Optimizing Linear Optics Quantum Gates

J. Eisert
Blackett Laboratory, Imperial College London, London SW7 2BW, United Kingdom
Institut für Physik, Universität Potsdam, 14469 Potsdam, Germany

(Received 11 October 2004; published 22 July 2005)

In this Letter, the problem of finding optimal success probabilities of linear optics quantum gates is linked to the theory of convex optimization. It is shown that by exploiting this link, upper bounds for the success probability of networks realizing single-mode gates can be derived, which hold in generality for postselected networks of arbitrary size, any number of auxiliary modes, and arbitrary photon numbers. As a corollary, the previously formulated conjecture is proven that the optimal success probability of a nonlinear sign shift without feedforward is $1/4$, a gate playing the central role in the scheme of Knill-Laflamme-Milburn for quantum computation. The concept of Lagrange duality is shown to be applicable to provide rigorous proofs for such bounds, although the original problem is a difficult nonconvex problem in infinitely many objective variables. The versatility of this approach is demonstrated.

DOI: 10.1103/PhysRevLett.95.040502 PACS numbers: 03.67.Lx, 02.10.Yn, 42.50.−p

Optical implementations of quantum information processing devices offer many advantages over implementations employing other physical systems. Photons are relatively prone to decoherence, and precise state control is possible with the help of linear optical elements. Moreover, although the required nonlinearities to do universal quantum computation are presently not available at the single-photon level, they can be effectively realized by means of measurements. This comes at the price of the scheme becoming probabilistic. It was one of the key insights in the field, proposed by Knill, Laflamme, and Milburn (KLM) [1], that quantum computation can be achieved in a near-deterministic way using only single-photon sources, linear optical elements, and photon counters [1–3]. For this to be possible, a significant overhead in resources is required [1,4]. At the basis of the construction of the original scheme, yet, is a gate that is implemented with some probability of success, the *nonlinear sign shift* (NSS) gate [1,2,5]. The best known success probability of this gate using static linear optics followed by postselection is one quarter; this can then be uplifted to close to unity using teleportation steps.

One of the central questions seems to be, therefore, how well can the elementary gates be performed with static linear optics networks? In particular, what are the upper bounds for success probabilities of energy-preserving gates of single modes? This seems a key question for two reasons: on the one hand, the success probability at the level of elementary gates is a quantity that determines the necessary and notably large overhead to achieve near-deterministic scalable quantum computation [1,6]. On the other hand, for small-scale applications such as quantum repeaters for the long-range distribution of entanglement, high fidelity of the quantum gates may often be the demanding requirement of salient interest. The achievable rates in entanglement distillation, say, may be of secondary importance compared to the very functioning of the scheme. In such contexts, one could be well advised to abandon some of the feedforward using quantum memories or delay lines, but rather postselect the outcomes.

The best known scheme to realize the NSS gate with postselected linear optics succeeds with a probability of a quarter. Later Knill showed that the success probability can at most reach one half [2]. This was an important step: it was not clear, yet, whether this bound was indeed tight. Aiming at tightening this bound, Scheel and Lütkenhaus made a further significant step, emphasizing that a network realizing a quantum gate can be thought of as one which is linked once to the input mode by a single beam splitter [7], based on a result by Reck and co-workers [8–10]. It was conjectured, based on a numerical analysis in a restricted setting, that the maximal success probability of this gate could be one quarter.

It is the aim of this Letter to link the question of success probabilities to the theory of convex optimization [11,12]. It turns out that convex optimization provides powerful analytical methods to prove the validity of bounds to optimal success probabilities, without having to resort to restrictions of generality. By doing that, we arrive at rigorous tight upper bounds for quantum gates in the framework of linear optics quantum computation with no feedforward on the level of elementary gates. In particular, it is proven that the NSS gate can, in fact, be optimally realized with a success probability of exactly $1/4$. *Nonlinear phase gates* and *equivalents in higher Fock layers* are also considered. These methods will turn out to provide helpful tools, although the original problem has infinite dimension and is nonconvex. The central difficulty here in the problem is that one cannot bound the size of the auxiliary network *a priori:* It may well be that large networks go in hand with a significant advantage [13].

Let us start by stating the considered setting: we aim at formulating a general recipe to find upper bounds for success probabilities of gates of single modes preserving the energy using (i) photon sources, (ii) photon counters followed by postselection, and (iii) static linear optical networks of any size, using an arbitrary number of auxiliary modes and photons and an arbitrary number of network elements, but without feedforward on the level of individual gates (in which case the unit probability as tight upper bound is already known from the KLM scheme [1]). We consider gates of the form

$$\begin{aligned} |\psi_{\text{in}}\rangle = y_0|0\rangle + y_1|1\rangle + \cdots + y_N|N\rangle &\mapsto U|\psi\rangle \\ &= y_0|0\rangle + y_1 e^{i\phi_1}|1\rangle + \cdots + y_N e^{i\phi_N}|N\rangle, \qquad (1) \end{aligned}$$

where $|n\rangle$ denote the state vectors of number states and $\phi_1, \ldots, \phi_N \in \mathbb{R}$. To set the phase $\phi_0 = 0$ merely corresponds to a change of the global phase and does not restrict generality. This includes the NSS gate, acting as

$$|\psi_{\text{in}}\rangle = y_0|0\rangle + y_1|1\rangle + y_2|2\rangle \mapsto y_0|0\rangle + y_1|1\rangle - y_2|2\rangle.$$

In a static linear optical realization of the quantum gates, the gate can be realized only with a nonunity success probability. Any network consisting of linear optical elements can be decomposed into three steps, as has been pointed out in Ref. [7] based on Ref. [8]: (i) A preparation of a distinguished auxiliary mode 2 and all (unboundedly many) other auxiliary modes jointly labeled 3 in some initial pure state. (ii) A unitary operation of the input on 1 and 2, reflecting an application of a central beam splitter with transmittivity $t \in [-1, 1]$ (a convenient convention) and phase $\varphi \in [0, 2\pi)$. (iii) A measurement of all modes labeled 2 and 3, associated with a state vector $|\eta\rangle$. As a consequence, any optimal static linear optical network of a single input mode is reflected as a map $p_{\max} U \rho_{\text{in}} U^\dagger = \langle\eta|(V_{1,2} \otimes \mathbb{1}_3)(\rho_{\text{in}} \otimes |\omega\rangle\langle\omega|)(V_{1,2}^\dagger \otimes \mathbb{1}_3)|\eta\rangle$, for all input states $\rho_{\text{in}} = |\psi_{\text{in}}\rangle\langle\psi_{\text{in}}|$ of the input mode labeled with 1, $V_{1,2}$ is the unitary of the central beam splitter characterized by a real transmittivity t and phase φ. Writing $|\omega\rangle = \sum_{k=0}^{n} x_{k+1}|k\rangle \otimes |\omega_k\rangle$ with real numbers $x_1, \ldots, x_{n+1}$, we require that $\sum_{k=0}^{n} x_{k+1} f_k^{(j)} \varepsilon_{k+1} = p^{1/2} e^{i\phi_j}$ for all $j = 0, \ldots, N$, with $\varepsilon_{k+1} = \langle\eta|k\rangle|\omega_k\rangle$ and $f_k^{(j)} = \langle j|\langle k|V_{1,2}|j\rangle|k\rangle = e^{i\varphi j} g_k^{(j)}$, where the real $g_k^{(j)}$, introduced for the convenience of notation, depend on $t \in [-1, 1]$.

The problem now is essentially to find the optimal transmittivity $t \in [-1, 1]$, phase $\varphi \in [0, 2\pi)$, state vectors $|\eta\rangle$ and $|\omega_0\rangle, \ldots, |\omega_n\rangle$ for the optimal $n \in \mathbb{N}$, and the optimal $x_1, \ldots, x_{n+1}$ in order to bound the optimal success probability. This is as such a very involved problem: The number n cannot be bounded from above, meaning we cannot *a priori* bound the required resources in the network. This makes it formally an infinite-dimensional problem. The function we consider is not convex, so we are expected to encounter infinitely many local maxima. So even numerically, without truncating the problem cannot be solved as such. In order to circumvene these difficulties, two central ideas are employed: We treat part of the objective variables as parameters in the problem, such that the remaining problem can be relaxed to a convex quadratic program. In this way we can exploit methods from convex optimization. For the resulting problem we make use of the ideas of Lagrange duality [12] and outer approximations, and are able to explicitly construct a family of solutions to the dual. Let us first clearly state the strategy.

(I) We consider the problem for each $t \in [-1, 1]$, $\varphi \in [0, 2\pi)$, each $n \in \mathbb{N}$, and all legitimate $\varepsilon_1, \ldots, \varepsilon_{n+1}$ as defined above. This choice is denoted as $(t, \varphi, n, \varepsilon_1, \ldots, \varepsilon_{n+1})$. (II) We formulate the remaining problem as a quadratic optimization program, which can be relaxed to a semidefinite program [12] in $x_1, \ldots, x_{n+1}$. (III) Then, we are in the position to establish the dual problem. (IV) A family of explicit constructions of solutions of the dual is presented. (V) These solutions can be simplified such that the dependence on the specific choice of $\varepsilon_1, \ldots, \varepsilon_{n+1}$ and φ and t can be eliminated. This is done by exploiting two key ideas: on the one hand, by using both families of solutions of the dual problem, dependent on $\varepsilon_1, \ldots, \varepsilon_{n+1}$ and φ, and, on the other hand, by appropriate convex outer approximations. These powerful methods allow us to identify rigorous general upper bounds for all numbers of auxiliary modes, even though the original problem is unbounded in size. In a sense, we approach the optimal solution "from the other side."

(I) The first simplification is that we may choose any $\varepsilon_1, \ldots, \varepsilon_{n+1}$ for some n satisfying $\sum_{k=1}^{n+1}(\alpha_k^2 + \beta_k^2) = 1$, denoting with $\alpha_k, \beta_k \in \mathbb{R}$ the real and imaginary parts of ε_k, $\varepsilon_k = \alpha_k + i\beta_k$. We introduce $e^{i(j\varphi - \phi_j)} = \xi^{(j)} + i\zeta^{(j)}$ with $\xi^{(j)}, \zeta^{(j)} \in \mathbb{R}$. Success of the gate requires that $\sum_{k=1}^{n+1} x_k(\alpha_k \xi^{(j)} - \beta_k \zeta^{(j)}) g_{k-1}^{(j)} = \sum_{k=1}^{n+1} x_k(\alpha_k \xi^{(l)} - \beta_k \zeta^{(l)}) g_{k-1}^{(l)}$, $\sum_{k=1}^{n+1} x_k(\beta_k \xi^{(j)} + \alpha_k \zeta^{(j)}) g_{k-1}^{(j)} = 0$, for $j, l = 0, \ldots, n$. This is already a major simplification: instead of maximizing the actual trace of the state, we set the imaginary part to zero and avoid a very involved additional quadratic constraint at this point, without losing generality. The square of the quantities of the first line in the previous equation is then the success probability.

(II) We optimize for all $(t, \varphi, n, \varepsilon_1, \ldots, \varepsilon_{n+1})$ over all weights $x_1, \ldots, x_{n+1}$ satisfying $\sum_{k=1}^{n+1} x_k^2 = x^T x = 1$. This freedom corresponds to the weights in the preparation of the initial state of the auxiliary modes. The fact that we cannot restrict the size of the linear optics network is here reflected by the fact that we optimize over all possible preparations, even over all n. In this form, however, we see that the problem is manageable. The constraint $x^T x = 1$ can be relaxed to the convex one $x^T x \leq 1$, so to

$$\begin{bmatrix} 1 & x^T \\ x & \mathbb{1}_{n+1,n+1} \end{bmatrix} \geq 0.$$

So, in general, the problem of assessing a bound for the optimal success probability can be reduced to the following *maximization problem* in the vector $x = (x_1, \ldots, x_{n+1})$. This maximization problem (but not the full problem) is found to be manifestly of the form of a so-called semidefinite optimization problem [12]. After a number of elementary steps, the maximization problem in this vector can be cast into the following form of a maximization problem in the real symmetric matrix $Z \in \mathbb{R}^{(n+3)\times(n+3)}$:

$$\text{maximize } -\text{tr}[F_0 Z],$$

$$\begin{aligned} \text{subject to } &\text{tr}[e_{a,a}Z] = 1, \quad a = 1, \dots, n+3, \\ &\text{tr}[(e_{a,b} + e_{b,a})Z] = 0, \quad a, b = 3, \dots, n+3, \quad a \neq b, \\ &\text{tr}[e_{1,a}Z] = \text{tr}[e_{a,1}Z] = 0, \quad a = 2, \dots, n+3, \\ &\text{tr}[F_j Z] = 0, \quad j = 1, \dots, 2N+2, \quad Z \geq 0. \end{aligned}$$

The square of the solution is an upper bound for the success probability. Here, $F_0 = \text{diag}(1, 0, \dots, 0)$. The matrices

$$F_j = 0_{1,1} \oplus \begin{bmatrix} 0 & (c^{(j)} - c^{(0)})^T \\ c^{(j)} - c^{(0)} & 0_{n+1,n+1} \end{bmatrix}, \quad j = 1, \dots, N,$$

$$F_{j+N+1} = 0_{1,1} \oplus \begin{bmatrix} 0 & (d^{(j)})^T \\ d^{(j)} & 0_{n+1,n+1} \end{bmatrix}, \quad j = 0, \dots, N,$$

correspond to the ones that ensure the proper realization of the gate on the level of the real part and the complex part, respectively. The matrix

$$F_{2N+2} = \mathbb{1}_{1,1} \oplus \begin{bmatrix} 0 & -(c^{(0)})^T/2 \\ -c^{(0)}/2 & 0_{n+1,n+1} \end{bmatrix}$$

finally links the constraints in the primal problem. Here, the abbreviations $c^{(j)} = ((\alpha_0 \xi^{(j)} - \beta_0 \zeta^{(j)}) g_0^{(j)}, \dots, (\alpha_n \xi^{(j)} - \beta_n \zeta^{(j)}) g_n^{(j)})$ and $d^{(j)} = ((\beta_0 \xi^{(j)} + \alpha_0 \zeta^{(j)}) g_0^{(j)}, \dots, (\beta_n \xi^{(j)} + \alpha_n \zeta^{(j)}) g_n^{(j)})$ are used for $j = 0, \dots, N$. The matrix $e_{a,b} \in \mathbb{R}^{(n+3)\times(n+3)}$ denotes the matrix all entries of which are zero, except an entry 1 at (a, b). The latter matrix F_{2N+2} can be replaced by

$$G = \mathbb{1}_{1,1} \oplus \begin{bmatrix} 0 & -\gamma c^{(0)}/2 \\ -\gamma (c^{(0)})^T/2 & 0_{n+1,n+1} \end{bmatrix}$$

with $\gamma \in [1, \infty)$ to be fixed later, such that $p_{\max}$ corresponds to the square of the optimal objective value for $\gamma = 1$, and is smaller for $\gamma > 1$. This seemingly irrelevant modification will turn out to be a helpful idea later on, to eliminate the dependence on the phase φ.

(III) We can now formulate the dual problem to this optimization problem delivering the bounds, as a solution can explicitly be constructed [14]. It can be shown that the dual problem can be written as follows, which is now a *minimization problem* in the objective vectors $z \in \mathbb{R}^{n+2}$, $v \in \mathbb{R}^{2N}$, and the matrix $V \in \mathbb{R}^{(n+3)\times(n+3)}$,

$$\text{minimize } q^T z,$$

$$\text{subject to } F_0 + \text{diag}(0, z_1, \dots, z_{n+2}) + \sum_{a=1}^{2N+1} v_a F_a + V + v_{2N+2} G \geq 0,$$

where $q = (1, \dots, 1)$, and matrix V has to be of the form $V = 0_{2,2} \oplus W$, with $W \in \mathbb{R}^{(n+1)\times(n+1)}$ being a real symmetric matrix satisfying $W_{a,a} = 0$ for all $a = 1, \dots, n+1$. In general, every solution of a dual problem to a semidefinite problem gives a bound to the optimal solution to the primal problem, as is not difficult to see [15]. Once we are able to construct a solution z of the dual for all values of $(t, \varphi, n, \varepsilon_1, \dots, \varepsilon_{n+1})$, we arrive at a rigorous upper bound for the primal problem. As such, $p_{\max} \leq (q^T z)^2/\gamma^2$ gives an upper general bound of the desired success probability.

(IV) We now explicitly construct a family of solutions, dependent on a single number $\delta \in \mathbb{R}$. The presented solutions may look like unlikely objects, yet they deliver the desired bounds. In order to construct a family of solutions, one has to find appropriate values for a matrix of arbitrary dimensions. This structure of the problem we encounter here is not only specific for the optimization problem at hand, but expected to be a generic feature in problems related to linear optics: roughly speaking, the intertwined quadratic problems originate from the auxiliary systems, whereas the polynomial constraints of high order are from the distinguished passive optical element.

In the construction, to start with, we choose $v_{2N+2} = 1$. For convenience let $w \in \mathbb{R}^{n+1}$ be defined as $w = (-\gamma/2 - \sum_{j=1}^N v_j) c^{(0)} + \sum_{j=1}^N v_j c^{(j)} + \sum_{j=0}^N v_{N+j+1} d^{(j)}$. We are free to choose $v_j = -\cos(j\varphi) s_j$ and $v_{N+j+1} = \sin(j\varphi) s_j$, $j = 1, \dots, N$, with functions s_j: $[-1, 1] \rightarrow \mathbb{R}^+$ yet to be specified. This freedom will later give rise to the outer approximation. Then, let us set $\gamma = 2\sum_{j=1}^N s_j [1 - \cos(\varphi j)] + 1$. This means that $\gamma \geq 1$, which is used to eliminate the unwanted dependence of φ. That is, $w_k/\alpha_k = (-1/2 + \sum_{j=1}^N s_j) g_k^{(0)} - \sum_{j=1}^N \cos(\phi_j) s_j g_k^{(j)}$. The matrix $W \in \mathbb{R}^{(n+1)\times(n+1)}$ is taken to be of the form $W_{a,b} = w_a w_b$ if $b \neq a$ and $W_{a,b} = 0$ if $b = a$. This construction yields a positive V. Finally, we choose $z_a = \alpha_{a-1}^2 \delta$ for $a = 2, \dots, n+2$ and $z_1 = \delta$. With this choice, indeed, $F_0 + \text{diag}(0, z_1, \dots, z_{n+2}) + \sum_{a=1}^{2N} v_a F_a + V + v_{2N+2} G \geq 0$ holds, so it is, in fact, a solution of the dual [15]. This choice, indeed, turns out to give the bounds.

(V) If we can now find functions $s_1, \dots, s_N$: $[-1, 1] \rightarrow \mathbb{R}^+$ such that there is a $\delta \in [0, 1]$ with $|w_k/\alpha_k| \leq \delta$ for all $k = 0, \dots, n$, we can, in fact, eliminate the dependence on $\alpha_1, \dots, \alpha_{n+1}$ and t, as we have then an outer approximation of the feasible set. The outer approximation defined by $|w_k/\alpha_k| \leq \delta$ takes care of the polynomial constraints in $t \in [0, 1]$ of arbitrary order.

We have then, indeed, established an upper bound: The above constructed solution yields $p_{\max} \leq (q^T z)^2/\gamma^2 \leq (q^T z)^2 \leq (\delta + \sum_{i=0}^{n+1} \alpha_i^2 \delta)^2 \leq 4\delta^2$, so $p_{\max} \leq 4\delta^2$ is a rigorous upper bound for the success probability. So finding an upper bound for the success probability amounts to finding solutions, possibly dependent on $t \in [-1, 1]$, for $s_1, \dots, s_N$ such that $|w_k/\alpha_k| \leq \delta$ is satisfied. This provides a general method that can be applied to all of the above considered gates. It is important to note that, although we had the freedom to construct this particular solution without caring whether this solution is unique or even optimal, this implies a rigorous bound for the primal problem and for the optimal success probability. This gives rise to a recipe for finding bounds for success probabilities for all the above quantum gates using linear optics.

The example of the NSS is, on the one hand, instructive to exemplify the general strategy, and, on the other hand, already the practically most important case. Here we have that $N = 2$ and $\phi_0 = 1$, $\phi_1 = 1$, and $\phi_2 = \pi$. For this case of $N = 2$, one finds $g_k^{(0)} = t^k$, $g_k^{(1)} = t^{k-1}(t^2 - k(1 - t^2))$, and $g_k^{(2)} = t^{k-2}(t^4 - 2kt^2(1 - t^2) + (1 - t^2)^2 k(k - 1)/2)$ using standard expressions for the unitaries of beam splitters in the number state basis. We now show that for each $t \in [-1, 1]$ we can find $s_1, s_2\colon [-1, 1] \to \mathbb{R}^+$ such that $|w_k/\alpha_k| \leq \delta$ is satisfied. More specifically, for all $t \in [-1, 1]$ we find s_1, s_2 such that $-1/4 \leq (-1/2 + s_1 + s_2)g_k^{(0)} - s_1 g_k^{(1)} + s_2 g_k^{(2)} \leq 1/4$ for all $k = 0, \ldots, \infty$, so we have that $\delta = 1/4$. Such a choice is given by

$$(s_1, s_2) = \frac{1}{4}\begin{cases} (1/(1-t), 0), & \text{if } t \in [-1, 1-\sqrt{2}), \\ (0, 1/(1+t^2)), & \text{if } t \in [1-\sqrt{2}, 0), \\ (1, 1/2), & \text{if } t \in [0, 1), \end{cases} \tag{2}$$

for all $k = 0, \ldots, \infty$. This can be shown with elementary methods, on the basis of only the functions in Eq. (2) such that $|w_k/\alpha_k| \leq \delta$ holds for $\delta = 1/4$ for all k. This finally demonstrates that the optimal success probability of a linear optical implementation of the NSS gate without feedforward is indeed 1/4: there are known schemes that fulfill this bound. This settles the question of the optimal success probability of this key quantum gate in this setting. This statement is interestingly completely independent of the network size, as long as it includes at least two auxiliary modes. The surprising result is that more resources do not help at all, and the smallest known functioning scheme can already be proven to be optimal. This unexpected outcome may also be taken as a further motivation to investigate cluster or graph state based approached, or to slightly leave the setting of linear optics [16].

The presented method can immediately be applied to assess optimal success probabilities of other quantum gates within the paradigm of linear optics [17]. The key point is that this method allows one to argue without having to restrict the amount of resources or the size of the specific network realizing a scheme. Statements on the distinguishability using auxiliary systems [18] are also accessible. As such, these ideas are hoped to be useful to contribute to finding linear optical schemes that make use of the minimal resources and to bringing linear optics quantum computation closer to feasibility.

Discussions with C. Emery, J. D. Franson, M. Knill, P. Kok, P. van Loock, N. Lütkenhaus, W. J. Munro, K. Nemoto, M. A. Nielsen, and S. Scheel are warmly acknowledged. This work was supported by the DFG (SPP 1078), the EU (IST-2001-38877), and the EURYI.

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature (London) **409**, 46 (2001).

[2] E. Knill, Phys. Rev. A **68**, 064303 (2003).

[3] T. B. Pittman, B. C. Jacobs, and J. D. Franson, Phys. Rev. Lett. **88**, 257902 (2002); J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, Nature (London) **426**, 264 (2003); A. B. U'Ren, K. Banaszek, and I. A. Walmsley, Quantum Inf. Comput. **3**, 480 (2003).

[4] N. Yoran and B. Reznik, Phys. Rev. Lett. **91**, 037903 (2003); M. A. Nielsen, *ibid.* **93**, 040503 (2004); D. E. Browne and T. Rudolph, *ibid.* **95**, 010501 (2005); S. D. Barrett and P. Kok, Phys. Rev. A **71**, 060310(R) (2005).

[5] K. Kojima, H. F. Hofmann, S. Takeuchi, and K. Sasaki, Phys. Rev. A **70**, 013810 (2004).

[6] T. C. Ralph, Phys. Rev. A **70**, 012312 (2004).

[7] S. Scheel and N. Lütkenhaus, New J. Phys. **6**, 51 (2004).

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

[9] U. Leonhardt and A. Neumaier, J. Opt. B **6**, L1 (2004).

[10] D. W. Berry, S. Scheel, B. C. Sanders, and P. L. Knight, Phys. Rev. A **69**, 031806(R) (2004); S. Scheel, K. Nemoto, W. J. Munro, and P. L. Knight, *ibid.* **68**, 032310 (2003).

[11] K. Audenaert, J. Eisert, E. Jane, M. B. Plenio, S. Virmani, and B. De Moor, Phys. Rev. Lett. **87**, 217902 (2001); F. Verstraete and H. Verschelde, *ibid.* **90**, 097901 (2003); E. M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001); A. Ambainis, H. Buhrman, Y. Dodis, and H. Roehrig, quant-ph/0304112; M. Jezek, J. Rehacek, and J. Fiurasek, Phys. Rev. A **65**, 060301(R) (2002); A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, *ibid.* **69**, 022308 (2004); J. Eisert, P. Hyllus, O. Gühne, and M. Curty, *ibid.* **70**, 062317 (2004).

[12] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).

[13] This puts straightforward approaches out of the question (e.g., first-order conditions of the Kuhn-Tucker type).

[14] The most general form of a semidefinite program is the maximization of $-\mathrm{tr}[ZF_0]$ subject to $\mathrm{tr}[ZF_i] = c_i$ for $i = 1, \ldots, n$ and $Z \geq 0$. In turn, its dual problem is the minimization of $c^T x$ subject to $F_0 + \sum_{i=1}^n x_i F_i \geq 0$.

[15] This is the content of so-called weak duality. In the notation of Ref. [14], let the matrix $Z \geq 0$ be the optimal solution of the primal problem, and the vector x be any solution of the dual. Then, $-\mathrm{tr}[ZF_0] \leq \sum_{i=1}^n x_i \,\mathrm{tr}[ZF_i] = c^T x$, delivering an upper bound.

[16] K. Nemoto and W. J. Munro, Phys. Rev. Lett. **93**, 250502 (2004); J. D. Franson, B. C. Jacobs, and T. B. Pittman, Phys. Rev. A **70**, 062302 (2004); A. Gilchrist, G. J. Milburn, W. J. Munro, and K. Nemoto, quant-ph/0305167.

[17] In order to exemplify the versatility of the approach, let us finally investigate two further quantum gates: for the *nonlinear phase shift* gate, acting as $y_0|0\rangle + y_1|1\rangle + y_2|2\rangle \mapsto y_0|0\rangle + y_1|1\rangle + e^{i\phi_2} y_2|2\rangle$ with some phase $\phi_2 \in [0, 2\pi)$, the presented method delivers immediately $p_{\max} \leq [3 - \cos(\pi - \phi_2)]^2/16$, consistent with $p_{\max} = 1$ for $\phi_2 = 0$ and $p_{\max} = 1/4$ for $\phi_2 = \pi$: it hence depends on the phase how difficult it is to implement the gate. For the *three photon gate* $y_0|0\rangle + y_1|1\rangle + y_2|2\rangle + y_3|3\rangle \mapsto y_0|0\rangle + y_1|1\rangle + y_2|2\rangle - y_3|3\rangle$ we find that $p_{\max} \leq 1/9$, indicating that, for higher Fock layers, the optimal success probabilities become even smaller.

[18] P. van Loock and N. Lütkenhaus, Phys. Rev. A **69**, 012302 (2004).