Optimizing Optical Quantum Logic Gates using Genetic Algorithms

Zhanhan Wu1, Sean D. Huver1, Dmitry Uskov2, Hwang Lee1, and Jonathan P. Dowling1
1Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803
2Department of Physics and Astronomy, Tulane University, New Orleans, Louisiana 70118

(Dated: February 1, 2008, V1.1)

We introduce the method of using an annealing genetic algorithm to the numerically complex problem of looking for quantum logic gates which simultaneously have highest fidelity and highest success probability. We first use the linear optical quantum nonlinear sign (NS) gate as an example to illustrate the efficiency of this method. We show that by appropriately choosing the annealing parameters, we can reach the theoretical maximum success probability (1/4 for NS) for each attempt. We then examine the controlled-z (CZ) gate as the first new problem to be solved. We show results that agree with the highest known maximum success probability for a CZ gate (2/7) while maintaining a fidelity of 0.9997. Since the purpose of our algorithm is to optimize a unitary matrix for quantum transformations, it could easily be applied to other areas of interest such as quantum optics and quantum sensors.

PACS numbers: 03.67.-a, 03.67.Lx, 42.50.Dv

Linear optics is an attractive candidate for building quantum computers in large part due to Knill, Laflamme and Milburn [1] and their scheme for non-deterministic quantum gates with projective measurement. The scheme provides a way to build elementary quantum gates with only linear optical elements. The trade off in this scheme is that we can only construct the gates with a certain success probability, i.e., they are non-deterministic. Therefore, one of the major tasks for this program is to figure out a general way to determine the design of gates which possess the highest success probability.

Any quantum gate or linear optical quantum state generator (LOQSG) can be viewed as a unitary transformation which transfers certain input states into desired output states. The goal of designing a LOQSG is to find a proper unitary matrix whose elements can then be implemented with linear optical devices [2, 3]. In this paper, we introduce genetic algorithms with a simulated annealing to the problem of optimizing a LOQSG. We first restate the problem so that it is suitable for genetic algorithms. We then briefly discuss the method of genetic algorithms with simulated annealing and a tunable control constraint. Using this method, we first test the efficiency of the algorithm with the nonlinear sign (NS) gate. We choose the NS gate due to its maximum success probability without feedforward having been theoretically proven to be 1/4 [4, 5]. We then investigate the CZ gate and try to obtain the global maximum for success probability while maintaining a high level of fidelity.

Any LOQSG can be represented as in Fig. 1. Suppose we have N input channels. They are composed with computational input states and ancilla channels. We want these inputs to be transformed to our expected output with projective measurement on the remaining ancilla ports. This process can be done by a linear optical device which we call a LOQSG. This device is an N dimensional unitary transformation. When a projective measurement determines a certain pattern of photons measured in some M < N of the modes, it is considered successful, which leads to a preparation of the desired state in the remaining modes. Therefore, the device is probabilistic and it can fail in two aspects. Firstly, the measurement does not give out the expected pattern, which leads to the failed measurement. This can be improved by increasing the success probability of the device. Secondly, it does not provide the expected output state in the computational channels even when the measurement works perfectly. We note this kind of failure as fidelity less than unity. Since we cannot measure the computational output during the computing process, we have to make sure that the fidelity is equal to one or numerically very very close to one.

The linear optical measurement-assisted transformation works as follows. We start from a computational input state \(|\psi^C\rangle\) of N - M modes, combined with ancilla state \(|\psi^A\rangle\) in M modes so that the input state \(|\Psi_{in}\rangle = |\psi^C\rangle \otimes |\psi^A\rangle\). The optical device induces a unitary transformation \(U\) of the \(|\Psi_{in}\rangle\) state. After that a number-resolving photocounting measurement is applied to the M ancilla modes. The latter is formally described by a Kraus POVM operator in ancilla modes \(\hat{P} = \text{[vacuum]}\langle k_{N-M+1}, k_{N-M+2}, ..., k_{N-M} \rangle\). The resulting transformation of the computational state \(|\psi^C_{in}\rangle\) is a contraction quantum map \(\hat{A}|\psi^C_{out}\rangle = \hat{A}|\psi^C_{in}\rangle \langle |\psi^C_{in}\rangle ||\psi^C_{in}\rangle\) [6], where the action of the linear operator \(\hat{A}\) is given by the following projection

\[
\hat{A}|\psi^C_{in}\rangle = \langle k_{N-M+1}, k_{N-M+2}, ..., k_{N-M} | \hat{U} |\psi^C_{in}\rangle \otimes |\psi^A_{in}\rangle. \tag{1}
\]
In the context of the LOQSG problem, operator $\hat{A}$ contains all the information of state transformation.

The optical interferometer is considered formally as canonical transformation of creation operators $a_i^\dagger \to U_{ij}a_j^\dagger$ induced by an $N \times N$ unitary matrix $U$. If the input state is given in the Fock representation as $|\Psi_{in}\rangle = |n_1, n_2, ..., n_N-M, 0\rangle \otimes |n_M+1, n_{M+2}, ..., n_N\rangle$, the unitary transformation $\hat{U}$ in equation (1) is given by

$$|\Phi_{out}\rangle = \hat{U}|\Psi_{in}\rangle = \prod_{i=1}^{N} \frac{1}{\sqrt{n_i!}} \left(\sum_{j=1}^{n_i} U_{ij}a_j^\dagger\right)^{n_i} |0\rangle. \tag{2}$$

Transformation of Eq. (2) is a high-dimensional irreducible representation of the $N \times N$ matrix of the optical transformation $U$ [8]. In Fock representation, matrix elements of $\langle n|\hat{U}|n'\rangle$ are calculated as permanents of matrix $U$ [6].

Now we specify main properties of the Eq. (1) relevant to numerical implementation of the optimization algorithm. In the computational Fock basis $|\rho\rangle$, the Eq. (1) is described by matrix $\hat{A}_{n_1, n_2}$, which is simply a submatrix of $\langle n|\hat{U}|n\rangle$. Thus $\hat{A}$ has a form of a set of polynomial functions in variables $u_{ij}$, computed using Eq. (2), so that Eq. (2) specifies explicit algebraic form of dependence of $\hat{A}$ on $\hat{U}$. If the total number of measured photons in ancilla modes $\sum_{i=M+1}^{N} n_i$ is the same as the number of input ancilla photons $\sum_{i=1}^{M} n_i$, then Eq. (1) leaves the number of computational photons invariant. Since the dual-rail computational basis is just a subset of all possible states in the computational modes, the transformation matrix $\hat{A}_{n_1, n_2}$ is in general a non-square matrix, mapping the Hilbert space of the computation basis to a larger Hilbert space. For example, $\hat{A}_{01, 02}$ for the CZ is a $10 \times 4$ matrix.

We now introduce the notion of operational fidelity of a transformation, which in general differs from the common measure of fidelity for a state transformation. From a physical point of view, the transformation $\hat{A}$ satisfies a 100% fidelity criteria if it is proportional to the target transformation operation $\hat{A}^T$, i.e., $\hat{A} \equiv \alpha \hat{A}^T$, where $\alpha$ is an arbitrary complex number. Since the target transformation is supposed to be a unitary gate, i.e., $\hat{A}^T \hat{A} = I$. The operational fidelity condition also requires that desired transformation $\hat{A}$ satisfies operational unitarity condition $\hat{A}^T \hat{A} = S I$, where $S = |\alpha|^2$ is the success probability of the transformation [9]. To formulate an algebraic estimate for the accuracy of the transformation, we consider complex rays $\beta \hat{A}$ and $\alpha \hat{A}^T$ as elements of complex projective space. The measure of closeness of elements in such a projective space is given by the Fubini-Study distance

$$\gamma(\hat{U}) = \arccos \sqrt{\frac{\langle \hat{A} | \hat{A}_{tar} \rangle \langle \hat{A}_{tar} | \hat{A} \rangle}{\langle \hat{A} | \hat{A} \rangle \langle \hat{A}_{tar} | \hat{A}_{tar} \rangle}}, \tag{3}$$

where the Hermitian inner product is $\langle A|B \rangle = Tr(AB^\dagger)/D_c$, and $D_c$ is the dimensionality of computational space.

In the numerical implementation of the optimization we used a nonsingular variable $F = \cos \gamma$, which we will refer to as fidelity in the rest of the paper.

Success probability $S$ of the transformation depends on the initial state $|\psi\rangle$. The upper bound of $S$ is determined by the operator norm $||\hat{A}||_{Max} = Max(|\langle \psi^\prime | \hat{A}^\dagger | \psi^\prime \rangle/\langle \psi | \psi \rangle|, and corresponding lower bound of $S$ is $||\hat{A}||_{Min} = Min(|\langle \psi | \hat{A}^\dagger | \psi^\prime \rangle/\langle \psi | \psi \rangle|$. As a measurement of the success probability, we use the Hilbert-Schmidt norm $||\hat{A}||_{HS} = \sqrt{Tr(\hat{A}^\dagger \hat{A})}/D_c$. It is easy to verify that $||\hat{A}||_{Min} \leq ||\hat{A}||_{HS} \leq ||\hat{A}||_{Max}$. As fidelity $F \to 1$, $||\hat{A}||_{Min} ||\hat{A}||_{Max} \to 1$ and $S$ becomes a well defined parameter equal to $||\hat{A}||_{HS}$. We will refer to the Hilbert-Schmidt norm $||\hat{A}||_{HS}$ as success probability, keeping in mind that such a definition may not correspond to a success probability of transformation of specific state initial state.

As an example, consider first the nonlinear sigm (NS) gate. The NS gate with 2 ancilla modes is as follows

$$\begin{align*}
\alpha|010\rangle + \beta|110\rangle + \gamma|210\rangle &\to \alpha|010\rangle + \beta|110\rangle - \gamma|210\rangle \quad (4)
\end{align*}$$

where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$. Because of conservation of the number of photons, the coefficient matrix $\hat{A}$ has a diagonal form (3×3) with entries given explicitly as functions of $U_{n,m}$.

$$\begin{align*}
|010\rangle : A_{11}(U) & = U_{22}, \quad \langle A_{11}(U) \rangle = \frac{3}{\Sigma_{i=1}^{3} |A_{ij}(U)|^2}, \tag{9}
|110\rangle : A_{22}(U) & = U_{12} U_{21} + U_{11} U_{22}, \quad \langle A_{22}(U) \rangle = \frac{3}{\Sigma_{i=1}^{3} |A_{ij}(U)|^2},
|210\rangle : A_{33}(U) & = U_{11} U_{21} + 2 U_{12} U_{22}. \quad \langle A_{33}(U) \rangle = \frac{3}{\Sigma_{i=1}^{3} |A_{ij}(U)|^2}.
\end{align*}$$

Then the operator success probability is

$$S(U) = \frac{1}{3} \sum_{i=1}^{3} |A_{ii}(U)|^2. \tag{8}$$

Since the target matrix $\hat{A}_{NS} = Diag(1, 1, -1)$, the fidelity of $\hat{A}$ is trivially calculated as

$$F(\hat{U}) = \frac{|A_{11}(\hat{U}) + A_{22}(\hat{U}) - A_{33}(\hat{U})|^2}{3 \Sigma_{i=1}^{3} |A_{ij}(\hat{U})|^2}, \tag{9}$$

Similarly, The CZ gate, constructed with four ancilla modes, can be represented as

$$\begin{align*}
\alpha|1010\rangle \otimes |1010\rangle + \beta|1001\rangle \otimes |1010\rangle &\to \alpha|1010\rangle \otimes |1010\rangle + \delta|0101\rangle \otimes |1010\rangle + \gamma|0101\rangle \otimes |1010\rangle - \delta|0101\rangle \otimes |1010\rangle. \tag{10}
\end{align*}$$

where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. In this case the matrix $\hat{A}$ is a $10 \times 4$ matrix, and the success probability has the form

$$S(\hat{U}) = \frac{1}{4} \sum_{i,j} |A_{ij}(\hat{U})|^2, \tag{11}$$

where $i=1,..,4$ and $j=1,..,10$. The corresponding fidelity function is

$$F(\hat{U}) = \frac{|A_{13}(\hat{U}) + A_{24}(\hat{U}) + A_{36}(\hat{U}) - A_{47}(\hat{U})|^2}{4 \Sigma_{i,j} |A_{ij}(\hat{U})|^2}. \tag{12}$$
where $A_{13}$, $A_{24}$, $A_{36}$ and $A_{47}$ are coefficients of output states $|1010\rangle$, $|1001\rangle$, $|0110\rangle$ and $|0101\rangle$, respectively.

If we write the unitary matrix as

$$U = U_0 \exp\left[\sum_{i=1}^{R} x_i g_i \right],$$

(13)

where $x$ is an $R$ dimensional vector which identifies the unitary matrix $\hat{U}$ in an R-D unitary matrix space with $g_i$ as basis $|i\rangle$. Each point of this space is a $N \times N$ unitary matrix. If we take Taylor expansion on the matrix exponential term, and truncate the polynomial with a proper error control, then we can represent the $\hat{U}$ as a polynomial function of $x$. Substituting it into the equations for success probability and fidelity, as derived in the last section, we have a polynomial function of $x$. Take $x$ as an individual or an abstract genotype in the language of genetics, we then have a formula ready for the genetic algorithm.

We decided to explore genetic algorithms for the following reasons. Firstly, the problem is to optimize a multidimensional nonlinear function (eight modes for CZ $|11\rangle$) with nonlinear constraints, i.e., a very large search space. In such cases there is no known traditional algorithm that has proven promising. (See Ref. [6] for other approaches.) Genetic Algorithms (GAs) are therefore an attractive candidate. Secondly, the GAs can handle any form of function to be optimized, which allows us to build a general scheme for our problem. In this paper, we focus on the NS and CZ gates, but the method can easily be adapted to any other LOQSG device. Thirdly, GAs are designed for searching the optimized results in a global space.

The first step in setting up a genetic algorithm is to choose a suitable fitness function. The simplest way is to select the fidelity given in Eqs. (9) and (12) as the fitness function. Once we get a maximum fidelity, we can substitute the corresponding $x$ into the function to obtain the success probability, Eqs. (8) and (11). This approach is simple to implement but we only search the fidelity without worrying about the success probability. Since we know the fidelity has to be close to one to have a reliable result, we can consider fidelity as a constraint and success probability as the fitness function. Many computer scientists are working on constrained genetic algorithms [12]. One of these is called the simulated annealing method [13]. Using this method, we need to reformulate the fitness function so that it contains the constraint. The new fitness function can be written as

$$\phi(x) = \alpha(F(x), T)S(x),$$

(14)

with

$$\alpha(F(x), T) = e^{-(1-F(x))/T},$$

(15)

where $F(x)$ is the fidelity function used as the constraint. The second parameter, referred to as the temperature $T$, is a function of the running time of the algorithm; $T$ tends to 0 (or very small values numerically) as execution proceeds. $S(x)$ is the success probability, and $\alpha(F(x), T)$ acts like a penalty so that the constraint can finally be satisfied. When the GA begins, we want the penalty to be small, i.e., $\alpha \approx 1$, so that the algorithm can search a bigger space to find the global maximum. When $T$ is large, which happens at the beginning of the execution, then $\alpha \approx 1$. As time goes on, $T \to 0$, then $\alpha \to 0$. It means the fitness tends to zero unless the constraint is satisfied, i.e., $F(x) \approx 1$. Therefore, at the end of the GA run, we can get the optimized success probability with a fidelity of one. The details of this simulated annealing genetic algorithm are described in Ref. [13].

The annealing genetic algorithm provides a way to search the global maximum of success probability of a LOQSG system and guarantees that fidelity is equal to one at the same
The disadvantage of this approach is that the efficiency delicately depends on the choice of the temperature annealing rate \( T \). In the following section, we investigate this problem using the NS gate as an example.

We use EOlib \cite{14} as the genetic algorithms framework. It provides a basic genetic operation template. In the case of the NS gate, we compared efficiency for different approaches. The comparison is shown in Fig. 2. The vertical axis denotes the success probability and the horizontal axis represents the events corresponding to each implementation. Each point indicates a complete run starting from a randomly selected population (a set of \( x \)). We rearrange the points in ascending order so that we can compare the efficiency easily. In Fig. 2, the black-triangle line indicates the result of a GA without any constraint, which takes fidelity as a fitness function, then calculates the success probability directly. The pink-square line describes the GA with a static-penalty when considering success probability as fitness function and fidelity as constraint, which is equivalent to setting the temperature in Eq. (13) to be a small constant (e.g., \( 10^{-5} \)). The third line with blue diamonds is the one which sets the annealing rate as \( T(t) = -\tan^{-1}(t) + \pi/2 \). The fourth green line sets the annealing rate as \( T(t) = 1/\sqrt{t} \), where \( t \) is time. From these plots, we can see that there are more chances for genetic algorithms with an annealing penalty to reach the global maximum of 1/4.

The most impressive result is that when choosing the annealing rate as \( 1/\sqrt{t} \), the simulation can get the global maximum for each run. This is also consistent with the practical estimation in Ref. [13].

Using this approach, we now design a CZ gate with high success probability. The results are shown in Fig. 3. They show a strong support for Knill’s highest known success probability (2/27) \cite{15} being the actual maximum success probability. Due to the complexity of the CZ gate, it is not known if an analytical proof for determining the maximum success probability is possible. As discussed, once we get the \( x \), we can construct the corresponding unitary matrix using Eq. (13). We can then discover the optimized design for the quantum circuit.

In this work, we introduced simulated annealing genetic algorithms to look for an optimized linear optical quantum state generator. By investigating the NS gate, which has been theoretically proven to have a maximum success probability of 1/4, we found that we can reach the global maximum for each run if we carefully choose the annealing rate. We have found what we believe to be an upper bound for Success Probability for the CZ gate, our results showing a plateau around the best known published results of 2/27 \( \approx 0.074 \). Our best result so far having been 0.0737. Since this approach is focused on searching the unitary matrix space to optimize the quantum circuits, it can be generalized to other devices, for instance, to quantum sensors.

We would like to acknowledge support from the Army Research Office and the Disruptive Technology Office.

* Electronic address: huver@phys.lsu.edu

* Electronic address: huver@phys.lsu.edu

---

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature 409, 46 (2001).
[2] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. 73, 58 (1994).
[3] P. Kok, W. Munro, K. Nemoto, T. Ralph, J. P. Dowling, and G. Milburn, Rev. Mod. Phys. 79, 135 (2007).
[4] S. Scheel and N. Lütkenhaus, New J. Phys. 6, 46 (2004).
[5] J. Eisert, Phys. Rev. Lett. 95, 040502 (2005).
[6] N. M. VanMeter, P. Lougovski, D. B. Uskov, K. Kieling, J. Eisert, and J. P. Dowling, quant-ph/0612154.
[7] K. Kraus, Lecture Notes: States, Effects and Operations (Springer, New York, 1983).
[8] A. Perelomov, Generalized Coherent States and Their Applications (Springer, Berlin, 1986).
[9] G. G. Lapaire, P. Kok, J. P. Dowling, and J. E. Sipe, Phys. Rev. A 68, 042314 (2003).
[10] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2000).
[11] For an eight modes CZ gate, we need an \( 8 \times 8 \) unitary matrix to execute the transformation. To generate such a complete matrix space, we need \( 8 \times 8 - 1 = 63 \) dimensions in general, which is the dimension of \( x \).
[12] C. A. Coello, Computer Methods in Applied Mechanics and Engineering 191, 1245 (2002).
[13] S. E. Carlson and R. Shonkwiler (1996), URL http://ftp.math.gatech.edu/~shenk/constrainedGA.pdf
[14] M. Keijzer, J. J. Merelo, G. Romero, and M. Schoenauer, Evolving Objects: A General Purpose Evolutionary Computation Library (Springer Verlag, 2002), pp. 229–241.
[15] E. Knill, Phys. Rev. A 66, 052306 (2002).