PHYSICAL REVIEW A **73**, 033821 (2006)

Optimal phase measurements with pure Gaussian states

Alex Monras
Grup de Física Teòrica & IFAE, Facultat de Ciències, Edifici Cn, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
(Received 26 September 2005; revised manuscript received 20 January 2006; published 29 March 2006)

We analyze the Heisenberg limit on phase estimation for Gaussian states. In the analysis, no reference to a phase operator is made. We prove that the squeezed vacuum state is the most sensitive for a given average photon number. We provide two adaptive local measurement schemes that attain the Heisenberg limit asymptotically. One of them is described by a positive operator-valued measure and its efficiency is exhaustively explored. We also study Gaussian measurement schemes based on phase quadrature measurements. We show that homodyne tomography of the appropriate quadrature attains the Heisenberg limit for large samples. This proves that this limit can be attained with local projective von Neuman measurements.

DOI: 10.1103/PhysRevA.73.033821 PACS number(s): 42.50.Dv, 42.50.Lc, 03.65.Ta, 03.67.Hk

I. INTRODUCTION

For a long time quantum phase measurements have been a field of attention for both theoretical and experimental research in quantum theory [1–18]. Many schemes have been proposed with the aim of optimally determining the phase shift produced by a physical process [16], usually focusing on quantized harmonic oscillators. Despite all the efforts, no attempt has ever settled the issue with full generality [1]. Many different approaches exist, but they often have a restricted range of applicability. There are various reasons for this situation. The most important one is the fact that a well defined phase operator does not exist on the whole Fock space. This seems to indicate that Heisenberg's uncertainty relation for phase and number may not be always taken at "face value" and a rederivation without the use of such a phase operator is called for (see [19] and references therein).

In this paper we are concerned with phase measurements on pure Gaussian states. First, we review the derivation of the Heisenberg limit using the theory of quantum state estimation [19,20], which does not require the existence of a phase operator. Then we propose a positive operator-valued measure (POVM) that attains the limit by an adaptive procedure [21–23], for an asymptotically large number of copies. We also determine the optimal Gaussian measurement for phase estimation in this asymptotic regime.

When performing phase measurements on Gaussian states, one usually implements them on light beams with a given average photon number, i.e., with a fixed energy. This constraint is always assumed when comparing different Gaussian states in order to determine their optimality for phase estimation. The techniques we use, namely, asymptotics in quantum statistics, are best suited for the problem at hand, where one usually has a light beam turned on for a large number of coherence times, which translates into having a large number of identical copies of such a Gaussian state.

II. QUANTUM PARAMETER ESTIMATION

In this section we review the theory of quantum parameter estimation [6,19,20,23–26], which leads to a useful generalization of the Heisenberg uncertainty relation.

Let us assume we have a family of pure states $\rho(\theta)$ labeled by a parameter θ. In our case the parameter is the phase. This family corresponds to a curve in Hilbert space, and translations along such a curve are generated by a Hermitian operator G in the usual manner. In our case the generator is the number operator $G=a^\dagger a$:

$$\rho(\theta)=\exp(-i\theta G)\rho(0)\exp(i\theta G). \tag{1}$$

Throughout this paper we use the following notation. We will write $|n\rangle$ to denote Fock states, i.e., states with a well defined number. With $\mathcal{U}_\theta$ we denote the unitary operator that yields the state $\rho(\theta)$ from the vacuum, namely

$$\rho(\theta)=\mathcal{U}_\theta|0\rangle\langle 0|\mathcal{U}_\theta^\dagger. \tag{2}$$

We define the vectors $|\phi_n\rangle$ as the Fock states transformed by $\mathcal{U}_\theta$,

$$|\phi_n\rangle=\mathcal{U}_\theta|n\rangle; \tag{3}$$

hence, $\rho(\theta)=|\phi_0\rangle\langle\phi_0|$.

Our aim is to determine the parameter θ by performing the best possible POVM measurement on the system. For this, we propose an estimator $\hat{\theta}(x)$, which is a function of the possible outcomes x of the measurement. Hence, the expectation value of such estimator is [27]

$$\mathbb{E}_\theta[\hat{\theta}]=\sum_x p(x|\theta)\hat{\theta}(x), \tag{4}$$

and its variance

$$\mathrm{Var}_\theta[\hat{\theta}]=\sum_x p(x|\theta)[\hat{\theta}(x)-\mathbb{E}_\theta[\hat{\theta}]]^2, \tag{5}$$

where the subscript θ means that the expectation value is taken with probability distribution $p(x|\theta)$. In the quantum case, they are given by the Born rule $p(x|\theta)=\mathrm{tr}[\rho(\theta)E_x]$, where $\{E_x\}$ are the elements of a POVM.

An estimator is called unbiased when

$$\mathbb{E}_{\theta}[\hat{\theta}] = \theta. \tag{6}$$

In this case, the variance is equivalent to the mean squared error (MSE):

$$\mathrm{MSE}[\hat{\theta}] \equiv \mathbb{E}[(\hat{\theta} - \theta)^2] = \mathrm{Var}[\hat{\theta}] + (\mathbb{E}[\hat{\theta}] - \theta)^2. \tag{7}$$

A well known theorem in statistics gives a lower bound to the variance of any unbiased estimator. It is the so-called Cramér-Rao bound [27],

$$\mathrm{Var}_{\theta}[\hat{\theta}] \geq \frac{1}{F(\theta)}, \tag{8}$$

where $F(\theta)$ is the Fisher information associated with the measurement

$$F(\theta) = \sum_{x \in X_+} p(x|\theta)\left(\frac{\partial \log p(x|\theta)}{\partial \theta}\right)^2, \tag{9}$$

and the sum runs over the set of *possible* outcomes X_+, i.e., those with $p(x|\theta) \neq 0$.

Furthermore, the Braunstein-Caves inequality [20] sets an upper bound on the Fisher information:

$$F(\theta) \leq H(\theta), \tag{10}$$

where $H(\theta)$ does not depend on the specific measurement being performed. $H(\theta)$ is sometimes regarded as the Quantum Fisher Information (QFI). It is defined as [6,20,21,23]

$$H(\theta) = \mathrm{tr}[\rho(\theta)\lambda(\theta)^2], \tag{11}$$

where λ is the symmetric logarithmic derivative (SLD), defined as the Hermitian operator that fulfills

$$\frac{\partial \rho(\theta)}{\partial \theta} = -i[G,\rho(\theta)] = \frac{1}{2}[\lambda(\theta)\rho(\theta) + \rho(\theta)\lambda(\theta)]. \tag{12}$$

Equations (8) and (10) set a fundamental bound on the variance of any unbiased estimator. This bound only depends on the geometrical properties of the curve $\rho(\theta)$. In our case, the SLD is

$$\lambda(\theta) = 2i(|\phi_0\rangle\langle\psi| - |\psi\rangle\langle\phi_0|) = -2i[G,\rho(\theta)], \tag{13}$$

where

$$|\psi\rangle = (\mathbb{1} - |\phi_0\rangle\langle\phi_0|)G|\phi_0\rangle. \tag{14}$$

The eigenvectors of λ are $|\phi^{\pm}\rangle = |\psi\rangle \pm i\sqrt{\langle\psi|\psi\rangle}|\phi_0\rangle$, with the corresponding eigenvalues given by $l_{\pm} = \pm 2\sqrt{\langle\psi|\psi\rangle}$. Note that $\langle\psi|\phi_0\rangle = 0$. Working out the value of H, we get

$$H(\theta) = 4\langle\psi|\psi\rangle = 4\langle\Delta G\rangle_{\theta}^2, \tag{15}$$

where $\langle\Delta G\rangle_{\theta}^2$ is defined, as usual, as $\langle\phi_0|G^2|\phi_0\rangle - \langle\phi_0|G|\phi_0\rangle^2$. Thus,

$$\mathrm{Var}_{\theta}[\hat{\theta}]\langle\Delta G\rangle_{\theta}^2 \geq \frac{1}{4}. \tag{16}$$

This expression is similar to Heisenberg's uncertainty relation for canonically conjugate variables, but has some advantages (see [19]). First of all, it has been derived without the use of any phase operator. In fact, the only operator we need is the phase shift generator, i.e., the number operator. On the other hand, it sets a lower bound on the variance of an estimator, whereas the standard uncertainty relation does not concern optimality but only variances obtained from measurements of observables (i.e., self-adjoint operators). In this sense, Eq. (16) is more general.

Now assume we have N identical copies of the same unknown state $\rho(\theta)$. In this case the collective state $\rho(\theta)^{\otimes N}$ is still a member of a one-parameter family, and it is straightforward to show that the corresponding QFI scales as N. Thus, combining Eqs. (8) and (10), with H replaced by the N-copy QFI and $\hat{\theta}$ denoting any unbiased estimator based on any measurement (collective or individual) on the N copies, we have

$$\mathrm{Var}_{N,\theta}[\hat{\theta}]\langle\Delta G\rangle_{\theta}^2 \geq \frac{1}{4N}, \tag{17}$$

where the subscript N stands for the number copies.

An important issue is the attainability of these bounds. The Braunstein-Caves inequality is known to be saturable [20,24], i.e., there exists a POVM that gives the equality in (10) [it is given by $\lambda(\theta)$; see Sec. IV for details]. Note, however, that in general $\lambda(\theta)$ is not constant, and the optimal POVM depends on the true value of θ. This suggests that, in order to attain the equality in (10) for large N, one needs an adaptive scheme. The measurement at a given step of the estimation scheme may need to be optimized using the data gathered from the previous steps. The Cramér-Rao bound is also known to be asymptotically saturable by the so-called maximum likelihood estimator (MLE). That is, for multiple identical measurements, the MLE has a variance that approaches the inverse of the Fisher information, thus attaining equality also in (8). All this seems to indicate that the optimal scheme should consist of at least two steps: A preliminary rough estimate of θ, using a vanishing fraction of copies (N^{α}, with $0 < \alpha < 1$), and a refinement using the remaining $\bar{N} = N - N^{\alpha}$ copies, where all measurements are identical in the last step.

Therefore, if $N \to \infty$, one can reasonably expect that

$$\lim_{N\to\infty} N\,\mathrm{Var}_{\bar{N},\theta}[\hat{\theta}] = \frac{1}{4\langle\Delta G\rangle_{\theta}^2}, \tag{18}$$

for the MLE, provided the optimal POVM measurement is performed in step two. Here the subscript $\bar{N}$ indicates that the variance is obtained from the outcomes of the last $\bar{N}$ identical measurements. This will be loosely written as

$$\mathrm{Var}_{\theta}[\hat{\theta}] \sim \frac{1}{4\langle\Delta G\rangle_{\theta}^2}. \tag{19}$$

We show below that (19) holds if the number of copies used in each of the two steps of the estimation scheme (i.e., α) is appropriately chosen.

A word should be said about asymptotically unbiased estimators. These are estimators whose expectation values converge to the parameter in the large sample limit, but have some vanishing bias, i.e.,

$$\mathbb{E}[\hat{\theta}] = \theta + O(N^{-\alpha}), \tag{20}$$

with $\alpha > 0$. In this case the correct quantifier of the sensitivity is the MSE. If $\alpha > 1/2$ the MSE is asymptotically equivalent to the variance, since

$$\text{MSE}[\hat{\theta}] = \text{Var}[\hat{\theta}] + O(N^{-2\alpha}), \tag{21}$$

and we note that the leading order contribution in inverse powers of N is solely given by $\text{Var}[\hat{\theta}]$.

III. OPTIMAL GAUSSIAN STATE

Next, we determine the optimal Gaussian state, which provides the highest sensitivity to phase measurements within our approach, i.e., the one that maximizes its QFI. The state is obtained by sequentially applying a series of operations on the vacuum: (*i*) a squeezing along a fixed direction (say Q), $S(r)=\exp[(r/2)(a^2-a^{\dagger 2})]$; (*ii*) a displacement, $D(\alpha)=\exp(\alpha a^\dagger - \alpha^* a)$ with real α; (*iii*) the unknown phase shift, $U(\theta)=\exp(-i\theta a^\dagger a)$. The state after applying these operations is

$$|\phi_0\rangle = U(\theta)D(\alpha)S(r)|0\rangle, \tag{22}$$

and its QFI is given by

$$H(\theta) = 4\langle \Delta a^\dagger a\rangle_\theta^2 = 4[|\alpha|^2(\cosh r - \sinh r)^2 + 2\sinh^2 r \cosh^2 r]. \tag{23}$$

Notice that actually H does not depend on the phase. The energy of the state (22) is

$$\langle n\rangle_\theta = \langle a^\dagger a\rangle_\theta = |\alpha|^2 + \sinh^2 r. \tag{24}$$

We aim at maximizing H with the constraint to have a fixed average number of photons. Using Lagrange multipliers, it is straightforward to find that the most sensitive choice is $\alpha=0$. As intuition dictates, the highest sensitivity is achieved by employing all available energy in squeezing the vacuum. Therefore, the optimal Gaussian state is $|\phi_0\rangle = U(\theta)S(r)|0\rangle$, for which

$$H = \cosh 4r - 1 = 8(\langle n\rangle^2 + \langle n\rangle). \tag{25}$$

Hence, the optimal phase estimation has a variance that goes as

$$\text{Var}[\hat{\theta}] \sim \frac{1}{8(\langle n\rangle^2 + \langle n\rangle)}. \tag{26}$$

IV. ASYMPTOTICALLY OPTIMAL FEW-OUTCOME MEASUREMENT

Following the derivation of the Braunstein-Caves inequality [20], one can check that the optimality conditions on the POVM elements $\{E_x\}$ are (see [24])

$$\text{Im}\{\text{tr}[\rho(\theta)E_x\lambda(\theta)]\} = 0,$$
$$\sqrt{\rho(\theta)}\lambda(\theta)\sqrt{E_x} = k_x\sqrt{\rho(\theta)}\sqrt{E_x}, \tag{27}$$

for all x that have nonvanishing probabilities, where k_x are some constants. Assuming pure states $\rho^2=\rho$, and the fact that any POVM can be decomposed into rank one operators (so that $\sqrt{E_x} \propto E_x$), one may write the second optimality condition as

$$\rho(\theta)\lambda(\theta)E_x = k_x\rho(\theta)E_x. \tag{28}$$

Finally, one can check that a sufficient condition for Eqs. (27) and (28) to hold is that $\{E_x\}$ project onto the eigenspaces of the SLD. Therefore, an optimal POVM can be chosen to have three elements, $\{E_+, E_-, E_0\}$, given by

$$E_\pm = \frac{1}{2\langle\psi|\psi\rangle}|\phi^\pm\rangle\langle\phi^\pm| \tag{29}$$

and

$$E_0 = \mathbb{1} - E_+ - E_-. \tag{30}$$

For completeness, we give the explicit form of the states $|\phi^\pm\rangle$. They are

$$|\phi^\pm\rangle = \sqrt{2}\sinh r\cosh r(\pm i|\phi_0\rangle - |\phi_2\rangle), \tag{31}$$

which, after normalizing, become

$$\frac{|\phi^\pm\rangle}{\||\phi^\pm\rangle\|} = \frac{\pm i|\phi_0\rangle - |\phi_2\rangle}{\sqrt{2}} = \mathcal{U}_\theta\frac{\pm i|0\rangle - |2\rangle}{\sqrt{2}}. \tag{32}$$

This measurement resembles tomography applied to the estimation of spin states lying close to the z axis of the Bloch sphere, where spin measurements along the x and y directions are optimal [21,28].

From the hermiticity of λ, it is straightforward to check that $\{E_+, E_-, E_0\}$ are positive and, furthermore, that they represent projective von Neuman measurements. The outcomes of these measurements will be used in the maximum likelihood analysis of the next section, where its corresponding expectation value and MSE will be discussed.

V. MAXIMUM LIKELIHOOD ESTIMATION

Assume one has performed step one, a series of nonoptimal measurements on a vanishingly small number of copies, N^α, of our optimal Gaussian state $|\phi_0\rangle$, and has obtained a rough guess $\hat{\theta}_0$ of its phase. Let us write $\hat{\theta}_0 = \theta - \delta\theta$, where $\delta\theta$ represents the error in this first step, which is assumed to be small. Let $|\hat{\phi}_0\rangle$ be the guessed state. We design (step two) the optimal measurement on the remaining $\bar{N} = N - N^\alpha$ copies as if $\hat{\theta}_0$ were the true phase.

Let $N_{\pm,0}$ be the number of times the outcomes $\pm, 0$ are obtained. The likelihood function for θ is [27]

$$\mathcal{L}(\theta) = \frac{\bar{N}!}{N_+!N_-!N_0!}p(+|\theta)^{N_+}p(-|\theta)^{N_-}p(0|\theta)^{N_0}. \tag{33}$$

Numerical maximization of this function yields the strict MLE. Recall that this is known to attain the Cramér-Rao bound in the limit of a large number of copies. We can, however, analytically compute an approximate estimator, which converges to the MLE in the large sample limit and has the same nice properties, i.e., it is asymptotically unbiased and attains the Cramér-Rao bound asymptotically. These properties will be checked explicitly.

Taylor expanding the true state around the first-step guess, we have

$$\rho(\theta)=\rho(\hat{\theta}_0)+\frac{1}{2}[\lambda(\hat{\theta}_0)\rho(\hat{\theta}_0)+\rho(\hat{\theta}_0)\lambda(\hat{\theta}_0)]\delta\theta+O(\delta\theta^2), \tag{34}$$

and the probabilities of obtaining the outcomes $x=\pm,0$ are

$$p(\pm|\theta)=\frac{1}{2}\pm\sqrt{\langle\hat{\psi}|\hat{\psi}\rangle}\delta\theta+O(\delta\theta^2), \tag{35}$$

$$p(0|\theta)=O(\delta\theta^2), \tag{36}$$

where $|\hat{\psi}\rangle$ is given by Eq. (14) by replacing $|\phi_0\rangle$ for $|\hat{\phi}_0\rangle$. With this, $\langle\hat{\psi}|\hat{\psi}\rangle=\langle\Delta n\rangle^2$. Thus,

$$\frac{\partial\log\mathcal{L}}{\partial\theta}=\sum_{x\in\{+,-,0\}}\frac{N_x}{p(x|\theta)}\frac{\partial p(x|\theta)}{\partial\theta}$$
$$=4\langle\Delta n\rangle\left(\frac{N_+-N_-}{2}-(N_++N_-)\langle\Delta n\rangle\delta\theta\right)+O(\delta\theta^2). \tag{37}$$

Equating this to zero yields the approximate MLE at leading order in $\delta\theta$:

$$\hat{\theta}_{\text{MLE}}\equiv\hat{\theta}_0+\frac{N_+-N_-}{2N_{\text{inf}}\langle\Delta n\rangle}. \tag{38}$$

Here $N_{\text{inf}}=N_++N_-$ is the total number of "informative" outcomes. This estimator is based on disregarding the improbable and noninformative outcomes $x=0$. Our approximate MLE is undefined, in the rare event that $N_{\text{inf}}=0$. In this case we may just keep the preliminary estimate and define $\hat{\theta}_{\text{MLE}}\equiv\hat{\theta}_0$.

Disregarding the unlikely $x=0$ outcomes, the estimator is just based on a binomial distribution, and for a fixed value of N_{inf} we have

$$\mathbb{E}_{\theta,N_{\text{inf}}}[\hat{\theta}_{\text{MLE}}]=\sum_{N_+=0}^{N_{\text{inf}}}\hat{\theta}_{\text{MLE}}\binom{N_{\text{inf}}}{N_+}q^{N_+}(1-q)^{N_-}=\hat{\theta}_0+\frac{2q-1}{2\langle\Delta n\rangle}, \tag{39}$$

where q is, to the relevant order, $q=1/2+\langle\Delta n\rangle\delta\theta+O(\delta\theta^2)$. This yields

$$\mathbb{E}_{\theta,N_{\text{inf}}}[\hat{\theta}_{\text{MLE}}]=\theta+O(\delta\theta^2). \tag{40}$$

Since $\delta\theta^2=O(N^{-\alpha})$, the expectation value is asymptotically unbiased [recall Eq. (20)]. It is also straightforward to compute the MSE of this estimator, for a given number N_{inf} of informative outcomes:

$$\text{MSE}_{\theta,N_{\text{inf}}}[\hat{\theta}_{\text{MLE}}]=\frac{1}{4\langle\Delta n\rangle^2N_{\text{inf}}}+O\left(\frac{\delta\theta^2}{\bar{N}}\right)+O(\delta\theta^4), \tag{41}$$

for $N_{\text{inf}}>0$, and

$$\text{MSE}_{\theta,0}[\hat{\theta}_{\text{MLE}}]=\delta\theta^2, \tag{42}$$

for $N_{\text{inf}}=0$.

To compute the full MSE of our scheme we just need to average over all possible values of N_{inf}. This yields (see Appendix A for details)

$$\text{MSE}_\theta[\hat{\theta}_{\text{MLE}}]=\sum_{N_{\text{inf}}=0}^{\bar{N}}\text{MSE}_{\theta,N_{\text{inf}}}[\hat{\theta}_{\text{MLE}}]\binom{\bar{N}}{N_{\text{inf}}}p^{N_{\text{inf}}}(1-p)^{\bar{N}-N_{\text{inf}}} \tag{43}$$

$$=\frac{1}{\bar{N}}\left(\frac{1}{4\langle\Delta n\rangle^2}+O(\delta\theta^2)\right), \tag{44}$$

where p is the probability of getting an informative outcome in a given measurement $[p=1-O(\delta\theta^2)]$. Clearly enough, by choosing $\alpha>1/2$ the error of the first step only contributes to subleading orders in the variance. If, e.g., $\alpha=2/3$, one sees that in the large N limit, the MSE$[\hat{\theta}]$ reduces to

$$\text{MSE}_\theta[\hat{\theta}_{\text{MLE}}]=\frac{1}{8(\langle n\rangle^2+\langle n\rangle)N}+O\left(\frac{1}{N^{4/3}}\right). \tag{45}$$

Hence, the optimal performance displayed in (26) is asymptotically attained.

VI. ASYMPTOTICALLY OPTIMAL GAUSSIAN MEASUREMENTS

Now we face the problem of determining the minimal phase variance that can be attained by means of dyne measurements. These consist of the simultaneous measurement of two conjugate variables, such as Q and P, or their phase generalizations $U(\theta')QU^\dagger(\theta')$ and $U(\theta')PU^\dagger(\theta')$. They are typically performed by splitting the signal state. Often this is done by means of a beamsplitter, the so-called eight-port homodyne detector [29]. Another possibility is heterodyne detection. In any case, the Arthurs and Kelly theorem (see, e.g., [29]) guarantees that further noise will appear in the detection process. In the eight-port homodyne detector, it appears as vacuum fluctuations entering through the unused port of the beamsplitter, which is sometimes called the *auxiliary port*. By introducing some sort of squeezed vacuum in the auxiliary port, one can reduce the noise in one quadrature, at the expense of increasing the noise in the corresponding conjugate quadrature. These measurements can be mathematically expressed as a covariant measurement [2,3] with POVM elements given by

$$E(\chi)=\frac{1}{2\pi}D(\chi)\sigma_0D^\dagger(\chi)=\frac{1}{2\pi}\sigma_\chi, \tag{46}$$

where $(q,p)\equiv\chi$ are the outcomes and represent coordinates in phase space, and $D(\chi)$ is the displacement operator $D(\chi)=\exp i(pQ-qP)$. The density matrix σ_0 is the ancilla entering through the auxiliary port of the beamsplitter. From the covariance of the measurement, σ_0 can be taken to be a squeezed vacuum state ($\langle q\rangle=\langle p\rangle=0$) without loss of generality. One can relate the probability of obtaining the outcome χ, $dp(\chi|\theta)$, with the fidelity [30], $\mathcal{F}[\sigma_\chi,\rho(\theta)]=\mathrm{tr}\sqrt{\rho(\theta)^{1/2}\sigma_\chi\rho(\theta)^{1/2}}$, as

$$dp(\chi|\theta)=\mathrm{tr}[\rho(\theta)E(\chi)]d^2\chi=\frac{d^2\chi}{2\pi}\mathcal{F}[\sigma_\chi,\rho(\theta)]^2. \quad (47)$$

When σ_0 is a Gaussian state, the measurement is said to be Gaussian and the outcomes are Gaussian distributed. Moreover, when one of the two states is pure, the fidelity can be easily expressed as $\mathcal{F}[\sigma_\chi,\rho(\theta)]^2=2\sqrt{\det M(\theta)}e^{-\chi^t M(\theta)\chi}$ [31], where $M(\theta)=(\gamma_0+\gamma_\theta)^{-1}$. The covariance matrices of the input and the auxiliary states are

$$\gamma_\theta=R^t(\theta)SR(\theta), \quad (48)$$

$$\gamma_0=R^t(\theta')TR(\theta'), \quad (49)$$

with

$$R(\theta)=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}, \quad (50)$$

$$S=\begin{pmatrix}s^2 & 0\\ 0 & 1/s^2\end{pmatrix},\quad T=\begin{pmatrix}t^2 & 0\\ 0 & 1/t^2\end{pmatrix}, \quad (51)$$

where $s\equiv e^{-r}$ and θ ($t\equiv e^{-r'}$ and θ') are the squeezing parameter and phase of the input state (ancilla). Recall that the diagonal matrix elements of the covariance matrices give the variances of the canonical observables $\langle\Delta Q\rangle^2=[\gamma_\theta]_{11}/2$ and $\langle\Delta P\rangle^2=[\gamma_\theta]_{22}/2$.

From (47), the probability density is

$$p(\chi|\theta)=\frac{\sqrt{\det M(\theta)}}{\pi}\exp[-\chi^t M(\theta)\chi]. \quad (52)$$

The Fisher information of such a measurement can be readily computed from (9):

$$F(\theta)=\frac{1}{2}\mathrm{tr}[M'M^{-1}M'M^{-1}], \quad (53)$$

where M' denotes the derivative of $M(\theta)$ wrt θ and M^{-1} its inverse. The θ dependency has been dropped for readability reasons. Further manipulations show that the Fisher information can be written as

$$F(\theta)=\frac{1}{2}\mathrm{tr}[\Gamma^{-1}\Sigma\Gamma^{-1}\Sigma], \quad (54)$$

with

$$\Gamma=S+R^t(\theta'-\theta)TR(\theta'-\theta)=(\cosh 2r+\cosh 2r')\mathbb{1} -(\sinh 2r+\sinh 2r'\cos\varphi)\boldsymbol{\sigma}_3+\sinh 2r'\sin\varphi\boldsymbol{\sigma}_1 \quad (55)$$

and

$$\Sigma=R^t(\pi/2)S+SR(\pi/2)=2\sinh 2r\boldsymbol{\sigma}_1, \quad (56)$$

where, as expected, $\varphi=2(\theta'-\theta)$ is the only relevant angular parameter, and $\boldsymbol{\sigma}_i$ $(i=1,2,3)$ are the standard Pauli matrices. With this, $F(\theta)$ reads as

$$F(\theta)=\frac{2\sinh^2 2r}{(\cosh 2r\cosh 2r'-\sinh 2r\sinh 2r'\cos\varphi+1)^2}\times(\cosh 2r\cosh 2r'-\sinh 2r\sinh 2r'\cos\varphi +\sinh^2 2r'\sin^2\varphi+1). \quad (57)$$

Here the dependence of φ on θ is implicit. To optimize the measurement, we find the values of θ' and r' that maximize the Fisher information. Imposing $\partial F/\partial\varphi=0$, we find three possible extremal points: $\varphi=0$ and $\varphi=\pm\varphi_0(s,t)$ [see Eq. (60) for the expression of φ_0]. We will drop the arguments of φ_0, where no confusion arises. The angle $\varphi=0$ is a maximum if the squeezing in the auxilliary port of the beamsplitter is weaker than a threshold, $1<t<t_{\mathrm{thr}}(s)$, given by

$$t_{\mathrm{thr}}(s)=\frac{1}{2s}\sqrt{s^4-1+\sqrt{s^8+14s^4+1}}. \quad (58)$$

In this case, this is the only extremal point of $F(\theta)$. When $t>t_{\mathrm{thr}}$ the trivial solution $\varphi=0$ becomes a minimum and the solutions $\varphi=\pm\varphi_0$ become maxima.

The maximum Fisher information below the threshold $(1<t<t_{\mathrm{thr}})$ is

$$F(\theta')=\frac{t^2(1-s^4)^2}{s^2(s^2+t^2)^2}=\frac{\sinh^2 2r}{\cosh^2(r-r')}. \quad (59)$$

As a function of t, the maximum of $F(\theta)$ is at the critical point $t=t_{\mathrm{thr}}$.

Over the threshold $(t>t_{\mathrm{thr}})$, the optimal angle is given by

$$\cos\varphi_0=\frac{2\cosh 4r'\sinh 2r+\cosh 2r'\sinh 4r}{(3+\cosh 4r)\sinh 2r'+2\cosh 2r\sinh 4r'}, \quad (60)$$

$\varphi_0\geqslant 0$ [recall the definition of s and t right after Eq. (51)]. According to the definition of φ, we have $\theta'=\theta\pm\varphi_0/2$. The maximum Fisher information is

$$F(\theta'\mp\varphi_0/2) =\sinh^2 2r\frac{3+\cosh 4r+8\cosh 2r\cosh 2r'+4\cosh 4r'}{4(\cosh 2r+\cosh 2r')^2}. \quad (61)$$

As $t\to\infty$ $(r'\to-\infty)$, the optimal angle becomes

$$\Phi(s)=\lim_{t\to\infty}\varphi_0(s,t)=\arccos\left(\frac{s^4-1}{s^4+1}\right)=\arccos(-\tanh 2r), \quad (62)$$

and the Fisher information approaches the value

$$\lim_{t\to\infty}F(\theta'\mp\varphi_0/2)=\frac{(1-s^4)^2}{2s^4}=\cosh 4r-1. \quad (63)$$

This is precisely the Heisenberg limit found in Sec. II [see also Eqs. (25) and (26)]. Hence, optimal phase sensitivity is attained with a POVM measurement given by Eq. (46) and an auxiliary state σ_0 characterized by $t\to\infty$ and $\theta'=\theta\pm\Phi/2$. This is an infinitely squeezed state phase-shifted wrt the true phase. In Sec. VIII we present a two-step adaptive implementation of this measurement that can be derived proceeding along the same lines as in Sec. V. This scheme also attains the Heisenberg limit.

VII. INTERPRETATION OF THE OPTIMAL GAUSSIAN MEASUREMENT

In this section we provide an interpretation of this apparently unphysical proposal of a Gaussian measurement with infinite squeezing in the auxiliary port. Introducing an infinitely squeezed ancilla in this port is equivalent to having zero noise in one quadrature, say P', at the expense of introducing infinite noise in the correspondingly conjugate one, Q'. This makes the readings of Q' to be just random noise with no information about the signal state. Therefore, one has only a measurement of P'. This strongly supports the idea that, after all, only homodyne detection is needed to implement the measurement under discussion. Later, we provide a mathematical justification of this interpretation, and show that the marginal POVM elements, i.e., those obtained by integrating Eq. (46) over q', become exactly the rank-one projectors of the P' measurement.

Let us start by giving the precise definition of the observables Q' and P':

$$Q' = U(\theta')QU^{\dagger}(\theta') = Q\cos\theta' - P\sin\theta', \tag{64}$$

$$P' = U(\theta')PU^{\dagger}(\theta') = Q\sin\theta' + P\cos\theta'. \tag{65}$$

This is just a rotation in phase space, thus the coordinates transform in the same manner,

$$\chi' = R(\theta')\chi, \tag{66}$$

and the displacement operators $D(\chi)$ transform as

$$D'(\chi') \equiv U(\theta')D(\chi')U^{\dagger}(\theta') = D(\chi). \tag{67}$$

Let $|Q;q\rangle$ and $|Q';q\rangle$ be the eigenstates of Q and Q', respectively, where q is the eigenvalue, so that $|Q';q\rangle = U(\theta')|Q;q\rangle$. Since σ_0 is

$$\sigma_0 = U(\theta')S(r')|0\rangle\langle 0|S^{\dagger}(r')U^{\dagger}(\theta'), \tag{68}$$

the POVM elements of Eq. (46) can be cast as

$$\begin{aligned} E(\chi) &= \frac{1}{2\pi}U(\theta')D(\chi')S(r')|0\rangle\langle 0|S^{\dagger}(r')D^{\dagger}(\chi')U^{\dagger}(\theta') \\ &= \int \frac{dq_1\,dq_2}{2\pi}|Q';q_1\rangle\langle Q';q_2| \\ &\quad\times e^{ip'(q_1-q_2)}\psi(q_1-q')\psi^*(q_2-q'), \end{aligned} \tag{69}$$

where we have introduced the squeezed-vacuum wave function $\psi(q)$, defined as $\psi(q) = \langle Q;q|S(r')|0\rangle$.

We can compute the marginal POVM elements $F(p')$ by integrating $E(\chi)$ over dq',

$$\begin{aligned} F(p') &= \int dq'\, E(\chi) \\ &= \int dq'\, E\big(R^{-1}(\theta')\chi'\big) \\ &= \int \frac{dq_1\,dq_2}{2\pi}|Q';q_1\rangle\langle Q';q_2| \\ &\quad\times e^{ip'(q_1-q_2)}\int dq'\,\psi(q_1-q')\psi^*(q_2-q'). \end{aligned} \tag{70}$$

Performing the dq' integral yields

$$\int dq'\,\psi(q_1-q')\psi^*(q_2-q') = \exp\left(-\frac{e^{2r'}}{4}(q_1-q_2)^2\right), \tag{71}$$

where we have used that [32]

$$\psi(q) = \frac{e^{r'/2}}{\pi^{1/4}}\exp\left(-\frac{e^{2r'}}{2}q^2\right). \tag{72}$$

With this, the marginal POVM reads as

$$\begin{aligned} F(p') &= \int \frac{dq_1\,dq_2}{2\pi}|Q';q_1\rangle\langle Q';q_2| \\ &\quad\times\exp\left(ip'(q_1-q_2) - \frac{e^{2r'}}{4}(q_1-q_2)^2\right), \end{aligned} \tag{73}$$

which in the limit $r' \to -\infty$ converges to

$$F(p') = |P';p'\rangle\langle P';p'|. \tag{74}$$

By rotating the canonical operators, Q' and P' we have diagonalized the covariance matrix of σ_0, and got rid of the correlations between Q and P. We thus put all the noise introduced by the auxiliary state into the Q' quadrature, and put all the information of the signal state into the P' quadrature. This enables us to use the marginal POVMs without loss of information, which in turn means that it suffices to measure the P' quadrature, i.e., the informative one.

Another interpretation follows from [29], where it is shown that introducing a squeezed state in the auxiliary port is equivalent to leaving the vacuum and tuning the beamsplitter to a transmittance/reflectance other than 50%/50%. The limit of infinite squeezing amounts to having a completely transmitting or reflecting mirror, which again converts heterodyning into homodyning.

VIII. ESTIMATION SCHEME

The analysis carried out so far tells us that optimal phase measurements can be performed by means of homodyne tomography of the appropriate quadrature $\theta' = \hat{\theta}_0 \pm \Phi(s)/2$, where $\Phi(s)$ is given by Eq. (62) and $\hat{\theta}_0$ is the estimate of the first step measurement, which is not assumed to be optimal. Let us call this quadrature P'. The outcomes will be Gaussian distributed with zero mean (recall that our signal state is a squeezed vacuum) and variance given by

$$\sigma^2(\theta',\theta) = \frac{1+s^4+(1-s^4)\cos 2(\theta'-\theta)}{4s^2}, \quad (75)$$

as can be seen by diagonalizing the M matrix in Eq. (52) through the transformation $\chi' = R(\theta')\chi$.

It is straightforward to check that the Fisher information provided by this distribution is exactly that given by Eq. (63).

The maximum likelihood analysis of the data yields the condition $\sigma^2(\theta',\hat{\theta}_{\mathrm{MLE}}) = \Sigma p_i^2/\bar{N}$, where $\{p_1,\ldots,p_{\bar{N}}\}$ is the set of outcomes corresponding to the $\bar{N}$ measurements of the second step. Thus

$$\hat{\theta}_{\mathrm{MLE}} = \theta' \mp \frac{1}{2}\arccos\left(\frac{4s^2\sum p_i^2/\bar{N} - 1 - s^4}{1-s^4}\right), \quad (76)$$

and the MLE is twice degenerate. There is, however, a trivial way to break this degeneracy by using the outcomes of the first step. The prescription is to choose the solution closest to the rough guess $\hat{\theta}_0$. This estimator has an asymptotically vanishing bias, whereas that corresponding to the other solution has a constant bias that goes roughly as

$$\mathbb{E}[\hat{\theta}] - \theta \simeq \arccos\left(\frac{4s^2\sigma^2(\theta',\theta) - 1 - s^4}{1-s^4}\right). \quad (77)$$

Therefore, as the prior estimation gets accurate with increasing $\bar{N}$, the separation of the two maxima of the likelihood function remain constant. By choosing the maximum closest to $\hat{\theta}_0$, one has an exponentially small error probability.

In summary, the optimal scheme goes as follows:

(1) Perform *any* phase measurement (not necessarily optimal) on N^α $(1/2 < \alpha < 1)$ copies of the signal state. From the outcomes, compute a preliminary estimation $\hat{\theta}_0$, which has a typical error $\delta\theta^2 \simeq 1/N^\alpha$.

(2) Measure the quadrature with $\theta' = \hat{\theta}_0 + \Phi(s)/2$, where $\Phi(s)$ is given in Eq. (62), on the remaining $\bar{N} = N - N^\alpha$ copies. This corresponds to setting the phase reference (the local oscillator) to θ'.

(3) The Maximum Likelihood Estimator is obtained by choosing the minus sign in Eq. (76).

Note that one could equivalently write $\theta' = \hat{\theta}_0 - \Phi(s)/2$ and choose the plus sign in Eq. (76). The variance of this estimator goes as

$$\mathrm{Var}[\hat{\theta}_{\mathrm{MLE}}] \sim \frac{1}{F} = \frac{2s^4}{(1-s^4)^2} = \frac{1}{8(\langle n\rangle^2 + \langle n\rangle)}, \quad (78)$$

and its bias vanishes asymptotically; therefore its MSE goes as

$$\mathrm{MSE}(\hat{\theta}_{\mathrm{MLE}}) \sim \frac{1}{8(\langle n\rangle^2 + \langle n\rangle)}, \quad (79)$$

which is the Heisenberg limit presented in Sec. II.

IX. CONCLUSIONS

We have seen that the phase-number Heisenberg inequality is valid, regardless of the existence of any phase operator, provided $\langle\Delta\theta\rangle^2$ is regarded as a phase estimation variance. It is just a consequence of the Cramér-Rao and Braunstein-Caves inequalities.

We have seen that asymptotically the Heisenberg limit (optimal estimation) can be attained by means of a (three-outcome) POVM adaptive measurement. The physical implementation of such a measurement is a demanding open problem. We have introduced an *approximate* MLE that provides an outstanding simplification of the maximum likelihood analysis. Asymptotically this approximate MLE and the exact one are shown to be equivalent.

We have shown that the Heisenberg limit can be also attained asymptotically by means of dyne measurements and, surprisingly, only one quadrature needs to be measured. This provides a remarkable economy of resources as compared to POVM measurements such as an eight-port homodyne detection with squeezing in the auxiliary port. The relative phase between the Local Oscillator in homodyne detection and the signal state has been computed and used to devise a two-step phase estimation scheme that attains the Heisenberg limit asymptotically.

Remarkably enough, optimality is achieved with local measurements. Hence, collective measurements prove of little use in the large sample limit. This is a consequence of quantum estimation theory applied to one-parameter problems.

Most of the schemes available so far succeed in attaining the behavior $\mathrm{Var}[\hat{\theta}] \sim \langle n\rangle^{-2}$ for $\langle n\rangle$ large enough. The exact behavior is usually not considered in the literature. This is, in our opinion, an important loophole in phase estimation, since one is usually interested in low energy states, where $\langle n\rangle$ is small. An exception is the paper by Yurke *et al.* [33], in which they introduce a scheme whose variance comes close to our bounds for large $\langle n\rangle$ [34]. However, our scheme is asymptotically optimal and beats, to the extent of our knowledge, any proposal presented up to this day.

ACKNOWLEDGMENTS

The early stage of this research was carried out at the Max-Planck Institut für Quantenoptik under the advice of Professor I. Cirac. Fruitful discussions with him, M. Wolf, and their Theory Group at the MPQ are gratefully acknowledged. Also, discussions with E. Bagan, R. Muñoz-Tapia, J. Calsamiglia, and M. Mitchell have proved fruitful.

APPENDIX A: AVERAGING OVER THE NUMBER OF INFORMATIVE OUTCOMES N_{inf}

In this appendix, we compute the average of $\mathbb{E}_{\theta,N_{\mathrm{inf}}}[(\hat{\theta}_{\mathrm{MLE}} - \theta)^2]$ over the number of informative outcomes N_{inf}, i.e.,

$$\mathbb{E}[(\hat{\theta}_{\text{MLE}}-\theta)^2] = \sum_{N_{\text{inf}}=0}^{\bar{N}} \mathbb{E}_{N_{\text{inf}}}[(\hat{\theta}_{\text{MLE}}-\theta)^2]\binom{\bar{N}}{N_{\text{inf}}} p^{N_{\text{inf}}}(1-p)^{\bar{N}-N_{\text{inf}}} = \frac{1}{4\langle\Delta n\rangle^2}\sum_{N_{\text{inf}}=1}^{\bar{N}}\frac{1}{N_{\text{inf}}}\binom{\bar{N}}{N_{\text{inf}}} p^{N_{\text{inf}}}(1-p)^{\bar{N}-N_{\text{inf}}} + \delta\theta^2(1-p)^{\bar{N}}. \tag{A1}$$

The first sum runs from 1 to $\bar{N}$, thus taking into account that the actual variance for the rare case $N_{\text{inf}}=0$ is considered separately. We can compute the expectation value $\langle 1/N_{\text{inf}}\rangle$ by expanding $1/N_{\text{inf}}$ around the expectation value of N_{inf}, which is $\langle N_{\text{inf}}\rangle = p\bar{N}$, and keeping terms up to second order. We obtain

$$\sum_{N_{\text{inf}}=1}^{\bar{N}}\frac{1}{N_{\text{inf}}}\binom{\bar{N}}{N_{\text{inf}}} p^{N_{\text{inf}}}(1-p)^{\bar{N}-N_{\text{inf}}} = \frac{1}{\bar{N}p}\left(1+\frac{1-p}{\bar{N}p}-(1-p)^{\bar{N}}\right). \tag{A2}$$

Inserting the value for $p=1+O(\delta\theta^2)$ yields

$$\mathbb{E}[(\hat{\theta}_{\text{MLE}}-\theta)^2] = \frac{1}{4\langle\Delta n\rangle^2\bar{N}} + O(\delta\theta^2). \tag{A3}$$

---

[1] W. Schelich and S. Barnett, *Quantum Phase and Phase Dependent Measurements*, Phys. Scr., T48 (1993).
[2] H. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, New York, 1982).
[3] A. Holevo, Lect. Notes Math. **1055**, 153, (1984).
[4] H. Yuen and M. Lax, IEEE Trans. Inf. Theory **IT-19**, 740 (1973).
[5] C. Helstrom, IEEE Trans. Inf. Theory **IT-14**, 234 (1968).
[6] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[7] D. Berry, H. M. Wiseman, and J. Breslin, Phys. Rev. A **63**, 053804 (2001).
[8] D. W. Berry, Ph.D. dissertation, The University of Queensland (2002), available at http://eprint.uq.edu.au/archive/00000972/
[9] L. Susskind and J. Glogower, Physics (N.Y.) **1**, 49 (1964).
[10] J. H. Shapiro, S. R. Shepard, and N. C. Wong, Phys. Rev. Lett. **62**, 2377 (1989).
[11] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (1991).
[12] D. T. Pegg and S. M. Barnett, Phys. Rev. A **39**, 1665 (1989).
[13] S. L. Braunstein, Phys. Rev. Lett. **69**, 3598 (1992).
[14] D. T. Smithey, M. Beck, J. Cooper, and M. C. Raymer, Phys. Rev. A **48**, 3159 (1993).
[15] M. Freyberger and W. Schleich, Phys. Rev. A **49**, 5056 (1994).
[16] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).
[17] H. N. Wiseman and R. B. Killip, Phys. Rev. A **57**, 2169 (1998).
[18] Z. Y. Ou, Phys. Rev. A **55**, 2598 (1997).
[19] S. Braunstein, C. Caves, and G. Milburn, Ann. Phys. **247**, 135 (1996).
[20] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[21] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).
[22] M. Hayashi and K. Matsumoto, e-print quant-ph/0308150 (2003), English translation of the manuscript that appeared in Surikaiseki Kenkyusho Kokyuroku No. 1055 (1998).
[23] K. Matsumoto, J. Phys. A **35**, 3111 (2002).
[24] O. Barndorff-Nielsen and R. Gill, J. Phys. A **33**, 4481 (2000).
[25] R. Gill, math.ST/0405571, 2004.
[26] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A **62**, 023815 (2000).
[27] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
[28] E. Bagan, A. Monras, and R. Munoz-Tapia, Phys. Rev. A **71**, 062318 (2005).
[29] U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, Cambridge, 1997).
[30] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[31] M. Wolf (private communication, 2005).
[32] E. Bagan (private communication).
[33] B. Yurke, S. L. McCall, and J. R. Klauder, Phys. Rev. A **33**, 4033 (1986).
[34] See Eq. (10.39) in Ref. [33].