Phase-sensitivity bounds for two-mode interferometers

Luca Pezzè,[1] Philipp Hyllus,[2] and Augusto Smerzi[1]
[1]QSTAR, INO-CNR and LENS, Largo Enrico Fermi 2, I-50125 Firenze, Italy
[2]Department of Theoretical Physics, University of the Basque Country EHU/UPV, P.O. Box 644, E-48080 Bilbao, Spain
(Received 29 August 2014; published 3 March 2015)

We provide general bounds of phase estimation sensitivity in linear two-mode interferometers. We consider probe states with a fluctuating total number of particles. With incoherent mixtures of states with different total number of particles, particle entanglement is necessary but not sufficient to overcome the shot noise limit. The highest possible phase estimation sensitivity, the Heisenberg limit, is established under general unbiased properties of the estimator. If coherences can be created, manipulated, and detected, the Heisenberg limit can only be set in the central limit, with a sufficiently large repetition of interferometric measurements.

DOI: 10.1103/PhysRevA.91.032103 PACS number(s): 03.65.Ta, 42.50.St, 42.50.Dv

I. INTRODUCTION

The problem of determining the ultimate phase sensitivity (often tagged as the Heisenberg limit) of linear interferometers has long puzzled the field [1–14] and still raises controversies [15–36]. The recent revival of interest is triggered by the current impressive experimental efforts in the direction of quantum phase estimation with ions [37], cold atoms [38], Bose-Einstein condensates [39], and photons [40], including possible applications to large-scale gravitational wave detectors [41]. Besides the possible technological applications, the problem is closely related to fundamental questions of quantum information, most prominently, regarding the role played by quantum correlations. Several works have focused on linear two-mode interferometers. In this case, it is widely accepted that when the number of particles in the input state is fixed and equal to N, there are two important bounds in the uncertainty of unbiased phase estimation. The shot-noise limit,

$$\Delta\theta_{\rm SN} = \frac{1}{\sqrt{m\,N}}, \qquad \text{for } (\Delta\hat{N})^2 = 0, \tag{1}$$

is the maximum sensitivity achievable with probe states containing only classical correlations among particles [15,16]. The factor m accounts for the number of independent measurements repeated with identical copies of the probe state. This bound is not fundamental. It can be surpassed by preparing the N particles of the probe in proper usefully entangled states [16]. The fundamental (Heisenberg) limit is given by

$$\Delta\theta_{\rm HL} = \frac{1}{\sqrt{m}\,N}, \qquad \text{for } (\Delta\hat{N})^2 = 0, \tag{2}$$

and it saturated [15,16] by maximally entangled (NOON) states [42–46].

It should be noticed that most of the theoretical investigations have been developed in the context of systems having a fixed, known, total number of particles. Yet, many experiments are performed in presence of finite number fluctuations, $(\Delta\hat{N})^2 > 0$. The consequences of such fluctuations (which may have a classical or a quantum nature) have not been investigated in great depth. It has been shown [17,18] that phase uncertainty bounds can be critically affected by the presence of coherences between different total number of particles in the probe state and/or the output measurement. However such quantum coherences do not play any role in two experimentally relevant cases: (i) in the presence of superselection rules, which are especially relevant for massive particles and forbid the existence of number coherences in the probe state; (ii) when the phase shift is estimated by measuring an arbitrary function of the number of particles in the output state of the interferometer, e.g., when the total number of particles is postselected by the measurement apparatus. The point (ii) is actually a ubiquitous condition in current atomic and optical phase estimation experiments.

In the absence of number coherences, or when coherences are present but irrelevant because of (ii), we can define a state as separable if it is separable in each subspace of a fixed number of particles [17]. A state is entangled if it is entangled in at least one subspace of fixed number of particles. With separable states, the maximum sensitivity of unbiased phase estimators is bounded by the shot noise

$$\Delta\theta_{\rm SN} = \frac{1}{\sqrt{m\,\langle\hat{N}\rangle}}, \qquad \text{for } (\Delta\hat{N})^2 > 0, \tag{3}$$

while the Heisenberg limit is given by [17]

$$\Delta\theta_{\rm HL} = \max\left[\frac{1}{\sqrt{m\langle\hat{N}^2\rangle}}, \frac{1}{m\langle\hat{N}\rangle}\right], \quad \text{for } (\Delta\hat{N})^2 > 0. \tag{4}$$

We point out that Eq. (4) cannot be obtained from Eq. (2) by simply replacing N with $\langle\hat{N}\rangle$ [47]. On the other hand, Eq. (4) reduces to Eq. (2) when number fluctuations vanish, $\langle\hat{N}^2\rangle = \langle\hat{N}\rangle^2 = N^2$. An example of phase estimation saturating the scaling $1/m\langle\hat{N}\rangle$ is obtained with the coherent cross product squeezed-vacuum probe state [50].

When the probe state and the output measurement contain number coherences, the situation becomes more involved. The notion of particle separability and entanglement becomes hazy since the number of parties stay in a quantum superposition. One can attempt to extend the notion of separability saying that a state is separable if the projection over each fixed-N subspace is separable [17]. However, within this definition, one loses the relation between separability and shot-noise limit: it is not difficult to find separable states that can achieve an arbitrary high sensitivity. Regarding the Heisenberg limit, it is still possible to show that Eq. (4) holds in the central limit ($m \gg 1$), at least. Outside the central limit (i.e., for a small number of measurements), we can only set the bound

$$\Delta\theta \geqslant \frac{1}{\sqrt{m\langle\hat{N}^2\rangle}}. \quad (5)$$

The crucial point is that the fluctuations $\langle\hat{N}^2\rangle$ can be made arbitrarily large even with a finite $\langle\hat{N}\rangle$ and we cannot rule out to the possibility that $\Delta\theta < 1/m\langle\hat{N}\rangle$. In general, no lower bound in terms of $\langle\hat{N}\rangle$ can be settled in this case.

This paper extends and investigates in detail the results and concepts introduced in Ref. [17]. In Sec. II we review the theory of multiparameter estimation with special emphasis on two-mode linear transformations. In Sec. III we discuss the relation between multiparticle entanglement of the probe state and the achievable phase sensitivity. In Sec. IV we show under which conditions the Heisenberg limit, Eq. (4), holds. We finally give an overview of the phase-sensitivity bounds recently discussed in the literature.

This paper focuses on the ideal noiseless case. It is worth pointing out that decoherence can strongly affect phase-uncertainty bounds. For several relevant noise models in quantum metrology as, for instance, particles losses, correlated or uncorrelated phase noise, phase-uncertainty bounds have been derived [43,51–56].

II. BASIC CONCEPTS

In the multiphase estimation problem, a probe state $\hat{\rho}$ undergoes a transformation that depends on the unknown vector parameter $\boldsymbol{\theta}$. The phase shift is estimated from measurements done on the transformed state $\hat{\rho}_{\text{out}}(\boldsymbol{\theta})$. The protocol is repeated m times by preparing identical copies of $\hat{\rho}$ and performing identical transformations and measurements. The aim of this section is to review the general theory of phase estimation for linear and lossless two-mode interferometers.

A. Probe state

A generic probe state with fluctuating total number of particles can be written as

$$\hat{\rho}_{\text{coh}} = \sum_k p_k |\psi_k\rangle\langle\psi_k| \quad (6)$$

with $p_k > 0$ and $\sum_k p_k = 1$, where

$$|\psi_k\rangle = \sum_N \sqrt{Q_{N,k}}\, |\psi_{N,k}\rangle \quad (7)$$

is a coherent superposition of states $|\psi_{N,k}\rangle$ with different number of particles. The coefficients $Q_{N,k}$ are complex numbers and the normalization condition $\langle\psi_k|\psi_k\rangle = 1$ implies $\sum_N |Q_{N,k}| = 1$. It is generally assumed that quantum coherences between states of different numbers of particles do not play any observable role [57]. In particular, with massive particles this is the consequence of superselection rules (SSRs) for the total number of particles [58]. In the presence of SSRs the only physical states are those that commute with the number of particles operator,

$$[\hat{\rho},\hat{N}] = 0. \quad (8)$$

A state satisfies this condition if and only if [59] it can be written as the incoherent mixture

$$\hat{\rho}_{\text{inc}} = \sum_N Q_N\, \hat{\rho}^{(N)}, \quad (9)$$

where $\hat{\rho}^{(N)}$ is a normalized ($\text{Tr}[\hat{\rho}^{(N)}] = 1$) state, $Q_N = \text{Tr}[\hat{\pi}_N \hat{\rho} \hat{\pi}_N]$ are positive numbers satisfying $\sum_N Q_N = 1$ and $\hat{\pi}_N$ are projectors on the fixed-N subspace. The existence of SSRs is the consequence of the lack of a phase reference frame (RF) [57]. However, the possibility that a suitable RF can be established in principle cannot be excluded [57]. If SSRs are lifted, then coherent superpositions of states with different numbers of particles become physically relevant.

B. Two-mode transformations

In the following we focus on linear lossless transformations involving two modes. These cover a large class of optical and atomic passive operations, including the beam splitters, Mach-Zehnder and Ramsey interferometers [11]. Most of the current phase-estimation experiments [37–40] are well described within a two-mode model.

Denoting by $\hat{a}_1$ and $\hat{a}_2$ ($\hat{b}_1$ and $\hat{b}_2$) the input (output) mode annihilation operators, we can generally write the mode transformation as

$$\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \mathbf{U} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}, \quad (10)$$

where $\mathbf{U}$ is a 2×2 matrix [3,60,61]. By imposing the conservation of the total number of particles, $\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2 = \hat{b}_1^\dagger\hat{b}_1 + \hat{b}_2^\dagger\hat{b}_2$, we obtain that $\mathbf{U}$ can be explicitly written as

$$\mathbf{U} = e^{-i\phi_0} \begin{bmatrix} e^{-i\phi_t}\cos\frac{\vartheta}{2} & -e^{-i\phi_r}\sin\frac{\vartheta}{2} \\ e^{i\phi_r}\sin\frac{\vartheta}{2} & e^{i\phi_t}\cos\frac{\vartheta}{2} \end{bmatrix}. \quad (11)$$

The matrix (11) is unitary, preserves bosonic and fermionic commutation relations between mode operators and its determinant is equal to $e^{-2i\phi_0}$. The most general two-mode transformation thus belongs to the $\text{U}(2) = \text{U}(1)\times\text{SU}(2)$ group (unitary matrices with determinant $|\det\mathbf{U}| = 1$). The coefficients ϑ is physically related to transmittance $t = \cos^2\vartheta/2$ and reflectance $r = \sin^2\vartheta/2$ of the transformation (11), ϕ_t and ϕ_r being the corresponding phases. The lossless nature is guaranteed by $t + r = 1$.

Using the Jordan-Schwinger representation of angular momentum systems in terms of mode operators [62], it is possible to find the operator $\hat{\text{U}}$ corresponding to the matrix (11). In other words, $\hat{b}_i = \hat{\text{U}}^\dagger\hat{a}_i\hat{\text{U}}$ for $i = 1,2$ is the transformation of mode operators (Heisenberg picture) and $\hat{\rho}_{\text{out}} = \hat{\text{U}}\hat{\rho}\hat{\text{U}}^\dagger$, $|\psi_{\text{out}}\rangle = \hat{\text{U}}|\psi_{\text{in}}\rangle$, is the equivalent transformation of statistical mixtures and quantum states, respectively (Schrödinger picture). One finds [3,60]

$$\hat{\text{U}}(\phi_0,\theta) = e^{-i\phi_0\hat{N}}\, e^{-i\theta\hat{J}_{\boldsymbol{n}}}, \quad (12)$$

where

$$\hat{N} = \hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2$$

is the number of particle operator,

$$\hat{J}_x = \frac{\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1}{2},$$
$$\hat{J}_y = \frac{\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_2^\dagger \hat{a}_1}{2i},$$
$$\hat{J}_z = \frac{\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2}{2},$$

are pseudospin operators, and $\hat{J}_{\boldsymbol{n}} = \alpha \hat{J}_x + \beta \hat{J}_y + \gamma \hat{J}_z$ (where α, β, and γ are the coordinates of the vector $\boldsymbol{n}$ in the Bloch sphere and satisfy $\alpha^2+\beta^2+\gamma^2=1$). The exact relation between the parameters of the matrix $\mathbf{U}$ [ϕ_τ, ϕ_ρ and ϑ in Eq. (11)] and those of the operator $\hat{U}$ [θ, α, β and γ in Eq. (12)] is given in Appendix A. The operators $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ satisfy angular momentum commutation relations. Since $[\hat{J}_{\boldsymbol{n}},\hat{N}]=0$, we can thus rewrite $\hat{J}_{\boldsymbol{n}} = \oplus_N \hat{J}_{\boldsymbol{n}}^{(N)}$, where $\hat{J}_{\boldsymbol{n}}^{(N)} = \hat{\pi}_N\, \hat{J}_{\boldsymbol{n}}\, \hat{\pi}_N$. Furthermore, $\hat{J}_{\boldsymbol{n}}^{(N)} = \sum_{l=1}^{N} \hat{\sigma}_{\boldsymbol{n}}^{(l)}/2$, where $\hat{\sigma}_{\boldsymbol{n}}^{(l)}$ is the Pauli matrix (along the direction $\boldsymbol{n}$ in the Bloch sphere, $\hat{\sigma}_{\boldsymbol{n}}^{(l)} = \alpha\hat{\sigma}_x^{(l)} + \beta\hat{\sigma}_y^{(l)} + \gamma\hat{\sigma}_z^{(l)}$) acting on the lth particle.

Equation (12) can be rewritten as

$$\hat{\mathrm{U}}(\theta_1,\theta_2) = e^{i\chi \hat{J}_s}[e^{-i\theta_1 \hat{a}_1^\dagger \hat{a}_1} e^{-i\theta_2 \hat{a}_2^\dagger \hat{a}_2}] e^{-i\chi \hat{J}_s}, \qquad (13)$$

where $\theta_1 = \phi_0 + \theta/2$, $\theta_2 = \phi_0 - \theta/2$, $\boldsymbol{s}$ is a direction perpendicular to $\boldsymbol{z}$ and $\boldsymbol{n}$, and $\cos\chi = \boldsymbol{n}\cdot\boldsymbol{z}$. θ_1 and θ_2 can be identified as the phases acquired in each mode a_1 and a_2 inside a Mach-Zehnder-like interferometer [with standard balanced beam splitters replaced by the transformation $e^{\pm i\chi \hat{J}_s}$, see Fig. 1(a)]. Both phases may be unknown. When setting one of the two phases to zero (or to any fixed known value), Eq. (13) reduces to different single-phase transformations:

(i) SU(2) transformations $e^{-i\theta \hat{J}_{\boldsymbol{n}}}$ ($\phi_0=0$) or, equivalently

$$\hat{\mathrm{U}}(\theta) = e^{+i\chi \hat{J}_s} e^{-i\theta \hat{J}_z} e^{-i\chi \hat{J}_s}, \qquad (14)$$

with notation analogous to Eq. (13) [see also Fig. 1(b)]. This depends only on the relative phase shift $\theta = \theta_1 - \theta_2$ among the two interferometer modes. This encompasses the beam splitter $e^{-i\theta \hat{J}_x}$, the relative phase shift $e^{-i\theta \hat{J}_z}$ and the Mach-Zehnder $e^{-i\theta \hat{J}_y}$ transformations.

(ii) U(1) transformations $e^{-i\phi_0 \hat{N}}$ ($\theta=0$), which can be understood as a phase shift equally imprinted on each of the two modes: $e^{-i\phi_0 \hat{N}} = e^{-i\phi_0 \hat{a}_1^\dagger \hat{a}_1} \otimes e^{-i\phi_0 \hat{a}_2^\dagger \hat{a}_2}$.

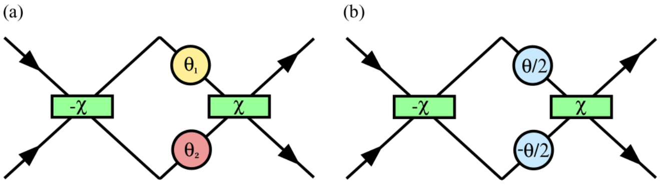

FIG. 1. (Color online) Schematic representation of (a) U(2) and (b) SU(2) interferometers. In the U(2) case, the general transformation is given by Eq. (13), where a_1 and a_2 are the modes inside the interferometer and the green squares represent $e^{\pm i\chi \hat{J}_s}$. In the SU(2) case, the general transformation is given by Eq. (14) and depends on the relative phase shift θ among the a_1 and a_2 modes.

C. Output measurement

A general quantum measurement scenario can be described by a set of bounded, non-negative Hermitian operators $\{\hat{E}(\varepsilon)\}_\varepsilon$ parametrized by ε and satisfying the completeness relation $\int d\varepsilon \hat{E}(\varepsilon) = \mathbb{1}$ [63]. The quantity ε labels the possible results of a measurement, which can be continuous (as here), discrete, or multivariate. The set of operators $\{\hat{E}(\varepsilon)\}_\varepsilon$ constitutes a positive-operator valued measure (POVM). Projective measurements are a special POVM class satisfying $\hat{E}(\varepsilon)\hat{E}(\varepsilon') = \hat{E}(\varepsilon)\delta(\varepsilon-\varepsilon')$. In the case of projective measurement, ε labels the possible eigenvalues of the measurement observable. In general, each outcome ε is characterized by a probability $P(\varepsilon|\boldsymbol{\theta}) = \mathrm{Tr}[\hat{E}(\varepsilon)\hat{\rho}_{\mathrm{out}}(\boldsymbol{\theta})]$, conditioned by the true value of the parameters. The positivity and Hermiticity of $\{\hat{E}(\varepsilon)\}_\varepsilon$ guarantee that $P(\varepsilon|\boldsymbol{\theta})$ are real and non-negative, the completeness guarantees that $\int d\varepsilon P(\varepsilon|\boldsymbol{\theta}) = 1$.

Generally speaking, a POVM $\{\hat{E}(\varepsilon)\}_\varepsilon$ may or may not contain coherences among different number of particles. A POVM does not contain number coherences if and only if all its elements $\hat{E}(\varepsilon)$ commute with the number of particles operator:

$$[\hat{E}(\varepsilon),\hat{N}] = 0, \quad \forall\varepsilon. \qquad (15)$$

or, equivalently [64],

$$\hat{E}(\varepsilon) = \sum_N \hat{E}_N(\varepsilon), \qquad (16)$$

where $\hat{E}_N(\varepsilon) \equiv \hat{\pi}_N\, \hat{E}(\varepsilon)\, \hat{\pi}_N$ acts on the fixed-N subspace and $\hat{\pi}_N$ are projectors. In the following we show that the class of incoherent POVMs (16) includes all measurements that are currently done experimentally.

In current phase estimation experiments, the phase shift is generally estimated by measuring an observable, which is a function $\hat{f}(\hat{N}_1,\hat{N}_2)$ of the particle number operators at the output ports of the interferometer. We can write $\hat{f}(\hat{N}_1,\hat{N}_2) = \sum_\varepsilon \varepsilon\, \hat{E}(\varepsilon)$, where $\hat{E}(\varepsilon)$ indicates the projector operator on the eigenstate of $\hat{f}(\hat{N}_1,\hat{N}_2)$ with eigenvalue $f(N_1,N_2)$. More explicitly,

$$\hat{E}(\varepsilon) = \sum_{N_1,N_2} \delta[f(N_1,N_2)-\varepsilon]\, |N_1,N_2\rangle\langle N_1,N_2|, \qquad (17)$$

which, by the change of variables $N = N_1 + N_2$ and $\mu = (N_1 - N_2)/2$ ($-N/2 \leqslant \mu \leqslant N/2$), rewrites as

$$\hat{E}(\varepsilon) = \sum_N \sum_\mu \delta[f(N,\mu)-\varepsilon]\, |N,\mu\rangle\langle N,\mu|. \qquad (18)$$

This POVM element (a projection operator, in this case) has precisely the form of Eq. (16). Notice that the information about the total number of particles is not necessarily included.

For instance, the POVM corresponding to the measurement of the relative number of particles can be written as $[f(N,\mu) = \mu]$

$$\hat{E}(\mu) = \sum_{N \geqslant 2\mu} |N,\mu\rangle\langle N,\mu|,$$

which, again, has the form of Eq. (16). Formally, we can write the relative number of particles operator as

TABLE I. The table summarizes the general two-mode transformation group for the phase estimation problem. The U(2) group is only relevant when number coherences are present in both the probe state and POVM.

	POVM with coherences	POVM without coherences
$\hat{\rho}$ with coherences	U(2)	SU(2)
$\hat{\rho}$ without coherences	SU(2)	SU(2)

$\hat{J}_z = \oplus_N \hat{J}_z^{(N)} = \sum_\mu \mu \hat{E}(\mu)$. This example can be straightforwardly generalized to any function of the relative particle number operator. Another example is the measurement of the number of particles in a single mode (for instance the "1" mode), we have $[f(N,\mu) = N/2 + \mu]$

$$\hat{E}(N_1) = \sum_{N,\mu} \delta[N/2 + \mu - N_1]\, |N,\mu\rangle\langle N,\mu|.$$

We recover a POVM of the form of Eq. (16) also in this case. Analogous results hold for any function of N_1 (or N_2), for instance the measurement of the parity operator [42,44,65,66] at one output port $[f(N,\mu) = (-1)^{N/2\pm\mu}]$.

We conclude this subsection pointing out that, given the probe state and interferometric transformation, there are good and bad choices of POVM. The worse ones are those for which the conditional probabilities of output results (see below) do not depend on the phase shift(s). In general, the adequacy of a certain POVM set for phase estimation is quantifies by the Fisher information (see Sec. II E): the best POVM being the one (not necessarily unique) that maximizes the Fisher information.

D. Conditional probabilities

For general U(2) transformations, Eq. (13), the conditional probabilities can be written as

$$P(\varepsilon|\theta_1,\theta_2) = \mathrm{Tr}[\hat{E}(\varepsilon)\hat{\mathrm{U}}(\theta_1,\theta_2)\hat{\rho}\hat{\mathrm{U}}^\dagger(\theta_1,\theta_2)]. \tag{19}$$

If the probe state and/or the POVM do not contain number coherences, i.e., $\hat{\rho}$ is given by Eq. (9) and/or $\hat{E}(\varepsilon)$ is given by Eq. (16), then (19) reduces to

$$P(\varepsilon|\theta) = \sum_N Q_N P(\varepsilon|N,\theta), \tag{20}$$

where $P(\varepsilon|N,\theta) = \mathrm{Tr}[\hat{E}(\varepsilon)e^{-i\theta\hat{J}_n}\hat{\rho}^{(N)}e^{+i\theta\hat{J}_n}]$. The derivation of Eq. (20) is detailed in Appendix B. Equation (20) depends only on θ, the relative phase shift among the two modes of the interferometer. We conclude that U(2) transformations are relevant only if the input state contains coherences among different number of particles and the output measurement is a POVM with coherences. In all other cases, the phase shift $e^{-i\phi_0\hat{N}}$ is irrelevant as the conditional probabilities are insensitive to ϕ_0. In this case, the mode transformation Eq. (13) restricts to the unimodular (i.e., unit determinant) subgroup SU(2). The SU(2) representation, while being not general, is widely used because, experimental measurements do not contain number coherences, as discussed in Sec. II C. Table I summarizes the general two-mode transformation group for the phase estimation problem, depending on the presence of number coherences in the probe state and/or POVM.

E. Multiphase estimation

Since U(2) transformations involve two phases, we recall here basic elements of two-parameter estimation theory [63,67]. The vector parameter $\boldsymbol{\theta} \equiv \{\theta_1,\theta_2\}$ is inferred from the values $\boldsymbol{\varepsilon} \equiv \{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_m\}$ obtained in m repeated independent measurements. The mapping from the measurement results into the two-dimensional parameter space is provided by the estimator function $\boldsymbol{\Theta}(\boldsymbol{\varepsilon}) \equiv [\Theta_1(\boldsymbol{\varepsilon}),\Theta_2(\boldsymbol{\varepsilon})]$. Its mean value is $\bar{\boldsymbol{\Theta}} \equiv [\bar{\Theta}_1,\bar{\Theta}_2]$, with $\bar{\Theta}_i = \int d\boldsymbol{\varepsilon}\, \mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})\Theta_i(\boldsymbol{\varepsilon})$ $(i = 1,2)$ and $\mathcal{L}(\boldsymbol{\varepsilon}|\theta) \equiv \prod_{l=1}^m P(\varepsilon_l|\boldsymbol{\theta})$ the likelihood function. We further introduce the covariance matrix $\mathbf{B}$ of elements

$$\mathbf{B}_{i,j} = \int d\boldsymbol{\varepsilon}\, \mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})(\Theta_i(\boldsymbol{\varepsilon}) - \bar{\Theta}_i)(\Theta_j(\boldsymbol{\varepsilon}) - \bar{\Theta}_j). \tag{21}$$

Notice that $\mathbf{B}$ is symmetric and its ith diagonal element is the variance $(\Delta\theta_i)^2$.

1. *Cramér-Rao bound*

Following a Cauchy-Schwarz inequality [68], we have [69]

$$(\boldsymbol{v}^\top \mathbf{b}\, \boldsymbol{u})^2 \leqslant m(\boldsymbol{v}^\top \mathbf{B}\boldsymbol{v})(\boldsymbol{u}^\top \mathbf{F}\boldsymbol{u}), \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in \mathbb{R}, \tag{22}$$

where $\mathbf{b}_{i,j} = \partial\bar{\Theta}_i/\partial\theta_j$ is the Jacobian matrix and

$$\mathbf{F}_{i,j} = \int d\boldsymbol{\varepsilon}\, \frac{1}{P(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}\left(\frac{\partial P(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}{\partial\theta_i}\right)\left(\frac{\partial P(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}{\partial\theta_j}\right) \tag{23}$$

the Fisher information matrix (FIM) [70], which is symmetric and non-negative definite. The FIM depends on the specific POVM via the conditional probability distribution. Note that $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{b}$ generally depend on $\boldsymbol{\theta}$ but we do not explicitly indicate this dependence, in order to simplify the notation. In the inequality (22) $\boldsymbol{u}$ and $\boldsymbol{v}$ are arbitrary real vectors: depending on $\boldsymbol{u}$ and $\boldsymbol{v}$ we thus have an infinite number of scalar inequalities. If the FIM is positive definite, and thus invertible, the specific choice $\boldsymbol{u} = \mathbf{F}^{-1}\mathbf{b}^\top\boldsymbol{v}$ in Eq. (22) leads to the vector parameter Cramér-Rao lower bound $\mathbf{B} \geqslant \mathbf{B}_{\mathrm{CR}}$ [71], in the sense that the matrix $\mathbf{B} - \mathbf{B}_{\mathrm{CR}}$ is non-negative definite [i.e., $\boldsymbol{v}^\top\mathbf{B}\boldsymbol{v} \geqslant \boldsymbol{v}^\top\mathbf{B}_{\mathrm{CR}}\boldsymbol{v}$ holds for any real vector $\boldsymbol{v}$], where

$$\mathbf{B}_{\mathrm{CR}} = \frac{\mathbf{b}\,\mathbf{F}^{-1}\,\mathbf{b}^\top}{m}. \tag{24}$$

This choice of $\boldsymbol{u}$ leads to a bound that is saturable by the maximum likelihood estimator (see Sec. II E 2) asymptotically in the number of measurements. As shown by the Cramér-Rao lower bound Eq. (24), optimal POVMs are the ones that minimize $\mathbf{F}^{-1}$.

In the two-parameter case, the FIM

$$\mathbf{F} = \begin{bmatrix} F_{1,1} & F_{1,2} \\ F_{1,2} & F_{2,2} \end{bmatrix} \tag{25}$$

is invertible if and only if $F_{1,1}F_{2,2} - F_{1,2}^2 \neq 0$, its inverse given by

$$\mathbf{F}^{-1} = \frac{1}{F_{1,1}F_{2,2} - F_{1,2}^2}\begin{bmatrix} F_{2,2} & -F_{1,2} \\ -F_{1,2} & F_{1,1} \end{bmatrix}. \tag{26}$$

Furthermore, if $\bar{\mathbf{\Theta}}_i$ does not depend on θ_j for $j \neq i$ (i.e., $\mathbf{b}$ is diagonal), the diagonal elements of $\mathbf{B}_{\mathrm{CR}}$ satisfy:

$$(\Delta\theta_i)^2_{\mathrm{CR}} = \frac{F_{j,j}\mathbf{b}^2_{i,i}}{m\left(F_{i,i}F_{j,j} - F^2_{i,j}\right)} \geqslant \frac{\mathbf{b}^2_{i,i}}{mF_{i,i}}, \tag{27}$$

with $i \neq j$, $i,j = 1,2$. For the two-parameter case, the inequality (27) can be immediately demonstrated by using $F_{1,1}F_{2,2} - F^2_{1,2} > 0$, which holds since $\mathbf{F}$ is non-negative definite and assumed here to be invertible.

In the estimation of a single parameter, the matrix $\mathbf{B}_{\mathrm{CR}}$ reduces to the variance $(\Delta\theta_{\mathrm{CR}})^2$. In the unbiased case $[\bar{\Theta}(\theta) = \theta]$, Eq. (24) becomes

$$\Delta\theta_{\mathrm{CR}} = \frac{1}{\sqrt{mF}}, \tag{28}$$

where $F = \int d\varepsilon \frac{1}{P(\varepsilon|\theta)}\left(\frac{dP(\varepsilon|\theta)}{d\theta}\right)^2$ is the (scalar) Fisher information (FI). By comparing Eq. (27) and Eq. (28), we see, as reasonably expected, that the estimation uncertainty of a unbiased multiparameter estimation is always larger, or at most equal, than the optimal uncertainty obtained for a single parameter (namely, when all other parameters are exactly known).

2. Maximum likelihood estimation

A main goal is to find the estimators saturating the Cramér-Rao bound. These are called estimators. While such estimators are rare, it is not possible to exclude, in general, that an efficient unbiased estimator may exist for any value of m. One of the most important estimators is the maximum likelihood (ML) $\mathbf{\Theta}_{\mathrm{ML}}(\boldsymbol{\varepsilon})$. It is defined as the value $\mathbf{\Theta}_{\mathrm{ML}}(\boldsymbol{\varepsilon})$, which maximizes the likelihood function:

$$\mathbf{\Theta}_{\mathrm{ML}}(\boldsymbol{\varepsilon}) = \arg[\max_{\boldsymbol{\varphi}} \mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\varphi})]. \tag{29}$$

It is possible to demonstrate, by using the law of large numbers and the central limit theorem, that, asymptotically in the number of measurements, the maximum likelihood is unbiased and normally distributed with covariance given by the inverse FIM [63,67,68]. Therefore, the specific choice of vector $\boldsymbol{u}$, which leads to the Cramér-Rao bound (24) is justified by the fact that the ML saturates this bound for a sufficiently large number of measurements.

3. Quantum Cramér-Rao bound

The FIM satisfies

$$\mathbf{F} \leqslant \mathbf{F}_{\mathrm{Q}}, \tag{30}$$

in the sense that the matrix $\mathbf{F}_{\mathrm{Q}} - \mathbf{F}$ is positive definite. The symmetric matrix $\mathbf{F}_Q$ is called the quantum Fisher information matrix and its elements are

$$[\mathbf{F}_{\mathrm{Q}}]_{i,j} = \tfrac{1}{2}\mathrm{Tr}[\hat{\rho}(\theta)(\hat{L}_i\hat{L}_j + \hat{L}_j\hat{L}_i)], \tag{31}$$

where the self-adjoint operator $\hat{L}_i$, called the symmetric logarithmic derivative (SLD) [63], is defined via

$$\frac{\partial\hat{\rho}(\boldsymbol{\theta})}{\partial\theta_i} = \frac{\hat{L}_i\hat{\rho}(\boldsymbol{\theta}) + \hat{\rho}(\boldsymbol{\theta})\hat{L}_i}{2}. \tag{32}$$

In particular, we have $\frac{dP(\varepsilon|\boldsymbol{\theta})}{d\theta_i} = \Re(\mathrm{Tr}[\rho(\boldsymbol{\theta})\hat{E}(\varepsilon)\hat{L}_i])$, $\Re(x)$ being the real part of x. Note also that the operator $\hat{L}_i$ (and also $\mathbf{F}_{\mathrm{Q}}$) generally depends on $\boldsymbol{\theta}$. Equation (30) holds for any FIM (invertible or not) and there is no guarantee that, in general, the equality sign can be saturated. Assuming that $\mathbf{F}$ and $\mathbf{F}_{\mathrm{Q}}$ are positive definite (and thus invertible) and combining Eq. (24)—in the unbiased case—with Eq. (30), we obtain the matrix inequality $\mathbf{B}_{\mathrm{CR}} \geqslant \mathbf{B}_{\mathrm{QCR}}$ [72], where

$$\mathbf{B}_{\mathrm{QCR}} = \frac{\mathbf{F}_{\mathrm{Q}}^{-1}}{m}. \tag{33}$$

This sets a fundamental bound, the quantum Cramér-Rao (QCR) [63], to the phase sensitivity achievable with unbiased estimators. The bound cannot be saturated, in general, in the multiparameter case.

In the single-parameter case we have $\Delta\theta_{\mathrm{CR}} \geqslant \Delta\theta_{\mathrm{QCR}}$, where the quantum Cramér-Rao reduces to

$$\Delta\theta_{\mathrm{QCR}} = \frac{1}{\sqrt{mF_Q[\hat{\rho}(\theta)]}}. \tag{34}$$

The (scalar) quantum Fisher information (QFI) can be written as

$$F_Q[\hat{\rho}(\theta)] = (\Delta\hat{L})^2, \tag{35}$$

where $\hat{L}$ is the θ-dependent SLD and we used $\mathrm{Tr}[\hat{\rho}(\theta)\hat{L}] = 0$. The equality $\Delta\theta_{\mathrm{CR}} = \Delta\theta_{\mathrm{QCR}}$ (or, equivalently $F = F_Q$) holds if the POVM $\{\hat{E}(\varepsilon)\}$ is made by the set of projector operators over the eigenvectors of the operator $\hat{L}$, as first discussed in Ref. [73]. The quantum Cramér-Rao is a very convenient way to calculate the phase uncertainty since it can be saturated and depends on the probe state and phase-encoding transformation.

III. SEPARABILITY AND ENTANGLEMENT

In this section we discuss the relation between phase-estimation sensitivity and the entanglement properties of the probe state. We first briefly review the case of states with a fixed number of particles [15,16,74,75]. For states with number fluctuations, the situation is more involved: a relation between particle entanglement and phase sensitivity can be established only in the incoherent case. In presence of number coherences, the definition of separability and entanglement between particles becomes a fuzzy concept because the number of parties stay in a quantum superposition.

A. Bounds on the QFI for states with a fixed number of particles

A state of N particle is called separable if it can be written as a convex sum of product states [76,77],

$$\hat{\rho}^{(N)}_{\mathrm{sep}} = \sum_k P_k \bigotimes_{i=1}^{N} |\phi_{k,i}\rangle\langle\phi_{k,i}|, \tag{36}$$

where $|\phi_{k,j}\rangle$ is the state of the ith particle $(i = 1,\ldots,N)$ and $P_k \geqslant 0$. A state is (multiparticle) entangled if it is not separable. One can further consider the case where only $\kappa \leqslant N$ particles are in an entangled state and classify multiparticle entangled states following Refs. [77–81]. A state of N particles is κ producible if it can be written as

$$\rho^{(N)}_{\kappa-\text{prod}} = \sum_k P_k \bigotimes_{i=1}^{M_k} \hat{\rho}_k^{(N_i)}, \tag{37}$$

where $\hat{\rho}_k^{(N_j)}$ is a state of $N_j \leqslant \kappa$ particles ($j = 1,\dots,M_k$) with $\sum_{j=1}^{M_k} N_j = N$. A state is κ-particle entangled if it is κ producible but not $(\kappa-1)$ producible. In other words, a state is κ-particle entangled if it contains at least one group of κ particles that are in an entangled state. Notice that, formally, a separable state is $\kappa = 1$ producible.

When the number of particles is fixed, there exists a precise relation between the entanglement properties of a probe state and the QFI: if the state is separable [i.e., can be written as in Eq. (36)] then the inequality

$$F_Q\big[\hat{\rho}^{(N)}_{\text{sep}}, \hat{J}^{(N)}_{\boldsymbol{n}}\big] \leqslant N \tag{38}$$

holds [15,16]. A QFI larger than N is a sufficient condition for entanglement and singles out the states that are useful for quantum interferometry [16], i.e., states that can be used to achieve a sub-shot-noise phase uncertainty. Furthermore, in Refs. [74,75], it has been shown that for κ-producible states the bound

$$F_Q\big[\rho^{(N)}_{\kappa-\text{prod}}; \hat{J}^{(N)}_{\boldsymbol{n}}\big] \leqslant s\kappa^2 + r^2 \tag{39}$$

holds, where $s = \lfloor N/\kappa \rfloor$ is the largest integer smaller than, or equal to, N/κ and $r = N - s\kappa$. Hence, a violation of the bound (39) proves $(\kappa+1)$-particle entanglement. For general states of a fixed number of particles, we have $F_Q[\hat{\rho}^{(N)}, \hat{J}^{(N)}_{\boldsymbol{n}}] \leqslant N^2$ [15,16], whose saturation requires N-particle entanglement [15,16,74,75].

B. Bounds on the QFI for states with a fluctuating number of particles, without number coherences

We here extend the definition of separability and entanglement to states of a fluctuating number of particles without number coherences. An incoherent mixture (9) is defined as separable if it can be written as [17]

$$\hat{\rho}_{\text{sep}} = \sum_N Q_N \hat{\rho}^{(N)}_{\text{sep}}, \tag{40}$$

where $\hat{\rho}^{(N)}_{\text{sep}}$ is a separable state of N particles, Eq. (36). Incoherent states that are not separable according to this definition are entangled. Similarly, an incoherent mixture is κ producible if [82]

$$\rho_{\kappa-\text{prod}} = \sum_N Q_N \rho^{(N)}_{\kappa-\text{prod}}, \tag{41}$$

where $\rho^{(N)}_{\kappa-\text{prod}}$ is a κ-producible state of N particles, Eq. (37).

For separable states without number coherences we obtain

$$F_Q[\hat{\rho}_{\text{sep}}, \hat{J}_{\boldsymbol{n}}] = \sum_N Q_N F_Q\big[\hat{\rho}^{(N)}_{\text{sep}}, \hat{J}^{(N)}_{\boldsymbol{n}}\big] \leqslant \sum_N Q_N N = \langle \hat{N} \rangle. \tag{42}$$

This inequality follows from Eq. (38) and the general relation

$$F_{\text{Q}}[\hat{\rho}_{\text{inc}}, \hat{J}_{\boldsymbol{n}}] = \sum_N Q_N F_{\text{Q}}\big[\hat{\rho}^{(N)}, \hat{J}^{(N)}_{\boldsymbol{n}}\big], \tag{43}$$

valid for states without number coherences, where $F_Q[\hat{\rho}^{(N)}, \hat{J}^{(N)}_{\boldsymbol{n}}]$ is the QFI calculated on the fixed-N subspace. To demonstrate Eq. (43) we recall that the QFI for unitary transformations can be written as [73],

$$F_Q[\hat{\rho}, \hat{J}_{\boldsymbol{n}}] = 2 \sum_{\substack{i,j \\ p_i+p_j \neq 0}} \frac{(p_i - p_j)^2}{p_i + p_j} |\langle i|\hat{J}_{\boldsymbol{n}}|j\rangle|^2, \tag{44}$$

where $p_j \geqslant 0$ and $\{|j\rangle\}$ is a basis of the Hilbert space, $\sum_j |j\rangle\langle j| = \mathbb{1}$, chosen such that $\hat{\rho} = \sum_j p_j |j\rangle\langle j|$. For states without number coherences, we have $\hat{\rho}_{\text{inc}} = \sum_N Q_N \sum_j p_j^{(N)} |j^{(N)}\rangle\langle j^{(N)}|$ where $\{|j^{(N)}\rangle\}$ is a basis on the fixed-N subspace. We obtain Eq. (43) by noticing that $\langle j^{(N)}|\hat{J}_{\boldsymbol{n}}|j'^{(N')}\rangle = \langle j^{(N)}|\hat{J}^{(N)}_{\boldsymbol{n}}|j'^{(N)}\rangle \delta_{N,N'}$, which follows since $\hat{J}_{\boldsymbol{n}}$ does not couple states with different number of particles. Similarly, it is possible to demonstrate that the SLD $\hat{L}$ is given by the sum of SLDs in each fixed-N subspace, $\hat{L} = \sum_N \hat{L}^{(N)}$. We thus conclude that when the input state does not have number coherences the Von Neumann measurement on the eigenstates of $\hat{L}^{(N)}$ for each value of N—which in particular does not have number coherences—is optimal: the corresponding FI saturates the QFI.

As a direct consequence of Eq. (42) and the quantum Cramér-Rao bound, the phase sensitivity achievable with separable states without number coherences satisfies $\Delta\theta \geqslant \Delta\theta_{\text{SN}}$, where

$$\Delta\theta_{\text{SN}} = \frac{1}{\sqrt{m\langle \hat{N} \rangle}}, \tag{45}$$

is the shot-noise or standard quantum limit. This brings us to the following result: An arbitrary state which fulfills the inequality

$$\chi^2 \equiv \frac{\langle \hat{N} \rangle}{F_Q[\hat{\rho}, \hat{J}_{\boldsymbol{n}}]} < 1, \tag{46}$$

for some direction $\boldsymbol{n}$, is necessarily particle entangled according to the definition given above. In other words, entanglement is a necessary resource for sub-shot-noise sensitivity in linear SU(2) interferometers when number coherences are not available. States $\hat{\rho}$ satisfying Eq. (46) are useful in a linear interferometer implemented by the transformation $\hat{J}_{\boldsymbol{n}}$, since, according to Eq. (34), they can provide a sub-shot-noise phase sensitivity.

The relation between the properties of a probe state without number coherences and the QFI can be further extended to the case of κ-particle entanglement. Using Eqs. (41) and (43), we have $F_Q[\rho_{\kappa-\text{prod}}; \hat{J}_{\boldsymbol{n}}] = \sum_N Q_N F_Q[\rho^{(N)}_{\kappa-\text{prod}}; \hat{J}^{(N)}_{\boldsymbol{n}}]$ and thus, using Eq. (39),

$$F_Q\big[\rho^{\text{inc}}_{\kappa-\text{prod}}; \hat{J}_{\boldsymbol{n}}\big] \leqslant \sum_N Q_N \left\{ \left(\frac{N}{\kappa}\right)\kappa^2 + \left[N - \left(\frac{N}{\kappa}\right)\kappa\right]^2 \right\}$$
$$= \langle \hat{s}\kappa^2 \rangle + \langle \hat{r}^2 \rangle.$$

The operators $\hat{s} = \lfloor \hat{N}/\kappa \rfloor$ and $\hat{r} = \hat{N} - \hat{s}\kappa$ commute with $\hat{N}$. The maximum value of the Fisher information is thus obtained for maximally entangled states ($\kappa = N$) and is

$$\max_{\hat{\rho}_{\text{inc}}} F_Q[\hat{\rho}_{\text{inc}}, \hat{J}_{\boldsymbol{n}}] = \langle \hat{N}^2 \rangle. \tag{47}$$

or, equivalently,

$$\min_{\hat{\rho}_{\rm inc}} \Delta\theta_{\rm QCR} = \frac{1}{\sqrt{m\langle \hat{N}^2\rangle}}. \tag{48}$$

Equations (47) and (48) are saturated by incoherent superpositions of NOON-like states $\hat{\rho} = \sum Q_N |{\rm NOON}\rangle_{\boldsymbol{n}}\langle{\rm NOON}|$, where

$$|{\rm NOON}\rangle_{\boldsymbol{n}} = \frac{|N,0\rangle_{\boldsymbol{n}} + |0,N\rangle_{\boldsymbol{n}}}{\sqrt{2}}, \tag{49}$$

and $|N,0\rangle_{\boldsymbol{n}}(|0,N\rangle_{\boldsymbol{n}})$ is the eigenstate of $\hat{J}_{\boldsymbol{n}}$ with eigenvalue $N/2(-N/2)$. By a proper choice of the Q_N distribution, $\langle\hat{N}^2\rangle$ can be an arbitrary function of $\langle\hat{N}\rangle$. Therefore, when fixing $\langle\hat{N}\rangle$, the bound Eq. (48) can be arbitrarily small, even zero for distribution having $\langle\hat{N}^2\rangle = +\infty$. This was first noticed in Ref. [18]. The significance of the bound Eq. (48) is the subject of a vivid debate in the recent literature, see Sec. IV.

C. Bounds on the QFI for states with a fluctuating number of particles, with number coherences

A possible definition of separability for states of a fluctuating number of particles has been proposed in Ref. [17]. There, a state with number coherences is called separable if it is separable in every fixed-N subspace, i.e., if the incoherent mixture $\sum_N \hat{\pi}_N \hat{\rho}_{\rm coh}\hat{\pi}_N$, obtained from $\hat{\rho}_{\rm coh}$ by projecting over fixed-N subspaces, has the form of Eq. (40). States that are not separable are called entangled. Yet, with this definition, there is no clear relation multiparticle entanglement and phase sensitivity. The main difficulty is due to the fact that the QFI of a state with number coherences is generally larger than the QFI of the incoherent counterpart of that state:

$$F_Q[|\psi\rangle,\hat{J}_{\boldsymbol{n}}] \geqslant F_Q[\hat{\rho}_{\rm inc},\hat{J}_{\boldsymbol{n}}], \tag{50}$$

where $|\psi\rangle = \sum_N \sqrt{Q_N}\,|\psi_N\rangle$ is a normalized pure state with coherences and $\hat{\rho}_{\rm inc} = \sum_N \hat{\pi}_N|\psi\rangle\langle\psi|\hat{\pi}_N = \sum_N |Q_N|\,|\psi_N\rangle\langle\psi_N|$ is obtained from $|\psi\rangle\langle\psi|$ by tracing out the number coherences. Note that $|\psi\rangle$ and $\hat{\rho}_{\rm inc}$ have the same number of particles distribution. Moreover, if $F_Q[|\psi\rangle,\hat{J}_{\boldsymbol{n}}] > F_Q[\hat{\rho}_{\rm inc},\hat{J}_{\boldsymbol{n}}]$ holds, then the saturation of $F_Q[|\psi\rangle,\hat{J}_{\boldsymbol{n}}]$ necessarily requires a POVM with number coherences. Indeed, the Fisher information obtained with POVMs without coherences is upper bounded by $F_Q[\hat{\rho}_{\rm inc},\hat{J}_{\boldsymbol{n}}]$, independently on the presence of number coherences in the probe state.

Equation (50) can be demonstrated using (i) $F_Q[|\psi\rangle,\hat{J}_{\boldsymbol{n}}] = 4(\Delta\hat{J}_{\boldsymbol{n}})^2_{|\psi\rangle}$ [16,73] and $F_Q[\hat{\rho}_{\rm inc},\hat{J}_{\boldsymbol{n}}] = \sum_N Q_N(\Delta\hat{J}_{\boldsymbol{n}}^{(N)})^2_{|\psi_N\rangle}$ [see Eq. (43)], where we have explicitly indicated the state on which the variance is calculated on (we will keep this notation where necessary and drop it elsewhere) and (ii) the Cauchy-Schwartz inequality

$$\left(\sum_N |Q_N|\langle\psi_N|\hat{J}_{\boldsymbol{n}}^{(N)}|\psi_N\rangle\right)^2 \leqslant \sum_N |Q_N|\langle\psi_N|\hat{J}_{\boldsymbol{n}}^{(N)}|\psi_N\rangle^2. \tag{51}$$

Note that in Eq. (51), the equality holds if and only if $\langle\psi_N|\hat{J}_{\boldsymbol{n}}^{(N)}|\psi_N\rangle$ is a constant independent of N. As a consequence of Eq. (50), we can find states with number coherences that are separable in each fixed-N subspace and have a QFI that can be arbitrarily larger than $\langle\hat{N}\rangle$ (and can thus overcome the shot-noise phase sensitivity). The two examples below illustrate this fact.

TABLE II. The table summarizes the upper bound to the Fisher information for separable states and SU(2) transformations. When number coherences are present in both the probe state and in the POVM there is no specific bounds for separable states. In this case we have $F \leqslant \langle\hat{N}^2\rangle$, which holds for general states (see Sec. III C 3).

	POVM with coherences	POVM without coherences
$\hat{\rho}$ with coherences	$F \leqslant F_Q \leqslant \langle\hat{N}^2\rangle$	$F \leqslant \langle\hat{N}\rangle$
$\hat{\rho}$ without coherences	$F \leqslant F_Q \leqslant \langle\hat{N}\rangle$	$F \leqslant F_Q \leqslant \langle\hat{N}\rangle$

We can give an operational meaning to the above definition of separability and entanglement for states with number coherences, if we restrict to SU(2) transformations and POVMs without number coherences. In this (nonoptimal) case, the Fisher information for separable states is $F \leqslant \langle\hat{N}\rangle$. Therefore, for coherent separable states and POVM without coherences, the inequality $\Delta\theta \geqslant \Delta\theta_{\rm SN}$ holds. States with number coherences, which, in a phase estimation experiment using POVMs without number coherences, overcome the shot-noise sensitivity, are necessarily entangled within the definition given above. The situation is summarized in Table II.

1. Example: MOON state

We consider the state (which we call MOON state)

$$|\psi\rangle = \sqrt{\frac{N}{N+M}}\,e^{i\phi}\,|M,0\rangle + \sqrt{\frac{M}{N+M}}\,|0,N\rangle, \tag{52}$$

with $N,M > 0$. For $N \neq M$ this state is separable in each subspace of a fixed number of particles and thus separable according to the definition [17] given above. If $N = M$ Eq. (52) reduces to the well-known NOON state [42–46], which is maximally entangled. The QFI along the z direction and given by

$$F_Q[|\psi\rangle,\hat{J}_z] = NM. \tag{53}$$

The average number of particles is $\langle\hat{N}\rangle = 2NM/(N+M)$. Therefore, $F_Q[|\psi\rangle,\hat{J}_z]/\langle\hat{N}\rangle = (N+M)/2$, and we have an example of a separable state (with number coherences) which has a QFI (arbitrarily, for $N+M > 2$) larger than the average number of particles.

By projecting Eq. (52) over fixed-N subspaces we obtain an incoherent mixture of separable states,

$$\hat{\rho}_{\rm inc} = \frac{N}{N+M}|M,0\rangle\langle M,0| + \frac{M}{N+M}|0,N\rangle\langle 0,N|. \tag{54}$$

Its QFI is maximum on the plane orthogonal to z and fulfills $F_Q[\hat{\rho}_{\rm inc},\hat{J}_{\boldsymbol{n}}] \leqslant \langle\hat{N}\rangle$, as expected [see Eq. (42)]. Notice also that $F_Q[|\psi\rangle,\hat{J}_z] \geqslant F_Q[\hat{\rho}_{\rm inc},\hat{J}_z]$, in agreement with Eq. (50).

2. Example: Coherence with the vacuum

An example similar to the one above has been discussed in Ref. [27] and highlights how the coherence with the vacuum state can increase the QFI (see also [25,26] for a discussion in the single-mode case). Let us take

$$|\psi\rangle = \sqrt{1 - \frac{\langle \hat{N}\rangle}{N}}\,|\text{vac}\rangle + \sqrt{\frac{\langle \hat{N}\rangle}{N}}\,e^{i\phi_N}\,|N,0\rangle, \tag{55}$$

where $|\text{vac}\rangle$ is the vacuum and $\langle \hat{N}\rangle \leqslant N$ is the average number of particles. The QFI for rotations around the $\hat{J}_z$ axis is $F_Q[|\psi\rangle,\hat{J}_z] = N\langle \hat{N}\rangle - \langle \hat{N}\rangle^2$. By properly choosing N, it is possible to reach arbitrary large values of the QFI. For instance $F_Q[|\psi\rangle,\hat{J}_z] > \langle \hat{N}\rangle^k$ for $N > \langle \hat{N}\rangle^{k-1} + \langle \hat{N}\rangle$ and any $k > 0$. In particular, Eq. (55) is, as above, an example of separable state [according to Eq. (40)], which has a QFI larger than $\langle \hat{N}\rangle$ [for $N > \langle \hat{N}\rangle + 1$]. We also have $F_Q[|\psi\rangle,\hat{J}_z] > \langle \hat{N}\rangle^2$ for $N > 2\langle \hat{N}\rangle$. Finally note that, as expected, the condition $F_Q[|\psi\rangle,\hat{J}_z] < \langle \hat{N}^2\rangle = N\langle \hat{N}\rangle$ is always fulfilled for $\langle \hat{N}\rangle > 0$.

3. *Upper bounds of the quantum Fisher information for states with number coherences*

In the following we derive upper bounds on the QFI for arbitrary states with number coherences. We consider different transformations:

(i) SU(2) transformation $\hat{U}(\theta) = e^{-i\theta \hat{J}_{\boldsymbol{n}}}$. In this case, using the convexity of the QFI [67] we obtain

$$F_Q[\hat{\rho}_{\text{coh}},\hat{J}_{\boldsymbol{n}}] \leqslant \sum_k p_k F_Q[|\psi_k\rangle,\hat{J}_{\boldsymbol{n}}] = 4\sum_k p_k (\Delta \hat{J}_{\boldsymbol{n}})^2_{|\psi_k\rangle}, \tag{56}$$

where the equality holds only for pure states. Furthermore,

$$4(\Delta \hat{J}_{\boldsymbol{n}})^2_{|\psi_k\rangle} \leqslant 4\sum_N |Q_{N,k}| \langle J^2_{\boldsymbol{n}}\rangle_{|\psi_{N,k}\rangle} \leqslant \sum_N |Q_{N,k}| N^2 = \langle \hat{N}^2\rangle_{|\psi_k\rangle}, \tag{57}$$

where the first inequality is saturated for $\langle \hat{J}_{\boldsymbol{n}}\rangle_{|\psi_k\rangle} = 0$ and the second inequality, due to $4\langle \hat{J}^2_{\boldsymbol{n}}\rangle_{|\psi_{N,k}\rangle} \leqslant N^2$, is saturated by a NOON-like state (49). Combining Eqs. (56) and (57), we obtain

$$F_Q[\hat{\rho}_{\text{coh}},\hat{J}_{\boldsymbol{n}}] \leqslant \text{Tr}[\hat{\rho}_{\text{coh}}\hat{N}^2], \tag{58}$$

where the equality is saturated by the coherent superposition of states (49) (note that $\langle \text{NOON}_{\boldsymbol{n}}|\hat{J}_{\boldsymbol{n}}|\text{NOON}_{\boldsymbol{n}}\rangle = 0$). We thus have

$$(\Delta\theta)^2_{\text{QCR}} \geqslant \frac{1}{m\text{Tr}[\hat{\rho}_{\text{coh}}\hat{N}^2]}. \tag{59}$$

(ii) U(2) transformations $\hat{U}(\phi_0) = e^{-i\phi_0 \hat{N}}$. From the convexity of the QFI [67], we obtain

$$F_Q[\hat{\rho}_{\text{coh}},\hat{N}] \leqslant 4\sum_k p_k (\Delta \hat{N})^2_{|\psi_k\rangle} \leqslant 4(\Delta \hat{N})^2_{\hat{\rho}_{\text{coh}}}, \tag{60}$$

where the second inequality follows from Cauchy-Schwarz. Therefore

$$(\Delta\phi_0)^2_{\text{QCR}} \geqslant \frac{1}{4m(\Delta \hat{N})^2_{\hat{\rho}_{\text{coh}}}}. \tag{61}$$

(iii) U(2) transformations $\hat{U}(\theta,\phi_0) = e^{-i\theta \hat{J}_{\boldsymbol{n}}} e^{-i\phi_0 \hat{N}}$. In this case both parameters are estimated at the same time and we apply the multiparameter estimation theory outlined above. The inequality (27) leads to

$$(\Delta\theta)^2_{\text{CR}} \geqslant \frac{1}{m F_Q[\rho_{\text{coh}},\hat{J}_{\boldsymbol{n}}]}, \quad (\Delta\phi_0)^2_{\text{CR}} > \frac{1}{m F_Q[\rho_{\text{coh}},\hat{N}]},$$

which can be further bounded using the inequalities (58) and (60). For pure states

$$\mathbf{F}_Q^{-1} = \frac{2}{\det[\mathbf{F_Q}]}\begin{pmatrix} 2(\Delta \hat{J}_{\boldsymbol{n}})^2 & \langle \hat{N}\rangle\langle \hat{J}_{\boldsymbol{n}}\rangle - \langle \hat{N}\hat{J}_{\boldsymbol{n}}\rangle \\ \langle \hat{N}\rangle\langle \hat{J}_{\boldsymbol{n}}\rangle - \langle \hat{N}\hat{J}_{\boldsymbol{n}}\rangle & (\Delta \hat{N})^2/2 \end{pmatrix},$$

where $\det[\mathbf{F_Q}] = 4(\Delta \hat{N})^2(\Delta \hat{J}_{\boldsymbol{n}})^2 - 4[\langle \hat{N}\hat{J}_{\boldsymbol{n}}\rangle - \langle \hat{N}\rangle\langle \hat{J}_{\boldsymbol{n}}\rangle]^2$. We thus have

$$(\Delta\phi_0)^2 \geqslant \frac{m^{-1}}{(\Delta \hat{N})^2 - [\langle \hat{N}\hat{J}_{\boldsymbol{n}}\rangle - \langle \hat{N}\rangle\langle \hat{J}_{\boldsymbol{n}}\rangle]^2/(\Delta \hat{J}_{\boldsymbol{n}})^2}, \tag{62}$$

which, in particular, is always larger than $1/m(\Delta \hat{N})^2$, and

$$(\Delta\theta)^2 \geqslant \frac{m^{-1}}{4(\Delta \hat{J}_{\boldsymbol{n}})^2 - 4[\langle \hat{N}\hat{J}_{\boldsymbol{n}}\rangle - \langle \hat{N}\rangle\langle \hat{J}_{\boldsymbol{n}}\rangle]^2/(\Delta \hat{N})^2}, \tag{63}$$

which is always larger than $1/4m(\Delta \hat{J}_{\boldsymbol{n}})^2$.

IV. HEISENBERG LIMIT

In this section we discuss the ultimate phase sensitivity achievable when fixing the average number of particles $\langle \hat{N}\rangle$ in the probe state. This is generally indicated as the Heisenberg limit. We focus on SU(2) transformations $e^{-i\theta \hat{J}_{\boldsymbol{n}}}$. In Sec. IV A we show that the Heisenberg limit for states and/or POVMs without number coherences is given by Eq. (4). For states with number coherences, the situation is more involved and a conclusive demonstration of Eq. (4) is not available. In Sec. IV B we demonstrate that Eq. (4) holds in the central limit, at least. Finally, in Sec. IV C we give an overview of the main results obtained in the literature regarding the Heisenberg limit for states with number coherences.

A. Heisenberg limit for states and/or POVMs without number coherences

The mean value and variance (for m independent measurements) of an arbitrary estimator $\Theta(\boldsymbol{\varepsilon})$ are

$$\bar{\Theta} = \int d\boldsymbol{\varepsilon}\; P(\boldsymbol{\varepsilon}|\theta)\,\Theta(\boldsymbol{\varepsilon}), \tag{64}$$

and

$$(\Delta\theta)^2 = \sum_{\boldsymbol{\varepsilon}} P(\boldsymbol{\varepsilon}|\theta)\,[\Theta(\boldsymbol{\varepsilon}) - \bar{\Theta}]^2, \tag{65}$$

respectively, where $\boldsymbol{\varepsilon} \equiv \{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_m\}$ and $P(\boldsymbol{\varepsilon}|\theta) = \prod_{i=1}^m P(\varepsilon_i|\theta)$. For states and/or POVMs without number coherences, taking $P(\varepsilon|\theta)$ as in Eq. (20), we can rewrite Eq. (64) as

$$\bar{\Theta} = \sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}\,\bar{\Theta}_{\boldsymbol{N}}, \tag{66}$$

where the sum extends over all possible sequences $\boldsymbol{N} \equiv \{N_1,N_2,\ldots,N_m\}$, $Q_{\boldsymbol{N}} \equiv \prod_{i=1}^m Q_{N_i}$ is the probability of the given sequence $[\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}} = 1]$,

$$\bar{\Theta}_{\boldsymbol{N}} \equiv \int d\boldsymbol{\varepsilon}\, P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta)\,\Theta(\boldsymbol{\varepsilon}), \tag{67}$$

and $P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta) \equiv \prod_{i=1}^{m} P(\varepsilon_i|N_i,\theta)$. Following analogous calculations, we can rewrite Eq. (65) as

$$(\Delta\theta)^2 = \sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}[\bar{\Theta}_{\boldsymbol{N}} - \bar{\Theta}]^2 + \sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}(\Delta\theta_{\boldsymbol{N}})^2, \tag{68}$$

where

$$(\Delta\theta_{\boldsymbol{N}})^2 \equiv \int d\boldsymbol{\varepsilon}\, P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta)\,[\Theta(\boldsymbol{\varepsilon}) - \bar{\Theta}_{\boldsymbol{N}}]^2 \tag{69}$$

is the variance of the estimator $\Theta(\boldsymbol{\varepsilon})$ for a given sequence $\boldsymbol{N}$. Since $\int d\boldsymbol{\varepsilon} P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta) = 1$, we can apply the Cramér-Rao theorem to set the bound

$$(\Delta\theta_{\boldsymbol{N}})^2 \geqslant \frac{b_{\boldsymbol{N}}^2}{F_{\boldsymbol{N}}(\theta)}, \tag{70}$$

where $b_{\boldsymbol{N}} \equiv \partial_\theta \bar{\Theta}_{\boldsymbol{N}}$ and

$$F_{\boldsymbol{N}}(\theta) \equiv \int d\boldsymbol{\varepsilon}\, \frac{1}{P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta)}\left(\frac{d}{d\theta}P(\boldsymbol{\varepsilon}|\boldsymbol{N},\theta)\right)^2 \tag{71}$$

is the Fisher information for the specific $\boldsymbol{N}$. Note that $F_{\boldsymbol{N}}(\theta) = \sum_{i=1}^{m} F_{N_i}(\theta)$, where $F_{N_i}(\theta)$ is the Fisher information calculated on the subspace of N_i particles,

$$F_{N_i}(\theta) \equiv \int d\varepsilon_i\, \frac{1}{P(\varepsilon_i|N_i,\theta)}\left(\frac{d}{d\theta}P(\varepsilon_i|N_i,\theta)\right)^2. \tag{72}$$

If all the numbers N_i are equal to N, we recover $F_{\boldsymbol{N}}(\theta) = mF_N(\theta)$ and thus the usual multiplication factor m. Note also that the Fisher information $F_{N_i}(\theta)$ is bounded as $F_{N_i}(\theta) \leqslant N_i^2$ [15,16], and thus $F_{\boldsymbol{N}}(\theta) \leqslant \sum_{i=1}^{m} N_i^2 = \boldsymbol{N}\cdot\boldsymbol{N} = \boldsymbol{N}^2$. Using this result and Eq. (70) we obtain

$$\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}(\Delta\Theta_{\boldsymbol{N}})^2 \geqslant \sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}} b_{\boldsymbol{N}}^2}{\boldsymbol{N}^2} \geqslant \sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}} b_{\boldsymbol{N}}^2}{\mathcal{S}(\boldsymbol{N})^2}, \tag{73}$$

where $\mathcal{S}(\boldsymbol{N}) \equiv \sum_{i=1}^{m} N_i$ is the sum of all values of N in the sequence. The last inequality in Eq. (73) is a consequence of $\boldsymbol{N}^2 \leqslant \mathcal{S}(\boldsymbol{N})^2$, which follows since all N_i are positive numbers. We now use the Cauchy-Schwarz inequality

$$\sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}} b_{\boldsymbol{N}}^2}{\mathcal{S}(\boldsymbol{N})^2} \sum_{\boldsymbol{N}'} Q_{\boldsymbol{N}'} \geqslant \left(\sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}}\, b_{\boldsymbol{N}}}{\mathcal{S}(\boldsymbol{N})}\right)^2. \tag{74}$$

Using the normalization of $Q_{\boldsymbol{N}}$ and Eq. (73), we have

$$\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}(\Delta\theta_{\boldsymbol{N}})^2 \geqslant \left(\sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}}\, b_{\boldsymbol{N}}}{\mathcal{S}(\boldsymbol{N})}\right)^2. \tag{75}$$

A second Cauchy-Schwarz inequality gives

$$\sum_{\boldsymbol{N}} \frac{Q_{\boldsymbol{N}}\, b_{\boldsymbol{N}}}{\mathcal{S}(\boldsymbol{N})} \sum_{\boldsymbol{N}'} Q_{\boldsymbol{N}'}\mathcal{S}(\boldsymbol{N}') \geqslant \left(\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}\sqrt{b_{\boldsymbol{N}}}\right)^2, \tag{76}$$

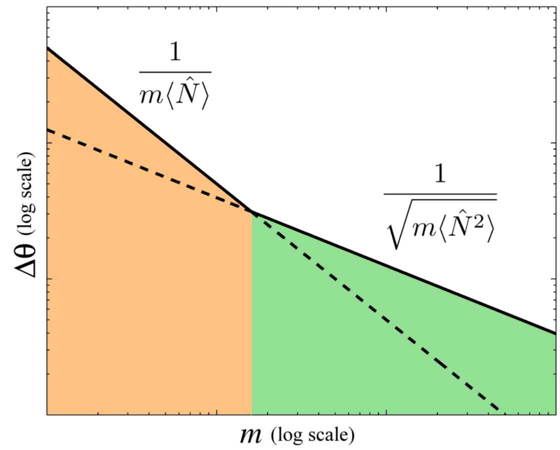

FIG. 2. (Color online) Schematic representation of Eq. (4) in log-log scale (solid line). In the left-hand side of the figure (orange region), $\Delta\theta_{\rm HL} = 1/m\langle\hat{N}\rangle$ (solid line), which is larger than $1/\sqrt{m\langle\hat{N}^2\rangle}$ (dashed line). In the right-hand side of the figure (green region), for $m \geqslant \langle\hat{N}^2\rangle/\langle\hat{N}\rangle^2$, $\Delta\theta_{\rm HL} = 1/\sqrt{m\langle\hat{N}^2\rangle}$ (solid line) is larger than $1/m\langle\hat{N}\rangle$ (dashed line)

where we note that $\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}\mathcal{S}(\boldsymbol{N}) = m\langle\hat{N}\rangle$, $b_{\boldsymbol{N}}$ are positive numbers, and $\langle\hat{N}\rangle = \sum_{\boldsymbol{N}} Q_{\boldsymbol{N}} N$ is the average number of particles. We thus have

$$\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}(\Delta\theta_{\boldsymbol{N}})^2 \geqslant \frac{\left(\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}\sqrt{b_{\boldsymbol{N}}}\right)^4}{(m\langle\hat{N}\rangle)^2}. \tag{77}$$

Finally, by using Eqs. (68) and (77), the sensitivity of the estimator can be bounded as

$$(\Delta\theta)^2 \geqslant \sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}(\bar{\Theta}_{\boldsymbol{N}} - \bar{\Theta})^2 + \frac{\left(\sum_{\boldsymbol{N}} Q_{\boldsymbol{N}}\sqrt{b_{\boldsymbol{N}}}\right)^4}{(m\langle\hat{N}\rangle)^2}. \tag{78}$$

This is the main result of this section. The first term in Eq. (78) is always positive and is characteristic of phase estimation with probe states of a nonfixed number of particles. It is equal to zero if and only if $\bar{\Theta}_{\boldsymbol{N}} = \bar{\Theta}$ for all possible sequences $\boldsymbol{N}$. Since sequences with m values of the same total number of particles N are possible, the condition $\bar{\Theta}_{\boldsymbol{N}} = \bar{\Theta}$ implies that the mean value of the estimator is the same (and equal to $\bar{\Theta}$) in each fixed-N subspace. Furthermore, a convenient situation is to have an unbiased estimator $\bar{\Theta} = \theta$ for all values of m. In this case [assuming that the first term in Eq. (78) is equal to zero, $\bar{\Theta}_{\boldsymbol{N}} = \theta$ for all possible sequences $\boldsymbol{N}$], we have [17]

$$\Delta\theta \geqslant \frac{1}{m\langle\hat{N}\rangle}. \tag{79}$$

We recall that this inequality holds when the estimator is unbiased in each fixed-N subspace, for all the values of m. If the estimator is biased in some fixed-N subspace, the more general, but less conclusive, inequality (78) holds.

1. Discussion

In Fig. 2 we schematically represent Eq. (4) as a function of the number of measurements m. We recall here that the first bound of Eq. (4), derived above, is generally not tight and is valid for estimators which are unbiased in each fixed-N subspace. The second bound in Eq. (4) is the optimal quantum Cramér-Rao bound, Eq. (48), for estimators which are globally unbiased ($\bar{\Theta} = \theta$). It can be saturated by the maximum likelihood estimator in the central limit (i.e., for $m \gtrsim m_{\rm cl}$ and a sufficiently large $m_{\rm cl}$) by using an incoherent mixture of NOON-like states.

Since $\langle \hat{N}^2 \rangle \geqslant \langle \hat{N} \rangle$, we can distinguish two regimes. For $m \leqslant \langle \hat{N}^2 \rangle / \langle \hat{N} \rangle^2$, the first bound in Eq. (4) is significant. In this regime, we can thus rule out, for states and/or POVM without coherences, the existence of estimators that are unbiased in each fixed-N subspace and saturate the Cramér-Rao bound: indeed, if such estimators would exist, we would have a violation of Eq. (4). The saturation of the Cramér-Rao bound is only possible for $m \geqslant \langle \hat{N}^2 \rangle / \langle \hat{N} \rangle^2$, where the second bound in in Eq. (4) dominates. In other words,

$$m_{\rm cl} \geqslant \frac{\langle \hat{N}^2 \rangle}{\langle \hat{N} \rangle^2}. \tag{80}$$

The larger is $\langle \hat{N}^2 \rangle$, the smaller is Eq. (48) but, at the same time, the larger is the number of repeated measurements needed to reach the central limit and saturate Eq. (48). If $\langle \hat{N}^2 \rangle \to \infty$, reaching the central limit requires an infinite number of measurements $m \to +\infty$, and, accordingly, the phase uncertainty vanishes, $\Delta\theta \to 0$.

2. *Example: Biased estimator*

We recall that Eq. (79) applies to estimators that are unbiased in each fixed-N subspace. If this is not the case, the bound $1/m\langle \hat{N} \rangle$ can be violated. This is explicitly shown in the following example. We consider the state

$$\hat{\rho} = (1-p)\,|{\rm vac}\rangle\langle{\rm vac}| + p\,|{\rm NOON}\rangle_{\boldsymbol{n}}\langle{\rm NOON}|, \tag{81}$$

where $|{\rm vac}\rangle$ is the vacuum and $|{\rm NOON}\rangle_{\boldsymbol{n}} = (|M,0\rangle_{\boldsymbol{n}} + |0,M\rangle_{\boldsymbol{n}})/\sqrt{2}$ is a NOON-like state [see Eq. (49)] of M particles. The average number of particles is $\langle \hat{N} \rangle = pM$. The QFI is

$$F_Q[\hat{\rho}, \hat{J}_{\boldsymbol{n}}] = pM^2 = \frac{\langle \hat{N} \rangle^2}{p}, \tag{82}$$

leading to the quantum Cramér-Rao bound

$$(\Delta\theta_{\rm QCR})^2 = \frac{1}{mF_Q[\hat{\rho}, \hat{J}_{\boldsymbol{n}}]} = \frac{p}{m\langle \hat{N} \rangle^2}. \tag{83}$$

We consider an optimal POVM (for which $F = F_Q$) and an estimator [83]

$$\Theta(\varepsilon) = \begin{cases} 0 & \text{if } N = 0 \\ \tilde{\Theta}(\varepsilon)/p & \text{if } N = M, \end{cases}$$

where ε is the result of a possible measurement in the fixed-N subspace and $\tilde{\Theta}(\varepsilon)$ is an arbitrary unbiased estimator. The estimator $\Theta(\varepsilon)$ is biased on each N subspace but it is globally unbiased, $\bar{\Theta} = \theta$. In this case Eq. (79) does not apply. We further assume that the estimator $\tilde{\Theta}(\varepsilon)$ saturates the Cramér-Rao bound for a single measurement ($m = 1$). In this case, using Eq. (68) [with $Q_0 = 1-p$, $Q_M = p$, $\bar{\Theta}_0 = 0$, $\bar{\Theta}_M = \theta/p$, $\Delta\theta_0 = 0$ and $\Delta\theta_M = \Delta\tilde{\Theta}/p$] we have

$$\begin{aligned}(\Delta\theta)^2 &= \sum_N Q_N(\bar{\Theta}_N - \bar{\Theta})^2 + \sum_N Q_N(\Delta\theta_N)^2 \\ &= \left(\frac{1-p}{p}\right)\theta^2 + \frac{(\Delta\tilde{\Theta})^2}{p} \\ &= \left(\frac{1-p}{p}\right)\theta^2 + \frac{1}{pF_Q[|{\rm NOON}\rangle_{\boldsymbol{n}}, \hat{J}_{\boldsymbol{n}}]} \\ &= \left(\frac{1-p}{p}\right)\theta^2 + \frac{p}{\langle \hat{N} \rangle^2}.\end{aligned}$$

The first term highlights the role of $\theta = 0$ as a sweet spot for the phase estimation. If $\theta = 0$ we may have a violation of Eq. (4) by an arbitrary small factor p, where p may even scale as $p \sim 1/\langle \hat{N} \rangle^k$, inversely proportional to an arbitrary power of $\langle \hat{N} \rangle$.

B. Some considerations about the Heisenberg limit for states and POVMs with number coherences

Here we show that the bound $1/m\langle \hat{N} \rangle$ applies in the fully coherent situation at least in the central limit. As discussed in Sec. III C, the optimal quantum Cramér-Rao bound is $\Delta\theta_{\rm QCR} = 1/\sqrt{m\langle \hat{N}^2 \rangle}$, which is uniquely saturated by a probe given by superpositions of pure NOON-like states of the form $|\psi\rangle = \sum_N \sqrt{Q_N}\,|{\rm NOON}\rangle_{\boldsymbol{n}}$ [see Eq. (49)]. Since $\langle {\rm NOON}_{\boldsymbol{n}}|\hat{J}_{\boldsymbol{n}}|{\rm NOON}_{\boldsymbol{n}}\rangle = 0$ for any N, the QFI can be written, in this case, as $F_Q = 4\langle \hat{J}_{\boldsymbol{n}}^2 \rangle$. In addition, since the operator $\hat{J}_{\boldsymbol{n}}$ commutes with $\hat{N}$, off-diagonal terms $N \neq N'$ in the density matrix $|{\rm NOON}\rangle_{\boldsymbol{n}}\langle{\rm NOON}|$ do not play any role in the calculation of the Cramér-Rao bound. Hence a mixture $\sum_N Q_N |{\rm NOON}\rangle_{\boldsymbol{n}}\langle{\rm NOON}|$ has the same QFI. It follows that the limit $\Delta\theta = 1/\sqrt{m\langle \hat{N}^2 \rangle}$ can be saturated for $m \geqslant m_{\rm cl}^{\rm inc}$ with a POVM without number coherences. It may happen, however, that the number of measurements for which the Cramér-Rao bound is saturated is different if number coherences in the state and POVM are used. For an asymptotically large m, the saturation in both cases is guaranteed by the Fisher theorem. In this regime, the results of Sec. IV A hold and we conclude that the Heisenberg limit Eq. (4) is valid also in the full coherent case. In particular, $\Delta\Theta \geqslant 1/m\langle \hat{N} \rangle$ is a general bound for sufficiently large m, even if states with coherences and POVMs with coherences are used.

A final remark concerns the uniqueness of the states that saturate the Cramér-Rao bound. Saturating $F_Q = \langle \hat{N}^2 \rangle$ requires that $4(\Delta \hat{J}_{\boldsymbol{n}}^{(N)})^2 = N^2$ for any N. Only states such that $\hat{\pi}_N \rho \hat{\pi}_N = Q_N |{\rm NOON}\rangle_{\boldsymbol{n}}\langle{\rm NOON}|$ satisfy this constraint. However, there are many such states for given $\langle \hat{N} \rangle$ and $\langle \hat{N}^2 \rangle$. This is because fixing $\langle \hat{N} \rangle$ and $\langle \hat{N}^2 \rangle$ corresponds to choosing two constraints on the distribution $\{Q_N\}$, in addition to the constraints $\sum_N Q_N = 1$ and $Q_N \geqslant 0$. $\{Q_N\}$ is in general not uniquely defined by these constraints. Even though different states can reach $F_Q = \langle \hat{N}^2 \rangle$, they may be characterized by different values of $m_{\rm cl}$, i.e., the minimal number where the central limit is reached, depending also on the estimator.

C. Overview of recent literature on the Heisenberg limit

The comparison between our results and the recent literature deserves a discussion. We recall that our definition of

TABLE III. Summary of the fundamental bounds of phase sensitivity discussed in this paper. For states and/or POVM without number coherence, the Heisenberg limit is given by the competition of two bounds [$\Delta\theta_{\rm HL}$, Eq. (4)], as explained in Sec. IV. For POVMs and states with number-coherences only the Cramèr-Rao bound $\Delta\theta_{\rm CR}$ applies. In this case and for SU(2) transformations, Eq. (4) holds at least in the central limit, as discussed in Sec. IV B.

	POVM with coherences	POVM without coherences
$\hat\rho_{\rm in}$ with coherences	$\Delta\theta \geqslant \Delta\theta_{\rm CR}$	$\Delta\theta \geqslant \Delta\theta_{\rm HL}$
$\hat\rho_{\rm in}$ without coherences	$\Delta\theta \geqslant \Delta\theta_{\rm HL}$	$\Delta\theta \geqslant \Delta\theta_{\rm HL}$

Heisenberg limit, Eq. (4), holds for two-mode transformations and unbiased estimators. In particular, it does not apply to states and POVM with number coherences outside the central limit, for which no conclusive results has been obtained so far. A summary of our findings is reported in Table III. In the literature, the problem of defining the Heisenberg limit for states and POVMs with number coherences has been tackled with different techniques, which we briefly discuss below. Overall, there is a strong indication in the literature that Eq. (4) is the general form of Heisenberg limit. While the literature leaves open the possibility to overcome the bound $1/m\langle\hat N\rangle$ at specific phase values, there is no proposal showing convincing evidence of sub-Heisenberg uncertainties [26].

Before presenting an overview of the literature, it is important to recall here that there are two models of phase estimation [84,85]: (i) The first model assumes that the phase shift θ is a fixed (nonrandom) unknown quantity. This is the framework discussed in this paper. It assumes that we can collect an arbitrary number of sequences $\boldsymbol\varepsilon = \{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_m\}$ of m measurements while keeping fixed the (unknown) phase shift. The phase sensitivity is given by the variance of the estimator $\Theta(\boldsymbol\varepsilon)$ (see Sec. II E):

$$(\Delta\Theta)^2_\theta = \int d\boldsymbol\varepsilon\, P(\boldsymbol\varepsilon|\theta)[\Theta(\boldsymbol\varepsilon) - \bar\Theta(\theta)]^2, \qquad (84)$$

where $\bar\Theta(\theta)$ is the θ-dependent mean value of the estimator. (ii) The second model assumes that the phase is a random variable with a probability distribution $P(\theta)$, called the prior. Parameter estimation based on this model is referred to as Bayesian estimation [85]. In this case, each sequence $\boldsymbol\varepsilon$ of m measurements is obtained with a random phase shift. The phase sensitivity is defined as the weighed mean-square error

$$(\Delta\Theta)^2_{\rm bay} = \int d\theta \int d\boldsymbol\varepsilon\, P(\boldsymbol\varepsilon,\theta)[\Theta(\boldsymbol\varepsilon) - \bar\Theta(\theta)]^2, \qquad (85)$$

where $P(\boldsymbol\varepsilon,\theta) = P(\boldsymbol\varepsilon|\theta)P(\theta)$ is the joint probability distribution of phase θ and experimental measurement $\boldsymbol\varepsilon$.

1. *Sweet spot phase estimation*

Let us consider here the estimation of a fixed phase shift with probe states and POVMs with number coherences. At certain phase values (indicated as "sweet spots" in Ref. [34]) the Cramér-Rao bound can be arbitrary small when fixing the average number of particles $\langle\hat N\rangle$ in the state. Nevertheless, in Ref. [34] it is shown that the sum of uncertainties at two nearby phase shifts, θ_1 and θ_2, is bounded when the phases are sufficiently far apart. For unbiased estimators, the inequality [34]

$$\frac{(\Delta\Theta)_{\theta_1} + (\Delta\Theta)_{\theta_2}}{2} \geqslant \frac{\kappa}{m(\langle\hat H\rangle - H_0)} \qquad (86)$$

holds, where $\hat H$ is the generator of phase shift (a phase encoding transformation $e^{-i\hat H\theta}$ is assumed), H_0 is the minimum eigenvalue of $\hat H$ populated in the probe state. The maximum value of κ is 0.074 reached when $|\theta_1 - \theta_2| \geqslant 0.83/m(\langle\hat H\rangle - H_0)$. In the special (which yet might be nonoptimal) case when the phase sensitivity $(\Delta\Theta)_\theta$ does not depend on θ, Eq. (86) implies $(\Delta\Theta)_\theta \geqslant \kappa/m(\langle\hat H\rangle - H_0)$ [34]. These results require the generator of phase shift to have a discrete spectrum and a finite lowest eigenvalue [34]. The bound (86) thus holds, for instance, for single-mode phase estimation, when the generator of phase shift is the number of particles operator, $\hat N$. Equation (86) does not hold for the two-mode case unless one imposes a bound on the total number of particles distribution. It should also be noticed that the bound found in Ref. [34] refers to the mean-square fluctuation of the estimator with respect to the true phase values [which coincides to Eq. (84) only if the estimator is unbiased]. For biased estimators, the bound (86) does not apply.

2. *Bayesian bounds*

Several works [29–33] have discussed the Heisenberg limit within the framework of Bayesian phase estimation, i.e., when the phase sensitivity is averaged over the prior, see Eq. (85). This approach might be considered as a generalization of the averaging over two phases discussed in the previous subsection [34]. In this case, the Heisenberg limit is found by making use of suitable Bayesian bounds. Using the Ziv-Zakai (Bayesian) bound [85] in the low prior information regime [e.g., when $P(\theta)$ is uniform a phase interval sufficiently wider than $1/m\langle\hat H\rangle$] it is possible to demonstrate that [29,30]

$$(\Delta\Theta)_{\rm bay} \geqslant \frac{\alpha}{m\langle\hat H\rangle}, \qquad (87)$$

where α is a constant [29,30] ($\alpha = 0.1548$ for an uniform prior distribution [35]). In the opposite regime, when the width of $P(\theta)$ is smaller than $1/m\langle\hat H\rangle$, the phase uncertainty is essentially determined by the prior distribution [29,35]. In this case, sub-Heisenberg uncertainties are possible but ineffective (i.e., the estimation process does not bring more information than a random guess of the phase within the prior itself [29]). In Refs. [29,30] the bound (87) was demonstrated by assuming $\hat H$ to have a finite lower bound in the spectrum, as in the single-mode case with $\hat H = \hat N$. The extension of Eq. (87) to unbounded Hamiltonians (and thus when $\hat H = \hat J_{\boldsymbol n}$) is discussed in Ref. [35].

In Refs. [31–33], using an entropic uncertainty relation, it was possible to show that

$$(\Delta\Theta)_{\rm bay} \geqslant \frac{\beta}{m\langle|\hat H - h|\rangle}, \qquad (88)$$

where h is an arbitrary eigenvalue of $\hat H$, which can have a discrete or continuous spectrum [33], and β, depending on the prior distribution $P(\theta)$, can be arbitrarily small for a sufficiently narrow prior ($\beta = 0.559$ for a completely random phase shift in a 2π interval [31]). The derivation of Eq. (88)

does not require $\hat{H}$ to be discrete, have integer eigenvalues, or have a finite lowest eigenvalue [33]. In particular, the bound applies to two-mode operators [32], i.e., when $\hat{H} = \hat{J}_{\mathbf{n}}$. In this case, we have $\langle|\hat{J}_{\mathbf{n}}|\rangle = \sum_N \sum_{\mu=-N/2}^{N/2} |\mu|\,|Q_{N,\mu}|^2$, where $-N/2 \leqslant \mu \leqslant N/2$ are eigenvalues of $\hat{J}_{\mathbf{n}}^{(N)}$ with eigenstate $|N,\mu\rangle$ and the mean value is calculated over $|\psi\rangle = \sum_{N,\mu} Q_{N,\mu}|N,\mu\rangle$, which is a state with number coherences. Using $|\mu| \leqslant N/2$ we obtain $\langle|\hat{J}_{\mathbf{n}}|\rangle \leqslant \langle\hat{N}\rangle/2$ and thus, from Eq. (88), $(\Delta\Theta)_{\mathrm{bay}} \geqslant \beta/m\langle\hat{N}\rangle$.

3. Coherent superposition with the vacuum

Reference [25] discusses a single-mode phase estimation reaching, at specific phase values, a phase uncertainty arbitrarily smaller than $1/m\langle\hat{N}\rangle$. This result is obtained from a calculation of the Fisher information for pure states having a large vacuum component (see a similar example in Sec. III C 2). In [25] it is argued that the maximum likelihood estimator might reach an arbitrary small phase uncertainty [25]. Results and claims similar to the one of Ref. [25] can be found in the early literature [5–7]. The bounds (87) and (88) do not apply to this case and therefore there are no analytical results in the literature that forbid the conclusions of Ref. [25] (and also of Refs. [5–7]). A detailed numerical analysis of the estimation protocol proposed in Ref. [25] can be found in Ref. [26], showing no violation of the Heisenberg limit. For a small number of measurements, m, the estimation protocol proposed in Ref. [25] is strongly biased [26]. In the large m limit, the estimation becomes unbiased and the sensitivity saturates the Cramér-Rao bound $(\Delta\theta)^2_\theta = 1/mF(\theta)$. However, the number of measurements needed to saturate an arbitrary small Cramér-Rao bound is so large that the Heisenberg limit $\Delta\theta = 1/m\langle\hat{N}\rangle$ is not surpassed [26]. Analogous conclusions were reported in Ref. [8], showing no violation of the Heisenberg limit for the proposals [5–7].

4. Two-mode squeezed vacuum proposal

In Ref. [21] it is argued that the two-mode squeezed vacuum state can be used to overcome the Heisenberg limit in a Mach-Zehnder interferometer with parity detection in a single-output. As discussed in Sec. II C, the measurement of the parity of the number of particles in a single output port does not contain number coherences. Therefore, Eq. (4) applies here, for estimators that are unbiased in each fixed-N subspace. It is worth analyzing this example in detail.

The two-mode squeezed vacuum state is

$$|\psi\rangle = \sum_{N=0}^{+\infty} \frac{e^{-i\psi N}(\tanh r)^N}{\cosh r}|N,N\rangle, \tag{89}$$

where r is a squeezing parameter. We have $\langle\hat{N}\rangle = 2\sinh^2 r$, $\langle\hat{N}^2\rangle = 2\langle\hat{N}\rangle(\langle\hat{N}\rangle + 1)$ and $(\Delta N)^2 = \sinh^2 2r = \langle\hat{N}\rangle(\langle\hat{N}\rangle + 2)$. The Heisenberg limit (4) is

$$(\Delta\theta)_{\mathrm{HL}} = \max\left[\frac{1}{\sqrt{2m\langle\hat{N}\rangle(\langle\hat{N}\rangle+1)}}, \frac{1}{m\langle\hat{N}\rangle}\right]. \tag{90}$$

This can be compared to the quantum Cramér-Rao bound for rotations around the $\mathbf{y}$ axis, [corresponding to the Mach-Zehnder interferometer transformation, with quantum Fisher information $F_Q = 4(\Delta\hat{J}_y)^2$]:

$$(\Delta\theta)_{\mathrm{QCR}} = \frac{1}{\sqrt{m\langle\hat{N}\rangle(\langle\hat{N}\rangle+2)}}. \tag{91}$$

In Ref. [21] the sensitivity was calculated with an error propagation formula, that matches Eq. (91) at $\theta = 0$. Equation (91) overcomes $\Delta\theta = 1/\sqrt{m}\langle\hat{N}\rangle$ that is often indicated as the Heisenberg limit [21]. While this appears as the natural extension of Eq. (2) to the case of fluctuating number of particles (by replacing N with $\langle\hat{N}\rangle$), it is not a fundamental bound.

In the large-m limit Eq. (91) is always (for the interesting case $\langle\hat{N}\rangle > 1$) larger than Eq. (90). The two-mode squeezed vacuum state is very useful to overcome the shot-noise limit but it does not really surpass the Heisenberg limit (even if POVMs with number coherences are used). The saturation of the Heisenberg limit in the large-m limit Eq. (90) can be obtained with the superposition of NOON states $\sum_{N=0}^{+\infty} \frac{e^{-i\psi N}(\tanh r)^N}{\cosh r} \frac{|N,0\rangle+|0,N\rangle}{\sqrt{2}}$.

In the small-m limit, when the second term in Eq. (90) dominates over the first one (for $m = 1$ in particular), we find $(\Delta\theta)_{\mathrm{QCR}} \leqslant 1/\langle\hat{N}\rangle$. The apparent contradiction between our results and Ref. [21] is solved by noticing that Eq. (91) is known to be saturable (by the maximum likelihood estimator) only in the central limit. There is no guarantee (and not shown in Ref. [21]) that an unbiased estimator saturating Eq. (91) for small-m values can be found. The results of our paper show that such an estimator (unbiased in each fixed-N subspace) cannot exist: the central limit is reached for $m \geqslant 2 + 1/\langle\hat{N}\rangle$, where the first term in Eq. (90) dominates over the second one. If the phase is estimated with a POVM with coherences (that is not the case discussed in [21]), then Eq. (90) holds in the central limit, at least.

V. CONCLUSIONS

Phase estimation will likely become the first many-body quantum technology where classical bounds are overcome thanks to quantum correlations. It is therefore interesting to define the fundamental quantum bound (generally indicated as the Heisenberg limits), which limit the sensitivity of phase-estimation experiments. In this paper we have set the Heisenberg limit, Eq. (4), under relevant experimental conditions: fluctuating number of particles, absence of number coherences in the probe state and/or in the measurement strategy, and unbiased estimations. In this case we have also demonstrated that particle entanglement (we have extended the concept of particle entanglement to the case of state with fluctuating number of particles) is necessary to overcome the classical—shot-noise—phase uncertainty. If the probe state and the output measurement contain coherences between different total number of particles, it is not possible to establish a relation between entanglement and phase sensitivity. In this case, the phase-sensitivity bound Eq. (4) can only be set in the central limit.

ACKNOWLEDGMENTS

We thank R. Demkowicz-Dobrzański for discussions. L.P. acknowledges support by MIUR through FIRB Project No. RBFR08H058. P.H. acknowledges the support of the ERC starting grant GEDENTQOPT and the EU project QUASAR. This work has been supported by the Seventh Framework Programme for Research of the European Commission, under FET-Open grant QIBEC (No. 284584).

APPENDIX A: GENERAL TWO-MODE TRANSFORMATIONS

It is possible to write the general transformation (11) as the product of four matrices [60]:

$$\mathbf{U} = \begin{bmatrix} e^{-i\phi_0} & 0 \\ 0 & e^{-i\phi_0} \end{bmatrix} \times \begin{bmatrix} e^{-i\psi/2} & 0 \\ 0 & e^{i\psi/2} \end{bmatrix} \times \begin{bmatrix} \cos\frac{\vartheta}{2} & -\sin\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} & \cos\frac{\vartheta}{2} \end{bmatrix} \times \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{+i\phi/2} \end{bmatrix},$$

where $\phi_t = (\psi + \phi)/2$ and $\phi_r = (\psi - \phi)/2$. Using the Jordan-Schwinger representation of angular momentum, we have [3],

$$\mathbf{U}_x = \begin{bmatrix} \cos\frac{\vartheta}{2} & -i\sin\frac{\vartheta}{2} \\ -i\sin\frac{\vartheta}{2} & \cos\frac{\vartheta}{2} \end{bmatrix} \leftrightarrow \hat{U}_x = e^{-i\vartheta\hat{J}_x}, \quad \text{(A1)}$$

$$\mathbf{U}_y = \begin{bmatrix} \cos\frac{\vartheta}{2} & -\sin\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} & \cos\frac{\vartheta}{2} \end{bmatrix} \leftrightarrow \hat{U}_y = e^{-i\vartheta\hat{J}_y}, \quad \text{(A2)}$$

$$\mathbf{U}_z = \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{+i\phi/2} \end{bmatrix} \leftrightarrow \hat{U}_z = e^{-i\phi\hat{J}_z}, \quad \text{(A3)}$$

and using $e^{i\phi_0\hat{N}}\,\hat{a}\,e^{-i\phi_0\hat{N}} = e^{-i\phi_0}\hat{a}$, Eq. (11) can be associated to

$$\hat{\mathrm{U}}(\phi_0,\theta) = e^{-i\phi_0\hat{N}}e^{-i\psi\hat{J}_z}e^{-i\vartheta\hat{J}_y}e^{-i\phi\hat{J}_z}. \quad \text{(A4)}$$

By using the Euler-Rodrigues formula, Eq. (A4) can be rewritten as Eq. (12), where

$$\cos\frac{\theta}{2} = \cos\frac{\vartheta}{2}\cos\frac{\phi+\psi}{2},$$

and $\hat{J}_{\boldsymbol{n}} = \alpha\hat{J}_{\boldsymbol{x}} + \beta\hat{J}_{\boldsymbol{y}} + \gamma\hat{J}_{\boldsymbol{z}}$, with

$$\alpha = \frac{\sin\frac{\vartheta}{2}\sin\frac{\phi-\psi}{2}}{\sqrt{1-\cos^2\frac{\vartheta}{2}\cos^2\frac{\phi+\psi}{2}}},$$

$$\beta = \frac{\sin\frac{\vartheta}{2}\cos\frac{\phi-\psi}{2}}{\sqrt{1-\cos^2\frac{\vartheta}{2}\cos^2\frac{\phi+\psi}{2}}},$$

$$\gamma = \frac{\cos\frac{\vartheta}{2}\sin\frac{\phi+\psi}{2}}{\sqrt{1-\cos^2\frac{\vartheta}{2}\cos^2\frac{\phi+\psi}{2}}}.$$

This encompasses, for instance, the beam splitter [Eq. (A1), for $\phi = \pi/2$, $\psi = -\pi/2$ and $\vartheta = \theta$], Mach-Zehnder [Eq. (A2), for $\phi = \psi = 0$ and $\vartheta = \theta$], and phase shift [Eq. (A3), for $\psi = \vartheta = 0$ and $\phi = \theta$] transformations.

APPENDIX B: DERIVATION OF EQ. (20)

States without number coherences. The incoherent probe Eq. (9) transforms according to Eq. (12) as

$$\hat{\rho}_{\rm out}(\phi_0,\theta) = \sum_{N=0}^{+\infty} Q_N\,\hat{\mathrm{U}}(\phi_0,\theta)\,\hat{\rho}^{(N)}\,\hat{\mathrm{U}}(\phi_0,\theta)^\dagger = \sum_{N=0}^{+\infty} Q_N\,e^{-i\theta\hat{J}_{\boldsymbol{n}}}\hat{\rho}^{(N)}e^{+i\theta\hat{J}_{\boldsymbol{n}}}, \quad \text{(B1)}$$

as a consequence of $[\hat{\rho}^{(N)},\hat{N}] = 0$. Equation (B1) is a function of θ and shows that only SU(2) transformations, $e^{-i\theta\hat{J}_{\boldsymbol{n}}}$, are relevant for states without number coherence. Equation (20) follows from Eq. (B1), independently from the presence of number coherences in the POVM.

POVMs without number coherences. For the case of states with coherences, Eq. (6), and POVM without number coherences, Eq. (16), we have

$$\begin{aligned} P(\varepsilon|\theta) &= \sum_N \mathrm{Tr}[\hat{\pi}_N\,\hat{E}_N(\varepsilon)\,\hat{\pi}_N\,\hat{\mathrm{U}}(\phi_0,\theta)\,\hat{\rho}_{\rm coh}\,\hat{\mathrm{U}}(\phi_0,\theta)^\dagger] \\ &= \sum_N \mathrm{Tr}[\hat{E}_N(\varepsilon)\,\hat{\mathrm{U}}(\phi_0,\theta)\,\hat{\pi}_N\hat{\rho}_{\rm coh}\hat{\pi}_N\,\hat{\mathrm{U}}(\phi_0,\theta)^\dagger] \\ &= \sum_N Q_N\mathrm{Tr}[\hat{E}_N(\varepsilon)\hat{\mathrm{U}}(\phi_0,\theta)\hat{\rho}^{(N)}\hat{\mathrm{U}}(\phi_0,\theta)^\dagger] \\ &= \sum_N Q_N\mathrm{Tr}\big[\hat{E}_N(\varepsilon)e^{-i\theta\hat{J}_{\boldsymbol{n}}^{(N)}}\hat{\rho}^{(N)}e^{+i\theta\hat{J}_{\boldsymbol{n}}^{(N)}}\big] \\ &= \sum_N Q_N\,P(\varepsilon|N,\theta), \end{aligned} \quad \text{(B2)}$$

where $P(\varepsilon|N,\theta) = \mathrm{Tr}[\hat{E}_N(\varepsilon)e^{-i\theta\hat{J}_{\boldsymbol{n}}^{(N)}}\hat{\rho}^{(N)}e^{+i\theta\hat{J}_{\boldsymbol{n}}^{(N)}}]$. To derive this result we have used the commutation relation $[\hat{\mathrm{U}},\hat{\pi}_N] = 0$ and $\hat{\pi}_N\hat{\rho}_{\rm coh}\hat{\pi}_N = Q_N\hat{\rho}^{(N)}$, where $\hat{\rho}^{(N)}$ is a density matrix defined on the fixed-N subspace. We have also used $\hat{\pi}_N\hat{\mathrm{U}}(\phi_0,\theta)\hat{\pi}_N = e^{-i\phi_0 N}e^{-i\theta\hat{J}_{\boldsymbol{n}}^{(N)}}$ due to $\hat{J}_{\boldsymbol{n}} = \oplus_N\hat{J}_{\boldsymbol{n}}^{(N)}$, where $\hat{J}_{\boldsymbol{n}}^{(N)}$ acts on the fixed-N subspace. When POVM as in Eq. (16) are used, we can therefore conclude, from Eq. (B2), that only SU(2) transformations are relevant and number coherences in the probe state do not play any role.

[1] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[2] R. S. Bondurant and J. H. Shapiro, Phys. Rev. D **30**, 2548 (1984).

[3] B. Yurke, S. L. McCall, and J. R. Klauder, Phys. Rev. A **33**, 4033 (1986).

[4] G. S. Summy and D. T. Pegg, Opt. Comm. **77**, 75 (1990).
[5] J. H. Shapiro, S. R. Shepard, and N. C. Wong, Phys. Rev. Lett. **62**, 2377 (1989).
[6] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (1991).
[7] W. P. Schleich, J. P. Dowling and R. J. Horowicz, Phys. Rev. A **44**, 3365 (1991).
[8] S. L. Braunstein, A. S. Lane, and C. M. Caves, Phys. Rev. Lett. **69**, 2153 (1992); S. L. Braunstein, J. Phys A: Math. Gen. **25**, 3813 (1992); Phys. Rev. Lett. **69**, 3598 (1992); A. S. Lane, S. L. Braunstein, and C. M. Caves, Phys. Rev. A **47**, 1667 (1993).
[9] M. J. W. Hall, J. Mod. Opt. **40**, 809 (1993).
[10] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[11] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys. Rev. A **50**, 67 (1994).
[12] Z. Hradil, Quant. Opt. **4**, 93 (1992).
[13] Z. Hradil, Phys. Rev. A **51**, 1870 (1995).
[14] Z. Y. Ou, Phys. Rev. Lett. **77**, 2352 (1996); Phys. Rev. A **55**, 2598 (1997).
[15] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[16] L. Pezzè and A. Smerzi, Phys. Rev. Lett. **102**, 100401 (2009).
[17] P. Hyllus, L. Pezzè, and A. Smerzi, Phys. Rev. Lett. **105**, 120501 (2010).
[18] H. F. Hofmann, Phys. Rev. A **79**, 033822 (2009).
[19] G. A. Durkin and J. P. Dowling, Phys. Rev. Lett. **99**, 070801 (2007).
[20] A. Monras, Phys. Rev. A **73**, 033821 (2006).
[21] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, Phys. Rev. Lett. **104**, 103602 (2010).
[22] Y. R. Zhang, G. R. Jin, J. P. Cao, W. M. Liu, and H. Fan, J. Phys. A **46**, 035302 (2013).
[23] J. Joo, W. J. Munro, and T. P. Spiller, Phys. Rev. Lett. **107**, 083601 (2011).
[24] M. Hayashi, Progress in Informatics **8**, 81 (2011).
[25] A. Rivas and A. Luis, New J. Phys. **14**, 093052 (2012).
[26] L. Pezzè, Phys. Rev. A **88**, 060101(R) (2013).
[27] F. Benatti and D. Braun, Phys. Rev. A **87**, 012340 (2013).
[28] M. Jarzyna and R. Demkowicz-Dobrzański, Phys. Rev. A **85**, 011801(R) (2012).
[29] V. Giovannetti and L. Maccone, Phys. Rev. Lett. **108**, 210404 (2012).
[30] M. Tsang, Phys. Rev. Lett. **108**, 230401 (2012).
[31] M. J. W. Hall, D. W. Berry, M. Zwierz, and H. M. Wiseman, Phys. Rev. A **85**, 041802(R) (2012).
[32] D. W. Berry, M. J. W. Hall, M. Zwierz, and H. M. Wiseman, Phys. Rev. A **86**, 053813 (2012).
[33] M. J. W. Hall and H. M. Wiseman, New J. Phys. **14**, 033040 (2012).
[34] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **108**, 260405 (2012).
[35] Y. Gao and H. Lee, J. Phys. A **45**, 415306 (2012).
[36] M. Zwierz, C. A. Perez-Delgado, and P. Kok, Phys. Rev. Lett. **105**, 180402 (2010); Phys. Rev. A **85**, 042112 (2012).
[37] V. Meyer *et al.*, Phys. Rev. Lett. **86**, 5870 (2001); D. Leibfried *et al.*, Science **304**, 1476 (2004); Nature (London) **438**, 639 (2005); C. F. Roos *et al.*, *ibid.* **443**, 316 (2006); W.-B. Gao *et al.*, Nature Phys. **6**, 331 (2010); T. Monz *et al.*, Phys. Rev. Lett. **106**, 130506 (2011).
[38] J. Appel *et al.*, PNAS **106**, 10960 (2009); A. Louchet-Chauvet *et al.*, New J. Phys. **12**, 065032 (2010); M. H. Schleier-Smith, I. D. Leroux, and V. Vuletic, Phys. Rev. Lett. **104**, 073604 (2010); Z. Chen, J. G. Bohnet, S. R. Sankar, J. Dai, and J. K. Thompson, *ibid.* **106**, 133601 (2011); S. Bernon *et al.*, New J. Phys. **13**, 065021 (2011); J. C. Bohnet *et al.*, Nature Photon. **8**, 731 (2014).
[39] J. Estève *et al.*, Nature (London) **455**, 1216 (2008); C. Gross *et al.*, *ibid.* **464**, 1165 (2010); M. F. Riedel *et al.*, *ibid.* **464**, 1170 (2010); B. Lücke *et al.*, Science **334**, 773 (2011); C. F. Ockeloen, R. Schmied, M. F. Riedel, and P. Treutlein, Phys. Rev. Lett. **111**, 143001 (2013); T. Berrada *et al.*, Nature Comm. **4**, 2077 (2013); S. van Frank *et al.*, *ibid.* **5**, 4099 (2014); H. Strobel *et al.*, Science **345**, 424 (2014).
[40] P. Walther *et al.*, Nature (London) **429**, 158 (2004); M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Nature **429**, 161 (2004); T. Nagata *et al.*, Science **316**, 726 (2007); M. Kacprowicz *et al.*, Nature Photon. **4**, 357 (2010); I. Afek, O. Ambar, and Y. Silberberg, Science **328**, 879 (2010); G. Y. Xiang *et al.*, Nature Photon. **5**, 43 (2011); R. Krischek *et al.*, Phys. Rev. Lett. **107**, 080504 (2011).
[41] R. Schnabel, N. Mavalvala, D. E. McClelland, and P. K. Lam, Nature Commun. **1**, 121 (2010); The LIGO Scientific Collaboration, Nature Phys. **7**, 962 (2011); H. Grote, K. Danzmann, K. L. Dooley, R. Schnabel, J. Slutsky, and H. Vahlbruch, Phys. Rev. Lett. **110**, 181101 (2013).
[42] J. J. Bollinger and W. M. Itano, D. J. Wineland and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).
[43] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).
[44] C. C. Gerry and R. A. Campos, Phys. Rev. A **68**, 025602 (2003); R. A. Campos, Christopher C. Gerry, and A. Benmoussa, *ibid.* **68**, 023810 (2003).
[45] L. Pezzè and A. Smerzi, Europhys. Lett. **78**, 30004 (2007).
[46] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, Phys. Rev. A **80**, 052114 (2009).
[47] There exist several protocols of phase estimation reaching $\Delta\theta \propto 1/\langle\hat{N}\rangle\sqrt{m}$ with states of a fluctuating number of particles and large m, see for instance Refs. [2,3,48–50].
[48] C. C. Gerry, Phys. Rev. A **61**, 043811 (2000).
[49] L. Pezzè and A. Smerzi, Phys. Rev. Lett. **110**, 163604 (2013).
[50] L. Pezzè and A. Smerzi, Phys. Rev. Lett. **100**, 073601 (2008).
[51] A. Shaji and C. M. Caves, Phys. Rev. A **76**, 032111 (2007).
[52] U. Dorner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Phys. Rev. Lett. **102**, 040403 (2009).
[53] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nature Phys. **7**, 406 (2011).
[54] R. Demkowicz-Dobrzanski, J. Kolodynski, and M. Guta, Nature Commun. **3**, 1063 (2012).
[55] U. Dorner, New J. Phys. **14**, 043011 (2012).
[56] M. Landini, M. Fattori, L. Pezzè, and A. Smerzi, New J. Phys. **16**, 113074 (2014).
[57] S. D. Bartlett, T. Rudolph, and R. Spekkens, Rev. Mod. Phys. **79**, 555 (2007); B. C. Sanders, S. D. Bartlett, T. Rudolph, and P. L. Knight, Phys. Rev. A **68**, 042329 (2003); K. Mølmer, *ibid.* **55**, 3195 (1997).

[58] G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev. **88**, 101 (1952).

[59] Let us consider a generic density matrix $\hat{\rho}$ as in Eq. (6) and recall that $|\psi_{N,k}\rangle$ is a state of N particles [it fulfills $\hat{N}|\psi_{N,k}\rangle = N|\psi_{N,k}\rangle$]. The commutator $[\hat{\rho},\hat{N}] = \sum_k p_k \sum_{N,N'} \sqrt{Q_{N,k}Q^*_{N',k}}(N'-N)|\psi_{N,k}\rangle\langle\psi_{N',k}|$ vanishes if and only if $\sqrt{Q_{N,k}Q^*_{N',k}} = |Q_{N,k}|\delta_{N,N'}$, *i.e.* if and only if $\hat{\rho}$ can be written as in Eq. (9), with $Q_N\hat{\rho}^{(N)} = \sum_k p_k|Q_{N,k}||\psi_{N,k}\rangle\langle\psi_{N,k}| = \hat{\pi}_N\hat{\rho}\hat{\pi}_N$.

[60] R. A. Campos, B. E. A. Saleh, and M. C. Teich, Phys. Rev. A **40**, 1371 (1989).

[61] U. Leonhardt, Phys. Rev. A **48**, 3265 (1993).

[62] J. Schwinger, in *Quantum Theory of Angular Momentum*, edited by L. Beidenharn and H. van Dam (Academic, New York, 1965).

[63] C. M. Helstrom, Phys. Lett. A **25**, 101 (2001); *Quantum detection and estimation theory* (Academic Press, New York, 1976).

[64] Since $\sum_N \hat{\pi}_N = \mathbb{1}$, we can rewrite the POVM $\hat{E}(\varepsilon)$ as $\hat{E}(\varepsilon) = \sum_{N,N'} \hat{\pi}_N \hat{E}(\varepsilon)\hat{\pi}_{N'}$. Using $\hat{N}\hat{\pi}_N = N\hat{\pi}_N$, we obtain $[\hat{E}(\varepsilon),\hat{N}] = \sum_{N,N'}(N'-N)\hat{\pi}_N\hat{E}(\varepsilon)\hat{\pi}_{N'}$. This vanishes if and only if $\hat{\pi}_N\hat{E}(\varepsilon)\hat{\pi}_{N'} = \delta_{N,N'}\hat{\pi}_N\hat{E}(\varepsilon)\hat{\pi}_N$, i.e., if and only if $\hat{E}(\varepsilon)$ has the form of Eq. (16).

[65] C. C. Gerry and J. Mimih, Cont. Phys. **51**, 497 (2010).

[66] K. P. Seshadreesan, S. Kim, J. P. Dowling and Hwang Lee, Phys. Rev. A **87**, 043833 (2013).

[67] For a review of the quantum theory of phase estimation see L. Pezzè and A. Smerzi, in *Atom Interferometry*, Proceedings of the International School of Physics Enrico Fermi, Course 188, edited by G. M. Tino and M. A. Kasevich (IOS Press, Amsterdam, 2014); arXiv:1411.5164.

[68] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, New Jersey, 1993), ch. 3.

[69] The inequality (22) is saturated if and only if

$$\boldsymbol{v}^\top[\boldsymbol{\Theta}(\boldsymbol{\varepsilon}) - \bar{\boldsymbol{\Theta}}] = h(\boldsymbol{\theta})\left[\frac{\partial}{\partial\boldsymbol{\theta}}\log\mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})\right]^\top\boldsymbol{u}, \quad \forall\boldsymbol{\varepsilon},\boldsymbol{\theta}, \quad \text{(B3)}$$

where the function $h(\boldsymbol{\theta})$ does not depend on $\boldsymbol{\varepsilon}$.

[70] The multiplication factor m, the number of independent measurements, follows from

$$\int d\boldsymbol{\varepsilon}\,\frac{1}{\mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}\left(\frac{\partial\mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}{\partial\theta_i}\right)\left(\frac{\partial\mathcal{L}(\boldsymbol{\varepsilon}|\boldsymbol{\theta})}{\partial\theta_j}\right) = m\mathbf{F}_{i,j}, \quad \text{(B4)}$$

as a direct consequence of $\int d\varepsilon\,\frac{\partial P(\varepsilon|\theta)}{\partial\theta} = 0$.

[71] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946); C. R. Rao, Bull. Calcutta Math. Soc. **37**, 81 (1945); M. Frechet, Revue de l'Institut International de Statistique **11**, 182 (1943).

[72] We use the well-known result that every positive definite matrix is invertible and its inverse is also positive definite. In particular, if $\mathbf{A} \geqslant \mathbf{B} > 0$, then $\mathbf{B}^{-1} \geqslant \mathbf{A}^{-1} > 0$.

[73] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[74] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzè, and A. Smerzi, Phys. Rev. A **85**, 022321 (2012).

[75] G. Tòth, Phys. Rev. A **85**, 022322 (2012).

[76] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[77] O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).

[78] O. Gühne, G. Tóth, and H. J. Briegel, New J. Phys. **7**, 229 (2005); O. Gühne and G. Tóth, Phys. Rev. A **73**, 052319 (2006).

[79] M. Seevinck and J. Uffink, Phys. Rev. A **65**, 012107 (2001); G. Tóth, O. Gühne, M. Seevinck, and J. Uffink, *ibid.* **72**, 014101 (2005).

[80] Z. Chen, Phys. Rev. A **71**, 052302 (2005).

[81] A. S. Sørensen and K. Mølmer, Phys. Rev. Lett. **86**, 4431 (2001).

[82] P. Hyllus, L. Pezzè, A. Smerzi, and G. Tóth, Phys. Rev. A **86**, 012337 (2012).

[83] R. Demkowicz-Dobrzański (private communication).

[84] M. G. A. Paris, Int. J. Quantum Inform. **7**, 125 (2009).

[85] *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, edited by H. L. Van Trees and K. L. Bell (Wiley-IEEE, Piscataway, 2007).