

Improved Quantum Magnetometry beyond the Standard Quantum Limit

J. B. Brask,¹ R. Chaves,^{2,5} and J. Kołodyński^{3,4}

¹*Département de Physique Théorique, University of Geneva, 1211 Geneva, Switzerland*

²*Institute for Physics, University of Freiburg, Rheinstrasse 10, D-79104 Freiburg, Germany*

³*ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

⁴*Faculty of Physics, University of Warsaw, 02-093 Warsaw, Poland*

⁵*Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany*

(Received 21 November 2014; published 22 July 2015)

Under ideal conditions, quantum metrology promises a precision gain over classical techniques scaling quadratically with the number of probe particles. At the same time, no-go results have shown that generic, uncorrelated noise limits the quantum advantage to a constant factor. In frequency estimation scenarios, however, there are exceptions to this rule and, in particular, it has been found that transversal dephasing does allow for a scaling quantum advantage. Yet, it has remained unclear whether such exemptions can be exploited in practical scenarios. Here, we argue that the transversal-noise model applies to the setting of recent magnetometry experiments and show that a scaling advantage can be maintained with one-axis-twisted spin-squeezed states and Ramsey-interferometry-like measurements. This is achieved by exploiting the geometry of the setup that, as we demonstrate, has a strong influence on the achievable quantum enhancement for experimentally feasible parameter settings. When, in addition to the dominant transversal noise, other sources of decoherence are present, the quantum advantage is asymptotically bounded by a constant, but this constant may be significantly improved by exploring the geometry.

DOI: [10.1103/PhysRevX.5.031010](https://doi.org/10.1103/PhysRevX.5.031010)

Subject Areas: Atomic and Molecular Physics,
Quantum Physics, Quantum Information

I. INTRODUCTION

High-precision parameter estimation is fundamental throughout science. Quite generally, a number of probe particles are prepared, then subjected to an evolution that depends on the quantity of interest, and finally measured. From the measurement results an estimate is then extracted. When the particles are classically correlated and non-interacting, as a consequence of the central limit theorem, the mean-squared error of the estimate decreases as $1/N$, where N is the number of particles (probe size). This best scaling achievable with a classical probe is known as the standard quantum limit (SQL) [1]. Quantum metrology aims to improve estimation by exploiting quantum correlations in the probe.

In an ideal setting without noise, it is well known that quantum resources allow for a quadratic improvement in precision over the SQL [1,2]; i.e., the mean-squared error of the estimate after a sufficient number of experimental repetitions can scale as $1/N^2$, yielding the so-called Heisenberg limit. Realistic evolution, however, always involves noise of some form, and although quantum

metrology has been demonstrated experimentally, e.g., for atomic magnetometry [3–9], spectroscopy [10,11], and clocks [12,13], there is currently much effort to determine exactly when, and by how much, quantum resources allow estimation to be improved in the presence of decoherence [14–33]. It is known that for most types of uncorrelated noise (acting independently on each probe particle) the asymptotic scaling is constrained to be SQL-like [17–24]. Specifically, when estimating a parameter ω , the mean-squared error obeys $\Delta^2\omega \geq r/\nu N$, where ν is the number of repetitions and r is a constant that depends on the evolution. If the evolution, which each probe particle undergoes, is independent of N , the scaling is constrained to be SQL-like. However, for frequency estimation this is not necessarily the case. In frequency estimation scenarios, such as those of atomic magnetometry [34–36], spectroscopy [37–39], and clocks [40–45], there are two relevant resources, the total number of probe particles N and the total time T available for the experiment [14,15]. The experimenter is free to choose the interrogation time $t = T/\nu$ and, in particular, t may be adapted to N . In this case, the time over which unitary evolution and decoherence act is different for each N and thus the evolution is not independent of N . Schematically, the no-go results for noisy evolution in this case become

$$\Delta^2\omega \geq \frac{r(t)}{NT/t} \xrightarrow{t(N)} \Delta^2\omega T \geq \frac{c[t(N)]}{N}, \quad (1)$$

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

with $c(t) = r(t)t$. Thus, if for some optimal choice of $t(N)$ the coefficient c decreases with N , although the no-go results may hold for any fixed evolution time, the bound does not imply SQL-like scaling. Note that the bound Eq. (1) is always achievable in the many-repetitions limit $\nu \rightarrow \infty$ [2], which corresponds to $T \gg t$ [14–16]. Although without noise it is optimal to take t as large as possible, i.e., $t = T$, for any noisy evolution the optimal t becomes finite because of noise dominating at large times. So the many-repetitions regime can always be ensured by considering sufficiently large T [24–27].

In frequency estimation scenarios, for the asymptotic scaling to be superclassical, c must vanish as $N \rightarrow \infty$, which is only possible if the evolution is such that decoherence can be neglected at short time scales, and the no-go theorems then do not apply [46]. This is also necessary for error-correction techniques, which utilize ancillary particles not sensing the parameter [29] or employ correcting pulses during the evolution [28], to surpass the SQL [22,23,46]. Without such additional resources—considering just interrogation-time optimization—the possibility of superclassical scaling has been demonstrated for non-Markovian [25,26] evolutions (for which the effective decoherence strength vanishes as $t \rightarrow 0$), as well as for dephasing directed along a direction perpendicular to the unitary evolution [27]. In the latter case, it was shown that an optimal variance scaling of $1/N^{5/3}$ can be obtained by choosing $t \propto 1/N^{1/3}$ [27]. This result was based on numerical analysis of the quantum Fisher information [47] and was shown to be saturable by Greenberger-Horne-Zeilinger (GHZ) states [48]. However, GHZ states of many particles are not easily generated in practice, and the Fisher information approach does not explicitly provide the required measurements. Thus, the question of whether the scaling is achievable in practically implementable metrology was left open.

In this paper, we argue that the transversal-noise model applies to atomic magnetometry, in particular, the experimental setting of [3], and study the quantum advantage attainable with use of one-axis-twisted spin-squeezed states (OATSSs) [49] and Ramsey-interferometry-like measurements [37–39], both of which are accessible with current experimental techniques. We explicitly show that the setup geometry plays an important role for the achievable quantum enhancement. A suboptimal choice leads to a constant factor of quantum enhancement, while superclassical precision scaling can be maintained for a more appropriate choice. We study the enhancement achievable with the numbers of the experiment [3] and demonstrate the advantage of modifying the geometry. We further consider the case of noise that is not perfectly transversal and find that, although the asymptotic precision scaling is then again SQL-like, the precision may be substantially enhanced by optimizing the geometry. As the previous results [27] were based on numerics, we also provide an analytical proof of the scaling for GHZ states in Appendix C.

II. MODEL

We consider a scheme in which N two-level quantum systems are used to sense a frequency parameter ω in an experiment of total duration T , divided into rounds of interrogation time t . We keep in mind that this can correspond to atomic magnetometry, in which the particles then represent the atoms with a spin precessing in a magnetic field at a frequency proportional to the field strength. As in Ref. [27], we describe the noisy evolution by a master equation of Lindblad form

$$\frac{\partial \rho}{\partial t} = \mathcal{H}(\rho) + \mathcal{L}(\rho). \quad (2)$$

Here, $\mathcal{H}(\rho) = -i[\hat{H}, \rho]$ is the unitary part of the evolution that encodes the parameter dependence. The Hamiltonian is given by

$$\hat{H} = \frac{\omega}{2} \sum_{k=1}^N \hat{\sigma}_z^{(k)}, \quad (3)$$

where $\hat{\sigma}_z^{(k)}$ is a Pauli operator acting on the k th particle (qubit). The Liouvillian $\mathcal{L}(\rho)$ describes the noise, which is uncorrelated on different qubits, so that $\mathcal{L} = \sum_k \mathcal{L}^{(k)}$, and for a single qubit we have

$$\mathcal{L}^{(k)} \rho = -\frac{\gamma}{2} [\rho - \alpha_x \hat{\sigma}_x^{(k)} \rho \hat{\sigma}_x^{(k)} - \alpha_y \hat{\sigma}_y^{(k)} \rho \hat{\sigma}_y^{(k)} - \alpha_z \hat{\sigma}_z^{(k)} \rho \hat{\sigma}_z^{(k)}], \quad (4)$$

where γ is the overall noise strength and $\alpha_{x,y,z} \geq 0$, with $\alpha_x + \alpha_y + \alpha_z = 1$. For $\alpha_z = 1$, Eq. (4) describes dephasing along the direction of the unitary, while $\alpha_x = 1$ (or equivalently $\alpha_y = 1$) corresponds to the transversal-dephasing noise. For $\alpha_x = \alpha_y = \alpha_z = 1/3$, we have an isotropic depolarizing channel.

Under this model, interrogation-time optimization leads to a quantum scaling advantage for transversal ($\alpha_x = 1$) but not for parallel ($\alpha_z = 1$) noise. This can be understood by looking at how the coefficient c in Eq. (1) behaves in the two cases. For short times, one can obtain bounds of the form Eq. (1) with [27]

$$c_z(\gamma, \omega, t) = 2\gamma + 2\gamma^2 t + O(t^2), \quad (5)$$

$$c_x(\gamma, \omega, t) = \frac{\gamma^2 \omega^3 t^3}{12} + O(t^5). \quad (6)$$

From this, we see that for parallel dephasing, interrogation-time optimization cannot prevent asymptotic SQL-like scaling, because c_z is bounded from below by the nonzero factor of 2γ . However, for perpendicular noise $c_x \rightarrow 0$ as $t \rightarrow 0$, and hence in this case, optimization may allow for superclassical scaling. In Ref. [27], it was found that taking $t = (3/\gamma\omega^2 N)^{1/3}$ leads to

$$\Delta^2 \omega T \geq \frac{3^{2/3}}{2} (\gamma \omega^2)^{1/3} \frac{1}{N^{5/3}}, \quad (7)$$

and that this bound is achievable with GHZ input states.

To see that the model is relevant in practice, we consider the atomic magnetometry experiment of Ref. [3] illustrated in Fig. 1. In this experiment, entanglement was demonstrated to enhance the sensitivity, but the precision scaling with N was not studied. The relevant magnetometer consists of a vapor of caesium atoms, which is subject to a strong external dc magnetic field B and used to sense a weak radio-frequency field B_{rf} perpendicular to B (note that in Ref. [3] two separate ensembles were used; this is not important for the present argument). The atoms are optically pumped into an extreme magnetic sublevel and may be treated as effective two-level systems with an energy splitting determined by B . With $B \gg B_{\text{rf}}$, the dominant noise is due to small variations in the dc magnetic field seen by different atoms (e.g., due to field inhomogeneities and atomic motion), which leads to fluctuations of the individual energy splittings. This corresponds to a dephasing noise that acts on each atom independently and is characterized by the spin-decoherence time T_2 [50]. As the experiment is conducted at a time scale much shorter than the ones of spontaneous emission and B -field fluctuations, other noise sources are suppressed. In particular, the spin-relaxation time T_1 can be taken infinite (see Appendix A for discussion) and collective noise can be neglected. The Larmor frequency of the strong field B is matched to the frequency of the weak field B_{rf} , and it is then convenient to describe the system in a rotating frame (RF). If ρ is the state of a single atom in the nonrotating frame and B is directed along the x axis (see Fig. 1), the state in the rotating frame reads $\rho_{\text{RF}} = e^{i\hat{H}_B t} \rho e^{-i\hat{H}_B t}$, where $\hat{H}_B = \kappa B \hat{\sigma}_x$ and κ is the coupling strength to the magnetic field. In such a Larmor-precessing frame, the master equation for the evolution may be written as (see also Appendix A) [35]

$$\frac{\partial \rho_{\text{RF}}}{\partial t} = -i\kappa B_{\text{rf}} [\hat{\sigma}_z, \rho_{\text{RF}}] - \frac{1}{T_2} (\rho_{\text{RF}} - \hat{\sigma}_x \rho_{\text{RF}} \hat{\sigma}_x), \quad (8)$$

where the first term can be understood as the effective free Hamiltonian in the rotating frame with the B_{rf} field pointing along z . The dephasing noise is directed along B and parametrized by T_2 . Since Eq. (8) is exactly of the form Eq. (2), it is clear that this experimental setting is captured by the previously stated model with $\omega = 2\kappa B_{\text{rf}}$ and transversal noise $\alpha_x = 1$, $\alpha_y = \alpha_z = 0$, $\gamma = 2/T_2$. We note that $B \gg B_{\text{rf}}$ is important for the noise to be transversal, which may imply that γ is large relative to ω . In particular, this is the case in Ref. [3], as we show below.

In Ref. [3], superclassical precision was demonstrated by initially aligning the collective spin of the atomic ensemble *along* B and reducing fluctuations of its component in the direction perpendicular to both B and B_{rf} via spin

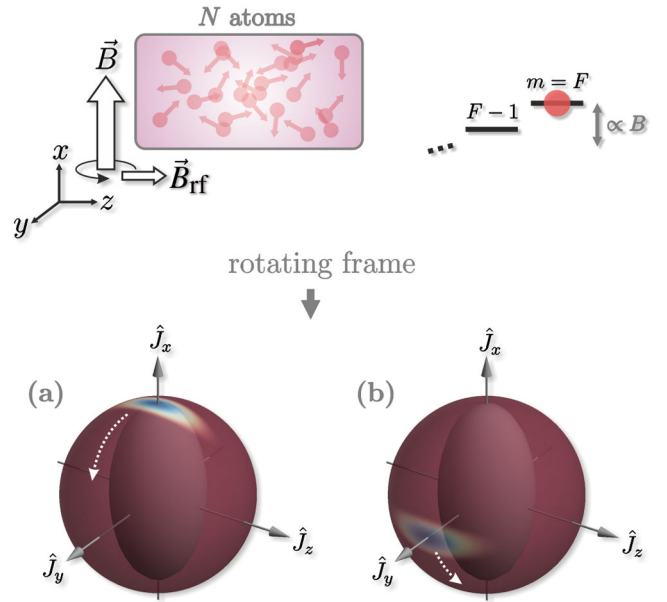


FIG. 1. Atomic magnetometry setup. An ensemble of atoms is placed in a strong magnetic field B which induces a level splitting between the magnetic sublevels. The atoms are used to sense a weak field B_{rf} in the plane perpendicular to B , which rotates in this plane with a frequency matched to the Larmor precession induced by B . We consider two cases for the state preparation and readout. Scenario (a) corresponds to the geometry of the experiment [3]. All atoms are initially pumped to an extreme magnetic sublevel $m = F$ creating a coherent spin state aligned with B . The state is then squeezed to make it more sensitive to the evolution induced by B_{rf} . In a frame rotating around B at the Larmor frequency, the state can be depicted as shown in the lower part. B_{rf} points then along z and induces a rotation around the z axis. The state is squeezed in y and B_{rf} is estimated from a measurement of the collective-spin component \hat{J}_y . Scenario (b) is similar, but the state is initially perpendicular to B . In the rotating frame, it is squeezed in x and \hat{J}_x is measured. The dominant noise in both cases comes from individual atomic motion causing variations in the effective magnetic field and hence the energy splitting. This results in uncorrelated dephasing noise in the direction of B ; the impact on the collective spin is schematically illustrated by the inner prolate spheroids. Importantly, the noise preserves the spin along B but shrinks it in the perpendicular directions.

squeezing [Fig. 1(a)]. Below, we study such a geometry along with another setting, in which the collective spin is initially *perpendicular* to both B and B_{rf} and its component along B is squeezed [Fig. 1(b)]. In principle, scenario (b) can be obtained from (a) by applying a $\pi/2$ pulse to the atomic ensemble before the evolution. In both cases, B_{rf} is estimated from a measurement of a component of the collective spin (in the rotating frame), read out, e.g., via the scheme of Ref. [3], which resembles a standard Ramsey-interferometry [37–39] measurement. We show that in (b), for one-axis-twisted spin-squeezed states, a superclassical scaling $1/N^{5/4}$ of the mean-squared error can be maintained,

thus demonstrating that a scaling quantum advantage is possible with feasible states and measurements. At the same time, we find that in (a) the quantum advantage is limited by a constant, which matches the bound for parallel dephasing [14,24]. As a consequence, for an atomic ensemble size and parameters matching the experiment of Ref. [3], (b) may considerably outperform (a).

As an aside, we note that when the true value of the estimated parameter is zero, the bound Eq. (7) vanishes. This does not mean that the precision is unbounded, but indicates that the bound gives no information in such a limit. One may then speculate whether the scaling can be further improved if ω can be made arbitrarily small in an adaptive manner. We discuss this issue in Appendix D.

III. COMPUTING PRECISION FOR SPECIFIC STATES AND MEASUREMENTS

To obtain results for the precision achievable within the above scenarios, we make use of error propagation and apply it for adequate choices of squeezed states and collective-spin-observable measurements. Generally, when a parameter ϕ is estimated based on the average of measuring an observable \hat{O} , and when the prior knowledge of ϕ is sufficiently tight, fluctuations in the estimate can be linearly related to the fluctuations in \hat{O} . Thus, for a system in a state ρ , in such a local estimation regime the mean-squared error of the estimate may be quantified as

$$\Delta^2\phi = \frac{(\Delta^2\hat{O})_\rho}{|\partial\langle\hat{O}\rangle_\rho/\partial\phi|^2} = \frac{\langle\hat{O}^2\rangle_\rho - \langle\hat{O}\rangle_\rho^2}{|\partial\langle\hat{O}\rangle_\rho/\partial\phi|^2}. \quad (9)$$

If the measurement is repeated, $\Delta^2\phi$ will additionally decrease inverse proportionally to the number of repetitions ν , which also ensures the above local regime as $\nu \rightarrow \infty$ and thus that Eq. (9) always holds. Here, we are interested in frequency estimation over a total time T with a single-round duration t , such that $\nu = T/t$. We therefore write the overall mean-squared error of the ω estimate as [2]

$$\Delta^2\omega T = t \frac{(\Delta^2\hat{O})_t}{|\partial\langle\hat{O}\rangle_t/\partial\omega|^2}, \quad (10)$$

which, like Eq. (1), is valid in the $T \gg t$ regime, i.e., for sufficiently large T . The expectation values in Eq. (10) can be evaluated by computing the expectation value of either the static operator in the time-evolved state or the time-evolved operator in the input state (analogously to the usual Schrödinger and Heisenberg pictures for unitary dynamics). Specifically, in terms of the Kraus representation of the evolution, one has

$$\langle\hat{O}\rangle_t = \sum_s \text{Tr}[\hat{O}K_s(t)\rho_0 K_s^\dagger(t)] = \sum_s \text{Tr}[K_s^\dagger(t)\hat{O}K_s(t)\rho_0], \quad (11)$$

where \hat{O} is the time-independent observable, ρ_0 is the input state, and K_s are the Kraus operators of the global channel. For independent channels acting on each qubit, $K_s = K_{s_1} \otimes \cdots \otimes K_{s_N}$, where the K_{s_i} are the Kraus operators acting on the i th qubit.

In subsequent sections, we determine the precisions attainable under our model Eq. (2) for specific input states and measurements. The model has four Kraus operators, which have the form

$$\begin{aligned} K_1 &= a_1\hat{\sigma}_y, & K_3 &= a_3\hat{\sigma}_z - ib_3\mathbb{1}, \\ K_2 &= a_2\hat{\sigma}_x, & K_4 &= a_4\hat{\sigma}_z - ib_4\mathbb{1}. \end{aligned} \quad (12)$$

Here, the coefficients a_i , b_i are real and depend on the frequency ω , the noise parameters γ , α_x , α_y , α_z , and the time t (see Appendix B). However, to simplify notation we suppress these dependences. Because of trace preservation, $\sum_s K_s^\dagger K_s = \mathbb{1}$, the coefficients must satisfy

$$a_1^2 + a_2^2 + a_3^2 + a_4^2 + b_3^2 + b_4^2 = 1. \quad (13)$$

For later calculations, it is useful to compute the evolution of both $\hat{\sigma}_x$ and $\hat{\sigma}_y$ under the Kraus map. For $\hat{\sigma}_x$, we have

$$\begin{aligned} K_1^\dagger\hat{\sigma}_x K_1 &= -a_1^2\hat{\sigma}_x, \\ K_2^\dagger\hat{\sigma}_x K_2 &= a_2^2\hat{\sigma}_x, \\ K_3^\dagger\hat{\sigma}_x K_3 &= (-a_3^2 + b_3^2)\hat{\sigma}_x + 2a_3b_3\hat{\sigma}_x, \\ K_4^\dagger\hat{\sigma}_x K_4 &= (-a_4^2 + b_4^2)\hat{\sigma}_x + 2a_4b_4\hat{\sigma}_x. \end{aligned} \quad (14)$$

Using Eq. (13), the evolution under the channel can then be written as (Pauli operators with no explicit time dependence are time independent)

$$\hat{\sigma}_x(t) = \sum_s K_s^\dagger\hat{\sigma}_x K_s = \xi_x\hat{\sigma}_x + \chi_x\hat{\sigma}_y, \quad (15)$$

where the coefficients $\xi_x = 1 - 2(a_1^2 + a_3^2 + a_4^2)$, $\chi_x = 2(a_3b_3 + a_4b_4)$ are again real and encode the full dependence of the evolved operator on time, frequency, and the noise parameters. They are given in Appendix B. Similarly, one obtains

$$\hat{\sigma}_y(t) = \sum_s K_s^\dagger\hat{\sigma}_y K_s = \xi_y\hat{\sigma}_y + \chi_y\hat{\sigma}_x, \quad (16)$$

with $\xi_y = 1 - 2(a_2^2 + a_3^2 + a_4^2)$, $\chi_y = -2(a_3b_3 + a_4b_4)$.

IV. BEATING THE SQL WITH REALISTIC STATES AND MEASUREMENTS

Several experiments have demonstrated superclassical sensitivity of magnetometry with atomic ensembles by squeezing the collective atomic spin [3–9]. Considering the perpendicular model noise, we now show that

spin-squeezed states and Ramsey-type measurements together with interrogation-time optimization are sufficient not only to reach precisions unattainable by classical protocols but also to maintain superclassical precision scaling with the particle number.

A. Collective spin

Ramsey interferometry performed on a collection of spin-1/2 particles (qubits) effectively corresponds to collective-spin measurements [37–39]. Here, we consider the components of collective spin along x and y ,

$$\hat{J}_x = \frac{1}{2} \sum_k \hat{\sigma}_x^{(k)}, \quad \hat{J}_y = \frac{1}{2} \sum_k \hat{\sigma}_y^{(k)}, \quad (17)$$

which specify the observables measured in scenarios (b) and (a) of Fig. 1, respectively. The evolution of \hat{J}_x under the model Eq. (2) follows directly from Eq. (15),

$$\hat{J}_x(t) = \sum_s K_s^\dagger \hat{J}_x K_s = \xi_x \hat{J}_x + \chi_x \hat{J}_y, \quad (18)$$

and similarly for \hat{J}_y using Eq. (16). The derivatives with respect to the estimated parameter then read

$$\frac{\partial \langle \hat{J}_x \rangle_t}{\partial \omega} = \frac{\partial \xi_x}{\partial \omega} \langle \hat{J}_x \rangle_0 + \frac{\partial \chi_x}{\partial \omega} \langle \hat{J}_y \rangle_0, \quad (19)$$

and similarly for \hat{J}_y after interchanging $x \leftrightarrow y$. We also compute [note that taking the square and evolving do not commute because the evolution is not unitary, i.e., $\hat{J}_x^2(t) \neq (\hat{J}_x(t))^2$]

$$\begin{aligned} \hat{J}_x^2(t) &= \sum_s K_s^\dagger \hat{J}_x^2 K_s = \frac{1}{4} \sum_{k,k'} \sum_s K_s^\dagger \hat{\sigma}_x^{(k)} \hat{\sigma}_x^{(k')} K_s \\ &= \frac{N}{4} + \frac{1}{4} \sum_{k \neq k'} \hat{\sigma}_x^{(k)}(t) \otimes \hat{\sigma}_x^{(k')}(t) \\ &= \frac{N}{4} + \frac{1}{4} \left(\sum_k \hat{\sigma}_x^{(k)}(t) \right)^2 - \frac{1}{4} \sum_k [\hat{\sigma}_x^{(k)}(t)]^2 \\ &= \frac{N}{4} + [\hat{J}_x(t)]^2 - \frac{1}{4} \sum_k (\xi_x \hat{\sigma}_x^{(k)} + \chi_x \hat{\sigma}_y^{(k)})^2 \\ &= \frac{N}{4} (1 - \xi_x^2 - \chi_x^2) + (\xi_x \hat{J}_x + \chi_x \hat{J}_y)^2, \end{aligned} \quad (20)$$

so that from Eqs. (18) and (20) we obtain the variance

$$\begin{aligned} (\Delta^2 \hat{J}_x)_t &= \frac{N}{4} (1 - \xi_x^2 - \chi_x^2) + \xi_x^2 (\Delta^2 \hat{J}_x)_0 \\ &\quad + \chi_x^2 (\Delta^2 \hat{J}_y)_0 + \xi_x \chi_x [\text{cov}(\hat{J}_x, \hat{J}_y)]_0, \end{aligned} \quad (21)$$

with cov denoting the covariance. The variance for \hat{J}_y is again obtained by just replacing $x \leftrightarrow y$.

For a specific initial state of the atomic ensemble with both its expectation values and variances known at $t = 0$, we can substitute the above expressions into Eq. (10), in order to quantify the precision attained in scenarios (a) and (b) of Fig. 1 for a given interrogation time t .

B. One-axis-twisted spin-squeezed states

There is no unique definition of spin squeezing [49], but generally spin-squeezed states are states in which fluctuations of the collective-spin component are reduced in a particular direction, when compared to the value they would have in a state with all individual spins aligned, i.e., in a coherent spin state (CSS), an eigenstate of the corresponding spin component with maximal eigenvalue. Spin-squeezed states are useful for metrology due to their enhanced sensitivity to any change of the collective spin in the squeezed direction, e.g., caused by precession in a magnetic field.

A number of experiments, in particular Ref. [3], employ the so-called two-axis-twisted spin-squeezed states, which can be generated by quantum nondemolition measurement of the collective atomic spin mediated by light. However, here we focus on one-axis-twisted spin-squeezed states because they are amenable to analytical treatment. As two-axis-twisted states allow for stronger suppression of the collective-spin variance in a particular direction, i.e., stronger squeezing, we expect them to attain precisions at least as good as those derived below for OATSSs. At the same time, quantum advantage with OATSSs in magnetometry has been demonstrated in the experiment of Ref. [9] using a Bose-Einstein condensate, and the generation of OATSSs using nitrogen-vacancy centers in diamond has been studied [51].

OATSSs are a particular kind of spin-squeezed states first introduced by Kitagawa and Ueda [52]. They can be produced by first preparing atoms in a CSS along one direction, and then applying an evolution with a Hamiltonian quadratic in one of the perpendicular spin components. For example, for spin-1/2 particles, one can start from an eigenstate of \hat{J}_x with eigenvalue $N/2$ (all spins aligned along x) and apply an evolution with a Hamiltonian proportional to \hat{J}_z^2 . This will generate a state with minimum uncertainty at an angle to both the y and z axes, which depends on the strength of the evolution. The state can then be rotated to align the direction of minimum uncertainty with one of the axes.

For scenarios (a) and (b) of Fig. 1, we consider two cases where the initial CSS is along either x or y , and the collective-spin component with minimum uncertainty is \hat{J}_y or \hat{J}_x , respectively. For scenario (a), the mean values of the collective spin are [52]

$$\langle \hat{J}_x \rangle_0 = \frac{N}{2} \cos^{N-1} \frac{\mu}{2}, \quad \langle \hat{J}_y \rangle_0 = \langle \hat{J}_z \rangle_0 = 0, \quad (22)$$

whereas the variances read

$$\begin{aligned}
(\Delta^2 \hat{J}_x)_0 &= \frac{N}{4} \left[N \left(1 - \cos^{2(N-1)} \frac{\mu}{2} \right) - \frac{1}{2} (N-1) A \right], \\
(\Delta^2 \hat{J}_y)_0 &= \frac{N}{4} \left[1 + \frac{1}{4} (N-1) [A - \sqrt{A^2 + B^2}] \right], \\
(\Delta^2 \hat{J}_z)_0 &= \frac{N}{4} \left[1 + \frac{1}{4} (N-1) [A + \sqrt{A^2 + B^2}] \right], \quad (23)
\end{aligned}$$

with μ being the squeezing parameter, $A = 1 - \cos^{N-2} \mu$, and $B = 4 \sin(\mu/2) \cos^{N-2}(\mu/2)$. We note that the covariance $[\text{cov}(\hat{J}_x, \hat{J}_y)]_0 = 0$ vanishes for this state. The equivalents of Eqs. (22) and (23) for scenario (b) are obtained by interchanging $x \leftrightarrow y$.

C. Mean-squared-error scaling under transversal noise

The mean-squared errors of estimation, which are achieved in scenarios (a) and (b) of Fig. 1, can be calculated by using Eqs. (22), (23), (19), (21) (and the equivalents for \hat{J}_y) and substituting into Eq. (10). The best precision is then obtained by optimizing the evolution time t and the squeezing μ for each N . The general expressions are rather involved, and we have been able to obtain their minima only numerically. However, any explicit choice of $t(N)$ and $\mu(N)$ provides a precision that is guaranteed to be attainable.

Specifically, for scenario (b) a choice that appears to be nearly optimal is $\mu = (\gamma/\omega)^{1/4} (N/4)^{-4/5}$ and $t = (\gamma\omega)^{-1/2} N^{-1/8}$ [53]. For this choice, we expand Eq. (10) in $1/N$ to find the expression for the asymptotic mean-squared error

$$\Delta^2 \omega_{(b)} T \underset{N \rightarrow \infty}{=} \frac{2\omega}{3} \frac{1}{N^{5/4}}. \quad (24)$$

Since the scaling is better than the $1/N$ of the SQL, this demonstrates that superclassical precision scaling is indeed possible with spin-squeezed states and Ramsey-type measurements in the presence of transversal noise.

The possibility for a large quantum enhancement depends strongly on the geometry. We can see this by comparing with scenario (a). There, we find that

$$\Delta^2 \omega_{(a)} T \underset{N \rightarrow \infty}{=} \frac{2\gamma}{N}, \quad (25)$$

which coincides with the best achievable precision for the parallel-noise setting [24] constrained by Eq. (5). As discussed in Appendix E, we find an analytical proof of Eq. (25) in the limit $\omega \rightarrow 0$, and strong numerical evidence for arbitrary ω , which indicates very clearly that no better precision can be achieved. The value is attainable by any choice of $\mu \propto 1/N^{s/(s+1)}$ and $t \propto 1/N^s$ with $s > 1$. Thus, for this geometry, under transversal noise only SQL-like scaling is possible and the quantum enhancement over classical, nonentangled strategies is bounded, while for

scenario (b), the quantum enhancement is unbounded with increasing N .

One can understand intuitively why scenario (b) is more robust than (a) from the pictures of the collective spin in Fig. 1. In both cases the dephasing is directed along x and so causes random rotations around the x axis. The effect of such rotations on the state in (a) is to smear the squeezed state into a more circular distribution at the pole. This reduces the sensitivity of a \hat{J}_y measurement to a small rotation around the z axis, which is the signal we want to detect. In scenario (b), the effect of x rotations is to smear the state along the equator. However, this does not affect the sensitivity of the \hat{J}_x measurement to a z rotation as strongly. Using the same picture, one can also understand why there is a finite optimal value of the squeezing parameter μ . When μ becomes large, the ellipse starts to stretch around the ball. Referring to scenario (b), a z rotation will then increase the projection of the state onto the x axis and hence the \hat{J}_x measurement loses sensitivity.

The difference between the two geometries shows up only for quantum strategies, that is, when squeezed states are employed. If the initial states are not squeezed but are simply CSS states along x for scenario (a) or y for scenario (b), then the precision takes the same form in both cases:

$$\Delta^2 \omega_{\text{CSS}} T = \frac{2t\gamma^2 \Gamma^2 \bar{\Gamma}^4 e^{-\gamma t \bar{\Gamma}} [1 - \cosh(\gamma t \bar{\Gamma}) + 2e^{\gamma t (\frac{\bar{\Gamma}}{\Gamma})^2}] \frac{1}{N}}{[2(1 - e^{-\gamma t \bar{\Gamma}}) - \gamma t \Gamma^2 \bar{\Gamma} (1 + e^{-\gamma t \bar{\Gamma}})]^2} \quad (26)$$

where $\Gamma = 2\omega/\gamma$ and $\bar{\Gamma} = \sqrt{1 - \Gamma^2}$. Thus, we can benchmark the quantum enhancement in either scenario against this classical value. In particular, we consider the numbers from Ref. [3]. In this experiment, $N \approx 10^{11}$, $T_2 \approx 30$ ms, $\kappa \approx 10^{10}$ (Ts) $^{-1}$, and $B_{\text{rf}} \approx 36$ fT, which gives $\gamma = 2/T_2 \approx 67$ Hz and $\omega = 2\kappa B_{\text{rf}} \approx 3.6 \times 10^{-3}$ Hz, and the measurement time is $t \approx 1$ ms. The experiment is not performed with OATSSs, but we compute the quantum enhancements that OATSSs would provide. We insert the numbers in the full expressions [from Eq. (10)] for $\Delta^2 \omega_{(a)} T$ and $\Delta^2 \omega_{(b)} T$ and vary the squeezing, which we quantify in dB [54]; see Fig. 2(a). The precision in scenario (a) saturates with increasing squeezing, and the best quantum enhancement attainable is $\Delta^2 \omega_{\text{CSS}} / \Delta^2 \omega_{(a)} \approx 8$. In scenario (b), on the other hand, the enhancement can reach $\Delta^2 \omega_{\text{CSS}} / \Delta^2 \omega_{(b)} \approx 2 \times 10^7$, corresponding to a 4500-fold improvement in precision. In this case, the precision does not saturate but is limited by the maximal squeezing attainable by the OATSS. This underlines the advantage offered by geometry. However, these maximal enhancements require rather prohibitive squeezings of $\gtrsim -18$ and -73 dB, respectively. If we restrict the squeezing to at most -8 dB as discussed in the outlook of Ref. [3], then scenario (b) provides an enhancement of $\Delta^2 \omega_{\text{CSS}} / \Delta^2 \omega_{(b)} \approx 6.4$,

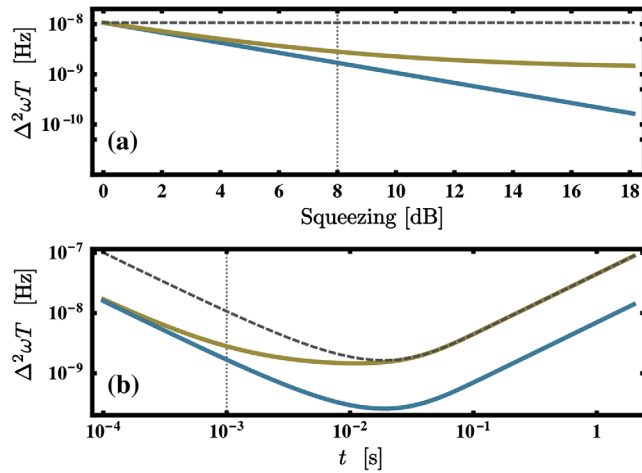


FIG. 2. Mean-squared error of estimation as a function of (a) squeezing and (b) interrogation time, for parameters corresponding to Ref. [3], $N = 10^{11}$, $\gamma = 67$ Hz, $\omega = 3.6 \times 10^{-3}$ Hz. In (a) the interrogation time is fixed at $t = 1$ ms, while in (b) the squeezing is fixed at -8 dB (indicated by the dotted lines, these values correspond to the numbers discussed in the text). The results for scenario Fig. 1(a) (yellow line), scenario Fig. 1(b) (blue line), and for CSS (dashed gray line) are shown.

corresponding to a factor of 2.5 in precision, while scenario (a) for the same numbers gives $\Delta^2\omega_{\text{CSS}}/\Delta^2\omega_{(a)} \approx 3.8$, corresponding to a factor of 1.95. However, the behavior in the two scenarios for varying interrogation time is very different, as explicitly shown in Fig. 2(b). The performance is similar at short times, but when the interrogation time is increased, the quantum advantage in scenario (a) is lost, while the advantage in scenario (b) is maintained. This is a nice feature of the optimized geometry, as in practice experimental constraints may impose a lower limit on t . As seen from Fig. 2(b), at a squeezing of -8 dB, the best precision in scenario (a) (optimizing t) is not significantly below the best precision attainable with a CSS, while the advantage of scenario (b) importantly remains even when experimental constraints do not allow for arbitrarily small t .

We note that, as for Eq. (7), the error Eq. (24) vanishes as $\omega \rightarrow 0$. We refer the reader to Appendix D for a discussion of this limit.

D. Nontransversal noise sources

In a realistic implementation, in addition to the dominant transversal noise, other sources of decoherence will be present. For example, in some setups different from Ref. [3], e.g., spin-exchange relaxation-free magnetometers [55,56], spin relaxation cannot be neglected. Typically [35,36], this is modeled as uncorrelated depolarizing noise [$\alpha_x = \alpha_y = \alpha_z = 1/3$ in Eq. (4)] with a strength dictated by the spin-relaxation time T_1 (see Appendix A). As opposed to the directional dephasing caused by spin decoherence, such noise is isotropic and always yields a parallel-noise

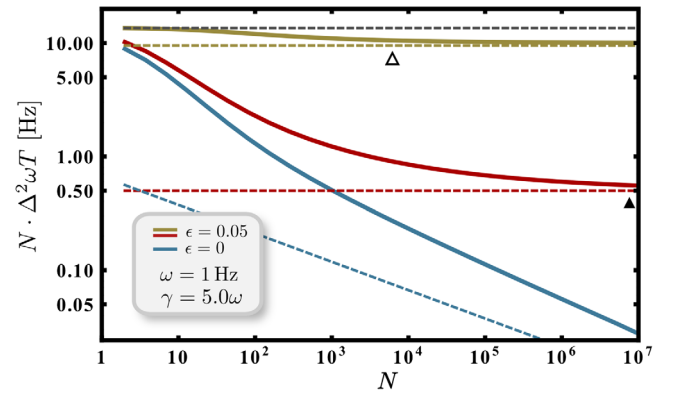


FIG. 3. The mean-squared error of estimation rescaled by the particle number such that SQL-like scaling is horizontal. The curves for scenario (a) (yellow line) and (b) (red line) with 5% of parallel noise are shown, and for scenario (b) under purely transversal noise (blue line), along with their asymptotes (dashed line), respectively, $2\gamma(1-\epsilon)N$ from (25), $2\gamma\epsilon$ from the parallel noise component, and $2\omega/3N^{5/4}$ from Eq. (24). We also show the performance Eq. (26) of a CSS without squeezing (gray dashed line). The locations where the OATSS strategies reach 90% of their asymptotic gain over this CSS strategy are indicated for scenarios (a) (open triangle) and (b) (closed triangle).

component independently of the geometry. In the case of Ref. [3], any spatial misalignment between B , B_{rf} , and the direction of squeezing, temporal mismatch between the B_{rf} rotations, and the Larmor frequency of B , or violation of the condition $B_{\text{rf}} \gg B$, may be phenomenologically included into such a depolarization model.

We assess the effect of such additional noise sources by considering a deviation from perfect transversality. In particular, we take a small component of dephasing directed along the z axis, such that $\alpha_x = 1 - \epsilon$ and $\alpha_z = \epsilon$. As discussed in Ref. [27], once any such parallel-dephasing contribution is present, the asymptotic scaling must return to its SQL-like behavior; that is,

$$\Delta^2\omega T \geq \frac{c_{xz}(\gamma, \epsilon)}{N}, \quad (27)$$

with $c_{xz}(\gamma, \epsilon)$ lower bounded by the minimum of Eq. (5), i.e., $c_z(\epsilon\gamma, \omega, t) \geq 2\epsilon\gamma$. For instance, for the depolarization-based spin-relaxation model, $c_{xz}(\gamma, \epsilon) \geq 8/(3T_1)$ (see Appendix A). We illustrate the resulting crossover behavior in Fig. 3.

Although the asymptotic scaling is now again SQL-like, geometry can strongly influence the achievable quantum gain and the effective N at which the crossover to SQL-like scaling happens. In Fig. 3, we show the mean-squared-error scalings attained in scenarios (a) and (b) using, respectively, OATSSs along x and y and measurements of \hat{J}_y and \hat{J}_x and compare them to a strategy without entanglement, simply using a CSS along x and measurement of \hat{J}_y , corresponding to the nonentangled strategy implemented in Ref. [3]. We

see that while the strategy in (b) can saturate the bound $2\gamma\epsilon/N$, the strategy in (a) only reaches $2\gamma(1-\epsilon)/N$ as imposed by Eq. (25). Thus, the mean-squared error of geometry (b) is a factor $\epsilon/(1-\epsilon)$ lower than (a), which may be significant when the noise is dominantly transversal. Furthermore, superclassical scaling persists over a larger range of N in geometry (b). In the Fig. 3, the locations where the OATSS strategies for (a) and (b) reach 90% of their asymptotic gain over the nonentangled CSS strategy are indicated. Clearly, the crossover happens at much larger N in scenario (b). As $\epsilon \rightarrow 0$, the crossover must go to infinity. To get an idea of the behavior, we can take the N at which the asymptotic bound $2\epsilon\gamma/N$ crosses the asymptote Eq. (24) for perfectly transversal noise. This intersection scales as $(\omega/\gamma)^4/\epsilon^4$. Thus, significant gain in precision by squeezing is attained over a larger range of N if the geometry is chosen correctly.

V. CONCLUSION AND OUTLOOK

For quantum metrology to be relevant in practical situations, it is important that good performance can be attained under realistic noise with states and measurements that are amenable to implementation in the laboratory. While recent results have shown that for many noise types precision scaling can only improve over the classical limit by a constant, here we demonstrate that under transversal dephasing, superclassical scaling can be preserved with experimentally accessible states and measurements, and we argue that this noise model is relevant to recent atomic magnetometry experiments. We show that the choice of geometry is important for the attainable quantum improvement both asymptotically and for parameter settings corresponding to recent experiments. Furthermore, we assess the robustness of the model to other nontransversal sources of noise and find that quantum enhancement could still be achieved for atomic ensembles of macroscopic size with an adequate choice of geometry.

Our results give a clear message that quantum-enhanced metrology maintains its relevance even in the presence of noise, and we hope that they will encourage the search for other practically motivated scenarios where quantum strategies provide an advantage. For instance, it has been suggested that the transversal-noise model applies also to nitrogen-vacancy centers in diamonds [28], and that one-axis-twisted spin-squeezed states could be prepared in such systems [51]. Very recently, a noise-robust magnetometry scheme employing SQUID junctions has been proposed [57]. Finally, in the Appendix D, we speculate about the potential of adaptive techniques that bias the estimated parameter towards the zero value, for which our current precision bounds fail. We expect that the question of what happens in this limit could be consistently resolved by employing Bayesian techniques [22,33,42], which account explicitly for the prior knowledge about the estimated parameter.

ACKNOWLEDGMENTS

We are grateful to Antonio Acín, Stefan Christensen, Rafał Demkowicz-Dobrzański, Marcin Jarzyna, and Wojciech Wasilewski for valuable feedback and discussions. J.B.B. acknowledges support from the Swiss National Science Foundation (QSIT director's reserve), SEFRI (COST action MP1006), and FP7 IP project SIQS. R. C. acknowledges support from the Excellence Initiative of the German Federal and State Governments (Grants No. ZUK 43 and ZUK 81), the Research Innovation Fund from the University of Freiburg, and the U.S. Army Research Office under Contracts No. W911NF-14-1-0098 and No. W911NF-14-1-0133 (Quantum Characterization, Verification, and Validation). J. K. acknowledges support from the FP7 IP project SIQS cofinanced by the Polish Ministry of Science and Higher Education, START scholarship granted by Foundation for Polish Science, as well as the ERC Consolidator Grant QITBOX, and EU MSCA Individual Fellowship Q-METAPP.

APPENDIX A: ATOMIC MASTER EQUATION

Following Ref. [3], we assume correlated effects in the atomic ensemble to play a role only at time scales longer than the interrogation time, i.e., the ensemble relaxation and decoherence times $T_1^*, T_2^* \gg t$ [50] (in Ref. [3], $t \approx 1$ ms). We may then describe the dynamics by a master equation where each atom evolves independently,

$$\frac{\partial \rho^N(t)}{\partial t} = \sum_{n=1}^N \mathcal{D}[\rho^N(t)], \quad (\text{A1})$$

where \mathcal{D} generates the evolution for a single atom (assumed to be the same for all atoms) and reads

$$\mathcal{D}[\rho(t)] = -i \frac{\Omega_B}{2} [\hat{\sigma}_x, \rho(t)] - i \frac{\omega}{2} [\hat{\sigma}_{yz}(t), \rho(t)] \quad (\text{A2})$$

$$+ \frac{1}{T_2} [\hat{\sigma}_x \rho(t) \hat{\sigma}_x - \rho(t)] \quad (\text{A3})$$

$$+ \frac{1}{T_1} \left(\frac{1}{3} \sum_{i=1}^3 \hat{\sigma}_i \rho(t) \hat{\sigma}_i - \rho(t) \right). \quad (\text{A4})$$

The terms (A2) represent the Hamiltonian part of the dynamics: $\Omega_B = 2\kappa B$ is the Larmor frequency of the strong field B (see Fig. 1), whereas $\omega = 2\kappa B_{\text{rf}}$ is the frequency corresponding to the weak field B_{rf} being sensed. Having Ref. [3] in mind, we allow B_{rf} to rotate in the yz plane, $\hat{\sigma}_{yz}(t) = \cos \theta_t \hat{\sigma}_z - \sin \theta_t \hat{\sigma}_y$.

The term (A3) represents dephasing. It is a consequence of B -field fluctuations independently affecting each atom (e.g., arising from the atomic motion and/or local field inhomogeneities). Equivalently, it may be derived by

considering an average description of a noiseless evolution but with Ω_B fluctuating according to a Gaussian distribution with variance $4/(T_2 t)$ (note that the variance diverges in the $t \rightarrow 0$ limit manifesting the Markovianity of the noise). Such fluctuations define the spin-decoherence time T_2 [50] and constitute the transversal [27] noise, which is the main focus of this paper.

The third term Eq. (A4) represents the spin-relaxation process occurring predominantly due to spin-destruction or spin-exchange atomic collisions [35] (the latter may be eliminated in the accordingly called spin-exchange relaxation-free magnetometers [55,56]). Note that Eq. (A4) effectively yields a depolarizing channel [as defined in Eq. (4)] with strength parametrized by the spin-relaxation time T_1 [36,50].

By moving to the rotating frame precessing with the Larmor frequency, in which $\rho_{\text{RF}}(t) = e^{(i\Omega_B t/2)\hat{\sigma}_x} \rho(t) e^{-(i\Omega_B t/2)\hat{\sigma}_x}$, and by using $e^{(i\theta/2)\hat{\sigma}_x} \hat{\sigma}_{y/z} e^{-(i\theta/2)\hat{\sigma}_x} = \cos \theta \hat{\sigma}_{y/z} \mp \sin \theta \hat{\sigma}_{y/z}$, we obtain the single-atom master equation in the RF as

$$\begin{aligned} \frac{\partial \rho_{\text{RF}}(t)}{\partial t} = & -i \frac{\omega}{2} [\sin \delta \theta_t \hat{\sigma}_y + \cos \delta \theta_t \hat{\sigma}_z, \rho_{\text{RF}}(t)] \\ & + \frac{1}{T_2} [\hat{\sigma}_x \rho_{\text{RF}}(t) \hat{\sigma}_x - \rho_{\text{RF}}(t)] \\ & + \frac{1}{T_1} \left(\frac{1}{3} \sum_{i=1}^3 \hat{\sigma}_i \rho_{\text{RF}}(t) \hat{\sigma}_i - \rho_{\text{RF}}(t) \right), \end{aligned} \quad (\text{A5})$$

with $\delta \theta_t = (\Omega_B t/2) - \theta_t$. As a consequence, if the Larmor frequency is chosen to exactly match the rotation of B_{rf} , i.e., $\theta_t = \Omega_B t/2$, and if we consider the limit $T_1 \rightarrow \infty$ in which the spin relaxation is completely ignored (as in Ref. [3]), then we indeed recover Eq. (8) in the main text.

Note that Eq. (A5) indicates that by *not* exactly matching θ_t with the Larmor frequency we nonetheless preserve the required transversal geometry between the dephasing noise Eq. (A3) and the ω -encoding part, so that the perpendicular-noise model Eq. (4) of Ref. [27], in principle, still applies. However, let us emphasize that for our analysis of squeezed states to be valid, the geometry of squeezing must always be adjusted for a particular choice of B_{rf} direction in the RF, e.g., depicted in Fig. 1 for B_{rf} chosen to be stationary and pointing along z in the RF. In an experiment, we expect there to be some mismatch between the B_{rf} -rotation frequency and Ω_B , which would constitute an extra source of *global* dephasing noise along x , i.e., decreasing the ensemble spin-decoherence time T_2^* neglected in Ref. [3]. In contrast, any such fluctuations felt locally by atoms would just lower T_2 in Eq. (A3), which would thus not impair the robustness of the setup.

On the other hand, any geometrical misalignment of the fields or any instabilities of θ_t can always be modeled in Eq. (A5) by lowering the effective spin-relaxation time T_1 . Importantly, the spin-relaxation

process may be interpreted as a nontransversal-noise source, which we study in Sec. IV D. Similarly to the parallel dephasing case [$\alpha_z = 1$ in Eq. (4)], which is known to asymptotically restrict the precision to $2\gamma/N$ [24], it has been shown [32] that depolarizing noise [$\alpha_x = \alpha_y = \alpha_z = 1/3$ in Eq. (4)] yields a SQL-like asymptotic precision $(4\gamma/3)/N$, thus giving just a 2/3-factor improvement as compared to the pure parallel-noise model.

A finite spin-relaxation time must thus impose an asymptotic SQL-like behavior, which we may estimate by taking $3/2$ of T_1 in Eq. (A5) to contribute solely to the parallel-dephasing component. Then, after substituting for $\epsilon = 2T_2/(3T_1 + 2T_2)$ and $\gamma = 2(3T_1 + 2T_2)/(3T_1 T_2)$, we can directly utilize Eq. (27) to obtain an asymptotic bound with $2\gamma\epsilon = 8/(3T_1)$ dictated by the spin relaxation. Crucially, the analysis of Sec. IV D thus shows that in the regime of $T_1 \gg T_2$, the optimal geometry of scenario (b) not only allows the counterbalance of the spin-decoherence effects, but also postpones the inevitable SQL-bounding impact of spin relaxation to much higher N (see the triangular marks in Fig. 3).

APPENDIX B: KRAUS OPERATORS

The map corresponding to evolution under the master equation (2) during time t can be written as a composite map of the form $\mathcal{E}_\omega^{\otimes N}$. Following Andersson *et al.* [58], the single-qubit maps are then given by

$$\mathcal{E}_\omega(\rho) = \sum_{i,j=0}^3 S_{ij} \tilde{\sigma}_i \rho \tilde{\sigma}_j, \quad (\text{B1})$$

where $\tilde{\sigma}_i$ are the normalized Pauli operators $\tilde{\sigma}_i = \hat{\sigma}_i/\sqrt{2}$ and $\hat{\sigma}_0$ denotes the identity. All elements of the matrix S are zero, except $S_{00} = A_+ + B_+$, $S_{11} = A_- + (\Gamma/\tilde{\alpha})B_-$, $S_{22} = A_- - (\Gamma/\tilde{\alpha})B_-$, $S_{33} = A_+ - B_+$, $S_{03} = i(\alpha_-/\tilde{\alpha})B_-$, $S_{01} = -i(\alpha_-/\tilde{\alpha})B_-$. Where we have defined $\Gamma = 2\omega/\gamma$, $\alpha_\pm = \alpha_x \pm \alpha_y$, and $\tilde{\alpha} = \sqrt{\alpha_-^2 - \Gamma^2}$, and the coefficients

$$A_\pm = \frac{1}{2} (1 \pm e^{-\gamma t \alpha_\pm}), \quad (\text{B2})$$

$$B_\pm = \frac{1}{2} e^{-(\gamma t/2)(1+\alpha_z-\tilde{\alpha})} (1 \pm e^{-\gamma t \tilde{\alpha}}). \quad (\text{B3})$$

A Kraus representation of the map \mathcal{E}_ω can be obtained by diagonalizing the matrix S . Denoting the eigenvalues and normalized eigenvectors of S by λ_i and \mathbf{v}_i , respectively, one can find a valid set of Kraus operators for the channel:

$$K_j = \sum_{i=1}^4 \sqrt{|\lambda_i|} (\mathbf{v}_i)_j \tilde{\sigma}_{j-1}, \quad (\text{B4})$$

with $j = 1, \dots, 4$, which gives the set in Eq. (12). The coefficients in Eq. (12) are rather involved, and we do not

explicitly state them here. Instead, we directly give the expressions for ξ_x , χ_x , ξ_y , and χ_y of Eqs. (15) and (16). For general noise, they read

$$\begin{aligned}\xi_x &= e^{-(\gamma t/2)(1+\alpha_z)} \left[\cosh\left(\frac{\gamma t}{2}\tilde{\alpha}\right) + \frac{\alpha_-}{\tilde{\alpha}} \sinh\left(\frac{\gamma t}{2}\tilde{\alpha}\right) \right], \\ \chi_y &= -e^{-(\gamma t/2)(1+\alpha_z)} \frac{\Gamma}{\tilde{\alpha}} \sinh\left(\frac{\gamma t}{2}\tilde{\alpha}\right),\end{aligned}\quad (\text{B5})$$

and

$$\begin{aligned}\xi_y &= e^{-(\gamma t/2)(1+\alpha_z)} \left[\cosh\left(\frac{\gamma t}{2}\tilde{\alpha}\right) - \frac{\alpha_-}{\tilde{\alpha}} \sinh\left(\frac{\gamma t}{2}\tilde{\alpha}\right) \right], \\ \chi_x &= e^{-(\gamma t/2)(1+\alpha_z)} \frac{\Gamma}{\tilde{\alpha}} \sinh\left(\frac{\gamma t}{2}\tilde{\alpha}\right).\end{aligned}\quad (\text{B6})$$

In the case of perfectly transversal noise, they further simplify, since $\alpha_x = 1$ implies $\alpha_z = 0$ and $\tilde{\alpha} = \sqrt{1 - \Gamma^2}$.

APPENDIX C: ANALYTICAL SCALING FOR GHZ STATES

Stemming from the error-propagation method [see Eq. (10)] utilized in the main text, we can also confirm the results of Ref. [27] analytically for the GHZ input states,

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0, \dots, 0\rangle + |1, \dots, 1\rangle), \quad (\text{C1})$$

by considering the parity operator in the x direction,

$$\hat{P}_x = \otimes_{k=1}^N \hat{\sigma}_x^{(k)}, \quad (\text{C2})$$

as the observable being measured. Similarly to the case of collective-spin operators and Eq. (18), we may utilize Eq. (15) to write the form of the parity operator at time t as

$$\hat{P}_x(t) = \sum_s K_s^\dagger \hat{P}_x K_s = \otimes_{k=1}^N (\xi_x \hat{\sigma}_x^{(k)} + \chi_x \hat{\sigma}_y^{(k)}). \quad (\text{C3})$$

In the computational basis $\{|0\rangle, |1\rangle\}^{\otimes N}$, such an operator just flips all of the qubits, and hence only the off-diagonal terms contribute when calculating its expectation value for a GHZ state Eq. (C1). Every $\hat{\sigma}_x$ contributes a factor of 1 while $\hat{\sigma}_y$ contributes a factor of $\pm i$. Thus, the expectation value of the measurement becomes

$$\langle \hat{P}_x \rangle_{\text{GHZ}, t} = \frac{1}{2} [(\xi_x + i\chi_x)^N + (\xi_x - i\chi_x)^N], \quad (\text{C4})$$

and, since $\hat{P}_x^2 = 1$, it follows that $\Delta^2 \hat{P}_x = 1 - \langle \hat{P}_x \rangle^2$.

We compute the mean-squared error of estimation via Eq. (10), after setting the interrogation time to $t = (3/\gamma\omega^2 N)^{1/3}$, as was found in Ref. [27] from numerical

analysis. Expanding the corresponding $\Delta^2 \omega T$ in $1/N$, we find the asymptotic scaling to read:

$$\begin{aligned}\Delta^2 \omega T &\stackrel{N \rightarrow \infty}{=} g(\gamma, \omega, N) (\gamma \omega^2)^{1/3} \frac{1}{N^{5/3}} \\ &\geq \frac{e^2}{3^{1/3}} (\gamma \omega^2)^{1/3} \frac{1}{N^{5/3}},\end{aligned}\quad (\text{C5})$$

where $g(\gamma, \omega, N)$ represents oscillating terms that are lower bounded by $e^2/3^{1/3}$. The constant prefactor here is larger than the prefactor $3^{2/3}/2$, which was numerically verified to be optimal—optimized over all possible measurements—for GHZ states [27]. Nevertheless, although this suggests that either parity measurement is suboptimal or the above interrogation time t dependence should be improved in the parity-based scenario, Eq. (C5) suffices to prove the superclassical precision scaling, $1/N^{5/3}$, as well as the $(\gamma \omega^2)^{1/3}$ behavior of the asymptotic coefficient.

APPENDIX D: NOTE ON VANISHING PARAMETER VALUE

For $\omega = 0$, both the GHZ-achievable bound Eq. (7) and the OATSS-based expression Eq. (24) vanish. This does not mean that the precision is unbounded for the two cases, but rather suggests that the results give no information in such a limit. It is therefore not clear what precision scaling can then be achieved.

In general, for the channel Eq. (B1) at $\omega = 0$, we get $\xi = 1$, $\chi = 0$, and $\partial \xi / \partial \omega = 0$, $\partial \chi / \partial \omega = (e^{-t\gamma} - 1)/\gamma$. For a GHZ state Eq. (C1) and parity measurement Eq. (C2), one can show utilizing Eq. (10) that

$$\Delta^2 \omega T \stackrel{N \rightarrow \infty}{=} \frac{t\gamma^2}{(1 - e^{-t\gamma})^2} \frac{1}{N^2}, \quad (\text{D1})$$

for fixed t . This is minimized at $t_{\text{opt}} = \kappa/\gamma$, where κ is a numerical constant. Similarly, at $\omega = 0$ for an OATSS along y squeezed in x [as in scenario (b) of the main text] with squeezing parameter $\mu = (N/4)^{-2/3}$, one finds

$$\Delta^2 \omega T \stackrel{N \rightarrow \infty}{=} \frac{5}{3 \times 2^{2/3}} \frac{t\gamma^2}{(1 - e^{-t\gamma})^2} \frac{1}{N^{5/3}}, \quad (\text{D2})$$

and again the optimal time is $t_{\text{opt}} = \kappa/\gamma$.

Thus, on the one hand, the local estimation approach that we employ above indicates that an improved scaling, even reaching the Heisenberg limit for GHZ states, is possible at the special $\omega = 0$ parameter value. Even if the value of ω is *a priori* nonzero, one might then think that the precision scaling can be improved by adopting an iterative, adaptive strategy [59–61]. By applying a bias (e.g., in the case of magnetometry, a magnetic field in opposite direction to the estimated field) to decrease the parameter after obtaining its first estimate, a better estimate is obtained with a precision that is less heavily constrained by bounds Eqs. (7) and (24),

due to the lower effective value of ω . On the other hand, the prior information on ω required to adjust the bias may scale prohibitively. We can compute the estimated mean-squared error for GHZ states and parity measurements (see above) and expand it in ω to obtain $\Delta^2\omega T \approx c(\gamma, t)/N^2 + O(\omega^2)$, with c given by Eq. (D1). For the Heisenberg-scaling term to dominate for a fixed t , the higher-order terms in the expansion must be negligible in comparison. However, we find (for even terms, odd terms vanish) that the k th term scales as $\omega^k N^{k-3}$, which implies that we need $\omega \ll N^{-(k-1)/k}$ to neglect the higher-order terms. For this to hold for all k , $\omega \ll 1/N$, which means that the prior information on ω must already be Heisenberg limited, as in the case of a decoherence-free local estimation scenario [22,33]. At first sight, this may indicate that such an adaptive scheme may not be successful for any prior distribution of finite width, and that the value of ω must be perfectly known and set to zero for the above improved scalings to be observed. However, recent results [42], based on the Bayesian approach to estimation, indicate that in the decoherence-free case the Heisenberg scaling is attained irrespectively of the prior knowledge of ω . Hence, we expect the transversal-noise model to behave similarly due to its decoherence-free-like regime at short interrogation times, which would then prove the above adaptive strategy to also be efficient.

APPENDIX E: BOUND FOR SCENARIO (a)

We give an analytical proof of Eq. (25) in the limit $\omega \rightarrow 0$ and strong numerical evidence for arbitrary ω .

Following the method outlined at the beginning of Sec. IV C, we find expressions for $\Delta^2\omega_{(a)}T$ and $\Delta^2\omega_{(b)}T$. We are looking for a lower bound on the former. From the expressions, it can be seen that, in the case $\omega = 0$,

$$(\Delta^2\omega_{(a)}T) - e^{-2t\gamma}(\Delta^2\omega_{(b)}T) = \frac{\gamma^2 t \coth(\frac{\gamma}{2}) \cos(\frac{\mu}{2})^{2-2N}}{N}. \quad (\text{E1})$$

Now, since both the prefactor $e^{-2t\gamma}$ and $\Delta^2\omega_{(b)}T$ are positive, any lower bound on this quantity is also a lower bound on $\Delta^2\omega_{(a)}T$. The t -dependent factor $t \coth(\gamma t/2)$ is lower bounded by $2/\gamma$ (attained when $t \rightarrow 0$), while $\cos(\mu/2)^{2-2N}$ is lower bounded by 1. It follows that $(\Delta^2\omega_{(a)}T)$ is lower bounded by $2\gamma/N$ as desired.

When $\omega > 0$, the expressions for $\Delta^2\omega_{(a)}T$ and $\Delta^2\omega_{(b)}T$ become significantly more complicated. However, we can again look at a quantity

$$M = (\Delta^2\omega_{(a)}T) - \left(\frac{A-B}{A+B}\right)^2 (\Delta^2\omega_{(b)}T), \quad (\text{E2})$$

where

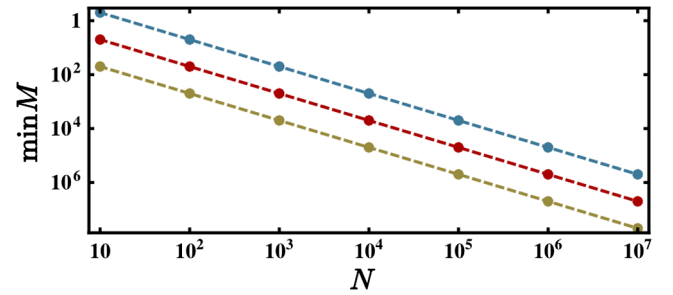


FIG. 4. Circles: Results of minimizing M in Eq. (E2) over μ, t versus N for $\{\gamma, \omega\} = \{10, 0.03\}$, $\{1.0, 0.3\}$, and $\{0.1, 0.03\}$ (top to bottom). Dashed lines: Plots of $2\gamma/N$ for the same values of γ .

$$A = \cosh\left(\frac{1}{2}t\sqrt{\gamma^2 - 4\omega^2}\right), \quad (\text{E3})$$

$$B = \gamma \sinh\left(\frac{1}{2}t\sqrt{\gamma^2 - 4\omega^2}\right) / \sqrt{\gamma^2 - 4\omega^2}. \quad (\text{E4})$$

Since $(A-B)^2/(A+B)^2$ and $\Delta^2\omega_{(b)}T$ are positive, a lower bound on M is again also a lower bound on $\Delta^2\omega_{(a)}T$. We have not been able to prove an analytical bound for M , but for given values of ω, γ, N we can numerically minimize M over μ and t . In Fig. 4, we plot the results of several such minimizations. As can be seen, the numerics give very clear evidence that $\min M = 2\gamma/N$ and hence $\Delta^2\omega_{(a)}T$ is lower bounded by $2\gamma/N$ as claimed.

-
- [1] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, *Science* **306**, 1330 (2004).
 - [2] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, *Phys. Rev. Lett.* **96**, 010401 (2006).
 - [3] W. Wasilewski, K. Jensen, H. Krauter, J. J. Renema, M. V. Balabas, and E. S. Polzik, *Quantum Noise Limited and Entanglement-Assisted Magnetometry*, *Phys. Rev. Lett.* **104**, 133601 (2010).
 - [4] M. Koschorreck, M. Napolitano, B. Dubost, and M. W. Mitchell, *Sub-Projection-Noise Sensitivity in Broadband Atomic Magnetometry*, *Phys. Rev. Lett.* **104**, 093602 (2010).
 - [5] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell, *Magnetic Sensitivity beyond the Projection Noise Limit by Spin Squeezing*, *Phys. Rev. Lett.* **109**, 253605 (2012).
 - [6] C. F. Ockeloen, R. Schmied, M. F. Riedel, and P. Treutlein, *Quantum Metrology with a Scanning Probe Atom Interferometer*, *Phys. Rev. Lett.* **111**, 143001 (2013).
 - [7] D. Sheng, S. Li, N. Dural, and M. V. Romalis, *Subfemtotesla Scalar Atomic Magnetometry Using Multipass Cells*, *Phys. Rev. Lett.* **110**, 160802 (2013).
 - [8] V. G. Lucivero, P. Anielski, W. Gawlik, and M. W. Mitchell, *Shot-Noise-Limited Magnetometer with Sub-Picotesla*

- Sensitivity at Room Temperature*, *Rev. Sci. Instrum.* **85**, 113108 (2014).
- [9] W. Muessel, H. Strobel, D. Linnemann, D. B. Hume, and M. K. Oberthaler, *Scalable Spin Squeezing for Quantum-Enhanced Magnetometry with Bose-Einstein Condensates*, *Phys. Rev. Lett.* **113**, 103004 (2014).
- [10] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, *Experimental Demonstration of Entanglement-Enhanced Rotation Angle Estimation Using Trapped Ions*, *Phys. Rev. Lett.* **86**, 5870 (2001).
- [11] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, *Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States*, *Science* **304**, 1476 (2004).
- [12] J. Appel, P. J. Windpassinger, D. Oblak, U. B. Hoff, N. Kjærgaard, and E. S. Polzik, *Mesoscopic Atomic Entanglement for Precision Measurements beyond the Standard Quantum Limit*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10960 (2009).
- [13] A. Louchet-Chauvet, J. Appel, J. J. Renema, D. Oblak, N. Kjærgaard, and E. S. Polzik, *Entanglement-Assisted Atomic Clock beyond the Projection Noise Limit*, *New J. Phys.* **12**, 065032 (2010).
- [14] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, *Improvement of Frequency Standards with Quantum Entanglement*, *Phys. Rev. Lett.* **79**, 3865 (1997).
- [15] D. Ulam-Orgikh and M. Kitagawa, *Spin Squeezing and Decoherence Limit in Ramsey Spectroscopy*, *Phys. Rev. A* **64**, 052106 (2001).
- [16] A. Shaji and C. M. Caves, *Qubit Metrology and Decoherence*, *Phys. Rev. A* **76**, 032111 (2007).
- [17] A. Fujiwara and H. Imai, *A Fibre Bundle over Manifolds of Quantum Channels and Its Application to Quantum Statistics*, *J. Phys. A* **41**, 255304 (2008).
- [18] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying, *Parameter Estimation of Quantum Channels*, *IEEE Trans. Inf. Theory* **54**, 5172 (2008).
- [19] M. Hayashi, *Comparison between the Cramer-Rao and the Mini-Max Approaches in Quantum Channel Estimation*, *Commun. Math. Phys.* **304**, 689 (2011).
- [20] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, *The Elusive Heisenberg Limit in Quantum Enhanced Metrology*, *Nat. Commun.* **3**, 1063 (2012).
- [21] J. Kołodyński and R. Demkowicz-Dobrzański, *Efficient Tools for Quantum Metrology with Uncorrelated Noise*, *New J. Phys.* **15**, 073043 (2013).
- [22] J. Kołodyński, *Precision Bounds in Noisy Quantum Metrology*, Ph.D. thesis, University of Warsaw, 2014.
- [23] R. Demkowicz-Dobrzański and L. Maccone, *Using Entanglement Against Noise in Quantum Metrology*, *Phys. Rev. Lett.* **113**, 250801 (2014).
- [24] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *General Framework for Estimating the Ultimate Precision Limit in Noisy Quantum-Enhanced Metrology*, *Nat. Phys.* **7**, 406 (2011).
- [25] Y. Matsuzaki, S. C. Benjamin, and J. Fitzsimons, *Magnetic Field Sensing beyond the Standard Quantum Limit under the Effect of Decoherence*, *Phys. Rev. A* **84**, 012103 (2011).
- [26] A. W. Chin, S. F. Huelga, and M. B. Plenio, *Quantum Metrology in Non-Markovian Environments*, *Phys. Rev. Lett.* **109**, 233601 (2012).
- [27] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, *Noisy Metrology beyond the Standard Quantum Limit*, *Phys. Rev. Lett.* **111**, 120401 (2013).
- [28] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, *Quantum Error Correction for Metrology*, *Phys. Rev. Lett.* **112**, 150802 (2014).
- [29] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, *Improved Quantum Metrology Using Quantum Error Correction*, *Phys. Rev. Lett.* **112**, 080801 (2014).
- [30] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, *Increasing Sensing Resolution with Error Correction*, *Phys. Rev. Lett.* **112**, 150801 (2014).
- [31] F. Fröwis, M. Skotiniotis, B. Kraus, and W. Dür, *Optimal Quantum States for Frequency Estimation*, *New J. Phys.* **16**, 083010 (2014).
- [32] S. I. Knysh, E. H. Chen, and G. A. Durkin, *True Limits to Precision via Unique Quantum Probe*, [arXiv:1402.0495](https://arxiv.org/abs/1402.0495).
- [33] M. Jarzyna and R. Demkowicz-Dobrzański, *True Precision Limits in Quantum Metrology*, *New J. Phys.* **17**, 013010 (2015).
- [34] M. Auzinsh, D. Budker, D. F. Kimball, S. M. Rochester, J. E. Stalnaker, A. O. Sushkov, and V. V. Yashchuk, *Can a Quantum Nondemolition Measurement Improve the Sensitivity of an Atomic Magnetometer?*, *Phys. Rev. Lett.* **93**, 173002 (2004).
- [35] I. K. Kominis, *Sub-Shot-Noise Magnetometry with a Correlated Spin-Relaxation Dominated Alkali-Metal Vapor*, *Phys. Rev. Lett.* **100**, 073002 (2008).
- [36] L. Liu, S. Cheng, B. Qi, Z. Xi, and D. Dong, *Precision Limit of Atomic Magnetometers in the Presence of Spin-Destruction Collisions*, *J. Phys. B* **48**, 035502 (2015).
- [37] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen, *Spin Squeezing and Reduced Quantum Noise in Spectroscopy*, *Phys. Rev. A* **46**, R6797 (1992).
- [38] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, *Squeezed Atomic States and Projection Noise in Spectroscopy*, *Phys. Rev. A* **50**, 67 (1994).
- [39] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, *Optimal Frequency Measurements with Maximally Correlated States*, *Phys. Rev. A* **54**, R4649 (1996).
- [40] V. Bužek, R. Derka, and S. Massar, *Optimal Quantum Clocks*, *Phys. Rev. Lett.* **82**, 2207 (1999).
- [41] A. Andrè, A. S. Sørensen, and M. D. Lukin, *Stability of Atomic Clocks based on Entangled Atoms*, *Phys. Rev. Lett.* **92**, 230801 (2004).
- [42] K. Macieszczak, R. Demkowicz-Dobrzański, and M. Fraas, *Bayesian Quantum Frequency Estimation in Presence of Collective Dephasing*, *New J. Phys.* **16**, 113002 (2014).
- [43] L. Ostermann, H. Ritsch, and C. Genes, *Protected State Enhanced Quantum Metrology with Interacting Two-Level Ensembles*, *Phys. Rev. Lett.* **111**, 123601 (2013).
- [44] J. Borregaard and A. S. Sørensen, *Near-Heisenberg-Limited Atomic Clocks in the Presence of Decoherence*, *Phys. Rev. Lett.* **111**, 090801 (2013).
- [45] E. M. Kessler, P. Kómár, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin, *Heisenberg-Limited*

- Atom Clocks based on Entangled Qubits*, *Phys. Rev. Lett.* **112**, 190403 (2014).
- [46] A. Acín, B. Bylicka, R. Demkowicz-Dobrzański, and J. Kołodyński (to be published).
- [47] S. L. Braunstein and C. M. Caves, *Statistical Distance and the Geometry of Quantum States*, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [48] D. M. Greenberger, M. A. Horne, and A. Zeilinger, *Bells Theorem, Quantum Theory, and Conceptions of the Universe* (Kluwer Academic Publishers, Dordrecht, 1989).
- [49] J. Ma, X. Wang, C. P. Sun, and F. Nori, *Quantum Spin Squeezing*, *Phys. Rep.* **509**, 89 (2011).
- [50] X. R. Wang, Y. S. Zheng, and S. Yin, *Spin Relaxation and Decoherence of Two-Level Systems*, *Phys. Rev. B* **72**, 121303 (2005).
- [51] S. D. Bennett, N. Y. Yao, J. Otterbach, P. Zoller, P. Rabl, and M. D. Lukin, *Phonon-Induced Spin-Spin Interactions in Diamond Nanostructures: Application to Spin Squeezing*, *Phys. Rev. Lett.* **110**, 156402 (2013).
- [52] M. Kitagawa and M. Ueda, *Squeezed Spin States*, *Phys. Rev. A* **47**, 5138 (1993).
- [53] Although ω is not perfectly known in advance, this choice is not a problem. Replacing $\omega \rightarrow \beta\omega$ in μ, t has the effect $\omega \rightarrow \omega/\beta$ in Eq. (24). Similarly, $\gamma \rightarrow \beta\gamma$ in μ, t leads to $\omega \rightarrow \omega/\beta$ in Eq. (24). The scaling remains unchanged.
- [54] We measure the squeezing by the CSS-relative squeezing parameter ξ_{rel} introduced in Refs. [37,38] and use Eqs. (112) and (139) of Ref. [49].
- [55] M. P. Ledbetter, I. M. Savukov, V. M. Acosta, D. Budker, and M. V. Romalis, *Spin-Exchange-Relaxation-Free Magnetometry with Cs Vapor*, *Phys. Rev. A* **77**, 033408 (2008).
- [56] W. C. Griffith, S. Knappe, and J. Kitching, *Femtotesla Atomic Magnetometry in a Microfabricated Vapor Cell*, *Opt. Express* **18**, 27167 (2010).
- [57] D. A. Herrera-Martí, D. Aharonov, and A. Retzker, *A Quantum Error Correction-Enhanced Magnetometer Overcoming the Limit Imposed by Relaxation*, arXiv:1410.7556.
- [58] E. Andersson, J. D. Cresser, and M. J. W. Hall, *Finding the Kraus Decomposition from a Master Equation and Vice Versa*, *J. Mod. Opt.* **54**, 1695 (2007).
- [59] H. M. Wiseman, D. W. Berry, S. D. Bartlett, B. L. Higgins, and G. J. Pryde, *Adaptive Measurements in the Optical Quantum Information Laboratory*, *IEEE J. Sel. Top. Quantum Electron.* **15**, 1661 (2009).
- [60] D. W. Berry and H. M. Wiseman, *Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry*, *Phys. Rev. Lett.* **85**, 5098 (2000).
- [61] A. Hentschel and B. C. Sanders, *Machine Learning for Precise Quantum Measurement*, *Phys. Rev. Lett.* **104**, 063603 (2010).