# Optimal input states and feedback for interferometric phase estimation

D. W. Berry,[1] H. M. Wiseman,[2] and J. K. Breslin[1]
[1]*Department of Physics, The University of Queensland, St. Lucia, Queensland 4072, Australia*
[2]*School of Science, Griffith University, Nathan, Queensland 4111, Australia*
(Received 20 December 2000; published 12 April 2001)

We derive optimal $N$-photon two-mode input states for interferometric phase measurements. Under canonical measurements the phase variance scales as $N^{-2}$ for these states, as compared to $N^{-1}$ or $N^{-1/2}$ for states considered by previous authors. We prove that it is not possible to realize the canonical measurement by counting photons in the outputs of the interferometer, even if an adjustable auxiliary phase shift is allowed in the interferometer. However, we introduce a feedback algorithm based on Bayesian inference to control this auxiliary phase shift. This makes the measurement close to a canonical one, with a phase variance scaling slightly above $N^{-2}$. With no feedback, the best result (given that the phase to be measured is completely unknown) is a scaling of $N^{-1}$. For optimal input states having up to four photons, our feedback scheme is the best possible one, but for higher photon numbers more complicated schemes perform marginally better.

DOI: 10.1103/PhysRevA.63.053804 PACS number(s): 42.50.Dv, 03.67.Hk, 42.50.Lc

## I. INTRODUCTION

Interferometry is the basis of many high-precision measurements, for example gravitational wave detection. Phase measurement sensitivity is limited by the complementarity between photon number and phase. This is most easily investigated for a Mach-Zehnder interferometer (see Fig. 1). The outputs of the interferometer are measured to give an estimate $\phi$ of the phase difference between the arms, $\varphi$. For simplicity we can take the phase shift $\varphi$ to be in one arm. We can then introduce a known phase shift $\Phi$ in the other arm in order to obtain a more accurate measurement.

The Mach-Zehnder interferometer achieves a phase sensitivity of $\Delta\phi \propto 1/\sqrt{N}$ when $N$ photons are fed into one arm. Several authors [1–5] have proposed ways of reducing the phase uncertainty to the Heisenberg limit of $1/N$. Except for that of Caves [1], these proposals involved input states where the total number of photons input to the interferometer is fixed, and has no noise. We will be considering this case for simplicity.

Most of these proposals [1–3] require that the input phase be zero or very small in order to obtain the $1/N$ scaling. Sanders and co-workers [4,5] considered canonical measurements, for which the $1/N$ scaling is independent of the input phase. Unfortunately it is not generally possible to realize these canonical measurements with photodetections at the output of the interferometer (see below). Recently we showed [6] that it is possible to obtain measurements that are very close to canonical using feedback. In Ref. [6] we also derived input states that minimize the Holevo phase variance. In this paper we expand our presentation of these results, and derive a number of additional results.

In general there are three areas available for optimization: the initial state, the feedback phase, and the final phase estimate. In Sec. II we rederive the optimal input states discussed in Ref. [6], and in addition show that these states have uncertainties scaling as $1/N$ for all common measures of uncertainty. In Sec. III we explain in detail the feedback scheme outlined in Ref. [6]. We also derive the optimal final phase estimates, demonstrating that those used in Ref. [6] are optimal. Finally we show, in Sec. VII, that the feedback scheme outlined in Ref. [6] is optimal for optimal states of up to four photons, and we numerically determine the optimal feedback for optimal states up to 12 photons. Even with optimal feedback and phase estimates, the phase variance is above the intrinsic phase variance for three or more photons, demonstrating that it is not possible to perform canonical interferometric measurements by photodetections, despite the added flexibility of the auxiliary phase $\Phi$.

## II. INITIAL STATE

We consider a Mach-Zehnder interferometer with two input beams with annihilation operators $\hat{a}$ and $\hat{b}$ as in Fig. 1. We use the Schwinger representation [2]

$$\hat{J}_x = (\hat{a}^\dagger \hat{b} + \hat{a}\hat{b}^\dagger)/2, \quad \hat{J}_y = (\hat{a}^\dagger \hat{b} - \hat{a}\hat{b}^\dagger)/2i, \tag{2.1}$$

$$\hat{J}_z = (\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b})/2, \quad \hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 . \tag{2.2}$$

Following the notation of Sanders and co-workers [4,5], we use the notation $|j\mu\rangle_z$ for the common eigenstate of $\hat{J}_z$ and $\hat{J}^2$ with eigenvalues $\mu$ and $j(j+1)$, respectively. This

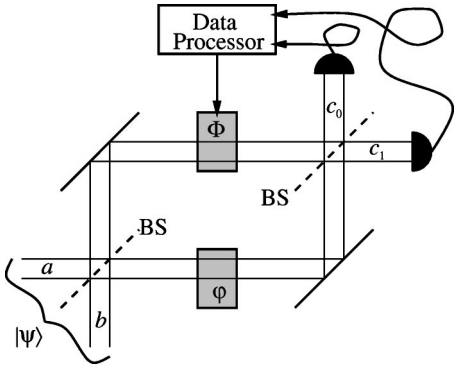


FIG. 1. The Mach-Zehnder interferometer, with the addition of a controllable phase $\Phi$ in one arm. The unknown phase to be estimated is $\varphi$. Both beam splitters (BS) are 50:50.

 

state corresponds to Fock states with $j+\mu$ and $j-\mu$ photons entering ports $a$ and $b$, respectively. Similarly we use the subscript $y$ for an eigenstate of $\hat{J}_y$. We can represent a general pure input state with $2j$ photons as $|\psi\rangle = \Sigma_{\mu=-j}^{j} \psi_\mu |j\mu\rangle_z$.

The Mach-Zehnder interferometer acts to transform the input field via the unitary operator $\hat{I}(\varphi)=\exp(-i\varphi\hat{J}_y)$ [5], and we wish to estimate the phase $\varphi$ by performing measurements on the transformed output state. The probability distribution for the phase estimate $\phi$ is, in general, given by $P(\phi)d\phi=\langle\psi|\hat{E}(\phi)|\psi\rangle d\phi$, where $\hat{E}(\phi)$ is the probability operator measure (POM), and $|\psi\rangle$ is the interferometer input state.

From Ref. [5], the optimal measurement strategy corresponds to projections onto the phase states $|j\phi\rangle$, given by

$$|j\phi\rangle=(2j+1)^{-1/2}\sum_{\mu=-j}^{j} e^{i\mu\phi}|j\mu\rangle_y. \tag{2.3}$$

The optimal POM is therefore the projection operator $\hat{E}(\phi)d\phi=\mathcal{N}|j\phi\rangle\langle j\phi|d\phi$, where $\mathcal{N}$ is a normalization factor. In terms of the $\hat{J}_y$ eigenstates the POM is given by

$$\hat{E}(\phi)d\phi=\frac{1}{2\pi}\sum_{\mu,\nu=-j}^{j} e^{i(\mu-\nu)\phi}|j\mu\rangle_y\langle j\nu|d\phi. \tag{2.4}$$

If we make the substitutions $\mu'=\mu+j$, $\nu'=\nu+j$, and $|\mu'\rangle=|j\mu\rangle_y$, the POM is

$$\hat{E}(\phi)d\phi=\frac{1}{2\pi}\sum_{\mu',\nu'=0}^{2j} e^{i(\mu'-\nu')\phi}|\mu'\rangle\langle\nu'|d\phi. \tag{2.5}$$

Written in this form the POM is identical to the POM for canonical measurements on a single mode, with an upper limit of $2j$ on the photon number.

In order to determine the optimal input state, we wish to minimize some measure of the phase uncertainty. The measure that we will be using is the Holevo phase variance, which is defined as [7]

$$V_\phi \equiv S_\phi^{-2}-1, \tag{2.6}$$

where $S_\phi\in[0,1]$ is the *sharpness* of the phase distribution, defined as

$$S_\phi\equiv|\langle e^{i\phi}\rangle|=\left|\int_{-\pi}^{\pi} P(\phi)e^{i\phi}d\phi\right|. \tag{2.7}$$

The Holevo variance is the natural quantifier for dispersion in a cyclic variable. If the variance is small, then it is easy to show that

$$V_\phi\simeq\int_{-\pi}^{\pi} 4\sin^2\left(\frac{\phi-\langle\phi\rangle}{2}\right)P(\phi)d\phi. \tag{2.8}$$

From this it is apparent that, provided there is no significant contribution to the variance from large $\phi$, this definition of the variance is equivalent to the usual definition. In addition, the Holevo phase variance approaches infinity in the limit of a flat distribution, which is not the case for the usual definition.

The optimal state in the single-mode case for this measure of the phase variance has been considered before [8,9], and there is a simple, analytic solution. The minimum Holevo phase variance is

$$V_\phi=\tan^2\left(\frac{\pi}{2j+2}\right)=\frac{\pi^2}{(2j+2)^2}+O(j^{-4}), \tag{2.9}$$

and the minimum uncertainty state for a mean phase of zero is

$$|\psi_{\rm opt}\rangle=\frac{1}{\sqrt{j+1}}\sum_{\mu'=0}^{2j}\sin\left[\frac{(\mu'+1)\pi}{2j+2}\right]|\mu'\rangle. \tag{2.10}$$

The effect of the interferometer (in the Schrödinger picture) is to change the input state to

$$\hat{I}(\varphi)|\psi_{\rm opt}\rangle=\frac{1}{\sqrt{j+1}}\sum_{\mu'=0}^{2j}\sin\left[\frac{(\mu'+1)\pi}{2j+2}\right]e^{-i\varphi\mu'}|\mu'\rangle, \tag{2.11}$$

which is the minimum uncertainty state for a mean phase of $-\varphi$. Because we have chosen the input phase to be zero, measuring the phase of the output state directly gives the interferometer phase shift $\varphi$.

To obtain the input state in terms of the eigenstates of $\hat{J}_z$, from Ref. [4], we have

$${}_y\langle j\mu|j\nu\rangle_z=e^{i(\pi/2)(\nu-\mu)}I^j_{\mu\nu}(\pi/2), \tag{2.12}$$

where $I^j_{\mu\nu}(\pi/2)$ are the interferometer matrix elements given by

$$I^j_{\mu\nu}(\pi/2)=2^{-\mu}\left[\frac{(j-\mu)!}{(j-\nu)!}\frac{(j+\mu)!}{(j+\nu)!}\right]^{1/2}P_{j-\mu}^{(\mu-\nu,\mu+\nu)}(0)$$

$$\text{for}\quad \mu-\nu>-1,\ \ \mu+\nu>-1, \tag{2.13}$$

where $P_n^{(\alpha,\beta)}(x)$ are the Jacobi polynomials, and the other matrix elements are obtained using the symmetry relations

$$I^j_{\mu\nu}(\phi)=(-1)^{\mu-\nu}I^j_{\nu\mu}(\phi)=I^j_{-\nu,-\mu}(\phi). \tag{2.14}$$

Using this, the state in terms of the eigenstates of $\hat{J}_z$ is

$$|\psi_{\rm opt}\rangle=\frac{1}{\sqrt{j+1}}\sum_{\mu,\nu=-j}^{j}\sin\left[\frac{(\mu+j+1)\pi}{2j+2}\right]$$
$$\times e^{i(\pi/2)(\mu-\nu)}I^j_{\mu\nu}(\pi/2)|j\nu\rangle_z. \tag{2.15}$$

An example of this state for 40 photons is plotted in Fig. 2. This state contains contributions from all the $\hat{J}_z$ eigenstates; however, the only significant contributions are from nine or ten states near $\mu=0$. The distribution near the center

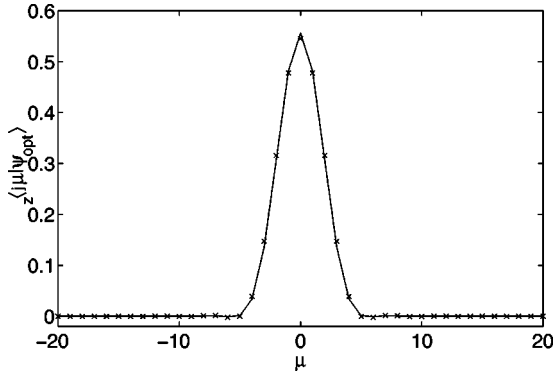


FIG. 2. The coefficients $_z\langle j\mu|\psi_{\rm opt}\rangle$ for states optimized for minimum phase variance under canonical measurements. All coefficients for a photon number of $2j=40$ are shown by the continuous line, and those near $\mu=0$ for a photon number of $2j=1200$ by crosses.

is fairly independent of the photon number. To demonstrate this, the distribution near the center for 1200 photons is also shown in Fig. 2.

In Ref. [2] it was shown that it is possible to generate a combination of two states near $\mu=0$. Since the optimal states described here have significant contributions from a small number of states near $\mu=0$, it seems likely that it should be possible to produce a reasonable approximation of them using a suitable modification of the apparatus described in Ref. [2].

In order to compare this state with $|j0\rangle_z$, where equal photon numbers are fed into both input ports (as considered in Refs. [3–5]), the exact phase variance for this case was calculated for a range of photon numbers. Since the phase is measured modulo $\pi$ for this state, rather than the Holevo phase variance the measure used for the variance was

$$V_\phi=(|\langle e^{2i\phi}\rangle|^{-2}-1)/4, \tag{2.16}$$

where the expectation value is determined using the POM above. The phase variances for this state and the optimal state are shown in Fig. 3. We also show the Holevo phase variance for the state $(|j0\rangle_z+|j1\rangle_z)/\sqrt{2}$ considered in Ref. [2]. This state is designed so that it can be considered modulo $2\pi$, and it has been claimed that this state has a phase uncertainty scaling as $N^{-1}$. The exact Holevo phase variance of the state where all the photons are incident on one port, $|jj\rangle_z$, is also shown for comparison.

As can be seen, the phase variance for $|j0\rangle_z$ scales down with the photon number much more slowly than the phase variance for optimal states, and even more slowly than the variance for $|jj\rangle_z$, which scales as $N^{-1}$. In fact, for the range of photon numbers considered, the phase variance for $|j0\rangle_z$ scales as $N^{-1/2}$. This seems to imply that the phase uncertainty scales as $N^{-1/4}$, a radically different result from the $N^{-1}$ scaling found in Refs. [3–5]. The state $(|j0\rangle_z+|j1\rangle_z)/\sqrt{2}$, considered in Ref. [2], is even worse than the state $|j0\rangle_z$. For this state, which is considered modulo $2\pi$, the phase variance does not scale down with the photon number, and is always of order unity.

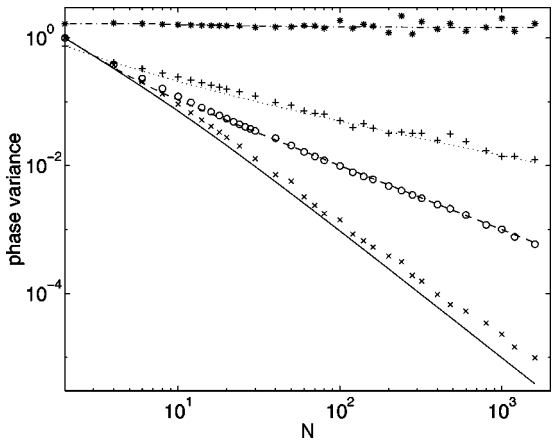


FIG. 3. Variances in the phase estimate vs the input photon number $2j$. The lines are exact results for canonical measurements on optimal states $|\psi_{\rm opt}\rangle$ (continuous line), on states with all photons incident on one input port $|jj\rangle_z$ (dashed line), on states with equal photon numbers incident on both input ports $|j0\rangle_z$ (dotted line), and on the state $(|j0\rangle_z+|j1\rangle_z)/\sqrt{2}$ (dash-dotted line). The crosses are numerical results for the adaptive phase measurement scheme on $|\psi_{\rm opt}\rangle$, the circles are those on $|jj\rangle_z$, the pluses are those on $|j0\rangle_z$, and the asterisks are those on a $(|j0\rangle_z+|j1\rangle_z)/\sqrt{2}$ input state. All variances for the $|j0\rangle_z$ state are for phase modulo $\pi$.

The reason for these discrepancies is that the results found in Refs. [3–5] are all based on the width of the central peak in the distribution, but the main contribution to the variance is from the *tails* of the distribution. To demonstrate this, the probability distribution for $|j0\rangle_z$ multiplied by $\sin^2\phi$, is plotted in Fig. 4. [Note that Eqs. (2.8) and (2.16) imply that $V_\phi\simeq\int\sin^2\phi P(\phi)d\phi$.] In this figure we can clearly see that the main contribution to the variance is from the tails. In practice this means that the error in the phase will be small most of the time, but there will be a significant number of results with large errors.

These results should not be taken to mean that the state $|j0\rangle_z$ is unusable, as this state has an uncertainty scaling as $N^{-1}$ for other measures of the phase uncertainty, for example the confidence interval and the Fisher length [5]. What these results mean is that care must be taken in analyzing phase data obtained using this state, as many simple data analysis techniques (for example the mean) will result in a large error.

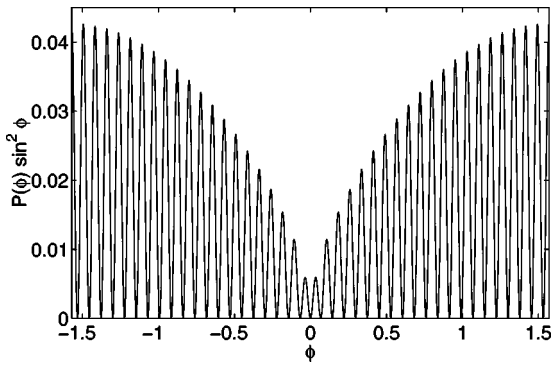


FIG. 4. The canonical phase probability distribution for $|j0\rangle_z$ multiplied by $\sin^2\phi$ for $2j=80$ photons.

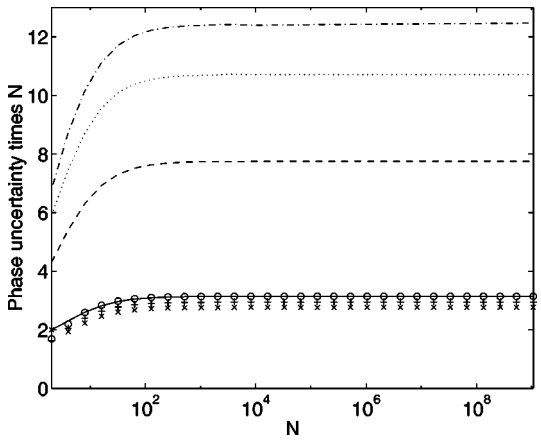


FIG. 5. The phase uncertainty of states optimized for minimum Holevo phase variance under several measures multiplied by *N*. The square root of the Holevo phase variance is shown by the continuous line, the square root of the standard variance is shown by circles, the inverse of the maximal value is shown by the dashed line, the Süssman measure is shown by the dotted line, the entropic length is shown by the dash-dotted line, the Fisher length is shown by the crosses, and the 67% confidence interval is shown by pluses.

In contrast, the optimal states we consider will give data that is easily analyzed, because the variance scales as the Heisenberg limit. The phase variance is a very stringent measure of the uncertainty, with inequalities such as [10,11]

$$\sqrt{2\pi e}\Delta X \geqslant L_H \geqslant \sqrt{2\pi e} L_F, \tag{2.17}$$

$$\frac{1}{\sqrt{1-C}}\Delta X \geqslant L_C, \tag{2.18}$$

where $\Delta X$ is the square root of the variance, $L_H$ is the entropic length [10], $L_F$ is the Fisher length, and $L_C$ is a $C \times 100\%$ confidence interval. These inequalities will also be true for the Holevo phase variance for sufficiently narrowly peaked distributions. The optimal states we consider have a Holevo phase variance scaling as $N^{-2}$, so the phase uncertainty scales as $N^{-1}$ for all these measures of uncertainty. In addition, we have found numerically that the phase uncertainty scales as $N^{-1}$, as estimated using the reciprocal-of-peak value $L_{\rm rp}$ and the Süssman measure $L_S$ (see Ref. [12] for a description of these). To demonstrate this, in Fig. 5 we plot all the common measures of uncertainty times *N*. In this figure only the Holevo phase variance and reciprocal-of-peak value were determined analytically. All other measures were determined using numerical integrals of the phase probability distribution. (The sum for the phase probability distribution can be evaluated analytically, and is therefore easily calculated).

It is clear from this plot that the asymptotic values for these measures are good approximations for photon numbers of order 100 or greater. For comparison, we have also calculated the corresponding measures for the state $|j0\rangle_z$. For this state the sum for the phase probability distribution cannot be evaluated analytically, but when the probability distribution is expressed as a sum, the integrals for all the measures except the entropic length can be determined. Because this state can only yield information about $\varphi$ modulo $\pi$, we have calculated all of these measures for the probability distribution of $\phi \bmod \pi$. For all of these measures except the variance, we have found that the state $|j0\rangle_z$ also yields an $N^{-1}$ scaling. Specifically, the asymptotic results compare to those for the optimal state $|\psi_{\rm opt}\rangle$, as shown in Table I. In addition we determined the scaling constants based on the Bessel function approximation given in Refs. [4,5], and the results agreed with those found using the exact calculations to the precision given in Table I.

TABLE I. The scaling constants for each of the measures of phase uncertainty that scale as $1/N$ for both the optimum state and $|j0\rangle_z$.

| Measure | $\|\psi_{\rm opt}\rangle$ | $\|j0\rangle_z$ |
|---|---|---|
| $NL_{\rm rp}$ | 7.752 | 6.875 |
| $NL_S$ | 10.711 | 12.305 |
| $NL_H$ | 12.415 | 35.79 |
| $NL_{2/3}$ | 2.948 | 3.071 |
| $NL_F$ | 2.766 | 1.414 |

From this we see that the optimal state is actually worse than the $|j0\rangle_z$ state under the reciprocal peak $L_{\rm rp}$ or the Fisher length $L_F$, but it is better under all the other measures of phase uncertainty. Although there is no clear distinction between the states on this basis, when one remembers that $|j0\rangle_z$ can only be used to measure phase modulo $\pi$, and that it yields a non-Heisenberg limited variance, the optimal state $|\psi_{\rm opt}\rangle$ is clearly better overall.

It is interesting to note that the coefficients we have found for the confidence interval (3.071) and the Fisher length (1.414) for the state $|j0\rangle_z$ differ from those found in Ref. [5], of 3.36 and 2, respectively. The slightly different scaling for the confidence interval appears to be due to the asymptotic Bessel function expansion used in Ref. [5]. The difference in the Fisher length is because the Bessel function approximation in Ref. [5] is not correctly normalized. The Bessel function approximation approximates the exact distribution over the interval $[-\pi/2,\pi/2]$, but the exact expression is normalized over the interval $[-\pi,\pi]$ and repeats modulo $\pi$. This means that the approximate expression in Ref. [5] is too small by a factor of 2. In our calculations we ensured that the distribution is correctly normalized over the interval $[-\pi/2,\pi/2]$.

Like the state $|j0\rangle_z$, the state $(|j0\rangle_z+|j1\rangle_z)/\sqrt{2}$ considered in Ref. [2] also suffers from the problem of yielding a significant number of results with large error. However, the problem is far worse in this case, so that the variance is always of order unity, as seen in Fig. 3. Although the central peak becomes narrower with the photon number, there are also peaks near $\pm\pi$, which although smaller than the main peak, do not become smaller with the photon number. An example of the phase distribution is shown in Fig. 6.

## III. FEEDBACK TECHNIQUE

The second problem is how to make these canonical measurements in practice. As we will show in Sec. VII, it is not

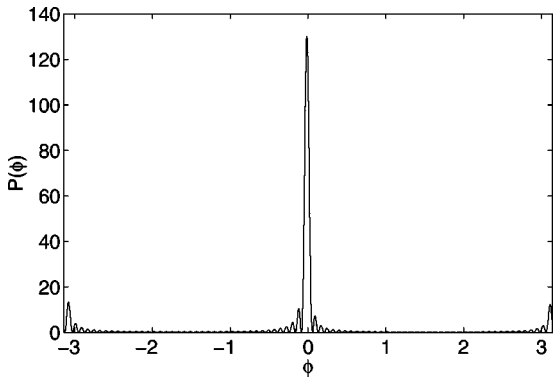


FIG. 6. The canonical phase probability distribution for $(|j0\rangle_z + |j1\rangle_z)/\sqrt{2}$ for $2j=80$ photons.

possible to make canonical measurements using photodetections at the output of the interferometer. It was shown [9,13,14] that in the single-mode case it is possible to make very good phase measurements by using feedback to an auxiliary phase shift. In order to apply the same principle here, we consider adaptive measurements where the phase we wish to measure, $\varphi$, is in one arm of the interferometer, and we introduce a known phase shift $\Phi$ into the other arm of the interferometer as in Fig. 1. After each photodetection we adjust this introduced phase shift in order to maximize the sharpness of the phase estimate after the next photodetection.

For this adaptive scheme to work, the feedback that adjusts $\Phi$ must act much faster than the average time between photon arrivals. For simplicity we make the assumption that the feedback is arbitrarily fast, which simply means that the phase $\Phi$ is assumed to have always been changed before the next detection occurs. It is the ability to change $\Phi$ during the passage of a single (two-mode) pulse that makes photon-counting measurements more general than a measurement of the output $\hat{J}_z$ considered in Refs. [2,3].

To describe this scheme, it is convenient to denote the result $u$ from the $m$th detection as $u_m$ (which is 0 or 1 depending on which output the photon is detected in), and the measurement record up to and including the $m$th detection as the binary string $n_m \equiv u_m \ldots u_2 u_1$. The input state after $m$ detections will be a function of the measurement record and $\varphi$, and we denote it as $|\psi(n_m,\varphi)\rangle$. In our calculations we do not normalize these states (except for the initial state), in order to express the state as a power series in $e^{i\varphi}$.

After the first beam splitter the operators for the two beams are (in the Heisenberg picture)

$$(\hat{a}+i\hat{b})/\sqrt{2}, \quad (i\hat{a}+\hat{b})/\sqrt{2}. \tag{3.1}$$

The two beams are then subjected to phase shifts of $\varphi$ and $\Phi$, so the operators become

$$e^{i\varphi}(\hat{a}+i\hat{b})/\sqrt{2}, \quad e^{i\Phi}(i\hat{a}+\hat{b})/\sqrt{2}. \tag{3.2}$$

Finally, the effect of the second beam splitter is the same as the first, giving the two operators

$$ie^{i(\varphi+\Phi)/2}(\hat{a}\sin\theta+\hat{b}\cos\theta),$$

$$ie^{i(\varphi+\Phi)/2}(\hat{a}\cos\theta-\hat{b}\sin\theta), \tag{3.3}$$

where $\theta=(\varphi-\Phi)/2$. Ignoring the unimportant initial phase factors, we can represent these as the operators $\hat{c}_0$ and $\hat{c}_1$, where

$$\hat{c}_u=\hat{a}\sin(\theta+u\pi/2)+\hat{b}\cos(\theta+u\pi/2). \tag{3.4}$$

The input state is then determined by the initial condition

$$|\psi(n_0,\varphi)\rangle=|\psi\rangle, \tag{3.5}$$

where $n_0$ is the empty string, and the recurrence relation

$$|\psi(u_m n_{m-1},\varphi)\rangle=\hat{c}_{u_m}(\varphi)\frac{|\psi(n_{m-1},\varphi)\rangle}{\sqrt{N+1-m}}. \tag{3.6}$$

These states are not normalized, and their norm represents the probability for the measurement record $n_m$ given $\varphi$. That is,

$$P(n_m|\varphi)=\langle\psi(n_m,\varphi)|\psi(n_m,\varphi)\rangle. \tag{3.7}$$

We can express an arbitrary input state with $N=2j$ photons as a sum of $\hat{J}_z$ eigenstates $|\psi\rangle=\Sigma_{\mu=-j}^{j}\psi_\mu|j\,\mu\rangle_z$. The state after $m$ detections will be (for $m\geqslant 1$) a function of $\varphi$. We denote it as follows:

$$|\psi(n_m,\varphi)\rangle=\sum_{\mu=-j+m/2}^{j-m/2}\psi_{\mu;m}(n_m,\varphi)|j-m/2;\mu\rangle_z. \tag{3.8}$$

Using the recurrence relation [Eq. (3.6)], we find that the functional form of $\psi_{\mu;m}(n_m,\varphi)$ is always

$$\psi_{\mu;m}(n_m,\varphi)=\sum_{k=-m/2}^{m/2}\psi_{\mu;m;k}(n_m)e^{ik\varphi}. \tag{3.9}$$

The recurrence relation for the coefficients $\psi_{\mu;m;k}$ is

$$\begin{aligned}\psi_{\mu;m;k}(n_m)=&\frac{e^{-i(\Phi_m-u_m\pi)/2}}{2\sqrt{2j-m+1}}[s_-\psi_{\mu-(1/2);m-1;k-(1/2)}(n_{m-1})\\&-is_+\psi_{\mu+(1/2);m-1;k-(1/2)}(n_{m-1})]\\&+\frac{e^{i(\Phi_m-u_m\pi)/2}}{2\sqrt{2j-m+1}}\\&\times[s_-\psi_{\mu-(1/2);m-1;k+(1/2)}(n_{m-1})\\&+is_+\psi_{\mu+(1/2);m-1;k+(1/2)}(n_{m-1})],\end{aligned} \tag{3.10}$$

where

$$s_\pm=\sqrt{j-\frac{m}{2}\pm\mu+1}. \tag{3.11}$$

We determine the feedback phase before the $m$th detection, $\Phi_m$, based on the measurement record $n_{m-1}$. To determine this phase, we need to determine the probability distribution for the unknown phase, $\varphi$. To determine this distribution we use Bayes' theorem

$$P(\varphi|u_m,n_{m-1})=\frac{P(\varphi|n_{m-1})P(u_m|\varphi,n_{m-1})}{P(u_m|n_{m-1})}. \tag{3.12}$$

Here $P(\varphi|n_{m-1})$ is the prior distribution for $\varphi$ before the $m$th detection. A similar Bayesian approach to interferometry has been considered before, and used in analyzing experimental results [15]. However, this was done only with nonadaptive measurements, and with all particles incident on one port.

Now the divisor $P(u_m|n_{m-1})$ is independent of the phase, and therefore only provides a normalizing factor to the phase distribution. Ignoring this normalization factor, the phase distribution changes with each step as

$$\begin{aligned}P(\varphi|n_m)&\propto P(\varphi|n_{m-1})P(u_m|\varphi,n_{m-1})\\&\propto P(\varphi|n_{m-1})\frac{\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle}{\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle}.\end{aligned} \tag{3.13}$$

This means that we must have

$$P(\varphi|n_m)\propto\langle\psi(n_m,\varphi)|\psi(n_m,\varphi)\rangle. \tag{3.14}$$

To show this, note first that this relation is trivial before any counts are made, because the distribution is flat, and the initial state is normalized, so there is no $\varphi$ dependence. Now, if Eq. (3.14) is true for $m=k$,

$$\begin{aligned}P(\varphi|n_{k+1})&\propto P(\varphi|n_k)\frac{\langle\psi(n_{k+1},\varphi)|\psi(n_{k+1},\varphi)\rangle}{\langle\psi(n_k,\varphi)|\psi(n_k,\varphi)\rangle},\\&\propto\langle\psi(n_k,\varphi)|\psi(n_k,\varphi)\rangle\\&\quad\times\frac{\langle\psi(n_{k+1},\varphi)|\psi(n_{k+1},\varphi)\rangle}{\langle\psi(n_k,\varphi)|\psi(n_k,\varphi)\rangle},\\&\propto\langle\psi(n_{k+1},\varphi)|\psi(n_{k+1},\varphi)\rangle,\end{aligned} \tag{3.15}$$

so Eq. (3.14) is true for $m=k+1$. Therefore, Eq. (3.14) is true for all $m$ by induction.

To minimize the Holevo phase variance, we wish to maximize the sharpness $S_\phi$. The closer $S$ is to unity, the sharper the distribution. After the result $u_m$ the sharpness of the posterior distribution is

$$\begin{aligned}S(u_m|n_{m-1})&=P(u_m|n_{m-1})^{-1}\\&\quad\times\left|\int P(\varphi|n_{m-1})P(u_m|\varphi,n_{m-1})e^{i\varphi}d\varphi\right|.\end{aligned} \tag{3.16}$$

We wish to know, given $|\psi\rangle$, $n_{m-1}$, and $P(\varphi|n_{m-1})$, what value to choose for $\Phi_m$ so as to gain the most information (on average) from the measurement, which gives the result $u_m$. One way to quantify this is to maximize the average sharpness of the posterior distribution of $\varphi$ after the result $u_m$, weighted by the probability of that result occurring. That is, we can take our maximand to be

$$\begin{aligned}M(\Phi_m)&=\sum_{u_m=0,1}P(u_m|n_{m-1})S(u_m|n_{m-1}),\\&=\sum_{u_m=0,1}\left|\int_0^{2\pi}P(\varphi|n_{m-1})P(u_m|\varphi,n_{m-1})e^{i\varphi}d\varphi\right|.\end{aligned} \tag{3.17}$$

Now, using

$$P(u_m|\varphi,n_{m-1})=\frac{\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle}{\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle}, \tag{3.18}$$

and Eq. (3.14), we find that

$$\begin{aligned}M(\Phi_m)&=\mathcal{N}_m(n_{m-1})^{-1}\sum_{u_m=0,1}\\&\quad\times\left|\int_0^{2\pi}\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle e^{i\varphi}d\varphi\right|,\end{aligned} \tag{3.19}$$

where $\mathcal{N}_m(n_{m-1})$ is a normalization factor.

In order to use this expression we require $\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle$ explicitly in terms of $u_m$. It is straightforward to show from Eq. (3.10) that

$$\begin{aligned}&\langle\psi(u_m n_{m-1},\varphi)|\psi(u_m n_{m-1},\varphi)\rangle\\&\quad=\tfrac{1}{2}[\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle\\&\qquad+\lambda_{m-1}(\varphi)e^{i\varphi}e^{-i(\Phi_m-u_m\pi)}\\&\qquad+\lambda^*_{m-1}(\varphi)e^{-i\varphi}e^{i(\Phi_m-u_m\pi)}],\end{aligned} \tag{3.20}$$

where $\lambda_{m-1}(\varphi)$ is defined by

$$\lambda_{m-1}(\varphi)=\sum_{n=-m+1}^{m-1}\lambda_{m-1;n}e^{in\varphi}, \tag{3.21}$$

where

$$\lambda_{m-1;n}=-(\xi_{m-1;n}+i\zeta_{m-1;n})/2, \tag{3.22}$$

and where

$$\begin{aligned}\xi_{m-1;n}=&\sum_{k,k'=-(m-1)/2}^{(m-1)/2}\ \sum_{\mu=-j+[(m-1)/2]}^{j-[(m-1)/2]}\\&\times\frac{\mu}{j-[(m-1)/2]}\psi_{\mu;m-1;k}\psi^*_{\mu;m-1;k'}\delta_{n,k-k'},\end{aligned} \tag{3.23}$$

$$\zeta_{m-1;n}=\sum_{k,k'=-(m-1)/2}^{(m-1)/2}\sum_{\mu=-j+[(m+1)/2]}^{j-[(m+1)/2]}\times\frac{\sqrt{\left(j-\frac{m-3}{2}+\mu\right)\left(j-\frac{m-1}{2}-\mu\right)}}{2\left(j-\frac{m-1}{2}\right)}\times(\psi_{\mu;m-1;k}\psi^{*}_{\mu+1;m-1;k'}+\psi^{*}_{\mu;m-1;-k}\psi_{\mu+1;m-1;-k'})\delta_{n,k-k'}. \quad (3.24)$$

Using this result it is possible to show that

$$M(\Phi_m)=\frac{1}{\mathcal{N}_m(n_{m-1})}\sum_{u_m=0,1}\times\left|\int_0^{2\pi}\frac{1}{2}[\langle\psi(n_{m-1},\varphi)|\psi(n_{m-1},\varphi)\rangle+\lambda_{m-1}(\varphi)e^{i\varphi}e^{-i(\Phi_m-u_m\pi)}+\lambda^{*}_{m-1}(\varphi)e^{-i\varphi}e^{i(\Phi_m-u_m\pi)}]e^{i\varphi}d\varphi\right|. \quad (3.25)$$

This can be expressed in the form

$$M(\Phi_m)=|a+be^{-i\Phi_m}+ce^{i\Phi_m}|+|a-be^{-i\Phi_m}-ce^{i\Phi_m}|, \quad (3.26)$$

where

$$a=\frac{\pi}{\mathcal{N}_m(n_{m-1})}\sum_{\mu=-j+[(m-1)/2]}^{j-[(m-1)/2]}\sum_{k=-(m-1)/2}^{(m-3)/2}\times\psi_{\mu;m-1;k}\psi^{*}_{\mu;m-1;k+1}, \quad (3.27)$$

$$b=\frac{\pi}{\mathcal{N}_m(n_{m-1})}\lambda_{m-1;-2}, \quad (3.28)$$

$$c=\frac{\pi}{\mathcal{N}_m(n_{m-1})}\lambda^{*}_{m-1;0}. \quad (3.29)$$

There is an analytical solution for the $\Phi_m$ that maximizes $M(\Phi_m)$. This solution gives three phases, $\Phi_0$ and $\Phi_\pm$, and the phase that is optimal must be found by substituting into Eq. (3.26). These phases are given by

$$\Phi_0=\arg(ba^{*}-c^{*}a) \quad (3.30)$$

and

$$\Phi_\pm=\arg\sqrt{\frac{c_2\pm\sqrt{c_2^2+|c_1|^2}}{c_1}}, \quad (3.31)$$

where

$$c_1=(a^{*}c)^2-(ab^{*})^2+4(|b|^2-|c|^2)b^{*}c, \quad (3.32)$$

$$c_2=-2i\,\mathrm{Im}(a^2b^{*}c^{*}). \quad (3.33)$$

Note that $M(\Phi)=M(\Phi+\pi)$, so that, in addition to the solution found by the above method, there will be another differing by $\pi$. It does not matter which of these we choose; it simply reverses the significance of the next results 0 and 1.

Since the original prior distribution $P(\varphi|n_0)$ is flat, there is no optimal phase $\Phi_1$, and we select this phase at random. At each following step, we determine the optimal feedback phase by the method described above, then determine the evolution of the state for that feedback phase. This process continues until all photons have been counted. The measurement record is then the binary string $n_{2j}=u_{2j}\ldots u_2u_1$ and the result is a posterior distribution $P(\varphi|n_{2j})$ that is proportional to $\langle\psi(n_{2j},\varphi)|\psi(n_{2j},\varphi)\rangle$, and is characterized by the $2j+1$ numbers $\psi_{0;2j;k}(n_{2j})$.

## IV. FINAL PHASE ESTIMATE

The third and final optimization problem is to determine what is the best final phase estimate to use. We define the best estimate to be that which minimizes the Holevo variance in the final phase estimate. This can be determined by summing over the $2^{2j}$ combinations of results, then averaging over the initial feedback phase. First, summing over the combinations of results gives us the probability distribution for the error in the phase estimate,

$$P(\phi|\varphi)=\sum_{[n_{2j}]=0}^{2^{2j}-1}P(n_{2j}|\varphi)\,\delta\{\phi-[\hat{\varphi}(n_{2j})-\varphi]\}, \quad (4.1)$$

where the square brackets in $[n_{2j}]$ denote the numerical value of this binary string interpreted as a binary number, and $\hat{\varphi}(n_{2j})$ is the final phase estimate.

Next we want to average over the initial feedback phase. Since it is only the differences between the feedback phases and the system phase that are significant, we can do this by averaging over the system phase. Performing this average gives

$$P(\phi)=\int d\varphi\,\frac{1}{2\pi}\sum_{[n_{2j}]=0}^{2^{2j}-1}P(n_{2j}|\varphi)\,\delta\{\phi-[\hat{\varphi}(n_{2j})-\varphi]\}, \quad (4.2)$$

$$=\frac{1}{2\pi}\sum_{[n_{2j}]=0}^{2^{2j}-1}P[n_{2j}|\hat{\varphi}(n_{2j})-\phi]. \quad (4.3)$$

From this probability distribution we can determine the exact phase variance for the measurement scheme. Evaluating $|\langle e^{i\phi}\rangle|$ we obtain

$$
\begin{aligned}
|\langle e^{i\phi}\rangle| &= \left| \int d\phi e^{i\phi} \frac{1}{2\pi} \sum_{[n_{2j}]=0}^{2^{2j}-1} P(n_{2j}|\hat{\varphi}(n_{2j})-\phi) \right| \\
&= \left| \frac{1}{2\pi} \sum_{[n_{2j}]=0}^{2^{2j}-1} e^{i\hat{\varphi}(n_{2j})} \right. \\
&\quad \left. \times \int d\phi e^{-i[\hat{\varphi}(n_{2j})-\phi]} P[n_{2j}|\hat{\varphi}(n_{2j})-\phi] \right| \\
&= \left| \frac{1}{2\pi} \sum_{[n_{2j}]=0}^{2^{2j}-1} e^{i\hat{\varphi}(n_{2j})} \int dx e^{-ix} P(n_{2j}|x) \right| \\
&= \left| \frac{1}{2\pi} \sum_{[n_{2j}]=0}^{2^{2j}-1} e^{i\hat{\varphi}(n_{2j})} \int dx \right. \\
&\quad \left. \times e^{-ix} \langle \psi(n_{2j},x)|\psi(n_{2j},x)\rangle \right|. \qquad (4.4)
\end{aligned}
$$

From this it is clear that the phase estimate that maximizes $|\langle e^{i\phi}\rangle|$, and therefore minimizes the phase variance, is

$$
\begin{aligned}
\hat{\varphi}(n_{2j}) &= \arg \int e^{i\varphi} \langle \psi(n_{2j},\varphi)|\psi(n_{2j},\varphi)\rangle d\varphi \\
&= \arg \int e^{i\varphi} P(\varphi|n_{2j}) d\varphi \\
&= \arg \sum_{k=-j}^{j-1} \psi_{0;2j;k} \psi^*_{0;2j;k+1}. \qquad (4.5)
\end{aligned}
$$

This has two important consequences. First, the exact phase variance for the feedback technique can be determined using

$$
|\langle e^{i\phi}\rangle| = \frac{1}{2\pi} \sum_{[n_{2j}]=0}^{2^{2j}-1} \left| \int e^{i\varphi} \langle \psi(n_{2j},\varphi)|\psi(n_{2j},\varphi)\rangle d\varphi \right|. \qquad (4.6)
$$

Second, from Eq. (3.19), it is clear that the last feedback phase is always optimal. Unfortunately it is not possible to prove in a similar way that the other feedback phases are optimal, as we will discuss in Sec. VII.

## V. STOCHASTIC METHOD

Because of the large number of measurement records that need to be evaluated, the exact phase variance can only be determined for small photon numbers (up to 20 or 30). For larger photon numbers it is necessary to determine the phase distribution stochastically. We choose the system phase $\varphi$ to be zero, and select the initial feedback phase $\Phi_1$ randomly. This introduces no bias, as only the differences between the system and feedback phases are significant.

The measurement results are chosen with probabilities determined using $\varphi=0$, and the final phase estimates are determined using Eq. (4.5). For an ensemble $\{\phi_n\}_{n=0}^{M}$ of $M$ final phase estimates, the Holevo phase variance can be estimated by

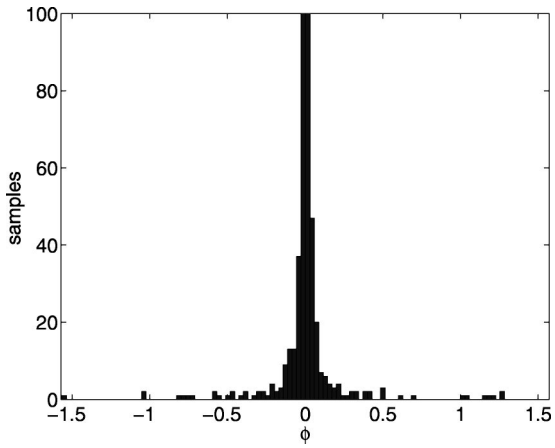


FIG. 7. The phase distribution that results from using adaptive phase measurements on an input state of $|j0\rangle_z$ for 800 photons. The vertical axis is cut off at 100 (the peak count is almost 500) to show the tails more clearly.

$$
V_\phi = \left| M^{-1} \sum_{n=1}^{M} e^{i\phi_n} \right|^{-2} - 1. \qquad (5.1)
$$

## VI. RESULTS

The results of using this adaptive phase measurement scheme on the four alternative input states are shown in Fig. 3. The phase variances for states up to $N=20$ (or $N=30$ for $|jj\rangle_z$) were determined exactly using Eq. (4.6), whereas those for larger photon numbers were determined using the stochastic method described in Sec. V. The phase variances are very close to the phase variances for canonical measurements for all of these input states.

For the optimal input states described in Sec. II, the scaling is close to $1/N^2$, but the phase variances differ relatively more from the canonical values for larger photon numbers. This indicates that the scaling is slightly worse than $1/N^2$, possibly $\log N/N^2$, as is the case for optimal single-mode phase measurements [14,13].

For $|j0\rangle_z$, the variances are very close to those for canonical measurements, scaling as $1/N^{1/2}$. If we look at the distribution of the phases resulting from these measurements, we find that there is a sharp peak, but a significant number of results with large error that produce the large variance (see Fig. 7). For the $|j0\rangle_z$ state we actually used a different feedback algorithm, one based on estimating the phase modulo $\pi$. The necessary changes in the relevant formulas are not difficult to derive.

For comparison, we have also considered the variance from two other measurement schemes. The first is a nonadaptive phase measurement introduced in Ref. [6], and defined by

$$
\Phi_m = \Phi_0 + \frac{m\pi}{N}. \qquad (6.1)
$$

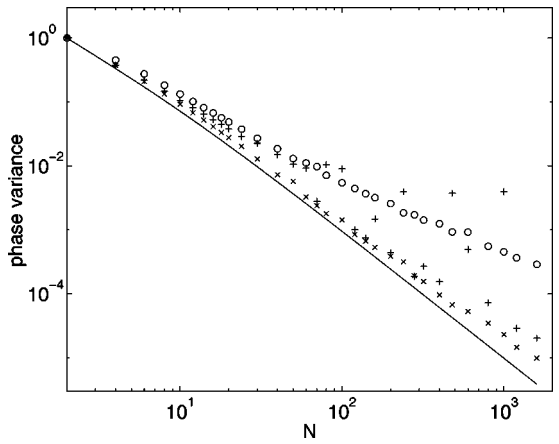


FIG. 8. The Holevo phase variance for optimal input states under various measurement schemes. The canonical phase variance is shown by the continuous line, the results for the adaptive measurement scheme of Sec. III by crosses, the nonadaptive measurement scheme of Eq. (6.1) by circles, and the feedback scheme of Eq. (6.2) by pluses.

This is analogous to heterodyne detection [14] on a single mode, in that the phase $\Phi$ equally weights all relevant values over the course of the measurement. The second scheme is a simple adaptive feedback scheme using a running estimate of the phase:

$$\Phi_m = \arg\langle e^{i\phi}\rangle = \arg \sum_{\mu=-j+[(m-1)/2]}^{j-[(m-1)/2]} \sum_{k=-(m-1)/2}^{(m-3)/2} \psi_{\mu;m-1;k}\psi^*_{\mu;m-1;k+1}. \tag{6.2}$$

This is motivated by the relative success of the analogous simple feedback scheme [14] for phase measurement of a single mode. The results of using these two measurement schemes, as well as the adaptive measurement scheme of Sec. III, on the optimal input states, are shown in Fig. 8.

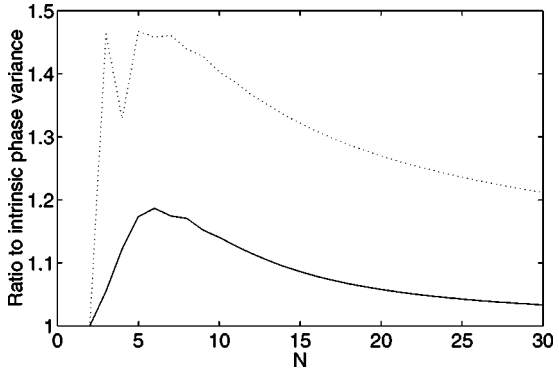


FIG. 9. The exact phase variance for $|jj\rangle_z$ input states under two different measurement schemes as a ratio of the canonical phase variance. The results for the adaptive measurement scheme of Sec. III are shown by the continuous line, and the nonadaptive measurement scheme of Eq. (6.1) by the dotted line.

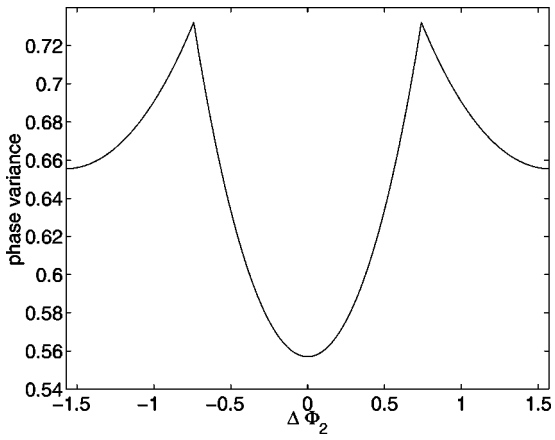


FIG. 10. The phase variance for the three-photon case as a function of the second feedback phase. The phase given here is relative to the second feedback phase given by the feedback scheme of Sec. III. The other feedback phases are given by this feedback scheme.

We find that the nonadaptive measurement scheme is far inferior to the adaptive measurement scheme of Sec. III, and the variance scales as $N^{-1}$. The simple adaptive feedback scheme also gives poor results. Although most of the phase results for this feedback scheme have small errors, there are a small number of results with very large errors. This also means that the results shown in Fig. 8 are fairly erratic, as for many of the results there were no phase samples with large errors, resulting in an underestimated phase variance.

We also considered a nonadaptive measurement scheme on a state with all photons in one port. The exact results for this case for $N$ up to 30 are shown in Fig. 9. The phase variance is not much more than the canonical phase variance, about 20% more for $N=30$ and still decreasing. This demonstrates that for this state there is relatively little improvement in using a more advanced feedback scheme for larger photon numbers. The largest improvement is about a 24% reduction in the variance for $N=3$.

## VII. OPTIMAL FEEDBACK

The next question is whether the adaptive measurement technique we have described is optimal. Note, first, that the initial feedback phase has no effect, because it is effectively averaged over by averaging over the system phase. Second, the last feedback phase is always optimal, as noted above. This means that for states with one or two photons the measurement technique must be optimal. In fact, for states considered in this paper the phase variance was equal to the canonical phase variance for one or two photons.

We will now restrict the discussion to optimal input states for simplicity. For three or four photons it was found that it is not possible to decrease the variance by altering the intermediate feedback slightly, showing that the feedback technique is locally optimal for the phase variance. For more than four photons it was possible to reduce the phase vari-

ance by varying the intermediate feedback phases, and so the feedback is not optimal.

In order to show that the feedback is globally optimal for three or four photons, it is necessary to test the entire phase range. There are three factors that reduce the number of phases that need be varied. The first two are as noted above: the first feedback phase has no effect (and so may be ignored), and the last feedback phase is always optimal (and so need not be varied). The third is that the contribution to the phase variance for a sequence of detections is independent of the first detection result. This is because changing the first feedback phase by $\pi$ reverses the significance of the first detection results, and the first feedback phase is arbitrary.

The consequences of these three factors are that for three photons the variation of only one feedback phase needs to be considered, and for four photons the variation of three feedback phases needs to be considered. The phase variance for the three-photon case, as the second feedback phase is varied from its value for our feedback technique is shown in Fig. 10. This figure shows that our feedback technique is globally optimal for three photons. Since the phase variance is above the canonical phase variance in this case, this demonstrates that it is not possible to perform canonical measurements using photodetection and feedback alone.

For four photons we need to vary the second feedback phase and the two third feedback phases. This case was tested with 100 steps in each of the three variables, and it was found that the feedback technique is globally optimal in this case also. Note that these results are only for the case of optimal input states. It would be a far more difficult problem to demonstrate that the feedback is optimal for general input states. Since optimal input states have the smallest phase variance, however, showing that the feedback is optimal or close to optimal in this case means that the feedback should be close to optimal for more arbitrary input states.

In order to see how far the phase variance could be improved for photon numbers above four, we optimized the feedback phases using function minimization techniques, and the results are shown in Fig. 11. Unfortunately the number of feedback phases increases exponentially with the photon number, making this technique infeasible very rapidly, and therefore only results up to $N=12$ are shown in Fig. 11. As can be seen, this optimization only gives minor improvements in the phase variance, with the maximum reduction in the phase variance being about 3.5%.

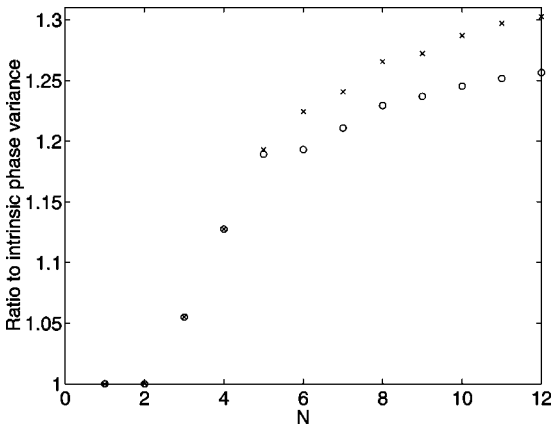


FIG. 11. The phase variance for the feedback scheme of Sec. III and a numerically optimized feedback scheme as a ratio to the intrinsic phase variance. The case for the feedback scheme of Sec. III is shown by crosses, and the numerically optimized case by circles.

## VIII. CONCLUSION

The state considered most in the past for attaining the Heisenberg limit in interferometry is the $|j0\rangle_z$ state (equal photon numbers in each input port). We have shown that this state cannot give a phase variance that is close to the Heisenberg limit. However, we have easily derived the optimal states that do give a Heisenberg-limited variance. These states only require significant contributions from about ten joint-number states near $|j0\rangle_z$. Since significant contributions are required from only a small number of $\hat{J}_z$ eigenstates, it should be possible to generate approximations of these states using an apparatus similar to that described in Ref. [2].

The variance of the optimal states under canonical measurements scales as $1/N^2$, in contrast to $1/N^{1/2}$, which is the scaling for the $|j0\rangle_z$ states. This anomalously high phase variance is due to the high tails in the phase distribution for the $|j0\rangle_z$ states. Other measures of spread give the same (Heisenberg-limited) scaling for both states. However, our optimal states also have the advantage that they can be used to measure phase modulo $2\pi$, not merely modulo $\pi$.

We have also shown that it is possible to approximate canonical measurements very closely using feedback and final phase estimates based on Bayesian inference. For one or two photons the measurements are optimal, and identical to canonical phase measurements for the input states considered in this paper. For larger photon numbers the measurements are extremely close to canonical phase measurements, with a phase variance scaling very close to $1/N^2$ for optimal input states. By contrast, naive (non-Bayesian) feedback schemes, and nonadaptive schemes, give far worse results.

For three- or four-photon optimal states, the feedback algorithm is the best possible, but the phase variance is above the canonical phase variance. This demonstrates that it is not possible to perform canonical measurements using photodetection of the interferometer outputs even with feedback. For small photon numbers above four it is possible to determine the optimal feedback by numerical function minimization techniques. However, the improvement over our original Bayesian algorithm is not great.

For relatively small photon numbers, all of the adaptive and nonadaptive algorithms considered should not be difficult to implement experimentally. Because the optimal input state, and the $|j0\rangle_z$ state, are difficult to create experimentally, initial experiments would probably be done with all photons incident on one input port (a $|jj\rangle_z$ state). For this state all measurement schemes give a phase variance scaling of $N^{-1}$. However, the adaptive scheme does give a variance up to 24% lower than the obvious nonadaptive scheme. In

this context, it is worth noting that for this input state the quantum statistics of the particles are irrelevant, so the experiment could be done with fermions (e.g., neutrons [15]) as well as bosons.

Finally, we mention that it should be possible to generalize our feedback approach for multichannel interferometry, involving more than two channels and more than one phase shift. This would hopefully yield physically realizable measurement schemes which are almost as good as the canonical measurements for such interferometers [16].

## ACKNOWLEDGMENT

We thank Michael Hall for enlightening discussions regarding measures of uncertainty.

[1] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).
[2] B. Yurke, S. L. McCall, and J. R. Klauder, Phys. Rev. A **33**, 4033 (1986).
[3] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[4] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).
[5] B. C. Sanders, G. J. Milburn, and Z. Zhang, J. Mod. Opt. **44**, 1309 (1997).
[6] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[7] A. S. Holevo, Lect. Notes Math. **1055**, 153 (1984).
[8] G. S. Summy and D. T. Pegg, Opt. Commun. **77**, 75 (1990).
[9] H. M. Wiseman and R. B. Killip, Phys. Rev. A **56**, 944 (1997).
[10] M. J. W. Hall, Phys. Rev. A **62**, 012107 (2000).
[11] C. E. Weatherburn, *Mathematical Statistics* (Cambridge University Press, Cambridge, England, 1946).
[12] I. Bialynicki-Birula, M. Freyberger, and W. Schleich, Phys. Scr. **T48**, 113 (1993).
[13] D. W. Berry and H. M. Wiseman, Phys. Rev. A **63**, 013813 (2001).
[14] H. M. Wiseman and R. B. Killip, Phys. Rev. A **57**, 2169 (1998).
[15] Z. Hradil *et al.*, Phys. Rev. Lett. **76**, 4295 (1996).
[16] B. C. Sanders, H. de Guise, D. J. Rowe, and A. Mann, J. Phys. A **32**, 7791 (1999).