APS physics

CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Quantum phase representation of Heisenberg limits and a minimally resourced quantum phase estimator

Scott Roger Shepard, Frederick Ira Moxley, III, and Jonathan P. Dowling

Phys. Rev. A **93**, 033805 — Published 2 March 2016

DOI: 10.1103/PhysRevA.93.033805

A Quantum Phase Representation of Heisenberg Limits and a Minimally Resourced Quantum Phase Estimator

Scott Roger Shepard[1], Frederick Ira Moxley III[2], and Jonathan P. Dowling[2,3]
[1] *Department of Electrical Engineering and Computer Science, College of Engineering & Science Louisiana Tech University, Ruston, LA 71272, USA*
[2] *Hearne Institute for Theoretical Physics, Department of Physics & Astronomy Louisiana State University, Baton Rouge, LA 70803, USA*
[3] *Beijing Computational Science Research Center, Beijing 100084, China*

Within the quantum phase representation we derive Heisenberg limits, in closed form, for N00N states and two other classes of states that can perform better in terms of local performance metrics relevant for multiply-peaked distributions. One of these can also enhance the super-resolution factor beyond that of a N00N state of the same power, at the expense of diminished fringe visibility. An accurate phase estimation algorithm, which can be applied to the minimally resourced apparatus of a standard interferometer, is shown to be resilient to the presence of additive white-Gaussian noise (AWGN). In the limit of no AWGN the algorithm achieves over nine digits of accuracy for the case of a four-photon N00N state — orders of magnitude below its Heisenberg limit.

PACS numbers: 42.50.St, 42.50.Dv, 42.50.Ex, 42.50.Lc

I. INTRODUCTION

Quantum phase estimation plays an important role in quantum computing and quantum sensing applications. With regard to quantum computing, phase estimation is often implemented via the inverse of the Quantum Discrete Fourier Transform (QDFT). Phase estimation via QDFT is an integral step in many quantum algorithms, including the assessment of periodicity in Shor's algorithm, etc. [1, 2]. Quantum computing is typically realized as quantum circuits implementing algorithms on individual qubits [3], however, these can also be realized as quantum multiparticle interferometers [4, 5] along the lines of Feynman's original ideas [6]. A mapping between quantum circuits, interferometers, and spectrometers [7], coupled with linear optics realizations [8], leads naturally to a quantum sensing perspective that is realizable via beam splitters, phase shifters and photodetectors.

Therein, in addition to the acheivement of phase measurement accuracies below the shot-noise limit [9], the quantum interference effects in electromagnetic fields have led to super-resolving phase measurements [10, 11] (a.k.a. super-resolution), which also circumvents the Rayleigh diffraction limit in lithography [12] and imaging [13]. In essence, N photons of a field at wavelength λ are utilized to perform "quantum sensing" with an effective wavelength of λ/N, while still utilizing sources, detectors and propagation properties associated with an actual wavelength λ. In the standard methods, the improvement factor N has been limited to date [14, 15, 16]. This limitation arose primarily because in order to observe these higher-order fringes, the standard schemes relied on coincidence detection methods. Thus, the measurement apparatus increases in complexity with N. It was thought that coincidence detection schemes of order N were required since the super-resolving fringes vanish in the output of a first-order interferometer. Fortunately, however, a method of extracting this higher-order phase information from a standard [17] interferometer has been found [18]. In this method (termed the phase function fitting algorithm, PFFA) the apparatus complexity is independent of the super-resolving improvement factor N. Herein, we examine the robustness of the PFFA to additive white-Gaussian noise (AWGN) — after first gleaning insights, and deriving N00N (and other state) limits within the quantum phase [19], [20] representation.

The quantum phase representation that is complementary to the measurement of the difference of photon numbers between two harmonic oscillators is useful for visualizing and calculating the phase information associated with a quantum state — although its apparatus has yet to be realized. Herein it is used to derive Heisenberg limits (without bounds or approximation) for three classes of states; and to show how an entanglement of N00N states with the vacuum state can surpass the N00N state in terms of two local performance measures appropriate for multiple-peaked distributions. Both measures, the HWHM and the square-root of the bin-variance, scale as $\sim 1/N$ (where N is the average photon number) for all three classes of states — which are Heisenberg limited in that sense. They differ however in the coefficient which multiplies $1/N$ and one of these can go to zero, but only at the expense of reduced fringe visibility.

The simplest phase estimation apparatus that we know how to realize is a standard quantum interferometer. Therein quantum resources are diminished by exploiting the multiphoton interferences inherent within the probability amplitudes of the quantum electromagnetic field itself. The quantum theory of an interferometer is based on the observation by Yurke et al. [21] that it is mathematically isomorphic to rotating a quantum state by an unknown angle and estimating that angle from the projection of the rotated state onto the z-component angular momentum eigenkets. This stems from Schwinger's observation that the algebra of two uncoupled harmonic oscillators can reproduce the algebra

of angular momentum [22]. In terms of eigenvalues we have $m = (n_u - n_d)/2$ for the eigenvalue associated with $\hat{J}_z$ and $j = (n_u + n_d)/2$ for the eigenvalue associated with $\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$, where n_u and n_d are the photon number eigenvalues of the two oscillators. The interferometers statistics are $P_m = |\Psi_m(\Phi)|^2$ which is a probability distribution in discrete m space, under relative interferometer-arm phase shift Φ and the underlying wavefunctions are $\Psi_m(\Phi) \equiv \langle m|\hat{D}_x(\Phi)|\psi\rangle$ where $\hat{D}_x(\Phi)$ is the analogous rotation about the x-axis by Φ, the unknown signal we wish to estimate.

The general theory of quantum angle measurement (complementary to the measurement of a single component of angular momentum) is described in [20], [23] but for states of unique j for all m (as for all states considered herein) the angle or relative-phase measurement, taken at a "snapshot" in absolute time, reduces to what one would naturally expect for complementary quantities: a Fourier transform between wavefunctions. Although this can only be formally justified on a product space of two oscillators — with dimension large enough to give a complete description of the quantum measurement (in terms of sets of commuting observables or equivalently: a collapsible wavefunction) for the states considered herein we can simply take a Fourier transform of the complementary wavefunction $\psi_{j,m} \equiv \langle j, m|\psi\rangle$, the magnitude square of which yeilds the quantum phase distribution, viz.,

$$\psi(\phi) = \sum_j \sum_{m=-j}^{j} \psi_{j,m} e^{im\phi} \text{ and } P(\phi) = \frac{|\psi(\phi)|^2}{2\pi}. \quad (1)$$

II. QUANTUM PHASE REPRESENTATION OF THREE CLASSES OF STATES

Although its apparatus has yet to be realized, the phase (or angle) representation is useful for visualizing and calculating the phase information associated with a quantum state. Herein we use it to derive Heisenberg limits (without bounds or approximation) for three classes of states — the latter two of which can be written in the following form (times a normalization constant)

$$|\psi\rangle = r_2(|2j_{\max}\rangle_u|0\rangle_d + |0\rangle_u|2j_{\max}\rangle_d) + r_1(|j_{\max}\rangle_u|0\rangle_d + |0\rangle_u|j_{\max}\rangle_d) + |0\rangle_u|0\rangle_d.$$

The first of the three classes considered are the N00N states; the second class (termed sub-states) are an entanglement of a N00N state of $j = j_{\max}/2$, with an equally likely superposition of a N00N state of $j = j_{\max}$ and the vacuum state ($r_2 = 1/\sqrt{2}$); and the third class (termed N00N-vac states) are comprised of an entanglement of a N00N state with the vacuum state ($r_1 = 0$). In all cases the cost function is the expected number of photons used, $N = 2\langle j\rangle$, and our metrics are local performance measures appropriate for multiple-peaked distributions (half-width-half-max, and bin-variance) as well as fringe visibility and other aspects of the probability distribution function (PDF) of the quantum phase measurement $P(\phi)$, including the number of peaks and their height.

The Fourier transform of the number-ket expansion coefficients of the N00N state readily yields $P(\phi) = (1/\pi)\,\mathrm{Cos}^2(j_{\max}\phi)$ which has $N = 2j_{\max}$ peaks (of height $1/\pi$) separated by perfect nulls (hence the fringe visibility $V \equiv (\max - \min)/(\max + \min)$ is always unity for N00N states in the phase representation, independent of $j_{\max}$) and we observe the super-resolving aspect of obtaining N identical peaks or fringes, in contrast to a single-peaked PDF which would arise in the case of a coherent state. As a consequence of the Fourier transform the periodicity (here the number of peaks, which is also the number of identical bins) is set by the minimal separation of values of m for which we have non-zero number-ket expansion coefficients (one for the coherent state, $2j_{\max}$ for the N00N state). The variance on a 2π interval of this multiple-peaked PDF is clearly not a useful performance measure, so we consider bin-variance: defined to be the variance of the PDF over one of these identical bins, renormalized to the bin width (for N00N states that is a $2\pi/N$ interval) and centered on that bin (to avoid branch-cut effects). For example, in the case of a N00N state, the bin-variance would be

$$\int_{-\pi/N}^{\pi/N} d\phi\ (N/\pi)\ \phi^2\ \mathrm{Cos}^2(N\phi/2). \quad (2)$$

Inherent to the utility of this metric (and to the use of super-resolution in general) is the assumption that we can correctly assess the bin to which any particular estimate corresponds — otherwise we make a bin error (the probability of which will be influenced by fringe visibility and how one tracks a dynamically varying unknown phase). We similarly consider the local (defined on one bin) half-width at half-max, HWHM, and obtain the N00N state results

$$\mathrm{HWHM} = \frac{\pi}{2\mathrm{N}}, \quad \text{bin-variance} = \frac{\pi^2/3 - 2}{\mathrm{N}^2}. \quad (3)$$

The square-root of the bin-variance and HWHM both scale as $1/N$ in terms of our cost function. Indeed all three classes of states considered herein follow a $1/N$ scaling but the difference in the coefficient multiplying $1/N$ varies widely and can even go to zero at the cost of diminished fringe visibility. Other differences, gleaned from the PDFs, will impact upon the probability of bin error in a practical system; but the fact that the other two classes of states can yield coefficients better than those in equation (3) indicates that state optimization remains a field open for investigation (e.g., the N00N state is not optimal, even in the above two metrics). The fact that the N00N states are also not optimal for interferometry is well known [24, 25] and the relationshp between the quantum phase measurement and the SU(2) interferometer statistics is delineated in [18] and [26].

Within the second class of states the strategy is to utilize a sub-harmonic to, in effect, delete alternate bins and

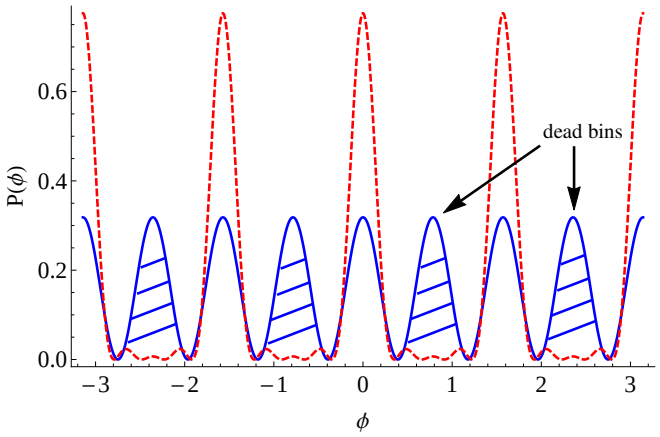

FIG. 1. (color online). N00N (blue solid) and $r_1 = 1$ sub-state (red dashed) Quantum Phase PDFs for N = 8.

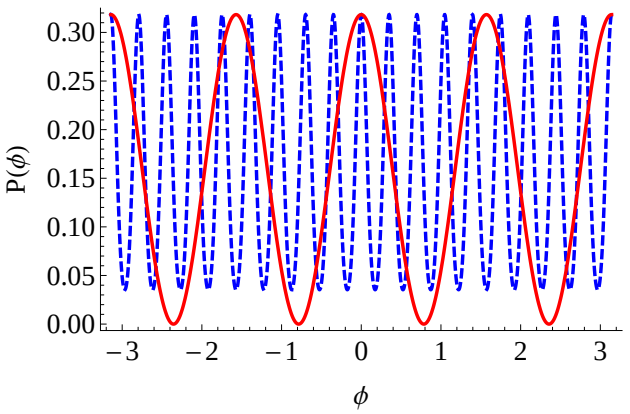

FIG. 2. (color online). N00N (red solid) and N00N-vac state (blue dashed) Quantum Phase PDFs for $N = 4$ (and $n = 8$).

sharpen the PDF within the remaining bins. Details of a bin dropping protocol in [27] show that the probability of a dropped bin achieves its minimal value of less than 3.164% at $r_1 = 1$. Fig. 1 compares the quantum phase distribution of an $N = 8$ sub-state ($r_1 = 1$, $j_{\max} = 8$) to that of a N00N state of $N = 8$ ($j_{\max} = 4$). The bin variance on the retained bin and the HWHM [27, 28] are readily derived in the phase representation and are always smaller than the N00N state limit of (1). Note in passing that the enhanced peaks of the sub-states also permit more rapid acquisition of useful phase estimates during the collection of the histograms which evolve into these PDFs (so these higher peaks can be useful for tracking a dynamic unknown phase).

The third class of states considered herein have $r_1 = 0$, and r_2 is a free parameter. These are entanglements of a N00N state (of $j = 2j_{\max} \to j_{\max}$) with the vacuum state, termed N00N-vac states. To clarify the physics we let $r_2 = 1/\sqrt{2n}$ and initially consider n to be an element of the set of integers. For $n = 1$ the N00N-vac state is equivalent to a sub-state of $r_1 = 0$. The number of bins for this class of states is $j_{\max}$ (the minimal distance in m of the N00N state component is $2j_{\max}$; but that distance is reduced to $j_{\max}$ via inclusion of the vacuum state) and there is no motivation for any bin dropping protocol.

The strategy for the N00N-vac states is to make $j_{\max}$ large (to reap the benefits of super-resolution) but at the same time reduce the probability of *actually* being found in the N00N state component, by also increasing n, since then we can simultaneously constrain $N \equiv 2\langle j \rangle$, which is $2\langle j \rangle = 2j_{\max}/(n+1)$ for these states. We find that the probability amplitude of the vacuum state can strongly interfere with the probability amplitude of the N00N state component even when the probability of actually being in that component is greatly diminished; and that the only fundamental tradeoff in this scheme is a slowly diminished fringe visibility as $V = (2\sqrt{2}\sqrt{n})/(2+n)$ [29]. In designing states for the algorithm of the next section the peaks are more important than the nulls. It can be shown that as we increase n the peaks of the N00N-vac PDFs come down and become equal to those of the N00N states when $n = 8$ (independent of $j_{\max}$) as illustrated for $N = 4$ in Fig. 2. As we continue to increase n the super-resolving enhancement aspect (i.e., the number of fringes or bins at a fixed N) continues to increase while the visibility diminishes, until at $n = 2(17+12\sqrt{2}) \approx 67.9411$ the min = max/2 and beyond this value of n the HWHM metric has no meaning. For n less than this value, the formula for HWHM [30] is better than 17.87 times smaller than the N00N state limit at $n = 67$. Similarly, the coefficient (multiplying N^{-2}) in the formula for the bin-variance of the N00N-vac states:

$$\frac{2(3 - 24\sqrt{2}\sqrt{n} + 2\pi^2 + 2n\pi^2)}{3(n+1)^3}, \tag{4}$$

is smaller than the N00N state limit (and is monotonically decreasing in n) for all $n \geq 1$. Bin-variance maintains its meaning, even when min > max/2, and its corresponding coefficient goes to zero as $n \to \infty$. With $V > 1/3$ at $n = 67$ the bin-variance coefficient is better than 569.9 times smaller than the N00N state limit.

III. THE PHASE FUNCTION FITTING ALGORITHM

The phase function fitting algorithm [18] consists of the following protocol. Under a given, but unknown, value of Φ one measures the interferometer's statistics and retains the $2j+1$ probabilities for each possible value of m: $P_m = |\Psi_m(\Phi)|^2$. To incorporate knowledge of the allowable quantum results one then calculates the interferometer's statistics for some dummy variable x: $f_m(x) = |\Psi_m(x)|^2$ (for the four-photon N00N state of our simulations herein, those are $2j + 1 = 5$ different functions). One then performs an LMS (Least Mean Square error) fit of these $2j+1$ functions to the measured $2j+1$ numbers to perform an optimal estimation of parameter x, thereby yielding our estimate of Φ.

In so doing, our simulations yield surprisingly good results: over 9 digits of phase accuracy, for a N00N state input of $j = 2$, when Φ ranges over a bin-width, here an interval of $\pi/4$ [31]. At first this might seem to be either: an impossibly good result (over nine digits of accuracy from only four photons would be orders of magnitude below the Heisenberg limit of 1/4); or a ridiculously trivial result (certainly similar extractions of information might be made from the ideally collected statistics of other measurements). There is no contradiction here, instead it teaches that:

1) *local performance measures are important characterizations of measurement statistics, but they do not uniquely identify what can be achieved when further processing (an algorithm) is applied;* and

2) *most significantly for phase estimation, it specifically demonstrates that the super-resolving information has not vanished from the measurement performed in a quantum interferometer.* These higher-order fringes only vanished in the standard (first-order) way of extracting information from the interferometer statistics. Since the higher-order information is uniquely extractable (over a non-trivial range) this information has not been destroyed in the measurement process. This significantly opens the door to the possibility of minimally resourced quantum interference sensors in which the quantum resources are dramatically reduced via classical signal processing that incorporates knowledge of the allowable quantum results.

This is not an impossible result, from 1) above, and in the static limit [18] there is always a sense in which an infinite number of photons is implied [32]; nor is it a trivial result because it cannot be applied to arbitrary quantum measurements. If for example, our apparatus consisted of a first-order interferometer (in which we subtract and average the two photodetector currents) the desired phase information vanishes entirely and no further processing can extract it. If instead we multiply those two currents (second-order coincidence detection) or do a parity measurement [33] then other limits are obtained. Note in passing that the PFFA does not require any presently unachievable amount of number resolving in the photodetectors. Indeed, in the four-photon ($j = 2$) example presented these need only discern between $0, 1, 2, 3$, or 4 counts (photodetection events within the observation time T).

The utility of any parameter estimation algorithm will lie in its ability to perform well under non-ideal conditions for the collection of measurement statics. To that end we herein test the algorithm's resilience to additive white-Gaussian noise (AWGN) in which five statistically independent AWGN processes are added to each of the the probabilities of the 5 different values of m, $\{+2, +1, 0, -1, -2\}$ for a N00N state input of $j = 2$ with Φ fixed. After each addition of the five independent noise samples, the PFFA is executed and its phase estimation error is calculated. Each dot in Fig. 3 represents the average error obtained from 40,000 such executions; and four such dots are presented for each value of noise power σ^2 to demonstrate the variation in the average error. The estimate is unbiased so there is some cancellation due to polarity in that average, which is useful in some applications. In other applications the absolute value of the error is more relevant and these are connected by the line in Fig. 3 (averaged over only 2,000 executions as these did not vary appreciably from run to run).

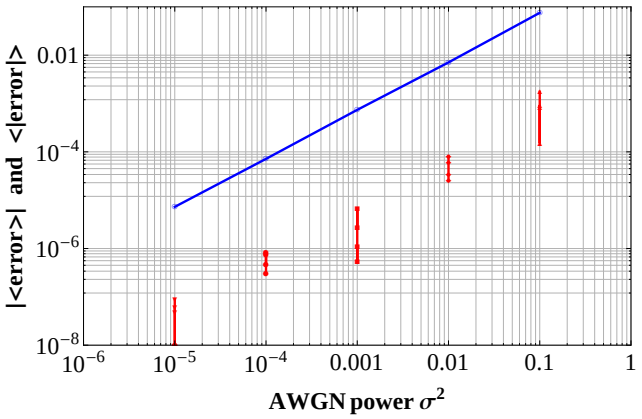

FIG. 3. (color online). Average PFFA error (dots) and average of the absolute value of the PFFA error (line); versus AWGN noise power.

IV. A COMPARISON TO BAYESIAN TECHNIQUES

It is important to note that the PFFA gives exceptional accuracy even in a single execution. Unlike a Bayesian protocol [34, 35] the PFFA does *not* improve its estimates in successive applications of the algorithm; and the 2,000 executions mentioned above were only for the sake of allowing the five statistically independent AWGN processes to combine in a sufficiently large number of ways so that we better test the algorithm's resilience to the presence of noise in its measurement statistics.

A Bayesian analysis [36] proceeds as follows: after an n-dimensional sample of data, m, one then updates a previous estimate of the PDF $P(\Phi)$; to $P(\Phi|m)$ in order to better estimate the parameter Φ via Bayes' theorem:

$$\begin{aligned} P(\Phi|m) &= P(\Phi)P(m|\Phi)/P(m) \\ &= P(\Phi)P(m|\Phi)/\int_{-\pi}^{\pi} d\Phi\ P(\Phi)P(m|\Phi) \end{aligned} \quad (5)$$

where, in our case, $P(m|\Phi) = \mathrm{P}_m(\Phi)$ are the interferometers' measurement statistics and the denominator ensures normalization of the updated PDF $P(\Phi|m)$. In contrast, when in the static limit (where Φ is fixed but unknown, as the above Bayesian strategy also assumes) in the normal application of the PFFA one would simply wait for

the histograms to sufficiently converge to $\mathrm{P}_m(\Phi)$ before executing the algorithm, once.

Nevertheless, as a step towards the more challenging problem of tracking a dynamic phase shift, it is of interest to compare the performance of the PFFA to that of a Bayesian protocol on a finite set of data collected within the static limit. For simplicity we will use a two-photon N00N state. The measurement statistics are therefore:

$$\mathrm{P}_{m=\pm 1/2}(\Phi) = (1/2)\ \mathrm{Cos}^2(\Phi);\ \mathrm{P}_{m=0}(\Phi) = \mathrm{Sin}^2(\Phi). \tag{6}$$

From successive applications of (5) and an apriori assumption of $P(\Phi) = 1/(2\pi)$; after n measurements (or one n-dimensional sample) the Bayesian would therefore obtain

$$P(\Phi|m) = c\ \mathrm{Cos}^{2(N_+)}(\Phi)\ \mathrm{Sin}^{2(N_0)}(\Phi)\ \mathrm{Cos}^{2(N_-)}(\Phi) \tag{7}$$

where N_+ is the number of counts for which $m = +1/2$; N_0 is the number of counts for which $m = 0$; N_- is the number of counts for which $m = -1/2$; and c is the normalization constant.

The rapidity with which trigonometric functions raised to a power, $p \sim O(10^2)$, can approach delta-functions is quite astounding — even more so when $p \sim O(10^3)$. On the other hand this can make the functions in (7) a bit cumbersome, such that in our simulations we had to abandon numerical integration of (7) as a means of obtaining c (which was $\sim O(10^{-246})$ for a run at $n = 1000$) in favor of integrating (7) in closed-form; resulting in hypergeometric functions (which accurately calculated $c \sim O(10^{-481})$ for a run at $n = 2000$). Likewise, although numerical schemes to obtain the peak of (7) sometimes failed to converge; we were able to accurately calculate $\langle\Phi\rangle$ numerically. Thus, $\langle\Phi\rangle$ will be used as the means for extracting an estimate from (7) and these are calculated on $[0, \pi/2]$ for a true value of $\Phi = \pi/3$ in our simulations (as any algorithm will have to know which bin it should use for this four-peaked N00N state).

The PFFA will use the three simple functions in (6) and perform an LMS fit of these to the same finite set of simulated counts via: $\mathrm{P}_{m=1/2} = N_+/n$; $\mathrm{P}_{m=0} = N_0/n$; and $\mathrm{P}_{m=-1/2} = N_-/n$. The "which bin" information is supplied to the PFFA in the form of an initial estimate of $\Phi = +1$, although any initial value on the open interval $(0, \pi/2)$ will converge to the same result.

Sufficient detail on the actual counts used will be provided so that the interested reader can readily verify our results. The simulation of any random process brings up some interesting issues of its own and means of doing so in a way in which one can control the spectrum (or equivalently the autocorrelation function) as well as the PDF have been suggested [18]. Herein, we will simply concern ourselves with the PDF and for a true value of $\Phi = \pi/3$ we have from (6) that we hope to generate sequences which will converge in mean to the measurement statistics of {1/8, 6/8, 1/8}. One way of generating the three counts is to process two Bernoulli trials; which we found to give the same conclusions as the method used for our presented results, which is to generate a discrete uniform distribution of eight integers and take six of these to represent the event $m = 0$; and the other two to represent the events $m = -1/2$ and $m = +1/2$.

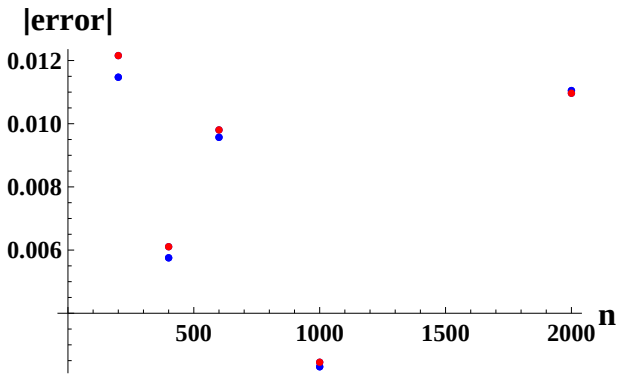

FIG. 4. (color online). Bayesian error (red upper dots) and PFFA error (blue lower dots) versus sample size n.

A discussion of the first two runs of only $n = 200$ samples each will illustrate some of the issues involved. The outcome of run 1 was $N_+ = 28$, $N_- = 24$ and $N_0 = 148$ which led to a Bayesian estimate error of $e_b \simeq .0121565$ and a PFFA error of $e_p \simeq .011472$. The outcome of run 2 however was $N_+ = 27$, $N_- = 23$ and $N_0 = 150$ which (by itself) would yield $e_b \simeq .000722269$ and $e_p \simeq 2.60236\ 10^{-13}$. Notice, in run 2 "we got lucky" in that $N_0/n = 6/8$ so that we hit one of our targeted measurement statistics exactly — this helped the Bayesian estimate but it helped the PFFA even more so. If we plotted the results of run 2 first in Fig. 4 it could very misleadingly appear that the PFFA holds an advantage for a small number of samples and this brings up some important issues. An experimentalist would never know that they "got lucky" and would continue to take more data in hopes that our estimates will improve in accuracy, in accordance with the Central Limit Theorem or the Weak Law of Large Numbers, etc., [36]. Moreover, run 2 did not occur first, it occurred second. Our experimentalist would report the estimates from run 1 for the first n=200 samples (the accuracies of which are as in Fig. 4) then adding the counts from run 2 (to those of run 1) we obtain $e_b \simeq .00610597$ and $e_p \simeq .00575451$; which correspond to the second set of points presented in Fig. 4 (for a total of $n = 400$ samples).

The accuracies improved because of run 2. Likewise, things can get worse, as they did by incorporating run 3 of $N_+ = 28$, $N_- = 25$ and $N_0 = 147$ for the points at $n = 600$; after which we increased the sample size to 400 in run 4 (improving the accuracies of both algorithms at $n = 1000$). In run 5 we increased the sample size to 1000, which changed the total counts to $N_+ = 130 + 113$, $N_- = 122 + 116$ and $N_0 = 748 + 771 = 1591$ (reducing the accuracies of both algorithms at $n = 2000$). Of course 2000 samples is not statistically large but at this point we can make some observations. We see that,

within the static limit, both algorithms can perform with similar accuracies for such small sample sizes. Secondly, even for states of relatively simple measurement statistics, as in (6), the burden on the classical signal processing aspects is increasing with sample size for the Bayesian analyst. In contrast, the PFFA operator could argue that if we are indeed in the static limit, then we can simply take as much time as is needed to acquire accurate measurement statistics since power is the relevant cost function (not energy). Thus, for future work, we suggest a comparison of the performance of the two methods in the tracking of a dynamic phase (but such an ambitious analysis is beyond the scope of the present paper).

V. CONCLUDING REMARKS

Within the quantum phase representation we have derived Heisenberg limits in closed form for three classes of states in terms of two local performance metrics: HWHM and bin-variance. All three classes of states are Heisenberg limited in that the HWHM and the square-root of the bin-variance scale as $1/N$ (where N is the expected number of photons) but the coefficients multiplying $1/N$ can vary. Two of these (termed sub-states and N00N-vac states) are demonstrated to perform better than the other class — the N00N states, in both metrics, as well as exhibit properties useful for tracking a dynamic phase. The N00N-vac states entangle the vacuum state with a N00N state, of relative probability equal to n. For fixed N, these enhance the super-resolution (number of fringes or bins) by a factor of $(n+1)/2$; and for large n they diminish the bin-variance as $1/n^2$ relative to the N00N state results at a diminished fringe visibility which scales only as $1/\sqrt{n}$. These higher-order (super-resolving) fringes vanish from the output signal of a first-order interferometer (in which one averages the difference of the two photodetector currents) so coincidence detection schemes of order N have been utilized. We discussed an algorithm which can extract this higher-order information from the apparatus of a standard interferometer (of complexity independent of N) by processing the information in a way that incorporates knowledge of the allowable quantum results. The algorithm provides over nine digits of phase estimation accuracy from a four-photon N00N state (over an unknown signal range of $\pi/4$) within a standard interferometer; and is shown herein to be fairly robust in the presence of additive white-Gaussian noise in the measurement statistics.

Acknowledgments

S.R.S. would like to acknowledge support from NASA. F.I.M. would like to acknowledge support from NASA EPSCoR & LaSPACE, Louisiana. J.P.D. would like to acknowledge support from the Air Force Office of Scientific Research and the Army Research Office.

[1] P.W. Shor, 35th Annual Symposium on Foundations of Computer Science, Santa Fe (1994).
[2] M. Wilde, Quantum Information Theory, Cambridge University Press (2013).
[3] C. G. Timpson, Quantum Information Theory and the Foundations of Quantum Mechanics, Oxford University Press (2013).
[4] R. Cleve, A. Ekert, C. Macchiavello, M. Mosca, Proc. R. Soc. Lond. A **454**, 339 (1998).
[5] A. Ekert, Physica Scripta **76**, 218 (1998).
[6] R. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[7] H. Lee, P. Kok, J.P. Dowling, J. Mod. Optic. **49**, 2325 (2002).
[8] E. Knill, R. Laflamme, G.J. Milburn, Nature **409**, 46 (2001).
[9] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[10] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).
[11] J. P. Dowling, Contemporary Physics **49**,125 (2008).
[12] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, J. P. Dowling, Phys. Rev. Lett. **85** , 2733 (2000).
[13] S. Guha and J. H. Shapiro, Quantum Communication, Measurement and Computing, AIP Conf. Proc. **1363**, 113 (2011).
[14] M. D'Angelo, M. V. Chekhova, and Y. Shih, Phys. Rev. Lett. **8701**, 013602 (2001).
[15] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Nature **13**, 161 (2004).
[16] P. Walther, J. W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni and A. Zeilinger, Nature **429**, 158 (2004).
[17] By "standard interferometer" we refer to the apparatus of a "first-order interferometer" in which the two photodetector currents are not necessarily subtracted. If they are subtracted then it is a "first-order interferometer;" if they are multiplied then it is a "second-order interferometer." If it is a higher-order interferometer then more beam splitters and more photodetectors are required in the apparatus than in the standard interferometer.
[18] S. R. Shepard, Nonlinear Analysis **71**, 1160 (2009).
[19] S. R. Shepard, Workshop on Squeezed States and Uncertainty Relations, NASA-CP-3135, 177 (1991).
[20] S. R. Shepard, arXiv:1410.0916 (2014).
[21] B. Yurke, S. L. Mc Call and J. R. Klauder, Phys. Rev. A **33**, 4033 (1986).
[22] J. Schwinger, U.S. Atomic Energy Report No. NYO-3071 U.S. GPO, Washington, D.C. (1952).
[23] S. R. Shepard, Workshop on Harmonic Oscillators, NASA-CP-3197, 109 (1992).
[24] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[25] G. Y. Xiang, H. F. Hofmann and G. J. Pryde, Nature Srep. **3** 2684 (2013).
[26] S. R. Shepard, arXiv:1511.09447 (2015).
[27] S. R. Shepard, F. I. Moxley and J. P. Dowling, arXiv:1403.2313 contains further information on the sub-states (2014).

[28] Sub-states have: (HWHM) (N) = $2\cos^{-1}[(4r_1 - 4\sqrt{6+4r_1+{r_1}^2})/(8\sqrt{2})]$ and (bin-variance) $(\mathrm{N}^2) = [128(27+13\sqrt{2})\mathrm{r}_1 + 144(3+\sqrt{2})\pi^2\mathrm{r}_1 + 36\pi^3(1+{\mathrm{r}_1}^2) - 27\pi(8\sqrt{2}-1+8{\mathrm{r}_1}^2)]/[36(4(3+\sqrt{2})\mathrm{r}_1 + 3\pi(1+{\mathrm{r}_1}^2))]$.

[29] For example, consider the ways in which we can make $\langle j\rangle = 2$ (i.e., an $N = 4$ photon state). A N00N state with $j_{\max} = 2$ would produce 4 bins; but since $1/4^{\text{th}}$ of 8 is also 2 we could (for the same $N = 4$ cost) produce twice as many bins with a N00N-vac state of $n = 3$ and $j_{\max} = 8$, in which case the peaks as well as the range (i.e., the value of max - min) for the N00N-vac state are larger than those of the N00N state (although $V < 1$ since the minimum no longer goes to zero). In the next section we will consider an algorithm which does not require the nulling of an interferometer — in such cases how close min comes to zero is of diminished importance. Moreover, we often are more concerned about the peaks of these PDFs because their enhancement permits more rapid acquisition of useful histograms (which is crucial while tracking a dynamic phase).

[30] N00N-vac states have: (HWHM) (N) = $[2/(n+1)]\cos^{-1}[1/2(\sqrt{2+2\sqrt{2n}+n}-\sqrt{2n})]$ and (bin-variance) $(\mathrm{N}^2) = 2(3-24\sqrt{2\mathrm{n}}+2\pi^2+2\mathrm{n}\pi^2)/[3(\mathrm{n}+1)^3]$.

[31] For Φ beyond a bin-width the algorithm can and sometimes does make a bin error (converging to what would be the right answer but displaced to the wrong branch of the periodic PDF) if we do not give it any apriori knowledge to avoid these.

[32] There is a sense in which an infinite number of photons have been used because we have extracted these virtually error-free estimates from the calculated (equivalent to ideally collected) quantum measurement statistics. In a laboratory environment it would take time (hence many photons) for the collected histograms to converge to the calculated PDFs. The unknown Φ cannot vary appreciably during this time for these static limit results to hold. This is the scenario within which most analyses are performed and it is how a fringe pattern is often measured. Also $N = 4$ in these simulations and power (i.e., N, not energy) is the relevant constraint. It is power that determines: the shot noise; the radiation pressure on a mirror; the cost of the laser; etc., so N is the relevant cost function. If we use more energy $E \sim NT$ (even towards an infinite number of photons) by collecting longer, that doesn't affect our cost and we can do so provided we can still make T small enough to track the dynamics. To go beyond this static limit one would also need to simulate the tracking of a dynamic signal, while acquiring and processing the histograms.

[33] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, J. P. Dowling, Phys. Rev. Lett. **104**, 103602 (2010).

[34] Z. Hradil, R. Myska, J. Perina, M. Zawisky, Y. Hasegawa and H. Rauch, Phys. Rev. Lett. **76**, 4295 (1996).

[35] L. Pezze, A. Smerzi, G. Khoury, J. F. Hodelin and D. Bouwmeester, Phys. Rev. Lett. **99**, 223602 (2007).

[36] A. W. Drake, Fundamentals of Applied Probability, McGraw Hill (1967).