Binary Projective Measurement via Linear Optics and Photon Counting

Masahiro Takeoka,[1,2] Masahide Sasaki,[1,2] and Norbert Lütkenhaus[3,4]

[1]*Quantum Information Technology Group, National Institute of Information and Communications Technology (NICT), 4-2-1 Nukui-kitamachi, Koganei, Tokyo 184-8795, Japan*

[2]*CREST, Japan Science and Technology Agency, 1-9-9 Yaesu, Chuoh-ku, Tokyo 103-0028, Japan*

[3]*Quantum Information Theory Group, Institute of Theoretical Physics, Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*

[4]*Institute for Quantum Computing, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada*

(Received 13 March 2006; published 26 July 2006)

We investigate the implementation of binary projective measurements with linear optics. This problem can be viewed as a single-shot discrimination of two orthogonal pure quantum states. We show that any two orthogonal states can be perfectly discriminated using only linear optics, photon counting, coherent ancillary states, and feedforward. The statement holds in the asymptotic limit of a large number of these physical resources.

DOI: 10.1103/PhysRevLett.97.040502 PACS numbers: 03.67.Hk, 03.65.Ta, 42.50.Dv

Projection measurements play an essential role in photonic quantum-information protocols. In these applications, generally, a projection onto superposition states or entangled states of optical fields is required. Physically, it is a highly nontrivial problem how to implement such a measurement.

One plausible approach is to use linear optics and classical feedforward associated with a partial measurement. For example, a universal quantum computation scheme for photonic-qubit states has been proposed, which utilizes only linear optics, photon counting, and highly entangled auxiliary states of n photons generated by probabilistic gate operations [1]. In principle, it works with unit success probability in the asymptotic limit of large n. It is, however, still a nontrivial question how to prepare entangled ancillae even for modest n.

In this Letter, we discuss the linear optics implementation of a measurement which affects a projection onto two orthogonal states $\{|\Psi\rangle, |\Phi\rangle\}$. This is equivalent to the problem of discriminating two orthogonal quantum signals $\{|\Psi\rangle, |\Phi\rangle\}$ unambiguously [2,3]. We show that, in the asymptotic limit of a large number of partial measurements, one can perfectly discriminate the two states with linear optics, photon counting, and feedforward, but without any nonclassical auxiliary states. Even in the worst case, the average error probability of discrimination approaches zero with the scaling factor of $N^{-1/3}$ where N is the number of the partial measurements. Note that the signal space is two dimensional, but $|\Psi\rangle$ and $|\Phi\rangle$ can be any physical states defined in a larger space, e.g., qubit states, continuous variable states, etc.

Before discussing a linear optics implementation, it is worth mentioning a result concerning the distinguishability of two orthogonal multipartite states via local operations and classical communication. Walgate *et al.* [4] showed that one can perfectly discriminate an arbitrary set of two orthogonal pure states via a series of local projective measurements. The measurement basis at each local site is chosen such that every possible remaining state after the measurement must be orthogonal to each other. This result means that if one shows a physical scheme that can exactly discriminate any two orthogonal single-mode states, its sequential application can achieve an exact discrimination of any two orthogonal multimode states. In the following, therefore, we concentrate on a discrimination of two single-mode states.

An arbitrary set of two orthogonal single-mode states is described by

$$|\Psi\rangle = \sum_{m=0}^{\infty} c_m |m\rangle_0, \qquad |\Phi\rangle = \sum_{m=0}^{\infty} d_m |m\rangle_0, \tag{1}$$

where $|m\rangle$ is an m-photon number state and $\langle\Psi|\Phi\rangle = \sum_{m=0}^{\infty} c_m^* d_m = 0$. Figure 1 is the schematic of the measurement apparatus. The states are equally split into N modes by $N-1$ asymmetric beam splitters [5],

$$\hat{B}_{N-1,0}(\theta_{N-1})\hat{B}_{N-2,0}(\theta_{N-2})\cdots\hat{B}_{1,0}(\theta_1)|0\rangle^{\otimes N-1}|\Psi\rangle_0 = e^{-\hat{a}_{N-1}^\dagger \hat{a}_0}\cdots e^{-\hat{a}_1^\dagger \hat{a}_0} e^{\hat{a}_0^\dagger \hat{a}_0 \ln(1/\sqrt{N})}|0\rangle^{\otimes N-1}|\Psi\rangle_0 \equiv \hat{N}_{\mathrm{BS}}|\Psi\rangle_0, \tag{2}$$

where $\hat{B}_{i,0}(\theta_i) = \exp[\theta_i(\hat{a}_i^\dagger \hat{a}_0 - \hat{a}_i \hat{a}_0^\dagger)]$ and $\tan\theta_i = 1/\sqrt{N-i}$. The input is symmetrically split to N modes with the effective power reflectance of $1/N$. Then, at each output port, one makes some measurement by using linear optics and photon counters, where the information about the measurement outcome is fed forward to design the next measurement. It should be noted that this is a generalized version of the scheme so-called "Dolinar receiver" [6–8] which was originally proposed as a physical model attaining the minimum error discrimination of the binary coherent signals $\{|\alpha\rangle, |-\alpha\rangle\}$.

We briefly sketch how two states are discriminated by such a scheme in the limit of $N \to \infty$ and then provide a rigorous proof. Suppose one inserts $|\Psi\rangle$ or $|\Phi\rangle$ into the first beam splitter. For sufficiently small $1/N$, the reflectance of multiphotons can be neglected. The states after

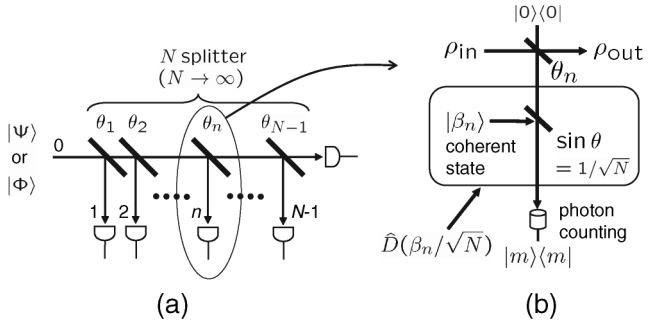

FIG. 1. (a) N splitter and (b) a measurement apparatus at each step. A displacement operation $\hat{D}(\beta_i/\sqrt{N})$ is realized by combining the signal with a coherent state local oscillator $|\beta_i/\sqrt{N}\sin\theta\rangle$ via a beam splitter with sufficiently small power reflectance of $\sin^2\theta$.

beam splitting are approximated to be $\hat{B}_{1,0}(\theta_1)|0\rangle_1|\Psi\rangle_0 \approx |0\rangle_1|\eta_0\rangle_0 + N^{-1/2}|1\rangle_1|\eta_1\rangle_0$ and $\hat{B}_{1,0}(\theta_1)|0\rangle_1|\Phi\rangle_0 \approx |0\rangle_1|\nu_0\rangle_0 + N^{-1/2}|1\rangle_1|\nu_1\rangle_0$, where $\langle\eta_0|\nu_0\rangle + \langle\eta_1|\nu_1\rangle/N \approx 0$, since a beam splitting operation is unitary. Then mode 1 is measured. The measurement here is required to maintain the orthogonality of any conditional outputs of $|\Psi\rangle$ and $|\Phi\rangle$. The local measurement satisfying this condition is described by a two-dimensional projective measurement,

$$|\pi_0\rangle = \mathcal{N}_{p0}\left\{|0\rangle + \frac{1}{X^*}(1 - \sqrt{1 + |X|^2})|1\rangle\right\} = \mathcal{N}_{p0}\{|0\rangle - [X/2 + O(X^3)]|1\rangle\}, \quad (3)$$

$$|\pi_1\rangle = \mathcal{N}_{p1}\{[X^*/2 + O(X^3)]|0\rangle + |1\rangle\}, \quad (4)$$

where $\mathcal{N}_{p0}$ and $\mathcal{N}_{p1}$ are the normalization factors and

$$X = \frac{2(\langle\nu_0|\eta_1\rangle\langle\eta_1|\nu_1\rangle - \langle\eta_0|\nu_1\rangle\langle\nu_1|\eta_1\rangle)}{\sqrt{N}(|\langle\eta_0|\nu_1\rangle|^2 - |\langle\eta_1|\nu_0\rangle|^2)}. \quad (5)$$

Here, we have assumed $|\langle\eta_0|\nu_1\rangle|^2 - |\langle\eta_1|\nu_0\rangle|^2 \neq 0$, which implies $X \propto 1/\sqrt{N}$, and thus we can take $X \ll 1$ in the limit of large N. The other case, i.e., $|\langle\eta_0|\nu_1\rangle|^2 - |\langle\eta_1|\nu_0\rangle|^2 = 0$, will be discussed later. Under this assumption, the projective measurement of Eqs. (3) and (4) can be implemented by the displacement operation $\hat{D}(\beta_1/\sqrt{N})$ and photon counting as shown in Fig. 1(b). Since both the signal and displacement are sufficiently weak, the corresponding measurement vectors are described by

$$\hat{D}^\dagger\left(\frac{\beta_1}{\sqrt{N}}\right)|0\rangle \approx e^{-|\beta_1|^2/2N}\left(|0\rangle - \frac{\beta_1}{\sqrt{N}}|1\rangle\right), \quad (6)$$

$$\hat{D}^\dagger\left(\frac{\beta_1}{\sqrt{N}}\right)|1\rangle \approx e^{-|\beta_1|^2/2N}\left(\frac{\beta_1^*}{\sqrt{N}}|0\rangle + |1\rangle\right), \quad (7)$$

which can be same as Eqs. (3) and (4) by choosing appropriate β_1.

The conditional states after the first measurement can be rewritten again as $|\Psi'\rangle = \sum_{m=0}^{\infty} c'_m|m\rangle$ and $|\Phi'\rangle = \sum_{m=0}^{\infty} d'_m|m\rangle$. Since $\hat{N}_{\rm BS}$ splits a state symmetrically, one can repeat the same procedure for the remaining state with the second beam splitter, the displacement operation $\hat{D}(\beta_2/\sqrt{N})$, where β_2 is conditioned on the previous measurement outcome, and a photon counter. After repeating the same procedure to modes 1 to $N-1$ with appropriate β_i's, the final states at mode 0 contain with dominating weight at most one photon and are still orthogonal to each other. As a consequence, applying the final (Nth) displacement and photon counting, one can exactly discriminate $|\Psi\rangle$ and $|\Phi\rangle$.

We now provide a detailed analysis on the above sketch for the case of finite resources (finite N). The effects of multiphoton reflections at each beam splitter are rigorously included, and an upper bound on the error rate is derived. We here assume that the average photon numbers of the input states $|\Psi\rangle$ and $|\Phi\rangle$ are finite and, moreover, that their photon number distributions decrease exponentially as $c_m \equiv \tilde{c}_m e^{-mx/2}$ with a real positive number x. The prior probabilities can be set to be equal without loss of generality. It is also assumed that the average powers of local oscillators always satisfy $|\beta_i|^2 \leq |C_{\beta_i}|^2 + O(1/N)$, where C_{β_i} is a complex constant independent of N.

After finishing a whole process of N measurement steps, one has a sequence of detected photon numbers. Because of the symmetry of the N beam splitting, the probability of detecting k photons at the ith measurement averaged over all possible patterns of the sequence is given by

$$P_k^{(i)} = |_i\langle k|\hat{D}_i(\beta_i/\sqrt{N})\hat{N}_{\rm BS}|\Psi\rangle_0|^2 \leq \frac{\langle\Psi_{\beta_i}|\hat{a}_0^{\dagger k}\hat{a}_0^k|\Psi_{\beta_i}\rangle}{N^k k!} + O\left(\frac{1}{N^{k+1}}\right) \leq C_k^{\max}/N^k + O(1/N^{k+1}), \quad (8)$$

where $|\Psi_{\beta_i}\rangle \equiv \hat{D}(C_{\beta_i})|\Psi\rangle$, whose photon number distribution decreases exponentially, and $C_k^{\max}$ is the maximum value of $\langle\Psi_{\beta_i}|\hat{a}_0^{\dagger k}\hat{a}_0^k|\Psi_{\beta_i}\rangle/k!$ for all i and possible inputs [9].

Let us now denote the patterns in which all the photon counters detect zero or one photon by "success" patterns and the others by "failure" patterns, since, in our scenario, the discrimination is hung up when the latter one is obtained. The probability of resulting the failure pattern $P_{\rm fail}$ is bounded by

$$P_{\rm fail} \leq [C_2^{\max}/N^2 + O(1/N^3)] \times N = C_2^{\max}/N + O(1/N^2), \quad (9)$$

which implies that $P_{\rm fail}$ approaches zero in the limit of large N, at least with the order of $1/N$.

Even if a successful pattern is obtained, one must accept a finite discrimination error depending on N, since the conditional states get slightly nonorthogonal after each measurement step. To see this, we revisit the first beam splitter $\hat{B}_{1,0}(\theta_1)$. Let us describe the states after beam splitting such that the orthogonal and nonorthogonal parts

are separated as

$$\hat{B}_{1,0}(\theta_1)|0\rangle|\Psi\rangle = |0\rangle|\eta_0\rangle + N^{-1/2}|1\rangle|\eta_1'\rangle + N^{-1}|2\rangle|\eta_2\rangle + \cdots = |0\rangle|\eta_0\rangle + N^{-1/2}|1\rangle|\eta_1\rangle + N^{-3/2}|1\rangle|\eta_r\rangle + \sum_{k=2}^{\infty} N^{-k/2}|k\rangle|\eta_k\rangle. \quad (10)$$

Note that one obtains the same expression for $|\Phi\rangle$ by replacing $|\eta_n\rangle$ with $|\nu_n\rangle$. The first two terms exactly satisfy the orthogonality $\langle\eta_0|\nu_0\rangle + \langle\eta_1|\nu_1\rangle/N = 0$ and the last terms represent the multiphoton reflection terms. Here, $|\eta_0\rangle = \sum_{m=0}^{\infty} c_m(1-1/N)^{m/2}|m\rangle$, $N^{-1/2}|\eta_1'\rangle = \sum_{m=1}^{\infty} c_m(m/N)^{1/2}(1-1/N)^{(m-1)/2}|m-1\rangle$, $N^{-1/2}|\eta_1\rangle = \sum_{m=1}^{\infty} c_m[1-(1-1/N)^m]^{1/2}|m-1\rangle$, and $N^{-3/2}|\eta_r\rangle = N^{-1/2}(|\eta_1'\rangle - |\eta_1\rangle)$ ($|\nu_n\rangle$'s are also obtained by replacing c_m with d_m). The terms $|\eta_r\rangle$, $|\nu_r\rangle$ and that for multiphoton reflections, which have been neglected in the previous discussion, cause the residual nonorthogonality. Note that the leading terms of all vectors $|\eta_k\rangle$'s and $|\nu_k\rangle$'s are independent of N. Denote the ith measurement operation as

$$\frac{{}_i\langle k|\hat{D}_i(\beta_i/\sqrt{N})\hat{B}_{i,0}(\theta_i)|0\rangle_i|\Psi\rangle}{|{}_i\langle k|\hat{D}_i(\beta_i/\sqrt{N})\hat{B}_{i,0}(\theta_i)|0\rangle_i|\Psi\rangle|} \equiv \hat{E}_k^{(i)}|\Psi\rangle. \quad (11)$$

Then the conditional outputs after detecting zero and one photons at the first measurement are given by

$$\hat{E}_0^{(1)}|\Psi\rangle = \mathcal{N}_0\left\{|\eta_0\rangle - \frac{\beta_1^*}{N}|\eta_1\rangle + \frac{1}{N^2}|\eta_{R_0}^{(1)}\rangle\right\}, \quad (12)$$

$$\hat{E}_1^{(1)}|\Psi\rangle = \mathcal{N}_1\left\{\beta_1|\eta_0\rangle + |\eta_1\rangle + \frac{1}{N}|\eta_{R_1}^{(1)}\rangle\right\}, \quad (13)$$

respectively, where $\mathcal{N}_0$ and $\mathcal{N}_1$ are the normalization factors and the third terms come from the higher order terms (higher than $1/N^{1/2}$) in ${}_i\langle k|\hat{D}_i(\beta_i/\sqrt{N})$ and $\hat{B}_{i,0}(\theta_i)|0\rangle_i|\Psi\rangle$. The same outputs are obtained for $|\Phi\rangle$ by replacing $|\eta_n\rangle$ with $|\nu_n\rangle$. The first two terms in Eqs. (12) and (13) can be exactly orthogonal to those of $|\Phi\rangle$ by choosing $\beta_1/\sqrt{N} = (1-\sqrt{1+|X|^2})/X^*$, where X is obtained by substituting $|\eta_0\rangle$, $|\eta_1\rangle$, $|\nu_0\rangle$, and $|\nu_1\rangle$ into Eq. (5). Since $X \propto 1/\sqrt{N}$, this choice of β_1 always satisfies the constraint on the average power of the local oscillator, $|\beta_1|^2 \le |C_{\beta_1}|^2 + O(1/N)$. However, we have to care of the fact that, in both events, the total conditional states $\hat{E}_k^{(1)}|\Psi\rangle$ and $\hat{E}_k^{(1)}|\Phi\rangle$ ($k = 0, 1$) are no longer orthogonal due to their third terms.

Now, suppose that the same strategy is applied to the choice of β_2 for the second measurement step. After the second measurement, the states are mapped into the new one with orthogonal and nonorthogonal terms, where the latter has two parts, i.e., contributions from the first and second measurements. Note that the leading order of prefactors of $|\eta_{R_k}^{(1)}\rangle$ with respect to $1/N$ does not change during the measurement process, as also the leading factors of $|\Psi\rangle$ do not change in the mapping in Eqs. (12) and (13). Eventually, after repeating $N-1$ measurement steps in a similar way, if every photon counter detected only zero or one photons, one obtains the conditional output consisting of the orthogonal term and $N-1$ nonorthogonal terms stemmed from each measurement as

$$|\Psi^{(N-1)}\rangle = \hat{E}^{(N-1)}\cdots\hat{E}^{(1)}|\Psi\rangle = |\eta^{(N-1)}\rangle + \frac{1}{N^2}\sum_{x=1}^{I^{(N-1)}}|H_0^{(i_x)}\rangle + \frac{1}{N}\sum_{y=1}^{J^{(N-1)}}|H_1^{(j_y)}\rangle, \quad (14)$$

where the first term is exactly orthogonal to that of $|\Phi^{(N-1)}\rangle$, while $|H_k^{(l)}\rangle$ is the residual nonorthogonal term coming from $|\eta_{R_k}^{(l)}\rangle$. $I^{(N-1)}$ and $J^{(N-1)}$ are the numbers of the events of detecting zero and one photon, respectively, and thus $I^{(N-1)} + J^{(N-1)} = N-1$.

Let us denote the final Nth measurement by $|D_k\rangle \equiv \hat{D}^\dagger(\beta_N/\sqrt{N})|k\rangle$ ($k = 0, 1$). Suppose that β_N is designed such that $|D_0\rangle$ and $|D_1\rangle$ are the same as the orthogonal terms in $|\Psi^{(N-1)}\rangle$ and $|\Phi^{(N-1)}\rangle$, respectively, up to the order of $1/N^{1/2}$ (the higher order terms contribute to the detection error). Then the error probability $P_{\mathrm{err}}^{D_1} = |\langle D_1|\Psi^{(N-1)}\rangle|^2$ is given by

$$P_{\mathrm{err}}^{D_1} = \left|\sum_{x=1}^{I^{(N)}}\frac{\langle D_1|H_0^{(i_x)}\rangle}{N^2} + \sum_{y=1}^{J^{(N)}}\frac{\langle D_1|H_1^{(j_y)}\rangle}{N}\right|^2, \quad (15)$$

where $I^{(N)} + J^{(N)} = N$. The leading order of $\langle D_1|H_k^{(j)}\rangle$ is independent of N for every j and k.

One can estimate the order of $J^{(N)}$ by counting the total amount of photons put into the system, where photons are supplied by the input state and N displacement operations. This configuration is mathematically converted into a simpler one with only two inputs, $\hat{D}(\beta_0)|\Psi\rangle$ and the coherent state $|\beta_{\mathrm{aux}}\rangle$, by adding some linear optics as illustrated in Fig. 2. Here, with the relation $\hat{D}_A(\alpha)\hat{D}_B(\beta)\hat{B}_{AB}(\theta) = \hat{B}_{AB}(\theta)\hat{D}_A(\alpha\cos\theta - \beta\sin\theta)\hat{D}_B(\alpha\sin\theta + \beta\cos\theta)$, one finds $|\beta_0|^2 = |\sum_{i=1}^N \beta_i/N|^2$ and $|\beta_{\mathrm{aux}}|^2 = \sum_{i=1}^N|\beta_i|^2/N - |\beta_0|^2$, where these are bounded as $|\beta_0|^2 = C_0 + O(1/N)$ and $|\beta_{\mathrm{aux}}|^2 = C_{\mathrm{aux}} + O(1/N)$ due to the constraint on $|\beta_i|^2$'s. C_0 and C_{aux} are constants independent of N.

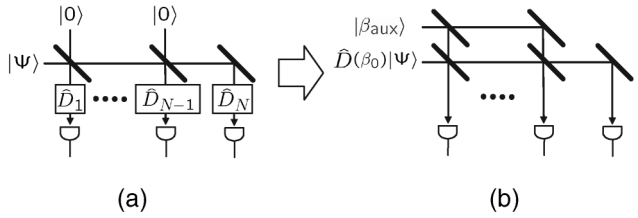

FIG. 2. The original scheme (a) can be transformed into (b) where the total input photon number is the sum of those of two input states.

The probability of having n photons in total decreases exponentially with respect to n as $P(n) = \sum_{m=0}^{n} P_{\mathrm{sig}}(n-m)P_{\mathrm{aux}}(m) = C_P e^{-nx} + O(1/N)$, where the last equality is derived from the fact that the photon number statistics of two inputs, $P_{\mathrm{sig}}(m)$ and $P_{\mathrm{aux}}(m)$, are exponential and Poissonian. Therefore, one can bound $J^{(N)}$ by a constant C_J with exponentially small exception as

$$\begin{aligned}\mathrm{Prob}[J^{(N)} &\le C_J + O(1/N) + N\epsilon] \\ &\ge 1 - C_P \exp\{-[C_J + O(1/N) + N\epsilon]\} \\ &\quad + O(1/N) \\ &= 1 - \tilde{C}_P e^{-N\epsilon} + O(1/N), \end{aligned} \tag{16}$$

where ϵ can be arbitrarily small for large N. Eventually, substituting it and $I^{(N)} \le N$ into Eq. (15), one obtains

$$\begin{aligned} P_{\mathrm{err}}^{D_1} &= \left| \frac{I^{(N)}}{N^2}\langle D_1|H_0\rangle_{\mathrm{av}} + \frac{J^{(N)}}{N}\langle D_1|H_1\rangle_{\mathrm{av}} \right|^2 \\ &\le C_E/N^2 + O(1/N^3) + \epsilon O(1/N) + \epsilon^2, \end{aligned} \tag{17}$$

where $\langle D_1|H_k\rangle_{\mathrm{av}} = \sum_i \langle D_1|H_k^{(i)}\rangle/L$ ($L = I^{(N)}$ and $J^{(N)}$ for $k = 0, 1$, respectively), and C_E is some constant independent of N. In a similar manner, the same bound is derived for $P_{\mathrm{err}}^{D_0} = |\langle D_0|\Phi^{(N-1)}\rangle|^2$. Then, summing over all detection patterns for both success and failure events, the average error probability is bounded as

$$\begin{aligned} P_{\mathrm{err}}^{\mathrm{tot}} &= \sum_s P(\#_s) P_{\mathrm{err}}^{\mathrm{succ}}(\#_s) + \sum_f P(\#_f)/2 \\ &\le \left(1 - \frac{C_2^{\max}}{N}\right)\frac{P_{\mathrm{err}}^{D_0} + P_{\mathrm{err}}^{D_1}}{2} + \frac{C_2^{\max}}{2N} + O\left(\frac{1}{N^2}\right) \\ &\le C/N + O(1/N^2) + O(1/N)\epsilon + \epsilon^2, \end{aligned} \tag{18}$$

where $\#_s$ and $\#_f$ represent the measurement sequence patterns for the success and failure events, respectively. $P(\#)$ is the probability to observe the pattern $\#$ where $P(\#_f)$ is bounded by Eq. (9). In the first sum, $P_{\mathrm{err}}^{\mathrm{succ}}(\#_s)$ is the average error probability for the success events, while $1/2$ in the second sum (the failure events) is the probability of randomly guessing the states. C is some constant. As a consequence, in the limit of $N \to \infty$, one can discriminate $|\Psi\rangle$ and $|\Phi\rangle$ with unit probability.

Finally, we discuss the case $|\langle\eta_0|\nu_1\rangle|^2 - |\langle\eta_1|\nu_0\rangle|^2 = 0$ in Eq. (5), in which the desirable local measurement cannot be implemented by a displacement and photon counting. Here, let us consider the projection measurement consisting of slightly perturbed vectors $|\Pi_0\rangle = \sqrt{1-\delta}|\Psi\rangle - \sqrt{\delta}|\Phi\rangle$ and $|\Pi_1\rangle = \sqrt{1-\delta}|\Phi\rangle + \sqrt{\delta}|\Psi\rangle$ with a perturbation parameter δ. One can design such a measurement by the previous strategy with the total error probability of $P_{\mathrm{err}}^{\mathrm{tot}} = C/N^{1-2\Delta} + O(1/N^{2-3\Delta}) + O(1/N^{1-3\Delta/2})\epsilon + O(N^{\Delta})\epsilon^2$, where $\Delta = -\log_N \delta$. This device can discriminate the original states $|\Psi\rangle$ and $|\Phi\rangle$ with the average error probability of

$$\begin{aligned} P_{\mathrm{err}}^{\mathrm{av}} &= 1 - (1 - P_{\mathrm{err}}^{\mathrm{tot}})(|\langle\Pi_0|\Psi\rangle|^2 + |\langle\Pi_1|\Phi\rangle|^2)/2 \\ &= C_1/N^{\Delta} + O(1/N^{2\Delta}) + C_2/N^{1-2\Delta} + O(1/N^{2-3\Delta}) \\ &\quad + O(1/N^{1-3\Delta/2})\epsilon + O(N^{\Delta})\epsilon^2. \end{aligned} \tag{19}$$

In the asymptotic limit of large N, this is minimized with $\Delta = 1/3$, and then we obtain $P_{\mathrm{err}}^{\mathrm{av}} = C/N^{1/3} + O(1/N^{2/3}) + O(1/N^{1/2})\epsilon + O(N^{1/3})\epsilon^2$, which still converges to zero.

In summary, we have proved that arbitrary two-dimensional projection measurement can be implemented by linear optics and feedfoward without using any nonclassical ancillary states. Since these linear optics tools are mostly available with current technology, the concrete linear optics circuit we showed can be directly applied for various quantum information protocols that require binary projection measurements. The remaining question is whether one can apply the same approach to the problem of more than three states discrimination.

We thank M. Ban, K. Tamaki, and P. van Loock for valuable discussions. M. T. also acknowledges the kind hospitality at the QIT group in Universität Erlangen-Nürnberg. This work was supported by the DFG under the Emmy-Noether program, the EU FET network RAMBOQ, and the network of competence QIP of the state of Bavaria.

Note added.—Details on derivation of some equations in this Letter can be found in [10].

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature (London) **409**, 46 (2001).
[2] P. van Loock and N. Lütkenhaus, Phys. Rev. A **69**, 012302 (2004).
[3] This is true only for the case when all physical operations during a whole measurement can be described by rank 1 operators. As will be shown in the text, our scheme corresponds to this case.
[4] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, Phys. Rev. Lett. **85**, 4972 (2000).
[5] P. van Loock and S. L. Braunstein, Phys. Rev. Lett. **84**, 3482 (2000).
[6] S. J. Dolinar, Research Laboratory for Electronics, MIT, Quarterly Progress Report No. 111, 1973 (unpublished), p. 115.
[7] J. M. Geremia, Phys. Rev. A **70**, 062303 (2004).
[8] M. Takeoka, M. Sasaki, P. van Loock, and N. Lütkenhaus, Phys. Rev. A **71**, 022318 (2005).
[9] One can show that $C_k^{\max}$ takes a finite value, by using the exponential decay of the number distribution of $|\Psi_\beta\rangle$ and the relation $\sum_{m=0}^{\infty} z^m(m+k)!/m! = k!/(1-z)^{k+1}$.
[10] M. Takeoka, M. Sasaki, and N. Lütkenhaus, quant-ph/0603074.