability of one photon being detected in the last mode. We will now try to find a unitary U that satisfies these constraints. Define $K_k := U_{k2}U_{53} + U_{k3}U_{52}, \forall k \in \{1, \dots, 4\}$.

First, let us consider the case where at least one of U_{51} , U_{52} and U_{53} is 0. Without loss of generality (wlog) we can label modes so that $U_{51} = 0$, because we can swap A for B and mode 1 for 2 without affecting entanglement. Then the equations in (12) can be rewritten as $\gamma_{kk} = U_{k1}K_k$ and $\gamma_{kj} = U_{k1}K_j + U_{j1}K_k$ for $k \neq j$. Since $\gamma_{11} = U_{11}K_1 = 0$ and $\gamma_{13} = U_{11}K_3 + U_{31}K_1 \neq 0$, then one and only one of U_{11} or K_1 can be equal to 0. First, assume that $U_{11} = 0$. Since $K_1 \neq 0$, from the constraints $\gamma_{12} = U_{21}K_1 = \gamma_{14} = U_{41}K_1 = 0$ and $\gamma_{24} = U_{21}K_4 + U_{41}K_2 \neq 0$, we see that there is no solution. Similarly, if $K_1 = 0$, then $U_{11} \neq 0$ and the constraints $\gamma_{12} = U_{11}K_2 = \gamma_{14} = U_{11}K_4 = 0$ and $\gamma_{24} = U_{21}K_4 + U_{41}K_2 \neq 0$ again results in no solution. Therefore there is no solution for which at least one of U_{51} , U_{52} , U_{53} is zero.

Next we assume $K_k \neq 0 \forall k \in \{1, \dots, 4\}$, with $U_{51}U_{52}U_{53} \neq 0$. Then solving for U_{k1} from $\gamma_{kk} = 0$ we get $U_{k1} = -U_{k2}U_{k3}U_{51}/K_k, \forall k \in \{1, \dots, 4\}$. Substituting this into the expression for γ_{kj} we get

$$\gamma_{kj} = \frac{U_{51}U_{52}U_{53}(U_{k2}U_{j3} - U_{j2}U_{k3})^2}{K_k K_j}, \quad (17)$$

for all $k, j \in \{1, \dots, 4\}, k \neq j$. The only way $\gamma_{12} = \gamma_{23} = 0$, is if $U_{12}U_{23} = U_{22}U_{13}$ and $U_{22}U_{33} = U_{32}U_{23}$. If $U_{22}U_{23} \neq 0$ then $U_{12}U_{33} = U_{13}U_{32}$, which means $\gamma_{13} = 0$ also, thus cannot be a solution. If only one of U_{22} or U_{33} is zero, assume $U_{2j} = 0$ where j is 2 or 3. But then $U_{1j} = U_{3j} = 0$ and again $\gamma_{13} = 0$. If both are zero, then $\gamma_{24} = 0$. Therefore, there is no solution with $K_k \neq 0 \forall k \in \{1, \dots, 4\}$.

Lastly, assume that at least one of the $K_k = 0$ and that $U_{51}U_{52}U_{53} \neq 0$; wlog, $K_1 = 0$. Then $U_{12} = -U_{13}U_{52}/U_{53}$ combined with the constraint $\gamma_{11} = U_{12}U_{13}U_{51} = 0$ means $U_{12} = U_{13} = 0$. This gives $\gamma_{1j} = U_{11}K_j, \forall j \in \{1, \dots, 4\}$. Since $\gamma_{12} = \gamma_{14} = 0$ and $\gamma_{13} \neq 0$, then $U_{11} \neq 0$, while $K_2 = K_4 = 0$. However, this implies $U_{22} = U_{23} = 0$ by a similar argument, further implying that $\gamma_{24} = 0$ and hence there is no solution.

We see that under no conditions is there a solution to the given equations where $\alpha \neq 0$ and $\beta \neq 0$.

This proves the claim for 5 modes. To see that it is true for any number of vacuum ancillas, notice that as long as there are no photons added, Eqs. (12) do not change other than the mode number 5 being replaced with the new detection ancilla. Each new case therefore gives rise

to the same constraints implied by Eqs. (16), with a lack of solution in the same way. Thus, vacuum ancillas can only increase the probability of creating a state if that probability was nonzero in the first place.

Finally, if we allow inputs other than completely unbunched, Eqs. (12) become even more restrictive. For example, if there were two photons in input mode 1 and one photon in input mode 2, then the matrix elements U_{i3} , U_{3i} would not appear in Eqs. (12), serving only to make the constraints harder to satisfy. \square

Corollary III.2. *In a passive linear optical setup using dual-rail encoding, ancillas and postselection, it is not possible to create a Bell state using 3 photon input.*

B. Optimal Bell state generation

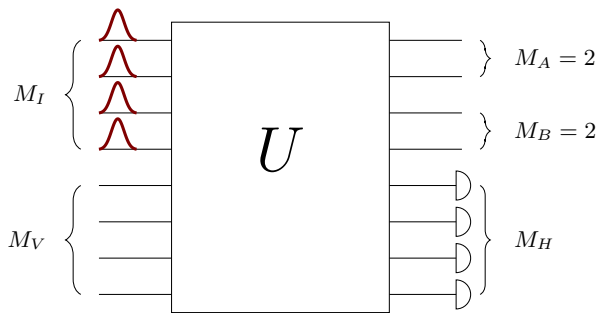


FIG. 3. The setup used in Section IIIB with four photons in eight modes; $M_I = n = 4$, $M_V = 4$, $M_A = M_B = 2$, $M_H = 4$. We give extensive numerical evidence for optimal Bell state generation using this setup when looking for specific Bell states as output.

The previous section showed that Bell state generation with non-zero success probability requires at least four photons. Two schemes which accomplish this task using four photons use six [10] and eight [12, 13] modes, with success probabilities of $2/27$ and $1/4$ respectively.

We performed a numerical search for a linear optical Bell state generator that gives a higher success probability. We used a gradient descent based optimization algorithm over $M = 8$ unitaries with $n = 4$ photon input. Numerical optimization was carried out in Python, using the BFGS algorithm from the SciPy library [20]. This algorithm finds local minima so it needs to be run many times with different seed unitaries, which were randomly selected according to the Haar measure.

The cost function we consider is based on the overlap with the desired Bell states. We allow for six different Bell states, which in the Fock basis after measurement correspond to $|B_{1,2}\rangle = (|1010\rangle \pm |0101\rangle)/\sqrt{2}$, $|B_{3,4}\rangle = (|1001\rangle \pm |0110\rangle)/\sqrt{2}$ and $|B_{5,6}\rangle = (|1100\rangle \pm |0011\rangle)/\sqrt{2}$, where the latter can be corrected to the usual dual-rail qubit encoding using a switch [13]. After detecting measurement pattern \underline{h} , the overlap between each of these

states with the post-selected state is calculated. We found that raising the overlap to the exponent 10 optimized the numerical efficacy, penalizing states far from a Bell state heavily. Multiplying by the probability of detection gives the target cost function to be minimized, $f(\underline{h}, U) = -\sum_{\underline{h}} P(\underline{h}, U) \sum_{k=1}^6 |\langle B_k | \psi(\underline{h}, U) \rangle|^{10}$.

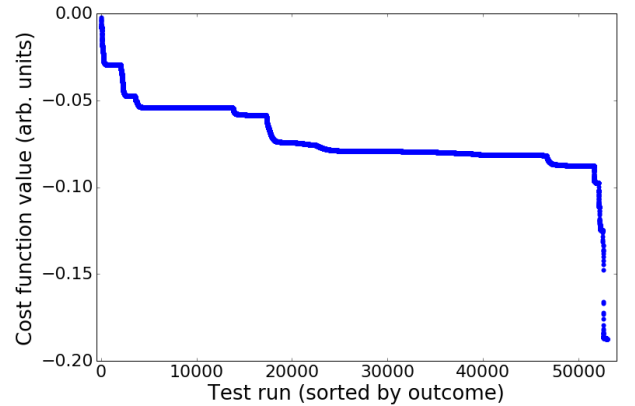


FIG. 4. Results of optimization looking for interferometers that generate Bell states with highest probability. The minimum found of ≈ -0.1875 is exactly bounded by the values of cost function for the known U_{Bell} interferometer as described in the text. Out of 53,000 test runs, 0.75 % of minima found were within 0.001 of the minimum corresponding to U_{Bell} , i.e. within numerical error. Besides the trials depicted in this graph, the cost function was also optimized with other parameters given to the optimizing algorithm as well as over the space of orthogonal matrices. Thus the number of test runs for which a better solution could not be found is more than 150,000.

Figure 4 shows the results of this minimization. The optimal known scheme, when evaluated for this cost function, gives a value of approximately -0.1875 . It produces one of these 6 Bell states with probability $1/32$ for 6 out of the 10 possible measurement patterns [13]. We can see from the figure that the minimum achieved by the numerical optimization over 53,000 trials is also approximately -0.1875 , thus giving solutions which are equivalent to the known scheme in terms of this parameter. While not a proof, this numerical evidence strongly suggests that the known scheme is optimal for generating the above set of Bell states. Other cost functions were also attempted, as well as other optimization libraries, but all gave the same results as the technique above.

We also investigated the case of non-orthogonal Bell states; for example, allowing $|00\rangle + |11\rangle$ as well as $|00\rangle + i|11\rangle$ as target states. The possibility of both of these states being generated from the same U for different measurement patterns was explored by running similar numerical optimizations rewarding such situations. We found no such unitary, which is an interesting result in itself.

Though the complexity of the problem grows quickly, we also looked at how the situation changes with higher numbers of input photons and modes. We numerically optimized over $n = 5$, $M = 10$ using a similar algorithm and no improved solutions were found over 5000 runs. Similarly, we checked $n = 6$, $M = 12$ over 1000 runs and here as well there was no improvement over the -0.1875 result for $n = 4$, $M = 8$.

IV. RANDOM UNITARIES

In this section, we move from the dual-rail qubit encoding of Section III to mode entanglement in Section V. First, we look at how much mode entanglement can be generated with random elements of the unitary group, which we can then use to compare with the dual-rail encoding. We do so by setting Alice and Bob's number of modes to 2, and numerically computing the average amount of entanglement over measurement patterns. Notice that this is different from the setting in Section III, where we aimed to get a maximally entangled Bell state with the highest possible probability. Here and in the rest of this work we will study this average entanglement, namely

$$\langle S(U) \rangle_H = \sum_{\underline{h}} P(\underline{h}, U) S(\rho_A(\underline{h}, U)). \quad (18)$$

The expectation over the unitary group (for fixed M and n) is then $\langle S \rangle_{H,U} = \int_{U(M)} dU \langle S(U) \rangle_H$, where dU is the normalized Haar measure.

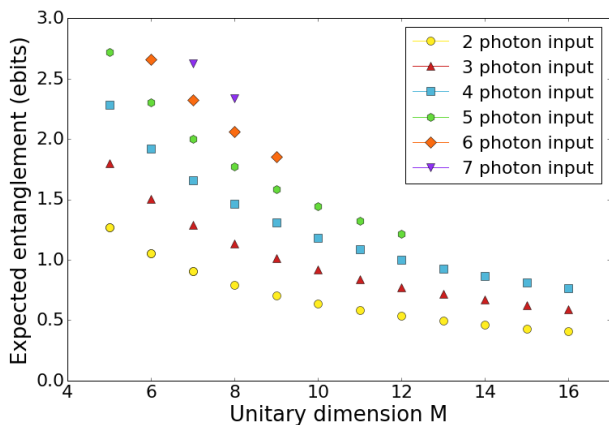


FIG. 5. The expectation, over the unitary group, of the average, over measurement patterns, mode entanglement versus the number of modes M , for various numbers of unbunched input photons. $M_A = M_B = 2$, and if the number of photons n is smaller than M , vacuum input modes are added. The number of heralding detectors is $M_H = M - M_A - M_B$. The entanglement for a single unitary U is averaged over all measurement patterns, and subsequently averaged over 100,000 randomly Haar-sampled unitaries U . Colours and symbols represent different number of input photons, with $2 \leq n \leq 7$.

Figure 5 shows the numerical results. We notice that often the average is higher than 1 ebit, which is the maximum we can achieve in dual-rail qubit encoding. Adding input photons for the same M increases the average entanglement, while adding vacuum ancillas decreases it. We see that the average entanglement of $n + 1$ photons in $M + 1$ modes can be lower than that for M and n (see $n = M = 5$ and 6). That is, we do not expect more average entanglement by adding a photon at the cost of adding another mode. Further, we note that even with 2 photons, there is more average entanglement generated than in the optimal Bell state generator with 4 photons. We explore this in more detail for a better comparison.

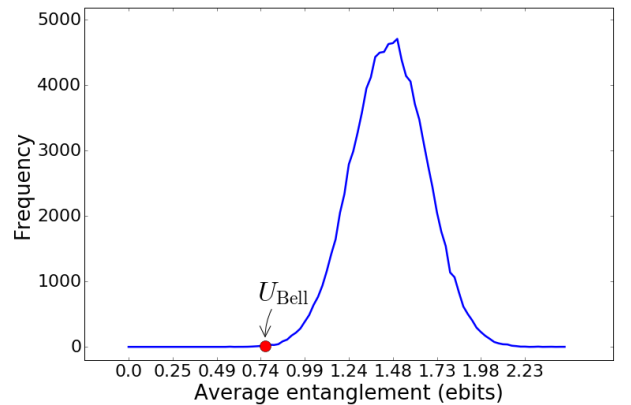


FIG. 6. Numerical evaluation of $\langle S(U) \rangle_H$ for 100,000 unitaries U chosen using the Haar measure in the case $M_A = M_B = 2$, $M_H = 4$, and $n = 4$ unbunched input photons. Average entanglement for a given U was calculated according to Equation (18) and then binned in one of 100 bins with a minimum of 0 and maximum obtained in the samples. The red dot marks the value of average entanglement that the Bell generating unitary from Section III B can give, denoted as U_{Bell} , if all of its output states were used.

In the usual Bell state generation scenario discussed in Section III B, if the measurement outcome indicates that the output state is outside of the qubit subspace, the output is discarded. Here we include the entanglement of the discarded states in accordance with Eq. (18). We compare the optimal Bell state generator to random unitaries with the same parameters; $M_A = M_B = 2$, $n = 4$ and $M = 8$.

In Fig. 6 we see the results of the comparison. Firstly, in Section III B we saw that the probability of getting a Bell state for a state correctable with a single switch is $3/16$ [13]. A Bell state gives a single ebit, and if all the other states are discarded, the average entanglement would be 0.1875 ebits. If all the outputs from this unitary were counted towards average entanglement as discussed in the previous paragraph (where Equation 18 is utilized), the entanglement obtained is marked on the Figure 6 as U_{Bell} . As we can see from the graphs, U_{Bell} gives a markedly lower amount of entanglement than what could

be generated on average with a random unitary on the same number of modes.

V. MODE ENTANGLEMENT

The previous section shows that, on average, random unitaries give significantly more mode entanglement than dual-rail encoding. We therefore turn our attention to the investigation of mode rather than qubit entanglement as defined in Section II.

Equation (6) states that the total system Hilbert space is a direct sum of Hilbert subspaces such that the sum of Alice and Bob's photon numbers is n_S , the number of photons left after heralding. Let $\rho_{AB} = |\psi_S(\underline{h}, U)\rangle \langle \psi_S(\underline{h}, U)|$ as in Eq. (5). Alice's reduced density matrix is

$$\begin{aligned} \rho_A(\underline{h}, U) &= \text{Tr}_B[\rho_{AB}(\underline{h}, U)] \\ &= \sum_{\underline{b}''} \langle \underline{b}'' | \left(\sum_{\underline{a}, \underline{b}, \underline{a}', \underline{b}'} C_{\underline{a}\underline{b}} \bar{C}_{\underline{a}'\underline{b}'} |\underline{a}\underline{b}\rangle \langle \underline{a}'\underline{b}'| \right) | \underline{b}'' \rangle \\ &= \sum_{\underline{a}, \underline{a}'} \left(\sum_{\underline{b}} C_{\underline{a}\underline{b}} \bar{C}_{\underline{a}'\underline{b}} \right) |\underline{a}\rangle \langle \underline{a}'|, \end{aligned} \quad (19)$$

where only the terms with $\|\underline{a}\|_1 = \|\underline{a}'\|_1$ are non-zero, because $\|\underline{b}\|_1 = \|\underline{b}'\|_1 = \|\underline{b}''\|_1$ and $n_S = \|\underline{a}\|_1 + \|\underline{b}\|_1 = \|\underline{a}'\|_1 + \|\underline{b}'\|_1$. Therefore, there exists a Fock basis ordering in which Alice's reduced state is block diagonal, which allows us to derive a bound on the entanglement (see Section VB 1). In the case that Alice has a single mode, this implies her state is diagonal in Fock basis. The total number of orthogonal states available to Alice is

$$\dim(\mathcal{H}_A^{n_S}) = \sum_{n_A=0}^{n_S} \binom{M_A + n_A - 1}{n_A} = \binom{M_A + n_S}{n_S}. \quad (21)$$

In Section V A, we find entanglement bounds when Alice only has one mode. The bound depends on the input state; if the input photons are bunched in a single mode, entanglement is unbounded as the number of photons increases. Surprisingly, if the input is unbunched, we find a constant bound independent of the number of Bob's modes and independent of the number of photons. More general bounds can be found, though they are also more loose. In Section VB 1 we give the bound on entanglement due to the block diagonal structure of Alice's reduced density matrix in Fock basis. In Section VB 2 we give a bound which is a consequence of the linearity of the mode transformations. Unlike in Sec. V A, neither of these bounds depend on the unitarity of the mode transformations, which we expect should affect the amount of entanglement that can be achieved. In Section V C we conjecture a general unitarity bound based on numerical evidence.

A. Entanglement when Alice has a single mode

1. Entanglement for bunched input can be unbounded

First, we show that mode entanglement is unbounded if we are not restricted to unbunched input.

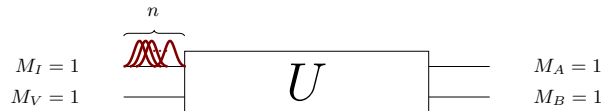


FIG. 7. The setup used in Section V A 1, where we consider only $M = 2$ modes. The input consists of all n photons bunched in the top mode; $M_I = M_V = M_A = M_B = 1$, $M_H = 0$. We prove that in this setup maximal entanglement grows as $\log n$. The special case where U is a balanced beamsplitter was analyzed in [21].

Proposition V.1. *Let the input into a $M = 2$ interferometer consist of n photons bunched in a single mode (see Figure 7). Then the entanglement across the two output modes is at most $O(\log n)$ ebits, which is achieved when U is a balanced beamsplitter.*

Proof. Parameterize the $M = 2$ unitary matrix U acting on Alice and Bob's single mode Hilbert spaces as

$$U = \begin{bmatrix} c & d \\ -d^* & c^* \end{bmatrix}, \quad (22)$$

where $|c|^2 + |d|^2 = 1$. The output state is

$$\begin{aligned} |n0\rangle &= \left(\hat{a}_1^\dagger \right)^n / \sqrt{n!} |0\rangle \\ &\mapsto \left(c\hat{a}_1^\dagger - d^*\hat{a}_2^\dagger \right)^n / \sqrt{n!} |0\rangle \\ &= \frac{1}{\sqrt{n!}} \sum_{k=0}^n \binom{n}{k} (c\hat{a}_1^\dagger)^k (-d^*\hat{a}_2^\dagger)^{n-k} |0\rangle \\ &= \frac{1}{\sqrt{n!}} \sum_{k=0}^n \binom{n}{k} c^k (-d^*)^{n-k} \sqrt{k!} \sqrt{(n-k)!} |k\rangle |n-k\rangle \\ &= \sum_{k=0}^n \sqrt{\binom{n}{k}} c^k (-d^*)^{n-k} |k\rangle |n-k\rangle. \end{aligned} \quad (23)$$

When Alice has only one mode, her reduced density matrix is diagonal in the Fock basis, so we can find the spectrum of her state directly from the above equations:

$$\lambda_k = \binom{n}{k} (|c|^2)^k (|d|^2)^{n-k} = \binom{n}{k} (|c|^2)^k (1 - |c|^2)^{n-k}. \quad (24)$$

This is a binomial distribution with a 'success' probability of $p = |c|^2$. The entropy of the binomial distribution for a fixed p is $1/2 \log_2 (2\pi en \cdot p \cdot (1-p)) + O(1/n)^*$.

* From, e.g., the de Moivre-Laplace Theorem

Thus we see that the entanglement bound is $O(\log n)$, where n is the number of photons. The constant prefactor is maximized for $p = |c|^2 = |d|^2 = 1/2$, whence the entropy of Alice's state is $1/2 \log_2(2\pi en \cdot 1/2 \cdot (1 - 1/2)) + O(1/n) = 1/2 \log_2(\pi en/2) + O(1/n)$. Finally, notice that solutions to Equation (22) where $|c|^2 = |d|^2 = 1/2$ are a family of balanced beamsplitters. \square

This is in stark contrast to the situation where the input is unbunched, where we will see in the next section that the entanglement is bounded by a constant.

2. Entanglement for unbunched input is bounded

We now consider situations where Alice only has one mode, Bob can have many, and we do not use any measurement. The following Lemma will be of use.

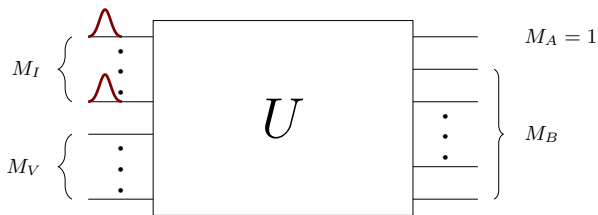


FIG. 8. The setup used in Section V A 2. The input is an unbunched state with $M_I = n$, with $M_V \geq 0$, $M_A = 1$, $M_B \geq 1$ and $M_H = 0$. We prove that entanglement for this setup is bounded by a constant.

Lemma V.2. Consider inputting a Fock state $|\underline{n}\rangle = |n_1 \dots n_M\rangle$ into an arbitrary interferometer that has M modes. Let $N = \max\{n_1, \dots, n_M\}$. Then the mean photon number in each output port is bounded by N [22].

Proof. Let $|\underline{n}\rangle$ be an arbitrary Fock state.

$$\begin{aligned}
 \langle \hat{n}_j \rangle &= \langle \underline{n} | \mathcal{U}^\dagger \hat{n}_j \mathcal{U} | \underline{n} \rangle \\
 &= \langle \underline{n} | \mathcal{U}^\dagger \hat{a}_j^\dagger \mathcal{U} \mathcal{U}^\dagger \hat{a}_j \mathcal{U} | \underline{n} \rangle \\
 &= \langle \underline{n} | \left(\sum_{j'} \hat{a}_{j'}^\dagger \overline{U_{jj'}} \right) \left(\sum_{j''} \hat{a}_{j''} U_{jj''} \right) | \underline{n} \rangle \\
 &= \sum_{j'} \sum_{j''} \overline{U_{jj'}} U_{jj''} \langle \underline{n} | \hat{a}_{j'}^\dagger \hat{a}_{j''} | \underline{n} \rangle \\
 &= \sum_{j'} \overline{U_{jj'}} U_{jj'} n_{j'} \tag{25}
 \end{aligned}$$

If, as hypothesized, $n_j \leq N$ for all modes j , then

$$\langle \hat{n}_j \rangle = \sum_{j'} |U_{j'j}|^2 n_{j'} \leq \sum_{j'} |U_{j'j}|^2 N = N. \tag{26}$$

\square

In the following calculations we shall assume that $n \rightarrow \infty$ as any bound on the entropy found for this infinite case would also hold for a finite one with the same set of constraints.

Lemma V.3. Let $\{p_j\}_{j=0}^\infty$ be a probability distribution subject to the constraint $\sum_j j p_j \leq N$. Then the entropy of this distribution is at most $\log((1+N)^{1+N}/N^N)$.

Proof. The entropy of the probability distribution $\{p_j\}_{j=0}^\infty$ is $S = -\sum_{j=0}^\infty p_j \log p_j$. We maximize this subject to the constraints $\sum_{j=0}^\infty j p_j = \bar{n} \leq N$ and $\sum_{j=0}^\infty p_j = 1$ using the method of Lagrange multipliers.

Let the Lagrangian be

$$L = S + (\lambda_0 + \log e) \left(\sum_{j=0}^\infty p_j - 1 \right) + \lambda_1 \left(\sum_{j=0}^\infty j p_j - \bar{n} \right). \tag{27}$$

Then $\partial L / \partial p_j = -\log p_j + \lambda_0 + \lambda_1 j$. Setting $\partial L / \partial p_j = 0$ gives $p_j = 2^{\lambda_0 + \lambda_1 j}$. Substituting the value of p_j into the constraints, we get

$$\sum_{j=0}^\infty j p_j = 2^{\lambda_0} 2^{\lambda_1} / (1 - 2^{\lambda_1})^2 = \bar{n} \tag{28}$$

$$\sum_{j=0}^\infty p_j = 2^{\lambda_0} / (1 - 2^{\lambda_1}) = 1 \tag{29}$$

This allows us to solve for λ_0 and λ_1 , giving

$$\lambda_0 = \log[1/(1 + \bar{n})], \quad \lambda_1 = \log[\bar{n}/(1 + \bar{n})]. \tag{30}$$

Notice that

$$\begin{aligned}
 S &= -\sum_j p_j \log p_j = -\sum_j p_j (\lambda_0 + \lambda_1 j) \\
 &= -\lambda_0 - \lambda_1 \bar{n} \\
 &= \log\left((1 + \bar{n})^{1 + \bar{n}} / \bar{n}^{\bar{n}} \right) \tag{31}
 \end{aligned}$$

The function above increases monotonically for $\bar{n} \geq 0$ and since $\bar{n} \leq N$ we get

$$S \leq \log((1 + N)^{1+N} / N^N). \tag{32}$$

\square

Corollary V.4. Let $\{p_j\}_{j=0}^\infty$ be some probability distribution subject to the constraint $\sum_j j p_j \leq N$, $N \in [0, 1]$. Then the entropy of this distribution is at most 2 ebits.

Theorem V.5. Let Alice have one output mode, $M_A = 1$, and Bob have $M_B = k$. Let the input be a single photon in each of the $k+1$ modes. Then the entanglement between Alice and Bob is bounded by 2 ebits for all k .

Proof. Alice's reduced density matrix is diagonal in the Fock basis, where each entry $\langle j | \rho_A | j \rangle$ corresponds to the probability that Alice's mode contains j photons.

\square

By Lemma V.2, this distribution satisfies the conditions of Corollary V.4. Thus the von Neumann entropy of this state is bounded by 2 for any k , as the bound which holds for $k \rightarrow \infty$ also holds for any finite k as well. \square

Notice that extra vacuum modes will not increase this limit on the entanglement as the limit is due to the expected number of photons in Alice's mode being at most 1. We see that despite the fact that the dimension of Alice's Hilbert space grows with the number of photons as $n + 1$, and Bob's can be even larger, the maximum entanglement is severely constrained to be less than 2 ebits.

Because we are interested in the average entanglement, the result will hold for heralding as well:

Corollary V.6. *Let Alice have one output mode, $M_A = 1$, while Bob and Harold have $M_B + M_H = k$. Let the input be a single photon in each of the $k + 1$ modes. Then the entanglement between Alice and Bob is bounded by 2 ebits for all k .*

Proof. No LOCC operation can increase the amount of entanglement in the system on average [23]. Therefore, $\langle S(U) \rangle_H = \sum_{\underline{h}} P(\underline{h}, U) S(\rho_A(\underline{h}, U)) \leq S(\rho_A(U))$, where $\rho_A(U)$ is Alice's reduced density matrix before any measurement, and by Theorem V.5, $S(\rho_A(U)) \leq 2$ ebits. \square

We can also examine inputs that have different numbers of bunched photons. If the highest number of photons in a single input mode is N , as per Lemma V.2, the expected number of photons in Alice's mode will then be bounded by N . Because the function which bounds the entropy, Eq. (32), is monotonically increasing, the entropy of Alice's (diagonal) state (p_0, \dots, p_n) is at most $\log((1 + N)^{1+N}/N^N)$ by Lemma V.3. In the extreme case where all the photons are bunched in a single mode, S scales as $O(\log(N + 1))$, consistent with Proposition V.1.

3. Entanglement when Bob also has a single mode

In this section we consider a similar setup to the previous section, except now we fix the number of Bob's modes to 1 and assign the rest to Harold. Recall that we are interested in generating the highest amount of entanglement between Alice and Bob on average, thus the probability of detection patterns must be taken into account. More precisely, we are looking for the maximum of $\langle S(U) \rangle_H = \sum_{\underline{h}} P(\underline{h}, U) S(\rho_A(\underline{h}, U))$. Some patterns might yield a state with high entanglement, but be very unlikely to occur. In a practical setting we might prefer states that are less entangled but we can generate more consistently.

We first prove a technical lemma that will be useful later.

Lemma V.7. *Given a probability distribution (p_0, \dots, p_n) such that $\sum_{j=0}^n j p_j = 1$, the sum*

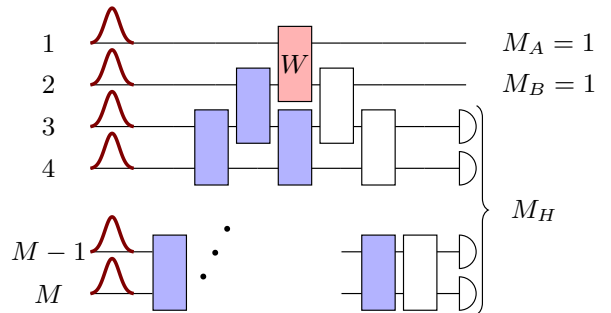


FIG. 9. The setup used in Section V A 3, where $M = n$, $M_A = M_B = 1$, and $M_H \geq 1$. An arbitrary M mode interferometer can be decomposed into $M(M + 1)/2$ two-mode interferometers [24, 25]. Note that this also applies to an arbitrary $M - 1$ mode sub-interferometer (blue). By focusing on the only component that entangles Alice and Bob (red), we show that the maximum entanglement is the $M = n = 2$ value of $\log 3$ ebits.

$\sum_{j=0}^n p_j \log(j + 2)$ is bounded by $\log 3$ which can be achieved by $p_1 = 1$ and $p_k = 0$ for $k \in \{0, 2, 3, \dots, n\}$.

Proof. Since $f(x) = \log(x + 2)$ is a concave function, by Jensen's inequality $\sum_{j=0}^n p_j f(j) \leq f(\sum_{j=0}^n p_j j) = f(1) = \log 3$, which is achieved by substituting $p_1 = 1$ and $p_k = 0$ for $k \in \{0, 2, 3, \dots, n\}$. \square

Theorem V.8. *Consider an interferometer with $M \geq 3$ modes, where both Alice and Bob have one mode and the other output modes are measured using photon counting detectors. Let the input be the $n = M$ unbunched Fock state. Then the maximal average entanglement that can be created between Alice and Bob is $\log 3$ ebits.*

Proof. First, notice that the average entanglement achievable by an $M = 2$ interferometer can be achieved for $M \geq 2$ by having modes 3 to M transform trivially, since photons in these modes will be detected with unit probability. Thus $\max \langle S(U_M) \rangle_H \geq \max \langle S(U_{M=2}) \rangle_H = \log 3$ ebits, $\forall M \geq 3$. The interferometer given in Section V D, Eq. (43) below achieves this.

Any $U \in U(M)$ can be decomposed as in Fig. 9. Then the bottom left triangle (colored blue in the figure) is a unitary $V \in U(M - 1)$. Since the input is unbunched, Lemma V.2 implies that each output from V has a mean photon number of 1. In particular, Bob's mode before beam splitter W (red in the figure) will satisfy $\sum_k k q_k = 1$ where k is the number of photons occurring with probability q_k . Since the remaining beam splitters (white in the figure) act only on Bob and Harold's systems, they have no effect on Alice's reduced state and can therefore be ignored.

Let the probability of detecting pattern \underline{h} be $p_{\underline{h}}$, and the probability of detecting a total of $n_H = \|\underline{h}\|_1$ photons

be $p_{n_H} = \sum_{\underline{h}: \|\underline{h}\|_1 = n_H} p_{\underline{h}}$. The average entanglement is

$$\begin{aligned}
\langle S(U) \rangle_H &= \sum_{\underline{h}} p_{\underline{h}} S(\rho_A(\underline{h})) \\
&= \sum_{n_H=0}^n p_{n_H} \sum_{\underline{h}: \|\underline{h}\|_1 = n_H} p_{\underline{h}}/p_{n_H} S(\rho_A(\underline{h})) \\
&\leq \sum_{n_H=0}^n p_{n_H} \sum_{\underline{h}: \|\underline{h}\|_1 = n_H} p_{\underline{h}}/p_{n_H} \log(n - n_H + 1) \\
&= \sum_{n_H=0}^{n-1} p_{n_H} \log(n - n_H + 1), \tag{33}
\end{aligned}$$

where we've used the fact that the entanglement of $S(\rho_A(\underline{h}))$ is upper bounded by the Schmidt rank $\log(n - n_H + 1)$.

As the photon number found in modes 1 and 2 is set before the beamsplitter W , if n_H photons have been detected, then there were already n_H photons in modes 3 through M . Alice contributes one photon through her mode to their joint system, which implies that Bob must contribute $n - n_H - 1$ photons through mode 2, occurring with probability q_{n-n_H-1} . Therefore $p_{n_H} = q_{n-n_H-1}$ and recall that Bob's probability distribution is constrained by $\sum_{k=0}^{n-1} k q_k = 1$. By Lemma V.7 $\sum_{n_H=0}^{n-1} q_{n-n_H-1} \log(n - n_H + 1) = \sum_{j=0}^{n-1} q_j \log(j + 2)$ is maximized for $j = n - n_H - 1 = 1$, that is $q_1 = 1$, yielding $\langle S(U) \rangle_H \leq \log 3$. This also implies that $n_H = n - 2$ photons are detected in the optimal situation. \square

Note that this agrees with the bound in Theorem V.5 found in the previous section, which follows from the entanglement measure property $\langle S(U) \rangle_H = \sum_{\underline{h}} P(\underline{h}, U) S(\rho_A(\underline{h}, U)) \leq S(\rho_A(U))$, where $\rho_A(U)$ is Alice's reduced density matrix before any measurement. Here the maximum entanglement is $\log 3 < 2$ ebits. Moreover, adding more vacuum input modes will not affect this bound, as this would only change Bob's expected number of photons before the beamsplitter W to be at most 1 instead of exactly 1 as per Lemma V.2.

B. Entanglement when Alice has many modes

In this section we give two bounds on entanglement for more general situations when Alice has more than one mode, based on the Schmidt rank of Alice's reduced state. They are independent of the input state or any interferometer transformation, depending only on the given number of photons and modes; we assume the latter is the same for both Alice and Bob. This generality comes at a price, however, in that the bounds loosen; we will discuss a conjectured tighter bound in the following section.

1. Dimensionality

By looking solely at the dimensions of Alice and Bob's Hilbert spaces, we can derive an entanglement bound as follows.

Proposition V.9. *Let Alice's and Bob's joint postselected state have a total of n_S photons. Let Alice and Bob have $M_A = M_B$ modes. The Schmidt rank, ω , is at most*

$$\omega = 2 \binom{M_A + \frac{n_S - 1}{2}}{\frac{n_S - 1}{2}} \quad n_S \text{ odd}, \tag{34}$$

$$\omega = 2 \frac{n_S + M_A}{n_S} \binom{M_A + \frac{n_S}{2} - 1}{\frac{n_S}{2} - 1} \quad n_S \text{ even}. \tag{35}$$

Proof. Let Alice's and Bob's joint state be $|\psi_S(\underline{h}, U)\rangle = \sum_{k,j} C_{kj}(\underline{h}, U) |k\rangle_A \otimes |j\rangle_B$, where we include the possibility of no measurement ($M_H = 0$). The Schmidt decomposition is achieved by a state dependent change of basis such that

$$|\psi_S(\underline{h}, U)\rangle = \sum_{q=1}^{\min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)} \lambda_q |q\rangle_A \otimes |q\rangle_B, \tag{36}$$

where $\{|q\rangle_{A,B}\}$ are orthonormal bases for A and B , respectively.

Writing this state in terms of Alice and Bob's photon numbers we have $|\psi_S(\underline{h}, U)\rangle = \sum_{n_A=0}^{n_S} |\psi_S^{n_A, n_B}(\underline{h}, U)\rangle$ with $n_B = n_S - n_A$. The overlap $\langle \psi_S^{n_A, n_B}(\underline{h}, U) | \psi_S^{n'_A, n'_B}(\underline{h}, U) \rangle = 0$ for $n_A \neq n'_A$, $n_B \neq n'_B$ as these states belong to different Hilbert subspaces in the direct sum. The reduced density matrix is block diagonal – each block corresponds to a different (n_A, n_B) combination. We may therefore consider each subspace individually, where the maximal Schmidt rank is $\min(\dim \mathcal{H}_A^{n_A}, \dim \mathcal{H}_B^{n_B})$. The total number of Schmidt coefficients is therefore at most

$$\begin{aligned}
\omega &= \sum_{n_A=0}^{n_S} \min\{\dim \mathcal{H}_A^{n_A}, \dim \mathcal{H}_B^{n_B}\} \\
&= \sum_{n_A=0}^{n_S} \min\left\{ \binom{M_A + n_A - 1}{n_A}, \binom{M_B + n_B - 1}{n_B} \right\}. \tag{37}
\end{aligned}$$

For $M_A = M_B$ this gives the result. \square

Since the entanglement is given by the number of nonzero Schmidt coefficients, this gives a bound on the entanglement $S \leq \log(\omega)$.

2. Linearity bound

Here we consider a bound due to the linearity of the interferometer transformations. In the following we do not assume anything about the form of the input Fock state, nor whether measurement occurs or not.

Proposition V.10. *Given an n photon Fock state as input to a M -mode linear optical device, with Alice and Bob having M_A and M_B output modes respectively, the maximal entanglement achievable between Alice and the rest of the modes for any state is bounded by n ebits.*

Proof. Starting with the arbitrary linear optical mode transformation in Eq.(4), we can group the sum into Alice’s modes and the ‘rest’:

$$\begin{aligned} \hat{a}_k^\dagger &\mapsto \sum_{j=1}^M \hat{a}_j^\dagger U_{jk} = \sum_{j=1}^{M_A} \hat{a}_j^\dagger U_{jk} + \sum_{j=M_A+1}^M \hat{a}_j^\dagger U_{jk} \\ &=: \hat{A}_k(U) + \hat{R}_k(U). \end{aligned} \quad (39)$$

The degree one polynomials $\hat{A}_k(U)$, $\hat{R}_k(U)$ in the creation operators are not canonical raising operators, because e.g. $[\hat{A}_k(U), \hat{A}_{k'}(U)] \neq \delta_{kk'}$. This means that different monomials in $\{\hat{A}_k(U)\}_k$ do not necessarily give rise to orthogonal states; however, this can only reduce the Schmidt rank of the resulting state.

An arbitrary input Fock state is of the form $\prod_{k=1}^M (\hat{a}_k^\dagger)^{n_k} / \sqrt{n_k!} |\text{vac}\rangle$, so that the output state is of the form

$$\prod_{k=1}^M \frac{1}{\sqrt{n_k!}} (\hat{A}_k(U) + \hat{R}_k(U))^{n_k} |\text{vac}\rangle \quad (40)$$

i.e. it is a product of n terms, not all of which are necessarily different. We can rewrite it as

$$\mathcal{N} \prod_{k=1}^n (\hat{A}_{j_k}(U) + \hat{R}_{j_k}(U)) |\text{vac}\rangle, \quad (41)$$

where $j_k \in \{1, \dots, M\}$ and \mathcal{N} is the necessary normalization. The highest number of monomial terms in this product is bounded by 2^n and after tracing out Bob and Harold this also bounds the number of monomial terms that can be in Alice’s reduced state. \square

Consider a balanced 50:50 beam splitter coupling one of Alice’s modes (say k) to one of Bob’s modes (say $k + M_A$). If Alice’s mode contained one input photon and Bob’s none, we get 1 ebit of entanglement. Proposition V.10 tells us we can only get up to n ebits using n photons, so as long as $n \leq M_A = M_B$, a beamsplitter coupling mode k with mode $k + M_A$ for $k = 1$ through n in this way would give us a state that achieves the bound.

The dimensionality (Section VB1) and linearity bounds above hold for all M and all n . We can find numerically the photon number $n_L(M_A) \in \mathbb{N}$, which depends on the number of Alice’s modes. For a given M_A it represents the number of photons up to which the linearity bound is smaller than the dimensionality bound. For $n > n_L(M)$, the dimensionality bound is a tighter limit on the entanglement (see Figure 10).

C. Hints of another bound

In this section we explore a potential bound that is motivated by numerical evidence (see Figure 10). While adding more photons to the interferometer increases the size of Alice’s and Bob’s Hilbert spaces, and according to the results from the previous section should allow for higher amounts of entanglement, we see that this is not what happens in general (assuming the number of modes that Alice and Bob have are fixed). Based on the analytical results from Section VA and the numerical evidence for all cases up to $n = 7$ photons and $M_A = M_B = 3$ modes, we make a conjecture that there is another bound which seems to arise from the unitarity of the mode transformation.

Conjecture V.1. For n unbunched photons input into an interferometer with M_A Alice and M_B Bob output modes with $n > M_A + M_B$, the average amount of entanglement, obtained over Harold’s measurements, is bounded above by the maximal average amount of entanglement achieved when $n = M_A + M_B$.

We provide numerical evidence supporting this “unitarity bound” for various numbers of input photons and modes. We assume that the input states are unbunched, ancillas and measurement are allowed, and Alice and Bob have the same number of modes; $M_I = n$, $M_V \geq 0$, $M_A = M_B \geq 1$ and $M_H \geq 0$.

Propositions V.9 and V.10 provide tight entanglement bounds when all input photons are kept in the system, i.e. when there is no detection. We know that it is possible to postselect states that exceed these bounds, but because we are interested in average entanglement these cases must be weighted with their heralding probabilities. Our findings are consistent with a generic trade off between these two quantities, leading to a bounded average entanglement.

Figure 10 shows the results of numerical optimization of the average entanglement given by Eq. (18) for various numbers of input photons and modes. We can see how the linearity and dimensionality bounds of Sec. VB are indeed limiting the entanglement. We also see the appearance of what looks like a third bound, seemingly when the number of photons is larger than the total number of modes in the system ($M_A + M_B$). This new behaviour is not captured by the bounds we have obtained and we conjecture that it is due to the unitarity of the interferometric transformation. This leads to the hypothesis that the maximum possible average entanglement, in situations with unbunched input and Alice and Bob have the same number of modes, can be reached using a $(M_A + M_B)$ -mode interferometer with $M_A + M_B$ photons.

D. Optimal interferometers

Finally, in this section we report some of the explicit interferometers (unitaries) that produce the optimal en-

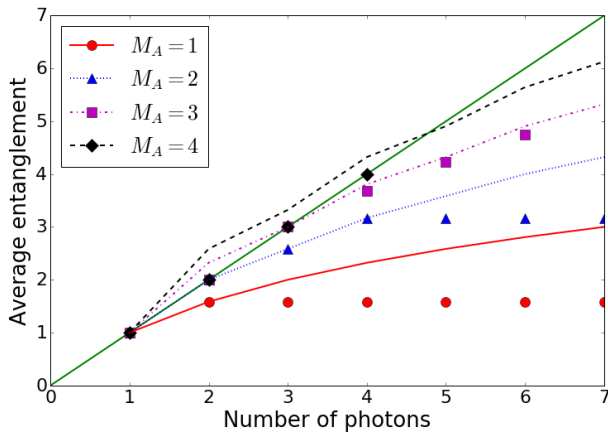


FIG. 10. Plot of the maximum average entanglement found through numerical optimization, along with the dimensionality and linearity bounds for $M_A = M_B$. The input are unbunched states. If $n > M_A + M_B$, the remaining $M_H = n - M_A - M_B$ modes contain detectors. The green solid (straight) line is the linearity bound. Other lines are dimensionality bounds for the value of M_A whose dots have the corresponding colour. The dots are values found through numerical optimization. We can see that the values of n_L for specific M_A s are: $n_L(1) = 1, n_L(2) = 2, n_L(3) = 3, n_L(4) = 4$. The dots are solutions returned by the optimization.

tanglement found for small number of modes.

In the case of $M_A = M_B = 1$ and a single photon $n = 1$, the well known balanced 50:50 beamsplitter is optimal,

$$BS_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \quad (42)$$

This is familiar, as in single-rail encoding it creates a Bell state. Let $\theta = \frac{1}{2} \arccos(1/\sqrt{3})$. When we input two photons, $n = 2$, with one in each mode, the unitary

$$BS_2 = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (43)$$

produces a state with $\log 3$ ebits of entanglement. Con-

jecture V.1 says that for all higher photon numbers, $\log 3$ will still be the maximum, achieved by using BS_2 between any pair of Alice and Bob's modes and identity on all the others (they are just routed straight to the detectors).

For $M_A = M_B = 2$, we have that all the optimal interferometers are actually combinations of BS_1 and BS_2 . An example for $M = n = 4$ is:

$$\begin{bmatrix} \cos \theta & 0 & 0 & \sin \theta \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ -\sin \theta & 0 & 0 & \cos \theta \end{bmatrix}, \quad (44)$$

where as before $\theta = \frac{1}{2} \arccos(1/\sqrt{3})$. This interferometer corresponds to a BS_2 beamsplitter between modes 1 and 4 and another BS_2 beamsplitter between modes 2 and 3, giving $\log 9 \approx 3.17$ ebits of entanglement. When $n = 3$, the optimal value of $\log 6 \approx 2.58$ ebits is achieved by using BS_2 on modes 1 and 4 and BS_1 on modes 2 and 3. For $n = 2$, the maximum of 2 ebits is achieved by two BS_1 beamsplitters, similar to $n = 4$ case. Finally, for $n = 1$ we just use a single BS_1 to achieve 1 ebit.

The data used to generate Figures 4, 5, and 6 is available for download at the University of Bristol data repository, data.bris, at <https://doi.org/10.5523/bris.1o09f3qf75gnc2018ko8kxklnf>.

ACKNOWLEDGMENTS

SS was supported by the Bristol Quantum Engineering Centre for Doctoral Training, EPSRC grant EP/L015730/1 and would like to thank P Birchall for bringing to our attention a more elegant proof of Lemma V.2. AM was supported by EPSRC Early Career Fellowship EP/L021005/1. PST was supported in part by the U.S. Army Research Office under contract W911NF-14-1-0133 and EPSRC First Grant EP/N014812/1.

-
- [1] J. L. O'Brien, A. Furusawa, and J. Vučković, *Nature Photonics* **3**, 687 (2010).
 - [2] M. Saffman, T. G. Walker, and K. Mølmer, *Reviews of Modern Physics* **82**, 2313 (2010).
 - [3] Z.-l. Xiang, S. Ashhab, J. Q. You, and F. Nori, *Reviews of Modern Physics* **85**, 623 (2013).
 - [4] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, *Reviews of Modern Physics* **79**, 135 (2007).
 - [5] R. Jozsa and N. Linden, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **459**, 2011 (2003).
 - [6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Reviews of Modern Physics* **81**, 865 (2009).
 - [7] J. S. Bell, *Physics* **1**, 195 (1964).
 - [8] A. Aspect, *Physics Letters A* **54**, 117 (1975).
 - [9] S. Popescu and D. Rohrlich, *Foundations of Physics* **24**, 379 (1994).
 - [10] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, *Science* **349**, 711 (2015).

- [11] M. Gimeno-Segovia, P. Shadbolt, D. E. Browne, and T. Rudolph, *Physical Review Letters* **115**, 020502 (2015).
- [12] J. Joo, P. L. Knight, J. L. O'Brien, and T. Rudolph, *Physical Review A* **76**, 052326 (2007).
- [13] Q. Zhang, X.-H. Bao, C.-Y. Lu, X.-Q. Zhou, T. Yang, T. Rudolph, and J.-W. Pan, *Physical Review A* **77**, 062316 (2008).
- [14] T. Rudolph, (2016), arXiv:1610.07128.
- [15] S. Aaronson and A. Arkhipov, *Theory of Computing* **9**, 143 (2013).
- [16] F. Benatti, R. Floreanini, and U. Marzolino, *Annals of Physics* **327**, 1304 (2012).
- [17] F. Benatti, R. Floreanini, and U. Marzolino, *Physical Review A* **85**, 042329 (2012).
- [18] S. Scheel, (2004), arXiv:quant-ph/0406127.
- [19] K. Kieling, *Linear optics quantum computing—construction of small networks and asymptotic scaling*, Ph.D. thesis, Imperial College London (2008).
- [20] E. Jones, T. Oliphant, P. Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” (2001–2017).
- [21] W. Vogel and J. Sperling, *Physical Review A* **89**, 052302 (2014).
- [22] J. P. Olson, K. R. Motes, P. M. Birchall, N. M. Studer, M. LaBorde, T. Moulder, P. P. Rohde, and J. P. Dowling, (2016), arXiv:1610.07128.
- [23] M. Horodecki, *Quantum Information & Computation* **1**, 3 (2001).
- [24] A. Hurwitz, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1897**, 71 (1897).
- [25] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Physical Review Letters* **73**, 58 (1994).