Quantum states for Heisenberg-limited interferometry

H. Uys and P. Meystre
Department of Physics and College of Optical Sciences, The University of Arizona, Tucson, Arizona 85721, USA
(Received 15 February 2007; published 9 July 2007)

The phase sensitivity of interferometers is limited by the so-called Heisenberg limit, which states that the optimum phase sensitivity is inversely proportional to the number of interfering particles N, a $1/\sqrt{N}$ improvement over the standard quantum limit. We have used simulated annealing, a global optimization strategy, to systematically search for quantum interferometer input states that approach the Heisenberg-limited uncertainty in estimates of the interferometer phase shift. We compare the performance of these states to that of other nonclassical states already known to yield Heisenberg-limited uncertainty.

DOI: 10.1103/PhysRevA.76.013804 PACS number(s): 42.50.Dv, 06.20.−f, 89.70.+c

I. INTRODUCTION

An important aspect of quantum metrology is the engineering of quantum states with which to achieve measurements whose precision is Heisenberg limited. In this limit the measurement uncertainty is inversely proportional to the number of interfering particles N, representing a $1/\sqrt{N}$ improvement over the standard quantum limit. Squeezed light has long been employed to beat the shot-noise limit [1,2] and a growing body of theoretical literature indicates that the Heisenberg limit is in principle achievable using more exotic quantum states as interferometer inputs [3–8]. Several proof-of-principle experimental realizations of such states have recently been carried out [9–13]. Other proposals to beat the standard quantum limit involve the use of feedback schemes [14,15] or multimode interferometry [16]. The potential superiority of atomic fermions over bosons in some applications of atom interferometry with quantum-degenerate atomic gases has also been pointed out [17,18].

This paper summarizes the results of a systematic search for input quantum states that lead to Heisenberg-limited interferometric detection of phase shifts. Using the global optimization method of simulated annealing we demonstrate the existence of numerous possibilities over-and-above those already proposed in the literature, and we evaluate and compare their performance.

Section II discusses our theoretical model of a Mach-Zehnder interferometer used to measure the relative phase shift ϕ accumulated during the propagation of single-mode optical or matter waves along its two arms. Section III introduces a likelihood function used to estimate that phase and discusses its asymptotic form in the limit of many measurements. Section IV summarizes our main results obtained using simulated annealing and Sec. V focuses on the prospects for the experimental realization of a quantum state of particular interest. Finally, Sec. VI is a summary and conclusion.

II. MACH-ZEHNDER INTERFEROMETER

We consider a Mach-Zehnder interferometer with two input ports A and B, see Fig. 1, characterized by bosonic annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$ and $\hat{b}$ and $\hat{b}^\dagger$, respectively. We restrict our investigation to a system with fixed particle number N, in which case its properties are conveniently described in terms of the angular momentum operators [5]

$$\hat{J}_x = \frac{\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}}{2}, \tag{1}$$

$$\hat{J}_y = \frac{\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a}}{2i}, \tag{2}$$

$$\hat{J}_z = \frac{\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}}{2}, \tag{3}$$

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \tag{4}$$

which obey the familiar commutation relations $[\hat{J}_i, \hat{J}_j] = \varepsilon_{ijk} i \hat{J}_k$, where ε_{ijk} is the Levi-Civita symbol, and $[\hat{J}^2, \hat{J}_i] = 0$. Choosing z as the quantization axis we work in the basis of eigenstates common to $\hat{J}^2$ and $\hat{J}_z$,

$$|j,m\rangle_z \equiv |n_a\rangle |n_b\rangle. \tag{5}$$

Here $|n_i\rangle$ is a Fock state with n_i particles in arm $I = A, B$. For brevity we drop the subscript z henceforth. The eigenvalues corresponding to $\hat{J}^2$ and $\hat{J}_z$ are $j(j+1)$ and m, respectively, where

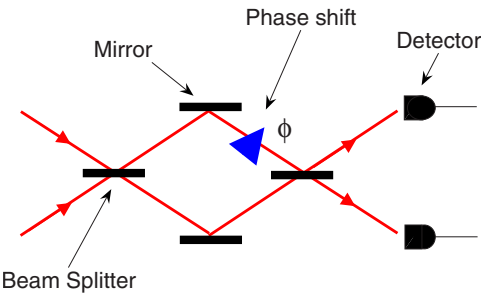

FIG. 1. (Color online) Schematic of a Mach-Zehnder interferometer, with relative phase shift ϕ resulting from propagation of an input field through its arms.

$$j = (n_a + n_b)/2 = N/2 \tag{6}$$

and

$$m = (n_a - n_b)/2. \tag{7}$$

In terms of these operators, the propagation of the input fields through the interferometer—which consists of three unitary transformations describing an input 50/50 beam splitter, the relative phase shift ϕ, and an output 50/50 beam splitter—reads [5]

$$|\psi_{\text{out}}\rangle = e^{-i(\pi/2)\hat{J}_x} e^{i\phi\hat{J}_z} e^{i(\pi/2)\hat{J}_x} |\psi_{\text{in}}\rangle \tag{8}$$

$$= e^{-i\phi\hat{J}_y} |\psi_{\text{in}}\rangle. \tag{9}$$

Our aim is to find input states

$$|\psi_{\text{in}}\rangle = \sum_{m=-N/2}^{N/2} \alpha_m |j,m\rangle \tag{10}$$

such that the uncertainty in the estimate of the phase ϕ is minimized.

Our restriction in this paper to fixed *total* particle number leads to considerable analytical and computational simplification, but the more general problem in which the *average* particle number is conserved is also of interest, and can in principle be carried out with the same techniques.

III. LIKELIHOOD FUNCTION

A number of approaches have been used as measures of the uncertainty in the estimate of the relative phase ϕ. Commonly the standard error propagation formula is used to express this phase uncertainty in terms of the mean square error of a measured observable such as the particle number difference [17]

$$\Delta\phi = \Delta\hat{J}_z / (d\hat{J}_z/d\phi). \tag{11}$$

Probability operator measures are also used [14], as well as information-theoretical measures such as the Shannon mutual information [19]. In this paper we estimate the relative phase following an operational approach based on Bayes' theorem [4,20–22]. Consider an experiment in which the probability amplitude of the ith basis state of an $(N+1)$-dimensional Hilbert space depends on some phase ϕ,

$$|\psi\rangle = \sum_{i=0}^{N} \alpha_i(\phi) |i\rangle. \tag{12}$$

The probability to measure $|i\rangle$ conditioned on that phase is $P(i|\phi) = |\alpha_i(\phi)|^2$, with

$$\sum_{i=0}^{N} P(i|\phi) = 1. \tag{13}$$

Bayes' theorem states that the probability that the phase shift has the value ϕ, conditioned on the outcome i, is

$$P(\phi|i) = \frac{P(\phi)P(i|\phi)}{P(i)}, \tag{14}$$

where $P(\phi)$ is the phase probability distribution *prior* to the measurement and $P(i)$ is the prior detection probability for the outcome i. Following a measurement with outcome i_1, the phase probability distribution becomes $P(\phi|i_1)$, which may now be used as the prior phase probability distribution for a second measurement [4], so that

$$P(\phi|i_1,i_2) = \frac{P(\phi|i_1)P(i_2|\phi)}{P(i_2)} = \frac{P(\phi)P(i_1|\phi)P(i_2|\phi)}{P(i_1)P(i_2)}. \tag{15}$$

Likewise, the phase probability distribution conditioned on the outcome of a sequence of M measurements is

$$P(\phi|i_1,i_2,\ldots,i_M) = \frac{P(\phi)P(i_1|\phi)P(i_2|\phi)\cdots P(i_M|\phi)}{P(i_1)P(i_2)\cdots P(i_M)}. \tag{16}$$

For a large number of measurements, $M \gg 1$, and assuming that the true phase shift is $\phi = \theta$, the number of times a factor $P(i|\phi)$ appears in the product (16) is approximately $P(i|\theta)M$. This motivates the introduction of a likelihood function for the phase shift to be ϕ, conditioned on its true value being θ, as [21,22]

$$P_M(\phi|\theta) = \frac{1}{\mathcal{N}} \prod_{i=0}^{N} P(i|\phi)^{P(i|\theta)M}, \tag{17}$$

where

$$\mathcal{N} = \int_{-\pi/2}^{\pi/2} d\phi' \prod_{i=0}^{N} P(i|\phi')^{P(i|\theta)M} \tag{18}$$

is a normalization constant. The likelihood function $P_M(\phi|\theta)$ has the desirable property that it possesses a maximum at the true value, θ, of the phase shift. This is easily shown by taking its derivative

$$\frac{dP_M(\phi|\theta)}{d\phi} = \sum_{i=0}^{N} \left(MP(i|\theta)P(i|\phi)^{[MP(i|\theta)-1]} \frac{dP(i|\phi)}{d\phi} \prod_{k\neq i}^{N} P(k|\phi)^{MP(k|\theta)} \right) = M \left(\prod_{k=0}^{N} P(k|\phi)^{MP(k|\theta)} \right) \sum_{i=0}^{N} \frac{P(i|\theta)}{P(i|\phi)} \frac{dP(i|\phi)}{d\phi}. \tag{19}$$

Evaluating Eq. (19) at θ, together with the normalization condition (13) gives then

$$\left.\frac{dP(\phi|\theta)}{d\phi}\right|_{\theta} = M\left(\prod_{k=0}^{N} P(k|\theta)^{MP(k|\theta)}\right)\sum_{i=0}^{N}\left.\frac{dP(i|\phi)}{d\phi}\right|_{\theta} = 0, \tag{20}$$

implying an extremum at θ. Taking the second derivative and again using normalization shows this extremum to be a maximum.

In order to estimate the phase uncertainty in the limit of large M we introduce the function

$$K(\phi|\theta) = \ln P(\phi|\theta). \tag{21}$$

Expanding then $K(\phi|\theta)$ around $\phi=\theta$ and accounting for the normalization condition (13) we find that $P(\phi|\theta)$ is approximately given by

$$P(\phi|\theta) \approx e^{K(\theta|\theta)-[(\phi-\theta)^2/2\sigma^2]}, \tag{22}$$

where

$$\sigma^2 = \frac{1}{M\sum_{i=0}^{N}\left(\left.\frac{dP(i|\phi)}{d\phi}\right|_{\theta}\right)^2 \Big/ P(i|\theta)} = \frac{1}{MF}, \tag{23}$$

and

$$F = \sum_{i=0}^{N}\left(\left.\frac{dP(i|\phi)}{d\phi}\right|_{\theta}\right)^2 \Big/ P(i|\theta) \tag{24}$$

is the so-called Fisher information, as shown in Appendix A [23]. For large M the exponential suppresses strongly those contributions to $P(\phi|\theta)$ for which $\phi \gtrsim \sigma$ so that $P(\phi|\theta)$ becomes Gaussian in that limit [20]. Asymptotically, the likelihood function is therefore completely characterized by its variance, or equivalently by its Fisher information. Equation (23) also shows that the phase uncertainty decreases as the inverse square root of the number of measurements.

The Fisher information plays an important role in information theory as it gives a lower limit to the variance of any estimator via the Cramer-Rao inequality [23]

$$\mathrm{var}(x) \geqslant \frac{1}{F}, \tag{25}$$

where $\mathrm{var}(x)$ is the mean square error of the random variable x being estimated and Eq. (25) is the defining relation for the Fisher information. (Note also that the Fisher information of M' independent and identically distributed samples is M' times the individual Fisher information.) Thus Eq. (22) indicates that the likelihood function $P_M(\phi|\theta)$ achieves the Cramer-Rao limit. It permits us to find input states of the interferometer of the form of Eq. (10) that result in an estimate of the phase shift with minimum uncertainty.

To illustrate how the likelihood function may be used to estimate a phase shift experimentally, consider a thought experiment using the Mach-Zehnder interferometer in Fig. 1. Each measurement counts the number of particles n_a exiting the interferometer in arm "A." Due to particle conservation, this is a direct measure of the quantum number m. Expanding the exit state of the field as

$$|\psi_{\mathrm{out}}(\phi)\rangle = \sum_{m} \alpha_m(\phi)|j,m\rangle, \tag{26}$$

each measurement yields a specific particle number $n_a^{(i)} = N/2 + m^{(i)}$ with an associated phase probability distribution

$$P(m^{(i)}|\phi) = |\alpha_m^{(i)}(\phi)|^2. \tag{27}$$

After M such measurements the conditional phase probability distribution takes the form of Eq. (16), which for sufficiently large M is a good approximation to the likelihood function Eq. (17)—up to the normalization constant as in Eq. (18). The maximum of this conditional phase probability distribution is an estimate of the phase shift and its variance gives the uncertainty.

An important consideration is the number of measurements needed for the conditional phase probability distribution, Eq. (16), to be an accurate representation of the likelihood function. This matter is not addressed in this paper where we use throughout the asymptotic form of the likelihood function, but has been investigated by Braunstein [24].

IV. RESULTS

This section summarizes results of a numerical search for optimum input states of the interferometer. This search employed the global optimization protocol of simulated annealing [25,26], whose main features are summarized in Appendix B.

To set the stage for this discussion, we first recall that several states have previously been proposed as good candidates for Heisenberg-limited interferometry. One such state is the balanced twin-Fock input state [6,8]

$$|\psi_{\mathrm{tw}}\rangle = |j,0\rangle \equiv |N/2\rangle_a|N/2\rangle_b, \tag{28}$$

a state that we use as a benchmark in the following discussion. It was suggested in Refs. [3,8] that improvements over that state can be achieved by using instead the state

$$|\psi_{\mathrm{di}}(q)\rangle = (1/\sqrt{2})(|j,q\rangle + |j,-q\rangle), \tag{29}$$

with $q=1$. This state, which we refer to as a di-Fock state in the following, presents the advantage of suppressing secondary peaks in the likelihood function, thus concentrating more probability density around the true value of the phase shift.

It has also been proposed that Heisenberg-limited phase sensitivity can be achieved with the so-called $N00N$ state [27]

$$\begin{aligned}|\psi_{N00N}\rangle &\equiv |\psi_{\mathrm{di}}(q=j)\rangle = (1/\sqrt{2})(|j,j\rangle + |j,-j\rangle)\\ &= (1/\sqrt{2})(|N\rangle_a|0\rangle_b + |0\rangle_a|N\rangle_b).\end{aligned} \tag{30}$$

Some disagreement exists in the current literature regarding the phase sensitivity of $N00N$ states, with some authors claiming that it in fact obeys shot-noise-limited sensitivity [8]. Mitchell *et al.* [13] as well as Walther *et al.* [12] have pointed out that $N00N$ states may be used to produce superresolving phase oscillations, with a period of $2\pi/N$, in inter-

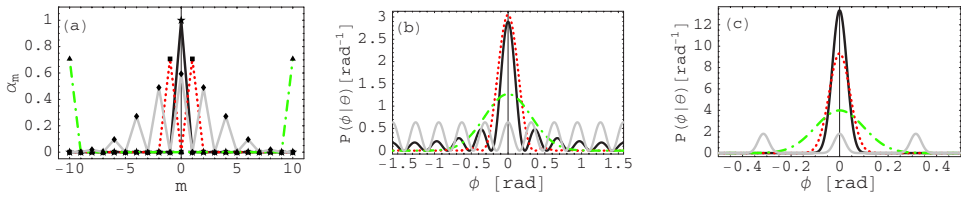

FIG. 2. (Color online) (a) Probability amplitudes α_m for the twin-Fock state (solid black line), the external $N00N$ state (green dotted-dashed line), the internal $N00N$ state Eq. (31) (solid gray line), and the di-Fock state $|\psi_{\mathrm{di}}(q=1)\rangle$ (red dotted line). Corresponding likelihood functions for (b) ($M=1$) and (c) ($M=10$). In the di-Fock state secondary peaks are already almost completely absent for $M=1$, thus concentrating more probability density around $\phi=0$. Secondary features are however suppressed for larger values of M in all cases, indicating that the twin-Fock state results in a slightly smaller phase uncertainty.

ferometric measurements. In agreement with this we will show that similar oscillations occur in the likelihood function, *if* Eq. (30) describes the state of the system *after the first beam splitter*. To be explicit we distinguish between external $N00N$ states for which Eq. (30) is the state *before* the first beam splitter, and internal $N00N$ states for which Eq. (30) is the state *after* the first beam splitter. A $N00N$ internal state is equivalent to using an input state

$$e^{-i\hat{J}_x\pi/2}|\psi_{N00N}\rangle. \tag{31}$$

It can also be achieved by using as an input the state $|N\rangle_a|0\rangle_b$ and replacing the first beam splitter with a nonlinear beam splitter with appropriate interaction time, as shown in [9].

Figure 2(a) shows the probability amplitudes α_m of the basis states $|j,m\rangle$ for the input states of Eqs. (28) and (29), as well as internal and external $N00N$ states, for a system with $N=20$ particles, and a relative phase $\theta=0$ between the two arms of the interferometer. The solid black line corresponds to the twin-Fock input, the dotted red line to the input state $|\psi_{\mathrm{di}}(q=1)\rangle$, the green dotted-dashed line to the $N00N$ external state, and the gray solid line to the $N00N$ internal state. The corresponding likelihood functions for $M=1$ are plotted in Fig. 2(b). Apart from the $N00N$ internal state the probability density is concentrated close to $\phi=0$ in all cases, but the di-Fock state seems more favorable as it results in a narrow distribution with no significant secondary peaks. However, this apparent advantage rapidly disappears for larger M, in which case the secondary peaks associated with the twin-Fock state are suppressed, leading to a slightly narrower distribution. Note the distribution corresponding to the $N00N$ external state remains considerably wider than the other candidates, indicating a larger uncertainty in the phase estimate. The likelihood function of the $N00N$ internal state rapidly oscillates with a period of $2\pi/N$ radians. This is consistent with the N-fold increase in phase oscillations observed in [12,13].

Figure 3 illustrates some of the large numbers of possible input states numerically obtained from the simulated annealing algorithm for a system of $N=100$ particles. Column A plots the probability amplitudes α_m of the input states; column B shows the corresponding likelihood functions with $M=1$ (broken red broken lines) and compares them to the likelihood function of the benchmark twin-Fock input (solid black lines); column C plots the situation for $M=10$ measurements. A remarkable feature of these results is that while these input states are very markedly different, their likelihood functions become almost indistinguishable for large M. Surprisingly perhaps the optimization procedure clearly shows the existence of a large number of local minima resulting in almost identical likelihood functions.

To demonstrate that all states identified by the simulated annealing algorithm indeed result in approximate Heisenberg-limited phase sensitivity, Fig. 4 shows a log-log plot of the inverse square root Fisher information, $1/\sqrt{F}$, as a function of particle number over a range of $50\leqslant N\leqslant 2500$. The solid red line is for twin-Fock states, the green dotted-dashed line for external $N00N$ states, the dotted blue line for the uppermost state in Fig. 3, and the dashed black line for the state in row 2 of Fig. 3. For clarity we refer to the latter state as a "Gaussian state" in the following. All lines are least square fits of the equation

$$\sigma=\frac{C}{N^{\beta}}, \tag{32}$$

where C and β are fit parameters.

Table I summarizes the results of this fit for the states of Fig. 3, the number referring to the row number in the figure. Up to differences of a few percent in the overall proportionality constant C, all of these states clearly satisfy the $1/N$ scaling characteristic of Heisenberg-limited sensitivity, the only notable exception being the external $N00N$ state, which (in agreement with Pezzé and Smerzi [8]) is shot-noise limited. On the other hand, the inverse square root Fisher information of the internal $N00N$ state does scale with the Heisenberg limit. Despite this, the rapid oscillations in the likelihood function seen in Figs. 2(b) and 2(c) allow a phase estimate only modulo $2\pi/N$. The consequent ambiguities imply that the internal $N00N$ state may not be useful for phase estimation when using the current Bayesian analysis unless one has *a priori* knowledge that the phase shift lies within a particular phase bin of width $2\pi/N$. We now discuss the candidates obtained by search algorithms in turn.

The state with the highest Fisher information that we found, an apparent global optimum, is shown in row 1 of Fig. 3. The envelope of its probability amplitudes α_i is a Gaussian with width $w=\sqrt{N}$. Despite the high Fisher information of that state, though, it produces a significant ambiguity in the determination of the phase estimate as secondary peaks persist even for $M=10$, as seen in column C. A feature not apparent on the scale in this figure is that the central peak

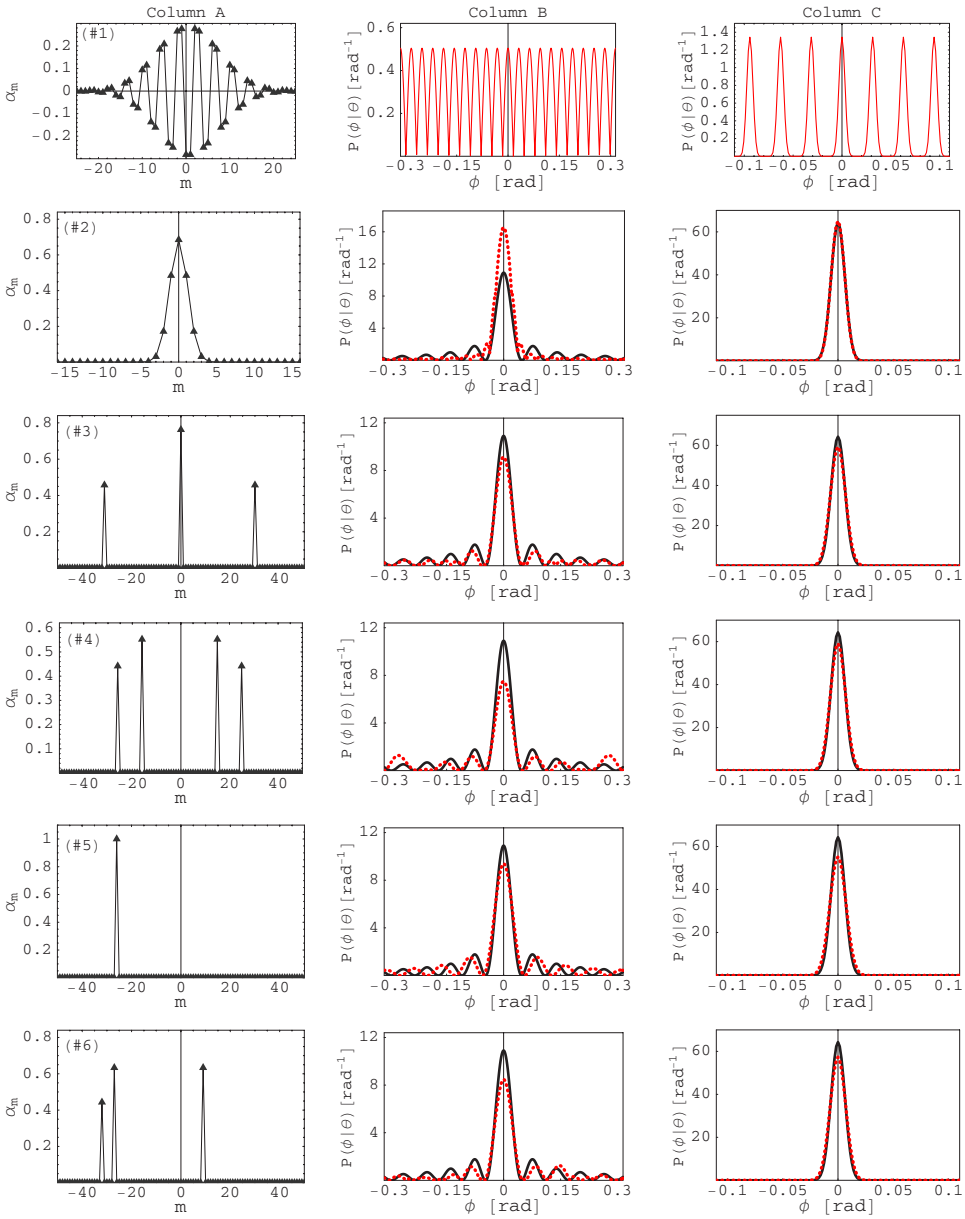

FIG. 3. (Color online) Column A: probability amplitudes for several possible input states with $N=100$ particles. Column B: Corresponding likelihood functions for $M=1$ (dotted red line), compared to the result for the benchmark twin-Fock input (solid black line). Column C: Likelihood functions for $(M=10)$.

is the absolute maximum and becomes increasingly dominant for increased M. Yet, as in the case of the internal $N00N$ state, the persistence of secondary peaks for relatively large sequences of measurements indicates that it may not be the most useful state in practice.

In the case of the "Gaussian state," second row of Fig. 3, the probability amplitudes α_m have a Gaussian distribution around the state $|j,m=0\rangle$,

$$\alpha_m = \exp(-m^2/\sigma'^2)/\mathcal{N}',$$

where $\mathcal{N}'$ is a normalization constant, and the standard deviation is $\sigma'=1.7$ for the example at hand. That state results in Heisenberg-limited sensitivity for $0<\sigma'\lesssim j$, with

$$\frac{1}{\sqrt{F}} \approx \frac{\sigma'}{N}. \tag{33}$$

The limit $\sigma' \to 0$ corresponding to the twin-Fock state.

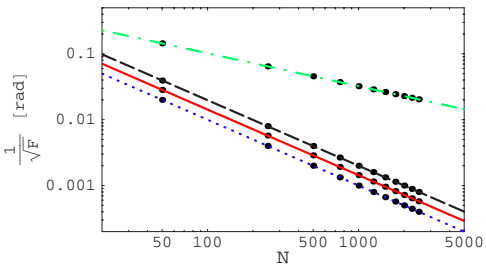

FIG. 4. (Color online) Scaling of the inverse square root Fisher information, $\sigma=1/\sqrt{F}$ with particle number for the state in row 1 of Fig. 3 (blue dotted line), twin-Fock state (solid red line), the "Gaussian state" (black dashed line), and the external $N00N$ state (green dotted-dashed line). The points are calculated using Eq. (24) while the lines correspond to least square fits, with the following values: external $N00N$ state, $\sigma=1.0/N^{1/2}$; Gaussian state, $\sigma=1.9/N$; twin-Fock state, $\sigma=1.4/N$; and the state in row 1 of Fig. 3, $\sigma=1.0/N$. Note the log-log scale.

TABLE I. Scaling of the phase uncertainty as a function of the particle number N for the twin state, the external $N00N$ and internal $N00N$ states, the di-Fock state, and the input states of Fig. 3.

State	$1/\sqrt{F}=C/N^{\beta}$
$\|j,0\rangle$	$1.4/N$
External $N00N$ state	$1.0/N^{1/2}$
Internal $N00N$ state	$1.0/N$
$\|\psi_{\mathrm{di}}(q=1)\rangle$	$2.0/N$
1	$1.0/N$
2	$1.9/N$
3	$1.5/N$
4	$1.5/N$
5	$1.7/N$
6	$1.6/N$

The state described in the third row of Fig. 3 is an example of a state we refer to as a tri-Fock state, and it has the form

$$|\psi_{\mathrm{tri}}(q)\rangle = (\alpha_-|j,-q\rangle + \alpha_0|j,0\rangle + \alpha_+|j,q\rangle)/\mathcal{N}', \qquad (34)$$

where $\mathcal{N}'$ is a normalization constant. We find numerically that it results in Heisenberg-limited sensitivity for any value of q.

The fourth row in Fig. 3 describes a state that is a superposition of four Fock states. For the $N=100$ particles considered in our simulations, and the state in row one of Fig. 3 aside, we have found states with superpositions containing up to approximately eight Fock states that result in Heisenberg-limited sensitivity.

We also found that di-Fock states, Eq. (29), with arbitrary q generally result in Heisenberg-limited or near Heisenberg-limited scaling for q as large as $q \lesssim 0.98j$. For larger q the state approaches the shot-noise limited external $N00N$ state.

Several general trends can be noted in the results of our search. First, we find that the scaling of Eq. (32) depends only weakly on the relative probability amplitudes α_M of the Fock states involved. Changing the relative amplitudes of these coefficients by factors as large as 3 typically results in changes in the coefficient C by a few percent only. The Gaussian state is a notable exception to this trend, and obeys instead the scaling equation (33).

Second, we found no inherent symmetry in the input states that result in the Heisenberg limit. This is illustrated by the states of rows 5–6 in Fig. 3. For example, the state of row 5 is an unbalanced twin-Fock state of the form

$$|\psi_{\mathrm{tw}}\rangle = |j,\gamma j\rangle, \qquad (35)$$

where γ is some fraction. We found numerically that the state resulted in Heisenberg-limited sensitivity for $0.02 \lesssim \gamma \lesssim 0.98$.

All states shown in Fig. 3 have real amplitudes. Allowing for complex amplitudes of the same magnitudes retains Heisenberg-limited or near Heisenberg-limited scaling, with $\beta \gtrsim 0.95$ in Eq. (32) and the coefficient C changed by only a few percent. Again the effect is more pronounced in the Gaussian state, where the change in C can be up to a factor of ~ 2. This is because that state has neighboring states occupied, and the number statistics of these states influence each other even for small phase shifts.

Due to the existence of numerous input states resulting in nearly identical uncertainties within the measurement scheme presented here, a more relevant criterion for the selection of an appropriate input state is likely to be its ease of experimental realization. We address this point in some more detail in the next section for the case of the "Gaussian state."

V. THE GAUSSIAN STATE

The "Gaussian" input state is a promising candidate for Heisenberg-limited interferometry for two reasons: (1) there is a simple experimental scheme available to generate it; (2) as we show below, the expectation value of $\hat{J}_z$ is phase dependent in that case (as opposed to the situation for twin-Fock input states) providing an alternative phase estimate to the direct measurement of the likelihood function, while still allowing Heisenberg-limited sensitivity.

A. Number statistics

As mentioned in Sec. III, the likelihood function can be experimentally reconstructed by multiplying the phase probability distributions $P(i|\phi)$ associated with a sequence of measurement outcomes. This is the approach that was adopted in the simulations carried out in Refs. [4,6] for twin-Fock input states. In that case however the average particle number difference remains zero for all relative phase shifts between the interferometer arms, and is therefore not a useful observable. The same is true for the majority of the states that we identified in our numerical optimization search. One way to circumvent this difficulty is to measure instead the variance of $\hat{J}_z$, an approach that still results in Heisenberg-limited estimates. However, as was pointed out in Ref. [6] for the case of the twin-Fock state, the signal-to-noise ratio is then small, $\langle \hat{J}_z^2\rangle / \sqrt{\langle \hat{J}_z^4\rangle - \langle \hat{J}_z^2\rangle^2} = \sqrt{2}$.

One advantage of using a Gaussian input state instead is that $\langle \hat{J}_z\rangle$ now depends on the relative phase ϕ. This is illustrated in Figs. 5(a)–5(c), which shows the probability distribution $P(m|\phi)$ at the exit of the interferometer for twin-Fock and Gaussian input states, and for phase shifts $\phi=0$, $\phi=\pi/100$, and $\phi=\pi/10$. The expectation value of $\hat{J}_z$ for the Gaussian state is clearly not equal to zero for nonzero phase shifts, and may therefore be used directly to estimate that shift. The uncertainty in $\hat{J}_z$, evaluated via the standard error propagation formula Eq. (11), is shown in Fig. 6 as a function of the number of interfering particles. Least square fits indicate that the uncertainties are Heisenberg limited, with $\Delta\phi=2.0/N$ for the Gaussian state and similarly $\Delta\phi=2.4/N$ in the case of a tri-Fock state with $q=1$.

B. Input state engineering

We have seen that the constraints on the relative phases of the complex probability amplitudes α_i of the input states are

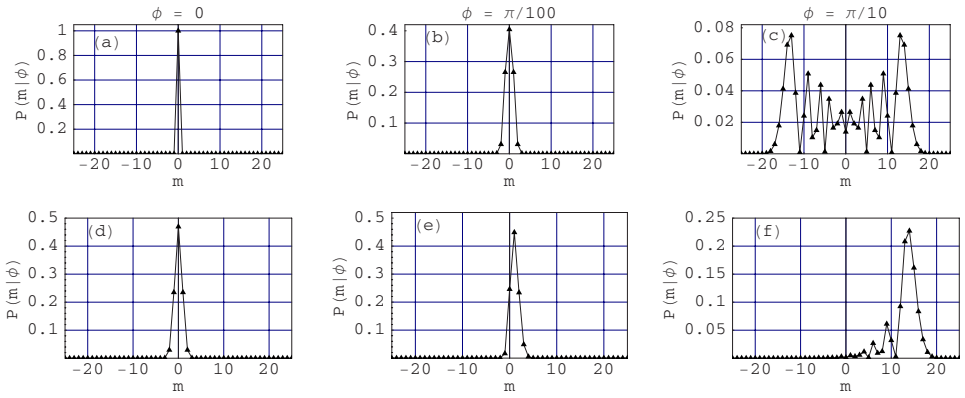

FIG. 5. (Color online) Number statistics for phase shifts of $\phi=0$, $\phi=\pi/100$, and $\phi=\pi/10$ in a system with $N=100$ particles for (a)–(c) a twin-Fock input state and (d)–(f) a Gaussian input state. The twin-Fock state populates number states symmetrically around $m=0$, leading to an expectation value of $\langle \hat{J}_z \rangle = 0$ for all values of ϕ. This is not the case for a Gaussian state.

surprisingly weak when estimating phase shifts ϕ via a reconstruction of the likelihood function. However, the number statistics of the field after passage through the interferometer depend critically on these relative phases. For example, if in the Gaussian state of Fig. 3 the components $|j,-1\rangle$ and $|j,+1\rangle$ are $\pi/2$ radians out of phase with $|j,0\rangle$, the output state $|\psi_{\text{out}}\rangle$ will be populated symmetrically around $|j,0\rangle$ independently of ϕ, so that $\langle \hat{J}_z \rangle = 0$ for all ϕ. Hence, some care must be taken in preparing the input states $|\psi_{\text{in}}\rangle$.

It is possible to generate a Gaussian input state from an initial twin-Fock state, Eq. (28), by subjecting it to the Hamiltonian

$$H_x = \hbar g \hat{J}_x, \tag{36}$$

where g is a coupling constant, for a time $\tau = 3\pi/4Ng$. The resulting state is precisely the state shown in row 2 of Fig. 3, but with the probability amplitudes of the components $|j,-1\rangle$ and $|j,+1\rangle$ $\pi/2$ radians out of phase with the component $|j,0\rangle$. These three components can be brought into phase via an additional evolution under the Hamiltonian

$$H_z = \hbar \chi \hat{J}_z^2 \tag{37}$$

for a time $\tau = 3\pi/2\chi$. The resulting state allows Heisenberg-limited phase estimation by measuring $\langle \hat{J}_z \rangle$, with the scaling law $\Delta\phi = 3.2/N$ as a function of the particle number N. We remark that while the time evolution (37) of the input state brings the three main components of $|\psi_{\text{in}}\rangle$ in phase with each other, that is not so for the other, weakly populated number states that comprise it. This results in a somewhat reduced performance compared to the Gaussian state with all components in phase.

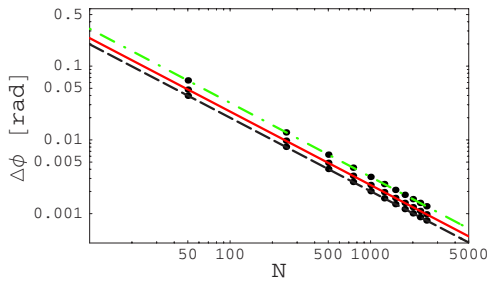

FIG. 6. (Color online) Scaling of the uncertainty in the estimate of the relative phase from a measurement of average particle number difference. The dashed black line is for a Gaussian state and the red solid line of a tri-Fock state, Eq. (34), with $q=1$, and the green dotted-dashed line represents the state engineered as described in Sec. V B. The points are numerically determined while the lines correspond to least square fits. Note the log-log scale.

Hamiltonians in the form of Eq. (37) have long been known to act as squeezing operators in interferometers [28]. In the case of photons, they can be implemented by inserting an optical Kerr medium into each arm of the interferometer [29]. In the case of charged particles they arise due to mutual phase modulation from Coulomb interaction between particles in each arm [30]. Atomic spins coupled to the polarization of an optical field [31,32] can lead to similar Hamiltonians for neutral atoms.

VI. CONCLUSION

We have used a global optimization scheme to systematically search for input states of a quantum-mechanical Mach-Zehnder interferometer that yield phase estimates with accuracy scaling like the inverse number of particles, the Heisenberg limit. Surprisingly perhaps, we find that a large number of states can achieve that limit. They typically consist of superpositions of a small number of Fock states, with few restrictions on the relative phase of their complex amplitudes on their symmetry. An input state of particular relevance consists of a Gaussian distribution of amplitudes around the state $|N/2\rangle_a |N/2\rangle_b$, due principally to its relative ease of realization with simple Hamiltonian dynamics.

ACKNOWLEDGMENTS

The authors would like to thank Lajos Diosi for insightful conversations regarding the relation between the likelihood function and information-theoretical concepts. This work is supported in part by the U.S. Office of Naval Research, by the National Science Foundation, by the U.S. Army Research Office, and by the National Aeronautics and Space Administration.

APPENDIX A: FISHER INFORMATION

Consider the deviation $\langle \delta A_\theta \rangle$ of an observable from its mean value $\langle A(\theta) \rangle$ at a fixed phase θ, as a function of the phase shift ϕ,

$$\langle \delta A_\theta \rangle = \sum_j [a_j - \langle A(\theta) \rangle] P(j|\phi). \tag{A1}$$

Then

$$\frac{d\langle \delta A_\theta \rangle}{d\phi} = \sum_j [a_j - \langle A(\theta) \rangle] \frac{dP(j|\phi)}{d\phi} \tag{A2}$$

$$= \sum_j [a_j - \langle A(\theta) \rangle] P(j|\phi) \frac{\ln P(j|\phi)}{d\phi}. \tag{A3}$$

Regrouping the factors in Eq. (A3) and squaring gives

$$\left(\frac{d\langle \delta A_\theta \rangle}{d\phi} \right)^2 = \left[\sum_j \{ [a_j - \langle A(\theta) \rangle] \sqrt{P(j|\phi)} \} \times \left(\sqrt{P(j|\phi)} \frac{\ln P(j|\phi)}{d\phi} \right) \right]^2 \tag{A4}$$

$$\leqslant \left[\sum_j P(j|\phi) \left(\frac{\ln P(j|\phi)}{d\phi} \right)^2 \right] \times \left(\sum_j [a_j - \langle A(\theta) \rangle]^2 P(j|\phi) \right), \tag{A5}$$

where we have used the Schwarz inequality in the last step. Noting that the second term in Eq. (A2) vanishes due to normalization we have $d\langle \delta A_\theta \rangle / d\phi = d\langle A \rangle / d\phi$. We can then rewrite inequality (A5) and evaluate it at θ to give the desired result,

$$\Delta\phi^2 = \frac{\Delta A^2}{(d\langle A \rangle / d\phi)^2} \geqslant \frac{1}{\sum_j P(j|\theta) \left(\left. \frac{\ln P(j|\phi)}{d\phi} \right|_\theta \right)^2}. \tag{A6}$$

This is the Cramer-Roa inequality, Eq. (25), in the current context. The denominator on the right-hand side of Eq. (A6) defines the Fisher information.

APPENDIX B: SIMULATED ANNEALING

Simulated annealing [25,26] is a mathematical approach to global optimization simulating the metallurgical process whereby an amorphous compound is successively heated and cooled while gradually lowering the average temperature in an attempt to enlarge the grain size of single crystals in the compound. The protocol is as follows:

(1) Initial conditions for the optimization parameters, $\mathbf{x} = (x_1, x_2, \ldots, x_r)$ are chosen, usually at random.

(2) With each allowed value of $\mathbf{x}$ is associated a pseudoenergy, $E(\mathbf{x})$, which is the quantity to be minimized.

(3) A new value, $\mathbf{x}' = \mathbf{x} + \mathbf{\Delta x}$, is then generated for the optimization parameters, and the change in energy $\Delta E = E(\mathbf{x}') - E(\mathbf{x})$ calculated.

(4) The new value of $\mathbf{x}$ always replaces the old if $p = e^{-\Delta E/kT} > 1$, or with probability p if $p < 1$. Here k is a constant of proportionality analogous to Boltzmann's constant in statistical mechanics, and T is a pseudotemperature.

(5) The process is repeated.

Accepting new parameter sets for which $p < 1$, i.e., uphill steps on the energy manifold, allows the algorithm to explore the whole parameter space instead of converging directly to the closest local minimum.

Two important considerations in this procedure are the method of choosing the next set of parameters $\mathbf{x}'$ and the annealing schedule, i.e., the protocol for gradually lowering the temperature with intermittent heating cycles until the system has frozen into, hopefully, a global minimum. Various approaches to these considerations have been discussed in the literature [25,33].

In our implementation the optimization parameters are the set of amplitudes α_m of the input state vector Eq. (10) and the energy the inverse square root of the Fisher information. We execute the simulated annealing algorithm not on a single vector $\bar{\alpha}$, but a population of vectors chosen at random. The pseudotemperature of the system is set by the average energy of the population,

$$T = \frac{1}{P} \sum_{i=1}^{P} 1/\sqrt{F_i}, \tag{B1}$$

where P is the number of state vectors in the population and F_i is the Fisher information of the ith state vector. Defining the temperature in this way self-regulates the cooling cycle. If a single global minimum exists, the algorithm will continue sampling until the majority of state vectors have fallen into the global minimum. On the other hand, if many local minima of comparable depth exist the algorithm will also continue to sample the parameter space until the majority of state vectors have found such local minima. As more local minima are found the system "cools down" by itself.

When the state vectors $\bar{\alpha}$ have converged near the minima and the step size $|\Delta\bar{\alpha}| = |\bar{\alpha}' - \bar{\alpha}|$ is fixed, all new steps $\bar{\alpha}'$ will be uphill, thus halting further convergence. To enable further convergence we therefore half the step size when the number of downhill steps found over several iterations drops below a threshold.

It may also happen in our approach that the system reaches an equilibrium condition in which the average number of uphill steps accepted become equal to the average number of downhill steps found. To ensure that the system continues to converge towards minima, we lower Boltzmann's constant by $k \to 0.9k$ if this point is reached. We take as an indicator that the system is near this point whenever the number of accepted uphill steps is larger than the number of downhill steps in a given iteration cycle.

To summarize the algorithm:

(1) Choose initial population at random and calculate pseudotemperature.

(2) Find new state vectors $\bar{\alpha}' = \bar{\alpha} + \Delta\bar{\alpha}$ and replaces the old if $p = e^{-\Delta E/kT} > 1$, or with probability p if $p < 1$.

(3) If the number of uphill steps accepted is greater than the number of downhill steps decrease $k \to 0.9k$.

(4) If the number of downhill steps found is less than the specified threshold reduce $\Delta\bar{\alpha}\rightarrow 0.5\Delta\bar{\alpha}$.

(5) Repeat the algorithm

We have implemented searches that assume either real or complex amplitudes. In addition, in some searches we imposed no restrictions on the symmetry of input states, while in others we forced the input states to be either symmetric or antisymmetric around $|N/2\rangle_a|N/2\rangle_b$.

In the case of real amplitudes, the initial population was chosen by generating a random number between $[-1,1]$ for each α_m and then normalizing the state vector. For complex amplitudes the magnitudes $|\alpha_m|$ were chosen at random between 0 and 1 and a complex phase between 0 and 2π.

We have used two different approaches to specifying the new state vectors $\bar{\alpha}'$ during each iteration. In the first one new state vectors were selected by changing each amplitude at random within an interval $\epsilon\alpha_m<\alpha'_m<(1+\epsilon)\alpha_m$, with $\epsilon<1$ and then renormalizing the state vector. In the case of complex amplitudes a new phase was also chosen as $\phi'_m=\phi_m+\Delta\phi$ where $-\varepsilon<\Delta\phi<\varepsilon$, while in the case of real amplitudes sign changes were allowed if $|\alpha_m|<0.02$ by choosing $-(1+\epsilon)|\alpha_m|<\alpha'_m<(1+\epsilon)|\alpha_m|$. In this approach, step (4) in the algorithm described above was rather insensitive to the values chosen for ϵ and ε. They were therefore taken to have fixed values $\epsilon=0.5$ and $\varepsilon=0.05$.

In the second approach each vector $\bar{\alpha}$ is specified by a set of angles such that the vector moves on a hyperspherical surface of radius 1 to ensure normalization. The next vector is chosen in a random direction on the hypersphere with the initial step size $\Delta\bar{\theta}=0.1$. It is decreased in successive iterations according to step (5) in the algorithm above.

In the first approach the state vectors in the population settle in many local minima of comparable depths, while in the second approach all state vectors converge to an apparent global minimum which is the state shown in row 1 of column A in Fig. 3.

[1] P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta, Phys. Rev. Lett. **59**, 2153 (1987).
[2] M. Xiao, L. A. Wu, and H. J. Kimble, Phys. Rev. Lett. **59**, 278 (1987).
[3] B. Yurke, Phys. Rev. Lett. **56**, 1515 (1986).
[4] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[5] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).
[6] T. Kim, O. Pfister, M. J. Holland, J. Noh, and J. L. Hall, Phys. Rev. A **57**, 4004 (1998).
[7] J. Combes and H. M. Wiseman, J. Opt. B: Quantum Semiclassical Opt. **7**, 14 (2005).
[8] L. Pezzé and A. Smerzi, Phys. Rev. A **73**, 011801(R) (2006).
[9] K. Mølmer and A. Sørensen, Phys. Rev. Lett. **82**, 1835 (1999).
[10] C. A. Sackett *et al.*, Nature (London) **404**, 256 (2000).
[11] D. Leibfried, M. D. Barret, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, Science **304**, 1476 (2004).
[12] P. Walther, J. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, Nature (London) **429**, 158 (2004).
[13] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Nature (London) **429**, 161 (2004).
[14] D. W. Berry, H. M. Wiseman, and J. K. Breslin, Phys. Rev. A **63**, 053804 (2001).
[15] D. Denot, T. Bschorr, and M. Freyberger, Phys. Rev. A **73**, 013824 (2006).
[16] J. Söderholm, G. Björk, B. Hessmo, and S. Inoue, Phys. Rev. A **67**, 053803 (2003).
[17] C. P. Search and P. Meystre, Phys. Rev. A **67**, 061601(R) (2003).
[18] T. Wang and J. Javanainen, Phys. Rev. A **75**, 013605 (2007).
[19] T. B. Bahder and P. A. Lopata, Phys. Rev. A **74**, 051801(R) (2006).
[20] S. L. Braunstein, Phys. Rev. Lett. **69**, 3598 (1992).
[21] Z. Hradil, Phys. Rev. A **51**, 1870 (1995).
[22] Z. Hradil, R. Myska, J. Perina, M. Zawisky, Y. Hasegawa, and H. Rauch, Phys. Rev. Lett. **76**, 4295 (1996).
[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, New York, 2006).
[24] S. L. Braunstein, J. Phys. A **25**, 3813 (1992).
[25] I. O. Bohachevsky, M. E. Johnson, and M. L. Stein, Technometrics **28**, 209 (1986).
[26] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1999).
[27] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).
[28] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993).
[29] M. Kitagawa and Y. Yamamoto, Phys. Rev. A **34**, 3974 (1986).
[30] M. Kitagawa and M. Ueda, Phys. Rev. Lett. **67**, 1852 (1991).
[31] G. A. Smith, A. Silberfarb, I. H. Deutsch, and P. S. Jessen, Phys. Rev. Lett. **97**, 180403 (2006).
[32] J. M. Geremia, J. K. Stockton, and H. Mabuchi, Phys. Rev. A **73**, 042112 (2006).
[33] D. Vanderbilt and S. G. Louie, J. Comput. Phys. **56**, 259 (1984).