PHYSICAL REVIEW A VOLUME 49, NUMBER 1 JANUARY 1994

Some limits to precision phase measurement

Samuel L. Braunstein
Department of Physics, Technion—Israel Institute of Technology, 32 000 Haifa, Israel
(Received 27 July 1993)

States near the vacuum having $\langle\hat{n}\rangle \ll 1$ are used for the precision determination of a phase shift. When this phase shift is sampled by enough of these near-vacuum states, its size may be extracted via data analysis. We calculate the achievable sensitivity of such schemes when we are limited in our resources to a mean total number of quanta $N_{\rm tot}$. We study in detail some examples of these near-vacuum schemes and show that in the absence of loss or noise their sensitivity approaches that of squeezed-state interferometry.

PACS number(s): 03.65.Bz, 42.50.Dv, 89.70.+c

I. INTRODUCTION

Given a mean total number of photons $N_{\rm tot}$, is there a limit to how precisely we can determine a shift in the phase of light? The traditional answer to this question has been that the particlelike nature of photons places shot noise as the limit on our ability to measure phase interferometrically, since intensity fluctuations are limited by photon-counting noise. This shot-noise limit to measuring a phase shift Φ is

$$\Delta\Phi \geq \frac{1}{2\sqrt{N_{\rm tot}}} .$$

Squeezed-state interferometry [1] is believed [2] to achieve a sensitivity scaling as

$$\Delta\Phi \simeq \frac{1}{N_{\rm tot}} , \tag{1}$$

with some $O(1)$ coefficient. Indeed, this method has been used to beat the shot-noise limit by several decibels [3].

There has been much work [4–11] in an attempt to beat squeezed-state interferometry by splitting the net resources $N_{\rm tot}$ into N independently measured packets and using data analysis. To determine how successful this is requires an understanding of the limits that data analysis places on our ability to extract information. To date, all work has concentrated on so-called "symmetric" schemes where N identical measurements are made on N identical quantum states, each having a mean number of photons $\langle\hat{n}\rangle$. For these symmetric schemes the constraint on resources reads

$$N_{\rm tot} = \langle\hat{n}\rangle N . \tag{2}$$

In this paper we study in some detail the symmetric schemes in the limit $\langle\hat{n}\rangle \to 0$, for which each packet used to "sample" the phase shift is near vacuum.

Shapiro, Shepard, and Wong (SSW) [4] were the first to propose that multiple measurement schemes aided by some kind of sophisticated data analysis may have advantages over single-shot measurement techniques. They studied how well a phase shift Φ could be determined in the state

$$|\psi_\Phi\rangle = e^{i\hat{n}\Phi}|\psi\rangle ,$$

given some fiducial state $|\psi\rangle$. Ordinarily we use interferometric methods to encode the phase shift Φ as an intensity. Instead, SSW looked at a measurement performed by a phase sensitive detector [12] based on the Susskind-Glogower operator $|\phi\rangle\langle\phi|$ [13], where

$$|\phi\rangle \equiv \frac{1}{\sqrt{2\pi}}\sum_{m=0}^{\infty} e^{im\phi}|m\rangle \tag{3}$$

in the number-state basis. SSW used crude statistical arguments to predict that, by splitting the resources for multiple measurements of the one-parameter family of fiducial states

$$|\psi\rangle \propto \sum_{m=0}^{M} \frac{1}{m+1}|m\rangle$$

and by employing maximum likelihood (ML) data analysis, a precision of

$$\Delta\Phi \simeq \frac{1}{N_{\rm tot}^2}$$

would be achievable. More careful calculations [9–11] have shown that this scheme approaches but never beats the sensitivity of Eq. (1) for squeezed-state interferometry, and so does not reach the sensitivity SSW claimed.

The scheme of SSW was extended by Shapiro and Shepard [5] by considering the two parameter family of fiducial states

$$|\psi\rangle \propto \sum_{m=0}^{M} \frac{1}{m+r}|m\rangle . \tag{4}$$

Interestingly, SSW's statistical argument suggests that the phase sensitivity for this variation would be arbitrarily good in the limit $r \to 0$. By contrast, Hradil and Shapiro [7,8] argued that in this limit there would be no sensitivity to measuring phase shifts whatsoever. We shall see that the truth lies somewhere in between these extremes. The limit $r \to 0$ corresponds to a fiducial state with a vanishingly small photon number ($\langle\hat{n}\rangle \to 0$). We consider these schemes as a general class and calculate the achievable sensitivity given the photon-number constraint in Eq. (2).

In addition to calculating the sensitivity of the scheme of Shapiro and Shepard for $r \to 0$, we consider a modification of a proposal by Dowling [6] to simplify the "manufacturability" of the original SSW fiducial state by replacing the sharp photon-number cutoff M by a smooth cutoff μ. This modified proposal is described by the fiducial state

$$|\psi\rangle \propto \sum_{m=0}^{\infty} \frac{e^{-m/\mu}}{m+r}|m\rangle \tag{5}$$

(with Dowling's original model corresponding to $r = 1$). We shall confirm, at least for the limit $r \to 0$, that the smooth cutoff does not degrade the scheme's sensitivity appreciably.

In Sec. II we discuss limitations to data analysis and review some recent results for the efficiency of maximum likelihood data analysis in the presence of resource constraints. This theoretical framework is applied to ideal phase measurement schemes in Sec. III. Section IV determines the sensitivity achievable in multiple measurement schemes in the limit of $\langle\hat{n}\rangle \to 0$ and sets it up as an optimization problem. Finally, in Sec. V we calculate the sensitivities achievable for the Shapiro-Shepard model and the modified Dowling model as described by the fiducial states given by Eqs. (4) and (5), respectively, in the limit $r \to 0$. We see that these sensitivities approach that of squeezed-state interferometry.

II. LIMITS TO DATA ANALYSIS

Although our primary interest is in precision phase measurement, the same concepts apply to the precision measurement of any parameter. For instance, suppose that we wish to find the parameter X that describes the location of the state

$$\psi_X(x) \equiv \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{(x-X)^2}{4\sigma^2}\right) ,$$

using N measurements of the position operator $\hat{x}$. Because our first measurement will disturb the system, the subsequent $N-1$ measurements must be made on identically prepared replicas of our quantum state. In this way, the question of quantum limitations partially maps onto the classical statistical problem of determining the parameter X via N independent samplings of the random variable x from the probability distribution $p(x|X)$; quantum theory tells us that $p(x|X) = |\psi_X(x)|^2$ and, in case we are interested in constraints, how to calculate $\langle\hat{n}\rangle$ from $\psi_X(x)$. Because, in our example, the state is Gaussian and the parameter is a so-called translation parameter [i.e., $p(x|X) = p(x-X)$], the most efficient form of data analysis is taking the mean value of the N measurements; our estimate $X_{\rm est}$ for the parameter X is then

$$X_{\rm est} = \frac{1}{N}\sum_{i=1}^{N} x_i ,$$

where x_i is the result of the ith sampling. (Note that when averaged over many N-sample trials, the mean of this estimate equals the actual value of the parameter, i.e., $\overline{X_{\rm est}} = X$.) The experimental uncertainty achieved from N measurements is easily calculated to be

$$\Delta X_{\rm est} = \frac{\sigma}{\sqrt{N}} .$$

A general result, valid even for non-Gaussian noise, is given by the Cramér-Rao lower bound (CRLB) [14]; starting with an arbitrary method of data analysis

$$X_{\rm est} = f(x_1, \ldots, x_N)$$

satisfying $\overline{X_{\rm est}} = X$, the CRLB states that

$$\Delta X_{\rm est} \geq \frac{1}{\sqrt{NF}} ,$$

where the Fisher information [14] is given by

$$F = \int dx\, \frac{1}{p(x|X)} \left(\frac{\partial p(x|X)}{\partial X}\right)^2 .$$

The CRLB is a limit to precision which no method of data analysis can beat.

What can we conclude from the CRLB? If there are no constraints on our resources then we can increase N without bound, in which case the CRLB tells us that there would be no lower bound to the precision $\Delta X_{\rm est}$. Unfortunately, the situation is significantly more complicated when constraints are included. In our case of symmetric schemes, as we increase N we must reduce the mean number of quanta available per state $\langle\hat{n}\rangle$ so as not to exceed the resources available in Eq. (2); this decrease in $\langle\hat{n}\rangle$, in turn, modifies the individual states measured and invariably reduces the Fisher information F. As a consequence, the main difficulty in studying these multiple-measurement schemes is to determine the optimal split in our available resources.

For Gaussian noise the CRLB is easily achieved for all N as seen above by using the mean of the data as our estimator. For non-Gaussian noise, however, this bound cannot be achieved in general for any method of data analysis. Thus, we must concentrate on specific methods. We use here the method called the maximum likelihood estimation; other methods are unlikely to be significantly more efficient. For the N data points $x_1, \ldots, x_N$, each distributed via $p(x|X)$, Bayes's theorem says that the "likelihood" for X will be

$$p(X|x_1, \ldots, x_N) \propto \prod_{i=1}^{N} p(x_i|X) ,$$

if we have no initial prejudice. The ML estimate $X_{\rm est}$ is the value of X that maximizes this likelihood function, and the standard deviation ΔX_{est} is a typical measure of our confidence in this estimate. Fisher's theorem [14] tells us that as $N \to \infty$ the likelihood function approaches a Gaussian distribution with standard deviation $\Delta X_{est} \to 1/\sqrt{NF}$. Thus ML estimation achieves the CRLB for large enough data sets and so is asymptotically efficient.

Braunstein [10,15] has argued that the optimal sensitivity based on ML estimation for a given mean number of quanta N_{tot} will occur just as this asymptotic regime is approached; he calls this regime the "knee" in ML estimation. He calculated the first $O(1/N)$ correction to Fisher's theorem to be

$$(\Delta X_{est})^2 = \frac{1}{NF}\left[1+\frac{N_{knee}}{N}+O\left(\frac{1}{N^2}\right)\right] , \tag{6}$$

where the quantity N_{knee} is a functional of the distribution of observations $p(x|X)$. Thus we can see that N_{knee} characterizes the sample size where the asymptotic behavior of Fisher's theorem is approached and so corresponds to Braunstein's knee. Thus, Braunstein predicts that the optimum multiple measurement sensitivity will occur at $N \simeq N_{knee}$ with

$$\frac{1}{N_{knee}F} \lesssim (\Delta X \mathrm{est})^2 \lesssim \frac{2}{N_{knee}F} . \tag{7}$$

III. APPLICATION TO IDEAL PHASE MEASUREMENTS

Consider the measurement scheme of SSW [4]. A state $|\psi\rangle$ is used to sample the phase shift Φ generating the state

$$e^{i\hat{n}\Phi}|\psi\rangle .$$

This state is then detected by some ideal device corresponding to the Susskind-Glogower operator so that the output statistics will be given by

$$p(\phi|\Phi) = |\langle\phi|e^{i\hat{n}\Phi}|\psi\rangle|^2 ,$$

where $|\phi\rangle$ is given by Eq. (3). For each state measured the detector's output will correspond to a random sampling from this probability distribution. By sampling the phase shift repeatedly a list of data $\phi_1, \ldots, \phi_N$ is obtained. In turn, this data is analyzed in order to extract the parameter Φ. All these measurements must, however, be done only with the resources available of a mean total number of quanta N_{tot}.

The results quoted in Sec. II state that, for an optimal split in N_{tot}, we need $N \simeq N_{knee}$ wave packets sampling the phase shift each with average photon number $\langle\hat{n}\rangle \simeq N_{tot}/N_{knee}$ where

$$N_{knee} = \frac{2}{F^2}\int_{-\pi}^{\pi} d\phi \left(\frac{p''^2}{p} - \frac{p'^4}{3p^3}\right) - 2 , \tag{8}$$

and the Fisher information is given by

$$F = \int_{-\pi}^{\pi} d\phi \frac{p'^2}{p} \tag{9}$$

[primes denote derivatives with respect to Φ and $p = p(\phi|\Phi)$]. Thus, the optimal sensitivity will be

$$\Delta\Phi \simeq \frac{1}{\sqrt{N_{knee}F}} .$$

IV. SMALL PEAK MODELS

In this section we shall consider the multiple measurement schemes which sample the phase shift Φ with states very near the vacuum state. Thus, we consider states such as

$$|\psi\rangle = A\left(|0\rangle + r|\psi_1\rangle\right) ,$$

with $r \ll 1$ and where $|\psi_1\rangle$ is orthogonal to the vacuum and normalized to unit probability; the normalization is $A = 1 + O(r^2)$ and the mean photon number is

$$\langle\hat{n}\rangle = r^2\langle\hat{n}\rangle_1[1+O(r^2)],$$

where the expectations $\langle\,\rangle_1$ are taken over the state $|\psi_1\rangle$. It is now easy to calculate the Fisher information and knee location to lowest order in r

$$F \simeq 2r^2\int \frac{d\phi}{2\pi}\left|\psi_1'(\phi)\right|^2 = 2r^2\langle\hat{n}^2\rangle_1,$$

$$N_{knee} \simeq \frac{4r^2}{F^2}\int \frac{d\phi}{2\pi}\left|\psi_1''(\phi)\right|^2 = \frac{\langle\hat{n}^4\rangle_1}{r^2(\langle\hat{n}^2\rangle_1)^2} .$$

These results tell us that the net cost is

$$N_{tot} \simeq \frac{\langle\hat{n}\rangle_1\langle\hat{n}^4\rangle_1}{(\langle\hat{n}^2\rangle_1)^2} \tag{10}$$

and that the phase sensitivity calculated at the optimal split in resources will be

$$(\Delta\Phi)^2 \simeq \frac{\langle\hat{n}^2\rangle_1}{2\langle\hat{n}^4\rangle_1} . \tag{11}$$

This is a surprising result as it does not depend on the phases in $|\psi_1\rangle$ at all. What if the phases do *not* conspire so as to give a single peak to the Susskind-Glogower distribution? Has the analysis broken down? No, the analysis is fine, however, we should recognize that such states will not be directly useful for our multiple measurement scheme where all measurements are the same. If a state has only fine structure in its sampling distribution, but no broad structure (i.e., no single peak), then the initial convergence of our data analysis will not work. Nonetheless, once we obtain enough information (from some other source) to locate the phase shift of interest to within the structure available, the data analysis will converge at the rates predicted above.

Given such simple expressions for the sensitivity and the cost we consider the question of optimizing our sensitivity over all choices of fiducial states in the limit that $\langle\hat{n}\rangle \to 0$. We seek the global minimum of the sensitivity functional

$$S[|\psi_1\rangle] \equiv (\Delta\Phi)^2 + \lambda\left(\frac{\langle\hat{n}\rangle_1\langle\hat{n}^4\rangle_1}{(\langle\hat{n}^2\rangle_1)^2} - N_{\rm tot}\right),$$

where the expression for $(\Delta\Phi)^2$ is given in Eq. (11) and λ is a Lagrange multiplier ensuring the mean total photon number constraint of Eq. (10). We incorporate the normalization of $|\psi_1\rangle$ explicitly by considering states of the form

$$|\psi_1\rangle_k = \sqrt{1 - \sum_{\substack{\ell=1\\ \ell\neq k}}^{\infty} a_\ell^2}\,|k\rangle + \sum_{\substack{\ell=1\\ \ell\neq k}}^{\infty} a_\ell|\ell\rangle\ ,$$

where the square root is sufficient without a sign since we have already noted that we may ignore the phases. Defining $m_p \equiv \langle\hat{n}^p\rangle_1$ we find that its derivative for $\ell \neq k$ is given by

$$\frac{\partial m_p}{\partial a_\ell} = 2a_\ell\left(\ell^p - m^p\right)\ ,$$

and from this we may write the derivative of the sensitivity functional with respect to a_ℓ (for $\ell \neq k$) in a different form for each k:

$$\frac{\partial S_k}{\partial a_\ell} = 2a_\ell\left[\frac{\ell^2 - k^2}{2m_4} - \frac{(\ell^4 - k^4)m_1}{2m_4^2} + \lambda\left(\frac{(\ell - k)m_4}{m_2^2} + \frac{(\ell^4 - k^4)m_1}{m_2^2} - \frac{2(\ell^2 - k^2)m_1m_4}{m_2^3}\right)\right]\ .$$

Let us fix k; then by setting this derivative to zero we obtain a quartic equation which is satisfied for all ℓ for which $a_\ell \neq 0$ so there are at most four Fock states contributing to the optimal state. Furthermore, since this quartic equation has no cubic term these four roots sum to zero so at most three ℓ will be positive. Finally, since $\ell = k$ is clearly one of these roots we are left with the most general solution for $|\psi_1\rangle_k$ as being

$$|\psi_1\rangle_k = \alpha_1|n_1\rangle + \alpha_2|n_2\rangle + \alpha_3|k\rangle\ ,$$

which if $\alpha_3 \neq 0$ then there would be an infinite number of solutions, one for each k. If instead, we suppose that there are at most a finite number of solutions then the most general form will become

$$|\psi_1\rangle = \alpha_1|n_1\rangle + \alpha_2|n_2\rangle\ .$$

Substituting this into Eq. (10) yields a complicated quadratic equation in the coefficients which we will not try to solve here. We do, however, note one special solution with

$$|\psi_1\rangle = |N_{\rm tot}\rangle\ ,$$

which has a sensitivity of

$$\Delta\Phi \simeq \frac{1}{\sqrt{2N_{\rm tot}}}\ .$$

As noted above, this sensitivity is not directly achievable in this case because the corresponding Susskind-Glogower distribution has no central peak, only fine structure. Nonetheless, this solution is a local minimum to the sensitivity achievable.

Instead of investigating the question of global optimization further, we will consider the two examples characterized by the fiducial states in Eqs. (4) and (5) and calculate their sensitivity in terms of the constraint $N_{\rm tot}$.

V. TWO SMALL PEAK MODELS

A. Shapiro Shepard model

The Shapiro and Shepard [5] fiducial state in the number state basis is given by

$$|\psi\rangle = A\sum_{m=0}^{M}\frac{1}{m + r}|m\rangle\ ,$$

where M and r are parameters and A is the normalization which to lowest orders in r is

$$A^2 = r^2\left[1 + O(r^2)\right]\ .$$

The mean photon number and photon uncertainty may also be expanded in this way yielding

$$\langle\hat{n}\rangle = r^2\left[\sum_{m=1}^{M} 1/m - 2r\sum_{m=1}^{M} 1/m^2 + O(r^2)\right],$$

$$(\Delta n)^2 = \langle\hat{n}^2\rangle + O(r^4) = r^2\left[M - 2r\sum_{m=1}^{M} 1/m + O(r^2)\right]\ ,$$

and the first few derivatives of the Susskind-Glogower phase distribution for this state are

$$p(\phi|\Phi) = \frac{A^2}{2\pi r^2}\left[1 + r\left(S_1 + S_1^*\right) + r^2\left(|S_1|^2 - S_2 - S_2^*\right) + O(r^3)\right],$$

$$p'(\phi|\Phi) = \frac{irA^2}{2\pi r^2}\left[S_0 - S_0^* + r\left(S_0S_1^* - S_1S_0^* - S_1 + S_1^*\right) + O(r^2)\right],$$

$$p''(\phi|\Phi) = \frac{-rA^2}{2\pi r^2}\left[S_{-1} + S_{-1}^* + r\left(S_{-1}S_1^* + S_1S_{-1}^* - 2|S_0|^2 - S_0 - S_0^*\right) + O(r^2)\right]\ ,$$

with derivatives over the parameter Φ and where we have introduced the symbol

$$S_k \equiv \sum_{m=1}^{M} \frac{e^{im(\Phi-\phi)}}{m^k} .$$

Using the sums in Appendix A we calculate the Fisher information from Eq. (9) about $r = 0$ giving

$$\begin{aligned} F &= 2r^2M\left[1 - r\left(1 + \frac{1}{M}\sum_{m=1}^{M}\frac{1}{m}\right) + O(r^2)\right] \\ &= 2r^2M\left[1 - r\left(1 + \frac{\gamma}{M} + \frac{\ln M}{M} + O(1/M^2)\right) + O(r^2)\right] . \end{aligned}$$

[We note that this result is consistent with the theorem $F \leq 4(\Delta n)^2$ [11].] Thus, so long as $r \ll 1$ the Fisher information can be well approximated by $F = 2r^2M$. As qualitatively predicted by Hradil and Shapiro [7,8] the Fisher information vanishes as $r \to 0$ so our ability to discover the peak location requires ever larger sample sizes. To see what efficiencies are actually attainable we now calculate the location of the knee in ML estimation.

The location of the knee in the ML estimation may be calculated from Eq. (8). The integrals involved are expanded around $r = 0$ just as was done for the Fisher information; the identities in Appendix A allow us to calculate

$$\begin{aligned} \int d\phi \frac{p''^2}{p} &= \frac{1}{3}r^2M(M+1)(2M+1)\left(1 - \frac{r}{6}\frac{22M^2+33M+17}{(M+1)(2M+1)} + O(r^2)\right), \\ \int d\phi \frac{p'^4}{p^3} &= \frac{1}{3}r^4M(2M^2+1)\left[1 + O(r)\right] , \end{aligned}$$

and so the knee location becomes

$$N_{\text{knee}} = \frac{(M+1)(2M+1)}{2r^2M}\left[1 + r\left(\frac{1}{6} + \frac{2}{M}\sum_{m=1}^{M}\frac{1}{m} - \frac{1}{(M+1)(2M+1)}\right) + O(r^2)\right] .$$

The sensitivity at the knee, from Eq. (11), is then given by

$$\Delta\Phi \simeq \frac{1}{\sqrt{(M+1)(2M+1)}}\left[1 + \frac{r}{2}\left(\frac{5}{6} - \frac{1}{M}\sum_{m=1}^{M}\frac{1}{m} + \frac{1}{(M+1)(2M+1)}\right) + O(r^2)\right] .$$

Note that this sensitivity is independent of r to lowest order; this implies that even though the Susskind-Glogower phase distribution approaches a flat probability distribution for $r \to 0$, multiple measurement schemes yield undiminished sensitivity for a mean photon-number constraint. Mathematically, this means that the limiting process is not *uniform*, i.e., that the limiting sensitivity is not the same as the sensitivity of the limiting distribution.

The last step is to translate this multiple measurement sensitivity at the knee, requiring N_{knee} measurements, into our constraint for N_{tot}. Assuming that both $M \gg 1$ and $r \ll 1$ we find

$$\begin{aligned} N_{\text{tot}} &= \langle\hat{n}\rangle N_{\text{knee}} \\ &\simeq M\left(\ln M + \gamma + \frac{3}{2}\frac{\ln M}{M} + O(1/M) + O(r)\right) , \end{aligned}$$

with a sensitivity of $\Delta\Phi \simeq 1/\sqrt{2}M$ corresponding to

$$\Delta\Phi \simeq \frac{1}{\sqrt{2}N_{\text{tot}}}(\text{logarithmic corrections}) .$$

B. Modified Dowling model

We consider now our modification of Dowling's model [6], which employs a smooth cutoff in place of the sharp SSW cutoff. Our modification consists of adding the variable r so that the fiducial state takes the form

$$|\psi\rangle = A\sum_{m=0}^{\infty}\frac{e^{-m/\mu}}{m+r}|m\rangle ,$$

with a normalization given by $A^2 = r^2\left[1 + O(r^2)\right]$. The mean photon number and variance are expanded around $r = 0$ yielding

$$\begin{aligned} \langle\hat{n}\rangle &= r^2\left[\sum_{m=1}^{\infty}\frac{e^{-2m/\mu}}{m} - 2r\sum_{m=1}^{\infty}\frac{e^{-2m/\mu}}{m^2} + O(r^2)\right], \\ (\Delta n)^2 &= \langle\hat{n}^2\rangle + O(r^4) \\ &= r^2\left[\frac{1}{e^{2/\mu}-1} - 2r\sum_{m=1}^{\infty}\frac{e^{-2m/\mu}}{m} + O(r^2)\right] . \end{aligned}$$

The Susskind-Glogower probability distribution and its derivatives are identical to those of the preceding subsection when we replace the sums S_k with those of D_k given in Appendix B. Following calculations similar to those in Sec. V A and assuming that both $\mu \gg 1$ and $r \ll 1$ we find that the overall phase sensitivity for this fiducial state is

$$\Delta\Phi \simeq \frac{1}{\sqrt{2}N_{\rm tot}}(\text{logarithmic corrections})\;;$$

so replacing the sharp cutoff with a smooth one has not diminished its convergence.

VI. CONCLUSION

While neglecting noise and losses entirely we have studied one small corner of the problem of determining the ultimate quantum limit to precision phase measurement. We considered strategies for determining the c-number phase shift Φ constrained by a mean total number of available photons $N_{\rm tot}$. One major difficulty in dealing with this constraint is that the resources may be split among many measurements followed by data analysis to extract Φ; this introduces extra freedom which must be optimized over. To date, only so-called symmetric schemes have been studied for which N identical measurements are made on N identically prepared quantum states each with mean photon number $\langle\hat{n}\rangle$. In this case our constraint reads

$$N_{\rm tot} = \langle\hat{n}\rangle N\;.$$

We studied such strategies in the limiting case where $\langle\hat{n}\rangle \to 0$. In this limit we calculated the phase sensitivity achievable, for maximum likelihood analysis of the measurements, and set up its optimization over all quantum states. The form of the state yielding a global optimum was derived and a local minimum to the sensitivity was found to be

$$\Delta\Phi \simeq \frac{1}{\sqrt{2}N_{\rm tot}}\;;$$

if this local minimum also turns out to be the global minimum then this sensitivity would correspond to the quantum limit to phase measurement in these near-vacuum schemes. Finally, we studied two near-vacuum schemes in detail which achieve this sensitivity up to logarithmic corrections.

These results are surprising since states with $\langle\hat{n}\rangle \to 0$ all have a vanishing Fisher information and so, as qualitatively predicted by Hradil and Shapiro [7,8], the information obtainable per state becomes vanishingly small implying a terrible overall sensitivity. Nonetheless, since a large number of states may be sampled, while still satisfying $N_{\rm tot} = \langle\hat{n}\rangle N$, we find that the net information gained can be substantial and that a sensitivity comparable with squeezed state interferometry is achievable.

APPENDIX A: SUMMATION IDENTITIES: THE SHAPIRO-SHEPARD MODEL

Recall the definition

$$S_k \equiv \sum_{m=1}^{M} \frac{e^{im(\Phi-\phi)}}{m^k}\,.$$

To calculate the integrals for the Fisher information $\int d\phi\, p'^2/p$ we need the sums

$$\int |S_0|^2 = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} \sum_{m,m'=1}^{M} e^{i(m-m')(\Phi-\phi)} = M,$$

$$\int S_0 S_1^* = \sum_{m=1}^{M} \frac{1}{m},$$

$$\int (S_0)^2 S_1^* = M - \sum_{m=1}^{M} \frac{1}{m}\,,$$

with

$$\sum_{m=1}^{M} \frac{1}{m} = \gamma + \ln M + \frac{1}{2M} + O\left(\frac{1}{M^2}\right)\,,$$

where $\gamma = 0.577\ldots$ is Euler's constant. Similarly, for the integral $\int d\phi\, p''^2/p$ we need the sums

$$\int |S_{-1}|^2 = \frac{1}{6}(2M^3 + 3M^2 + M),$$

$$\int S_{-1} S_0^* = \frac{1}{2}(M^2 + M),$$

$$\int (S_{-1})^2 S_1^* = \frac{1}{36}(2M^3 + 3M^2 - 5M),$$

$$\int S_{-1}|S_0|^2 = \frac{1}{6}(M^3 - M)\,.$$

Finally, the integral involved in calculating $\int d\phi\, p'^4/p^3$ is

$$\int |S_0|^4 = \frac{1}{3}M(2M^2+1)\,.$$

APPENDIX B: SUMMATION IDENTITIES: THE MODIFIED DOWLING MODEL

Instead of S_k we make use of a sum with a smooth cutoff

$$D_k \equiv \sum_{m=1}^{\infty} \frac{e^{[i(\Phi-\phi)-1/\mu]m}}{m^k}.$$

To calculate the integrals for the Fisher information $\int d\phi\, p'^2/p$ we need the sums

$$\int |D_0|^2 = \sum_{m=1}^{\infty} e^{-2m/\mu} = \frac{1}{e^{2/\mu}-1},$$

$$\int D_0 D_1^* = \sum_{m=1}^{\infty} \frac{e^{-2m/\mu}}{m} = -\ln(1-e^{-2/\mu}),$$

$$\int (D_0)^2 D_1^* = \sum_{m=1}^{\infty} e^{-2m/\mu} - \sum_{m=1}^{\infty} \frac{e^{-2m/\mu}}{m}.$$

Similarly, for the integral $\int d\phi\, p''^2/p$ we need the sums

$$\int |D_{-1}|^2 = \sum_{m=1}^{\infty} m^2 e^{-2m/\mu} = \frac{2e^{4/\mu}}{\left(e^{2/\mu}-1\right)^3} - \frac{e^{2/\mu}}{\left(e^{2/\mu}-1\right)^2},$$

$$\int D_{-1} D_0^* = \sum_{m=1}^{\infty} m\, e^{-2m/\mu} = \frac{e^{2/\mu}}{\left(e^{2/\mu}-1\right)^2},$$

$$\int (D_{-1})^2 D_1^* = \frac{1}{6}\sum_{m=1}^{\infty} m^2 e^{-2m/\mu} - \frac{1}{6}\sum_{m=1}^{\infty} e^{-2m/\mu},$$

$$\int D_{-1}|D_0|^2 = \frac{1}{2}\sum_{m=1}^{\infty} m^2 e^{-2m/\mu} - \frac{1}{2}\sum_{m=1}^{\infty} m\, e^{-2m/\mu}.$$

Finally, the integral involved in calculating $\int d\phi\, p'^4/p^3$ is

$$\int |D_0|^4 = \sum_{m=1}^{\infty} m^2 e^{-2m/\mu} - 2\sum_{m=1}^{\infty} m\, e^{-2m/\mu} + \sum_{m=1}^{\infty} e^{-2m/\mu}.$$

[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. Xiao, L.-A. Wu, and H. J. Kimble, Phys. Rev. Lett. **59**, 278 (1987).
[4] J. H. Shapiro, S. R. Shepard, and N. W. Wong, Phys. Rev. Lett. **62**, 2377 (1989).
[5] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (1991).
[6] J. D. Dowling, Opt. Commun. **86**, 119 (1991).
[7] Z. Hradil and J. H. Shapiro, Quantum Opt. **4**, 31 (1992).
[8] Z. Hradil, Phys. Rev. A **46**, R2217 (1992).
[9] S. L. Braunstein, A. S. Lane, and C. M. Caves, Phys. Rev. Lett. **69**, 2153 (1992).
[10] S. L. Braunstein, Phys. Rev. Lett. **69**, 3598 (1992).
[11] A. S. Lane, S. L. Braunstein, and C. M. Caves, Phys. Rev. A **47**, 1667 (1993).
[12] Currently we have no idea how to implement this kind of measurement in the laboratory; but presumably such an implementation would contain some phase reference for which we must pay from our net resources.
[13] L. Susskind and J. Glogower, Physics **1**, 49 (1964).
[14] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946), p. 500.
[15] S. L. Braunstein, J. Phys. A **25**, 3813 (1992).