Quantum gates using linear optics and postselection

E. Knill*
MS B256, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Received 5 August 2002; published 14 November 2002)

Recently it was realized that linear optics and photodetectors with feedback can be used for theoretically efficient quantum information processing. The first of three steps toward efficient linear optics quantum computation is to design a simple postselected gate that implements a nonlinear phase shift on one mode. Here a computational strategy is given for finding postselected gates for bosonic qubits with helper photons. A more efficient conditional sign flip gate is obtained. What is the maximum efficiency for such gates? This question is posed and it is shown that the probability of success cannot be 1.

DOI: 10.1103/PhysRevA.66.052306 PACS number(s): 03.67.Lx, 42.50.−p

I. INTRODUCTION

Now that we know that linear optics and photodetectors are sufficient for quantum information processing [1,2], it is necessary to investigate how the required schemes can be realized more efficiently. One promising direction is to use superpositions of squeezed or coherent states for encoding qubits [2,3]. In this paper, it is shown how the postselected gates at the foundation of the constructions in [1] can be found and improved. Other relevant work in this direction includes [4–7], where networks suitable for experimental realization are given. The focus of this paper is on what can be done in principle. To that end, a systematic method is given for finding postselected gates based on a combination of algebraic solution finding, exploitation of known symmetries, and numerical optimization. By using the method, a conditional sign flip for bosonic qubits that succeeds with probability 1/13.5 using two helper photons is found. This improves the one in [1], which succeeds with probability 1/16. To conclude this paper, the following question is considered: What is the optimum probability of success for any number of helper photons? (See the problem at the end of the paper.) A characterization of states that can be obtained from helper photons with passive linear optics and no postselection is given. This characterization implies that the probability cannot be 1, a result related to known bounds on Bell measurements [8,9].

II. PRELIMINARIES

The physical system of interest consists of optical modes, each of whose state space is spanned by the number states $|0\rangle,|1\rangle,|2\rangle,\ldots$. If more than one mode is used, they are distinguished by labels. For example, $|k\rangle_{\mathsf{r}}$ is the state with k photons in the mode labeled r. The Hermitian transpose of this state is denoted by ${}^{\mathsf{r}}\langle k|$. The vacuum state for a set of modes has each mode in the state $|0\rangle$ and is denoted by $|\mathbf{0}\rangle$. The annihilation operator for mode r is written as $\mathbf{a}^{(\mathsf{r})}$ and the creation operator as $\mathbf{a}^{\dagger(\mathsf{r})}=(\mathbf{a}^{(\mathsf{r})})^{\dagger}$. Recall that $\mathbf{a}^{\dagger(\mathsf{r})}|m\rangle=\sqrt{m+1}|m+1\rangle$. Labels are omitted when no ambiguity results. Hamiltonians that are at most quadratic in creation and annihilation operators generate the group of linear optics transformations. Among these, the ones that preserve the particle number are called passive linear. Every passive linear optics transformation can be achieved by a combination of beam splitters and phase shifters. If U is passive linear, then $U\mathbf{a}^{\dagger(\mathsf{r})}|\mathbf{0}\rangle=\Sigma_s u_{sr}\mathbf{a}^{\dagger(\mathsf{s})}|\mathbf{0}\rangle$, where u_{sr} defines a unitary matrix $\hat{u}$. Conversely, for every unitary matrix $\hat{u}$, there is a corresponding passive linear optics transformation [10]. For the remainder of this paper, all linear optics transformations are assumed to be passive.

III. CONDITIONAL PHASE SHIFTS

A conditional phase shift by θ on two modes is the map $\mathrm{CS}_\theta:|ab\rangle\rightarrow e^{i(ab)\theta}|ab\rangle$ for $0\leq a,b\leq 1$. These phase shifts can be used to implement conditional sign flips on two bosonic qubits. A bosonic qubit $\mathsf{Q}_{\mathsf{r},\mathsf{s}}$ is defined by identifying a qubit's logical $|0\rangle$ with $|01\rangle_{\mathsf{rs}}$ and logical $|1\rangle$ with $|10\rangle_{\mathsf{rs}}$. The modes r and s can be two distinct spatial modes or the two polarizations of one spatial mode. A key gate in quantum information processing is the controlled-NOT (see, for example, [11]), which "flips" the second qubit if the first qubit is in the state $|1\rangle$. A gate that is equivalent to the controlled-NOT up to one-qubit transformations is the conditional sign flip, which changes the sign of $|11\rangle$. To realize the conditional sign flip between two bosonic qubits $\mathsf{Q}_{1,2}$ and $\mathsf{Q}_{3,4}$, apply CS_{180° to modes $\mathsf{1}$ and $\mathsf{3}$. The bosonic qubit controlled-NOT can then be implemented using conditional sign flips and single qubit rotations, which are realizable with beam splitters. Note that all one-qubit gates can be realized with linear optics on bosonic qubits.

In [1], conditional sign flips were implemented indirectly using a postselected (referred to there as "nondeterministic") realization of the map

$$\mathrm{NS}:\alpha|0\rangle+\beta|1\rangle+\gamma|2\rangle\rightarrow\alpha|0\rangle+\beta|1\rangle-\gamma|2\rangle \qquad (1)$$

that succeeds with probability 1/4. The realization of NS requires one helper photon and two ancilla modes. The implementation of CS_{180° in [1] requires two instances of NS, resulting in a probability of success of 1/16. The goal is to implement CS_{180° more efficiently directly using two helper photons. One helper photon can be shown to be insufficient by means of the same algebraic method about to be

*Electronic address: knill@lanl.gov

used. Let modes 1 and 2 contain the state to which CS_θ is to be applied. The basic scheme is to start with two ancilla modes 3 and 4 initialized with one photon each, apply a linear optics transformation to modes $1,2,\ldots,k$ with $k \geqslant 4$, measure all but the first two of these modes, and accept only a predetermined outcome, say where one photon is detected in each of modes 3 and 4 and none in the added modes. Let $\hat{u}$ be the unitary matrix associated with the linear optics transformation, with u_{sr} the entries of $\hat{u}$. The postselected final state is determined completely by the 4×4 upper left submatrix V of $\hat{u}$ with entries $V_{rs}=u_{sr}$ for $s,r\leqslant 4$.

It is necessary to consider the effects of the scheme on the initial states $|00\rangle_{12},|01\rangle_{12},|10\rangle_{12},|11\rangle_{12}$. Since photon number is conserved, we have, without renormalization,

$$|00\rangle_{12}\rightarrow \alpha_{0000}|00\rangle_{12}, \tag{2}$$

$$|01\rangle_{12}\rightarrow \alpha_{0101}|01\rangle_{12}+\alpha_{0110}|10\rangle_{12}, \tag{3}$$

$$|10\rangle_{12}\rightarrow \alpha_{1010}|10\rangle_{12}+\alpha_{1001}|01\rangle_{12}, \tag{4}$$

$$|11\rangle_{12}\rightarrow \alpha_{1111}|11\rangle_{12}+\alpha_{1120}|20\rangle_{12}+\alpha_{1102}|02\rangle_{12}. \tag{5}$$

To be successful, the amplitudes have to satisfy

$$\alpha_{0110}=\alpha_{1001}=\alpha_{1120}=\alpha_{1102}=0, \tag{6}$$

$$\alpha_{1010}=\alpha_{0101}=\alpha_{0000}, \tag{7}$$

$$\alpha_{1111}=e^{i\theta}\alpha_{0000}. \tag{8}$$

The probability of success is $|\alpha_{0000}|^2$. The amplitudes are polynomials of the coefficients of V. For example, $\alpha_{0000}=v_{33}v_{44}+v_{34}v_{43}$. More generally, define $p_s=\Sigma_r v_{rs}\mathbf{a}^{\dagger(\mathsf{r})}$. If the initial state in mode s has d_s photons, then the output state is given by $\Pi_s p_s^{d_s}|\mathbf{0}\rangle$. Let $P=\Pi_s p_s^{d_s}$. Thus, P is a polynomial of the $\mathbf{a}^{\dagger(\mathsf{r})}$. If β is the coefficient of the monomial $\Pi_t(\mathbf{a}^{\dagger(\mathsf{t})})^{m_t}$ in P, then the output amplitude for having m_t photons in each mode t is given by $\sqrt{\Pi_t(m_t!)}\beta$. The coefficient β is a polynomial of the coefficients of the p_s. This shows that the output amplitudes α_{abcd} are polynomials of the v_{rs}.

The first step for constructing CS_θ is to solve Eqs. (6)–(8), which are polynomial identities in the v_{rs}. Before showing how to reduce the difficulty of doing that, let us see how to proceed from there. Since there are 16 free complex variables, the solution will have a number of remaining free variables that must be chosen to optimize the probability of success $|\alpha_{0000}|^2$ and to satisfy one more constraint: The solution must be an (explicit) matrix V that needs to be extended to a unitary matrix $\hat{u}$. This is possible if and only if the maximum singular value (that is the square root of the maximum eigenvalue of $V^\dagger V$) is at most 1. The extension is not unique. One can set the first four columns of $\hat{u}$ to the matrix with orthonormal columns,

$$X=\begin{pmatrix} V \\ (\mathbf{I}-V^\dagger V)^{1/2}\end{pmatrix}, \tag{9}$$

and then complete the last four columns with any orthonormal basis of the orthogonal complement of the space spanned by the columns of X. The maximum singular value constraint ensures that one can compute the square root in the expression for X. If some of the singular values of V are equal to 1, then fewer than four additional columns and rows can be used.

The singular value constraint cannot be easily achieved using algebraic methods. In principle, one can reparametrize the matrix V to guarantee the constraint, for example by using the polar decomposition and an Euler angle representation of unitary matrices. In the case in which CS_θ is to be applied to the "left" modes of a pair of bosonic qubits, the singular value constraint can be removed by exploiting a rescaling symmetry. In this situation, there are two additional modes contributed by the bosonic qubits. The total number of photons is always four. Let V be a matrix whose coefficients satisfy the identities for the α_{abcd}. Let $\lambda=\lambda(V)$ be the maximum singular value of V and consider the matrix

$$V_e=\frac{1}{\lambda}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & V\end{pmatrix}, \tag{10}$$

where $\mathbf{I}$ in the upper left corner of V_e is a 2×2 identity matrix whose indeces are associated with the two other ("right") qubit modes. V_e has maximum singular value 1 and can be extended to a unitary $\hat{u}_e$ as before. The claim is that if the resulting optics operation is applied to the pair of bosonic qubits with the same postselection procedure, it has the intended effect with probability $1/\lambda^8$. To see that this is true, first observe that $V'=\lambda V_e$ satisfies the polynomial equations obtained by requiring that the operation works correctly for the pair of bosonic qubits. The amplitudes [as in Eqs. (2)–(5)] that occur in these equations are polynomials which are either homogeneous linear in the coefficients of a given column of V' or independent of them. (A polynomial is homogeneous of degree d if each monomial has total degree exactly d in the variables.) This is because the input states have at most one photon in each mode. Because each input state under consideration has exactly four photons, the amplitudes are all homogeneous of degree 4 in the coefficients of V'. This implies that multiplying V' by δ scales the amplitudes by δ^4. Since the equations to be satisfied are homogeneous linear in the amplitudes, every scalar multiple of the matrix also satisfies the equations.

With the observation of the previous paragraph, instead of trying to satisfy the singular value constraint, one can recalculate the probability of success by dividing V's probability of success by λ^8 before optimizing. Note that this works even if $\lambda<1$. In computer experiments using naive optimization methods (see below), this usually led to solutions V with $\lambda=1$ for $\theta=180°$ and $\theta=90°$.

To simplify solving the equations for the α_{abcd}, one can use scaling symmetries to standardize V. Since each of the α_{abcd} is homogeneous in the variables associated with any one row or column of V, the equations of the form $\alpha_{abcd}=0$ are satisfied for any rescaling of a row or column. The nonzero α_{abcd} are homogeneous of degree 1 in the third and

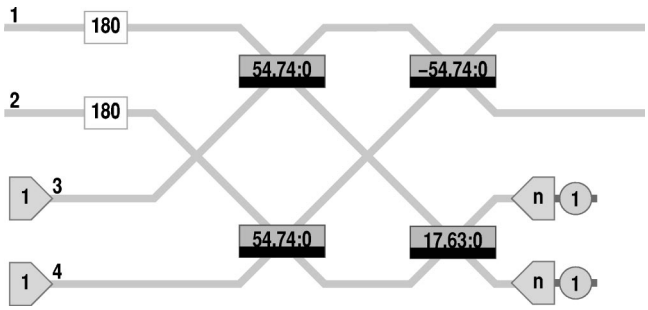

FIG. 1. Optical network realizing $CS_{180°}$. The notation is as explained in [1]. The lines denote the time lines of optical modes. Modes 1 and 2 are the input modes; the first gates applied to them are linear phase shifts $\exp(i\pi \mathbf{a}^\dagger \mathbf{a})$ that map $|1\rangle \rightarrow -|1\rangle$. Modes 3 and 4 are prepared in their one-photon states; the elements applied to these modes at the end of the operation are photon counters, where the outcome is conditioned on the indicated classical output (in circles). The elements that mix the modes in the middle are beam splitters, whose action $\exp(-i(e^{i\phi}\mathbf{a}^{\dagger(\mathrm{A})}\mathbf{a}^{(\mathrm{B})} + e^{-i\phi}\mathbf{a}^{(\mathrm{A})}\mathbf{a}^{\dagger(\mathrm{B})})\theta)$ is determined by the two phases $\theta;\phi$ given in degrees. In each instance, modes A and B are the top and bottom modes, respectively, at both the input and the output of the beam splitter.

fourth column (because of the presence of the helper photons at the input in modes 3 and 4) and in the third and fourth row (because of the postconditioning on detecting exactly one photon in each of modes 3 and 4). Because Eqs. (7) and (8) are homogeneous linear in the amplitudes, rescaling these rows or columns preserves the identities. The nonzero α_{abcd} satisfy that they are of equal degree and homogeneous in the first column and (separately) in the first row. This is due to the fact that when a photon is present in mode 1 at the input, this is designed to be the case at the output too. Thus multiplying the first column by δ and the first row by $1/\delta$ does not change the values. Similarly, this rescaling can be used on the second column and row.

The scaling rules of the previous paragraph can be used to introduce unconstrained scaling variables and standardize the entries of V. For example, one can take $v_{13}=v_{24}=v_{33}=v_{44}=v_{43}=1$. Note that this choice implies that solutions where any one of these variables is 0 are not easily found. It may therefore be necessary to try solving with some of the variables set to 0. For example, the CS gate of [1] after translation into the form used here satisfies $v_{43}=0$. I did not find any solutions satisfying this constraint with better probability of success.

MATHEMATICA was used to solve the equations (any other computer algebra system would do equally well). The strategy was to solve linear equations first and then to simplify expressions. MATHEMATICA notes can be found in the appendix of the online version of this paper [12] and include formulas for the solution found. The solution can be expressed in terms of the remaining variables of the last two columns of V and one additional variable. After some experimentation, it seemed that in all the best solutions, $v_{11}=v_{22}$. This was exploited in the final version of the optimization procedure, implemented in MATLAB (the programs are available by request). Briefly, the function to be optimized takes as input the remaining free complex variables $(v_{14},v_{23},v_{34},l_1)$, and a nonredundant subset of the scaling variables. To avoid infinities, one can provide the logarithms of the scaling variables as inputs. The scales can be taken to be real since phases have no effect on the probability of success. The function can then be optimized using random starting points. With the optimization procedures provided by MATLAB, it was found useful to randomly perturb the point returned and repeat until the solution no longer changes significantly. This procedure routinely finds the same optimum. For $\theta=180°$, it was possible to guess the algebraic numbers to which it converged. Here is a version of the matrix found, which turns out to be unitary,

$$V_{180°}=\begin{pmatrix} -1/3 & -\sqrt{2}/3 & \sqrt{2}/3 & 2/3 \\ \sqrt{2}/3 & -1/3 & -2/3 & \sqrt{2}/3 \\ -\sqrt{3+\sqrt{6}}/3 & \sqrt{3-\sqrt{6}}/3 & -\sqrt{(3+\sqrt{6})/2}/3 & \sqrt{1/6-1/(3\sqrt{6})} \\ -\sqrt{3-\sqrt{6}}/3 & -\sqrt{3+\sqrt{6}}/3 & -\sqrt{1/6-1/(3\sqrt{6})} & -\sqrt{(3+\sqrt{6})/2}/3 \end{pmatrix}. \tag{11}$$

The probability of success is 2/27. The matrix can be systematically decomposed into elementary beam splitter and phase-shift operators [10]. An optical network realizing it is shown in Fig. 1. The implementation uses fewer elements (four beam splitters, four modes, two photon counters, probability of success 2/27) than the solution in [1] (six beam splitters, six modes, two photon counters, two photodetectors, probability of success 1/16). As before, the counters must be able to distinguish between zero, one, or more than one photon.

A matrix that can be extended to obtain $CS_{90°}$ by postselection was also obtained,

$$V_{90°}=\begin{pmatrix} -0.3202+0.0418i & -0.2520-0.3226i & 0.2883 & -0.1292-0.7221i \\ -0.2520-0.3226i & -0.3202+0.0418i & -0.1292-0.7221i & 0.2883 \\ -0.3216+0.7210i & -0.1711-0.1725i & 0.2469 & 0.3322+0.3285i \\ -0.1711-0.1725i & -0.3216+0.7210i & 0.3322+0.3285i & 0.2469 \end{pmatrix}. \tag{12}$$

The probability of success for this solution is 1/19.37. A ''nice'' beam splitter decomposition of this matrix was not found. This is partly due to the fact that because only two of the singular values are (close to) 1, at least two extra modes must be added for the unitary completion. The simplest method of decomposing a 6×6 unitary matrix normally requires 15 beam splitters.

It is an open problem to determine whether the above solutions are indeed optimal, as is suggested by the results of the numerical experiments.

IV. BOUNDS ON CONDITIONAL PHASE SHIFTS?

To obtain bounds on the probability of success of a phase-shift gate implemented with helper photons, one can attempt to use a characterization of the states obtained in the output modes after tracing out the helper modes. The purpose of this section is to obtain such a characterization and to show that the phase-shift gate cannot be implemented with probability of success 1. For obtaining a bound, the initial state of the modes that the gate is applied to can be chosen arbitrarily. Assume that this is a state obtained by applying linear optics to prepared single photons. In this case, the final state after a linear optics transformation is given by

$$|\psi_f\rangle=\prod_{k=1}^{n}(\alpha_{k1}\mathbf{a}^{\dagger(1)}+\cdots+\alpha_{km}\mathbf{a}^{\dagger(\mathsf{m})})|\mathbf{0}\rangle. \tag{13}$$

The goal is to show that after tracing out modes $m'+1,\ldots,m$, the state in the remaining modes is a mixture of states of the form

$$\prod_{k=1}^{n}(\beta_{k0}+\beta_{k1}\mathbf{a}^{\dagger(1)}+\cdots+\beta_{km'}\mathbf{a}^{\dagger(\mathsf{m}')})|\mathbf{0}\rangle. \tag{14}$$

In fact, this is the case even if the final state before tracing out is also of this form, which is more general than the form in Eq. (13). To be explicit, add to the factors in the expression for $|\psi_f\rangle$ any constant terms α_{k0} so that

$$|\psi_f\rangle=\prod_{k=1}^{n}(\alpha_{k0}+\alpha_{k1}\mathbf{a}^{\dagger(1)}+\cdots+\alpha_{km}\mathbf{a}^{\dagger(\mathsf{m})})|\mathbf{0}\rangle. \tag{15}$$

First trace out mode m. Given a set of states $\{|\gamma\rangle_{\mathsf{m}}\}_\gamma$ from which one can form a partition of unity $\mathbf{I}=\int d\mu(\gamma)|\gamma\rangle_{\mathsf{m}}^{\mathsf{m}}\langle\gamma|$ for some measure μ, the state of the remaining modes $1\cdots\mathsf{m}-1$ can be expressed as a mixture of the (unnormalized) states ${}^{\mathsf{m}}\langle\gamma|\psi_f\rangle$. Choosing as the set of states the coherent states and using the fact that for these states ${}^{\mathsf{m}}\langle\gamma|\mathbf{a}^{\dagger(\mathsf{m})}={}^{\mathsf{m}}\langle\gamma|\overline{\gamma}$, the mixture consists of states of the form

$$\prod_{k=1}^{n}(\alpha_{k0}+\overline{\gamma}\alpha_{km}+\alpha_{k1}\mathbf{a}^{\dagger(1)}+\cdots+\alpha_{k(m-1)}\mathbf{a}^{\dagger(\mathsf{m}-1)})|\mathbf{0}\rangle_{1\cdots(\mathsf{m}-1)}. \tag{16}$$

Iterating this procedure proves the desired result.

Consider the conditional sign-flip gate. With this gate and using a few beam splitters, one can map the state $|1100\rangle$ to the state $1/\sqrt{2}(|1100\rangle+|0011\rangle)=1/\sqrt{2}(\mathbf{a}^{\dagger(1)}\mathbf{a}^{\dagger(2)}+\mathbf{a}^{\dagger(3)}\mathbf{a}^{\dagger(4)})$, an entangled photon state. By the above, before postselection on a measurement of the other modes and with n helper photons, the state can be written as a mixture of products of linear expressions in the creation operators. Therefore, to obtain a bound on the probability of success, it suffices to obtain a bound on the overlap of (normalized) such states with the Bell state. Because the normalized overlap of $(\mathbf{a}^{\dagger(1)}+\mathbf{a}^{\dagger(3)})(\mathbf{a}^{\dagger(2)}+\mathbf{a}^{\dagger(4)})$ with the Bell state is $1/\sqrt{2}$, the bound on the probability of success thus obtained can be no smaller than 1/2. It is clear that the probability of success cannot be made equal to 1: The polynomial $xy+uv$ associated with the creation operators in the Bell state cannot be factored.

A problem suggested by the above is as follows.

Problem. What is the maximum probability of success for implementing CS_θ using linear optics with at most k independently prepared helper photons and postselection from photon counters without feedback?

It was shown that for $\theta=180°$, a probability of success of 1 is not possible, but for $k\geqslant 2$, 1/13.5 can be realized. A variant of the problem asks the same question for conditional sign shifts of two bosonic qubits (a four-mode operation). Other directions for investigation are to determine what improvements are possible if active linear optics operations can be used, or if initial states such as prepared entangled photon pairs [6] or photon number states like $|2\rangle$ are available.

ACKNOWLEDGMENTS

Thanks to A. Imamoğlu and I. Wilson-Rae for motivation and stimulating discussions. This work was partially supported by the NSA and the DOE (Contract No. W-7405-ENG-36).

[1] E. Knill, R. Laflamme, and G. Milburn, Nature (London) **409**, 46 (2001).
[2] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310/1–21 (2001).
[3] T.C. Ralph, W.J. Munro, and G.J. Milburn, e-print quant-ph/0110115.
[4] T.C. Ralph, A.G. White, W.J. Munro, and G.J. Milburn, e-print quant-ph/0108049.
[5] T. Rudolph and J.-W. Pan, e-print quant-ph/0108056.
[6] T.B. Pittman, B.C. Jacobs, and J.D. Franson, e-print quant-ph/0107091.
[7] T.B. Pittman, B.C. Jacobs, and J.D. Franson, e-print quant-ph/0109128.
[8] N. Lütkenhaus, J. Calsamiglia, and K.-A. Suominen, Phys. Rev. A **59**, 3295 (1999).
[9] J. Calsamiglia, e-print quant-ph/0108108.

[10] M. Reck, A. Zeilinger, H.J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[11] A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).

[12] E. Knill, A note on linear optics gates by post-selection, Technical Report LAUR-01-5973, Los Alamos National Laboratory, 2001, e-print quant-ph/0110144.