Adaptive phase estimation with two-mode squeezed vacuum and parity measurement

Zixin Huang, Keith R. Motes, Petr M. Anisimov, Jonathan P. Dowling, and Dominic W. Berry

Phys. Rev. A **95**, 053837 — Published 12 May 2017

DOI: 10.1103/PhysRevA.95.053837

Adaptive phase estimation with two-mode squeezed-vacuum and parity measurement

Zixin Huang,[1,*] Keith R. Motes,[2] Petr M. Anisimov,[3] Jonathan P. Dowling,[4] and Dominic W. Berry[2]

[1] *School of Physics, University of Sydney, Sydney, NSW 2006, Australia*
[2] *Department of Physics and Astronomy, Macquarie University, Sydney NSW 2113, Australia*
[3] *Los Alamos National Laboratory, New Mexico 87545*
[4] *Hearne Institute for Theoretical Physics and Department of Physics & Astronomy, Louisiana State University, Baton Rouge, LA 70803*

A proposed phase-estimation protocol based on measuring the parity of a two-mode squeezed-vacuum state at the output of a Mach-Zehnder interferometer shows that the Cramér-Rao sensitivity is sub-Heisenberg [Phys. Rev. Lett. **104**, 103602 (2010)]. However, these measurements are problematic, making it unclear if this sensitivity can be obtained with a finite number of measurements. This sensitivity is only for phase near zero, and in this region there is a problem with ambiguity because measurements cannot distinguish the sign of the phase. Here, we consider a finite number of parity measurements, and show that an adaptive technique gives a highly accurate phase estimate regardless of the phase. We show that the Heisenberg limit is reachable, where the number of trials needed for mean photon number $\bar{n} = 1$ is approximately one hundred. We show that the Cramér-Rao sensitivity can be achieved approximately, and the estimation is unambiguous in the interval $(-\pi/2, \pi/2)$.

I. INTRODUCTION

Phase estimation and optical interferometry are the basis for many precision measurement applications. Coherent light based interferometry is most commonly used but its sensitivity for phase estimation is limited by the shot-noise limit, $\Delta\varphi^2 \geq \bar{n}^{-1}$, where φ is the unknown phase, and $\bar{n}$ is the mean number of photons used to perform the estimation [1]. This is not a problem in the case of limitless resources or in the case of samples that can withstand large doses of radiation. However, in order to achieve a finer precision given a finite amount of resources, one has to resort to interferometry with quantum states of light, such as N00N states [2] with parity measurements [3, 4], in order to achieve sub-shot-noise or the Heisenberg limit (HL) to sensitivity of phase estimation. Squeezed vacuum, which is the brightest experimentally available nonclassical light, has received much attention. In particular, the two-mode squeezed-vacuum (TMSV) which is the simplest two-mode state that contains strong photon-number entanglement, offers a notable improvement in phase estimation precision when compared to coherent states [4–7].

Significant advances have been made in quantum-enhanced phase sensitivity [8] and the meaning of the Heisenberg limit has been thoroughly examined [9, 10]. Yet, a proposed phase estimation scheme dips below the HL in the case of an infinite number of parity measurements [5]. That scheme is based on measuring the parity of the state of light at the output of a Mach-Zehnder interferometer (MZI), as shown in Fig. 1, with two-mode squeezed-vacuum input. A parity measurement focuses on whether the output photon number is odd or even, rather than the the actual number itself. It turned out that this particular scheme using TMSV input has sub-Heisenberg sensitivity, due to the fact that the photon number uncertainty for the state of light inside of the MZI is greater than the average photon number used for the measurement [11, 12].

The Heisenberg limit, $\Delta\varphi^2 \geq 1/(MN^2)$ for states with N fixed total number of photons such as twin-Fock [13] or N00N states [2] and M copies of the state, is a rigorous lower limit for *local* phase sensitivity for such states [14]. The measure of estimation error $\Delta\varphi$ is the root-mean-square error in estimating the parameter φ. However, for states with well-defined mean photon number but undefined total photon number, such as the TMSV used here, it is now understood that the Heisenberg limit so defined is not a lower limit.

Here we concern ourselves primarily with the Cramér-Rao bound, which is provably the ultimate limit of phase sensitivity [15]. For unbiased estimators, the estimation error is lower bounded by $\Delta\varphi^2 \geq 1/(MF)$, where F is the quantum Fisher information for the state and M is the number of measurements on copies of the state used in obtaining the estimate. When there is a unique maximally likely estimate, the bound is achievable in the asymptotic limit that $M \to \infty$. Hence, if we saturate this limit, as we do in this work with parity detection, then our measurement scheme is optimal. Of course, in practice the number of measurements that can be performed is limited, so the purpose of this paper is to investigate how $\Delta\varphi^2 M$ changes with M. Unlike in Ref. [5] where the Cramér-Rao bound sensitivity is expected in the limit of an infinite number of parity measurements, we consider a finite parity measurement record. We use the terminology "detection" to denote measurement on an individual copy of the state, and "measurement record" to denote the list of parity detections.

In previous schemes considered [4, 6, 7], it was unclear as to how an unambiguous estimate could be made, given that the parity of the output is symmetric around

* zixin.huang@sydney.edu.au

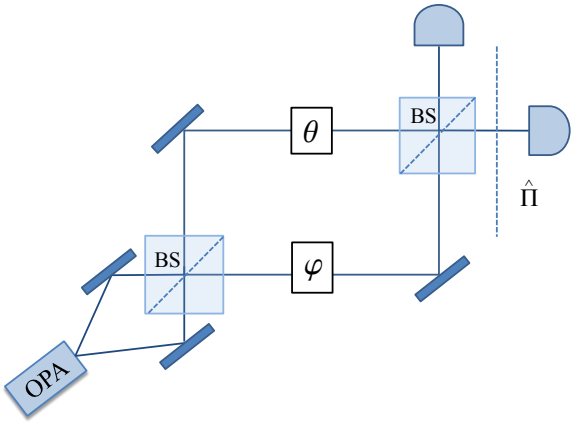

FIG. 1. Two-mode squeezed-vacuum states are generated at the input of the Mach-Zehnder interferometer (MZI) by an optical parametric amplifier (OPA), where φ is the phase being measured and θ is a controllable phase. The parity signal at the output of the MZI is then measured with photon-number-resolving detectors. Since TMSV states always have even photon numbers, the parity signal can be detected by performing photon-counting at only one output.

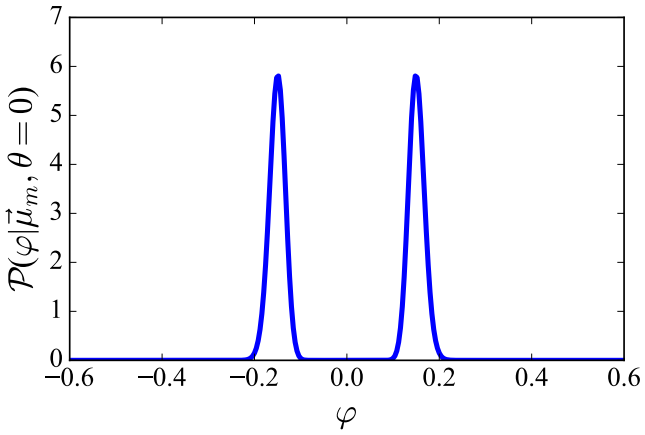

FIG. 2. The probability distribution for the phase φ for an example measurement record, generated for an actual phase of $\varphi = 0.15$ and the control phase $\theta = 0$. The record consists of $M = 512$ parity detections for $\bar{n} = 3$, of which 466 are even.

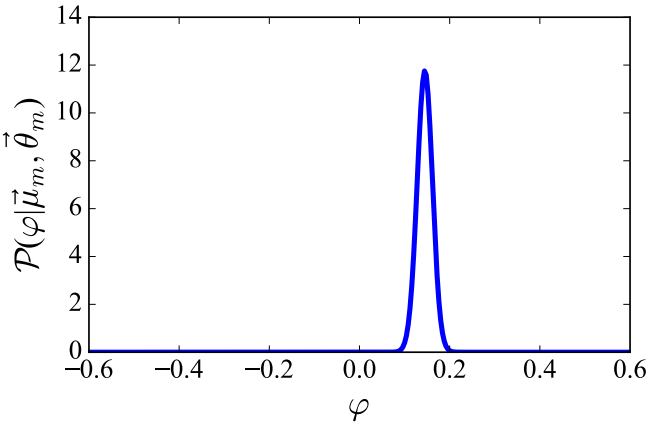

FIG. 3. The probability distribution for the phase φ for an example measurement record, generated for an actual phase of $\varphi = 0.15$ and θ being adaptive. The record consists of $M = 512$ parity detections for $\bar{n} = 3$. There is no ambiguity in the sign of φ.

the origin. If one infers the value of the measured phase φ by only considering the statistics (number of odd/even outcomes) of a static interferometer, there is ambiguity in the sign of the phase estimate. Figure 2 shows the probability distribution for the phase φ for an example measurement record, generated for an actual phase of 0.15. The distribution has two maxima distributed symmetrically around 0.

In addition, the most sensitive region is confined to a small interval where the relative phase between the two arms of the interferometer is zero, where the ambiguity of the sign is most problematic. Moving away from this region, the phase sensitivity decays quickly. In this paper we introduce an adaptive feedback scheme based on Bayesian techniques to eliminate this ambiguity; an example of the resulting probability distribution for the phase is shown in Fig. 3. By implementing an adaptable control phase θ we extend the region where the measurement is accurate to the interval $(-\pi/2, \pi/2)$. We show that with a sufficient number of detections, the method achieves Heisenberg-limited sensitivity. In this paper we define the Heisenberg limit as $\Delta\varphi^2 \equiv 1/(M\bar{n}^2)$ [1].

In a Mach-Zehnder interferometer, the achievable Fisher information is often phase-dependent (see for example Refs. [16–18]), and reaches the QFI for only one specific value of the phase. Adaptive phase feedback can overcome this issue by driving the measured phase to the optimal value. The phase-adaptive technique we adopt here is based on that in Refs. [19, 20]. This method has been experimentally demonstrated on both a photonic [21] and electron spin platform [22]. Both show an improved estimation precision that beats the shot noise limit. The technique uses Bayesian tools for calculating the probability distribution of the phase: it is known that the Bayesian estimator asymptotically saturates the Cramér-Rao bound. For example, this has been shown for the estimation of single parameter qubit gates [23] and for measuring the phase of the local oscillator of a TMSV state [24].

The structure of the paper is as follows. In Sec. II we describe the model of the TMSV state and discuss the problematic features associated with parity measurements. In Sec. III we detail the adaptive technique used to resolve these issues, and present results for the accuracy of the measurement scheme. We discuss the effects of noise in Sec. IV, then conclude in Sec. V.

II. MODEL

We consider a phase estimation scheme with a two-mode squeezed-vacuum input state which is commonly generated in unseeded optical parametric amplifiers. A TMSV state is ideally a superposition of twin Fock states,

$$|\psi_{\bar{n}}\rangle = \sum_{n=0}^{\infty} \sqrt{p_n(\bar{n})}\,|n,n\rangle\,, \tag{1}$$

where $p_n(\bar{n}) = (1-t_{\bar{n}})t_{\bar{n}}^n$, $t_{\bar{n}} = 1/\,(1+2/\bar{n})$, and $\bar{n}$ is the average photon number in the state [25].

Parity based phase estimation was originally introduced in quantum optics by Gerry [26] and is based on the parity of the photon number detected at the output of the MZI. Since the photon number in a TMSV state is always even, the parity signals are the same in both the output ports, and therefore can be detected by performing photon-counting at only one port. It turns out that parity detection is sufficient to achieve the Cramér-Rao bound in the case where the state is path symmetric [27], which is the case here. Propagation of the light through a MZI imprints phase information on the state that is retrieved by measuring parity at the output of the MZI. The expected value of the parity signal $\langle\hat{\Pi}\rangle$ for a TMSV based phase estimation scheme,

$$\langle\hat{\Pi}\rangle = \frac{1}{\sqrt{1+\bar{n}(\bar{n}+2)\sin^2(\theta-\varphi)}}, \tag{2}$$

was obtained in Ref. [5]. Equation (2) was obtained by summing the expected parity value $\langle\hat{\Pi}\rangle_n$ for each twin-Fock state $|n,n\rangle$, weighted by their respective probability of occurring,

$$\langle\hat{\Pi}\rangle = (1-t_{\bar{n}})\sum_{n=0}^{\infty} t_{\bar{n}}^n\,\langle\hat{\Pi}\rangle_n\,, \tag{3}$$

where $\langle\hat{\Pi}\rangle_n = (-1)^n\mathcal{L}_n[\cos(2(\theta-\varphi+\frac{\pi}{2}))]$ and $\mathcal{L}_n$ is the Legendre polynomial of order n [4].

A straightforward method for determining the magnitude of the unknown phase φ is as follows. If we let the controllable phase be $\theta = 0$ and we know $\bar{n}$ at the input, then we send the TMSV through the MZI of Fig. 1, which interrogates the unknown phase φ. Then, we perform parity measurements at the output, which returns either an even or odd outcome with probabilities $\mathcal{P}_\mathrm{e}$ and $\mathcal{P}_\mathrm{o}$ (as defined below), respectively. This allows us to determine the parity signal $\hat{\Pi}$.

This parity measurement can be implemented with photon-number-resolving detectors or homodyne detection [28, 29]. Using balanced homodyne detection and detectors with 99% quantum efficiency, phase estimation precision above the shot noise limit has been experimentally demonstrated with squeezed vacuum [16].

Inferring the parity of a state disregards the actual number of photons detected and focuses on whether this number is even or odd. Since $\mathcal{P}_e + \mathcal{P}_o = 1$ and the expectation value of a state's parity is $\langle\hat{\Pi}\rangle = \mathcal{P}_e - \mathcal{P}_o$, the probabilities of detecting an even or odd photon number are

$$\mathcal{P}_\mathrm{e} = \frac{1}{2}(1+\langle\hat{\Pi}\rangle), \qquad \mathcal{P}_\mathrm{o} = \frac{1}{2}(1-\langle\hat{\Pi}\rangle). \tag{4}$$

When the estimate is unbiased [30], the precision of the estimate is lower-bounded by the Cramér-Rao bound, $\Delta\varphi^2 \le 1/M\mathcal{F}(\varphi)$, where M is the number of times the estimation is repeated and $\mathcal{F}(\varphi)$ is the Fisher information [15]. In this case, the Fisher information is

$$\begin{aligned}\mathcal{F}(\varphi) &= \frac{1}{\mathcal{P}_\mathrm{e}}\left(\frac{\partial\mathcal{P}_\mathrm{e}}{\partial\varphi}\right)^2 + \frac{1}{\mathcal{P}_\mathrm{o}}\left(\frac{\partial\mathcal{P}_\mathrm{o}}{\partial\varphi}\right)^2 \\ &= \frac{1}{1-\langle\hat{\Pi}\rangle^2}\left(\frac{\partial\langle\hat{\Pi}\rangle}{\partial\varphi}\right)^2 \\ &= \frac{\cos^2(\theta-\varphi)\bar{n}(\bar{n}+2)}{\left[1+\bar{n}(\bar{n}+2)\sin^2(\theta-\varphi)\right]^2}.\end{aligned} \tag{5}$$

The Fisher information is maximized for $\theta-\varphi = 0$. Therefore, the ultimate precision of this estimation scheme is $1/M\bar{n}(\bar{n}+2)$, which is sub-Heisenberg if the bound can be achieved. When $F \ge \bar{n}$, the scheme performs better than the shot-noise limit. This would be achieved if $|\varphi-\theta|$ is smaller than approximately $\bar{n}^{-1/4}$. As $|\varphi-\theta|$ increases, the sensitivity decays quickly.

One can estimate φ from the statistics of the detections. In order to choose the estimate, it is useful to determine a probability density function for the phase based on the detection results. When the controllable phase θ is constant, this probability density can be determined via Bayes' theorem to be

$$\mathcal{P}(\varphi) \propto \mathcal{P}_\mathrm{e}^{\ell}(\varphi)\mathcal{P}_\mathrm{o}^{M-\ell}(\varphi), \tag{6}$$

where M is the number of parity detections and ℓ is the number of even results. In this expression φ is being used as a dummy variable for the system phase. The estimate of the actual system phase could be chosen to be, for example, the maximum of this probability distribution.

However, this approach cannot distinguish the sign of the phase. For example, for $\bar{n} = 3$ and $\theta = 0$, if 512 parity measurements are performed and 466 turn out to be even, the probability distribution for the phase is proportional to

$$\mathcal{P}(\varphi) \propto \mathcal{P}_\mathrm{e}^{466}(\varphi)\mathcal{P}_\mathrm{o}^{46}(\varphi). \tag{7}$$

It has two maxima in the interval $[-\pi/2, \pi/2]$: one at $\varphi = 0.18$ and the other at $\varphi = -0.18$ (see Fig. 2). This ambiguity comes about because the term containing φ is squared, so $\mathcal{P}(\varphi) = \mathcal{P}(-\varphi)$. Therefore 0.18 and -0.18 are equally valid estimates and cannot be distinguished.

One can model adaptive phase estimation with TMSV and parity detection numerically in the following way. Choose an average photon number $\bar{n}$ and an actual value

of the system phase φ, then generate a measurement record of finite length M using the probabilities $\mathcal{P}_{\rm e}$ and $\mathcal{P}_{\rm o}$. Each detection event is simulated from $\mathcal{P}_{\rm e}(\theta,\varphi)$ by generating a random real number between 0 and 1, then recording an even (odd) event if this number is less (more) than $\mathcal{P}_{\rm e}(\theta,\varphi)$. After the detection, the controllable phase θ is updated according to a method described in the next section. The new value of θ is used to recalculate $\mathcal{P}_{\rm e}(\theta,\varphi)$ to simulate the next detection. At the end an estimate of the phase is determined based on the measurement results. Note that the values of θ and the final estimate of φ are determined entirely from the measurement results (without using the actual value of the phase), just as they would be in an actual experiment. With this adaptive approach the ambiguity is eliminated, as illustrated in Fig. 3.

III. THE MEASUREMENT SCHEME

In the adaptive technique, the controlled phase θ is adjusted based upon previous detection results and controlled phases. Intuitively, this method works by maximizing $|\langle e^{i\varphi}\rangle|$ thereby reducing the appearance of multiple peaks [19]. Since

$$\langle e^{i\varphi}\rangle = \int_0^{2\pi} e^{i\varphi}\mathcal{P}(\varphi)\,d\varphi, \tag{8}$$

$|\langle e^{i\varphi}\rangle|$ is maximum when $\mathcal{P}(\varphi)$ is a delta function, and equals zero when $\mathcal{P}(\varphi)$ is flat. Here we use a variation of the method in Ref. [19] where we maximize $|\langle e^{i2\varphi}\rangle|$, as explained next.

The adaptive technique uses the latest probability distribution $\mathcal{P}(\varphi)$ to calculate the next controlled phase. Initially the distribution is flat, because there is no phase information. The initial controlled phase is therefore chosen to be random. Then, after each detection, the probability distribution is updated using Bayes' theorem

$$\mathcal{P}(\varphi|\vec{\mu}_m,\vec{\theta}_m) \propto \mathcal{P}(\mu|\varphi,\vec{\theta}_m)\mathcal{P}(\varphi|\vec{\mu}_{m-1},\vec{\theta}_{m-1}), \tag{9}$$

where μ denotes whether the detection is even or odd, $\vec{\mu}_m = (\mu_1,\mu_2,..,\mu_m)$ is the vector of successive detection results, and $\vec{\theta}_m = (\theta_1,\theta_2,..,\theta_m)$ is the vector of the corresponding controlled phases. The lower case m is used to indicate how many detections have been obtained so far within a series of M detections used to obtain a phase estimate. Therefore, μ_m indicates the m'th detection result, and θ_m indicates the controlled phase used for that detection (so after m detections the controlled phase is changed to θ_{m+1}).

The probabilities can be expressed as a Fourier series,

$$\mathcal{P}(\varphi|\vec{\mu}_m,\vec{\theta}_m) = \frac{1}{2\pi}\sum_{k=-K}^{K} a_k e^{ik\varphi}, \tag{10}$$

$\bar{n}$	No. of terms
1	10
2	10
3	15
5	20
8	25

TABLE I. For a TMSV state of mean photon $\bar{n}$, the cut-off for the number of Legendre polynomial terms used in Eq. (3).

where K is highest order of the Fourier coefficient. The coefficient of the term $e^{ik\varphi}$ is denoted a_k, which depends on $\vec{\mu}_m$ and $\vec{\theta}_m$ (the dependence is not given explicitly for simplicity). For example, $\mathcal{P}(\varphi) = \frac{1}{2\pi}[1+\cos(2\varphi)]$ can be written as

$$\mathcal{P}(\varphi) = \frac{1}{2\pi} + \frac{1}{4\pi}e^{i2\varphi} + \frac{1}{4\pi}e^{-i2\varphi}, \tag{11}$$

so $a_0 = 1$, $a_2 = a_{-2} = 1/2$, $a_1 = a_{-1} = 0$, and $a_k = 0$ for $|k| > 2$. Therefore we can take $K = 2$. As in this example, for the states we consider we only need even k because the probability distribution repeats modulo π. The Fourier representation of the probability distribution is particularly useful because the coefficients give the expectation values of exponentials as $a_{-k} = \langle e^{ik\varphi}\rangle$. Therefore $a_0 = \langle 1\rangle$ and is therefore 1 for a normalized distribution, as it is in this example. The coefficient $a_{-2} = \langle e^{i2\varphi}\rangle$ is one that we use below.

Before the first detection, Eq. (10) contains only one term, $a_0 = 1$. After each detection result given by the probabilities in Eq. (4), the Fourier coefficients a_k in Eq. (10) are updated using Eq. (9), which again uses Eq. (4). However, to exactly represent Eq. (4) in terms of Fourier coefficients, an infinite sum over the Legendre polynomial terms is needed in Eq. (3). In order to perform the numerical calculations, we truncated this infinite sum. It was found that the cut-off needed depended on $\bar{n}$; the values used are given in Table I.

We adjust the controlled phase θ_m based on the previous detection results and controlled phases. The value for θ_m is the one which maximizes the average sharpness of the probability distribution for φ after the next detection, across the interval $(-\pi/2,\pi/2)$. Here we take the sharpness to be $s(\theta) \equiv |\langle e^{i2\varphi}\rangle|$, which differs from the case in Ref. [19] where the sharpness is $|\langle e^{i\varphi}\rangle|$. We make this choice because the probability distribution for the parity detection results has a period of π instead of 2π.

The explicit expression for the average sharpness is

$$s_{\rm av}(\theta_m) = \frac{1}{2\pi}\sum_{\mu=\{+1,-1\}}\left|\int_0^{\pi} e^{i2\varphi}\prod_{k=1}^{m}\mathcal{P}(\mu_k|\varphi,\vec{\theta}_k)\,d\varphi\right|. \tag{12}$$

This expression corresponds to the average sharpness for the measurement results weighted by their probability of occurring. Maximizing this expression

yields the highest average accuracy of the phase estimates after the next detection [31]. The integral over φ simply yields the coefficient a_{-2}. Therefore, this integral may be obtained by summing the absolute value of a_{-2} for $\mathcal{P}(\varphi|\vec{\theta}_m, \vec{\mu}_{m-1}, \mu_m = \text{even})$ and $\mathcal{P}(\varphi|\vec{\theta}_m, \vec{\mu}_{m-1}, \mu_m = \text{odd})$ after the next detection. No analytical formula exists for calculating the optimal θ_m, so it was determined numerically [32].

At the end of each measurement record, the estimate is taken to be the argument of a_{-2}, corresponding to $\arg(\langle e^{2i\varphi}\rangle)$. This estimate of the phase is optimal for a measure of the measurement accuracy based on $|\langle\cos(2(\hat{\varphi}-\varphi)\rangle|$ [31], where $\hat{\varphi}$ is the value of an individual estimate. Here we are using the mean-square error (MSE), for which this estimate is not exactly optimal. (The MSE is an estimate of the error that is equal to the variance if the measurement is unbiased, and also appropriately penalises biased estimates.) However, this estimate is close to optimal for narrowly peaked distributions. In that case, the cosine function can be accurately approximated by expanding to second order, so

$$|\langle\cos(2(\hat{\varphi}-\varphi)\rangle| \approx 1 - 2\langle(\hat{\varphi}-\varphi)^2\rangle\,. \tag{13}$$

Therefore the estimate that maximizes $|\langle\cos(2(\hat{\varphi}-\varphi)\rangle|$ approximately minimizes the MSE $\langle(\hat{\varphi}-\varphi)^2\rangle$. This estimate of the phase is also close to unbiased for narrowly peaked distributions. This is because $\langle e^{2i(\hat{\varphi}-\varphi)}\rangle = 1$. For narrowly peaked distributions, the exponential can be expanded to first order, giving $1 + \langle 2i(\hat{\varphi}-\varphi)\rangle \approx 1$, which implies that $\langle\hat{\varphi}\rangle \approx \varphi$. Therefore, the Cramér-Rao bound should hold approximately.

Another feature of this estimate of the phase, is that together with the random initial controlled phase, it ensures that the measurement scheme is covariant. That is, the probability distribution for the error in the estimate is independent of the system phase. In order to fairly evaluate the overall performance of a measurement scheme, one should determine the MSE averaged over the system phase. When the measurement scheme is covariant, this averaging is unnecessary, because the MSE is independent of the system phase.

In order to estimate the performance of this adaptive scheme, measurement records were generated numerically and the controlled phases θ_m were calculated. After each sequence of M detection events, one obtains an estimate for the phase, $\hat{\varphi}$. The sequence of M detection events and resulting estimate were simulated J times, and the value of $\Delta\varphi$ was estimated by taking the root-mean-square error of the J estimates. Since simulating a large number of estimates is computationally intensive with large M, J was chosen such that the error in estimating $\Delta\varphi$ is approximately less than 3% of $\Delta\varphi$, keeping the simulation run time tractable.

It was found that in order to achieve the same precision in calculating $\Delta\varphi^2$, as M increases, J can be reduced. For example, for $M \in [64, 128]$, $J = 10^6$ phase estimates were performed to calculate $\Delta\varphi^2$, whereas when $M = 3096$, only $J = 5000$ estimates were necessary.

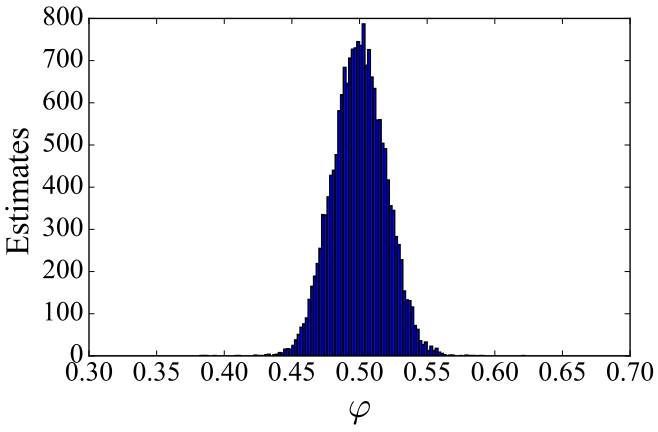

FIG. 4. The distribution of estimates for φ, where $M = 256$ detections were used to obtain each estimate. The plot contains results from $J = 20000$ measurement records which were numerically generated for $\varphi = 0.5$ and $\bar{n} = 3$. The distribution has a MSE $\Delta\varphi^2 = 3.81 \times 10^{-4}$.

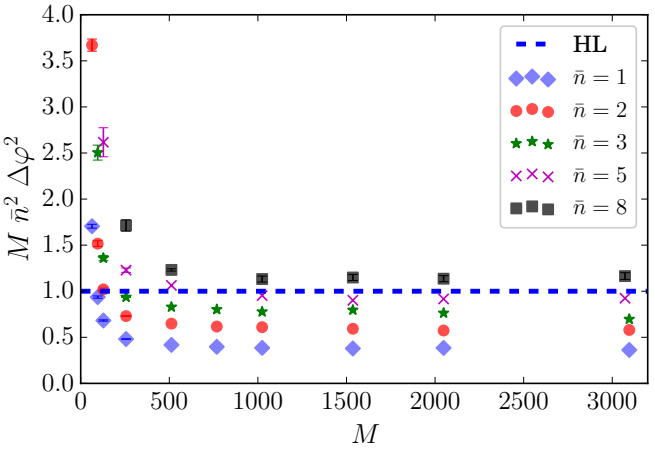

FIG. 5. The ratio of the phase MSE to the Heisenberg limit versus the measurement record length M, for TMSV states with a range of mean photon numbers: $\bar{n} = 1$ (blue diamonds), $\bar{n} = 2$ (red circles), $\bar{n} = 3$ (green stars), $\bar{n} = 5$ (purple crosses) and $\bar{n} = 8$ (black squares). Error bars are shown only if they are larger than the marker size. The Heisenberg limit (blue dashed line) is plotted for comparison (it is 1 because all values are shown as a ratio to the Heisenberg limit).

As an example, $J = 20000$ measurement records were generated for $\varphi = 0.5$ and $\bar{n} = 3$, where $M = 256$ detections were used for each estimate. Figure 4 shows the distribution of the estimates. There is a spread in the phase estimates with a MSE $\Delta\varphi^2 = 3.81 \times 10^{-4}$.

In Fig. 5 we show the ratio of the MSE to the Heisenberg limit against the length of the measurement record M. For most of the data points, the error bar is smaller than the marker. For $1.0 \leq \bar{n} \leq 5.0$, the MSEs beat the HL. (Since the HL is $\Delta\varphi^2 = 1/(\bar{n}^2 M)$, when multiplied by $M\bar{n}^2$, the HL equals 1 on this plot.) In Fig. 6, we show the ratio of the MSE to the quantum Cramér-Rao bound; as we expect, the MSEs asymptotically approach

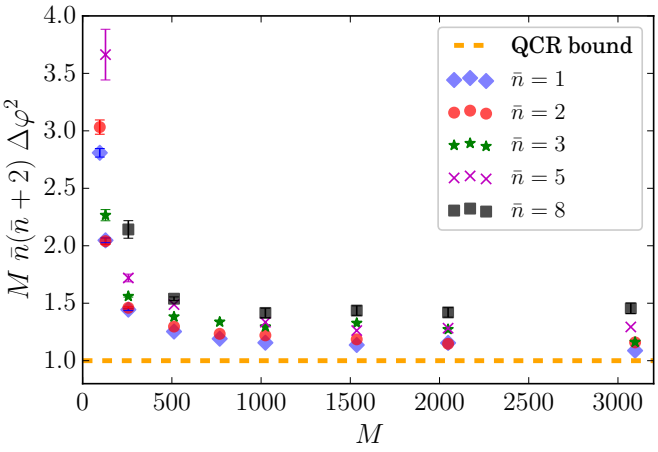

FIG. 6. The ratio of the phase MSE to the quantum Cramér-Rao bound versus the measurement record length M, for TMSV states with a range of mean photon numbers: $\bar{n}=1$ (blue diamonds), $\bar{n}=2$ (red circles), $\bar{n}=3$ (green stars), $\bar{n}=5$ (purple crosses) and $\bar{n}=8$ (black squares). The quantum Cramér-Rao bound (dashed orange line) is shown for comparison. The error bars are shown only if they are larger than the size of the markers.

this limit. Evidently, the smaller the mean photon number, the faster the MSEs converge to the bound. For $\bar{n}=1$, when $M=3096$, the phase MSE is larger than the Cramér-Rao bound by less than 10%. For larger $\bar{n}$, the relative difference from the Cramér-Rao bound is larger for the same M. Larger M would be needed to achieve agreement within 10%.

IV. MEASUREMENT SCHEME WITH NOISE

A. Losses

In reality photon-number-resolving detectors are not 100% efficient, and photonic states in an interferometer are subjected to loss. If losses are equal in both arms, then the inefficiency of the system can be combined and described by a single parameter η, where $1-\eta$ is the probability of losing a photon. To model detector inefficiency we use [33],

$$\mathcal{P}_{\rm D}(t|s)=\binom{s}{t}\eta^t(1-\eta)^{s-t}, \tag{14}$$

which gives the conditional probability of detecting t photons given that s photons were present.

The Fourier coefficients for the parity signal in the presence of loss are calculated as follows. First, calculate how the state $|n,n\rangle$ is transformed by the interferometer, according to

$$|n,n\rangle\rightarrow\frac{1}{2^{2n}}\sum_{\sigma=0}^{2n}\sqrt{\sigma!(2n-\sigma)!}\sum_{k=0}^{n}\frac{(-1)^k}{k!(n-k)!}e^{2ik\phi}\sum_{\ell=\max(0,\sigma-2n+2k)}^{\min(2k,\sigma)}\frac{(-1)^\ell(2k)!(2n-2k)!}{(2k-\ell)!\ell!(2n-2k-\sigma+\ell)!(\sigma-\ell)!}\,|\sigma,2n-\sigma\rangle\,. \tag{15}$$

Second, sum over n with the weightings $p_n(\bar{n})$. Third, apply Eq. (14) to obtain the new photon number distribution in the presence of loss, and finally, sum the even/odd-photon number events to give the parity signal. The net result of this calculation is that when η is less than 1, the visibilities of the signal peaks as a function of φ are reduced; this effect becomes more pronounced as $\bar{n}$ is increased.

Figures 7 and 8 show the phase MSE for $\bar{n}=1$ and $\bar{n}=3$ respectively. The phase MSE multiplied by $\bar{n}^2M$ is plotted against M for a range of values of η. The error bars are not shown if they are smaller than the marker size. For $\bar{n}=1$, when $\eta=0.95$, Heisenberg limit precision can be reached when M is approximately one thousand, whereas for $\bar{n}=3$, the variance does not reach this limit even for $\eta=0.99$. As is evident in the plot, when losses are equal to 10% for $\bar{n}=3$, the MSE increases by a factor of more than 10, whereas if single photon states or coherent states are used (MSE bounded by the shot-noise limit), the MSE would only increase by a factor of $1/\eta$. Therefore, in order to observe error below the Heisenberg limit (or even the shot-noise limit), the system must be highly efficient.

Another type of detector noise, coarse graining [34, 35], has a similar effect to inefficiency, since it means that the detection is unable to resolve the exact photon number. However, if photon numbers are grouped into bins, then the parity signal is completely destroyed and the scheme has no phase resolution.

B. Dark counts

Let the rate of dark counts be denoted r. That is, at the detectors at the two outputs of the interferometer, the probability of n dark counts follows the Poisson distribution

$$\mathcal{P}_{\rm dark}(n)=e^{-r}\frac{r^n}{n!}. \tag{16}$$

The parity is estimated just from the parity of the signal at a single output of the interferometer. If the number of dark counts is even, then there is no change in the parity,

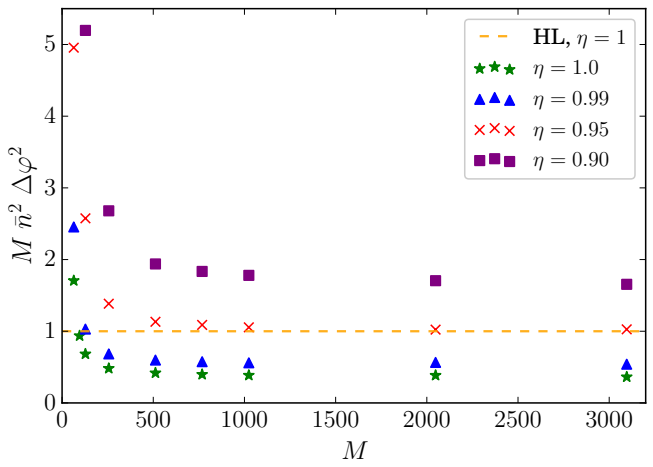

FIG. 7. The ratio of the phase MSE to the Heisenberg limit is plotted against the measurement record length M for $\bar{n} = 1$ and a range of levels of loss. The results shown are: $\eta = 1$ (green stars), $\eta = 0.99$ (blue triangles), $\eta = 0.95$ (red crosses), and $\eta = 0.90$ (purple squares) are shown. The Heisenberg limit for $\eta = 1$ (dashed orange line) is plotted for comparison. The error bars are not shown if they are smaller than the markers.

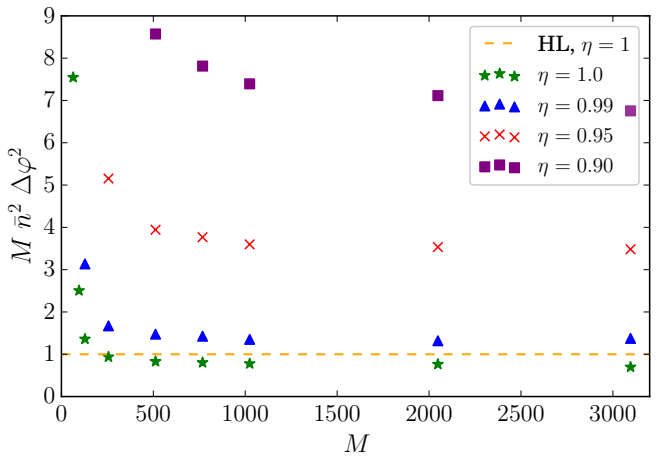

FIG. 8. The ratio of the phase MSE to the Heisenberg limit is plotted against the measurement record length M for $\bar{n} = 3$ and a range of levels of loss. The results shown are: $\eta = 1$ (green stars), $\eta = 0.99$ (blue triangles), $\eta = 0.95$ (red crosses), and $\eta = 0.90$ (purple squares) are shown. The Heisenberg limit for $\eta = 1$ (HL, dashed orange line) is plotted for comparison. The error bars are not shown if they are smaller than the markers.

but if the number of dark counts is odd then the parity is changed. Therefore the probability of registering the

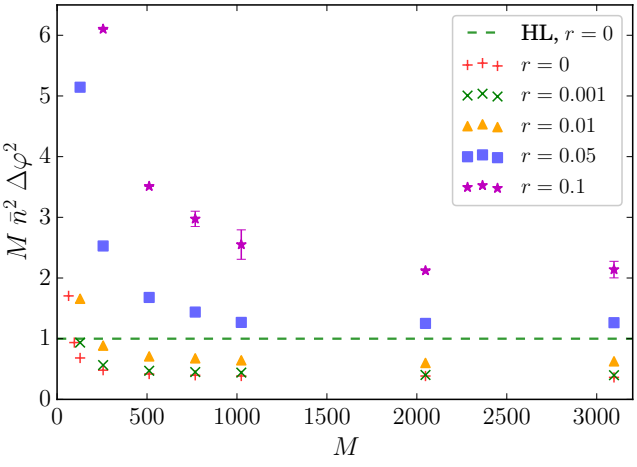

FIG. 9. The ratio of the phase MSE to the Heisenberg limit is plotted against the measurement record length M for $\bar{n} = 1$ and a range of dark count rates. The results shown are: $r = 0$ (red plusses), $r = 0.001$ (green crosses), $r = 0.01$ (orange triangles), $r = 0.05$ (blue squares), and $r = 0.1$ (purple stars). The Heisenberg limit for $r = 0$ (HL, dashed green line) is plotted for comparison. The error bars are not shown if they are smaller than the markers.

parity as even becomes

$$\begin{aligned}
\mathcal{P}'_{\rm e} &= \left(\sum_{n\ \text{even}} \mathcal{P}_{\rm dark}(n)\right)\mathcal{P}_{\rm e} + \left(\sum_{n\ \text{odd}} \mathcal{P}_{\rm dark}(n)\right)\mathcal{P}_{\rm o} \\
&= \mathcal{P}_{\rm e} + \left(\sum_{n\ \text{odd}} \mathcal{P}_{\rm dark}(n)\right)[\mathcal{P}_{\rm o} - \mathcal{P}_{\rm e}] \\
&= \frac{1}{2}(1 + \langle\hat{\Pi}\rangle) - e^{-r}\left(\sum_{n\ \text{odd}} \frac{r^n}{n!}\right)\langle\hat{\Pi}\rangle \\
&= \frac{1}{2}(1 + \langle\hat{\Pi}\rangle) - e^{-r}\frac{1}{2}(e^{r} - e^{-r})\langle\hat{\Pi}\rangle \\
&= \frac{1}{2}(1 + e^{-2r}\langle\hat{\Pi}\rangle). \qquad (17)
\end{aligned}$$

Similarly, we obtain

$$\mathcal{P}'_{\rm o} = \frac{1}{2}(1 - e^{-2r}\langle\hat{\Pi}\rangle). \qquad (18)$$

That is, the effect of the dark counts is simply to reduce the visibility. According to Ref. [36], where applicable, the fractional dark count rate compared to the maximum count rate ranges between 10^{-8} to 10^{-2}. The rate r is the number of dark counts at each detector during the measurement of the TMSV state, which involves detection of multiple photons. If we assume that each detector needs to detect on the order of 10 photons for small $\bar{n}$, then r is in the range 10^{-7} to 10^{-1}.

Figures 9 and 10 shows the phase MSE for $\bar{n} = 1$ and $\bar{n} = 3$ in the presence of dark counts. The phase MSE multiplied by $\bar{n}^2 M$ is plotted against M for $r = 0.001$, $r = 0.01$, $r = 0.05$ and $r = 0.1$. The error bars are not shown if they are smaller than the marker size. Evidently

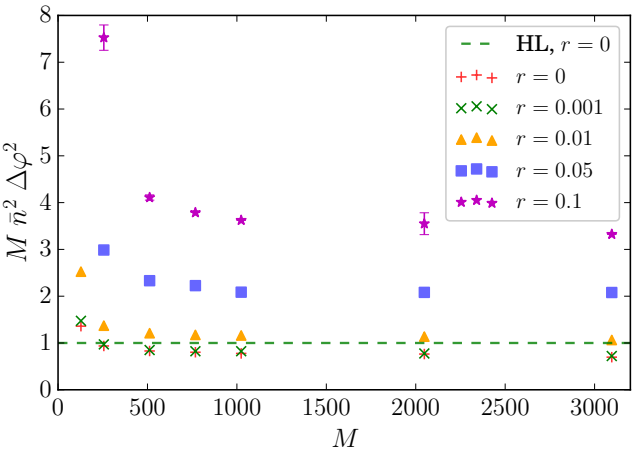

FIG. 10. The ratio of the phase MSE to the Heisenberg limit is plotted against the measurement record length M for $\bar{n} = 3$ and a range of dark count rates. The results shown are: $r = 0$ (red plusses), $r = 0.001$ (green crosses), $r = 0.01$ (orange triangles), $r = 0.05$ (blue squares), and $r = 0.1$ (purple stars). The Heisenberg limit for $r = 0$ (HL, dashed green line) is plotted for comparison. The error bars are not shown if they are smaller than the markers.

for both $\bar{n}$, when $r = 0.001$ the phase MSE is barely larger than those of the ideal case when $r = 0$. For this reason we do not plot the data points for smaller values of r.

For $\bar{n} = 1$, when $r = 0.01$, the Heisenberg limit can be reached when M is approximately 200. For $\bar{n} = 3$ with $r = 0.01$, the Heisenberg limit can still be reached, but M needs to be over 3000. When $r = 0.05$, the phase MSEs do not reach this bound, though for $\bar{n} = 3$, the precision is still better than the shot noise limit. (In Fig. 10 the shot noise limit is a straight line at $M\bar{n}^2\Delta\varphi^2 = 3$.)

The effect of dark counts can be predicted analytically by considering the Fisher information. For parity signals with the probability distribution in Eqs. (17) and (18), the Fisher information is

$$\begin{aligned}\mathcal{F}_r(\varphi) = {} & \bar{n}^2(\bar{n}+2)^2 e^{-4r}\sin^2(\theta-\varphi)\cos^2(\theta-\varphi) \\ & \times [\bar{n}(\bar{n}+2)\sin^2(\theta-\varphi)+1]^{-2} \\ & \times [\bar{n}(\bar{n}+2)\sin^2(\theta-\varphi) - e^{-4r}+1]^{-1}. \end{aligned} \quad (19)$$

A notable feature of this result is the exponential dependence on the dark count rate. Regardless of the value of $\bar{n}$, the dark counts will have a significant effect when the rate is comparable to 1. When $\bar{n} = 1$, the Fisher information of the parity scheme is lower than the shot noise limit when r is larger than ≈ 0.05, which is consistent with the simulation.

In comparison, the dark count rates have a much less significant effect on coherent states. If a coherent state with mean photon number $\bar{n}$ (which has shot-noise limited precision) is used in an MZI, dark counts simply increase the mean intensities at the two outputs,

$$\bar{n}_\pm = \bar{n}[1 \pm \cos(\theta-\varphi)]/2 + r. \quad (20)$$

The probabilities of detecting n photons at the two output ports are given by the Poisson distributions

$$\mathcal{P}_\pm(n) = e^{-\bar{n}_\pm}\frac{\bar{n}_\pm^n}{n!}. \quad (21)$$

This gives the Fisher information

$$\begin{aligned}\mathcal{F}_{\rm SQL}(\varphi) &= \sum_{n=0}^{\infty}\left(\frac{(\partial\mathcal{P}_+(n)/\partial\varphi)^2}{\mathcal{P}_+(n)} + \frac{(\partial\mathcal{P}_-(n)/\partial\varphi)^2}{\mathcal{P}_-(n)}\right) \\ &= \frac{\bar{n}^2(n+2r)\sin^2(\theta-\varphi)}{(n+2r)^2 - \bar{n}^2\cos^2(\theta-\varphi)}, \end{aligned} \quad (22)$$

This Fisher information is maximized at $\theta - \varphi = \pi/2$, where its value is $\bar{n}^2/(\bar{n}+2r)$. It is only decreased by a factor of $1 + 2r/\bar{n}$ from the case without dark counts. Provided the dark count rate is small compared to $\bar{n}$, it will not have a significant effect on the accuracy.

V. CONCLUSION

Schemes for phase measurement often consider N00N states, or equal photon numbers in both input ports for an interferometer [1, 13, 28, 37]. The states that are most commonly produced experimentally are the TMSV states. For these states, it was previously shown that parity detection yields a phase sensitivity, estimated from the Cramér-Rao bound, beyond the Heisenberg limit. Since then, it was found that parity detection attains the Cramér-Rao bound for a wide range of states including TMSV states [38].

However, the ability to actually achieve estimation with mean-square error near the Cramér-Rao bound was uncertain, because the probability distribution for the measurements is not well behaved. The sign of the phase is ambiguous, because for any nonzero result, the probability distribution for the phase is symmetric about zero. Moreover, the phase measurement is most sensitive in a small region about zero, where the ambiguity of the sign is most problematic (because widely separated peaks would be easier to distinguish).

In this work we showed that it is possible to perform adaptive measurements on a moderate number of copies of the state, and achieve small mean-square error, that is close to the Cramér-Rao bound in the interval $(-\pi/2, \pi/2)$. In particular, we find that it is possible to obtain mean-square error below the Heisenberg limit. It was found that the higher the mean photon number is in the TMSV state, the larger the number of trials is needed to reach the Heisenberg limit. For mean photon number $\bar{n} = 1$, approximately one hundred trials is sufficient, whereas for $\bar{n} = 5$, one thousand is required. As the mean number of the TMSV state is increased, the number of parity detections needed to approximate the Cramér-Rao bound also increases.

Those results are for an ideal system with no noise. When loss is present, we find that the total efficiency of

the system must be very close to one to observe precision below the Heisenberg limit: for mean photon number $\bar{n} = 3$, an efficiency larger than 0.99 would be necessary. We also find that in the presence of dark counts where the rate is lower than approximately 0.01, the Heisenberg limit can still be reached.

ACKNOWLEDGMENTS

DWB is funded by an ARC Future Fellowship (FT100100761) and an ARC Discovery Project (DP160102426). ZH is supported by the Australian Postgraduate Award scheme. KRM acknowledges the Australian Research Council Centre of Excellence for Engineered Quantum Systems (Project number CE110001013). JPD acknowledges support from the Air Force Office (Grant No. FA9550-13-10098), the Army Research Office (Grant No. W911NF-13-1-0381), the National Science Foundation (Grant No. 1403105), and the Northrop Grumman Corporation.

[1] J. P. Dowling, Contemp. Phys. **49**, 125 (2008).
[2] H. Lee, P. Kok, and J. P. Dowling, J. Mod. Opt. **49**, 2325 (2002).
[3] C. C. Gerry and J. Mimih, Contemp. Phys. **51**, 497 (2010).
[4] R. A. Campos, C. C. Gerry, and A. Benmoussa, Phys. Rev. A **68**, 023810 (2003).
[5] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, Phys. Rev. Lett. **104**, 103602 (2010).
[6] Y. Gao, P. M. Anisimov, C. F. Wildfeuer, J. Luine, H. Lee, and J. P. Dowling, JOSA B **27**, A170 (2010).
[7] K. P. Seshadreesan, P. M. Anisimov, H. Lee, and J. P. Dowling, N. J. Phys. **13**, 083026 (2011).
[8] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photon. **5**, 222 (2011).
[9] M. Zwierz, C. A. Pérez-Delgado, and P. Kok, Phys. Rev. A **85**, 042112 (2012).
[10] M. J. W. Hall, D. W. Berry, M. Zwierz, and H. M. Wiseman, Phys. Rev. A **85**, 041802 (2012).
[11] H. F. Hofmann, Phys. Rev. A **79**, 033822 (2009).
[12] P. Hyllus, L. Pezzé, and A. Smerzi, Phys. Rev. Lett. **105**, 120501 (2010).
[13] M. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[14] G. A. Durkin and J. P. Dowling, Phys. Rev. Lett. **99**, 070801 (2007).
[15] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Progress in Optics **60**, 345 (2015).
[16] A. A. Berni, T. Gehring, B. M. Nielsen, V. Händchen, M. G. Paris, and U. L. Andersen, Nat. Photon. **9**, 577 (2015).
[17] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[18] Z. Huang, C. Macchiavello, and L. Maccone, Phys. Rev. A **94**, 012101 (2016).
[19] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[20] D. W. Berry, H. M. Wiseman, and J. K. Breslin, Phys. Rev. A **63**, 053804 (2001).
[21] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature **450**, 393 (2007).
[22] C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, Nat. Nanotechnol. **11**, 247 (2016).
[23] B. Teklu, S. Olivares, and M. G. Paris, J. Phys. B **42**, 035502 (2009).
[24] S. Olivares and M. G. Paris, J. Phys. B **42**, 055506 (2009).
[25] C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2005).
[26] C. C. Gerry, Phys. Rev. A **61**, 043811 (2000).
[27] K. P. Seshadreesan, S. Kim, J. P. Dowling, and H. Lee, Phys. Rev. A **87**, 043833 (2013).
[28] K. P. Seshadreesan, P. M. Anisimov, H. Lee, and J. P. Dowling, N. J. Phys. **13**, 083026 (2011).
[29] W. N. Plick, P. M. Anisimov, J. P. Dowling, H. Lee, and G. S. Agarwal, N. J. Phys. **12**, 113025 (2010).
[30] D. W. Berry, M. J. W. Hall, M. Zwierz, and H. M. Wiseman, Phys. Rev. A **86**, 053813 (2012).
[31] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, Phys. Rev. A **80**, 052114 (2009).
[32] H. T. Dinani and D. W. Berry, Phys. Rev. A **90**, 023856 (2014).
[33] K. R. Motes, J. P. Dowling, and P. P. Rohde, Phys. Rev. A **88**, 063822 (2013).
[34] S. Raeisi, P. Sekatski, and C. Simon, Phys. Rev. Lett. **107**, 250401 (2011).
[35] D. S. Tasca, L. Rudnicki, R. M. Gomes, F. Toscano, and S. P. Walborn, Phys. Rev. Lett. **110**, 210502 (2013).
[36] R. H. Hadfield, Nat. Photon. **3**, 696 (2009).
[37] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photon. **5**, 222 (2011).
[38] K. P. Seshadreesan, S. Kim, J. P. Dowling, and H. Lee, Phys. Rev. A **87**, 043833 (2013).