Conditional linear-optical measurement schemes generate effective photon nonlinearities

G. G. Lapaire,[1] Pieter Kok,[2] Jonathan P. Dowling,[2] and J. E. Sipe[1]
[1]*Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7*
[2]*Quantum Computing Technologies Group, Section 367, Jet Propulsion Laboratory, California Institute of Technology, Mail Stop 126-347, 4800 Oak Grove Drive, Pasadena, California 91109, USA*
(Received 26 May 2003; published 14 October 2003)

We provide a general approach for the analysis of optical state evolution under conditional measurement schemes, and identify the necessary and sufficient conditions for such schemes to simulate unitary evolution on the freely propagating modes. If such unitary evolution holds, an effective photon nonlinearity can be identified. Our analysis extends to conditional measurement schemes more general than those based solely on linear optics.

DOI: 10.1103/PhysRevA.68.042314 PACS number(s): 03.67.Lx, 42.50.−p, 42.65.−k

I. INTRODUCTION

One of the main problems that optical quantum computing has to overcome is the efficient construction of two-photon gates [1]. We can use Kerr nonlinearities to induce a phase shift in one mode that depends on the photon number in the other mode, and this nonlinearity is sufficient to generate a universal set of gates [2]. However, passive Kerr media have typically small nonlinearities (of the order of 10^{-16} cm^2 s V^{-1} [3]). We can also construct large Kerr nonlinearities using slow light, but these techniques are experimentally difficult [4].

On the other hand, we can employ linear optics with projective measurements. The benefit is that linear-optical schemes are experimentally much easier to implement than Kerr-media approaches, but the downside is that the measurement-induced nonlinearities are less versatile and the success rate can be quite low (especially when inefficient detectors are involved). However, Knill, Laflamme, and Milburn (KLM) [5] showed that with sufficient ancilla systems, these linear-optical quantum computing (LOQC) devices can be made near-deterministic with only polynomial resources. This makes linear-optics a viable candidate for quantum computing. Indeed, many linear optical schemes and approaches have been proposed since [6–11], and significant experimental progress has already been made [12,13].

The general working of a device that implements linear-optical processing with projective measurements is shown in Fig. 1. The computational input and the ancilla systems add up to N optical modes that are subjected to a unitary transformation U, which is implemented with beam splitters, phase shifters, etc. This is called an optical N-port device. In order to induce a transformation of interest on the computational input, the output is conditioned on a particular measurement outcome of the ancilla system. For example, one can build a single-photon quantum nondemolition detector with an optical N-port device [14]. In general, N-port devices have been studied in a variety of applications [15].

The class of such devices of interest here is that in which a unitary evolution on the computational input is effected. To date these devices have been proposed and studied on a more-or-less case-by-case basis. Our approach is to address this class in a more general way, and identify the conditions that such a device must satisfy to implement a unitary evolution on the computational input. Once that unitary evolution is established, an effective photon nonlinearity associated with the device can be identified.

In this paper, we present necessary and sufficient conditions for the unitarity of the optical transformation of the computational input, and we derive the effective nonlinearities that are associated with some of the more common optical gates in LOQC. We begin Sec. II by introducing the formalism. In Secs. III–V, we examine the transformation equation under the assumption that it is unitary. We show that there are two necessary and sufficient conditions for the transformation to be unitary and we provide a simple test condition. In Sec. VI, we expand the formalism and conditions to include measurement dependent output processing (see Fig. 2), which is used in several schemes. In Sec. VII, we show how the formalism can be applied to quantum computing gates. We choose as examples two quantum gates already proposed, the conditional sign flip of Knill, Laflamme, and Milburn [5], and the polarization-encoded controlled-NOT (CNOT) of Pittman *et al*. [7]. Our concluding remarks are presented in Sec. VIII, where we note that our main results extend to devices where the unitary transformation U is more general than those implementable with linear optics alone.

II. THE GENERAL FORMALISM

We consider a class of optical devices that map the computational input state onto an output state, conditioned on a particular measurement outcome of an ancilla state (see Fig. 1). We introduce a factorization of the entire Hilbert space into a space $\mathcal{H}_C$ involving the input computing channels (i.e., both ‘‘target’’ and ‘‘control’’ in a typical quantum gate), and a Hilbert space $\mathcal{H}_A$ involving the input ancilla channels,

$$\mathcal{H}=\mathcal{H}_C\otimes\mathcal{H}_A .$$

We assume that the input computing and ancilla channels are uncorrelated and unentangled, so we can write the full initial density operator as $\rho\otimes\sigma$, where ρ is the initial density operator for the computing channels, and σ is the initial density operator for the ancilla channels.

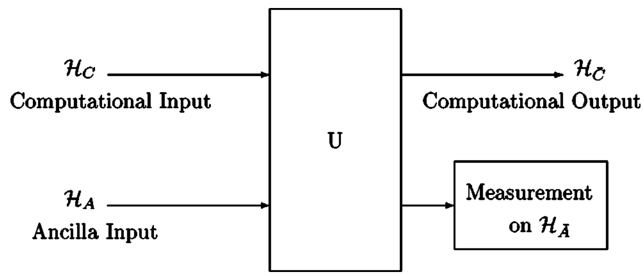

FIG. 1. A schematic diagram of a basic conditional measurement device. The input computational channels $\mathcal{H}_C$ and input ancilla channels $\mathcal{H}_A$ undergo unitary evolution. The measurement performed in the ancilla output space, $\mathcal{H}_{\bar{A}}$, indicates the success or failure of the computation.

Let U be the unitary operator describing the premeasurement evolution of the optical multiport device. At the end of this process we have a full density operator given by $U(\rho \otimes \sigma)U^\dagger$. In anticipation of the projective measurement, it is useful to introduce a new factorization of the full Hilbert space into an output computing space $\mathcal{H}_{\bar{C}}$ and a new ancilla space $\mathcal{H}_{\bar{A}}$,

$$\mathcal{H} = H_{\bar{C}} \otimes \mathcal{H}_{\bar{A}}.$$

The Von Neumann projective measurements of interest are described by projector-valued measures (or PVMs) of the type $\{\bar{P}, I - \bar{P}\}$, where I is the identity operator for the whole Hilbert space, and the projector $\bar{P}$ is of the form

$$\bar{P} = I_{\bar{C}} \otimes \sum_{k} s_{\bar{k}} |\bar{k}\rangle\langle\bar{k}|, \tag{1}$$

where $I_{\bar{C}}$ is the identity operator in $\mathcal{H}_{\bar{C}}$, and we use Roman letters with an overbar, e.g., $|\bar{k}\rangle$, to label a set of orthonormal states, $\langle\bar{k}|\bar{l}\rangle = \delta_{\bar{k}\bar{l}}$, spanning the Hilbert space $\mathcal{H}_{\bar{A}}$; each $s_{\bar{k}}$ is equal to zero or unity. The number of nonzero $s_{\bar{k}}$ identifies the rank of the projector $\bar{P}$ in $\mathcal{H}_{\bar{A}}$. "Success" is defined as a measurement outcome associated with the projector $\bar{P}$, and the probability of success is thus

$$d(\rho) \equiv \mathrm{Tr}_{\bar{C},\bar{A}}[U(\rho \otimes \sigma)U^\dagger \bar{P}]. \tag{2}$$

Clearly, in general $d(\rho)$ depends on the ancilla density operator σ, the unitary evolution U, and the projector $\bar{P}$, as well as on ρ. However, we consider the first three of these quantities fixed by the protocol of interest and thus only display the dependence of the success probability on the input density operator ρ. In the event of a successful measurement, the output of the channels associated with $\mathcal{H}_{\bar{C}}$ is identified as the computational result, and it is described by the reduced density operator

$$\bar{\rho} = \frac{\mathrm{Tr}_{\bar{A}}[\bar{P} U(\rho \otimes \sigma)U^\dagger \bar{P}]}{\mathrm{Tr}_{\bar{C},\bar{A}}[U(\rho \otimes \sigma)U^\dagger \bar{P}]}. \tag{3}$$

For any ρ with $d(\rho) \neq 0$, this defines a so-called completely positive (CP), trace preserving map $\mathcal{T}$ that takes each ρ to its associated $\bar{\rho}$: $\bar{\rho} = \mathcal{T}(\rho)$, relating density operators in $\mathcal{H}_C$ to density operators in $\mathcal{H}_{\bar{C}}$. It will be convenient to write $\mathcal{T}(\rho) = \mathcal{V}(\rho)/d(\rho)$, where

$$\mathcal{V}(\rho) \equiv \mathrm{Tr}_{\bar{A}}[\bar{P} U(\rho \otimes \sigma)U^\dagger \bar{P}] \tag{4}$$

is a linear (non-trace-preserving) CP map of density operators in $\mathcal{H}_C$ to positive operators in $\mathcal{H}_{\bar{C}}$ that is defined for all density operators ρ in $\mathcal{H}_C$. We restrict ourselves to density operators ρ over a subspace $\mathcal{S}_C$ of $\mathcal{H}_C$. This is usually the subspace in which the quantum gate operates.

As an example, consider the gate that turns the computational basis into the Bell basis. In terms of polarization states, the subspace $\mathcal{S}_C$ might be spanned by the computational basis $\{|H,H\rangle, |H,V\rangle, |V,H\rangle, |V,V\rangle\}$ (whereas $\mathcal{H}_C$ is spanned by the full Fock basis). The Bell basis on $\mathcal{S}_C$ is then given by $\{|\Psi^+\rangle, |\Psi^-\rangle, |\Phi^+\rangle, |\Phi^-\rangle\}$, where

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle \pm |V,H\rangle)$$

and

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|H,H\rangle \pm |V,V\rangle).$$

This gate is very important in quantum information theory, because it produces maximal entanglement, and its inverse can be used to perform Bell measurements. Both functions are necessary in, e.g., quantum teleportation [16]. However, it is well known that such gates cannot be constructed deterministically, and we therefore need to include an ancilla state σ and a projective measurement. We consider gates such as these in this paper.

Suppose the subspace $\mathcal{S}_C$ is spanned by a set of vectors labeled by Greek letters, e.g., $|\alpha\rangle$. We can then write

$$\rho = \sum_{\alpha,\beta} |\alpha\rangle \rho^{\alpha\beta} \langle\beta|, \tag{5}$$

where $\rho^{\alpha\beta} \equiv \langle\alpha|\rho|\beta\rangle$. We identify a convex decomposition of the ancilla density operator σ as

$$\sigma = \sum_i p_i |\chi_i\rangle\langle\chi_i|,$$

where the normalized (but not necessarily orthogonal) vectors $|\chi_i\rangle$ are elements of $\mathcal{H}_A$, and the p_i are all non-negative and sum to unity,

$$\sum_i p_i = 1.$$

We can then use Eq. (3) to write down an expression for the matrix elements of $\bar{\rho}$. Note that it is possible to work with the eigenkets of σ so that $\{|\chi_i\rangle\}$ is an orthonormal set; however, this does not simplify the analysis so we do not introduce the restriction. Furthermore, dealing with nonorthogonal states in the ancilla convex decomposition may be more convenient, depending on the system of interest. Choosing an orthonormal basis of $\mathcal{H}_{\bar{C}}$ that we label by Greek letters with overbars, e.g., $|\bar{\alpha}\rangle$, we find

$$\bar{\rho}^{\bar{\alpha}\bar{\delta}}=\sum_{\beta,\gamma}\sum_{i,\bar{k}}[W_{\bar{k},i}^{\bar{\alpha}\beta}(\rho)]\rho^{\beta\gamma}[W_{\bar{k},i}^{\bar{\delta}\gamma}(\rho)]^*, \tag{6}$$

where

$$W_{\bar{k},i}^{\bar{\alpha}\beta}(\rho)=s_{\bar{k}}\sqrt{\frac{p_i}{d(\rho)}}(\langle\bar{k}|\langle\bar{\alpha}|)U(|\beta\rangle|\chi_i\rangle).$$

Note that

$$\sum_{\bar{\alpha}}\sum_{i,\bar{k}}[W_{\bar{k},i}^{\bar{\alpha}\gamma}(\rho)]^*[W_{\bar{k},i}^{\bar{\alpha}\beta}(\rho)]=\delta_{\gamma\beta},$$

which is confirmed by

$$\mathrm{Tr}_{\bar{C}}(\bar{\rho})=\sum_{\bar{\alpha}}\bar{\rho}^{\bar{\alpha}\bar{\alpha}}=\sum_{\beta}\rho^{\beta\beta}=\mathrm{Tr}_C(\rho). \tag{7}$$

This last equation follows immediately from Eq. (3), since $\mathcal{T}$ is a trace preserving CP map and $\mathrm{Tr}_C(\rho)=1$.

In this paper, we consider a special class of maps that constitute a unitary transformation on the computational subspace $\mathcal{S}_C$. In particular, such transformations include the CNOT, the controlled sign flip (C-SIGN), and the controlled bit flip. These are not the only useful maps in linear-optical quantum computing, but they arguably constitute the most important class. Before we continue, we introduce the following definition.

Definition. We call a CP map $\rho\to\bar{\rho}=\mathcal{T}(\rho)$ an *operationally unitary transformation* on density operators ρ over a subspace $\mathcal{S}_C$ if and only if (1) for each ρ over the subspace $\mathcal{S}_C$ we have $d(\rho)\neq 0$ and (2) for each ρ defined by Eq. (5) over the subspace $\mathcal{S}_C$, the map $\mathcal{T}(\rho)$ yields a $\bar{\rho}$ given by

$$\bar{\rho}=\sum_{\alpha,\beta}|\bar{\nu}_\alpha\rangle\rho^{\alpha\beta}\langle\bar{\nu}_\beta|, \tag{8}$$

where the $|\bar{\nu}_\alpha\rangle$ are fixed vectors in $\mathcal{H}_{\bar{C}}$ satisfying $\langle\bar{\nu}_\alpha|\bar{\nu}_\beta\rangle=\langle\alpha|\beta\rangle=\delta_{\alpha\beta}$.

This forms the obvious generalization of usual unitary evolution, since it maintains the inner products of vectors under the transformation. Much of our concern in this paper is in identifying the necessary and sufficient conditions for a general map $\mathcal{T}(\rho)$ of Eqs. (3) and (6) to constitute an operationally unitary map. We begin in the following section by considering what can be said about such maps.

III. CONSEQUENCES OF OPERATIONAL UNITARITY

In this section we restrict ourselves to CP maps $\mathcal{T}(\rho)$ that are operationally unitary [see Eqs. (5) and (8)] for density operators ρ over a subspace $\mathcal{S}_C$ of $\mathcal{H}_C$. The linearity of such maps implies that the convex sum of two density operators is again a density operator:

$$\rho_c=x\rho_a+(1-x)\rho_b,$$

with $0\leqslant x\leqslant 1$. Applying Eqs. (5) and (8) to the three density operators ρ_a, ρ_b, and ρ_c it follows immediately that

$$\bar{\rho}_c=x\bar{\rho}_a+(1-x)\bar{\rho}_b. \tag{9}$$

Now a second expression for $\bar{\rho}_c$ can be worked out by using the defining relation (3) directly,

$$\begin{aligned}\bar{\rho}_c=\mathcal{T}(\rho_c)&=\frac{\mathcal{V}(\rho_c)}{d(\rho_c)}\\&=\frac{x\mathcal{V}(\rho_a)+(1-x)\mathcal{V}(\rho_b)}{xd(\rho_a)+(1-x)d(\rho_b)}\\&=\frac{xd(\rho_a)\bar{\rho}_a+(1-x)d(\rho_b)\bar{\rho}_b}{xd(\rho_a)+(1-x)d(\rho_b)},\end{aligned} \tag{10}$$

where in the second line we have used the linearity of $\mathcal{V}(\rho)$, Eq. (4), and $d(\rho)$, Eq. (2), and in the third line we have used the corresponding relations for $\bar{\rho}_a$ in terms of ρ_a, and $\bar{\rho}_b$ in terms of ρ_b. Setting the right-hand sides of Eqs. (9) and (10) equal, we find

$$x(1-x)[d(\rho_b)-d(\rho_a)](\bar{\rho}_a-\bar{\rho}_b)=0. \tag{11}$$

Since it is easy to see from Eqs. (5) and (8) that if ρ_a and ρ_b are distinct then $\bar{\rho}_a$ and $\bar{\rho}_b$ are as well; choosing $0<x<1$ it is clear that the only way the operator equation (11) can be satisfied is if $d(\rho_a)=d(\rho_b)$. But since this must hold for *any* two density operators acting over $\mathcal{S}_C$, we have established the following condition.

Condition. If a map $\mathcal{T}(\rho)$ is operationally unitary for ρ (acting on a subspace $\mathcal{S}_C$), then $d(\rho)$ is independent of ρ: $d(\rho)=d$, for all ρ acting on that subspace.

With this result in hand we can simplify Eq. (6) for a map that is operationally unitary, writing

$$\bar{\rho}^{\bar{\alpha}\bar{\delta}}=\sum_{\beta,\gamma}\sum_{J}w_J^{\bar{\alpha}\beta}\rho^{\beta\gamma}(w_J^{\bar{\delta}\gamma})^*, \tag{12}$$

where now

$$w_J^{\bar{\alpha}\beta}=w_{\bar{k},i}^{\bar{\alpha}\beta}=s_{\bar{k}}\sqrt{\frac{p_i}{d}}(\langle\bar{k}|\langle\bar{\alpha}|)U(|\beta\rangle|\chi_i\rangle)$$

is independent of ρ; we have also introduced a single label J to refer to the pair of indices $\bar{k},i$. A further simplification arises because the condition of operational unitarity guarantees that the subspace $\mathcal{S}_{\bar{C}}$ of $\mathcal{H}_{\bar{C}}$, over which the range of density operators $\bar{\rho}$ generated by $\mathcal{T}(\rho)$ act as ρ ranges over $\mathcal{S}_C$, has the same dimension as $\mathcal{S}_C$. We can thus adopt a set of orthonormal vectors $|\bar{\alpha}\rangle$ that span that subspace $\mathcal{S}_{\bar{C}}$, and the matrices $w_J^{\bar{\alpha}\beta}$ are square.

At this point we can formally construct a unitary map on $\mathcal{S}_C$: $\tilde{\rho}\equiv\mathcal{U}(\rho)$, which is isomorphic in its effect on density operators ρ with our operationally unitary map $\bar{\rho}=\mathcal{T}(\rho)$. We do this by associating each $|\overline{\alpha}\rangle$ with the corresponding $|\alpha\rangle$, introducing a density operator $\tilde{\rho}$ acting over $\mathcal{S}_C$, and putting

$$\tilde{\rho}^{\alpha\delta} \equiv \overline{\rho}^{\overline{\alpha}\overline{\delta}},$$

$$M_J^{\alpha\beta} \equiv w_J^{\overline{\alpha}\beta}. \tag{13}$$

The unitary map $\tilde{\rho} \equiv \mathcal{U}(\rho)$ is defined by the CP map

$$\tilde{\rho}^{\alpha\delta} = \sum_{\beta,\gamma} \sum_J M_J^{\alpha\beta}\rho^{\beta\gamma}(M_J^{\delta\gamma})^*,$$

or simply

$$\tilde{\rho} = \sum_J M_J \rho M_J^\dagger. \tag{14}$$

This is often what is done implicitly when describing an operationally unitary map, and we will see examples later in Sec. VII; here we find this strategy useful to simplify our reasoning below.

Since the map $\tilde{\rho} \equiv \mathcal{U}(\rho)$ is unitary it can be implemented by a unitary operator *M*,

$$\tilde{\rho} = M\rho M^\dagger,$$

where $M^\dagger = M^{-1}$. Thus $(M_1, M_2, \dots)$ and $(M,0,0,\dots)$, where we add enough copies of the zero operator so that the two lists have the same number of elements, constitute two sets of Kraus operators that implement the same map $\tilde{\rho} \equiv \mathcal{U}(\rho)$. From Nielsen and Chuang [17] we have the following theorem.

Theorem. Suppose $\{E_1, \dots, E_n\}$ and $\{F_1, \dots, F_m\}$ are Kraus operators giving rise to CP linear maps $\mathcal{E}$ and $\mathcal{F}$, respectively. By appending zero operators to the shorter list of elements we may ensure that $m=n$. Then $\mathcal{E}=\mathcal{F}$ if and only if there exist complex numbers u_{jk} such that $E_j = \Sigma_k u_{jk} F_k$, and u_{jk} is an $m \times m$ unitary matrix.

Hence, $(M_1, M_2, \dots)$ must be related to $(M,0,0,\dots)$ by a unitary matrix, and each M_J is proportional to the single operator *M*. This proof carries over immediately to the operationally unitary map $\mathcal{T}(\rho)$ under consideration, and we have the following condition.

Condition. If a map $\mathcal{T}(\rho)$ is operationally unitary for ρ acting over a subspace $\mathcal{S}_C$, then for fixed $\overline{k}$ and i the square matrix defined by

$$w_{\overline{k},i}^{\overline{\alpha}\beta} = s_{\overline{k}} \sqrt{\frac{p_i}{d}} (\langle \overline{k}|\langle\overline{\alpha}|) U(|\beta\rangle|\chi_i\rangle),$$

with $\overline{\alpha}$ labeling the row and β the column, either vanishes or is proportional to all other nonvanishing matrices identified by different $\overline{k}$ and i. We can thus define a matrix $w^{\overline{\alpha}\beta}$ proportional to all the nonvanishing $w_{\overline{k},i}^{\overline{\alpha}\beta}$ such that we can write our map (12) as

$$\overline{\rho}^{\overline{\alpha}\overline{\delta}} = \sum_{\beta,\gamma} w^{\overline{\alpha}\beta}\rho^{\beta\gamma}(w^{\overline{\delta}\gamma})^*. \tag{15}$$

It is in fact easy to show that the two *necessary* conditions we have established here for a map $\mathcal{T}(\rho)$ to be an operationally unitary transformation are also *sufficient* conditions to guarantee that it is. We show this in Sec. V. First, however, we establish a simple way of identifying whether or not $d(\rho)$ is independent of ρ.

IV. THE TEST CONDITION

In this section we consider a general map $\mathcal{T}(\rho)$ of the form of Eq. (3), and seek a simple condition equivalent to the independence of $d(\rho)$ on ρ for all ρ acting over $\mathcal{S}_C$. To do this we write $d(\rho)$ of Eq. (2) by taking the complete trace over $\mathcal{H}_C$ and $\mathcal{H}_A$ rather than over $\mathcal{H}_{\overline{C}}$ and $\mathcal{H}_{\overline{A}}$,

$$d(\rho) = \mathrm{Tr}_{C,A}[U(\rho\otimes\sigma)U^\dagger \overline{P}]$$
$$= \mathrm{Tr}_{C,A}[(\rho\otimes\sigma)U^\dagger \overline{P} U] = \mathrm{Tr}_C(\rho T),$$

where we have introduced a *test operator T* over the Hilbert space $\mathcal{H}_C$ as

$$T = \mathrm{Tr}_A(\sigma U^\dagger \overline{P} U),$$

which does not depend on ρ. The operator *T* is clearly Hermitian; it is also a positive operator, since the probability for success $d(\rho) \geqslant 0$ for all ρ. We can now identify a condition for $d(\rho)$ to be independent of ρ:

Theorem. $d(\rho)$ is independent of ρ, for density operators ρ acting over a subspace $\mathcal{S}_C$ of $\mathcal{H}_C$, if and only if the test operator *T* is proportional to the identity operator $I_{\mathcal{S}_C}$ over the subspace $\mathcal{S}_C$. We refer to this condition on *T* as the *test condition.*

Proof. The sufficiency of the test condition for a $d(\rho)$ independent of ρ is clear. Necessity is easily established by contradiction: Suppose that $d(\rho)$ were independent of ρ but *T* not proportional to $I_{\mathcal{S}_C}$. Then at least two of the eigenkets of *T* must have different eigenvalues; call those eigenkets $|\mu_a\rangle$ and $|\mu_b\rangle$. It follows that $d(\rho_a) \neq d(\rho_b)$, where $\rho_a = |\mu_a\rangle\langle\mu_a|$ and $\rho_b = |\mu_b\rangle\langle\mu_b|$, in contradiction with our assumption.

When the test condition is satisfied we denote the single eigenvalue of *T* over $\mathcal{S}_C$ as τ, i.e., $T = \tau I_{\mathcal{S}_C}$. Then $d(\rho) = \tau$, and τ is identified as the probability that the measurement indicated success. For any given protocol the calculation of the operator *T* gives an easy way to identify whether or not $d(\rho)$ is independent of ρ.

V. NECESSARY AND SUFFICIENT CONDITIONS

We can now identify necessary and sufficient conditions for a map $\overline{\rho} = \mathcal{T}(\rho)$ to be an operationally unitary map for ρ acting on a subspace $\mathcal{S}_C$ of $\mathcal{H}_C$. These are as follows.

(1) The test condition is satisfied: Namely, the operator

$$T = \mathrm{Tr}_A(\sigma U^\dagger \overline{P} U)$$

is proportional to the identity operator $I_{\mathcal{S}_C}$ over the subspace $\mathcal{S}_C$.

(2) Each matrix

$$w_{\bar{k},i}^{\bar{\alpha}\beta}=s_{\bar{k}}\sqrt{\frac{p_i}{\tau}}(\langle\bar{k}|\langle\bar{\alpha}|)U(|\beta\rangle|\chi_i\rangle),$$

identified by the indices $\bar{k}$ and i, with row and column labels $\bar{\alpha}$ and β, respectively, either vanishes or is proportional to all other such nonvanishing matrices; here τ is the eigenvalue of T.

The necessity of the first condition follows because it is equivalent to the independence of $d(\rho)$ on ρ, which was established above as a necessary condition for the transformation to be operationally unitary, as was the second condition given here. So we need only demonstrate sufficiency, which follows immediately: If the first condition is satisfied then $d(\rho)=\tau$ is independent of ρ, and if the second is satisfied then, from Eq. (12), we can introduce a single matrix $w^{\bar{\alpha}\beta}$ such that Eq. (15) is satisfied. Then

$$\sum_{\bar{\alpha}}\bar{\rho}^{\bar{\alpha}\bar{\alpha}}=\sum_{\beta,\gamma}\rho^{\beta\gamma}\sum_{\bar{\alpha}}(w^{\bar{\alpha}\gamma})^*w^{\bar{\alpha}\beta}.$$

Now the Hermitian matrix

$$Y^{\gamma\beta}\equiv\sum_{\bar{\alpha}}(w^{\bar{\alpha}\gamma})^*w^{\bar{\alpha}\beta}$$

must in fact be the unit matrix: $Y^{\gamma\beta}=\delta_{\gamma\beta}$, otherwise we would not have

$$\sum_{\bar{\alpha}}\bar{\rho}^{\bar{\alpha}\bar{\alpha}}=\sum_{\beta}\rho^{\beta\beta}$$

for an arbitrary ρ over $\mathcal{S}_C$, and we know our general map $\bar{\rho}=\mathcal{T}(\rho)$ satisfies that condition [see Eq. (7)]. Thus $w^{\bar{\alpha}\beta}$ is a unitary matrix, and from the form of Eq. (15) of the map from ρ to $\bar{\rho}$ it follows immediately that the map is operationally unitary [see Eqs. (5) and (8)].

The physics of the two necessary and sufficient conditions given above is intuitively clear, and indeed the results we have derived here could have been guessed beforehand. For if the probability for success $d(\rho)$ of the measurement were dependent of the input density operator ρ, by monitoring the success rate in an assembly of experiments all characterized by the same input ρ, one could learn something about ρ, and we would not expect operationally unitary evolution in the presence of this kind of gain of information. And the independence of the nonvanishing matrices $w_{\bar{k},i}^{\bar{\alpha}\beta}$ on $\bar{k}$ and i, except for overall factors, can be understood as preventing the "mixedness" of both the input ancilla state σ and the generally high rank projector $\bar{P}$, from degrading the operationally unitary transformation and leading to a decrease in purity.

If a map is found to be operationally unitary, we can introduce the formally equivalent unitary operator M on $\mathcal{H}_C$, as in Eq. (13), which can then be written in terms of an effective action operator Q,

$$M=e^{-iQ/\hbar}. \tag{16}$$

The operator Q can be determined simply by diagonalizing M, and its form reveals the nature of the Hamiltonian evolution simulated by the conditional measurement process. We can define an effective Hamiltonian H_{eff} that characterizes an effective photon nonlinearity acting through a time $t_{eff}$by putting $H_{eff}\equiv Q/t_{eff}$, where t_{eff} can be taken as the time of operation of the device.

In a special but common case, the input ancilla state is pure and the projector $\bar{P}$ is of unit rank in $\mathcal{H}_{\bar{A}}$. For cases such as this there is only one matrix $w^{\bar{\alpha}\beta}$ in the problem, and thus there is only a single necessary and sufficient condition for the map to be operationally unitary.

Condition. In the special case of a projector $\bar{P}$ of rank 1 in $\mathcal{H}_{\bar{A}}$, where $\bar{P}=I_{\bar{C}}\otimes|\bar{K}\rangle\langle\bar{K}|$, and a pure input ancilla state, $\sigma=|\chi\rangle\langle\chi|$, then map $\bar{\rho}=\mathcal{T}(\rho)$ is operationally unitary for ρ acting on a subspace $\mathcal{S}_C$ of $\mathcal{H}_C$ if and only if T satisfies the test condition. Here

$$T=\langle\chi|U^\dagger\bar{P}U|\chi\rangle,$$

which is an operator in $\mathcal{H}_C$. If it does satisfy this condition, then the transformation is given by

$$\bar{\rho}^{\bar{\alpha}\bar{\delta}}=\sum_{\beta,\gamma}w^{\bar{\alpha}\beta}\rho^{\beta\gamma}(w^{\bar{\delta}\gamma})^*, \tag{17}$$

where

$$w^{\bar{\alpha}\beta}=\sqrt{\frac{1}{\tau}}(\langle\bar{K}|\langle\bar{\alpha}|)U(|\beta\rangle|\chi\rangle),$$

and τ is the single eigenvalue of T over $\mathcal{S}_C$.

VI. GENERALIZATION TO INCLUDE FEED-FORWARD PROCESSING

Suppose that the measurement outcome of the ancilla does not yield the desired result, but that it signals that the output can be transformed by simply applying a (deterministic) unitary mode transformation on the output (see Fig. 2). This is called feed-forward processing and is widely used. For example, in teleportation, Alice sends Bob a classical message which allows him to correct for "wrong" outcomes of Alice's Bell measurement. Here, we can explicitly take into account feed-forward processing.

Suppose the projective measurement is characterized by a set of projectors, each identifying a different detection signature $\{\bar{P}_{(1)},\bar{P}_{(2)},\ldots\bar{P}_{(N)},\bar{P}_\perp\}$, where

$$\bar{P}_\perp=I-\sum_{L=1}^{N}\bar{P}_{(L)},$$

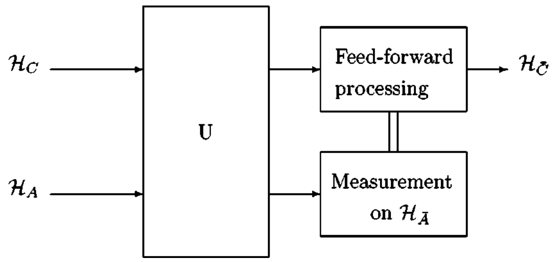

FIG. 2. A schematic diagram of a conditional measurement device that incorporates feed-forward processing. The double line connecting the two small boxes represents a classical channel that carries the measurement result. Based on the outcome, the appropriate processing is performed on the output channel.

and

$$\bar{P}_{(L)} = I_{\bar C} \otimes \sum_{\bar k} s_{L,\bar k} |\bar k\rangle\langle \bar k|.$$

All the $s_{L,\bar k}$ are equal to zero or unity, such that

$$\bar{P}_{(L)}\bar{P}_{(L')} = \bar{P}_{(L)}\delta_{LL'}.$$

Here success arises if the measurement outcome is associated with *any* of the operators $\bar{P}_{(L)}$. And if outcome L is achieved, then the computational output is processed by application of the unitary operator $\bar{V}_{(L)}$ acting over $\mathcal{H}_{\bar C}$. The probability of achieving outcome L is

$$d_{(L)}(\rho) \equiv \mathrm{Tr}_{\bar C,\bar A}[U(\rho\otimes\sigma)U^\dagger \bar{P}_{(L)}]$$

and if outcome L is achieved the feed-forward processed computational output is then

$$\bar{\rho}_{(L)} = \frac{\bar{V}_{(L)}[\mathrm{Tr}_{\bar A}\{\bar{P}_{(L)}U(\rho\otimes\sigma)U^\dagger \bar{P}_{(L)}\}]\bar{V}^\dagger_{(L)}}{\mathrm{Tr}_{\bar C,\bar A}[U(\rho\otimes\sigma)U^\dagger \bar{P}_{(L)}]},$$

which defines a map $\bar{\rho}_{(L)} = \mathcal{T}_{(L)}(\rho)$ for those ρ for which $d_{(L)}(\rho)\neq 0$. In this more general case we define the *set* of maps $\{\mathcal{T}_{(L)}\}$ to be operationally unitary for density operators ρ over the subspace $\mathcal{S}_C$ when (1) for each ρ over the subspace $\mathcal{S}_C$ at least one of the $d_{(L)}(\rho)\neq 0$, and (2) for each ρ over the subspace $\mathcal{S}_C$, for each L for which $d_{(L)}(\rho)\neq 0$ the map $\mathcal{T}_{(L)}(\rho)$ yields a $\bar{\rho}_{(L)}$ of the form of Eq. (8), independent of L.

The kind of arguments we have presented above can be extended to show that the necessary and sufficient conditions for such a set of maps to be operationally unitary for density operators ρ over the subspace $\mathcal{S}_C$ are the following.

(1) Test conditions are satisfied: The operators

$$T_{(L)} = \mathrm{Tr}_A(\sigma U^\dagger \bar{P}_{(L)} U)$$

are each proportional to the identity operator I_{SC} over the subspace $\mathcal{S}_C$. The proportionality constants $\tau_{(L)}$ need not be the same for all L.

(2) Omitting matrices associated with any L for which $\tau_{(L)}=0$, each matrix

$$w^{\bar\alpha\beta}_{L,\bar k,i} = s_{L,\bar k}\sqrt{\frac{p_i}{\tau_{(L)}}}\sum_{\bar\lambda} \bar{V}^{\bar\alpha\bar\lambda}_{(L)}(\langle \bar k|\langle\bar\lambda|)U(|\beta\rangle|\chi_i\rangle),$$

identified by the indices L, $\bar k$, and i, with row and column labels $\bar\alpha$ and β respectively, either vanishes or is proportional to all other such nonvanishing matrices.

The probability of success is $\Sigma_L \tau_{(L)} = \tau$. This expanded formalism applies to the feed-forward schemes discussed by Pittman *et al*. [13] and the teleportation schemes of Gottesman and Chuang [6]. In devices such as these, a measurement provides classical information that is used in the subsequent evolution of the output state.

In a common special case, the input ancilla state is pure, $\sigma=|\chi\rangle\langle\chi|$, and each of the projectors $\bar{P}_{(L)}$ is of unit rank in $\mathcal{H}_{\bar A}$, $\bar{P}_{(L)} = I_{\bar C}\otimes|\bar k_L\rangle\langle\bar k_L|$. Here the two necessary and sufficient conditions for the set of maps to be operationally unitary for density operators ρ over the subspace $\mathcal{S}_C$ simplify to the following.

(1) All the operators

$$T_{(L)} = \langle\chi|U^\dagger \bar{P}_{(L)} U|\chi\rangle$$

over $\mathcal{H}_C$ satisfy the test condition.

(2) Omitting matrices associated with any L for which $\tau_{(L)}=0$, each matrix

$$w^{\bar\alpha\beta}_{L} = \frac{1}{\sqrt{\tau_{(L)}}}\sum_{\bar\lambda} \bar{V}^{\bar\alpha\bar\lambda}_{(L)}(\langle \bar k_L|\langle\bar\lambda|)U(|\beta\rangle|\chi\rangle),$$

identified by the indices L, with row and column labels $\bar\alpha$ and β respectively, either vanishes or is proportional to all other such nonvanishing matrices.

If these conditions are met, then the operationally unitary transformation is given by

$$\bar\rho^{\bar\alpha\bar\delta} = \sum_{\beta,\gamma} w^{\bar\alpha\beta}_{L}\rho^{\beta\gamma}(w^{\bar\delta\gamma}_{L})^*,$$

which is independent of L.

Another extension of the standard Von Neumann, or projection, measurements is to the class of measurements described by more general positive-operator-valued measures or POVMs. These can be used to describe more complicated measurements, often resulting from imperfections in a designed PVM. Our analysis can be generalized to POVMs by expanding the ancilla space, and then describing the POVMs by PVMs in this expanded space. In some instances operationally unitarity might still be possible; in others, the extension would allow us to study of the effect of realistic limitations such as detector loss and the lack of single-photon resolution.

VII. EXAMPLES

In this section we will apply the formalism developed above to two proposed optical quantum gates for LOQC. The straightforward calculation of the effects of these gates presented in the original publications makes it clear that they are operationally unitary; our purpose here is merely to illustrate how the approach we have introduced here is applied.

To evaluate the test operators $\overline{T}_{(L)}$ and matrix elements $w_{L,\overline{k},i}^{\alpha\overline{\beta}}$ it is useful to have expression for quantities such as $Ua_{\Omega}U^{\dagger}$, where we use capital Greek letters as subscripts on the letter a to denote annihilation operators for input (computing and ancilla) channels; similarly, we use $a_{\overline{\Delta}}$ to denote annihilation operators for output (computing and ancilla) channels. We now characterize the unitary transformation U by a set of quantities $\mathsf{U}^{*}_{\Omega\overline{\Delta}}$ that give the complex amplitude for an output photon in mode $\overline{\Delta}$ given an input photon in mode Ω. That is,

$$U(a_{\Omega}^{\dagger}|\mathrm{vac}\rangle)=\sum_{\overline{\Delta}} \mathsf{U}^{*}_{\Omega\overline{\Delta}}(a_{\overline{\Delta}}^{\dagger}|\mathrm{vac}\rangle), \tag{18}$$

where $|\mathrm{vac}\rangle$ is the vacuum of the full Hilbert space $\mathcal{H}$. Since only linear optical elements are involved we have $U^{\dagger}|\mathrm{vac}\rangle=|\mathrm{vac}\rangle$, and it further follows from Eq. (18) that

$$Ua_{\Omega}^{\dagger}U^{\dagger}=\sum_{\overline{\Delta}} \mathsf{U}^{*}_{\Omega\overline{\Delta}}a_{\overline{\Delta}}^{\dagger}, \tag{19}$$

or

$$Ua_{\Omega}U^{\dagger}=\sum_{\overline{\Delta}} \mathsf{U}_{\Omega\overline{\Delta}}a_{\overline{\Delta}}. \tag{20}$$

Using the commutation relations satisfied by the creation and annihilation operators, it immediately follows that the matrix $\mathsf{U}_{\Omega\overline{\Delta}}$, which identifies the unitary transformation U, is itself a unitary matrix. Certain calculations can be simplified by its diagonalization, but for the kind of analysis of few photon states that we require this is not necessary. We will need to express, in terms of few photon states with respect to the decomposition $\mathcal{H}_{\overline{C}}\otimes\mathcal{H}_{\overline{A}}$, the result of acting with U on few photon states of the decomposition $\mathcal{H}_{C}\otimes\mathcal{H}_{A}$; this follows directly from Eq. (19). For example, denoting by $|1_{\Omega_1}2_{\Omega_2}\rangle$ the state with one photon in mode Ω_1 and two in mode Ω_2, we have

$$\begin{aligned}U|1_{\Omega_1}2_{\Omega_2}\rangle&=Ua_{\Omega_1}^{\dagger}\frac{(a_{\Omega_2}^{\dagger})^2}{\sqrt{2}}|\mathrm{vac}\rangle\\&=\frac{1}{\sqrt{2}}(Ua_{\Omega_1}^{\dagger}U^{\dagger})(Ua_{\Omega_2}^{\dagger}U^{\dagger})(Ua_{\Omega_2}^{\dagger}U^{\dagger})|\mathrm{vac}\rangle\\&=\frac{1}{\sqrt{2}}\sum_{\overline{\Delta}_1,\overline{\Delta}_2,\overline{\Delta}_3}\mathsf{U}^{*}_{\Omega_1\overline{\Delta}_1}\mathsf{U}^{*}_{\Omega_2\overline{\Delta}_2}\mathsf{U}^{*}_{\Omega_2\overline{\Delta}_3}\\&\quad\times(a_{\overline{\Delta}_1}^{\dagger}a_{\overline{\Delta}_2}^{\dagger}a_{\overline{\Delta}_3}^{\dagger})|\mathrm{vac}\rangle,\end{aligned} \tag{21}$$

and doing the sums in the last line allows us to indeed accomplish our goal.

A. KLM conditional sign flip

The first example we consider is the conditional sign flip discussed by Knill, Laflamme, and Milburn [5]. Note that in this case the input ancilla state is pure, there is no feed-forward processing, and the projector $\overline{P}$ is of unit rank in $\mathcal{H}_{\overline{A}}$. The necessary and sufficient conditions for the transformation to be operationally unitary are those of the special case discussed in Sec. V. The gate consists of one computational input port (labeled 1) and two ancilla input ports (2 and 3). The projective measurement is performed on two output ports (b,c) and the one remaining port is the computational output (a). The subspace $\mathcal{S}_C$ is spanned by the Fock states $|0\rangle$, $|1\rangle$, and $|2\rangle$ in each optical mode.

The premeasurement evolution, which is done *via* beam splitters and a phase shifter, is given by the unitary transformation U and is characterized by the matrix

$$\mathsf{U}=\mathsf{U}^{*}=\begin{bmatrix}1-\sqrt{2} & 2^{-1/4} & (3/\sqrt{2}-2)^{1/2}\\ 2^{-1/4} & 1/2 & 1/2-1/\sqrt{2}\\ (3/\sqrt{2}-2)^{1/2} & 1/2-1/\sqrt{2} & \sqrt{2}-1/2\end{bmatrix}. \tag{22}$$

The ancilla input state is

$$|\chi\rangle=a_2^{\dagger}|\mathrm{vac}_A\rangle, \tag{23}$$

denoting a single photon in the 2 mode, where $|\mathrm{vac}_A\rangle$ denotes the vacuum of $\mathcal{H}_A$. The projective measurement operator is given by

$$\overline{P}=I_{\overline{C}}\otimes|\overline{K}\rangle\langle\overline{K}|=I_{\overline{C}}\otimes a_b^{\dagger}|\mathrm{vac}_{\overline{A}}\rangle\langle\mathrm{vac}_{\overline{A}}|a_b$$

which corresponds to the detection of one and only one photon in mode b, and zero photons in mode c. The basis states that define the subspace $\mathcal{S}_C$ are

$$|0\rangle=|\mathrm{vac}_C\rangle,\quad |1\rangle=a_1^{\dagger}|\mathrm{vac}_C\rangle,\quad |2\rangle=\frac{(a_1^{\dagger})^2}{\sqrt{2}}|\mathrm{vac}_C\rangle,$$

and the basis states of $\mathcal{H}_{\overline{C}}$ are

$$|\overline{0}\rangle=|\mathrm{vac}_{\overline{C}}\rangle,\quad |\overline{1}\rangle=a_a^{\dagger}|\mathrm{vac}_{\overline{C}}\rangle,\quad |\overline{2}\rangle=\frac{(a_a^{\dagger})^2}{\sqrt{2}}|\mathrm{vac}_{\overline{C}}\rangle.$$

In order to evaluate the test function, we first write

$$U^{\dagger}\overline{P}U=\sum_{\overline{\alpha}} U^{\dagger}(a_b^{\dagger}|\mathrm{vac}_{\overline{A}}\rangle\otimes|\overline{\alpha}\rangle)(\langle\overline{\alpha}|\otimes\langle\mathrm{vac}_{\overline{A}}|a_b)U$$

and look at the matrix elements

$$\begin{aligned}&(\langle\alpha|\otimes\langle\chi|)U^{\dagger}\overline{P}U(|\chi\rangle\otimes|\beta\rangle)\\&\quad=\sum_{\overline{\alpha}} (\langle\alpha|\otimes\langle\mathrm{vac}_A|a_2)U^{\dagger}(a_b^{\dagger}|\mathrm{vac}_{\overline{A}}\rangle\otimes|\overline{\alpha}\rangle)\\&\quad\quad\times(\langle\overline{\alpha}|\otimes\langle\mathrm{vac}_{\overline{A}}|a_b)U(a_2^{\dagger}|\mathrm{vac}_A\rangle\otimes|\beta\rangle)\end{aligned} \tag{24}$$

over the computational subspace, $\mathcal{S}_C$. The calculation is straightforward. Applying the operator U on each of the states $a_2^\dagger|\mathrm{vac}_A\rangle\otimes|\beta\rangle$ gives the following states in the $\mathcal{H}_C^- \otimes \mathcal{H}_A^-$ decomposition:

$$U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|0\rangle)=\left(2^{-1/4}a_a^\dagger+\frac{1}{2}a_b^\dagger+\left[\frac{1}{2}-\frac{1}{\sqrt{2}}\right]a_c^\dagger\right)|\mathrm{vac}\rangle$$

$$U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|1\rangle)=\left(2^{-1/4}a_a^\dagger+\frac{1}{2}a_b^\dagger+\left[\frac{1}{2}-\frac{1}{\sqrt{2}}\right]a_c^\dagger\right)\times\left([1-\sqrt{2}]a_a^\dagger+2^{-1/4}a_b^\dagger+\left[\frac{3}{\sqrt{2}}-2\right]^{1/2}a_c^\dagger\right)|\mathrm{vac}\rangle,$$

$$U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|2\rangle)=\frac{1}{\sqrt{2}}\left(2^{-1/4}a_a^\dagger+\frac{1}{2}a_b^\dagger+\left[\frac{1}{2}-\frac{1}{\sqrt{2}}\right]a_c^\dagger\right)\times\left([1-\sqrt{2}]a_a^\dagger+2^{-1/4}a_b^\dagger+\left[\frac{3}{\sqrt{2}}-2\right]^{1/2}a_c^\dagger\right)^2|\mathrm{vac}\rangle,$$

and we can then separately evaluate the terms in sum (24), noting that the nonzero elements are

$$|(\langle\bar{0}|\otimes\langle\mathrm{vac}_{\bar{A}}|a_b)U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|0\rangle)|^2=\tfrac{1}{4},$$

$$|(\langle\bar{1}|\otimes\langle\mathrm{vac}_{\bar{A}}|a_b)U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|1\rangle)|^2=\tfrac{1}{4},$$

$$|(\langle\bar{2}|\otimes\langle\mathrm{vac}_{\bar{A}}|a_b)U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|2\rangle)|^2=\tfrac{1}{4}.$$

The test operator T is then

$$T=\frac{1}{4}\left[|\mathrm{vac}_C\rangle\langle\mathrm{vac}_C|+a_1^\dagger|\mathrm{vac}_C\rangle\langle\mathrm{vac}_C|a_1+\frac{(a_1^\dagger)^2}{\sqrt{2}}|\mathrm{vac}_C\rangle\langle\mathrm{vac}_C|\frac{(a_1)^2}{\sqrt{2}}\right]=\frac{1}{4}I_{\mathcal{S}C},$$

and is indeed a multiple of the unit operator in the computational input space. The probability of a success-indicating measurement is 1/4, independent of the computational input state. Since this test condition is satisfied, transformation (17) is operationally unitary. The terms of the transformation matrix $w^{\bar{\alpha}\beta}$ can be calculated noting that the nonzero $\langle\bar{K}|\langle\bar{\alpha}|U|\beta\rangle|\chi\rangle$ terms are

$$\langle\bar{K}|\langle\bar{0}|U|0\rangle|\chi\rangle=\tfrac{1}{2},$$

$$\langle\bar{K}|\langle\bar{1}|U|1\rangle|\chi\rangle=\tfrac{1}{2},$$

$$\langle\bar{K}|\langle\bar{2}|U|2\rangle|\chi\rangle=-\tfrac{1}{2},$$

and since $\tau=1/4$ the nonzero elements of the transformation matrix are

$$w^{\bar{0}0}=1,$$

$$w^{\bar{1}1}=1,$$

$$w^{\bar{2}2}=-1,$$

which corresponds to the conditional sign flip, since with probability 1/4 the gate takes the input state $|\psi\rangle=\alpha_0|0\rangle+\alpha_1|1\rangle+\alpha_2|2\rangle$ and produces the state $|\bar{\psi}\rangle=\alpha_0|\bar{0}\rangle+\alpha_1|\bar{1}\rangle-\alpha_2|\bar{2}\rangle$.

This map can be seen to exhibit an effective nonlinear interaction between the photons, since the formally equivalent unitary map (see Sec. III) is characterized by the unitary operator M (13),

$$|\tilde{\psi}\rangle=\alpha_0|0\rangle+\alpha_1|1\rangle-\alpha_2|2\rangle=M(\alpha_0|0\rangle+\alpha_1|1\rangle+\alpha_2|2\rangle),$$

which can be written in terms of an effective action operator Q (16), where we can take

$$Q=\frac{\pi\hbar}{2}(5\hat{n}-\hat{n}^2),$$

where $\hat{n}$ is the photon number operator. But such an effective action operator exists only if we restrict ourselves to the three-dimensional subspace $\mathcal{S}_C$, spanned by the kets $|0\rangle$, $|1\rangle$, and $|2\rangle$. Consider an attempt to expand this subspace to that spanned by the kets $(|0\rangle,|1\rangle,|2\rangle,|3\rangle)$. The device guarantees that a computational input of three photons can only produce a computational three-photon output, since a successful measurement requires the detection of one and only one photon in the ancilla space. The test operator is therefore still diagonal in the photon number basis. However, we find

$$|(\langle\bar{3}|\otimes\langle\mathrm{vac}_{\bar{A}}|a_b)U(a_2^\dagger|\mathrm{vac}_A\rangle\otimes|3\rangle)|^2=(2\sqrt{2}-\tfrac{5}{2})^2,$$

and thus the test operator T is no longer a multiple of the unit operator in this enlarged subspace. In this larger space the probability of a success-indicating measurement is dependent on the input, and the map is not operationally unitary.

B. Polarization encoded CNOT

The second example is the polarization-encoded Gottesman-Chuang protocol discussed by Pittman *et al.* [7]. In this case the input ancilla state is pure, there is feed-forward processing, and there are several projectors $\bar{P}_{(L)}$ of unit rank in $\mathcal{H}_{\bar{A}}^-$. The necessary and sufficient conditions for the transformation to be operationally unitary are therefore those of the special case discussed in Sec. VI. The device has two computational input ports (labeled a and b) and four ancilla input ports (1–4). A projective measurement is made on four output ports (p,q,n,m) while the two remaining ports are the computational output (5 and 6). A photon of horizontal polarization represents a logical 0, and a vertically polarized photon represents a logical 1. We use the same notation as Pittman *et al.* [7]. For example, $|H(V)_a\rangle$ represents a horizontally (vertically) polarized photon in port "a" and the Hadamard transformed modes are $|F(S)_a\rangle=\frac{1}{2}[|H_a\rangle \pm|V_a\rangle]$. The four basis states of the computational input are $|00\rangle=|H_a\rangle|H_b\rangle$, $|01\rangle=|H_a\rangle|V_b\rangle$, $|10\rangle=|V_a\rangle|H_b\rangle$, $|11\rangle=|V_a\rangle|V_b\rangle$ and the output states are labeled as $|\overline{00}\rangle=|H_5\rangle|H_6\rangle$, $|\overline{01}\rangle=|H_5\rangle|V_6\rangle$, $|\overline{10}\rangle=|V_5\rangle|H_6\rangle$, $|\overline{11}\rangle=|V_5\rangle|V_6\rangle$. The input ancilla state is

$$|\chi\rangle=\tfrac{1}{2}(|H_1\rangle|H_4\rangle|H_2\rangle|H_3\rangle+|H_1\rangle|V_4\rangle|H_2\rangle|V_3\rangle)+\tfrac{1}{2}(|V_1\rangle|H_4\rangle|V_2\rangle|V_3\rangle+|V_1\rangle|V_4\rangle|V_2\rangle|H_3\rangle),$$

and the measurement projectors $\overline{P_{(L)}}=I_{\overline{C}}\otimes|\overline{k_L}\rangle\langle\overline{k_L}|$ represent the 16 possible success outcomes:

$$|\overline{k_1}\rangle=|F_p\rangle|F_q\rangle|F_n\rangle|F_m\rangle=\tfrac{1}{4}(|H_p\rangle+|V_p\rangle)(|H_q\rangle+|V_q\rangle)(|H_n\rangle+|V_n\rangle)(|H_m\rangle+|V_m\rangle),$$

$$|\overline{k_2}\rangle=|F_p\rangle|F_q\rangle|F_n\rangle|S_m\rangle=\tfrac{1}{4}(|H_p\rangle+|V_p\rangle)(|H_q\rangle+|V_q\rangle)(|H_n\rangle+|V_n\rangle)(|H_m\rangle-|V_m\rangle),$$

$$\vdots$$

$$|\overline{k_{15}}\rangle=|S_p\rangle|S_q\rangle|S_n\rangle|F_m\rangle=\tfrac{1}{4}(|H_p\rangle-|V_p\rangle)(|H_q\rangle-|V_q\rangle)(|H_n\rangle-|V_n\rangle)(|H_m\rangle+|V_m\rangle),$$

$$|\overline{k_{16}}\rangle=|S_p\rangle|S_q\rangle|S_n\rangle|S_m\rangle=\tfrac{1}{4}(|H_p\rangle-|V_p\rangle)(|H_q\rangle-|V_q\rangle)(|H_n\rangle-|V_n\rangle)(|H_m\rangle-|V_m\rangle).$$

The polarizing beam splitters perform a unitary evolution on the input ports, characterized by the set of quantities $\mathsf{U}^*_{\Omega\overline{\Delta}}$. One can summarize the evolution of modes in $\mathcal{H}_C\otimes\mathcal{H}_A$ to modes in $\mathcal{H}_{\overline{C}}\otimes\mathcal{H}_{\overline{A}}$ with the following linear map:

$$|H_1\rangle\rightarrow|H_p\rangle,\quad |V_1\rangle\rightarrow-i|V_q\rangle,$$

$$|H_2\rangle\rightarrow|H_5\rangle,\quad |V_2\rangle\rightarrow|V_5\rangle,$$

$$|H_3\rangle\rightarrow|H_6\rangle,\quad |V_3\rangle\rightarrow|V_6\rangle,$$

$$|H_4\rangle\rightarrow|H_m\rangle,\quad |V_4\rangle\rightarrow-i|V_n\rangle,$$

$$|H_a\rangle\rightarrow|H_q\rangle,\quad |V_a\rangle\rightarrow-i|V_p\rangle,$$

$$|H_b\rangle\rightarrow|H_n\rangle,\quad |V_b\rangle\rightarrow-i|V_m\rangle,$$

since $\mathsf{U}^*_{H_1H_p}=1$, $\mathsf{U}^*_{V_1V_q}=-i$, etc. As in the previous example, to evaluate the test operators, we first look at the terms

$$U(|00\rangle|\chi\rangle)=\frac{|H_q\rangle|H_n\rangle}{2}[|H_p\rangle|H_m\rangle|H_5\rangle|H_6\rangle-i|H_p\rangle|V_n\rangle|H_5\rangle|V_6\rangle-i|V_q\rangle|H_m\rangle|V_5\rangle|V_6\rangle-|V_q\rangle|V_n\rangle|V_5\rangle|H_6\rangle],$$

$$U(|01\rangle|\chi\rangle)=\frac{-i|H_q\rangle|V_m\rangle}{2}[|H_p\rangle|H_m\rangle|H_5\rangle|H_6\rangle-i|H_p\rangle|V_n\rangle|H_5\rangle|V_6\rangle-i|V_q\rangle|H_m\rangle|V_5\rangle|V_6\rangle-|V_q\rangle|V_n\rangle|V_5\rangle|H_6\rangle],$$

$$U(|10\rangle|\chi\rangle)=\frac{-i|V_p\rangle|H_n\rangle}{2}[|H_p\rangle|H_m\rangle|H_5\rangle|H_6\rangle-i|H_p\rangle|V_n\rangle|H_5\rangle|V_6\rangle-i|V_q\rangle|H_m\rangle|V_5\rangle|V_6\rangle-|V_q\rangle|V_n\rangle|V_5\rangle|H_6\rangle],$$

$$U(|11\rangle|\chi\rangle)=\frac{-|V_p\rangle|V_m\rangle}{2}[|H_p\rangle|H_m\rangle|H_5\rangle|H_6\rangle-i|H_p\rangle|V_n\rangle|H_5\rangle|V_6\rangle-i|V_q\rangle|H_m\rangle|V_5\rangle|V_6\rangle-|V_q\rangle|V_n\rangle|V_5\rangle|H_6\rangle].$$

The matrix elements of interest are now

$$(\langle\alpha|\otimes\langle\chi|)U^\dagger\overline{P_{(L)}}U(|\chi\rangle\otimes|\beta\rangle)=\sum_{\overline{\alpha}}(\langle\alpha|\otimes\langle\chi|)U^\dagger|\overline{k_L}\rangle|\overline{\alpha}\rangle\langle\overline{\alpha}|\langle\overline{k_L}|U(|\chi\rangle\otimes|\beta\rangle)\qquad(25)$$

and the nonzero terms of the sum in Eq. (25) are

$$|\langle\overline{00}|\langle\overline{k_L}|U(|\chi\rangle\otimes|00\rangle)|^2=\tfrac{1}{16},$$

$$|\langle\overline{01}|\langle\overline{k_L}|U(|\chi\rangle\otimes|\overline{01}\rangle)|^2=\tfrac{1}{16},$$

$$|\langle\overline{11}|\langle\overline{k_L}|U(|\chi\rangle\otimes|\overline{10}\rangle)|^2=\tfrac{1}{16},$$

$$|\langle\overline{10}|\langle\overline{k_L}|U(|\chi\rangle\otimes|\overline{11}\rangle)|^2=\tfrac{1}{16},$$

for all L. The test functions $\{T_{(L)}\}$ are then

$$T_{(L)}=\tfrac{1}{64}[|H_a\rangle|H_b\rangle\langle H_b|\langle H_a|+|H_a\rangle|V_b\rangle\langle V_b|\langle H_a|+|V_a\rangle|H_b\rangle\langle H_b|\langle V_a|+|V_a\rangle|V_b\rangle\langle V_b|\langle V_a|]=\tfrac{1}{64}I_{\mathcal{SC}}$$

and are indeed multiples of the unit operator in the computational input space. In this scheme $\tau_{(L)}=1/64$, and the probability of success is the sum of the individual probabilities of the 16 detection outcomes, $\Sigma_L\tau_{(L)}=1/4$. The terms of the transformation matrices $w_L^{\alpha\beta}$ can be calculated noting that the nonzero $\langle\overline{k_L}|\langle\overline{\lambda}|U|\beta\rangle|\chi\rangle$ terms are

$$\langle\overline{k_L}|\langle\overline{00}|U|00\rangle|\chi\rangle=e^{i\phi_{L,0}}/8,$$

$$\langle\overline{k_L}|\langle\overline{01}|U|01\rangle|\chi\rangle = e^{i\phi_{L,1}}/8,$$

$$\langle\overline{k_L}|\langle\overline{11}|U|10\rangle|\chi\rangle = e^{i\phi_{L,2}}/8,$$

$$\langle\overline{k_L}|\langle\overline{10}|U|11\rangle|\chi\rangle = e^{i\phi_{L,3}}/8,$$

where $e^{i\phi_{L,0}}=1, e^{i\phi_{1,1}}=-1, e^{i\phi_{2,1}}=1,\ldots,e^{i\phi_{16,3}}=1$ are phase factors of ± 1. For this transformation to be operationally unitary, the $w_L^{\bar{\alpha}\beta}$ matrices must all be proportional to each other. In certain outcomes, single-qubit operations (π-phase shifts) are required to correct the phase factors so that the transformation is operationally unitary and the desired output is produced. The feed-forward processing matrices $\overline{V}_{(L)}^{\bar{\alpha}\bar{\lambda}}$ represent these single-qubit operations. Setting

$$\overline{V}_{(L)}^{\overline{00},\overline{00}} = e^{i\phi_{L,0}}, \quad \overline{V}_{(L)}^{\overline{01},\overline{01}} = e^{i\phi_{L,1}},$$

$$\overline{V}_{(L)}^{\overline{11},\overline{11}} = e^{i\phi_{L,2}}, \quad \overline{V}_{(L)}^{\overline{10},\overline{10}} = e^{i\phi_{L,3}},$$

with all other elements equal to zero gives the appropriate corrections. The nonzero transformation matrix elements are then

$$w_L^{\overline{00},00} = 1,$$

$$w_L^{\overline{01},01} = 1,$$

$$w_L^{\overline{11},10} = 1,$$

$$w_L^{\overline{10},11} = 1$$

for all L. Since the 16 evolution matrices are identical, the proportionality condition is satisfied. The transformation is then

$$\overline{\rho}^{\bar{\alpha}\bar{\delta}} = \sum_{\beta,\gamma} w_L^{\bar{\alpha}\beta} \rho^{\beta\gamma} (w_L^{\bar{\delta}\gamma})^*$$

which is the CNOT operation. This gate takes the input state $|\psi\rangle = \alpha_0|00\rangle + \alpha_1|01\rangle + \alpha_2|10\rangle + \alpha_3|11\rangle$ and produces the state $\alpha_0|\overline{00}\rangle + \alpha_1|\overline{01}\rangle + \alpha_2|\overline{11}\rangle + \alpha_3|\overline{10}\rangle$ with probability 1/4. Again, this map exhibits an effective nonlinear interaction between the photons since the formally equivalent unitary map is characterized by a nonlinear effective action operator Q (16). In this case one could choose

$$Q = \frac{\pi\hbar}{2}(3 + a_b^\dagger(1-\hat{n}_b) + (1-\hat{n}_b)a_b)\hat{n}_a.$$

where the operators are associated with the vertical polarization of the respective mode. Again, however, the operational unitarity is restricted to the subspace. Suppose we expand the computational subspace to include an extra photon in one of the input modes. As an example, consider the special state $|S\rangle = |H_a\rangle|H_b\rangle|H_b\rangle$. The form of the projectors indicates that the detection events involve one and only one photon in the appropriate modes. Evaluating the corresponding test operator elements we find

$$|\langle\overline{\alpha}|\langle\overline{k_L}|U(|\chi\rangle\otimes|S\rangle)|^2 = 0,$$

since the extra photon inhibits a success-indicating measurement result. The evolution cannot be operationally unitary in this expanded subspace because the test operator is no longer proportional to the unit operator.

VIII. CONCLUSION

In this paper we introduced a general approach to the investigation of conditional measurement devices. We considered an important class of optical N-port devices, including those employing projectors of rank greater than unity, mixed input ancilla states, multiple success outcomes, and feed-forward processing. We also sketched how more general POVMs, rather than PVMs, could be included. The necessary and sufficient conditions for these devices to simulate unitary evolution have been derived. They are not surprising, and indeed from a physical point of view are fairly obvious. But to our knowledge they have not been discussed in this general way before. One of the conditions is that the probability of each successful outcome must be independent of the input density operator. Whether or not this holds can be checked by evaluating a set of test operators over the input computational Hilbert space, which is easily done for any proposed device. In the special case of only one successful outcome there is only one test operator to be computed; furthermore, if the ancilla state is pure and the success projector of rank 1, then the passing of a test condition by that single test operator guarantees that the map is operationally unitary. In the case of more than one successful outcome it is a necessary consequence of operational unitarity that each of the test operators pass the test condition. This is not sufficient to imply operational unitarity in the multiple projector case unless the proportionality condition is also satisfied. The proportionality condition can often be satisfied by introducing feed-forward processing.

Besides application in the analysis of particular proposed devices, we believe the general framework presented here will be useful in exploring the different types of premeasurement evolution and measurements that might be useful in the design, optimization, and characterization of such devices. In particular, the conditional sign flip and polarization-encoded CNOT devices we considered functioned as operationally unitary maps only over the input computational subspaces for which they were originally proposed. So while effective photon nonlinearities could be introduced, the degree to which they are physically meaningful is somewhat limited. An outstanding issue, perhaps even of interest more from the general perspective of nonlinear optics than from that of quantum computer design, is the study of potential devices that provide effective photon nonlinearities over much larger input computational subspaces. The question remains: to what extent are such devices possible in theory and feasible in practice?

Finally, we note that only in Sec. VII did we assume that the premeasurement unitary evolution U is associated with linear elements in an optical system. The more general framework of the earlier sections may find application in describing other proposed devices for quantum information processing that involve conditional measurement schemes in the presence of more complicated interactions [18].

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Walter C. Sumner Foundation. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. In addition, P.K. acknowledges the United States National Research Council. Support was also received from the Advanced Research and Development Activity, the National Security Agency, the Defense Advanced Research Projects Agency, and the Office of Naval Research. We would like to thank Alexei Gilchrist, James Franson, and Gerard Milburn for stimulating discussions.

[1] P. Kok and S.L. Braunstein, Phys. Rev. A **62**, 064301 (2000).
[2] S. Lloyd and S.L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999); G.M. D'Ariano, C. Macchiavello, and L. Maccone, Fortschr. Phys. **48**, 573 (2000).
[3] A.L. Gaeta and R.W. Boyd, in *Atomic, Molecular and Optical Physics Handbook*, edited by G.W.F. Drake (AIP, Woodbury, NY, 1996), p. 809; R.W. Boyd, J. Mod. Opt. **46**, 367 (1999).
[4] M.D. Lukin and A. Imamoğlu, Phys. Rev. Lett. **84**, 1419 (2000).
[5] E. Knill, R. Laflamme, and G.J. Milburn, Nature (London) **409**, 46 (2001).
[6] D. Gottesman and I.L. Chuang, Nature (London) **402**, 390 (1999).
[7] T.B. Pittman, B.C. Jacobs, and J.D. Franson, Phys. Rev. A **64**, 062311 (2001).
[8] T.C. Ralph, A.G. White, W.J. Munro, and G.J. Milburn, Phys. Rev. A **65**, 012314 (2001).
[9] X. Zou, K. Pahlke, and W. Mathis, Phys. Rev. A **65**, 064305 (2002).
[10] M. Koashi, T. Yamamoto, and N. Imoto, Phys. Rev. A **63**, 030301 (2001).
[11] S. Scheel, K. Nemoto, W.J. Munro, and P.L. Knight, e-print quant-ph/0305082.
[12] T.B. Pittman, B.C. Jacobs, and J.D. Franson, Phys. Rev. A **66**, 052305 (2002).
[13] T.B. Pittman, B.C. Jacobs, and J.D. Franson, Phys. Rev. Lett. **88**, 257902 (2002).
[14] P. Kok, H. Lee, and J.P. Dowling, Phys. Rev. A **66**, 063814 (2002).
[15] H. Paul, P. Törma, T. Kiss, and I. Jex, Phys. Rev. Lett. **76**, 2464 (1996); H.Y. Fan and M. Xiao, Quantum Semiclassic. Opt. **9**, 53 (1997); P. Kok and S.L. Braunstein, Phys. Rev. A **63**, 033812 (2001).
[16] C.H Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[17] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000), p. 372.
[18] See, e.g., A. Beige, D. Braun, B. Tregenna, and P.L. Knight, Phys. Rev. Lett. **85**, 1762 (2000); A. Beige, Phys. Rev. A **67**, 020301(R) (2003).