Quantum-limited metrology and Bose-Einstein condensates

Sergio Boixo,[1] Animesh Datta,[2,3] Matthew J. Davis,[4] Anil Shaji,[5,*] Alexandre B. Tacla,[5] and Carlton M. Caves[5]

[1]*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA*

[2]*Institute for Mathematical Sciences, Imperial College, 53 Prince's Gate, London SW7 2PG, United Kingdom*

[3]*QOLS, The Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BW, United Kingdom*

[4]*School of Mathematics and Physics, ARC Centre of Excellence for Quantum-Atom Optics,*
The University of Queensland, Queensland 4072, Australia

[5]*Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131-0001, USA*

We discuss a quantum-metrology protocol designed to estimate a physical parameter in a Bose-Einstein condensate of N atoms, and we show that the measurement uncertainty can decrease faster than $1/N$. The $1/N$ scaling is usually thought to be the best possible in any measurement scheme. From the perspective of quantum information theory, we outline the main idea that leads to a measurement uncertainty that scales better than $1/N$. We examine in detail some potential problems and challenges that arise in implementing such a measurement protocol using a Bose-Einstein condensate. We discuss how some of these issues can be dealt with by using lower-dimensional condensates trapped in nonharmonic potentials.

I. INTRODUCTION

In quantum metrology, the description "Heisenberg-limited scaling" refers to the best possible scaling of the measurement uncertainty with the resources put into a measurement. The phrase arises not from Heisenberg uncertainty relations, but from uncertainty relations of the Mandelstam-Tamm type [1],

$$\delta\gamma \langle \Delta^2 K \rangle^{1/2} \geq \frac{1}{2}, \tag{1.1}$$

in units with $\hbar = 1$. The uncertainty $\delta\gamma$ in a parameter γ that, in part, determines the state of a quantum system is related to the standard deviation of the operator K that generates translations of the state along a path parameterized by γ. A sequence of logical and mathematical steps is needed to provide a rigorous connection between the problem of measurement uncertainty in quantum metrology and uncertainty relations of the Mandelstam-Tamm type. The pioneering work of Helstrom [2], Holevo [3], Braunstein, Caves, and Milburn [4,5], and others laid out and elucidated these steps. We summarize them below for the sake of completeness.

The discussion in this paper is restricted to single-parameter estimation. The first step in estimating the value of a parameter is to identify an elementary physical system that is sensitive to changes in the parameter, just as one would choose a balance to measure weight or a thermometer to measure temperature. One or more of these elementary systems make up the measuring device or *probe*. The measurement uncertainty is a property of this measuring device. In quantum metrology this means that we expect the measurement uncertainty to depend on the initial state of the quantum probe, its evolution, and the measurement made on the probe to extract information about the parameter. The

quantum Cramér-Rao bound quantifies the idea that the optimal measurement uncertainty is inversely proportional to the change in the state of the probe corresponding to small changes in the value of the parameter

$$(\delta\gamma)^2 \geq \frac{1}{(ds_{\mathrm{DO}}/d\gamma)^2} = \frac{1}{\Im(\gamma,t)}. \tag{1.2}$$

Here ds_{DO} denotes a distance element in the space of density operators of the probe and $\Im(\gamma,t)$ is the quantum Fisher information. The uncertainty in determining γ is quantified by the units-corrected root-mean-square deviation of one's estimate of the parameter, γ_{est}, from the true value γ,

$$\delta\gamma = \left\langle \left(\frac{\gamma_{\mathrm{est}}}{|\partial\langle\gamma_{\mathrm{est}}\rangle/\partial\gamma|} - \gamma \right)^2 \right\rangle^{1/2}. \tag{1.3}$$

In classical statistics, the Cramér-Rao bound on measurement uncertainty is given by

$$(\delta\gamma)^2 \geq \frac{1}{\mathcal{I}(\gamma)}, \tag{1.4}$$

where

$$\mathcal{I}(\gamma) \equiv \left\langle \left(\frac{\partial}{\partial\gamma} \ln p(\zeta|\gamma) \right)^2 \right\rangle, \tag{1.5}$$

called the *Fisher information*, is an average over the probability distribution $p(\zeta|\gamma)$ for a random variable ζ and is a measure of the information that ζ can provide about γ. The classical Cramér-Rao bound (1.4) can generally be achieved only asymptotically in a large number of trials, i.e., independent measurements of ζ. The requirement of many trials to achieve the Cramér-Rao bound is important, but as a purely classical effect, it is not germane to our discussion of quantum limits, so we do not consider it further in the remainder of this paper.

---

*shaji@unm.edu

If the probe used to estimate the value of γ were a classical system, then ζ would label the possible states of the probe at the end of the measurement, with $p(\zeta|\gamma)$ being the probability of finding the probe in each of these states. For a quantum probe in a state $\rho(\gamma,t)$ at the end of the measurement process, ζ labels the possible outcomes of a measurement performed on the probe, which is described by POVM elements $E(\zeta)$, with $\int d\zeta E(\zeta)=\mathbb{1}$ and $p(\zeta|\gamma,t)$ $=\mathrm{tr}[E(\zeta)\rho(\gamma,t)]$. The classical Fisher information $\mathcal{I}(\gamma,t)$, defined using $p(\zeta|\gamma,t)$, clearly depends on the choice of POVM. The quantum Fisher information, which is independent of the choice of POVM, is therefore defined as

$$\mathfrak{I}(\gamma,t) \equiv \max_{E(\zeta)} \mathcal{I}(\gamma,t). \tag{1.6}$$

The maximization over all possible measurements in the above equation is a rather daunting prospect, but it can be shown that [2–5]

$$\mathfrak{I}(\gamma,t) = \mathrm{tr}[\rho(\gamma,t)\mathfrak{L}^2(\gamma,t)] = \langle \mathfrak{L}^2(\gamma,t) \rangle. \tag{1.7}$$

The symmetric logarithmic derivative, $\mathfrak{L}(\gamma,t)$, is the Hermitian operator defined implicitly by the equation

$$\frac{1}{2}(\mathfrak{L}\rho + \rho\mathfrak{L}) = \frac{\partial\rho}{\partial\gamma}. \tag{1.8}$$

We now make two simplifying assumptions. First, we assume that translations in the parameter are generated by a unitary operator. This allows us to characterize the translations in terms of a Hermitian generator $K(\gamma,t)$ defined by

$$\frac{\partial\rho(\gamma,t)}{\partial\gamma} = -i[K(\gamma,t),\rho(\gamma,t)]. \tag{1.9}$$

Second, we assume that the state of the probe is pure, which implies $\rho^2(\gamma,t)=\rho(\gamma,t)$. Under these two assumptions, we can identify the symmetric logarithmic derivative as

$$\mathfrak{L}(\gamma,t) = 2\frac{\partial\rho}{\partial\gamma} = -2i[K(\gamma,t),\rho(\gamma,t)], \tag{1.10}$$

and the quantum Fisher information becomes

$$\mathfrak{I}(\gamma,t) = 4\,\mathrm{tr}(\rho K^2 - \rho K\rho K) = 4\langle \Delta^2 K(\gamma,t) \rangle. \tag{1.11}$$

Thus, using Eq. (1.2), we obtain

$$\delta\gamma \geq \frac{1}{2\langle \Delta^2 K(\gamma,t) \rangle^{1/2}}, \tag{1.12}$$

which is a rigorous statement of the Mandelstam-Tamm uncertainty relation (1.1). Our two simplifying assumptions can be relaxed [2–5], but we do not need the more general forms of the Cramér-Rao bound in this paper.

We can further simplify Eq. (1.12) by noting that the variance of a Hermitian operator is bounded from above by $\langle \Delta^2 K \rangle \leq \|K\|^2/4$, where $\|\cdot\|$ is defined as the difference between the largest and smallest eigenvalues of a Hermitian operator (this is a seminorm for Hermitian operators). The quantum Cramér-Rao bound then becomes

$$\delta\gamma \geq \frac{1}{\|K(\gamma,t)\|}. \tag{1.13}$$

We mentioned the elementary quantum systems, sensitive to γ, that are used to build the probe. The number N of such elementary units of the probe can be regarded as the most significant resource that goes into a measurement scheme. The differences between the tensor-product state space of a composite quantum system of N probe units and the Cartesian-product state space of an equivalent classical composite system is the motivation for investigating whether a composite quantum probe offers advantages over classical ones in the relationship between $\delta\gamma$ and N.

From Eq. (1.13) we see that theoretically the N dependence of the bound on $\delta\gamma$ comes solely from the dependence of the generator K on N. In any particular quantum-metrology scheme, however, this bound might not be achievable, and additional dependence of $\delta\gamma$ on N can come from the nature of the state of the probe as well. To see these dependences clearly and to understand what "Heisenberg-limited scaling" means, we view quantum metrology from the perspective of quantum information theory using the language of quantum circuits in Sec. II. We also explain how one can construct measurement protocols in which $\delta\gamma$ scales with N in a manner not thought to be possible until recently. Section III examines in some detail how such an enhanced metrology protocol might be implemented in a Bose-Einstein condensate (BEC) of N atoms and considers various problems and issues that might arise in a BEC realization of the proposed metrology scheme.

II. QUANTUM METROLOGY FROM AN INFORMATION-THEORETIC PERSPECTIVE

In this section, we follow Giovannetti *et al.* [6] in using quantum circuits to describe and analyze metrology protocols. From this perspective, we first look at a couple of well-known measurement schemes that were considered in [6]—Ramsey interferometry (Sec. II A) and interferometry using a Schrödinger-cat state (superposition of macroscopically distinct states) (Sec. II B)—with the aim of generalizing these circuits to new protocols that were introduced in [7,8]. In these initial discussions of Ramsey interferometry and cat-state interferometry, we assume that the elementary quantum systems that make up the probe are qubits. The quantization axis is taken to be along the z direction of a Bloch-sphere representation, with the standard basis states along this direction denoted as $|0\rangle$ and $|1\rangle$. Despite the notation the actual qubits need not be spin-1/2 particles; they could very well be atoms in which only two energy levels are relevant or a variety of other suitable systems. In the subsequent general discussions of linear and nonlinear interferometries (Secs. II C and II D), we allow the probe units to be any quantum system. It turns out, however, that optimal sensitivities are always attained by using only two levels of each unit, so in the end we can always regard the probe units as qubits.

In all the quantum circuits depicted in this section, we use $N=3$ probe units as an example.

A. Ramsey interferometry

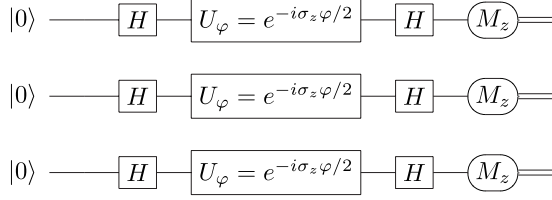A typical Ramsey interferometer, such as the one in [9], can be represented by the quantum circuit in Fig. 1. In this

FIG. 1. Quantum circuit for Ramsey interferometry.

FIG. 2. Quantum circuit for cat-state interferometry.

measurement protocol, each of the N qubits that make up the probe evolves independently. All the qubits are initialized in the state $|0\rangle$, which might represent the ground state in Ramsey interferometry using atoms. The Hadamard gate H puts each of the qubits in an equal superposition of the two basis states, $(|0\rangle+|1\rangle)/\sqrt{2}$. The parameter-dependent evolution of the quantum probe is generated by the Hamiltonian

$$H_{\text{Ramsey}} = \gamma \sum_{j=1}^{N} \sigma_{z;j}/2 = \gamma J_z, \tag{2.1}$$

where $\sigma_{z;j}$ denotes the σ_z operator acting on the jth probe qubit and J_z is the z component of the "total angular momentum" for all the qubits.

Evolution under this Hamiltonian for a time t introduces a relative phase $\varphi \equiv \gamma t$ between the two components of the superposition, changing the state of the probe qubits to $(e^{-i\varphi/2}|0\rangle+e^{i\varphi/2}|1\rangle)/\sqrt{2}$. The last set of Hadamard gates changes the parameter-dependent phases in the superposition into amplitude (population) information. Thus the state of the probe qubits just before the readout is $\cos(\varphi/2)|0\rangle + \sin(\varphi/2)|1\rangle$. The final readout in Ramsey interferometry is done by measuring each of the qubits along the z direction. This leads to a measured signal

$$\langle J_z \rangle \equiv \left\langle \frac{1}{2}\sum_{j=1}^{N}\sigma_{z;j} \right\rangle = \frac{1}{2}N\cos\varphi. \tag{2.2}$$

The variance in the signal is

$$\langle \Delta^2 J_z \rangle = \frac{1}{4}N\langle \Delta^2 \sigma_z \rangle = \frac{1}{4}N\sin^2\varphi. \tag{2.3}$$

The uncertainty in the estimate of γ from the measured signal in Eq. (2.2) is

$$\delta\gamma_{\text{Ramsey}} = \frac{\langle \Delta^2 J_z \rangle^{1/2}}{|d\langle J_z\rangle/d\gamma|} = \frac{1}{t\sqrt{N}}. \tag{2.4}$$

For the Ramsey Hamiltonian (2.1), the generator of translations in γ is tJ_z. Thus, according to the quantum Cramér-Rao bound (1.13), the measurement uncertainty is bounded from below by

$$\delta\gamma \geq \frac{1}{t\|J_z\|} = \frac{1}{tN\|\sigma_z/2\|} = \frac{1}{tN}. \tag{2.5}$$

The Ramsey interferometer described here does not achieve the best measurement uncertainty given by the quantum Cramér-Rao bound. The Hamiltonian H_{Ramsey} that governs the evolution of the probe qubits is fixed by the choice of physical systems that are the qubits. Given a choice of probe

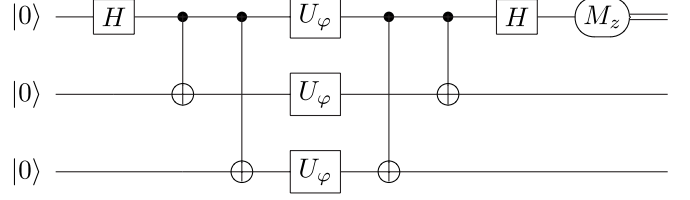qubits, however, we still have the freedom to choose an optimal initial state for the probe and an optimal measurement of the qubits to minimize the measurement uncertainty. It turns out that the best possible scaling for the measurement uncertainty can be achieved if the probe is initialized in an entangled "Schrödinger-cat" state [10,11].

B. Cat-state interferometry

The quantum circuit that uses a probe initialized in a Schrödinger-cat state is depicted in Fig. 2. The Hadamard gate on the first qubit, followed by the controlled-NOT gates to the remaining qubits, initializes the probe in the state $|\text{cat}\rangle = (|0\ldots0\rangle+|1\ldots1\rangle)/\sqrt{2}$. This state is often referred to as the Schrödinger-cat state because when the number of qubits is large, it is a superposition of two macroscopically distinct states.

The probe qubits evolve under the same Hamiltonian [Eq. (2.1)] as in Ramsey interferometry. The parameter-dependent evolution of the probe for a duration t changes the probe state to $(e^{-iN\varphi/2}|0\ldots0\rangle+e^{iN\varphi/2}|1\ldots1\rangle)/\sqrt{2}$, where $\varphi = \gamma t$. After the parameter-dependent evolution, one way to handle the readout, depicted in the circuit above, is to subject the qubits to a sequence of gates that kick the phases picked up by the two components of the cat state into amplitudes on the first qubit, so that the state of the probe just before readout is $[\cos(N\varphi/2)|0\rangle+\sin(N\varphi/2)|1\rangle]\otimes|0\rangle^{\otimes(N-1)}$. The readout of the probe can then be performed by measuring the σ_z operator on the first qubit. This leads to a measured signal and variance given by

$$\langle \sigma_{z;1} \rangle = \cos N\varphi \quad \text{and} \quad \langle \Delta^2 \sigma_{z;1} \rangle = \sin^2 N\varphi. \tag{2.6}$$

The frequency of the γ-dependent fringe in cat-state interferometry is N times greater than the frequency of the signal in ordinary Ramsey interferometry. This leads to an enhanced sensitivity in the estimate of γ in cat-state interferometry, which achieves the Cramér-Rao bound:

$$\delta\gamma_{\text{cat}} = \frac{\langle \Delta^2 \sigma_{z;1} \rangle^{1/2}}{|d\langle \sigma_{z;1}\rangle/d\gamma|} = \frac{1}{tN}. \tag{2.7}$$

C. Heisenberg-limited metrology with linear Hamiltonians

We can put our interferometry circuits in a general setting by considering the case in which the probe units are arbitrary systems and the probe Hamiltonian is of the form
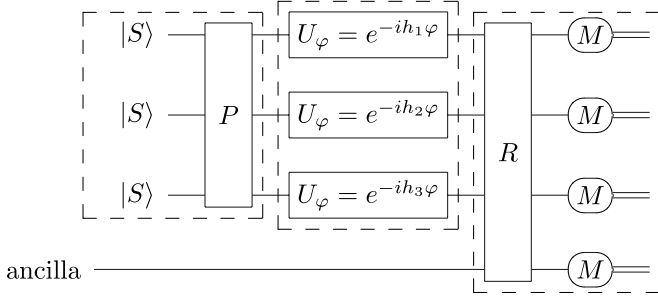
FIG. 3. Quantum circuit for a general linear interferometer.

$$H_{\text{linear}} = \gamma h_{\text{linear}} = \gamma \sum_{j=1}^{N} h_j. \qquad (2.8)$$

Here the operators h_j denote identical couplings to the probe units; the use of independent couplings to the parameter is the source of our appelation "linear" for this Hamiltonian. The generator of translations in γ is $K(\gamma,t) = t h_{\text{linear}}$, so the quantum Cramér-Rao bound (1.13) on the uncertainty in a determination of γ takes the form

$$\delta\gamma \geq \frac{1}{t \|h_{\text{linear}}\|} = \frac{1}{tN(\Lambda - \lambda)}, \qquad (2.9)$$

where Λ and λ are the largest and smallest eigenvalues of the single-unit operators h_j. Achieving the Cramér-Rao bound only requires using two levels of each unit, the eigenstates $|\Lambda\rangle$ and $|\lambda\rangle$ corresponding to the largest and smallest eigenvalues of the operators h_j, so we can always regard the units as qubits with $|0\rangle = |\Lambda\rangle$ and $|1\rangle = |\lambda\rangle$.

The quantum circuit that represents a measurement protocol of this sort is drawn in Fig. 3. The dashed boxes highlight the three stages of this protocol: probe preparation, dynamics, and readout. All the probe units begin in a standard state $|S\rangle$. The arbitrary unitary operator P can then prepare any initial state as input to the dynamics. In the dynamics stage, the gates U_φ imprint information about the parameter on the probe. The final readout stage includes an arbitrary unitary interaction R among the probe units and with an arbitrary ancilla system. This unitary followed by measurements on each subsystem in a standard basis can be used to perform any quantum measurement. The quantum Cramér-Rao bound (2.9) applies to all circuits of the above form. Indeed, the bound actually applies to somewhat more general situations in which the unitary operator R is interleaved with the gate dynamics and the results of ancilla measurements are fed back onto the probe [7].

If the preparation unitary P is omitted from the circuit, making the input to the dynamics a product state, then the uncertainty in the generator of γ displacements is bounded by $\langle \Delta^2 K \rangle^{1/2} \leq t\sqrt{N}(\Lambda - \lambda)/2$. The resulting bound on measurement uncertainty, from Eq. (1.12), is

$$\delta\gamma \geq \frac{1}{t\sqrt{N}(\Lambda - \lambda)} \equiv \delta\gamma_{\text{QNL}}. \qquad (2.10)$$

This bound, a general form of that for standard Ramsey interferometry, is called the *quantum noise limit* (QNL) or

the *shot-noise limit*. The optimal $1/t\sqrt{N}$ sensitivity for product-state inputs can be achieved by using initial state $|S\rangle = (|\Lambda\rangle + |\lambda\rangle)/\sqrt{2}$ for each unit and by making a final product measurement of an equatorial-plane spin component on each unit (in the qubit Bloch sphere formed from $|0\rangle = |\Lambda\rangle$ and $|1\rangle = |\lambda\rangle$).

One can achieve the Cramér-Rao bound (2.9) by operating the circuit in a way that takes advantage of entangled input states. The preparation operator is chosen to take the initial product of standard states to the "catlike" state $(|\Lambda, \ldots, \Lambda\rangle + |\lambda, \ldots, \lambda\rangle)/\sqrt{2}$. In the dynamics stage, this "cat-like" initial state is subject to a period of parameter-dependent evolution that changes it to $(e^{-iN\Lambda\varphi}|\Lambda, \ldots, \Lambda\rangle + e^{iN\lambda\varphi}|\lambda, \ldots, \lambda\rangle)/\sqrt{2}$. The readout process kicks back the differential phase shift into amplitude information, which produces fringes with frequency proportional to $N(\Lambda - \lambda)$, thus achieving the optimal measurement uncertainty,

$$\delta\gamma = \frac{1}{tN(\Lambda - \lambda)} \equiv \delta\gamma_{\text{HL}}, \qquad (2.11)$$

of the Cramér-Rao bound (2.9). This optimal measurement uncertainty, a general form of that for cat-state interferometry, is often called the *Heisenberg limit*.

The general quantum-metrology scheme considered in this subsection indicates that probe preparation gives an enhancement of $1/\sqrt{N}$ over the case where the probe qubits are initialized in a product state. Readout has already been optimized to take advantage of this entangled input, so we conclude that when the parameter-dependent dynamics acts independently on the probe qubits, Heisenberg-limited scaling is indeed the $1/N$ scaling. *The one remaining way of exploring whether the $1/N$ scaling can be improved is to consider more general dynamics* [7,8,12–17]; we turn to that possibility in the next subsection.

D. Heisenberg-limited metrology with nonlinear Hamiltonians

A generalized quantum-metrology scheme in which the dynamics of the probe is generated by a Hamiltonian that includes all k-body couplings between the probe qubits was first considered in [7]. This nonlinear coupling Hamiltonian has the form

$$H_{\text{nonlinear}} = \gamma h_{\text{nonlinear}} = \gamma \left(\sum_{j=1}^{N} h_j \right)^k = \gamma \sum_{j_1, \ldots, j_k = 1}^{N} h_{j_1} h_{j_2} \cdots h_{j_k}. \qquad (2.12)$$

The generator of translations in γ is $K(\gamma,t) = t h_{\text{nonlinear}}$, so the quantum Cramér-Rao bound for this dynamics is

$$\delta\gamma \geq \frac{1}{tN^k(\Lambda_{\max} - \Lambda_{\min})}, \qquad (2.13)$$

where $\Lambda_{\max}$ and $\Lambda_{\min}$ are functions of Λ and λ, the largest and smallest eigenvalues, respectively, of the single-unit operators h_j. For instance, if both Λ and λ are positive, then $\Lambda_{\max} = \Lambda^k$ and $\Lambda_{\min} = \lambda^k$ for all values of k. The other possible signs of Λ and λ are discussed in [8]; they all lead to a $1/N^k$ scaling.
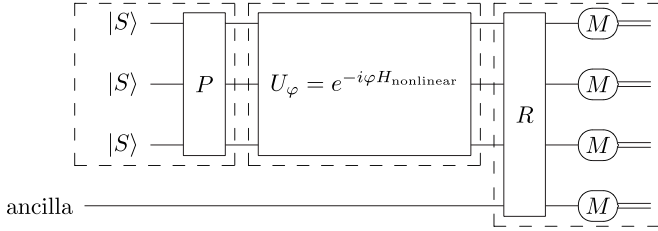
FIG. 4. Quantum circuit for a general nonlinear interferometer.

The quantum circuit for metrology with nonlinear Hamiltonians has the form shown in Fig. 4. This circuit has the same overall form as that for linear quantum metrology, with the same three stages highlighted by the dashed boxes. The only difference comes in the dynamics stage, where the gate that imprints information about the parameter on the probe involves simultaneous coupling to all the probe units.

To achieve the $1/N^k$ scaling made available by using a nonlinear Hamiltonian, the probe units have to be initialized in an entangled state that is very much like a cat state. Experimental limitations up until now have precluded making such catlike states for large numbers of systems. To avoid this difficulty, Boixo *et al.* [8] analyzed the performance of quantum-metrology protocols employing nonlinear Hamiltonians when the initial state of the probe is a product state. In this case the optimal measurement uncertainty scales as

$$\delta\gamma \sim \frac{1}{tN^{k-1/2}}. \tag{2.14}$$

The factors multiplying this scaling depend on the particular nonlinear coupling Hamiltonian [8]. It is noteworthy that the optimal $1/tN^{k-1/2}$ sensitivity can be achieved using product measurements of equatorial spin components in the effective qubit space formed from $|\Lambda\rangle$ and $|\lambda\rangle$. The key point is that for a k-body coupling Hamiltonian, the use of a product-state input costs only a factor of $\sim\sqrt{N}$ relative to the optimal sensitivity (2.13). The quantum noise limit and the Heisenberg limit of linear metrology are a special case of this $\sqrt{N}$ loss of sensitivity when using input product states as opposed to an optimal entangled state.

With two-body couplings and an initial product state for the probe, a measurement uncertainty scaling as $1/N^{3/2}$ is possible. Since two-body coupling between all probe units is not an especially onerous requirement for a probe system, the prospect of improving upon the $1/N$ Heisenberg scaling motivates us to investigate candidate systems for such metrology schemes. In the next section we consider a BEC as such a candidate system with the aim of developing a detailed, realistic, and viable proposal for an experiment that achieves better than $1/N$ scaling for the measurement uncertainty in quantum single-parameter estimation.

E. Role of entanglement

The more general point of view provided by nonlinear quantum metrology allows us to see exactly what benefit entanglement bestows on quantum metrology. Entanglement

permits one to marshal the available resources in a quantum-metrology protocol into an initial state that can achieve the best possible scaling for the measurement uncertainty as laid out by the quantum Cramér-Rao bound. The use of an appropriately entangled input state purchases a sensitivity boost by a factor of $\sim 1/\sqrt{N}$ relative to the use of an optimal initial product state. Initial entanglement is, however, not necessary for getting to or improving upon the $1/N$ Heisenberg scaling.

When specialized to qubits, with $h_j = \sigma_{z;j}/2$, and to quadratic couplings, the nonlinear probe Hamiltonian of Eq. (2.12) becomes $H_{\text{nonlinear}} = \gamma J_z^2$. This Hamiltonian generates entanglement during the dynamics stage of the protocol. Despite the evidence from the quantum Cramér-Rao bound that entanglement only helps in the initial state, one might reasonably ask whether this dynamically generated entanglement plays a role in improving upon the $1/N$ scaling. The J_z^2 probe Hamiltonian was analyzed in detail in [8]. The optimal initial product state is $[\cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle]^{\otimes n}$. Evolution under the J_z^2 Hamiltonian for a short time $t \ll \gamma^{-1}$, followed by a measurement of J_y, gives a measurement precision $\delta\gamma = 2/N^{3/2}$, which is the optimal precision for this Hamiltonian and initial state. As a consequence of the J_z^2 evolution, however, the probe qubits become entangled and suffer from an associated "phase dispersion" that makes the measurement uncertainty large for separable measurements when γt becomes large. Far from being an aid, the generated entanglement seems only to make it impossible to achieve the $1/N^{3/2}$ sensitivity using product measurements.

In [18] it was pointed out that in addition to the J_z^2 Hamiltonian, $H_{\text{nonlinear}} = \gamma N J_z$ can also be used to get an optimal measurement uncertainty $1/N^{3/2}$ when the probe qubits all start off in the state $(|0\rangle + |1\rangle)/\sqrt{2}$. The NJ_z Hamiltonian does not produce entanglement between the probe qubits, nor does it produce phase dispersion. Indeed, the NJ_z Hamiltonian acts like a linear coupling whose strength is enhanced by a factor of N. In this case it is clearly the *dynamics* alone that leads to the enhanced scaling for the measurement uncertainty, since there is no entanglement between the probe qubits at any stage in the metrology protocol.

On physical grounds, the NJ_z Hamiltonian cannot be a fundamentally linear coupling whose strength is enhanced by addition of qubits, but rather must come naturally from quadratic couplings to the parameter. Such a coupling does appear in the Hamiltonian for a two-mode BEC, as was pointed out in [18]. We introduce this BEC implementation in the next section and discuss it in some detail.

III. BOSE-EINSTEIN CONDENSATE AS A QUANTUM PROBE

A. Nonlinear BEC interferometry

The many-body Hamiltonian for a dilute Bose gas consisting of atoms of mass m in a trapping potential $V(\boldsymbol{r})$ at zero temperature, in second-quantized notation, is given by [19–22]

$$\hat{H} = \int d\boldsymbol{r} \left(\frac{\hbar^2}{2m} \nabla \hat{\psi}^\dagger \cdot \nabla \hat{\psi} + V(\boldsymbol{r}) \hat{\psi}^\dagger \hat{\psi} + \frac{1}{2} g \hat{\psi}^\dagger \hat{\psi}^\dagger \hat{\psi} \hat{\psi} \right), \tag{3.1}$$

where $\hat{\psi}^\dagger(\boldsymbol{r})$ and $\hat{\psi}(\boldsymbol{r})$ are creation and annihilation field operators that obey bosonic commutation relations,

$$[\hat{\psi}(\boldsymbol{r}), \hat{\psi}^{\dagger}(\boldsymbol{r}')] = \delta^{(3)}(\boldsymbol{r} - \boldsymbol{r}'),$$

$$[\hat{\psi}(\boldsymbol{r}), \hat{\psi}(\boldsymbol{r}')] = [\hat{\psi}^{\dagger}(\boldsymbol{r}), \hat{\psi}^{\dagger}(\boldsymbol{r}')] = 0, \quad (3.2)$$

and the coupling constant g, for a dilute gas in which the interparticle spacing is much larger than the scattering length, is related to the s-wave scattering length a by

$$g = \frac{4\pi\hbar^2 a}{m}. \quad (3.3)$$

In a zero-temperature BEC, to a very good approximation, all the atoms are in the ground state $\psi_N(\boldsymbol{r})$, which is the N-dependent ground-state solution (normalized to unity) of the time-independent Gross-Pitaevskii (GP) equation for a trapping potential $V(\boldsymbol{r})$ and a scattering term with coefficient g,

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) + g(N-1)|\psi_N|^2\right)\psi_N = \mu_N\psi_N. \quad (3.4)$$

Here μ_N is the chemical potential. At this level of approximation, the expansion of the field operator in terms of modal annihilation operators can be truncated to just one term,

$$\hat{\psi}(\boldsymbol{r}) = \psi_N(\boldsymbol{r})\hat{a}, \quad (3.5)$$

where $\hat{a}$ annihilates a particle with wave function $\psi_N(\boldsymbol{r})$. The number operator $\hat{a}^{\dagger}\hat{a}$ for this single mode can be treated as the c-number N because the number of atoms is a constant. The Hamiltonian then reduces to the c-number mean-field energy for this single mode, $H = E_0 N + \frac{1}{2}g\,\eta_N N(N-1)$, where

$$E_0 = \int d\boldsymbol{r}\left(\frac{\hbar^2}{2m}|\nabla\psi_N|^2 + V(\boldsymbol{r})|\psi_N|^2\right) \quad (3.6)$$

is the single-particle kinetic plus trapping energy, and the quantity

$$\eta_N = \int d\boldsymbol{r}\,|\psi_N(\boldsymbol{r})|^4 \quad (3.7)$$

is a measure of the inverse volume occupied by the ground-state wave function. The product $g\,\eta_N$, which has units of energy, is a scattering strength normalized by this effective volume. The average number density in the atomic cloud is $N\eta_N$.

So far we have assumed that all the atoms in the BEC are in a single atomic state, but as mentioned earlier, we want these atoms to be two-level systems, or qubits, in order for them to serve as the probe units in the quantum-metrology protocols we are interested in. We therefore consider two-mode BECs in which the atoms can occupy one of two internal states, labeled $|1\rangle$ and $|2\rangle$. These two states are typically hyperfine levels of the atoms. In practice, the atoms are cooled to form the BEC while they are all in the same internal state, and then an external field is used to drive transitions between the two levels to achieve the desired coherent superposition of atomic population between the two levels. The effect we are looking for is the difference between the integrated nonlinear phase shifts experienced by the two levels, the difference being due to the different scattering inter-

actions experienced by the two levels. This differential integrated phase shift is detected by driving a second transition between the levels, which transfers the phase information into the populations of the two levels.

For an initial analysis of this scenario in this section, we make three simplifying assumptions.

(1) The external field that drives the transitions between the two states $|1\rangle$ and $|2\rangle$ acts only for a short time compared to the phase-shift dynamics that leads to the estimate of the parameter we are interested in. We therefore treat these transitions as effectively instantaneous and do not include the driving field in the Hamiltonian.

(2) The collisions between the atoms are elastic. Thus the only allowed scattering processes are $|1\rangle|1\rangle \to |1\rangle|1\rangle$, $|2\rangle|2\rangle \to |2\rangle|2\rangle$, and $|1\rangle|2\rangle \to |1\rangle|2\rangle$, with scattering coefficients g_{11}, g_{22}, and g_{12}, where $g_{\alpha\beta} = 4\pi\hbar^2 a_{\alpha\beta}/m = g_{\beta\alpha}$, with Greek letters used to label the internal states.

These first two assumptions imply that the many-body Hamiltonian takes the form

$$\hat{H} = \sum_{\alpha}\int d\boldsymbol{r}\left(\frac{\hbar^2}{2m}\nabla\hat{\psi}_{\alpha}^{\dagger}\cdot\nabla\hat{\psi}_{\alpha} + V(\boldsymbol{r})\hat{\psi}_{\alpha}^{\dagger}\hat{\psi}_{\alpha}\right)$$

$$+ \frac{1}{2}\sum_{\alpha,\beta}g_{\alpha\beta}\int d\boldsymbol{r}\,\hat{\psi}_{\beta}^{\dagger}\hat{\psi}_{\alpha}^{\dagger}\hat{\psi}_{\alpha}\hat{\psi}_{\beta}, \quad (3.8)$$

where $\hat{\psi}_{\alpha}(\boldsymbol{r})$ is the field annihilation operator for internal state α. In writing this Hamiltonian, we assume that any energy splitting between the two internal states has been removed by going to an interation picture. Our third assumption is by far the most problematical of the three.

(3) The two modes retain the same spatial wave function $\psi_N(\boldsymbol{r})$ as they evolve. Since the atoms that form the initial BEC are all in the state $|1\rangle$, in the mean-field approximation they all share the spatial wave function $\psi_N(\boldsymbol{r})$, which is the N-dependent ground-state solution of the time-independent GP Eq. (3.4) with scattering coefficient g_{11}. Immediately after the nearly instantaneous action of the external field, the wave function for both internal states is $\psi_N(\boldsymbol{r})$. We further assume that the second internal state is chosen so that it sees the same trapping potential $V(\boldsymbol{r})$. Even though the two internal states have identical initial wave functions and experience identical trapping potentials, their wave functions will gradually become different because of the difference in their scattering lengths. What we are assuming now is that the integrated nonlinear phase shifts that we are interested in accumulate on a time scale that is shorter than the time scale for the two wave functions to differentiate spatially. Thus, for the present, we take the two wave functions to be identical. We return to the question of the time scale for differentiation of the two wave functions at the end of this section, in Sec. III F.

Using the third assumption, we can write the field annihilation operators as

$$\hat{\psi}_{\alpha}(\boldsymbol{r}) = \psi_N(\boldsymbol{r})\hat{a}_{\alpha}. \quad (3.9)$$

Since the total number of atoms is fixed, we can treat the total number operator,

$$\hat{N} \equiv \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2, \tag{3.10}$$

as a c-number N. We can then put the two-mode Hamiltonian in the form

$$\hat{H} = E_0 N + \frac{1}{2} \eta_N \sum_{\alpha,\beta} g_{\alpha\beta} \hat{a}_\beta^\dagger \hat{a}_\alpha^\dagger \hat{a}_\alpha \hat{a}_\beta$$

$$= H_0 + \gamma_1 \eta_N (N-1) \hat{J}_z + \gamma_2 \eta_N \hat{J}_z^2, \tag{3.11}$$

where E_0 and η_N are as in Eqs. (3.6) and (3.7). The operator $\hat{J}_z$ is defined by

$$\hat{J}_z \equiv \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2), \tag{3.12}$$

and we have also introduced a c-number energy,

$$H_0 = E_0 N + \frac{1}{4}\left(\frac{1}{2}(g_{11}+g_{22}) + g_{12}\right)\eta_N N^2 - \frac{1}{4}(g_{11}+g_{22})\eta_N N, \tag{3.13}$$

which includes the common-mode part of the mean-field scattering energy. Finally, we define two coupling constants that characterize the interaction of the two modes,

$$\gamma_1 \equiv \frac{1}{2}(g_{11}-g_{22}) \text{ and } \gamma_2 \equiv \frac{1}{2}(g_{11}+g_{22}) - g_{12}. \tag{3.14}$$

The Hamiltonian (3.11) is often called the Josephson approximation.

The common-mode energy H_0 in Eq. (3.11) can be ignored because its only effect is to introduce an overall phase in the evolved state of the probe. In the other two terms, we have $(N-1)\hat{J}_z$ and $\hat{J}_z^2$ couplings, suggesting that we might be able to measure the coupling constants γ_1 and γ_2 with an accuracy that scales as $1/N^{3/2}$ with the number of atoms in the BEC.

To see how this works out, suppose the first optical pulse puts each atom in a superposition $c_1|1\rangle + c_2|2\rangle$, where c_1 and c_2 can be assumed to be real (i.e., the first optical pulse performs a rotation about the y axis of the Bloch sphere). For short times, we can make a linear approximation to $\hat{J}_z^2$ in the Josephson Hamiltonian, i.e., we can set $\hat{J}_z^2 = (\langle\hat{J}_z\rangle + \Delta\hat{J}_z)^2 \simeq \langle\hat{J}_z\rangle^2 + 2\langle\hat{J}_z\rangle\Delta\hat{J}_z$, with $\langle\hat{J}_z\rangle = N(c_1^2 - c_2^2)/2$. The linear approximation amounts to neglecting the phase dispersion and corresponding entanglement produced by the $\hat{J}_z^2$ term. We need not make any such short-time approximation for the $(N-1)\hat{J}_z$ term. Up to irrelevant phases, the resulting evolution is a rotation of each atom's state about the z axis of the Bloch sphere with angular velocity

$$\frac{\eta_N}{\hbar}[(N-1)\gamma_1 + N(c_1^2 - c_2^2)\gamma_2] \simeq \frac{(N-1)\eta_N}{\hbar}[\gamma_1 + (c_1^2 - c_2^2)\gamma_2]$$

$$\equiv \Omega_N, \tag{3.15}$$

where in the second form, we approximate N as N−1. Under these circumstances, the BEC acts like a linear Ramsey interferometer whose rotation rate is enhanced by a factor of

$(N-1)\eta_N$, leading to a sensitivity that scales as $1/\sqrt{N}(N-1)\eta_N \simeq 1/N^{3/2}\eta_N$. If $\gamma_2 = 0$, the optimal initial state has $c_1 = c_2 = 1/\sqrt{2}$, but if $\gamma_1 = 0$, the optimal choice is $c_1 = \cos(\pi/8)$ and $c_2 = \sin(\pi/8)$ [8].

Achieving a $1/N^{3/2}$ scaling requires that η_N have no dependence on N. As noted above, however, η_N^{-1} is a measure of the volume occupied by the ground-state wave function ψ_N. As atoms are added to a BEC, the wave function spreads because of the repulsive scattering of the atoms, thereby reducing η_N as N increases. To pin down how the measurement accuracy scales with N, we need to determine how η_N behaves as a function of N.

B. Two critical atom numbers

Since we first create a BEC of N atoms all in hyperfine state $|1\rangle$, before putting them in a superposition of states $|1\rangle$ and $|2\rangle$, we can focus on the N dependence of η_N for a single-mode BEC of atoms in state $|1\rangle$. Thus, in this subsection and the next two, we deal with the single-mode GP Eq. (3.4) with $g = g_{11}$ and $a = a_{11}$.

An obvious strategy to suppress the N dependence of η_N is to constrain the BEC within a hard-walled trap so that it cannot expand as more atoms are added. BECs effectively confined to two or one dimensions and held in power-law trapping potentials along these dimensions are the sort found in real experiments. Thus we look at the dependence of η_N on N for a BEC that is *loosely* trapped in d dimensions, referred to as *longitudinal* (L) dimensions, and *tightly* trapped in $D = 3-d$ dimensions, referred to as *transverse* (T) dimensions. We assume that in the longitudinal dimensions, the atoms are trapped in a power-law potential of the form

$$V_L(\mathbf{r}) = \frac{1}{2}kr^q, \quad q = 1, 2, \dots, \tag{3.16}$$

and that in the transverse dimensions, the trapping potential is harmonic,

$$V_T(\boldsymbol{\rho}) = \frac{1}{2}m\omega_T^2\rho^2. \tag{3.17}$$

The parameter q characterizes the hardness of the longitudinal trapping potential. We deal with a three-dimensional (3D) trap by setting $D = 0$, meaning there are no transverse dimensions.

When N is small, the mean-field scattering energy is negligible compared to the atomic kinetic energy of the atoms and the trapping potential energy. In this situation, the scattering term in the GP equation can be neglected, and the ground-state wave function is the solution of the Schrödinger equation for the trapping potential $V_L(\mathbf{r}) + V_T(\boldsymbol{\rho})$. As more atoms are added to the BEC, the repulsive scattering term in Eq. (3.4) comes into play and causes the wave function to spread. We define two critical atom numbers, N_L and N_T, which characterize the onset of spreading in the longitudinal and transverse dimensions. The lower critical atom number, N_L, is defined as the atom number at which the scattering term in the GP equation is as large as the longitudinal kinetic-energy term and thus characterizes when the wave

function begins to spread in the longitudinal dimensions. The upper critical atom number, N_T, is defined as the atom number at which the scattering term is as large as the transverse kinetic energy and thus characterizes when the wave function begins also to spread in the transverse dimensions. The notion of an upper critical atom number only makes sense for one-dimensional (1D) and two-dimensional (2D) traps and not for $d=3$.

For small atom number, i.e., $N \ll N_L$, as just noted, the scattering term in the GP equation can be neglected, and the ground-state solution of the GP equation is the N-independent product ground state of the Schrödinger equation,

$$\psi_0(\boldsymbol{\rho}, \boldsymbol{r}) = \chi_0(\boldsymbol{\rho}) \phi_0(\boldsymbol{r}). \qquad (3.18)$$

Here $\chi_0(\boldsymbol{\rho})$ is the Gaussian ground state for the transverse dimensions,

$$\chi_0(\boldsymbol{\rho}) = \frac{1}{(2\pi\rho_0^2)^{D/4}} \exp\left(-\frac{\rho^2}{4\rho_0^2}\right), \qquad (3.19)$$

whose corresponding probability density has half-width

$$\rho_0 \equiv \left(\frac{\hbar}{2m\omega_T}\right)^{1/2}, \qquad (3.20)$$

and $\phi_0(\boldsymbol{r})$ is the bare ground state for the loosely confined longitudinal dimensions. We can estimate the half-width of ϕ_0 by equating the trapping potential energy (PE) and the kinetic energy (KE) per dimension, i.e., $kr_0^q/2 = \hbar^2/2mr_0^2$, which gives

$$r_0 \equiv \left(\frac{\hbar^2}{mk}\right)^{1/(q+2)}. \qquad (3.21)$$

In accordance with our assumptions, we assume that r_0 is much larger than ρ_0. A hard-walled trap in the longitudinal dimensions corresponds to the limit $q \to \infty$ with r_0 held constant.

The trapped BECs we consider are thus characterized by three length scales: (i) the scattering length a; (ii) the bare transverse trap half-width ρ_0; and (iii) the bare longitudinal trap half-width r_0. Typical values, which we use for estimates in the following, are $a = 10$ nm, $\rho_0 = 1$ μm, and $r_0 = 100$ μm. For ^{87}Rb atoms (which have $a = a_{11} = 5.3$ nm), the corresponding transverse trap frequency is $\nu_T = 58$ Hz; we can also identify an approximate longitudinal trap frequency,

$$\nu_L = \frac{\omega_L}{2\pi} \equiv \frac{1}{2\pi} \frac{\hbar}{mr_0^2} \simeq 10^{-2} \text{ Hz}, \qquad (3.22)$$

associated with the bare longitudinal ground state.

Whenever the wave function is a product of transverse and longitudinal wave functions, η_N is also a product, $\eta_N = \eta_T \eta_L$. When $N \ll N_L$, $\eta_N \equiv \eta_0$ is independent of N since

$$\eta_T = \int d^D\rho |\chi_0(\boldsymbol{\rho})|^4 = \frac{1}{(4\pi)^{D/2} \rho_0^D}, \qquad (3.23)$$

$$\eta_L = \int d^d r |\phi_0(\boldsymbol{r})|^4 \simeq \frac{1}{V_d r_0^d}, \qquad (3.24)$$

where V_d is the volume of a unit sphere in d dimensions ($V_1 = 2$, $V_2 = \pi$, and $V_3 = 4\pi/3$), The lower critical atom number, N_L, is defined by setting

$$\frac{\hbar^2}{2mr_0^2} \simeq (\text{longitudinal KE}) \simeq (\text{scattering term})$$

$$\simeq (N_L - 1) g \eta_0 \simeq \frac{\hbar^2}{2m} (N_L - 1) \frac{1}{\beta_d} \frac{a}{\rho_0^D r_0^d}, \qquad (3.25)$$

where

$$\beta_d \equiv \frac{V_d}{2(4\pi)^{(d-1)/2}} \qquad (3.26)$$

is a geometric factor ($\beta_1 = 1$, $\beta_2 = \sqrt{\pi}/4$, $\beta_3 = 1/6$). Thus we define

$$N_L - 1 \equiv \beta_d \frac{r_0}{a} \left(\frac{\rho_0}{r_0}\right)^D. \qquad (3.27)$$

For the typical length scales mentioned above, the lower critical atom number is about 1700 for a 3D trap, 45 for a 2D trap, and 2 for a 1D trap. The small value of N_L for a 1D trap is the reason we retain the -1 wherever it appears in our discussion of atom numbers, even though it could be dropped in most situations.

For $N_L \lesssim N \ll N_T$, the tight confinement in the transverse dimensions means that the wave function continues to be a product,

$$\psi_N(\boldsymbol{\rho}, \boldsymbol{r}) = \chi_0(\boldsymbol{\rho}) \phi_N(\boldsymbol{r}), \qquad (3.28)$$

but with the longitudinal wave function satisfying a GP equation,

$$\left(-\frac{\hbar^2}{2m} \nabla_L^2 + V_L(\boldsymbol{r}) + g(N-1)\eta_T |\phi_N|^2\right) \phi_N = \mu_L \phi_N, \qquad (3.29)$$

where $\mu_L = \mu_N - D\hbar\omega_T/2$ is the longitudinal part of the chemical potential. As atoms are added to the trap in this intermediate regime, the wave function spreads in the longitudinal dimensions. We can estimate the longitudinal half-width r_N by noticing that $\eta_N = \eta_T \eta_L$, where η_T is given by Eq. (3.23) and

$$\eta_L = \int d\boldsymbol{r} |\phi_N(\boldsymbol{r})|^4 \simeq \frac{1}{V_d r_N^d}, \qquad (3.30)$$

and then equating the attractive longitudinal trapping PE to the repulsive scattering term,

$$\frac{1}{2}kr_N^q \simeq (\text{longitudinal PE}) \simeq (\text{scattering term})$$

$$\simeq (N-1)g\,\eta_N \simeq \frac{\hbar^2}{2m}(N-1)\frac{1}{\beta_d}\frac{1}{\rho_0^D r_N^d}a$$

$$= \frac{\hbar^2}{2m}\frac{N-1}{N_L-1}\frac{r_0^{d-2}}{\rho_N^d}, \tag{3.31}$$

where we have used Eq. (3.27) in the last step. This leads us to define

$$\frac{r_N}{r_0} \equiv \left(\frac{N-1}{N_L-1}\right)^{1/(d+q)}. \tag{3.32}$$

We now define the upper critical atom number by setting

$$\frac{\hbar^2}{2m\rho_0^2} \simeq (\text{transverse KE}) \simeq (\text{scattering term})$$

$$\simeq (N_T-1)g\,\eta_{N_T} = \frac{\hbar^2}{2m}\frac{N_T-1}{N_L-1}\frac{r_0^{d-2}}{r_T^d}, \tag{3.33}$$

where r_T is the longitudinal half-width at the upper critical atom number,

$$\frac{r_T}{r_0} \equiv \left(\frac{N_T-1}{N_L-1}\right)^{1/(d+q)}. \tag{3.34}$$

Using Eq. (3.33) and the definitions in Eqs. (3.34) and (3.27) we end up with the definition

$$N_T-1 \equiv (N_L-1)\left(\frac{r_0}{\rho_0}\right)^{2(d+q)/q} = \beta_d\frac{\rho_0}{a}\left(\frac{r_0}{\rho_0}\right)^{d(q+2)/q}. \tag{3.35}$$

We stress that the notion of an upper critical atom number only makes sense for 1D and 2D traps and not for $d=3$. Using the typical values mentioned above, we have that the upper critical atom number for a harmonic longitudinal trap ($q=2$) is about 4×10^9 for a 2D trap and about 10^6 for a 1D trap; for a hard longitudinal trap ($q\to\infty$), N_T is about 4×10^5 for a 2D trap and about 10^4 for a 1D trap. Using Eq. (3.35) we can rewrite the longitudinal radius in Eq. (3.32) as

$$\frac{r_N}{r_0} = \left(\frac{r_0}{\rho_0}\right)^{2/q}\left(\frac{N-1}{N_T-1}\right)^{1/(d+q)}. \tag{3.36}$$

It should be noted that

$$\frac{r_T}{\rho_0} = \left(\frac{a}{\rho_0}\frac{N_T-1}{\beta_d}\right)^{1/d}. \tag{3.37}$$

For a 1D trap, this gives $r_T=a(N_T-1)$, making the relation between r_T and N_T independent of the parameters of the trap. Another way of thinking about Eq. (3.37) is that the number density at the upper critical atom number,

$$\frac{N_T}{\beta_d\rho_0^D r_T^d} \simeq \frac{1}{a\rho_0^2}, \tag{3.38}$$

is independent of the properties of the longitudinal trap, with typical value 10^{14} cm^{-3}.

As the atom number increases from N_T, the transverse kinetic energy becomes unimportant compared to the transverse trapping energy and the scattering term. The wave function continues to spread in the longitudinal dimensions and also spreads in the transverse dimensions, with the longitudinal and transverse radii, r_N and ρ_N, given by

$$\frac{1}{2}kr_N^q \simeq \frac{1}{2}m\omega_T^2\rho_N^2 \simeq (\text{scattering term})$$

$$\simeq (N-1)g\,\eta_N \simeq (N-1)g\frac{1}{V_D\rho_N^D V_d r_N^d}, \tag{3.39}$$

which leads us to define in the regime $N\gg N_T$,

$$\frac{r_N}{r_0} \equiv \left(\frac{r_0}{2\rho_0}\frac{\rho_N}{\rho_0}\right)^{2/q}, \tag{3.40}$$

$$\left(\frac{\rho_N}{\rho_0}\right)^{5-d+2d/q} \equiv \frac{4(4\pi)^{D/2}2^{2d/q}}{V_D}\frac{N-1}{N_T-1}. \tag{3.41}$$

C. Renormalization of the nonlinear interaction terms and the sensitivity scaling

The estimates in the previous subsection tell us how η_N scales with atom number. For atom numbers smaller than the lower critical atom number, η_N has the constant value η_0, a consequence of the fact that the repulsive scattering has negligible effect on the atomic wave function. In the intermediate regime of atom numbers, i.e., for atom numbers between N_L and N_T, the wave function expands in the longitudinal dimensions, making η_N scale as

$$\eta_N \sim \frac{1}{r_N^d} \sim \left(\frac{N_L-1}{N-1}\right)^{d/(d+q)}. \tag{3.42}$$

For atom numbers above the upper critical atom number, as the wave function spreads in all dimensions, η_N scales as

$$\eta_N \sim \frac{1}{\rho_N^D r_N^d} \sim \frac{1}{\rho_N^{3-d+2d/q}} \sim \left(\frac{N_T-1}{N-1}\right)^{(3-d+2d/q)/(5-d+2d/q)}. \tag{3.43}$$

In the measurement schemes we contemplate, the uncertainties in determining γ_1 and γ_2 scale as

$$\delta\gamma_{1,2} \sim \frac{1}{\sqrt{N(N-1)\,\eta_N}} \sim \frac{1}{N^\xi}, \tag{3.44}$$

where in the final form we neglect 1 compared to N. For atom numbers below N_L, the scaling exponent ξ is 3/2; for $N_L\ll N\ll N_T$, it takes on the value

$$\xi = \frac{3}{2} - \frac{d}{d+q} = \frac{d+3q}{2(d+q)}, \tag{3.45}$$

and for $N\gg N_T$, ξ is given by

$$\xi = \frac{3}{2} - \frac{3-d+2d/q}{5-d+2d/q}. \tag{3.46}$$

For atom numbers above N_T, harmonic 1D and 2D traps have $\xi=9/10$, a hard-walled 1D trap has $\xi=1$, and a hard-walled
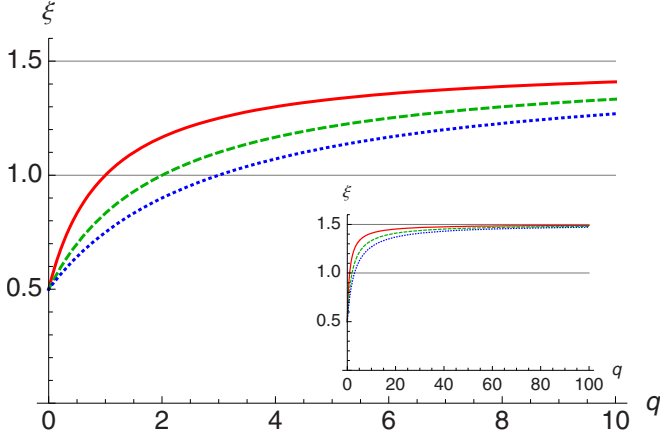
FIG. 5. (Color online) Sensitivity scaling exponent $\xi = (d+3q)/2(d+q)$ in the intermediate regime of atom numbers, $N_L \ll N \ll N_T$, plotted as a function of hardness parameter q of the longitudinal trapping potential for 1D (red, solid), 2D (green, dashed), and 3D (blue, dotted) traps in the intermediate regime of atom numbers (for 3D traps, there is no upper critical atom number N_T). To achieve super-$1/N$ scaling ($\xi > 1$) requires $q > d$. A harmonic 1D trap has super-$1/N$ scaling $\xi = 7/6$, a harmonic 2D trap has Heisenberg scaling $\xi = 1$, and a 3D harmonic trap has sub-$1/N$ scaling $\xi = 9/10$. This sub-$1/N$ scaling for 3D harmonic traps is still markedly better than the QNL scaling of $\xi = 1/2$. A hard-walled trap ($q \to \infty$) in any dimension has $\xi = 3/2$.

2D trap has $\xi = 7/6$. Our main interest is the intermediate regime of Eq. (3.45). The scaling exponent in this regime is plotted in Fig. 5 as a function of q for 1D, 2D, and 3D traps.

D. Thomas-Fermi approximations

Although we have determined how the scaling exponent behaves with d and q, we can do a better job of evaluating η_N, determining more precisely the constants in front of the scaling, by using the Thomas-Fermi (TF) approximation. In the intermediate regime of atom numbers, the wave function is the product (3.28), with the longitudinal wave function $\phi_N(\mathbf{r})$ satisfying the GP Eq. (3.29) in d dimensions. When N is much larger than N_L, we can ignore the kinetic-energy term in the reduced GP equation, which gives the TF probability distribution,

$$|\phi_N(\mathbf{r})|^2 = \frac{\mu_L - kr^q/2}{(N-1)g\,\eta_T}. \tag{3.47}$$

Since $|\phi_N(\mathbf{r})|^2$ must be positive, the radial extent of the BEC in the longitudinal dimensions is bounded by $\tilde{r}_N$ such that

$$\mu_L = \frac{1}{2}k\tilde{r}_N^q. \tag{3.48}$$

Normalization yields

$$1 = \int d^d r |\phi_N(\mathbf{r})|^2 \equiv I_1(N,d,q). \tag{3.49}$$

Integrals over TF probability distributions, such as I_1, are defined and evaluated in the Appendix. Using Eq. (A.7), we find

$$\frac{\tilde{r}_N}{r_0} = \left(\frac{d+q}{q}\frac{N-1}{N_L-1}\right)^{1/(d+q)} \tag{3.50}$$

[cf. Eq. (3.32)] or, equivalently, using Eqs. (A.1) and (A.4),

$$\begin{aligned}
\frac{\mu_L}{(N-1)g\,\eta_T} &= \frac{d+q}{q}\frac{1}{V_d\tilde{r}_N^d} \\
&= \frac{1}{V_d r_0^d}\left(\frac{d+q}{q}\right)^{q/(d+q)}\left(\frac{N_L-1}{N-1}\right)^{d/(d+q)}.
\end{aligned} \tag{3.51}$$

Now, from Eq. (A.8), we can find

$$\begin{aligned}
\eta_L = I_2(N,d,q) &= \frac{2q}{d+2q}\frac{\mu_L}{(N-1)g\,\eta_T} \\
&= \frac{1}{V_d r_0^d}\frac{2q}{d+2q}\left(\frac{d+q}{q}\right)^{q/(d+q)}\left(\frac{N_L-1}{N-1}\right)^{d/(d+q)}
\end{aligned} \tag{3.52}$$

and thus determine $\eta_N = \eta_T \eta_L$ [cf. Eq. (3.42)]. Numerical computation of η_N in the intermediate regime indicates that this expression is quite accurate in spite of the approximations that went into obtaining it.

When N is much larger than N_T, we can again use a TF approximation, this time ignoring the kinetic-energy term in the 3D GP Eq. (3.4), which gives the probability distribution

$$|\psi_N(\boldsymbol{\rho},\mathbf{r})|^2 = \frac{\mu_N - m\omega_T^2\rho^2/2 - kr^q/2}{(N-1)g}. \tag{3.53}$$

Positivity of this distribution requires that $\rho \leq \tilde{\rho}_N$ and $r \leq \tilde{r}_N(\rho)$, where

$$\frac{1}{2}m\omega_T^2\tilde{\rho}_N^2 = \mu_N, \tag{3.54}$$

$$\frac{1}{2}k\tilde{r}_N^q(\rho) = \frac{1}{2}m\omega_T^2(\tilde{\rho}_N^2 - \rho^2). \tag{3.55}$$

The extent of the atomic cloud in the longitudinal direction is characterized by $\tilde{r}_N \equiv \tilde{r}_N(0)$, i.e., $\tilde{r}_N/r_0 = [(r_0/2\rho_0)(\tilde{\rho}_N/\rho_0)]^{2/q}$. Normalization yields

$$1 = \int d^D\rho\, d^d r |\psi_N(\boldsymbol{\rho},\mathbf{r})|^2 \equiv K_1(N,d,q). \tag{3.56}$$

Using Eq. (A.11), we find

$$\left(\frac{\tilde{\rho}_N}{\rho_0}\right)^{5-d+2d/q} = \frac{1}{dJ_{1+d/q}(D,2)J_1(d,q)}\frac{4(4\pi)^{D/2}2^{2d/q}}{S_{D-1}}\frac{N-1}{N_T-1} \tag{3.57}$$

[cf. Eq. (3.41)]. We can also write

$$\frac{\mu_N}{(N-1)g} = \frac{1}{4(4\pi)^{D/2}\rho_0^D V_d r_0^d}\left(\frac{\rho_0}{r_0}\right)^{2q/d}\left(\frac{\tilde{\rho}_N}{\rho_0}\right)^2\frac{N_T-1}{N-1}$$

(3.58)

and

$$\eta_N \equiv K_2(N,d,q) = \frac{2(d+q)}{(D/2+d/q+1)(d+2q)}\frac{\mu_N}{(N-1)g}.$$

(3.59)

The scaling of η_N agrees with that in Eq. (3.43).

E. Bose-condensed ^{87}Rb atoms

A good candidate for implementing the generalized metrology protocol is a BEC made of rubidium (^{87}Rb) atoms [23–26]. Initially the atoms in the condensate are in the ground electronic state. In most experiments [24–26], the $|F=1;M_F=-1\rangle \equiv |1\rangle$ state is trapped and cooled to the condensation point. Once the atoms in $|1\rangle$ have accumulated in the condensate ground state, a two-photon drive is used to couple the $|1\rangle$ state to the $|F=2;M_F=+1\rangle \equiv |2\rangle$ state. The two-photon drive involves applying both microwave and radio-frequency electromagnetic fields to the condensate. Both hyperfine states are not cooled simultaneously to form a condensate in a superposition of the two states because the lifetime of atoms in the $|2\rangle$ state in a trap is much shorter than the lifetime of atoms in the $|1\rangle$ state. As mentioned earlier, we assume that the driving field that initializes the atoms in a desired superposition of the $|1\rangle$ and $|2\rangle$ states is instantaneous in comparison to the dynamics that is part of the parameter-estimation process.

The s-wave scattering lengths for the three processes, $|1\rangle|1\rangle \rightarrow |1\rangle|1\rangle$, $|2\rangle|2\rangle \rightarrow |2\rangle|2\rangle$, and $|1\rangle|2\rangle \rightarrow |1\rangle|2\rangle$ are nearly degenerate for ^{87}Rb, with the ratios $\{a_{22}:a_{12}:a_{11}\}=\{0.97:1:1.03\}$ ($a_{11}=5.31$ nm) [23]. These values for the scattering lengths mean that $\gamma_2=(g_{11}+g_{22})/2-g_{12}=0$ for ^{87}Rb. Therefore, the ^{87}Rb BEC realizes the generalized quantum-metrology protocol with just the $\gamma_1(N-1)\eta_N\hat{J}_z$ coupling in Hamiltonian (3.11).

The optimal initial state for metrology with the $\gamma_1(N-1)\eta_N\hat{J}_z$ coupling is the one in which all atoms are initialized in the equatorial plane of the Bloch sphere, say, in the $+1$ eigenstate of σ_x, $(|1\rangle+|2\rangle)/\sqrt{2}$. The quantity estimated by this measurement scheme is $\gamma_1=(g_{11}-g_{22})/2$, which is small, but finite in the case of a ^{87}Rb BEC. Once the probe is initialized in the optimal initial state, we let it evolve for a time t under the influence of the $\gamma_1(N-1)\eta_N\hat{J}_z$ Hamiltonian, which simply rotates the state of each atom about the z axis of the Bloch sphere with angular velocity $\gamma_1(N-1)\eta_N/\hbar$. At the end of this evolution, we measure an equatorial component of $\hat{J}$ ($\hat{J}_y$ for short evolution times), which is achieved by a $\pi/2$ pulse about the desired equatorial axis, followed by a measurement of $\hat{J}_z$, i.e., of the difference in the populations of the two internal states.

Precision experiments with two-component ^{87}Rb BECs in modestly nonspherical harmonic traps have been reported in [27,28]. These experiments were carried out with atom num-bers in excess of 100 000 and thus operated in the full TF regime well above the upper critical atom number.

F. Differentiation of the spatial wave functions

The strongest assumption we made in obtaining Hamiltonian (3.11) was that the wave functions for the two modes remain identical throughout the duration of the proposed measurement scheme. Here we examine this assumption more carefully.

We noted earlier that even if the two modes see the same trapping potential, the difference in their scattering lengths will cause the two wave functions to evolve differently [24]. The initial effect of the difference in scattering lengths is to produce a relative phase between $|1\rangle$ and $|2\rangle$. This relative phase depends on the local density within the condensate. The integrated (or average) part of the relative phase provides the signal for our measurement protocol, whereas the residual position-dependent part of the relative phase reduces the visibility of the fringes on which the signal relies. For our protocol to succeed, we need the integrated phase to accumulate more rapidly than the residual position-dependent phase. Yet a further effect is that the position-dependent phases drive differences between the atomic densities associated with the two hyperfine levels, but as this occurs on a longer time scale than the accumulation of the position-dependent phase shift, we do not consider it here.

We can analyze this scenario in the following way. Initially all atoms are in the state $\psi_N(\boldsymbol{\rho},\boldsymbol{r})|1\rangle$. After the first optical pulse, the state becomes $\psi_N(\boldsymbol{\rho},\boldsymbol{r})(c_1|1\rangle+c_2|2\rangle)$, where c_1 and c_2 are the amplitudes to be in the hyperfine states. We can assume that c_1 and c_2 are real, i.e., that the initial optical pulse produces a rotation about the y axis of the Bloch sphere. The different scattering lengths make the wave functions for the two modes evolve differently, so that after a time t, the atomic state becomes $c_1\psi_{N,1}(\boldsymbol{\rho},\boldsymbol{r},t)|1\rangle+c_2\psi_{N,2}(\boldsymbol{\rho},\boldsymbol{r},t)|2\rangle$, where the wave functions for the two modes evolve according to time-dependent coupled GP equations

$$i\hbar\frac{\partial\psi_{N,\alpha}}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V + (N-1)\sum_\beta g_{\alpha\beta}c_\beta^2|\psi_{N,\beta}|^2\right)\psi_{N,\alpha}.$$

(3.60)

The second optical pulse is a $\pi/2$ pulse about an equatorial axis of the Bloch sphere. For the discussion here, we assume that this rotation is about the x axis so that subsequent counting of the populations of the two hyperfine levels is equivalent to measuring $\hat{J}_y$ before the second optical pulse. The state after the pulse is

$$c_1\psi_{N,1}\frac{1}{\sqrt{2}}(|1\rangle-i|2\rangle) + c_2\psi_{N,2}\frac{1}{\sqrt{2}}(-i|1\rangle+|2\rangle)$$

$$= \frac{1}{\sqrt{2}}(c_1\psi_{N,1}-ic_2\psi_{N,2})|1\rangle - \frac{i}{\sqrt{2}}(c_1\psi_{N,1}+ic_2\psi_{N,2})|2\rangle.$$

(3.61)

The corresponding probabilities to be in the two states,

$$p_{1,2} = \frac{1}{2}[1 \mp 2c_1 c_2 \, \text{Im}(\langle \psi_{N,2} | \psi_{N,1} \rangle)], \qquad (3.62)$$

are determined by the overlap of the two spatial wave functions,

$$\langle \psi_{N,2} | \psi_{N,1} \rangle = \int d^D \rho \, d^d r \, \psi_{N,2}^* \psi_{N,1}. \qquad (3.63)$$

In a ground-breaking set of experiments, Anderson *et al.* [28] measured the position-dependent phase shifts in a two-component ^{87}Rb BEC, trapped in a modestly nonspherical trap, and saw the associated reduction in fringe visibility. The details of the experiment were shown to be well accounted for by numerical integrations of the two-component GP Eq. (3.60) with a loss term included. The experiment was carried out with atom number $N \simeq 1.5 \times 10^5$, well above the upper critical atom number.

To compare the time scales for the integrated and position-dependent phase shifts in our protocol, we assume that we are operating in the intermediate regime of atom numbers, i.e., $N_L \lesssim N \ll N_T$. In this regime, the wave functions for the two modes factor into transverse and longitudinal wave functions, i.e.,

$$\psi_{N,\alpha}(\boldsymbol{\rho}, \boldsymbol{r}, t) = \chi_0(\boldsymbol{\rho}) \phi_{N,\alpha}(\boldsymbol{r}, t), \quad \alpha = 1, 2, \qquad (3.64)$$

where χ_0 is the time-independent Gaussian ground state in the transverse dimensions and the longitudinal wave functions obey time-dependent coupled longitudinal GP equations,

$$i\hbar \frac{\partial \phi_{N,\alpha}}{\partial t} = \left(-\frac{\hbar^2}{2m} \nabla_L^2 + V_L + \eta_T (N-1) \sum_\beta g_{\alpha\beta} c_\beta^2 |\phi_{N,\beta}|^2 \right) \phi_{N,\alpha}. \qquad (3.65)$$

To estimate the time scales, we assume that N is large enough relative to the lower critical atom number to justify the TF approximation in the longitudinal dimensions, thus allowing us to ignore the kinetic-energy terms in the coupled GP equations. With these assumptions, the probability densities do not change with time, i.e.,

$$|\phi_{N,\alpha}(\boldsymbol{r}, t)|^2 = |\phi_N(\boldsymbol{r}, 0)|^2 \equiv q_0(\boldsymbol{r}), \qquad (3.66)$$

and the evolution under the coupled GP equations only introduces a phase,

$$\phi_{N,\alpha}(\boldsymbol{r}, t) = \sqrt{q_0} \, \exp\left[-\frac{it}{\hbar} \left(V_L + \eta_T (N-1) q_0 \sum_\beta g_{\alpha\beta} c_\beta^2 \right) \right]. \qquad (3.67)$$

This gives an overlap $\langle \psi_{N,2} | \psi_{N,1} \rangle = \langle \phi_{N,2} | \phi_{N,1} \rangle = \int d^d r \, q_0 e^{-i\delta\theta(\boldsymbol{r})}$, where the relative phase is given by

$$\delta\theta(\boldsymbol{r}) = \frac{\eta_T (N-1) q_0(\boldsymbol{r}) \Delta g t}{\hbar} = \Omega_N t \left(1 + \frac{q_0(\boldsymbol{r}) - \eta_L}{\eta_L} \right), \qquad (3.68)$$

with

$$\Delta g \equiv c_1^2 (g_{11} - g_{12}) - c_2^2 (g_{22} - g_{12}) = \gamma_1 + (c_1^2 - c_2^2) \gamma_2. \qquad (3.69)$$

In the second equality of Eq. (3.68), we have separated out the integrated phase shift, which has angular frequency

$$\Omega_N \equiv \frac{(N-1) \eta_N \Delta g}{\hbar} = \omega_L \frac{\Delta g}{g_{11}} \frac{q}{d+2q} \left(\frac{q}{d+q} \frac{N-1}{N_L - 1} \right)^{q/(d+q)} \qquad (3.70)$$

[cf. Eqs. (3.15) and (3.22)], leaving the residual position-dependent phase shift as a correction. The final expression for Ω_N uses the TF approximation to evaluate η_N in the intermediate regime. For the ^{87}Rb protocol outlined in Sec. III E, in which γ_2 is essentially zero, we choose $c_1^2 = c_2^2 = 1/2$ in order to maximize the fringe visibility in Eq. (3.62).

It is worth emphasizing how this approach based on coupled GP equations differs from use of Josephson Hamiltonian (3.11). Although the GP equations yield a position-dependent phase, which cannot be obtained from the Josephson Hamiltonian, this comes at a price: the integrated relative phase in Eq. (3.68) amounts to making the linear approximation to $\hat{J}_z^2$ described in the paragraph containing Eq. (3.15). The linear approximation is essential because the $\hat{J}_z^2$ coupling does not preserve product states, whereas the GP equations assume a product state. It means that the GP equations miss the phase dispersion generated by the $\hat{J}_z^2$ coupling and the associated dynamically generated entanglement.

We can now write the overlap as

$$\langle \psi_{N,2} | \psi_{N,1} \rangle = e^{-i\Omega_N t} \int d^d r \, q_0 e^{-i\Omega_N t (q_0 - \eta_L)/\eta_L}$$

$$\simeq e^{-i\Omega_N t} \exp\left(-\frac{\Omega_N^2 t^2}{2\eta_L^2} \int d^d r \, q_0 (q_0 - \eta_L)^2 \right), \qquad (3.71)$$

where the second expression comes from expanding the exponential inside the integral to second order and then converting to an equivalent Gaussian at the same order. The contribution from the first-order term vanishes since $\eta_L = \int d^d r \, q_0^2$. We can identify a time scale τ_{pd} for the position-dependent phase as the time set by the half-width of the Gaussian, i.e.,

$$\Omega_N \tau_{\text{pd}} \equiv \eta_L \left(\int d^d r \, q_0 (q_0 - \eta_L)^2 \right)^{-1/2} = \sqrt{\frac{2(d+3q)}{d}}. \qquad (3.72)$$

The final form comes from using TF approximation (3.47) for the density q_0 and the results in the Appendix to evaluate the integral.

What this result means is that to retain good fringe visibility, our protocol will generally be restricted to operating well within the first fringe. One can expect, however, that as the longitudinal trap becomes more hard walled, the TF density becomes more and more flat topped, eventually approaching a box, with the result that the residual position-

dependent phase shift becomes smaller and smaller. This expectation is borne out by Eq. (3.72), which reports that τ_{pd} gets larger as the hardness parameter q increases; e.g., for a 1D trap with $q=10$, $\Omega_N \tau_{\mathrm{pd}} \simeq 8$.

To investigate further this way of reducing the effect of the position-dependent phase requires numerical simulations and more accurate approximation procedures, both of which we have undertaken. Initial results, to be reported elsewhere, suggest that things turn out better than is suggested by the crude approximations that go into Eq. (3.72).

G. Other practical considerations

1. Loss of atoms

For a one-component BEC, the total number of atoms for a given trap is limited by three-body losses. This process is usually the most significant loss channel, with all other losses being negligible. For a two-component BEC, however, things are different because other loss channels, such as inelastic two-atom (spin-exchange) collisions, become significant even when the number of atoms in the trap is such that three-body collisions are unimportant. Just as in the case of three-body collisions, the spin-exchange collisions can be considered as a process that leads to loss of atoms from the trap.

Spin-exchange collisions in a two-component BEC of ^{87}Rb atoms in the two hyperfine levels we are interested in were considered in [27,28]. The effect of inelastic spin-exchange interactions was modeled by including non-Hermitian potentials in the coupled GP Eqs. (3.60),

$$-\frac{i\hbar}{2}(N-1)\Gamma_{12}c_2^2|\psi_2|^2 \quad \text{for mode 1,} \qquad (3.73)$$

$$-\frac{i\hbar}{2}(N-1)(\Gamma_{12}c_1^2|\psi_1|^2 + \Gamma_{22}c_2^2|\psi_2|^2) \quad \text{for mode 2.} \qquad (3.74)$$

The loss constants in ^{87}Rb were measured to be $\Gamma_{12}=0.780(19)\times 10^{-13}$ cm^3/s and $\Gamma_{22}=1.194(19) \times 10^{-13}$ cm^3/s. If we assume that the wave functions are the same for the two hyperfine states, as in the short-time analysis of Sec. III F, the integrated effect of the spin-exchange losses across the atomic cloud is characterized by a decay constant

$$\Gamma \equiv \frac{(N-1)\eta_N(\Gamma_{12}+\Gamma_{22}c_2^2)}{2}. \qquad (3.75)$$

We can get an idea of the importance of spin-exchange losses by comparing Γ to the angular frequency Ω_N for the integrated phase shift. The ratio of interest for comparing coherent and decoherent processes is thus

$$\frac{\Gamma}{\Omega_N} = \frac{\hbar(\Gamma_{12}+\Gamma_{22}/2)}{2\gamma_1} = \frac{m}{4\pi\hbar}\frac{\Gamma_{12}+\Gamma_{22}/2}{a_{11}-a_{22}} \simeq \frac{1}{19}, \qquad (3.76)$$

where we specialize to the case $c_1^2=c_2^2=1/2$ relevant to the ^{87}Rb protocol and the final estimate applies to that protocol. This ratio indicates that the proposed protocol can obtain an estimate of γ_1 with better than $1/N$ scaling before atom losses degrade the sensitivity.

It is also to be noted that an advantage of using a measurement scheme that uses product states is that loss of atoms from the BEC does not change the sensitivity scaling since loss of particles from a product state does not damage any coherence. There is a decay in the signal strength given by a factor $e^{-\Gamma t}$, which would require us to complete the experiment before too many atoms are lost, but the ratio (3.76) provides a window for doing this. The discussion in Sec. III F suggests, however, that differentiation of the spatial wave functions for the two modes becomes a limiting factor on the duration of the experiment before loss of atoms becomes an important consideration.

2. Number uncertainties

In real experiments the number of atoms in a BEC is not known to arbitrary precision as we have assumed so far. Thus we have to consider what happens when the number of atoms in the BEC is not fixed from trial to trial.

To analyze this situation, let $p(N_0)$ denote the probability that the number of atoms participating in our measurement protocol is N_0. The final step in the protocol is to count the number of atoms in the two hyperfine levels. The difference between the two counts is used to estimate the parameter, here denoted as γ; the sum can be used to refine the estimate of the number of atoms that participated in the protocol.

We let N_1' and N_2' be the number of atoms that would be counted by an ideal counting procedure. We generally work in terms of the total number of atoms, $N_0=N_1'+N_2'$, and the difference, $m'=(N_1'-N_2')/2$, normalized by a factor of two to match the eigenvalues of $\hat{J}_z$. Quantum mechanics gives the conditional probability $q(m'|N_0,\gamma)$ for a measurement of $\hat{J}_z$.

The counting is not completely precise, so we introduce independent conditional probabilities, $p(N_1|N_1')$ and $p(N_2|N_2')$, for counting N_1 and N_2 atoms in the two levels, given the ideal counts. We can think of these two probabilities as describing processes in which condensate atoms are missed or noncondensate atoms are counted by mistake. In addition, in a complete analysis of the protocol, we would need to include the loss of atoms, discussed in the previous subsection, in this analysis. As already noted, we are mainly interested in the total number of atoms counted, $N=N_1+N_2$, and the normalized difference, $m=(N_1-N_2)/2$. In the absence of a better model, we assume, to illustrate the effect of number uncertainties, that $p(N_1|N_1')$ and $p(N_2|N_2')$ are independent Gaussian random processes, with mean $\bar{N}_j=N_j'$ and variance $\Delta^2 N_j=\sigma^2$. Under this assumption, N and m become independent Gaussian random processes, described by conditional probabilities $p(N|N_0)$ and $p(m|m')$, which have $\bar{N}=N_0$, $\bar{m}=m'$, $\Delta^2 N=2\sigma^2$, and $\Delta^2 m=\sigma^2/2$.

The probability this model gives us directly is

$$p(N,m,m',N_0|\gamma) = p(N|N_0)p(m|m')q(m'|N_0,\gamma)p(N_0). \qquad (3.77)$$

The probability we need in order to evaluate the sensitivity of our protocol is the conditional probability for m, given the parameter γ and the measured total number of atoms, N,

$$p(m|N,\gamma) = \frac{p(N,m|\gamma)}{p(N|\gamma)} = \frac{\displaystyle\sum_{m',N_0} p(N,m,m',N_0|\gamma)}{\displaystyle\sum_{m,m',N_0} p(N,m,m',N_0|\gamma)}$$

$$= \sum_{m',N_0} p(m|m')q(m'|N_0,\gamma)p(N_0|N). \quad (3.78)$$

In the final form, $p(N_0|N) = p(N|N_0)p(N_0)/p(N)$ is the conditional probability for N_0 atoms to have participated in the protocol, given the measured total count N. It quantifies the refinement in the knowledge of N_0 provided by the total count.

The quantities that go into determining the sensitivity are the mean and second moment of m, calculated from the probability (3.78),

$$\bar{m}_{N,\gamma} = \sum_{N_0} \langle \hat{J}_z \rangle_{N_0,\gamma}\, p(N_0|N), \quad (3.79)$$

$$\overline{(m^2)}_{N,\gamma} = \frac{1}{2}\sigma^2 + \sum_{N_0} (\langle \hat{J}_z \rangle_{N_0,\gamma}^2 + (\Delta^2 \hat{J}_z)_{N_0,\gamma})p(N_0|N),$$

$$(3.80)$$

where $\langle \hat{J}_z \rangle_{N_0,\gamma} = \overline{(m')}_{N_0,\gamma}$ and $(\Delta^2 \hat{J}_z)_{N_0,\gamma} = (\Delta^2 m')_{N_0,\gamma}$ are the mean and variance of $\hat{J}_z$ calculated from the quantum-mechanical probabilities. If σ is much less than the initial uncertainty in N_0, which is itself somewhat less than N_0 (depending on the care taken in loading the trap), then the measured total count N gives a very good, improved estimate of the number of atoms that participated in the protocol; under these circumstances, the probability $p(N_0|N)$ is peaked at the measured value N, with half-width given very nearly by σ. The quantum-mechanical expectation values vary over a range from $-N_0/2$ to $+N_0/2$, so as long as $\sigma \ll N$, we can evaluate the averages over $p(N_0|N)$ at the mean value N with little error, thus giving mean $\bar{m}_{N,\gamma} = \langle \hat{J}_z \rangle_{N,\gamma}$ and variance $(\Delta^2 m)_{N,\gamma} = \sigma^2/2 + (\Delta^2 \hat{J}_z)_{N,\gamma}$. The resulting measurement uncertainty in determining γ,

$$\delta\gamma^2 = \frac{(\Delta^2 m)_{N,\gamma}}{|\partial \bar{m}_{N,\gamma}/\partial \gamma|^2} = \frac{\sigma^2/2 + (\Delta^2 \hat{J}_z)_{N,\gamma}}{|\partial \langle \hat{J}_z \rangle_{N,\gamma}/\partial \gamma|^2}, \quad (3.81)$$

has the quantum-mechanical scaling and nearly the optimal sensitivity, provided we can count atoms to better than $\sqrt{N}$, i.e., $\sigma \lesssim \sqrt{N}$. Ultimately, what this result expresses is that the variance of the measurement of $\hat{J}_z$ in our protocol is of order $\sqrt{N}$, so we need to know the number of atoms to this same accuracy.

IV. CONCLUSION

This paper serves two purposes. The first is to extend the discussion of Heisenberg-limited quantum metrology from

its traditional focus on a $1/N$ scaling for measurement uncertainty. Our discussion centers on the role of the dynamics of an N-qubit quantum probe in determining the quantum Cramér-Rao bound for single-parameter estimation. Looking at quantum metrology using the language of quantum circuits makes it easy to see that abandoning the usual independent couplings of the parameter to the qubits in favor of nonlinear couplings can yield scalings better than $1/N$. With k-body couplings, it is possible to achieve sensitivities scaling as $1/N^k$. Although the $1/N^k$ scaling requires entangled input states, $1/N^{k-1/2}$ scalings can be obtained with initial product states. Thus a sensitivity scaling as $1/N^{3/2}$ can be achieved if quadratic couplings to the parameter can be engineered; moreover, particular quadratic couplings yield this sensitivity even though the state remains unentangled under the dynamics, thus showing that super-$1/N$ scaling can be achieved without any entanglement.

The second purpose of this paper is to show that a two-component BEC is a promising candidate system for a proof-of-principle experiment that demonstrates scaling better than $1/N$. A simplified analysis of the system, based on strong assumptions, but followed by a more detailed analysis of the realm of applicability of those assumptions, shows that such an experiment might indeed be realizable. This work motivates further, yet more detailed analyses and numerical simulations of the experiment. We have undertaken such further investigations of the proposed metrology scheme, and this further work, to be reported elsewhere, supports the conclusions reached in this paper. Our numerical studies include computing the ground-state solution of the time-independent GP equation for different values of N in order to find the exact dependence of η_N on N. Numerical integration of the time-dependent coupled two-mode GP Eqs. (3.60) is then used to compute the expected signal [Eq. (3.62)] in order to compare it with the theoretical prediction in Eq. (3.71).

The quantity that is measured in the proposed metrology protocol is essentially a constant. Estimating a constant using sophisticated quantum measurement schemes is interesting only as a proof of principle because there is nothing to preclude estimating the same constant using much simpler, classical measurement techniques. Since the measured quantity is a constant, we have the time to perform whatever number of repetitions of the simplest estimation procedure is required to achieve the desired accuracy. Metrology protocols of the type described here are relevant in circumstances where there are constraints on the available time or on the available number of qubits. The available time can be constrained, for example, because the quantity that is being measured is changing, as in the case of gravitational-wave detection or magnetometry. There can be further time constraints placed by decoherence of the probe qubits. In such scenarios, picking the optimal metrology scheme with the best measurement uncertainty, given the constraints, is of primary importance [29]. For our proposal using a BEC, one possibility is to work around a broad Feshbach resonance that makes the scattering lengths sensitive to external magnetic fields. We might then be able to use our scheme for high-precision magnetometry.

APPENDIX: INTEGRALS OVER THOMAS-FERMI DISTRIBUTIONS

In the intermediate TF regime, i.e., $N_L \ll N \ll N_T$, we need to do integrals over the TF probability density (3.47),

$$
\begin{aligned}
I_l(N,d,q) &\equiv \int d^d r |\phi_N(\boldsymbol{r})|^{2l} \\
&= \left(\frac{k/2}{(N-1)g\,\eta_T} \right)^l \int d\Omega_{d-1} \int_0^{\tilde{r}_N} r^{d-1} dr (\tilde{r}_N^q - r^q)^l \\
&= \left(\frac{k/2}{(N-1)g\,\eta_T} \right)^l \tilde{r}_N^{d+ql} S_{d-1} J_l(d,q) \\
&= \left(\frac{\mu_L}{(N-1)g\,\eta_T} \right)^l \tilde{r}_N^d S_{d-1} J_l(d,q),
\end{aligned}
\tag{A.1}
$$

where $S_{d-1} = d V_d$ is the area of a unit sphere in $d-1$ dimensions and

$$
\begin{aligned}
J_l(d,q) &\equiv \int_0^1 du\, u^{d-1} (1 - u^q)^l \\
&= -\frac{\pi}{\sin(l\pi)\Gamma(-l)} \frac{\Gamma(d/q)}{q\Gamma(d/q + l + 1)}
\end{aligned}
\tag{A.2}
$$

for $l > -1$.

It is easy to see that

$$
\frac{J_{x+l}(d,q)}{J_x(d,q)} = \frac{(x+1)\cdots(x+l)}{(d/q + x + 1)\cdots(d/q + x + l)}.
\tag{A.3}
$$

Combined with $J_0 = 1/d$, this gives, when l is a nonnegative integer,

$$
J_l(d,q) = \frac{l!\, q^l}{d(d+q)(d+2q)\cdots(d+lq)}.
\tag{A.4}
$$

Notice also that

$$
J_l(d,2) = \int_0^{\pi/2} dv\, \sin^{d-1} v \, \cos^{2l+1} v = \frac{\Gamma(d/2)\Gamma(l+1)}{2\Gamma(d/2 + l + 1)}.
\tag{A.5}
$$

Now we use

$$
\frac{k/2}{(N-1)g\,\eta_T} S_{d-1} = \frac{d}{r_0^{d+q}} \frac{N_L - 1}{N - 1}
\tag{A.6}
$$

to write

$$
I_1(N,d,q) = d J_1(d,q) \left(\frac{\tilde{r}_N}{r_0} \right)^{d+q} \frac{N_L - 1}{N - 1} = \frac{q}{d+q} \left(\frac{\tilde{r}_N}{r_0} \right)^{d+q} \frac{N_L - 1}{N - 1}
\tag{A.7}
$$

and

$$
I_l(N,d,q) = I_1(N,d,q) \frac{J_l(d,q)}{J_1(d,q)} \left(\frac{\mu_L}{(N-1)g\,\eta_T} \right)^{l-1}.
\tag{A.8}
$$

In the upper TF regime, i.e., $N \gg N_T$, we need to do integrals over the TF probability density (3.53),

$$
\begin{aligned}
K_l(N,d,q) &\equiv \int d^D \rho\, d^d r |\psi_N(\boldsymbol{\rho},\boldsymbol{r})|^{2l} = \left(\frac{k/2}{(N-1)g} \right)^l \int d\Omega_{D-1} \int_0^{\tilde{\rho}_N} \rho^{D-1} d\rho \int d\Omega_{d-1} \int_0^{\tilde{r}_N(\rho)} r^{d-1} dr (\tilde{r}_N^q(\rho) - r^q)^l \\
&= \left(\frac{m\omega_T^2/2}{(N-1)g} \right)^l \left(\frac{m\omega_T^2}{k} \right)^{d/q} \tilde{\rho}_N^{D+2(l+d/q)} S_{D-1} S_{d-1} J_{l+d/q}(D,2) J_l(d,q) \\
&= \left(\frac{\mu_N}{(N-1)g} \right)^l \left(\frac{m\omega_T^2}{k} \right)^{d/q} \tilde{\rho}_N^{D+2d/q} S_{D-1} S_{d-1} J_{l+d/q}(D,2) J_l(d,q).
\end{aligned}
\tag{A.9}
$$

Now we use $m\omega_T^2/k = r_0^{q+2}/4\rho_0^4$ and

$$
\frac{m\omega_T^2/2}{(N-1)g} \left(\frac{m\omega_T^2}{k} \right)^{d/q} = \frac{1}{32\pi\beta_d 2^{2d/q}} \frac{1}{\rho_0^{5-d+2d/q}} \frac{N_T - 1}{N - 1}
\tag{A.10}
$$

to write

$$K_1(N,d,q) = dJ_{1+d/q}(D,2)J_1(d,q)\frac{S_{D-1}}{4(4\pi)^{D/2}2^{2d/q}}\left(\frac{\widetilde{\rho}_N}{\rho_0}\right)^{5-d+2d/q}\frac{N_T-1}{N-1} \tag{A.11}$$

and

$$K_l(N,d,q) = K_1(N,d,q)\frac{J_{l+d/q}(d,q)J_l(d,q)}{J_{1+d/q}(d,q)J_1(d,q)}\left(\frac{\mu}{(N-1)g}\right)^{l-1}. \tag{A.12}$$

[1] I. E. Mandelstam and I. Tamm, J. Phys. (USSR) **9**, 249 (1945).

[2] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[3] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

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

[5] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. **247**, 135 (1996).

[6] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[7] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Phys. Rev. Lett. **98**, 090401 (2007).

[8] S. Boixo, A. Datta, S. T. Flammia, A. Shaji, E. Bagan, and C. M. Caves, Phys. Rev. A **77**, 012317 (2008).

[9] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deleglise, U. B. Hoff, M. Brune, J.-M. Raimond, and S. Haroche, Nature (London) **446**, 297 (2007).

[10] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).

[11] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).

[12] A. Luis, Phys. Lett. A **329**, 8 (2004).

[13] J. Beltran and A. Luis, Phys. Rev. A **72**, 045801 (2005).

[14] A. Luis, Phys. Rev. A **76**, 035801 (2007).

[15] A. M. Rey, L. Jiang, and M. D. Lukin, Phys. Rev. A **76**, 053617 (2007).

[16] S. Choi and B. Sundaram, Phys. Rev. A **77**, 053613 (2008).

[17] M. J. Woolley, G. J. Milburn, and C. M. Caves, New J. Phys. **10**, 125018 (2008).

[18] S. Boixo, A. Datta, M. J. Davis, S. T. Flammia, A. Shaji, and C. M. Caves, Phys. Rev. Lett. **101**, 040403 (2008).

[19] A. J. Leggett, Rev. Mod. Phys. **73**, 307 (2001).

[20] A. L. Fetter, in *Bose-Einstein Condensation in Atomic Gases*, edited by M. Inguscio, S. Stringari, and C. E. Wieman (IOS Press, Amsterdam, 1999).

[21] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).

[22] J. Rogel-Salazar, S. Choi, G. H. C. New, and K. Burnett, J. Opt. B: Quantum Semiclassical Opt. **6**, R33 (2004).

[23] J. E. Williams, Ph.D. thesis, University of Colorado, 1999.

[24] M. R. Matthews, D. S. Hall, D. S. Jin, J. R. Ensher, C. E. Wieman, E. A. Cornell, F. Dalfovo, C. Minniti, and S. Stringari, Phys. Rev. Lett. **81**, 243 (1998).

[25] D. S. Hall, M. R. Matthews, J. R. Ensher, C. E. Wieman, and E. A. Cornell, Phys. Rev. Lett. **81**, 1539 (1998).

[26] D. S. Hall, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Phys. Rev. Lett. **81**, 1543 (1998).

[27] K. M. Mertes, J. W. Merrill, R. Carretero-Gonzalez, D. J. Frantzeskakis, P. G. Kevrekidis, and D. S. Hall, Phys. Rev. Lett. **99**, 190402 (2007).

[28] R. Anderson, C. Ticknor, A. Sidorov, and B. V. Hall, Phys. Rev. A **80**, 023603 (2009).

[29] A. Shaji and C. M. Caves, Phys. Rev. A **76**, 032111 (2007).