Experimental Demonstration of Adaptive Quantum State Estimation

Ryo Okamoto,[1,2] Minako Iefuji,[1,2] Satoshi Oyama,[1,2] Koichi Yamagata,[3] Hiroshi Imai,[4] Akio Fujiwara,[3,*] and Shigeki Takeuchi[1,2,†]

[1]*Research Institute for Electronic Science, Hokkaido University, Kita-ku, Sapporo 001-0020, Japan*
[2]*The Institute of Scientific and Industrial Research, Osaka University, Mihogaoka 8-1, Ibaraki, Osaka 567-0047, Japan*
[3]*Department of Mathematics, Osaka University, 1-1 Machikaneyama, Toyonaka, 560-0043 Osaka, Japan*
[4]*University of Pavia, Dipartimento di Fisica A. Volta, Via Bassi 6, 27100 Pavia, Italy*

(Received 30 January 2012; published 28 September 2012)

The first experimental demonstration of an adaptive quantum state estimation (AQSE) is reported. The strong consistency and asymptotic efficiency of AQSE have been mathematically proven [A. Fujiwara, J. Phys. A **39**, 12489 (2006)]. In this Letter, the angle of linear polarization of single photons, the phase parameter between the right and the left circularly polarization, is estimated using AQSE, and the strong consistency and asymptotic efficiency are experimentally verified. AQSE will provide a general useful method in both quantum information processing and metrology.

DOI: 10.1103/PhysRevLett.109.130404 PACS numbers: 03.65.Wj, 03.67.−a, 42.50.Dv, 42.50.Ex

Quantum theory is inherently statistical. This entails repetition of experiments over a number of identically prepared quantum objects, for example, quantum states, if one wants to know the *true state* or the *true value* of the parameter that specifies the quantum state [1–4]. Such an estimation procedure is particularly important for quantum communication and quantum computation [5] and is, also, indispensable to quantum metrology [6–10]. In applications, one needs to design the estimation procedure in such a way that the estimated value of the parameter should be close to the true value (consistency), and that the uncertainty of the estimated value should be as small as possible (efficiency) for a given limited number of samples. In order to realize these requirements, Nagaoka advocated an adaptive quantum state estimation (AQSE) procedure [11,12], and recently, Fujiwara proved the strong consistency and asymptotic efficiency for AQSE [13,14].

In this Letter, we report the first experimental demonstration of AQSE using photons. The angle of a half wave plate (HWP) that initializes the linear polarization of input photons is estimated using AQSE. A sequence of AQSE is carried out with 300 input photons and the sequence is repeated 500 times for four different settings of HWP. The statistical analysis of these results verifies the strong consistency and asymptotic efficiency of AQSE. Recently, it has been mathematically proven that the precision of AQSE outperforms the conventional state tomography [15]. It is thus expected that AQSE will provide a useful methodology in the broad area of quantum information processing, communication, and metrology.

Let us first explain AQSE, in detail. For simplicity, we restrict ourselves to a one-dimensional *quantum statistical model* $\mathcal{S} = \{\rho_\theta; \theta \in \Theta(\subset \mathbb{R})\}$, a smooth parametric family of density operators on a Hilbert space $\mathcal{H}$ having a one-dimensional parameter θ. Our aim is to estimate the true value of θ by means of a certain quantum estimation scheme. An *estimator* is represented by a pair $(M, \check{\theta})$, where $M = \{M(x); x \in \mathcal{X}\}$ is a positive operator-valued measure (POVM) that takes values on a set $\mathcal{X}$, and $\check{\theta}: \mathcal{X} \to \Theta$ is a map that gives the estimated value $\check{\theta}(x)$ from each observed data $x \in \mathcal{X}$. The observed data $x \in \mathcal{X}$ have a probability density

$$f(x; \theta, M) := \mathrm{Tr}\rho_\theta M(x), \tag{1}$$

which depends on both the parameter θ and the measurement M.

In traditional statistics, it is often the case to confine our attention to unbiased estimators. An estimator $(M, \check{\theta})$ is called *unbiased* if

$$E_\theta[M, \check{\theta}] = \theta \tag{2}$$

is satisfied for all $\theta \in \Theta$, where $E_\theta[\cdot]$ denotes the expectation with respect to the density, Eq. (1). It is well-known [16] that an unbiased estimator $(M, \check{\theta})$ satisfies the quantum Cramér–Rao inequality $V_\theta[M, \check{\theta}] \geq (J_\theta)^{-1}$, where $V_\theta[\cdot]$ denotes the variance, and J_θ is the quantum Fisher information of the model $\mathcal{S}$ defined by $J_\theta := \mathrm{Tr}\rho_\theta L_\theta^2$, where L_θ is the symmetric logarithmic derivative defined by the self-adjoint operator satisfying the equation $\frac{d\rho_\theta}{d\theta} = \frac{1}{2}(L_\theta\rho_\theta + \rho_\theta L_\theta)$.

In quantum statistics, however, it is regarded that unbiasedness is too restrictive a requirement, and we usually weaken the condition to a *local* one. An estimator $(M, \check{\theta})$ is called *locally unbiased* [17] at a given point $\theta_0 \in \Theta$ if the condition [Eq. (2)] is satisfied around $\theta = \theta_0$ up to the first order of the Taylor expansion, that is, if $E_{\theta_0}[M, \check{\theta}] = \theta_0$ and $\frac{d}{d\theta}E_\theta[M, \check{\theta}]|_{\theta=\theta_0} = 1$ hold. Clearly, an estimator is unbiased if and only if it is locally unbiased at all $\theta \in \Theta$. A crucial observation is that an estimator $(M, \check{\theta})$ that is locally unbiased at θ_0, also satisfies the quantum Cramér–Rao inequality

$$V_{\theta_0}[M, \check{\theta}] \geq (J_{\theta_0})^{-1} \tag{3}$$

at $\theta = \theta_0$, and that the lower bound in Eq. (3) is achievable for any one-dimensional quantum statistical model $\mathcal{S}$. To put it differently, the best locally unbiased estimator (LUE) for the parameter θ at $\theta = \theta_0$ is the one that satisfies $V_{\theta_0}[M, \check{\theta}] = (J_{\theta_0})^{-1}$.

Here, we encounter a difficulty which often becomes the target of criticism: since the best LUE for estimating the parameter θ depends, in general, on the unknown parameter θ itself, the estimation strategy based on LUEs would be infeasible. In a different yet analogous context, Cochran [18] ingeniously described this kind of dilemma as follows: ''You tell me the value of θ and I promise to design the best experiment for estimating θ.''

To surmount this difficulty, Nagaoka [11,12] advocated an AQSE scheme as follows. Suppose that, by prior investigation of the quantum statistical model $\mathcal{S}$, one has the list of optimal LUEs $(M(\cdot;\theta), \check{\theta}(\cdot;\theta))$ for each $\theta \in \Theta$. One begins with an arbitrary initial guess $\hat{\theta}_0 \in \Theta$, and applies the measurement $M(\cdot;\hat{\theta}_0)$ that is optimal at $\hat{\theta}_0$. Suppose the data x_1 are observed, one then applies the maximum likelihood method to the likelihood function $L_1(\theta) = f(x_1;\theta, M(\cdot;\hat{\theta}_0))$, to obtain the next guess $\hat{\theta}_1$. At stage $n(\geq 2)$, one applies the measurement $M(\cdot;\hat{\theta}_{n-1})$, where $\hat{\theta}_{n-1}$ is the maximum likelihood estimator (MLE) obtained at the previous stage. The likelihood function is then given by $L_n(\theta) := \prod_{i=1}^{n} f(x_i;\theta, M(\cdot;\hat{\theta}_{i-1}))$, where x_i is the observed data at stage i, and one obtains the nth MLE $\hat{\theta}_n$ that maximizes $L_n(\theta)$. It is quite natural to expect that the sequence $\hat{\theta}_n$ of MLEs would converge to the true value of the parameter θ. In fact, under certain regularity conditions, it can be shown that the sequence $\hat{\theta}_n$ is strongly consistent and asymptotically efficient [13,14].

Now, let us discuss the implementation of AQSE using photons (Fig. 1). Here, the unknown parameter is the angle θ of HWP0, which determines the phase ϕ between right and left circularly polarizations of input photons by the relation $\phi = 4\theta$. An arbitrary linear polarization can be described using right and left circular polarizations as follows:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|R\rangle + e^{i\phi}|L\rangle) = \cos\left(\frac{\phi}{2}\right)|H\rangle + \sin\left(\frac{\phi}{2}\right)|V\rangle. \tag{4}$$

By changing the angle of the half wave plate (HWP1), we can adjust the measurement basis. For such measurements, the POVM having optimal estimation capability is given by

$$M(\theta) = (M(1;\theta), \quad M(2;\theta)) = (|\xi\rangle\langle\xi|, \quad I - |\xi\rangle\langle\xi|), \tag{5}$$

where $\langle\xi| = (\cos(2\theta + \frac{\pi}{4}), \sin(2\theta + \frac{\pi}{4}))$. By applying the POVM $M(\theta)$ to the input state $|\psi(\theta)\rangle := |\psi\rangle$, one obtains the probability distribution on $\mathcal{X} := \{1, 2\}$, which is isomorphic to the fair coin flipping.

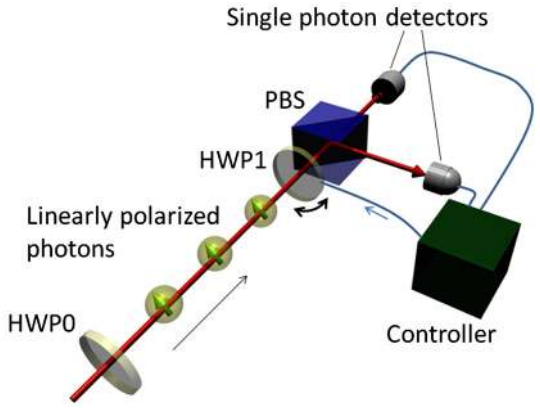

FIG. 1 (color online). Schematic of adaptive quantum state estimation. Photons are linearly polarized with a polarization direction determined by HWP0. The polarization is analyzed by HWP1 and the polarizing beam splitter (PBS). The controller sets HWP1 to an angle calculated on the basis of the photon measurement results.

The drawback to realizing this measurement is that the optimal POVM $M(\theta)$ depends on the unknown value of the parameter θ [19]. We can avoid this drawback by adopting an AQSE as follows. We begin by setting the initial log-likelihood function to be $l_0(\theta) = 0$, and then start inputting and detecting photons one by one. For nth photon, we apply the measurement $M(\hat{\theta}_{n-1})$ which depends on the latest MLE $\hat{\theta}_{n-1}$. Let $x_n \in \mathcal{X}$ be the outcome indicating which detector has been lit. The log-likelihood function is then updated by the formula

$$l_n(\theta) := l_{n-1}(\theta) + \log\langle\psi(\theta)|M(x_n;\hat{\theta}_{n-1})|\psi(\theta)\rangle, \tag{6}$$

and the nth MLE is given by $\hat{\theta}_n = \mathrm{argmax}_\theta l_n(\theta)$. Let us denote the true value of the parameter θ by θ^t. It is known that the sequence $\hat{\theta}_n$ of MLEs converges to the true value θ^t with probability one (strong consistency) and that the distributions of the random variables $\sqrt{n}(\hat{\theta}_n - \theta^t)$ converge to the normal distribution $N(0, J_{\theta^t}^{-1})$ (asymptotic efficiency), where J_θ denotes the quantum Fisher information of the parameter θ [13,14], which turns out to be 16 for our model [Eq. (4)].

The experimental setup is shown in Fig. 2(a). Single photons at 780 nm are generated from a heralded single photon source [20], consisting of a CW diode pump laser (wavelength: 402 nm) and a 3 mm long BBO crystal (Type I). A pair of a signal photon (780 nm) and a trigger photon (830 nm) is created via spontaneous parametric down-conversion. The detector (DT, SPCM-AQR, Perkin Elmer) after an interference filter (IF1, center wavelength 830 nm) outputs an electric pulse (width 30 ns) when it detects a trigger photon and the electric pulse heralds the generation of a signal photon, which is coupled to a polarization maintaining fiber (PMF) after an interference filter (IF2, center wavelength 780 nm, width 4 nm). The

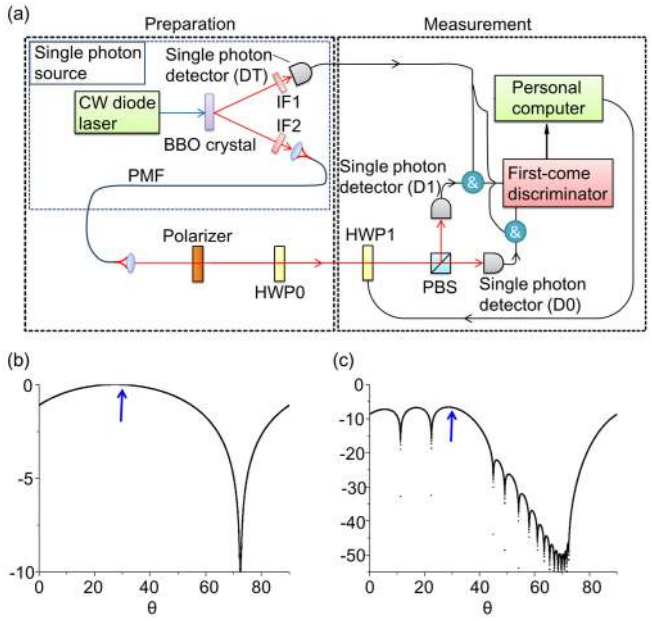

FIG. 2 (color online). (a) Schematic of the experimental setup. (b)(c) An example showing the update of a log-likelihood function. The second term $\log\langle\psi(\theta)|M(x_n;\hat{\theta}_{n-1})|\psi(\theta)\rangle$ in Eq. (6) is shown in panel (b), and the updated $l_n(\theta)$ is shown in panel (c). The (blue) arrows in (b) and (c) indicate the true value θ^t.

polarization of photons are then initialized to be horizontal using a polarizer (extinction ratio 10^{-5}). The target parameter θ^t was set using HWP0. The polarization state of the photon was analyzed by HWP1 and a polarizing beam splitter (PBS). After passing through the PBS, photons are guided to single photon detectors (D0 and D1, SPCM-AQR, Perkin Elmer) on each PBS output port. The outputs of single photon detectors are gated by the rise of the heralding signal and connected to the *first-come discriminator*, consisting of a homemade electric circuit. When the discriminator receives the first signal from one of the detectors (D0 or D1) after the measurement for $(n-1)$th photon starts, the discriminator informs which detector has been clicked. The minimum pulse interval of 2.5 ns can be discriminated. Note that the discriminator ignores the case when it receives the pulses from both the detectors within 2.5 ns. The angle of HWP1 for measuring the nth photon is determined by calculating the discretized MLE $\hat{\theta}_n$, the maximizer of the log-likelihood function, Eq. (6), chosen from among the 10 000 points that divide the domain $[0, \pi/2)$ of the parameter θ into equal parts [Figs. 2(b) and 2(c)]. When the change of HWP1 angle is completed, the measurement for the next (nth) photon will be started. In a sequence of AQSE, the above mentioned procedure is carried out up to 300 input photons ($n = 300$). For four different HWP0 angles $\theta = 0$, 30, 60, and 78.3 deg, we repeated the sequence for 500 times ($r = 500$).

Let us first observe the strong consistency for the sequence $\hat{\theta}_n$ of MLEs for the parameter θ of HWP0.

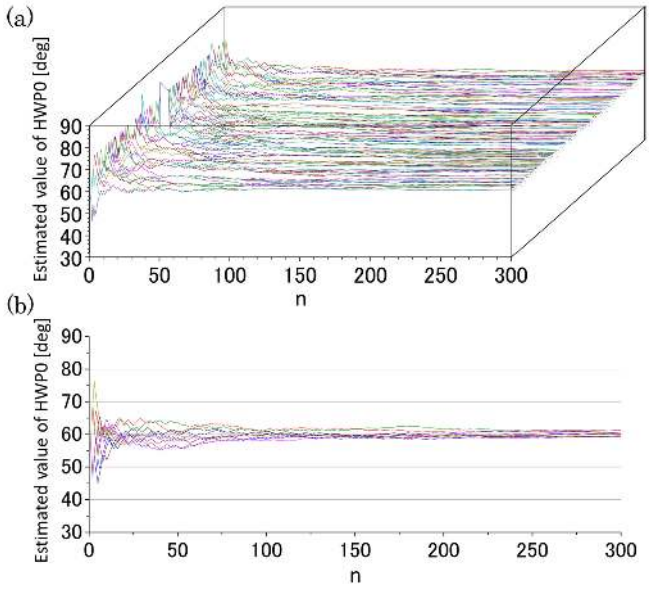

FIG. 3 (color online). (a) Trajectories of estimated HWP0 angles against the number n of photons for $r = 500$ repetitions is shown in a three-dimensional plot. (b) The first 10 curves are superposed in a two-dimensional graph.

Figure 3(a) shows 500 trajectories of estimated HWP0 angle $\hat{\theta}_n$ against the number n of photons when the true value θ^t of the parameter is set to be 60 degrees. The curves correspond to independent runs of adaptive estimation. Evidently, each curve of $\hat{\theta}_n$ approaches the true value θ^t, which is in accord with the mathematical result that $\hat{\theta}_n \to \theta^t$ almost surely as $n \to \infty$, even though the curves are dissimilar to each other, reflecting the genuine statistical nature of quantum system. The convergence to the true value is clear in Fig. 3(b) where first 10 trajectories in Fig. 3(a) are superposed.

We next test the hypothesis that the MLE $\hat{\theta}_n$ follows a normal distribution for large n. More concretely, we will investigate if the random variable $\sqrt{nJ_\theta}(\hat{\theta}_n - \bar{\theta})$ follows the standard normal distribution $N(0, 1)$, i.e., $\sqrt{nJ_\theta}(\hat{\theta}_n - \bar{\theta}) \sim N(0, 1)$, where $\bar{\theta}$ is the sample average of MLEs $\hat{\theta}_n$ over sufficiently many independent trials. A goodness of fit test [21] was carried out as follows: (1) The real axis was divided into 23 intervals (bins) $\{I_b\}_{b=0}^{22}$, where $I_1, \ldots, I_{21}$ are disjoint partitions of the interval $[-3.5, 3.5]$ of equal width, and $I_0 = (-\infty, -3.5)$, $I_{22} = (3.5, +\infty)$. In reality, these bins were slightly shifted by $\delta/10\,000$, where $\delta := \sqrt{nJ_\theta}\pi/20\,000$ is the scaled resolution of the estimator $\hat{\theta}_n$, so that the data $\sqrt{nJ_\theta}(\hat{\theta}_n - \bar{\theta})$ did not fall on the boundaries of the bins. (2) The test-statistic $X^2 := \sum_{b=0}^{22} \frac{(N_b - r p_b)^2}{r p_b}$ was calculated, where N_b is the number of observed data which fell into bth bin, p_b the theoretical probability of falling a datum into bth bin under the null hypothesis $N(0, 1)$, and r the number of repetitions of adaptive estimation procedure. (3) The test-statistic X^2

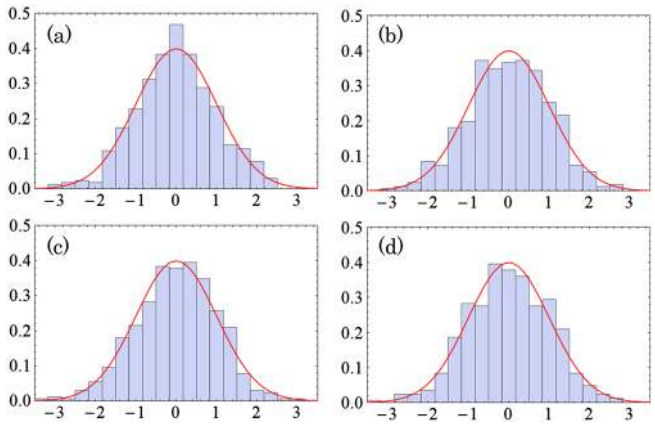

FIG. 4 (color online). Histogram of the observed data obtained by $r = 500$ independent experiments of adaptive estimation scheme, each using $n = 300$ photons. These histograms were taken for four different true values of (a) 0 deg, (b) 30 deg, (c) 60 deg, and (d) 78.3 deg.

was analyzed using the chi-square distribution $\chi^2_{23\text{-}p}$ of degree 23-p, where $p = 2$ degrees of freedom ought to be subtracted because of the normalization and the use of sample average $\bar{\theta}$.

Figure 4 shows the histogram of the observed data obtained by $r = 500$ independent experiments of adaptive estimation scheme, each using $n = 300$ photons. The true values θ^t of the parameter θ of HWP0 are set to be 0, 30, 60, and 78.3 degrees. The density function of the standard normal distribution $N(0, 1)$ is also plotted as the solid curve. All the experimental data agree with the standard normal distribution. To be precise, the values of the test statistic X^2 are (a) 16.8 (b) 15.7 (c) 12.8 (d) 16.2, and the null hypothesis is accepted with 10% significance level in each case.

Having obtained the strong evidence that the distribution of the MLE has converged quite well to a normal distribution at $n = 300$, we finally proceed to the estimation of confidence intervals [21] for the mean μ and variance v, assuming that $\sqrt{n}(\hat{\theta}_n - \mu) \sim N(0, v)$. The confidence intervals for μ and v are obtained by the standard procedure based on the statistical laws that $\sqrt{\frac{r}{\bar{V}}}(\bar{\theta} - \mu) \sim T_{r-1}$ and $\frac{r-1}{(v/n)}\bar{V} \sim \chi^2_{r-1}$. Here $\bar{V}$ is the unbiased variance of MLEs $\hat{\theta}_n$ over r trials, and T_{r-1} the t-distribution of degree $r - 1$.

Table I summarizes the results for $r = 500$ with 90% confidence level. Recall that the asymptotic efficiency asserts that $\mu \simeq \theta^t$ and $v \simeq J^{-1}_{\theta^t} (= 0.0625)$. Since the precision of the present experiment is about ± 0.2 degrees [22], we conclude that the estimated values of μ and v listed in Table I are in excellent agreement with the theoretical values.

It should be noted that the purpose of our AQSE is completely different from "adaptive measurements" proposed by Berry and Wiseman [23]. Their scheme was devised to estimate the phase difference between the two arms of an interferometer using a special N-photon two-mode state, approximating the canonical measurement proposed by Sanders and Milburn [24], and is not applicable to general quantum state estimation problems. By contrast, our AQSE is a general purpose estimation scheme applicable to any quantum statistical model using n identical copies of an unknown state. AQSE may also be used in verifying the achievability of the Cramér–Rao version of the Heisenberg limit $O(1/N^2)$ [25] by applying the scheme to the n-i.i.d. extension $\rho_\theta^{\otimes n}$ of an N-photon phase-shift model ρ_θ on $\mathcal{H} \simeq (\mathbb{C}^2)^{\otimes N}$. (See also Ref. [26] for estimating a unitary channel under noise.) Incidentally, AQSE is based on the Cramér–Rao type point estimation theory and is free from the choice of *a priori* distribution which matters in Bayesian statistics such as adaptive Bayesian quantum tomography [27].

TABLE I. Confidence intervals for the mean μ and the variance v. C.L. means confidence level.

θ^t (deg)	μ (deg) (90% C.L.)	v (90% C.L.)
0.0	-0.15 ± 0.06	[0.054, 0.067]
30.0	29.90 ± 0.06	[0.055, 0.067]
60.0	60.00 ± 0.06	[0.056, 0.068]
78.3	78.27 ± 0.06	[0.055, 0.068]

In summary, we have verified both the strong consistency and asymptotic efficiency of AQSE by experimentally estimating the angle of linear polarization of photons. Since AQSE has been mathematically proven to outperform the conventional estimation scheme such as the state tomography [15], we plan to apply AQSE to multiparameter cases and compare the performance with other protocols using a fixed measurement basis [28]. It will also be intriguing to apply AQSE to enhance the performance of quantum metrological experiments beating the standard quantum limit [6–9].

We would like to thank Professor Nagaoka for helpful discussion. This work was supported in part by Quantum Cybernetics project of JSPS, Grant-in-Aid from JSPS, JST-CREST project, FIRST Program of JSPS, Special Coordination Funds for Promoting Science and Technology, Research Foundation for Opto-Science and Technology, and the GCOE program.

*fujiwara@math.sci.osaka-u.ac.jp
†takeuchi@es.hokudai.ac.jp

[1] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
[2] M. G. A. Paris, Int. J. Quantum. Inform. **07**, 125 (2009).
[3] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Rev. Lett. **102**, 010404 (2009).
[4] A. Hentschel and B. C. Sanders , Phys. Rev. Lett. **107**, 233601 (2011).

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[6] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[7] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, Science **316**, 726 (2007).
[8] R. Okamoto, H. F. Hofmann, T. Nagata, J. L. O'Brien, K. Sasaki, and S. Takeuchi, New J. Phys. **10**, 073033 (2008).
[9] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Nature Photon. **5**, 43 (2010).
[10] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs and J. J. L. Morton, Science **324**, 1166 (2009).
[11] H. Nagaoka, in *Proceedings of the International Symposium on Information Theory, 1988* (unpublished).
[12] H. Nagaoka, in *Proceeding of the 12th Symposium on Information Theory and Its Application, 1989* (unpublished).H. Nagaoka, in *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2005), p. 125.
[13] A. Fujiwara, J. Phys. A **39**, 12489 (2006).
[14] A. Fujiwara, J. Phys. A **44**, 079501 (2011).
[15] K. Yamagata, Int. J. Quantum. Inform. **09**, 1167 (2011).
[16] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[17] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[18] W. G. Cochran, J. Am. Stat. Assoc. **68**, 771 (1973).
[19] Note that any fixed POVM of the form of Eq. (5) is optimal for almost all values of the parameter ϕ if we treat only the pure state model [Eq. (4)]. However, if we treat mixed state models, the dependence of optimal POVM to the parameter becomes crucial [8].
[20] C. K. Hong and L. Mandel, Phys. Rev. Lett. **56**, 58 (1986).
[21] R. V. Hogg, J. W. McKean, and A. T. Craig, *Introduction to Mathematical Statistics* (Pearson Prentice Hall,, Upper Saddle River, 2005).
[22] The precision of the rotation stage for HWP1 and the accuracy of the polarization basis states limited the total precision of the experimental setup to ± 0.2 degree.
[23] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[24] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).
[25] H. Imai and A. Fujiwara, J. Phys. A **40**, 4391 (2007).
[26] A. Fujiwara and H. Imai, J. Phys. A **41**, 255304 (2008).
[27] F. Huszár and N. M. T. Houlsby, Phys. Rev. A **85**, 052120 (2012).
[28] Y. I. Bogdanov, G. Brida, M. Genovese, S. P. Kulik, E. V. Moreva, and A. P. Shurupov, Phys. Rev. Lett. **105**, 010404 (2010).