

Quantum mechanical description of linear optics

Johannes Skaar,^{a)} Juan Carlos García Escartín, and Harald Landro
Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

(Received 13 January 2004; accepted 3 June 2004)

Linear optical networks with any number of modes are described quantum mechanically. We reformulate the conventional formal description based on input–output operator relations and obtain an intuitive description based on the possible paths taken by the photons. The effect of a linear optical network on localized photons is treated within the same formalism. The potential and limitations of linear optics for application in quantum information processing are briefly discussed.

© 2004 American Association of Physics Teachers.

[DOI: 10.1119/1.1775241]

I. INTRODUCTION

Linear optical networks have important applications in the telecommunication and optical sensor industry. In particular, the possibility of manufacturing advanced filters, including bandpass filters and dispersion compensation filters, has attracted much attention during the last three decades. Sophisticated filter responses can be made of discrete components such as optical couplers and phase shifters,¹ or with quasi-periodical structures such as fiber Bragg gratings,² and optical thin-film filters.³

Recently, there has been a growing interest in linear optical networks in the context of quantum information technology.⁴ It has been demonstrated that universal (non-deterministic) quantum computation is possible when linear optical networks are combined with single photon detectors. Due to the increasing interest in linear optics as a candidate for universal quantum computation, it is useful to have a general and self-contained introduction to this field which is appropriate for students, teachers, and potential researchers. Although the quantum theory of linear multiports is well-established,^{5–7} the theory may be difficult to understand intuitively because it is based on a formal input–output operator relation. We will demonstrate that this conventional theory of linear optics can be reformulated in an intuitive way: The probability amplitude for a given event is given by a sum over all possible paths that the photons may take through the network. Each term in the sum represents the classical transmission coefficients that the photons pick up on their way. This path formulation is analogous to the Feynman rules of quantum electrodynamics.⁸

In conventional descriptions of linear optical networks, the photons are usually assumed to be monochromatic. However, in all practical experiments the optical signals have finite duration. Therefore, to understand the behavior of an optical network in practical experiments, a time-domain formulation is necessary. We will therefore also analyze the effect of the multiport on localized photons in the time domain. Finally, we review some important properties and limitations for applications in the field of quantum information processing.

II. CLASSICAL DESCRIPTION

An optical component is linear if the output fields are linearly related to the input fields. Examples of such components include beam splitters, couplers, phase shifters, polarization rotators, polarizers, wave retarders, filters, or networks of these components.

We start by reviewing the classical theory of linear optical networks. Consider a linear multiport with N input modes and N output modes (see Fig. 1). The different input and output modes do not have to be associated with different physical ports. Input and output modes with the same index may share the same physical port because they propagate in different directions. Input (or output) modes with different indices also may share the same physical port if they are separated, e.g., in frequency or polarization. We denote the complex, classical fields in the input and output modes by a_i and b_i , respectively, where i is an integer between 1 and N . Because the multiport is assumed to be linear, the output fields must be linearly related to the input fields,

$$b_i = \sum_{j=1}^N S_{ij} a_j, \quad (1)$$

or in vector notation

$$\mathbf{b} = \mathbf{S}\mathbf{a}. \quad (2)$$

The matrix \mathbf{S} is the scattering matrix associated with the network. The scattering parameter S_{ij} is the transmission coefficient from input j to output i (note the order).

We restrict ourselves to lossless networks, so that the output power is equal to the input power:

$$\mathbf{a}^\dagger \mathbf{a} = \mathbf{b}^\dagger \mathbf{b} = (\mathbf{S}\mathbf{a})^\dagger \mathbf{S}\mathbf{a} = \mathbf{a}^\dagger \mathbf{S}^\dagger \mathbf{S}\mathbf{a}, \quad (3)$$

or

$$\mathbf{a}^\dagger (\mathbf{S}^\dagger \mathbf{S} - \mathbf{I}) \mathbf{a} = 0. \quad (4)$$

Equation (4) is valid for any vector \mathbf{a} , and it follows that \mathbf{S} is unitary:

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}. \quad (5)$$

This property is realized by noting that $\mathbf{M} \equiv \mathbf{S}^\dagger \mathbf{S} - \mathbf{I}$ is Hermitian. Thus \mathbf{M} can be diagonalized in an orthonormal basis $|i\rangle$, that is, $\mathbf{M} = \sum_i m_i |i\rangle\langle i|$ in Dirac notation. Because \mathbf{a} is arbitrary, it may be chosen to be equal to one of the eigenvectors $|i\rangle$. Equation (4) then gives $m_i = 0$, which implies that $\mathbf{M} = 0$.

Most electromagnetic components are reciprocal, so for coinciding inputs and outputs, the scattering matrix is symmetrical ($\mathbf{S} = \mathbf{S}^T$).⁷ The symmetry applies to reciprocal components such as beam splitters, couplers, thin-film filters, fiber Bragg gratings, and networks of such components, while components based on the Faraday effect (Faraday rotators,

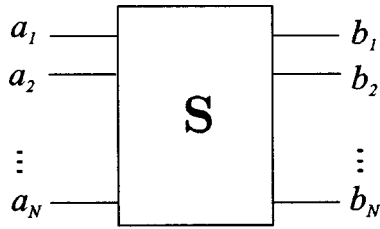


Fig. 1. A linear optical multiport with N input and N output modes.

optical isolators, Faraday mirrors, and circulators) do not have symmetric scattering matrices in general.

As an example, consider a beam splitter (see Fig. 2). Because we have a total of four physical ports with two polarization modes in each, a general scattering matrix would have dimension 8×8 . However, because TE and TM polarizations (electric field orthogonal or parallel to the plane of incidence) do not couple to each other, we can treat the two polarizations separately. If we assume for instance TE polarization, the scattering matrix has dimension 4×4 . In this description, the input and output modes share the same physical ports. Because the beam splitter is reciprocal, this 4×4 scattering matrix must be symmetric. Moreover, one input mode couples only to two output modes, so many of the elements in this 4×4 matrix are zero. It is therefore convenient to use the convention that the beam splitter has only two input modes a_1 and a_2 , and two output modes b_1 and b_2 (see Fig. 2):

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}. \quad (6)$$

Because $\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}$, we obtain:

$$|S_{11}|^2 + |S_{21}|^2 = 1, \quad (7a)$$

$$|S_{12}|^2 + |S_{22}|^2 = 1, \quad (7b)$$

$$S_{11}S_{12}^* = -S_{21}S_{22}^*. \quad (7c)$$

By taking the absolute value of both sides of Eq. (7c) and using Eqs. (7a) and (7b), we see that

$$|S_{11}| = |S_{22}|, \quad (8a)$$

$$|S_{12}| = |S_{21}| = \sqrt{1 - |S_{11}|^2}. \quad (8b)$$

Thus the scattering matrix can be written as

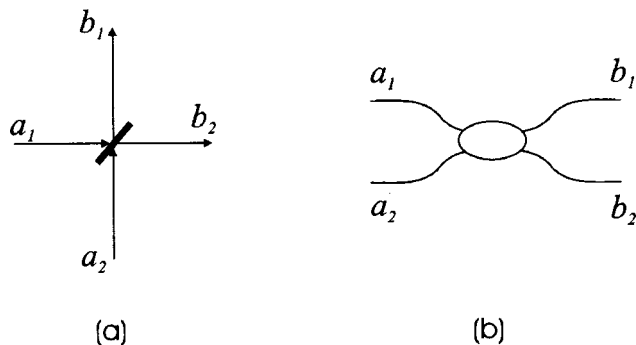


Fig. 2. (a) A beam splitter and (b) an optical coupler. In principle, the two optical components are equivalent.

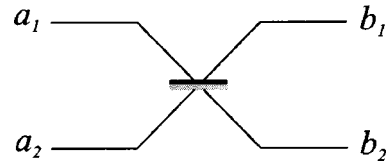


Fig. 3. An asymmetrical beam splitter described by Eq. (11). Reflection off the grey surface results in a π phase shift, while transmission from either side or reflection off the black surface results in no phase shift.

$$\mathbf{S} = \begin{bmatrix} \sqrt{\eta}e^{i\phi_{11}} & \sqrt{1-\eta}e^{i\phi_{12}} \\ \sqrt{1-\eta}e^{i\phi_{21}} & \sqrt{\eta}e^{i\phi_{22}} \end{bmatrix}, \quad (9)$$

where $0 \leq \eta \leq 1$ is a real constant that describes the power transmission from input 1 to output 1: $\eta = |S_{11}|^2$. The phases ϕ_{ij} are real, and satisfy

$$\phi_{11} + \phi_{22} = \pi + \phi_{12} + \phi_{21} \quad (10)$$

as a result of Eq. (7c). These phases depend on the choice of reference plane at the four ports.

It is easy to realize that with an appropriate shift of reference plane, $\phi_{12} = \phi_{21} = 0$, which gives $\phi_{11} + \phi_{22} = \pi$. In addition, we may assume that the beam splitter is asymmetric such that $\phi_{11} = 0$ and $\phi_{22} = \pi$. This asymmetry is present, for example, if the modes with index 1 propagate in a high-index material, while the modes with index 2 propagate in a low-index material. The beam splitter is in this case the plane boundary between the two materials. The scattering matrix becomes

$$\mathbf{S} = \begin{bmatrix} \sqrt{\eta} & \sqrt{1-\eta} \\ \sqrt{1-\eta} & -\sqrt{\eta} \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}. \quad (11)$$

In Eq. (11) we have introduced the parametrization $\eta = \cos^2 \theta$. A drawing of the asymmetric beam splitter is given in Fig. 3, using the convention given in Ref. 9.

III. QUANTUM MECHANICAL DESCRIPTION

In this section we consider the multiport from a quantum mechanical perspective. The theory is based on conventional field quantization in the (discrete) frequency-domain. In Sec. IV we will see how the theory can be formulated for localized photons (time-domain description) so that mode separation in time is possible as well.

A single mode of the electromagnetic field is described classically as a harmonic oscillator with a certain frequency ω . Quantization means that the classical oscillator is replaced by a quantum mechanical oscillator. The classical amplitude a becomes a quantum mechanical operator. This operator and its adjunct satisfy the commutator relation $[a, a^\dagger] = 1$. The Hamiltonian associated with the mode is

$$H = \hbar \omega (a^\dagger a + 1/2), \quad (12)$$

with orthonormal eigenstates $|n\rangle$ and eigenvalues $E_n = \hbar \omega (n + 1/2)$. The eigenvectors satisfy

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad (13a)$$

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \quad (13b)$$

so we can express the n photon state as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle. \quad (14)$$

An optical component with N input and N output modes is described by the operators a_i (inputs) and b_i (outputs). The input modes are independent, so the associated operators obey the commutator relations

$$[a_i, a_j^\dagger] = \delta_{ij}, \quad (15a)$$

$$[a_i, a_j] = 0. \quad (15b)$$

Because the transition from the classical to the quantum mechanical description consists in replacement of the classical fields by operators, it is clear that the relation between b_i and a_i still can be written as

$$b_i = \sum_{j=1}^N S_{ij} a_j, \quad (16)$$

where S_{ij} is the same scattering matrix as in the classical description. Equation (16) is consistent with the fact that $[b_i, b_j^\dagger] = \delta_{ij}$:

$$\begin{aligned} [b_i, b_j^\dagger] &= \left[\sum_{k=1}^N S_{ik} a_k, \sum_{l=1}^N S_{jl}^* a_l^\dagger \right] \\ &= \sum_{k,l=1}^N S_{ik} S_{jl}^* [a_k, a_l^\dagger] = \delta_{ij}, \end{aligned} \quad (17)$$

where the last equality follows from Eq. (15a) and the fact that \mathbf{S} is unitary.

The component is now completely described by the relation between the input and output operators a_i and b_i . For readers used to the Schrödinger picture, this relation might be surprising because the Schrödinger equation tells us that dynamics is represented by unitary evolution of the state vector:

$$|\psi_b\rangle = U|\psi_a\rangle, \quad (18)$$

where U is a unitary operator that takes the input state $|\psi_a\rangle$ to the output state $|\psi_b\rangle$. In this picture, operators and observables are constant. On the other hand, Eq. (16) refers to the Heisenberg picture, where the state vectors are constant while the operators evolve. Of course, the two pictures predict the same measurement statistics for any observable. In the Schrödinger picture, where the dynamics is described by Eq. (18), the expectation value of an observable (operator) A on the output state is

$$\langle A \rangle = \langle \psi_b | A | \psi_b \rangle = \langle \psi_a | U^\dagger A U | \psi_a \rangle. \quad (19)$$

Because the Heisenberg picture must give the same result, the operator A must evolve as

$$A \rightarrow U^\dagger A U, \quad (20)$$

where the state vector is constant.

We still do not know the operator U because we have not identified the Hamiltonian of the multiport. Nevertheless, from Eq. (16) we know the dynamics of the annihilation and creation operators:¹⁰

$$a_i \rightarrow b_i = \sum_{j=1}^N S_{ij} a_j \equiv U^\dagger a_i U, \quad (21a)$$

$$a_i^\dagger \rightarrow b_i^\dagger = \sum_{j=1}^N S_{ij}^* a_j^\dagger \equiv U^\dagger a_i^\dagger U. \quad (21b)$$

This equation can be used for determining the probabilities for detecting a certain number of photons at certain outputs.

For simplicity, take the beam splitter Eq. (11) as an example. Assume that input 1 is in the n -photon state $|n\rangle$ while input 2 is in the vacuum state, that is, the total input state is $|n0\rangle \equiv |n\rangle|0\rangle$. If we use the Heisenberg picture where the state is constant and the operators evolve, we obtain the following expectation value for the number of detected photons in output 1:

$$\begin{aligned} \langle n0 | b_1^\dagger b_1 | n0 \rangle &= \langle n0 | (\sqrt{\eta} a_1^\dagger + \sqrt{1-\eta} a_2^\dagger) \\ &\quad \times (\sqrt{\eta} a_1 + \sqrt{1-\eta} a_2) | n0 \rangle = \eta n. \end{aligned} \quad (22)$$

This result is not surprising because the power reflection of the beam splitter is η . A more interesting result is obtained if we calculate the expectation value of the product of the number of detected photons in output 1 and 2, $\langle n0 | b_2^\dagger b_2 b_1^\dagger b_1 | n0 \rangle$. If we use the same procedure as above, we find with the help of Eq. (15) that the expectation value is $\eta(1-\eta)n(n-1)$. This expression vanishes for $n=1$; thus a photon cannot be detected in both outputs as long as there is only one photon in the inputs. This non-classical result demonstrates that the photon cannot be split using this linear device.

The Heisenberg picture is convenient for computing the expectation values of photon numbers. In many cases, however, we are mainly interested in the probabilities for detecting a certain number of photons in the outputs. To calculate the output state, we return to the Schrödinger picture where the operators are constant and the states evolve. We assume the input state $|10\rangle$, that is, a single incident photon in mode 1 with the other mode in the vacuum state, and we want to find the output state $|\psi_b\rangle$. The dynamics due to the beam splitter can be described by the unitary operator U , that is, $|\psi_b\rangle = U|10\rangle$. With the help of Eq. (14) we obtain:

$$U|10\rangle = U a_1^\dagger |00\rangle = U a_1^\dagger U^\dagger U |00\rangle = U a_1^\dagger U^\dagger |00\rangle. \quad (23)$$

The last equality results from the fact that the beam splitter has no effect on the vacuum $|00\rangle$; no photons at the input give no photons at the output.¹¹ Although we are in the Schrödinger picture, it is perfectly valid to use the last part of Eq. (21b), which can be viewed as a mathematical transformation of a_i^\dagger :

$$U^\dagger a_i^\dagger U = \sum_{j=1}^N S_{ij}^* a_j^\dagger. \quad (24)$$

We need the inverse transformation $U a_i^\dagger U^\dagger$. The substitution $U \rightarrow U^\dagger \equiv U^{-1}$ means evolution under the inverse operator, that is, exchanging the role of the inputs and the outputs. This exchange implies that $\mathbf{S} \rightarrow \mathbf{S}^\dagger$. Thus,

$$U a_i^\dagger U^\dagger = \sum_{j=1}^N S_{ji} a_j^\dagger. \quad (25)$$

If we substitute Eq. (25) into Eq. (23) and use the scattering matrix Eq. (11), we find

$$U|10\rangle = (\sqrt{\eta} a_1^\dagger + \sqrt{1-\eta} a_2^\dagger) |00\rangle = \sqrt{\eta} |10\rangle + \sqrt{1-\eta} |01\rangle. \quad (26a)$$

Similarly we can calculate

$$U|01\rangle = \sqrt{1-\eta}|10\rangle - \sqrt{\eta}|01\rangle, \quad (26b)$$

$$U|11\rangle = \sqrt{2\eta}\sqrt{1-\eta}[|20\rangle - |02\rangle] + [1-2\eta]|11\rangle, \quad (26c)$$

$$U|02\rangle = (1-\eta)|20\rangle + \eta|02\rangle - \sqrt{2\eta}\sqrt{1-\eta}|11\rangle, \quad (26d)$$

etc. In obtaining Eq. (26) we have used the fact that the creation operators associated with different input modes commute, see Eq. (15b).

Finally, we will demonstrate how to calculate the output state in a general linear network with several modes. Initially, we limit ourselves to the case with three input and three output modes, and we wish to calculate the output state when the input is in the state $|210\rangle$, that is, mode 1 has two incident photons, mode 2 has one, and mode 3 has zero. We find:

$$\begin{aligned} U|210\rangle &= U \frac{(a_1^\dagger)^2}{\sqrt{2!}} \frac{a_2^\dagger}{\sqrt{1!}} |000\rangle \\ &= \frac{(Ua_1^\dagger U^\dagger)^2}{\sqrt{2!}} \frac{Ua_2^\dagger U^\dagger}{\sqrt{1!}} U|000\rangle \\ &= \frac{(\sum_{k=1}^N S_{k1} a_k^\dagger)^2}{\sqrt{2!}} \frac{\sum_{k=1}^N S_{k2} a_k^\dagger}{\sqrt{1!}} |000\rangle \\ &= \sum_{k,l,m=1}^N \frac{S_{k1} S_{l1} S_{m2}}{\sqrt{2!1!}} a_k^\dagger a_l^\dagger a_m^\dagger |000\rangle. \end{aligned} \quad (27)$$

All terms in Eq. (27) contain three creation operators, that is, the superposition state at the output consists only of three-photon terms such as $|021\rangle$, $|300\rangle$, $|111\rangle$, etc. This fact implies that no photons are lost or created, as we would expect for a linear and lossless network. First, consider the term with $|300\rangle$, that is, $k=l=m=1$. The probability amplitude of this term (its coefficient) becomes $S_{11}S_{11}S_{12}\sqrt{3!}/2!$. If we ignore the normalization $\sqrt{3!}/2!$, this term represents the product of the transmission coefficients along the paths of the three photons; two photons are transmitted from input 1 to output 1 and one photon is transmitted from input 2 to output 1. Second, consider the term with $|201\rangle$. Here there are several possible set of indices (k,l,m) giving $|201\rangle$: $(1,1,3)$, $(1,3,1)$, and $(3,1,1)$. The associated probability amplitude becomes $S_{11}S_{11}S_{32} + S_{11}S_{31}S_{12} + S_{31}S_{11}S_{12}$. Note that an extra normalization factor does not appear in this case because $\sqrt{2!}/2! = 1$. The three terms in the amplitude represent the possible ways the photons can go. For example, the last term $S_{31}S_{11}S_{12}$ means that the first photon is transmitted from input 1 to output 3, the second from 1 to 1, and the third from 2 to 1.

This procedure can of course be extended to any number of modes and photons. More generally, we have found the following rules that can be applied for determining the quantum mechanical operation of any linear optical network.

(1) The state $|n_1, n_2, \dots, n_N\rangle$ becomes a superposition of terms of the form $|m_1, m_2, \dots, m_N\rangle$, where the total number of photons is conserved, $m_1 + m_2 + \dots + m_N = n_1 + n_2 + \dots + n_N$.

(2) The probability amplitude for the transition $|n_1, n_2, \dots, n_N\rangle \rightarrow |m_1, m_2, \dots, m_N\rangle$ is

$$\sum_{\text{paths}} C \cdot S_{\text{path photon } 1} \cdots S_{\text{path photon } n_1+n_2+\dots+n_N}, \quad (28)$$

where

$$C = \sqrt{\frac{m_1! m_2! \cdots m_N!}{n_1! n_2! \cdots n_N!}}.$$

The terms in the sum represent the possible paths taken by the photons.

Using this method it is simple to verify Eq. (26). For example:

$$\begin{aligned} U|11\rangle &= \sqrt{2}\sqrt{\eta}\sqrt{1-\eta}|20\rangle - \sqrt{2}\sqrt{\eta}\sqrt{1-\eta}|02\rangle \\ &\quad + [\sqrt{1-\eta}\sqrt{1-\eta} - \sqrt{\eta}\sqrt{\eta}]|11\rangle \\ &= \sqrt{2}\sqrt{\eta}\sqrt{1-\eta}[|20\rangle - |02\rangle] + [1-2\eta]|11\rangle. \end{aligned} \quad (29)$$

The probability of detecting one photon in both outputs is $|1-2\eta|^2$. For a 50:50 beam splitter, $1-2\eta=0$, so $U|11\rangle = [|20\rangle - |02\rangle]/\sqrt{2}$. In this case the two paths leading to $|11\rangle$ interfere destructively. It is interesting to note that the state $[|20\rangle - |02\rangle]/\sqrt{2}$ cannot be written as a product state, so the output modes are entangled.

IV. LOCALIZED PHOTONS

In practical experiments the photons are more or less localized. Although a localized input pulse contains different frequencies, we will find that the effect of the multiport still can be analyzed in a similar fashion as we did previously.

Localized photons contain necessarily a continuous band of frequencies. We denote the continuous frequency destruction and creation operators in mode i by $a_i(\omega)$ and $a_i^\dagger(\omega)$, respectively. The operators are normalized such that

$$[a_i(\omega), a_j^\dagger(\omega')] = \delta_{ij} \delta(\omega - \omega'), \quad (30)$$

similarly to Eq. (15a). Localized photons in a pulse centered about time t_0 can be created by the photon-wavepacket creation operator,⁵

$$a_{i,t_0}^\dagger = \int d\omega \xi(\omega, t_0) a_i^\dagger(\omega), \quad (31)$$

where the function $\xi(\omega, t_0)$ for example can be a Gaussian pulse shape:

$$\xi(\omega, t_0) = (2\pi\Delta^2)^{-1/4} \exp\left[i(\omega - \omega_0)t_0 - \frac{(\omega - \omega_0)^2}{4\Delta^2} \right]. \quad (32)$$

In Eq. (32), ω_0 and Δ are the central frequency and pulse bandwidth, respectively. Usually $\Delta \ll \omega_0$, so the range of integration in Eq. (31) can be extended to the entire frequency axis from $-\infty$ to ∞ . With the help of Eq. (30) we find the commutator relation,

$$[a_{i,t_0}, a_{j,t_1}^\dagger] = \delta_{ij} \exp\left(-\frac{\Delta^2(t_1 - t_0)^2}{2} \right), \quad (33)$$

showing that two pulses in the same mode can be treated as independent if they are sufficiently separated in time, that is,

$$t_1 - t_0 \gg 1/\Delta. \quad (34)$$

In principle, even if Eq. (34) is not fulfilled, we can still define orthogonal pulses by the standard Gram-Schmidt procedure.

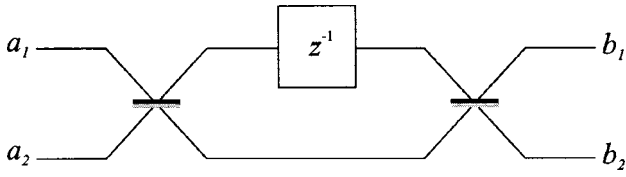


Fig. 4. A filter structure consisting of two couplers and a delay element. The labels refer to the classical fields or the Heisenberg-picture operators.

We are interested in the dynamics of a linear multiport for the case with localized photons. Equation (25) is obviously valid for the continuous frequency creation operators as well, so

$$U a_{i,t_0}^\dagger U^\dagger = \sum_{j=1}^N \int d\omega \xi(\omega, t_0) S_{ji} a_j^\dagger(\omega). \quad (35)$$

For the special case of frequency-independent scattering parameters S_{ij} , we simply find

$$U a_{i,t_0}^\dagger U^\dagger = \sum_{j=1}^N S_{ji} a_{j,t_0}^\dagger, \quad (36)$$

analogously to Eq. (25). This result demonstrates that the methods in Sec. III can be applied directly to this case.

Also for the case of frequency-dependent scattering parameters, a similar method can be used to analyze the effect of the multiport on the input state. As an example, consider the linear filter given in Fig. 4 consisting of a delay element sandwiched between two couplers, both with power reflection $\eta = 1/2$. From Eq. (11) we find the classical scattering matrix, which in this case depends on frequency:

$$\begin{aligned} \mathbf{S} &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} z^{-1}+1 & z^{-1}-1 \\ z^{-1}-1 & z^{-1}+1 \end{bmatrix}. \end{aligned} \quad (37)$$

Here $z^{-1} \equiv \exp(i(\omega - \omega_0)\tau)$ represents a delay τ . Note that the scattering parameters can be viewed as a polynomial function of z^{-1} : $S_{ij}(\omega) = S_{ij}^{(0)} + S_{ij}^{(1)} z^{-1}$.

For simplicity, let the input state be $a_{1,t_0}^\dagger |00\rangle$, that is, one photon in input 1 in a pulse centered about t_0 . More inter-

esting transformations with several interfering photons can be analyzed along the same lines. We define $t_1 \equiv t_0 + \tau$, and assume that Eq. (34) is fulfilled so that the pulses centered about t_0 and t_1 do not overlap. The output state becomes

$$\begin{aligned} U a_{1,t_0}^\dagger |00\rangle &= U a_{1,t_0}^\dagger U^\dagger U |00\rangle \\ &= \sum_{j=1}^2 \int d\omega \xi(\omega, t_0) S_{j1}(\omega) a_j^\dagger(\omega) |00\rangle. \end{aligned} \quad (38)$$

In Eq. (38), there are four terms due to the two terms in the scattering parameters. If we use the fact that $z^{-1} \xi(\omega, t_0) = \xi(\omega, t_1)$, we find with the help of Eq. (31) that

$$U a_{1,t_0}^\dagger |00\rangle = \frac{1}{2} (a_{1,t_0}^\dagger - a_{2,t_0}^\dagger + a_{1,t_1}^\dagger + a_{2,t_1}^\dagger) |00\rangle. \quad (39)$$

This transformation also can be found by treating the different pulses as independent modes. If we consider two pulses (at time t_0 and t_1), we have four modes in total (see Fig. 5). If we also take the extra input (c) and output (d) line in Fig. 5 into account, we obtain the transformation

$$\begin{bmatrix} b_{1,t_0} \\ b_{2,t_0} \\ b_{1,t_1} \\ b_{2,t_1} \\ d \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 & 0 & \sqrt{2} \\ -1 & 1 & 0 & 0 & \sqrt{2} \\ 1 & 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} a_{1,t_0} \\ a_{2,t_0} \\ a_{1,t_1} \\ a_{2,t_1} \\ c \end{bmatrix}, \quad (40)$$

from which it is straightforward to verify Eq. (39) using the method in Sec. III. Alternatively, we can find Eq. (39) simply by tracing the four possible paths for the photon through the network in Fig. 5, and determining the associated transmission coefficients.

V. APPLICATIONS IN QUANTUM INFORMATION PROCESSING

A quantum bit (qubit) is a general superposition state $\alpha|0\rangle_L + \beta|1\rangle_L$, where $|\alpha|^2 + |\beta|^2 = 1$, and $|0\rangle_L$ and $|1\rangle_L$ are some logical, orthonormal basis states. The basis states can in principle be encoded with any two physical, orthonormal states. A popular encoding in the photonic case is dual-rail encoding, where two optical modes are used to represent

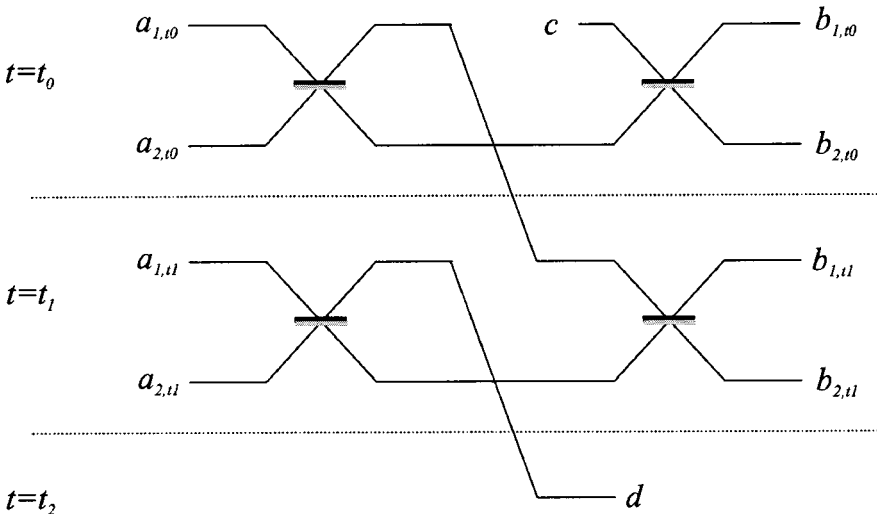


Fig. 5. Equivalent multiport when the pulses are considered as independent modes. The input c is the pulse after the delay element at $t = t_0$, and the output d is the pulse after the delay element at $t = t_2$, where $t_2 = t_1 + \tau$.

each qubit. In this encoding, $|0\rangle_L \equiv |10\rangle$ and $|1\rangle_L \equiv |01\rangle$, where for example, $|10\rangle$ means one photon in mode 1 and none in mode 2. This encoding is separable in the sense that each qubit is encoded using a certain number of associated modes.

More generally, assume that N optical modes are used to encode n qubits. The state space associated with n qubits has dimension 2^n , so a general quantum circuit on n qubits can be described by a $2^n \times 2^n$ unitary matrix V . This circuit has 2^{2^n} degrees of freedom.¹² On the other hand, a linear optical network with N modes has only N^2 degrees of freedom. Hence, to implement any quantum circuit on n qubits (universal logic), it is a necessary condition that

$$N \geq 2^n. \quad (41)$$

In other words, the required number of modes scales exponentially with the number of qubits.

An encoding that satisfies $N = 2^n$ is the “one-hot” encoding where all N modes but one are in the vacuum state. The logical states $|000\rangle_L$, $|001\rangle_L$, and $|010\rangle_L$ would then correspond to the number states $|1000000\rangle$, $|0100000\rangle$, and $|0010000\rangle$, respectively. In this encoding, there is only one photon in the network. In the logical, computational basis, V therefore corresponds directly to \mathbf{S} . Thus any V can be implemented (universal logic). The main problem with the one-hot encoding is that it is not separable; if we add one qubit to the circuit, the representation of the other qubits and the entire circuit is changed.

It would be interesting to check if universal quantum logic is possible in the separable case (using linear optics). For the case of a circuit with only $n = 1$ qubits, Eq. (41) becomes $N \geq 2$, suggesting that any single-qubit gate might be possible to implement using a linear network with two modes. Indeed, a separable dual-rail encoding is the same as one-hot in this case, so any single-qubit circuit can be implemented.

For a quantum circuit with $n \geq 2$, the situation is different. The set of two-qubit gates is universal.⁴ Thus if any two-qubit gate could be implemented, it would be possible to implement an arbitrary n -qubit circuit. In a separable representation with a constant number of modes for each qubit, Eq. (41) would be violated for a sufficiently large number of qubits n . In other words, for $n \geq 2$, some gates are not possible, even if Eq. (41) is satisfied.

A possible way to get around the degrees of freedom problem is to introduce ancillary modes. The ancilla inputs must be prepared in some standard state, and the ancilla outputs may be measured or left free. If we measure the ancillas, we can choose to accept or reject the output depending on the measurement results (postselection). It turns out that universal logic is possible using postselection,^{9,13} but the circuit becomes nondeterministic. The inclusion of photodetectors introduces a form of nonlinearity.¹⁴

Any quantum circuit can be made using one-qubit gates (attainable with linear optics) and a two-qubit gate called controlled-not gate (CNOT).⁴ This gate flips the value of the second qubit if the first one is set to $|1\rangle_L$:

$$\alpha|00\rangle_L + \beta|01\rangle_L + \gamma|10\rangle_L + \delta|11\rangle_L \xrightarrow{\text{CNOT}} \alpha|00\rangle_L + \beta|01\rangle_L + \gamma|11\rangle_L + \delta|10\rangle_L. \quad (42)$$

The subscript L indicates logical states. Much effort has been put into designing an efficient CNOT gate or similar gates

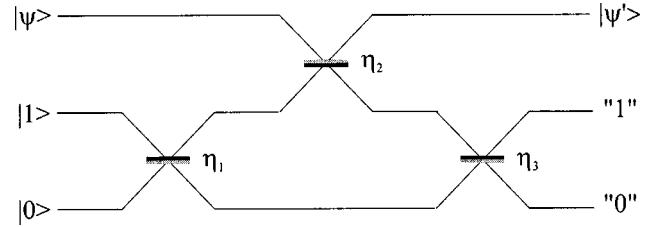


Fig. 6. NS gate as implemented in Ref. 9.

that can easily provide a CNOT gate, including the conditional sign flip gate (CS gate)¹³ and the nonlinear sign shift gate (NS gate).^{9,13} We will give an example for the simplest case of interest in the search of a CNOT gate: the NS gate defined by the following nonlinear transformation on the signal state:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle \xrightarrow{\text{NS}} |\psi'\rangle = \alpha|0\rangle + \beta|1\rangle - \gamma|2\rangle. \quad (43)$$

This gate can be implemented nondeterministically using a multiport with three input and output modes; the signal mode and two ancilla modes (see Fig. 6).

The ancilla inputs are assumed to be in the states $|1\rangle$ and $|0\rangle$, respectively. The output of the signal mode is accepted only if one photon is detected in the first ancilla output and none in the other. By using the method described in Sec. III, we will now analyze the circuit in Fig. 6, and show that the beam splitter reflectivities η_1 , η_2 , and η_3 can be chosen to obtain the transformation Eq. (43) on the signal state.

The network can be seen as three multiports in cascade, one for each of the beam splitters, with the scattering matrices

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\eta_1} & \sqrt{1-\eta_1} \\ 0 & \sqrt{1-\eta_1} & -\sqrt{\eta_1} \end{bmatrix}, \quad (44a)$$

$$\mathbf{S}_2 = \begin{bmatrix} -\sqrt{\eta_2} & \sqrt{1-\eta_2} & 0 \\ \sqrt{1-\eta_2} & \sqrt{\eta_2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (44b)$$

$$\mathbf{S}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\eta_3} & \sqrt{1-\eta_3} \\ 0 & \sqrt{1-\eta_3} & -\sqrt{\eta_3} \end{bmatrix}. \quad (44c)$$

These are variations of the beam splitter scattering matrix Eq. (11), with an extra row and column for the input that does not go through the beam splitter. The first multiport to act on the input is the leftmost one, so the scattering matrix of the entire multiport is

$$\mathbf{S} = \mathbf{S}_3 \mathbf{S}_2 \mathbf{S}_1. \quad (45)$$

It is now easy to find the probability amplitudes for the different transitions. According to Eq. (43) we want $|010\rangle \rightarrow c|010\rangle$, $|110\rangle \rightarrow c|110\rangle$, and $|210\rangle \rightarrow -c|210\rangle$ for some complex number c satisfying $|c| \leq 1$. The case $|c| < 1$ corresponds to nondeterministic operation. By determining the product of scattering parameters that the photons pick up on their way, and summing over the possible paths, we obtain the following:

- Transition $|010\rangle \rightarrow c|010\rangle$:

$$c = S_{22} = \sqrt{1-\eta_1}\sqrt{1-\eta_3} + \sqrt{\eta_1}\sqrt{\eta_2}\sqrt{\eta_3}. \quad (46)$$

- Transition $|110\rangle \rightarrow c|110\rangle$:

$$\begin{aligned} c &= S_{11}S_{22} + S_{12}S_{21} \\ &= -\sqrt{\eta_2}[\sqrt{1-\eta_1}\sqrt{1-\eta_3} + \sqrt{\eta_1}\sqrt{\eta_2}\sqrt{\eta_3}] \\ &\quad + \sqrt{\eta_1}(1-\eta_2)\sqrt{\eta_3} \\ &= -c\sqrt{\eta_2} + \sqrt{\eta_1}(1-\eta_2)\sqrt{\eta_3}. \end{aligned} \quad (47)$$

- Transition $|210\rangle \rightarrow -c|210\rangle$:

$$\begin{aligned} -c &= S_{11}S_{11}S_{22} + S_{21}S_{11}S_{12} + S_{11}S_{21}S_{12} \\ &= -2\sqrt{\eta_1}\sqrt{\eta_2}(1-\eta_2)\sqrt{\eta_3} + \eta_2(\sqrt{1-\eta_1}\sqrt{1-\eta_3} \\ &\quad + \sqrt{\eta_1}\sqrt{\eta_2}\sqrt{\eta_3}) \\ &= -2\sqrt{\eta_1}\sqrt{\eta_2}(1-\eta_2)\sqrt{\eta_3} + c\eta_2. \end{aligned} \quad (48)$$

Here, S_{ij} are the scattering parameters, that is, the elements of \mathbf{S} , and we have numbered the modes from the top. Note that the normalization factor C , as defined in Sec. III, is 1 in all cases. We follow Ref. 9, solve Eq. (47) for c and substitute the result into Eq. (48) and find

$$\eta_2 = (\sqrt{2} - 1)^2. \quad (49)$$

If we substitute η_2 into Eqs. (47) and (46), we obtain η_1 and η_3 . The maximum value for $|c|$ is achieved when

$$\eta_1 = \eta_3 = \frac{1}{4 - 2\sqrt{2}}, \quad (50)$$

yielding $c = 1/2$. In other words, every time one photon is detected in the first ancilla and none in the other, the desired NS operation on the signal state has been successful. On average this will happen 1/4 of the time. To achieve a successful operation in practice, the interferometer in Fig. 6 must be carefully stabilized.

For larger multiports the analysis can become tedious, but has the advantage of being easily implemented on a computer, thus providing a tool for the study of more complex linear optical multiports.

VI. SUMMARY

We have analyzed linear optical networks using basic concepts from quantum mechanics. The input–output operator method is shown to be equivalent to a simple formulation based on the superposition of paths taken by the photons. Moreover, we demonstrated that the theory of linear multiports also can be applied directly to the case with orthogonal time-domain modes (light pulses). Finally, we have discussed some applications in the field of quantum information processing, where this method is particularly useful.

¹Electronic mail: johannes.skaar@iet.ntnu.no

²C. K. Madsen and J. H. Zhao, *Optical Filter Design and Analysis: A Signal Processing Approach* (Wiley, New York, 1999).

³T. Erdogan, “Fiber grating spectra,” *J. Lightwave Technol.* **15**, 1277–1294 (1997).

⁴J. A. Dobrowolski, “Numerical methods for optical thin films,” *Opt. Photonics News* **8**, 24–33 (June 1997).

⁵M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge U.P., New York, 2000).

⁶R. Loudon, *The Quantum Theory of Light* (Oxford U.P., New York, 2000).

⁷C. W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics* (Springer, Berlin, 2000).

⁸H. A. Haus, *Electromagnetic Noise and Quantum Optical Measurements* (Springer, New York, 2000).

⁹F. Mandl and G. Shaw, *Quantum Field Theory* (Wiley, New York, 1993).

¹⁰T. C. Ralph, A. G. White, W. J. Munro, and G. J. Milburn, “Simple scheme for efficient linear optics quantum gates,” *Phys. Rev. A* **65**, 012314-1–012314-6 (2002).

¹¹Although a_i is not Hermitian and therefore not an observable, its dynamics must be $a_i \rightarrow U^\dagger a_i U$. This dynamics can be realized by introducing the Hermitian operators $X_i \equiv (a_i^\dagger + a_i)/2$ and $Y_i \equiv i(a_i^\dagger - a_i)/2$. Because X_i and Y_i evolve as $X_i \rightarrow U^\dagger X_i U$ and $Y_i \rightarrow U^\dagger Y_i U$, so does $a_i = X_i + iY_i$.

¹²Because the beam splitter is lossless, it cannot create photons from vacuum, that is, $U|00\rangle = e^{i\alpha}|00\rangle$ for some phase α . For any input state, this phase factor $e^{i\alpha}$ of the output state is global, and consequently we can set $\alpha = 0$ with no loss of generality.

¹³A unitary matrix V with $2^m \times 2^m$ complex elements has only 2^{2m} free (real) parameters due to the condition $V^\dagger V = I$.

¹⁴E. Knill, R. Laflamme, and G. J. Milburn, “A scheme for efficient quantum computation with linear optics,” *Nature (London)* **409**, 46–52 (2001).

¹⁵S. Scheel, K. Nemoto, W. J. Munro, and P. L. Knight, “Measurement-induced nonlinearity in linear optics,” *Phys. Rev. A* **68**, 032310-1–032310-13 (2003).