Achieving the Heisenberg limit in quantum metrology using quantum error correction

Sisi Zhou,[1,2] Mengzhen Zhang,[1,2] John Preskill,[3] and Liang Jiang[1,2]

[1]*Department of Applied Physics and Physics, Yale University, New Haven, Connecticut 06511, USA*
[2]*Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*
[3]*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

We study the fundamental limits on precision for parameter estimation using a quantum probe subject to Markovian noise. The best possible scaling of precision δ with the total probing time t is the Heisenberg limit (HL) $\delta \propto 1/t$, which can be achieved by a noiseless probe, but noise can reduce the precision to the standard quantum limit (SQL) $\delta \propto 1/\sqrt{t}$. We find a condition on the probe's Markovian noise such that SQL scaling cannot be surpassed if the condition is violated, but if the condition is satisfied then HL scaling can be achieved by using quantum error correction to protect the probe from damage, assuming that noiseless ancilla qubits are available, and fast, accurate quantum processing can be performed. When the HL is achievable, the optimal quantum code can be found by solving a semidefinite program. If the condition is violated but the noise channel acting on the probe is close to one that satisfies the condition, then approximate HL scaling can be realized for an extended period before eventually crossing over to SQL scaling.

Quantum metrology concerns the task of estimating a parameter, or several parameters, characterizing the Hamiltonian of a quantum system. This task is performed by preparing a suitable initial state of the system, allowing it to evolve for a specified time, performing a suitable measurement, and inferring the value of the parameter(s) from the measurement outcome. Quantum metrology is of great importance in science and technology, with wide applications including frequency spectroscopy, magnetometry, accelerometry, gravimetry, gravitational wave detection, and other high-precision measurements [1–9].

Quantum mechanics places a fundamental limit on measurement precision, called the Heisenberg limit (HL), which constrains how the precision of parameter estimation improves as the total probing time t increases. According to HL, the scaling of precision with t can be no better than $1/t$; equivalently, precision scales no better than $1/N$ with the total number of probes N used in an experiment. For a noiseless system, HL scaling is attainable in principle by, for example, preparing an entangled "cat" state of N probes [10–13]. In practice, though, in many cases environmental decoherence imposes a more severe limitation on precision; instead of HL, precision scales like $1/\sqrt{N}$, called the standard quantum limit (SQL), which can be achieved by using N independent probes [14–17].

Quantum error correction (QEC) is a method for reducing noise in quantum channels and quantum processors [18–20]. In principle, QEC enables a noisy quantum computer to simulate faithfully an ideal quantum computer, with reasonable overhead cost, if the noise is not too strong or too strongly correlated. But the potential value of QEC in quantum metrology has not yet been fully fleshed out, even as a matter of principle. A serious obstacle for applications of QEC to sensing is that it may in some cases be exceedingly hard to distinguish the signal arising from the Hamiltonian evolution of the probe system from the effects of the noise acting on the probe. Nevertheless, it has been shown that QEC can be invoked to achieve HL scaling under suitable conditions [21–24], and experiments demonstrating the efficacy of QEC in a room-temperature hybrid spin register have recently been conducted [25].

As is the case for quantum computing, we should expect positive (or negative) statements about improving metrology via QEC to be premised on suitable assumptions about the properties of the noise and the capabilities of our quantum hardware. But what assumptions are appropriate, and what can be inferred from these assumptions? In this paper, we assume that the probes used for parameter estimation are subject to noise described by a Markovian master equation [18, 26, 27], where the strength and structure of this noise is beyond the experimentalist's control. However, aside from the probe system, the experimentalist also has noiseless ancilla qubits at her disposal, and the ability to apply noiseless quantum gates which act jointly on the ancilla and probe; she can also perform perfect ancilla measurements, and reset the ancillas after measurement. Furthermore, these quantum gates and measurements can be executed as quickly as needed (though the Markovian description of the probe's noise is assumed to be applicable no matter how fast the processing). We endow the experimentalist with these powerful tools because we wish to address, as a matter of principle, how effectively QEC can overcome the deficiencies of the noisy probe system. In making these assumptions we are emboldened by the example of NV centers in diamond [25], where sensing of a magnetic field by an electron spin can be enhanced using a quantum code which takes advantage of the long coherence time of a nearby (ancilla) nuclear spin.

In accord with these assumptions, we adopt the sequential scheme for quantum metrology [28–30] (see Fig. 1(a)). In this scheme, a single noisy probe senses the unknown parameter for many rounds, where each round lasts for a short time interval dt, and the total number of rounds is t/dt, where t is the total sensing time. In between rounds, an arbitrary (noiseless) quantum operation can be applied instantaneously, which acts jointly on the probe and the noiseless ancillas. The rapid operations between rounds empower us to perform QEC, suppressing the damaging effects of the noise on the probe. Note that this sequential scheme can simulate a parallel scheme (Fig. 1(b)), in which N probes simultaneously sense the parameter for time t/N.

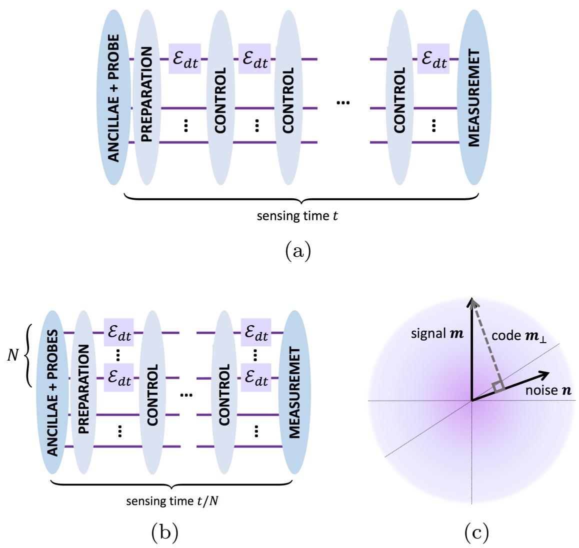

FIG. 1. (a) The sequential scheme. One probe sequentially senses the parameter for time t, with quantum controls applied every dt. (b) The parallel scheme. N probes sense the parameter for time t/N in parallel. The parallel scheme can be simulated by the sequential scheme. (c) Bloch sphere sketch illustrating the relation between the signal Hamiltonian, the noise and the QEC code in the two-dimensional case.

Previous studies have shown that whether HL scaling can be achieved by using QEC to protect a noisy probe depends on the algebraic structure of the noise. For example, if the probe is a qubit (two-dimensional quantum system), then HL scaling is possible when detecting a σ_z signal in the presence of bit-flip (σ_x) errors, but not for dephasing (σ_z) noise acting on the probe, even if arbitrary quantum controls and feedback are allowed [16]. For this example, we say that σ_x noise is "perpendicular" to the σ_z signal, while σ_z noise is "parallel" to the signal. In some previous work on improving metrology using QEC, perpendicular noise has been assumed [21, 22], but this assumption is not necessary — for a qubit probe, HL scaling is achievable for any noise channel with just one Hermitian jump operator L, except in the case where the signal Hamiltonian H commutes with L [30].

In this paper, we extend these results to any finite-dimensional probe, finding the necessary and sufficient condition on the noise for achievability of HL scaling. This condition is formulated as an algebraic relation between the signal Hamiltonian whose coefficient is to be estimated and the Lindblad operators which appear in the master equation describing the evolution of the probe. We prove that (1) if the signal Hamiltonian can be expressed as a linear combination of the identity operator I, the Lindblad operators L_k, their Hermitian conjugates $L_k^\dagger$ and the products $L_k^\dagger L_j$ for all k, j, then SQL scaling cannot be surpassed. (2) Otherwise HL scaling is achievable by using a QEC code such that the effective "logical" evolution of the probe is noiseless and unitary. This QEC technique can also be applied to the parallel scheme (Fig. 1(b)), where sensing is conducted using a cat state of N noiseless effective probes, each protected by the code [6].

The necessary condition for HL scaling is derived from the quantum Cramr-Rao bound [31–33]

$$\delta\hat{\omega} \geq 1/\sqrt{\mathcal{F}(\rho_\omega(t))}; \tag{1}$$

here $\hat{\omega}$ denotes any unbiased estimator for the parameter ω, and $\delta\hat{\omega}$ is that estimator's standard deviation. $\mathcal{F}(\rho_\omega(t))$ is the quantum Fisher information (QFI) of the state $\rho_\omega(t)$; this state is obtained by preparing an initial state ρ_{in} of the probe, and then evolving this state for total time t, where the evolution is governed by the ω-dependent probe Hamiltonian $H(\omega)$, the Markovian noise acting on the probe, and our fast quantum controls. We show that $\mathcal{F}(\rho_\omega(t))$ is at most asymptotically linear in t when the Hamiltonian $H(\omega)$ is contained in the linear span (denoted $\mathcal{S}$) of I, L_k, $L_k^\dagger$, and $L_k^\dagger L_j$, which means that SQL scaling cannot be surpassed in this case.

To prove the sufficient condition for HL scaling, we show that a QEC code achieving HL scaling can be explicitly constructed if $H(\omega)$ is not in the linear span $\mathcal{S}$. Furthermore, we find that searching for the optimal code, which maximizes the QFI, can be formulated as a semidefinite program that can be efficiently solved numerically, and can be solved analytically in some special cases. Our sufficient condition cannot be satisfied if the noise channel is full rank, and is therefore not applicable for generic noise. However, for noise which is ϵ-close to meeting our criterion, using the QEC code ensures that HL scaling can be maintained approximately for a time $O(1/\epsilon)$, before crossing over to asymptotic SQL scaling.

Results

Necessary and sufficient condition for HL. We denote the d-dimensional Hilbert space of our probe by $\mathcal{H}_P$, and we assume the state ρ_p of the probe evolves according to a time-homogeneous Lindblad master equation of the form (with $\hbar = 1$) [18, 26, 27],

$$\frac{d\rho_p}{dt} = -i[H, \rho_p] + \sum_{k=1}^{r} (L_k \rho_p L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho_p\}), \tag{2}$$

where H is the probe's Hamiltonian, $\{L_k\}$ are the Lindblad jump operators and we say that r is the "rank" of the noise channel acting on the probe. The Hamiltonian depends on a parameter ω, and our goal is to estimate ω. For simplicity we'll suppose $H = \omega G$, but our arguments actually apply more generally. If $H(\omega)$ is not a linear function of ω, by treating G as the derivative of $H(\omega)$ with respect to ω, the coding scheme we describe below can be repeated many times if necessary, using the latest estimate of ω after each round to adjust the code used in the next round. Eventually, ω is determined to sufficient accuracy that the function $H(\omega)$ can be well approximated by a linear function, and then our analysis applies. The asymptotic scaling of precision with the total probing time is not affected by the preliminary adaptive rounds [34].

We denote by $\mathcal{H}_A$ the d-dimensional Hilbert space of a noiseless ancilla system, whose evolution is determined solely by our fast and accurate quantum controls. Over the small time interval dt, during which no controls are applied, the ancilla evolves trivially, and the joint state ρ of probe and ancilla evolves according to the quantum channel

$$\begin{aligned}\mathcal{E}_{dt}(\rho) =& \rho - i\omega[G,\rho]dt + \\ & \sum_{k=1}^{r}(L_k\rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k,\rho\})dt + O(dt^2),\end{aligned} \tag{3}$$

where G, L_k are shorthand for $G\otimes I$, $L_k\otimes I$ respectively. We assume that this time interval dt is sufficiently small that corrections higher order in dt can be neglected. In between rounds of sensing, each lasting for time dt, control operations acting on $\rho\in\mathcal{H}_P\otimes\mathcal{H}_A$ are applied instantaneously.

Our conclusions about HL and SQL scaling of parameter estimation make use of an algebraic condition on the master equation which we will refer to often, and it will therefore be convenient to have a name for this condition. We will call it the HNLS condition, or simply HNLS, an acronym for "Hamiltonian not in Lindblad span." We denote by $\mathcal{S}$ the linear span of the operators $I, L_k, L_k^\dagger, L_k^\dagger L_j$ (for all k and j ranging from 1 to r), and say that the Hamiltonian H obeys the HNLS condition if H is not contained in $\mathcal{S}$. Now we can state our main conclusion about parameter estimation using fast and accurate quantum controls as Theorem 1.

Theorem 1. *Consider a finite-dimensional probe with Hamiltonian $H=\omega G$, subject to Markovian noise described by a Lindblad master equation with jump operators $\{L_k\}$. Then ω can be estimated with HL (Heisenberg-limited) precision if and only if G and $\{L_k\}$ obey the HNLS (Hamiltonian-not-in-Lindblad-span) condition.*

Qubit probe. To illustrate how Theorem 1 works, let's look at the case where the probe is a qubit [30]. Suppose one of the Lindblad operators is $L_1\propto \boldsymbol{n}\cdot\boldsymbol{\sigma}$, where $\boldsymbol{n}=\boldsymbol{n_r}+i\boldsymbol{n_i}$ is a normalized complex 3-vector and $\boldsymbol{n_r}$, $\boldsymbol{n_i}$ are its real and imaginary parts, so that $L_1^\dagger L_1\propto(\boldsymbol{n}^*\cdot\boldsymbol{\sigma})(\boldsymbol{n}\cdot\boldsymbol{\sigma}) = I + 2(\boldsymbol{n_i}\times\boldsymbol{n_r})\cdot\boldsymbol{\sigma}$. If $\boldsymbol{n_r}$ and $\boldsymbol{n_i}$ are not parallel vectors, then $\boldsymbol{n_r}$, $\boldsymbol{n_i}$ and $\boldsymbol{n_i}\times\boldsymbol{n_r}$ are linearly independent, which means that I, L_1, $L_1^\dagger$, and $L_1^\dagger L_1$ span the four-dimensional space of linear operators acting on the qubit. Hence HNLS cannot be satisfied by any qubit Hamiltonian, and therefore parameter estimation with HL scaling is not possible according to Theorem 1. We conclude that for HL scaling to be achievable, $\boldsymbol{n_r}$ and $\boldsymbol{n_i}$ must be parallel, which means that (after multiplying L_1 by a phase factor if necessary), we can choose L_1 to be Hermitian. Moreover, if L_1 and L_2 are two linearly independent Hermitian traceless Lindblad operators, then $\{I, L_1, L_2, L_1L_2\}$ span the space of qubit linear operators and HL scaling cannot be achieved. In fact, for a qubit probe, HNLS can be satisfied only if there is a single Hermitian (not necessarily traceless) Lindblad operator L, and the Hamiltonian does not commute with L.

We will describe below how to achieve HL scaling for any master equation that satisfies HNLS, by constructing a two-dimensional QEC code which protects the probe from the Markovian noise. To see how the code works for a qubit probe, suppose $G=\frac{1}{2}\boldsymbol{m}\cdot\boldsymbol{\sigma}$ and $L\propto\boldsymbol{n}\cdot\boldsymbol{\sigma}$ where $\boldsymbol{m}$ and $\boldsymbol{n}$ are unit vectors in $\mathbb{R}^3$ (see Fig. 1(c)). Then the basis vectors for the QEC code may be chosen to be

$$|C_0\rangle = |\boldsymbol{m}_\perp,+\rangle_P\otimes|0\rangle_A\,,\quad |C_1\rangle = |\boldsymbol{m}_\perp,-\rangle_P\otimes|1\rangle_A\,; \tag{4}$$

here $|0\rangle_A$, $|1\rangle_A$ are basis states for the ancilla qubit, and $|\boldsymbol{m}_\perp,\pm\rangle_P$ are the eigenstates with eigenvalues ±1 of $\boldsymbol{m}_\perp\cdot\boldsymbol{\sigma}$ where $\boldsymbol{m}_\perp$ is the (normalized) component of $\boldsymbol{m}$ perpendicular to $\boldsymbol{n}$. In particular, if $\boldsymbol{m}\perp\boldsymbol{n}$ (perpendicular noise), then $|C_0\rangle = |\boldsymbol{m},+\rangle_P\otimes|0\rangle_A$ and $|C_1\rangle = |\boldsymbol{m},-\rangle_P\otimes|1\rangle_A$, the coding scheme previously discussed in [21–24].

In the case of perpendicular noise, we estimate ω by tracking the evolution in the code space of a state initially prepared as (in a streamlined notation) $|\psi(0)\rangle = (|+,0\rangle+|-,1\rangle)/\sqrt{2}$; neglecting the noise, this state evolves in time t to

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\omega t/2}|+,0\rangle + e^{i\omega t/2}|-,1\rangle\right). \tag{5}$$

If a jump then occurs at time t, the state is transformed to

$$|\psi'(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\omega t/2}|-,0\rangle + e^{i\omega t/2}|+,1\rangle\right). \tag{6}$$

Jumps are detected by performing a two-outcome measurement which projects onto either the span of $\{|+,0\rangle,|-,1\rangle\}$ (the code space) or the span of $\{|-,0\rangle,|+,1\rangle\}$ (orthogonal to the code space), and when detected they are immediately correcting by flipping the probe. Because errors are immediately corrected, the error-corrected evolution matches perfectly the ideal evolution (without noise), for which HL scaling is possible.

When the noise is not perpendicular to the signal, then not just the jumps but also the Hamiltonian evolution can rotate the joint state of probe and ancilla away from the code space. However, after evolution for the short time interval dt the overlap with the code space remains large, so that the projection onto the code space succeeds with probability $1 - O(dt^2)$. Neglecting $O(dt^2)$ corrections, then, the joint probe-ancilla state rotates noiselessly in the code space, at a rate determined by the component of the Hamiltonian evolution along the code space. As long as this component is nonzero, HL scaling can be achieved.

As we'll see, this reasoning can be extended to any finite-dimensional probe satisfying HNLS. First, though, we'll discuss why HL is impossible when HNLS fails.

Non-achievability of HL when HNLS fails. The Cramér-Rao bound Eq. (1) is a powerful tool for assessing the efficacy of parameter estimation. Though it is challenging to compute the maximum attainable quantum Fisher information (QFI) for arbitrary quantum channels, useful upper bounds on QFI can be derived, which provide lower bounds on the precision of quantum metrology [16, 17, 28, 30, 35, 36]. The quantum channel describing the joint evolution of probe and ancilla has a Kraus operator representation

$$\mathcal{E}_{dt}(\rho) = \sum_k K_k \rho K_k^\dagger, \tag{7}$$

and in terms of these Kraus operators we define

$$\begin{aligned} \alpha_{dt} &= \sum_k \dot{K}_k^\dagger \dot{K}_k = \dot{\boldsymbol{K}}^\dagger \dot{\boldsymbol{K}}, \\ \beta_{dt} &= i \sum_k \dot{K}_k^\dagger K_k = i \dot{\boldsymbol{K}}^\dagger \boldsymbol{K}, \end{aligned} \tag{8}$$

where we express the Kraus operators in vector notation $\boldsymbol{K} := (K_0, K_1, \cdots)^T$, and the over-dot means the derivative with respect to ω. If ρ_{in} is the initial joint state of probe and ancilla at time 0, and $\rho(t)$ is the corresponding state at time t, then the upper bound on the QFI

$$\mathcal{F}(\rho(t)) \leq 4\frac{t}{dt}\|\alpha_{dt}\| + 4\left(\frac{t}{dt}\right)^2 \|\beta_{dt}\| \left(\|\beta_{dt}\| + 2\sqrt{\|\alpha_{dt}\|}\right), \tag{9}$$

($\|\cdot\|$ denotes the operator norm) derived by the "channel extension method", holds for any choice of ρ_{in}, even when fast and accurate quantum controls are applied during the evolution [30]. This upper bound on the QFI provides a lower bound on the precision $\delta\hat{\omega}$ via Eq. (1).

Kraus representations are not unique — for any matrix u satisfying $u^\dagger u = I$, $\boldsymbol{K}' = u\boldsymbol{K}$ represents the same channel as $\boldsymbol{K}$. Hence, we can tighten the upper bound on the QFI by minimizing the RHS of Eq. (9) over all such valid Kraus representations. We see that

$$\dot{\boldsymbol{K}}' = u\left(\dot{\boldsymbol{K}} - ih\boldsymbol{K}\right), \quad \dot{\boldsymbol{K}}'^\dagger = \left(\dot{\boldsymbol{K}} - ih\boldsymbol{K}\right)^\dagger u^\dagger \tag{10}$$

where $h = iu^\dagger \dot{u}$. Therefore, to find α_{dt} and β_{dt} providing the tightest upper bound on the QFI, it suffices to replace $\dot{\boldsymbol{K}}$ by $\dot{\boldsymbol{K}} - ih\boldsymbol{K}$ and to optimize over the Hermitian matrix h.

To evaluate the bound for asymptotically large t, we expand α_{dt}, β_{dt}, h in powers of $\sqrt{dt}$:

$$\alpha_{dt} = \alpha^{(0)} + \alpha^{(1)}\sqrt{dt} + \alpha^{(2)}dt + O(dt^{3/2}), \tag{11}$$

$$\beta_{dt} = \beta^{(0)} + \beta^{(1)}\sqrt{dt} + \beta^{(2)}dt + \beta^{(3)}dt^{3/2} + O(dt^2), \tag{12}$$

$$h = h^{(0)} + h^{(1)}\sqrt{dt} + h^{(2)}dt + h^{(3)}dt^{3/2} + O(dt^2). \tag{13}$$

We show in the Supplementary Methods that the first two terms in α_{dt} and the first four terms in β_{dt} can all be set to zero by choosing a suitable h, assuming that HNLS is violated. We therefore have $\alpha_{dt} = O(dt)$ and $\beta_{dt} = O(dt^2)$, so that the second term in the RHS of Eq. (9) vanishes as $dt \to 0$:

$$\mathcal{F}(\rho(t)) \leq 4\|\alpha^{(2)}\|t, \tag{14}$$

proving that SQL scaling cannot be surpassed when HNLS is violated (the necessary condition in Theorem 1). We require the probe to be finite dimensional in the statement of Theorem 1 because otherwise the norm of α_{dt} or β_{dt} could be infinite. The theorem can be applied to the case of a probe with an infinite-dimensional Hilbert space if the state of the probe is confined to a finite-dimensional subspace even for asymptotically large t.

QEC code for HL scaling when HNLS holds. Our discussion of the qubit probe indicates how a QEC code can be used to achieve HL scaling for estimating the parameter ω. The code allows us to correct quantum jumps whenever they occur, and in addition the noiseless error-corrected evolution in the code space depends nontrivially on ω. Similar considerations apply to higher-dimensional probes. Let Π_C denote the projection onto the code space. Jumps are correctable if the code satisfies the error correction conditions [18–20], namely:

$$\begin{aligned} &(1)\ \Pi_C L_k \Pi_C = \lambda_k \Pi_C,\ \forall k, \\ &(2)\ \Pi_C L_k^\dagger L_j \Pi_C = \mu_{kj} \Pi_C,\ \forall k, j. \end{aligned} \tag{15}$$

The error-corrected joint state of probe and ancilla evolves according to the unitary channel

$$\frac{d\rho}{dt} = -i[H_{\text{eff}}, \rho] \tag{16}$$

where $H_{\text{eff}} = \Pi_C H \Pi_C$, and there is a code state for which the evolution depends nontrivially on ω provided that

$$(3)\ \Pi_C G \Pi_C \neq \text{constant} \times \Pi_C. \tag{17}$$

To prove the sufficient condition in Theorem 1, we need to show that a code with properties (1)–(3) can be constructed whenever HNLS is satisfied. For further justification of these conditions see the Methods.

To see how the code is constructed, note that the d-dimensional Hermitian matrices form a real Hilbert space where the inner product of two matrices A and B is defined to be $\mathrm{tr}(AB)$. Let $\mathcal{S}$ denote the subspace of Hermitian matrices spanned by I, $L_k + L_k^\dagger$, $i(L_k - L_k^\dagger)$, $L_k^\dagger L_j + L_j^\dagger L_k$ and $i(L_k^\dagger L_j - L_j^\dagger L_k)$ for all k, j. Then G has a unique decomposition into $G = G_\parallel + G_\perp$, where $G_\parallel \in \mathcal{S}$ and $G_\perp \perp \mathcal{S}$.

If HNLS holds, then $G_\perp$ is nonzero. It must also be traceless, in order to be orthogonal to I, which is contained in $\mathcal{S}$. Therefore, using the spectral decomposition, we can write $G_\perp = \frac{1}{2}\left(\mathrm{tr}\,|G_\perp|\right)(\rho_0 - \rho_1)$, where ρ_0 and ρ_1 are trace-one positive matrices with orthogonal support and $|G_\perp| := \sqrt{G_\perp^2}$. Our QEC code is chosen to be the two-dimensional subspace of $\mathcal{H}_S \otimes \mathcal{H}_A$ spanned by $|C_0\rangle$ and $|C_1\rangle$, which are normalized purifications of ρ_0 and ρ_1 respectively, with orthogonal support in $\mathcal{H}_A$. (If the probe is d-dimensional, a d-dimensional ancilla can purify its state.) Because the code basis states have orthogonal support on $\mathcal{H}_A$, it follows that, for any O acting on $\mathcal{H}_P$,

$$\langle C_0|O \otimes I|C_1\rangle = 0 = \langle C_1|O \otimes I|C_0\rangle, \tag{18}$$

and furthermore

$$\mathrm{tr}\big((|C_0\rangle\langle C_0| - |C_1\rangle\langle C_1|)(O \otimes I)\big) = \mathrm{tr}\big((\rho_0 - \rho_1)O\big) = \frac{2\,\mathrm{tr}(G_\perp O)}{\mathrm{tr}|G_\perp|}. \tag{19}$$

Because the RHS of Eq. (19) vanishes for any O in the span $\mathcal{S}$, our code satisfies the conditions (1) and (2). Condition (3) is also satisfied, because $\langle C_0|G|C_0\rangle - \langle C_1|G|C_1\rangle = 2\,\mathrm{tr}(G_\perp^2)/\mathrm{tr}|G_\perp| > 0$, which means that the Hamiltonian is not constant when projected onto the code space. Thus we have demonstrated the existence of a code with properties (1)–(3).

Code optimization. When HNLS is satisfied, we can use our QEC code, along with fast and accurate quantum control, to achieve noiseless evolution of the error-corrected probe, governed by the effective Hamiltonian $H_{\mathrm{eff}} = \Pi_C H \Pi_C$ where Π_C is the orthogonal projection onto the code space. When the evolution is noiseless, the QFI for the state at time t is

$$\mathcal{F}(\rho(t)) = 4t^2\Big[\mathrm{tr}(\rho_{\mathrm{in}}\dot{H}_{\mathrm{eff}}^2) - \big(\mathrm{tr}(\rho_{\mathrm{in}}\dot{H}_{\mathrm{eff}})\big)^2\Big]. \tag{20}$$

where ρ_{in} is the initial state at time $t = 0$. The QFI is maximized by choosing the initial pure state

$$|\psi_{\mathrm{in}}\rangle = \frac{1}{\sqrt{2}}(|\lambda_{\mathrm{min}}\rangle + |\lambda_{\mathrm{max}}\rangle), \tag{21}$$

where $|\lambda_{\mathrm{min}}\rangle$, $|\lambda_{\mathrm{max}}\rangle$ are the eigenstates of $\dot{H}_{\mathrm{eff}} = G_{\mathrm{eff}}$ with the minimal and maximal eigenvalues; with this choice the QFI is

$$\mathcal{F}(\rho(t)) = t^2\big(\lambda_{\mathrm{max}} - \lambda_{\mathrm{min}}\big)^2. \tag{22}$$

By measuring in the appropriate basis at time t, we can estimate ω with a precision which saturates the Cramér-Rao bound, realizing HL scaling.

Because the optimal initial state is a superposition of just two eigenstates of G_{eff}, a two-dimensional QEC code suffices for achieving the best possible precision. For a code with basis states $\{|C_0\rangle, |C_1\rangle\}$, the effective Hamiltonian is

$$G_{\mathrm{eff}} = |C_0\rangle\langle C_0|G_\perp|C_0\rangle\langle C_0| + |C_1\rangle\langle C_1|G_\perp|C_1\rangle\langle C_1|; \tag{23}$$

here we have ignored the contribution due to $G_\parallel$, which is an irrelevant additive constant if the code satisfies condition (2). We have seen how to construct a code for which

$$\lambda_{\mathrm{max}} - \lambda_{\mathrm{min}} = 2\,\frac{\mathrm{tr}(G_\perp^2)}{\mathrm{tr}|G_\perp|}. \tag{24}$$

It is possible, though, that a larger value of this difference of eigenvalues could be achieved using a different code, improving the precision by a constant factor (independent of the time t).

To search for a better code, with basis states $\{|C_0\rangle, |C_1\rangle\}$, define

$$\tilde{\rho}_0 = \mathrm{tr}_A(|C_0\rangle\langle C_0|), \quad \tilde{\rho}_1 = \mathrm{tr}_A(|C_1\rangle\langle C_1|), \tag{25}$$

and consider

$$\tilde{G} = \tilde{\rho}_0 - \tilde{\rho}_1. \tag{26}$$

Conditions (1)-(2) on the code imply

$$\mathrm{tr}(\tilde{G}O) = 0,\ \forall O \in \mathcal{S}, \tag{27}$$

and we want to maximize

$$\lambda_{\mathrm{max}} - \lambda_{\mathrm{min}} = \mathrm{tr}(G_{\mathrm{eff}}\tilde{G}) = \mathrm{tr}(G_\perp\tilde{G}), \tag{28}$$

over matrices $\tilde{G}$ of the form Eq. (26) subject to Eq. (27). Note that $\tilde{G}$ is the difference of two normalized density operators, and therefore satisfies $\mathrm{tr}|\tilde{G}| \leq 2$. In fact, though, if $\tilde{G}$ obeys the constraint Eq. (27), then the constraint is still satisfied if we rescale $\tilde{G}$ by a real constant greater than one, which increases $\mathrm{tr}(G_\perp\tilde{G})$; hence the maximum of $\mathrm{tr}(G_\perp\tilde{G})$ is achieved for $\mathrm{tr}|\tilde{G}| = 2$, which means that $\tilde{\rho}_0$ and $\tilde{\rho}_1$ have orthogonal support.

Now recall that $G_\perp = \frac{1}{2}\left(\mathrm{tr}\,|G_\perp|\right)(\rho_0 - \rho_1)$ is also (up to normalization) a difference of density operators with orthogonal support, and obeys the constraint Eq. (27). The quantity to be maximized is proportional to

$$\mathrm{tr}\left[(\rho_0 - \rho_1)(\tilde{\rho}_0 - \tilde{\rho}_1)\right] = \mathrm{tr}(\rho_0\tilde{\rho}_0 + \rho_1\tilde{\rho}_1 - \rho_0\tilde{\rho}_1 - \rho_1\tilde{\rho}_0). \tag{29}$$

If ρ_0 and ρ_1 are both rank 1, then the maximum is achieved by choosing $\tilde{\rho}_0 = \rho_0$ and $\tilde{\rho}_1 = \rho_1$. Conditions (1)-(2) are satisfied by choosing $|C_0\rangle$ and $|C_1\rangle$ to be purifications of ρ_0 and ρ_1 with orthogonal support on $\mathcal{H}_A$. Thus we have recovered the code we constructed previously. If ρ_0 or ρ_1 is higher rank, though, then a different code achieves a higher maximum, and hence better precision for parameter estimation.

Geometrical picture. There is an alternative description of the code optimization, with a pleasing geometrical interpretation. As discussed in the Methods, the optimization can be formulated as a semidefinite program with a feasible dual program. By solving the dual program we find that, for the optimal QEC code, the QFI is

$$\mathcal{F}(\rho(t)) = 4t^2 \min_{\tilde{G}_\parallel \in \mathcal{S}} \|G_\perp - \tilde{G}_\parallel\|^2. \tag{30}$$

In this sense, the QFI is determined by the minimal distance between $G_\perp$ and $\mathcal{S}$ (see Fig. 2(b)).

We can recover the solution to the primal problem from the solution to the dual problem. We denote by $\tilde{G}_\parallel^\diamond$ the choice of $\tilde{G}_\parallel \in \mathcal{S}$ that minimizes Eq. (30), and we define

$$\tilde{G}^\diamond := G_\perp - \tilde{G}_\parallel^\diamond. \tag{31}$$

Then $\tilde{G}^*$ which maximizes Eq. (28) has the form

$$\tilde{G}^* = \tilde{\rho}_0^\diamond - \tilde{\rho}_1^\diamond, \tag{32}$$

where $\tilde{\rho}_0^\diamond$ is a density operator supported on the eigenspace of $\tilde{G}^\diamond$ with the maximal eigenvalue, and $\tilde{\rho}_1^\diamond$ is a density operator supported on the eigenspace of $\tilde{G}^\diamond$ with the minimal eigenvalue. The minimization in Eq. (30) ensures that $\tilde{G}^*$ of this form can be chosen to obey the constraint Eq. (27).

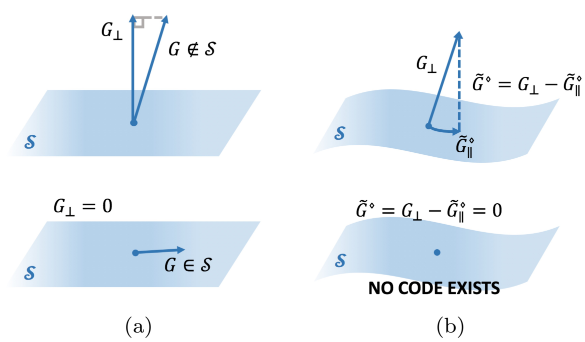

FIG. 2. (a) An illustration of obtaining $G_\perp$ from G in the Hilbert space of Hermitian matrices equipped with the Hilbert?Schmidt norm $\sqrt{\mathrm{tr}(O \cdot O)}$. $G_\perp \neq 0$ if and only if $G \notin \mathcal{S}$. (b) An illustration of obtaining the optimal QEC code from $G_\perp$ in the linear space of Hermitian matrices equipped with the operator norm $\|O\| = \max_{|\psi\rangle} \langle\psi| O |\psi\rangle$. No code achieving HL scaling exists if and only if $G_\perp = 0$.

In the noiseless case ($\mathcal{S} = \mathrm{span}\{I\}$), the minimum in Eq. (30) occurs when the maximum and minimum eigenvalues $G_\perp - \tilde{G}_\parallel$ have the same absolute value, and then the operator norm is half the difference of the maximum and minimum eigenvalues of $G_\perp$. Hence we recover the result Eq. (22). When noise is introduced, $\mathcal{S}$ swells and the minimal distance shrinks, lowering the QFI and reducing the precision of parameter estimation. If HNLS fails, then the minimum distance is zero, and no QEC code can achieve HL scaling, in accord with Theorem 1.

Example: Kerr effect with photon loss. To illustrate how the optimization procedure works, consider a bosonic mode with the nonlinear (Kerr effect [37]) Hamiltonian

$$H(\omega) = \omega(a^\dagger a)^2, \tag{33}$$

where our objective is to estimate ω. In this case the probe is infinite dimensional, but suppose we assume that the occupation number $n = a^\dagger a$ is bounded: $n \leq \bar{n}$, where $\bar{n}$ is even. The noise source is photon loss, with Lindblad jump operator $L \propto a$. Can we find a QEC code that protects the probe against loss and achieves HL scaling for estimation of ω?

To solve the dual program, we find real parameters α, β, γ, δ which minimize the operator norm of

$$\tilde{n}^2 := n^2 + \alpha n + \beta a + \gamma a^\dagger + \delta, \tag{34}$$

where $n \leq \bar{n}$. Since a and $a^\dagger$ are off-diagonal in the occupation number basis, we should set β and γ to zero for the purpose of minimizing the difference between the largest and smallest eigenvalue of $\tilde{n}^2$. After choosing α such that $\tilde{n}^2$ is minimized at $n = \bar{n}/2$, and choosing δ so that the maximum and minimum eigenvalues of $\tilde{n}^2$ are equal in absolute value and opposite in sign, we have

$$\left(\tilde{n}^2\right)^\diamond = \left(n - \frac{1}{2}\bar{n}\right)^2 - \frac{1}{8}\bar{n}^2, \tag{35}$$

which has operator norm $\|\left(\tilde{n}^2\right)^\diamond\| = \bar{n}^2/8$; hence the optimal QFI after evolution time t is $\mathcal{F}(\rho(t)) = t^2\bar{n}^4/16$, according to Eq. (30). For comparison, the minimal operator norm is $\bar{n}^2/2$ for a noiseless bosonic mode with $n \leq \bar{n}$. We see that loss reduces the precision of our estimate of ω, but only by a factor of 4 if we use the optimal QEC code. HL scaling can still be maintained.

To find the code states, we note that the eigenstate of $\left(\tilde{n}^2\right)^\diamond$ with the lowest eigenvalue $-\bar{n}^2/8$ is $|n = \bar{n}/2\rangle$, while the largest eigenvalue $+\bar{n}^2/8$ has the two degenerate eigenstates $|n = 0\rangle$ and $|n = \bar{n}\rangle$. The code condition (2) requires that both code vectors have the same expectation value of $L^\dagger L \propto n$, and we therefore may choose

$$|C_0\rangle = |n/2\rangle_P \otimes |0\rangle_A, \; |C_1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_P + |n\rangle_P\right) \otimes |1\rangle_A \tag{36}$$

as the code achieving optimal precision.

Approximate error correction. Generic Markovian noise is full rank, which means that the span $\mathcal{S}$ is the full Hilbert space $\mathcal{H}_P$ of the probe; hence the HNLS criterion of Theorem 1 is violated for any probe Hamiltonian $H(\omega)$, and asymptotic SQL scaling cannot be surpassed. Therefore, for any Markovian noise model that meets the HNLS criterion, the HL scaling achieved by our QEC code is not robust against generic small perturbations of the noise model.

We should therefore emphasize that a substantial improvement in precision can be achieved using a QEC code even in cases where HNLS is violated. Consider in particular a Markovian master equation with Lindblad operators divided into two sets $\{L_k\}$ (L-type noise) and $\{J_m\}$ (J-type noise), where the J-type noise is parametrically weak, with noise strength

$$\epsilon := \Big\| \sum_m J_m^\dagger J_m \Big\|. \quad (37)$$

If we use the optimal code that protects against L-type noise, then the joint logical state of probe and ancilla evolves according to a modified master equation, with Hamiltonian $H_{\rm eff} = \Pi_C H \Pi_C$, and effective Lindblad operators $J_{m,j}$ acting within the code space, where

$$\Big\| \sum_{m,j} J_{m,j}^\dagger J_{m,j} \Big\| \leq \epsilon. \quad (38)$$

(See the Methods for further discussion.)

This means that the state of the error-corrected probe deviates by a distance $O(\epsilon t)$ (in the L^1 norm) from the (effectively noiseless) evolution in the absence of J-type noise. Therefore, using this code, the QFI of the error-corrected probe increases quadratically in time (and the precision $\delta\hat{\omega}$ scales like $1/t$) up until an evolution time $t \propto 1/\epsilon$, before crossing over to asymptotic SQL scaling.

Discussion

Noise limits the precision of quantum sensing. Quantum error correction can suppress the damaging effects of noise, thereby improving the fidelity of quantum information processing and quantum communication, but whether QEC improves the efficacy of quantum sensing depends on the structure of the noise and the signal Hamiltonian. Unless suitable conditions are met, the QEC code that tames the noise might obscure the signal as well, nullifying the advantages of QEC.

Our study of quantum sensing using a noisy probe has focused on whether the precision δ of parameter estimation scales asymptotically with the total sensing time t as $\delta \propto 1/t$ (Heisenberg limit) or $\delta \propto 1/\sqrt{t}$ (standard quantum limit). We have investigated this question in an idealized setting, where the experimentalist has access to noiseless ancillas and can apply quantum controls which are arbitrarily fast and accurate, and we have also assumed that the noise acting on the probe is Markovian. Under these assumptions, we have found the general criterion for HL scaling to be achievable, the Hamiltonian-not-in Lindblad span (HNLS) criterion. If HNLS is satisfied, a QEC code can be constructed which achieves HL scaling, and if HNLS is violated, then SQL scaling cannot be surpassed.

In the case where HNLS is satisfied, we have seen that the QEC code achieving the optimal precision can be chosen to be two-dimensional, and we have described an algorithm for constructing this optimal code. The precision attained by this code has a geometrical interpretation in terms of the minimal distance (in the operator norm) of the signal Hamiltonian from the "Lindblad span" $\mathcal{S}$, the subspace spanned by I, L_k, $L_k^\dagger$, and $L_k^\dagger L_j$, where $\{L_k\}$ is the set of Lindblad jump operators appearing in the probe's Markovian master equation.

Many questions merit further investigation. We have focused on the dichotomy of HL vs. SQL scaling, but it is also worthwhile to characterize constant factor improvements in precision that can be achieved using QEC in cases where HNLS is violated [38]. We should clarify the applications of QEC to sensing when quantum controls have realistic accuracy and speed. Finally, it is interesting to consider probes subject to non-Markovian noise. In that case, tools such as dynamical decoupling [39–42] can mitigate noise, but just as for QEC, we need to balance desirable suppression of the noise against undesirable suppression of the signal in order to formulate the most effective sensing strategy.

Note added: During the preparation of the manuscript, the authors became aware of a related work by R. Demkowicz-Dobrza?ski, *et al.* [43]. The necessity part of Theorem 1 has been proven using similar methods in that work.

Methods

The QEC condition. Here we consider the quantum channel Eq. (3), which describes the joint evolution of a noisy quantum probe and noiseless ancilla over time interval dt. Suppose that a QEC code obeys the conditions (1) and (2) in Eq. (15), where Π_C is the orthogonal projector onto the code space. We will construct a recovery operator such that the error-corrected time evolution is unitary to linear order in dt, governed by the effective Hamiltonian $H_{\rm eff} = \omega \Pi_C G \Pi_C$.

For a density operator $\rho = \Pi_C \rho \Pi_C$ in the code space, conditions (1) and (2) imply

$$\Pi_C \mathcal{E}_{dt}(\rho) \Pi_C = \rho - i\omega[\Pi_C G \Pi_C, \rho] dt + \sum_{k=1}^{r} (|\lambda_k|^2 - \mu_{kk}) \rho dt + O(dt^2), \quad (39)$$

$$\Pi_E \mathcal{E}_{dt}(\rho) \Pi_E = \sum_{k=1}^{r} (L_k - \lambda_k) \rho (L_k^\dagger - \lambda_k^*) dt + O(dt^2), \quad (40)$$

where $\Pi_E = I - \Pi_C$. When acting on a state in the code space, $\Pi_E \mathcal{E}_d t(\cdot) \Pi_E$ is an operation with Kraus operators

$$K_k = (I - \Pi_C) L_k \Pi_C \sqrt{dt}, \quad (41)$$

which obey the normalization condition

$$\begin{aligned} \sum_{k=1}^{r} K_k^\dagger K_k &= \sum_{k=1}^{r} \Pi_C L_k^\dagger (I - \Pi_C) L_k \Pi_C dt \\ &= \sum_{k=1}^{r} \left(\mu_{kk} - |\lambda_k|^2\right) dt, \end{aligned} \quad (42)$$

where we have used conditions (1) and (2). Therefore, if ρ is in the code space, then a recovery channel $\mathcal{R}_E(\cdot)$ such that

$$\mathcal{R}_E(\Pi_E\mathcal{E}_{dt}(\rho)\Pi_E) = -\sum_{k=1}^{r}(|\lambda_k|^2-\mu_{kk})\rho dt+O(dt^2) \quad (43)$$

can be constructed, provided that the operators $\{L_k - \lambda_k\}_{k=1}^{r}$ satisfy the standard QEC conditions [18–20]. Indeed these conditions are satisfied because $\Pi_C(L_k^\dagger - \lambda_k^*)(L_j - \lambda_j)\Pi_C = (\mu_{kj} - \lambda_k^*\lambda_j)\Pi_C$, for all k, j. Therefore, the quantum channel

$$\mathcal{R}(\sigma) = \Pi_C\sigma\Pi_C + \mathcal{R}_E(\Pi_E\sigma\Pi_E) \quad (44)$$

completely reverses the effects of the noise. The channel describing time evolution for time dt followed by an instantaneous recovery step is

$$\mathcal{R}(\mathcal{E}_{dt}(\rho)) = \rho - i\omega[\Pi_C G\Pi_C, \rho]dt + O(dt^2), \quad (45)$$

a noiseless unitary channel with effective Hamiltonian $\omega\Pi_C G\Pi_C$ if $O(dt^2)$ corrections are neglected.

The dependence of the Hamiltonian on ω can be detected, for a suitable initial code state $\rho_{\rm in}$, if and only if $\Pi_C G\Pi_C$ has at least two distinct eigenvalues. Thus for nontrivial error-corrected sensing we require condition (3): $\Pi_C G\Pi_C \neq \text{constant} \times \Pi_C$.

Robustness of the QEC scheme. We consider the following quantum channel, where the "J-type noise," with Lindblad operators $\{J_m\}_{m=1}^{r_2}$, is regarded as a small perturbation:

$$\mathcal{E}_{dt}(\rho) = \rho - i\omega[G, \rho]dt + \sum_{k=1}^{r_1}(L_k\rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\})dt + \sum_{m=1}^{r_2}(J_m\rho J_m^\dagger - \frac{1}{2}\{J_m^\dagger J_m, \rho\})dt + O(dt^2). \quad (46)$$

We assume that the "L-type noise," with Lindblad operators $\{L_k\}_{k=1}^{r_1}$, obeys the QEC conditions (1) and (2), and that $\mathcal{R}$ is the recovery operation which corrects this noise. By applying this recovery step after the action of $\mathcal{E}_{dt}$ on a state ρ in the code space, we obtain a modified channel with residual J-type noise.

Suppose that $\mathcal{R}$ has the Kraus operator decomposition $\mathcal{R}(\sigma) = \sum_{j=1}^{s} R_j\sigma R_j^\dagger$, where $\sum_{j=1}^{s} R_j^\dagger R_j = I$. We also assume that $R_j = \Pi_C R_j$, because the recovery procedure has been constructed such that the state after recovery is always in the code space. Then

$$\mathcal{R}(\mathcal{E}_{dt}(\rho)) = \rho - i\omega[\Pi_C G\Pi_C, \rho]dt + \sum_{m=1}^{r_2}\sum_{j=1}^{s}(J_{m,j}^{(C)}\rho J_{m,j}^{(C)\dagger} - \frac{1}{2}\{J_{m,j}^{(C)\dagger}J_{m,j}^{(C)}, \rho\})dt + O(dt^2), \quad (47)$$

where $\{J_{m,j}^{(C)} = \Pi_C R_j J_m \Pi_C\}$ are the effective Lindblad operators acting on code states.

The trace (L^1) distance [18] between the unitarily evolving state Eq. (45) and the state subjected to the residual noise Eq. (47) is bounded above by

$$\Big\|\sum_{m=1}^{r_2}\sum_{j=1}^{s} J_{m,j}^{(C)\dagger}J_{m,j}^{(C)}\Big\|dt = \Big\|\Pi_C\Big(\sum_{m=1}^{r_2} J_m^\dagger J_m\Big)\Pi_C\Big\|dt \leq \Big\|\sum_{m=1}^{r_2} J_m^\dagger J_m\Big\|dt \quad (48)$$

to first order in dt. If the noise strength

$$\epsilon := \Big\|\sum_{m=1}^{r_2} J_m^\dagger J_m\Big\| \quad (49)$$

of the Lindblad operators $\{J_m\}_{m=1}^{r_2}$ is low, the evolution is approximately unitary when $t \ll 1/\epsilon$. In this sense, the QEC scheme is robust against small J-type noise. .

Code optimization as a semidefinite program. Optimization of the QEC code can be formulated as the following optimization problem:

$$\begin{aligned}&\text{maximize } \operatorname{tr}(\tilde{G}G_\perp)\\ &\text{subject to } \operatorname{tr}(|\tilde{G}|) \leq 2 \text{ and } \operatorname{tr}(\tilde{G}) = \operatorname{tr}(\tilde{G}L_k)\\ &\qquad = \operatorname{tr}(\tilde{G}L_k^\dagger L_j) = 0,\ \forall j, k.\end{aligned} \quad (50)$$

This optimization problem is convex (because $\operatorname{tr}(|\bullet|)$ is convex) and satisfies the Slater's condition, so it can be solved by solving its Lagrange dual problem [44]. The Lagrangian $L(\tilde{G}, \lambda, \nu)$ is defined for $\lambda \geq 0$ and $\nu_k \in \mathbb{R}$:

$$L(\tilde{G}, \lambda, \nu) = \operatorname{tr}(\tilde{G}G_\perp) - \lambda(\operatorname{tr}(|\tilde{G}|) - 2) + \sum_k \nu_k \operatorname{tr}(E_k\tilde{G}), \quad (51)$$

where $\{E_k\}$ is any basis of $\mathcal{S}$. The optimal value is obtained by taking the minimum of the dual

$$\begin{aligned}g(\lambda, \nu) &= \max_{\tilde{G}} L(\tilde{G}, \lambda, \nu)\\ &= \max_{\tilde{G}} \operatorname{tr}\big((G_\perp + \sum_k \nu_k E_k)\tilde{G} - \lambda|\tilde{G}|\big) + 2\lambda\\ &= \begin{cases} 2\lambda & \lambda \geq \|G_\perp + \sum_k \nu_k E_k\| \\ \infty & \lambda \leq \|G_\perp + \sum_k \nu_k E_k\| \end{cases}\end{aligned} \quad (52)$$

over λ and $\{\nu_k\}$, where $\|\bullet\| = \max_{|\psi\rangle} |\langle\psi|\bullet|\psi\rangle|$ is the operator norm. Hence the optimal value of the primal problem is

$$\min_{\lambda,\nu} g(\lambda, \nu) = 2\min_{\nu_k}\|G_\perp + \sum_k \nu_k E_k\| = 2\min_{\tilde{G}_\parallel \in \mathcal{S}}\|G_\perp - \tilde{G}_\parallel\|. \quad (53)$$

The optimization problem Eq. (53) is equivalent to the following semidefinite program (SDP) [44]:

$$\begin{aligned}&\text{minimize } s\\ &\text{subject to } \begin{pmatrix} sI & G_\perp + \sum_k \nu_k E_k \\ G_\perp + \sum_k \nu_k E_k & sI \end{pmatrix} \succeq 0\end{aligned} \quad (54)$$

for variables $\nu_k \in \mathbb{R}$ and $s \geq 0$. Here "$\succeq 0$" denotes positive semidefinite matrices. SDPs can be solved using the Matlab-based package CVX [45].

Once we have the solution to the dual problem we can use it to find the solution to the primal problem. We denote by $\lambda^\diamond$ and $\nu^\diamond$ the values of λ and ν where $g(\lambda, \nu)$ attains its minimum, and define

$$\tilde{G}^\diamond = G_\perp + \sum_k \nu_k^\diamond E_k. \tag{55}$$

The minimum $g(\lambda^\diamond, \nu^\diamond)$ matches the value of the Lagrangian $L(\tilde{G}, \lambda^\diamond, \nu^\diamond)$ when $\tilde{G} = \tilde{G}^*$ is the value of $\tilde{G}$ which maximizes $\mathrm{tr}\left(\tilde{G}G_\perp\right)$ subject to the constraints. This means that

$$\mathrm{tr}\left(\tilde{G}^*\tilde{G}^\diamond\right) = 2\|\tilde{G}^\diamond\|. \tag{56}$$

Since we require $\mathrm{tr}(\tilde{G}^*) = 0$ and $\mathrm{tr}|\tilde{G}^*| = 2$, and because minimizing $g(\lambda, \nu)$ enforces that the maximum and minimal eigenvalues of $\tilde{G}^\diamond$ have the same absolute value and opposite sign, we conclude that

$$\tilde{G}^* = \tilde{\rho}_0^\diamond - \tilde{\rho}_1^\diamond, \tag{57}$$

where $\tilde{\rho}_0^\diamond$ is a density operator supported on the eigenspace of $\tilde{G}^\diamond$ with the maximal eigenvalue, and $\tilde{\rho}_1^\diamond$ is a density operator supported on the eigenspace of $\tilde{G}^\diamond$ with the minimal eigenvalue. A $\tilde{G}^*$ of this form which satisfies the constraints of the primal problem is guaranteed to exist.

ACKNOWLEDGMENTS

We thank Fernando Brandão, Yanbei Chen, Steve Girvin, Linshu Li, Mikhail Lukin, Changling Zou for inspiring discussions. We acknowledge support from the ARL-CDQI (W911NF-15-2-0067), ARO (W911NF-14-1-0011, W911NF-14-1-0563), ARO MURI (W911NF-16-1-0349), AFOSR MURI (FA9550-14-1-0052, FA9550-15-1-0015), NSF (EFMA-1640959), Alfred P. Sloan Foundation (BR2013-049), and Packard Foundation (2013-39273). The Institute for Quantum Information and Matter is an NSF Physics Frontiers Center with support from the Gordon and Betty Moore Foundation.

[1] Caves, C. M. Quantum-mechanical noise in an interferometer. *Physical Review D* **23**, 1693 (1981).

[2] Wineland, D., Bollinger, J., Itano, W., Moore, F. & Heinzen, D. Spin squeezing and reduced quantum noise in spectroscopy. *Physical Review A* **46**, R6797 (1992).

[3] Holland, M. & Burnett, K. Interferometric detection of optical phase shifts at the heisenberg limit. *Physical review letters* **71**, 1355 (1993).

[4] McKenzie, K., Shaddock, D. A., McClelland, D. E., Buchler, B. C. & Lam, P. K. Experimental demonstration of a squeezing-enhanced power-recycled michelson interferometer for gravitational wave detection. *Physical review letters* **88**, 231102 (2002).

[5] Lee, H., Kok, P. & Dowling, J. P. A quantum rosetta stone for interferometry. *Journal of Modern Optics* **49**, 2325–2338 (2002).

[6] Bollinger, J., Itano, W. M., Wineland, D. & Heinzen, D. Optimal frequency measurements with maximally correlated states. *Physical Review A* **54**, R4649 (1996).

[7] Leibfried, D. *et al.* Toward heisenberg-limited spectroscopy with multiparticle entangled states. *Science* **304**, 1476–1478 (2004).

[8] Valencia, A., Scarcelli, G. & Shih, Y. Distant clock synchronization using entangled photon pairs. *Applied Physics Letters* **85**, 2655–2657 (2004).

[9] de Burgh, M. & Bartlett, S. D. Quantum methods for clock synchronization: Beating the standard quantum limit without entanglement. *Physical Review A* **72**, 042301 (2005).

[10] Giovannetti, V., Lloyd, S. & Maccone, L. Quantum-enhanced measurements: beating the standard quantum limit. *Science* **306**, 1330–1336 (2004).

[11] Giovannetti, V., Lloyd, S. & Maccone, L. Quantum metrology. *Physical review letters* **96**, 010401 (2006).

[12] Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nature photonics* **5**, 222–229 (2011).

[13] Chaves, R., Brask, J., Markiewicz, M., Kołodyński, J. & Acín, A. Noisy metrology beyond the standard quantum limit. *Physical review letters* **111**, 120401 (2013).

[14] Huelga, S. F. *et al.* Improvement of frequency standards with quantum entanglement. *Physical Review Letters* **79**, 3865 (1997).

[15] Demkowicz-Dobrzański, R. *et al.* Quantum phase estimation with lossy interferometers. *Physical Review A* **80**, 013825 (2009).

[16] Escher, B., de Matos Filho, R. & Davidovich, L. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. *Nature Physics* **7**, 406–411 (2011).

[17] Demkowicz-Dobrzański, R., Kołodyński, J. & Guţă, M. The elusive heisenberg limit in quantum-enhanced metrology. *Nature Communications* **3**, 1063 (2012).

[18] Nielsen, M. A. & Chuang, I. L. *Quantum computation and quantum information* (Cambridge university press, 2010).

[19] Knill, E. & Laflamme, R. Theory of quantum error-correcting codes. *Physical Review A* **55**, 900 (1997).

[20] Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics*, vol. 68, 13–58 (2009).

[21] Kessler, E. M., Lovchinsky, I., Sushkov, A. O. & Lukin, M. D. Quantum error correction for metrology. *Physical*

review letters **112**, 150802 (2014).

[22] Arrad, G., Vinkler, Y., Aharonov, D. & Retzker, A. Increasing sensing resolution with error correction. *Physical review letters* **112**, 150801 (2014).

[23] Dür, W., Skotiniotis, M., Froewis, F. & Kraus, B. Improved quantum metrology using quantum error correction. *Physical Review Letters* **112**, 080801 (2014).

[24] Ozeri, R. Heisenberg limited metrology using quantum error-correction codes. *arXiv:1310.3432* (2013).

[25] Unden, T. *et al.* Quantum metrology enhanced by repetitive quantum error correction. *Physical Review Letters* **116**, 230502 (2016).

[26] Gorini, V., Kossakowski, A. & Sudarshan, E. C. G. Completely positive dynamical semigroups of n-level systems. *Journal of Mathematical Physics* **17**, 821–825 (1976).

[27] Lindblad, G. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics* **48**, 119–130 (1976).

[28] Demkowicz-Dobrzański, R. & Maccone, L. Using entanglement against noise in quantum metrology. *Physical review letters* **113**, 250801 (2014).

[29] Yuan, H. Sequential feedback scheme outperforms the parallel scheme for hamiltonian parameter estimation. *Physical Review Letters* **117**, 160801 (2016).

[30] Sekatski, P., Skotiniotis, M., Kołodyński, J. & Dür, W. Quantum metrology with full and fast quantum control. *arXiv:1603.08944* (2016).

[31] Helstrom, C. W. *Quantum detection and estimation theory* (Academic press, 1976).

[32] Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Physical Review Letters* **72**, 3439 (1994).

[33] Holevo, A. S. *Probabilistic and statistical aspects of quantum theory*, vol. 1 (Springer Science & Business Media, 2011).

[34] Fujiwara, A. Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *Journal of Physics A Mathematical General* **39**, 12489–12504 (2006).

[35] Fujiwara, A. & Imai, H. A fibre bundle over manifolds of quantum channels and its application to quantum statistics. *Journal of Physics A: Mathematical and Theoretical* **41**, 255304 (2008).

[36] Kolodynski, J. *Precision bounds in noisy quantum metrology*. Ph.D. thesis, University of Warsaw (2014).

[37] Walls, D. F. & Milburn, G. J. *Quantum optics* (Springer Science & Business Media, 2007).

[38] Herrera-Martí, D. A., Gefen, T., Aharonov, D., Katz, N. & Retzker, A. Quantum error-correction-enhanced magnetometer overcoming the limit imposed by relaxation. *Physical review letters* **115**, 200501 (2015).

[39] Viola, L., Knill, E. & Lloyd, S. Dynamical decoupling of open quantum systems. *Physical Review Letters* **82**, 2417 (1999).

[40] Uhrig, G. S. Keeping a quantum bit alive by optimized π-pulse sequences. *Physical Review Letters* **98**, 100504 (2007).

[41] Maze, J. *et al.* Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature* **455**, 644–647 (2008).

[42] Biercuk, M. J. *et al.* Optimized dynamical decoupling in a model quantum memory. *Nature* **458**, 996–1000 (2009).

[43] Demkowicz-Dobrzański, R., Czajkowski, J. & Sekatski, P. Adaptive quantum metrology under general markovian noise. *arXiv:1704.06280* (2017).

[44] Boyd, S. & Vandenberghe, L. *Convex optimization* (Cambridge university press, 2004).

[45] Grant, M. & Boyd, S. Cvx: Matlab software for disciplined convex programming. URL http://cvxr.com/cvx/.

SUPPLEMENTARY METHODS

Linear Scaling of the QFI

Here we prove that the QFI scales linearly with the evolution time t in the case where the HNLS condition is violated. We follow the proof in Ref. [30], which applies when the probe is a qubit, and generalize their proof to the case where the probe is d-dimensional.

We approximate the quantum channel

$$\mathcal{E}_{dt}(\rho) = \rho - i\omega[G,\rho]dt + \sum_{k=1}^{r}(L_k\rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k,\rho\})dt + O(dt^2) \tag{58}$$

by the following one:

$$\tilde{\mathcal{E}}_{dt}(\rho) = \sum_{k=0}^{r} K_k\rho K_k^\dagger, \tag{59}$$

where $K_0 = I + \big(-i\omega G - \frac{1}{2}\sum_{k=1}^{r} L_k^\dagger L_k\big)dt$ and $K_k = L_k\sqrt{dt}$ for $k \geq 1$. The approximation is valid because the distance between $\mathcal{E}_{dt}$ and $\tilde{\mathcal{E}}_{dt}$ is $O(dt^2)$ and the sensing time is divided into $\frac{t}{dt}$ segments, meaning the error $O(\frac{t}{dt}\cdot dt^2) = O(tdt)$ introduced by this approximation in calculating the QFI vanishes as $dt \to 0$. Next we calculate the operators $\alpha_{dt} = (\dot{\boldsymbol{K}} - ih\boldsymbol{K})^\dagger(\dot{\boldsymbol{K}} - ih\boldsymbol{K})$ and $\beta_{dt} = i(\dot{\boldsymbol{K}} - ih\boldsymbol{K})^\dagger\boldsymbol{K}$ for the channel $\tilde{\mathcal{E}}_{dt}(\rho)$, and expand these operators as a power series in $\sqrt{dt}$:

$$\begin{aligned}\alpha_{dt} &= \alpha^{(0)} + \alpha^{(1)}\sqrt{dt} + \alpha^{(2)}dt + O(dt^{3/2}),\\ \beta_{dt} &= \beta^{(0)} + \beta^{(1)}\sqrt{dt} + \beta^{(2)}dt + \beta^{(3)}dt^{3/2} + O(dt^2).\end{aligned} \tag{60}$$

We will now search for a Hermitian matrix h that sets low-order terms in each power series to zero.

Expanding h as $h = h^{(0)} + h^{(1)}\sqrt{dt} + h^{(2)}dt + h^{(3)}dt^{3/2} + O(dt^2)$ in $\sqrt{dt}$, and using the notation $(K_0, K_1, \cdots, K_r)^T = \boldsymbol{K} = \boldsymbol{K}^{(0)} + \boldsymbol{K}^{(1)}dt^{1/2} + \boldsymbol{K}^{(2)}dt$, we find

$$\alpha^{(0)} = \boldsymbol{K}^{(0)\dagger}h^{(0)}h^{(0)}\boldsymbol{K}^{(0)} = \sum_{k=0}^{r}\left|h_{0k}^{(0)}\right|^2 I = 0 \Longrightarrow h_{0k}^{(0)} = 0,\ 0 \leq k \leq r. \tag{61}$$

Therefore $h^{(0)}\boldsymbol{K}^{(0)} = \boldsymbol{0}$ and $\alpha^{(1)} = \beta^{(0)} = 0$ are automatically satisfied. Then,

$$\beta^{(1)} = -\boldsymbol{K}^{(0)\dagger}h^{(1)}\boldsymbol{K}^{(0)} = -h_{00}^{(1)}I = 0 \Longrightarrow h_{00}^{(1)} = 0. \tag{62}$$

and

$$\begin{aligned}\beta^{(2)} &= i\dot{\boldsymbol{K}}^{(2)\dagger}\boldsymbol{K}^{(0)} - \boldsymbol{K}^{(1)\dagger}h^{(0)}\boldsymbol{K}^{(1)} - \boldsymbol{K}^{(0)\dagger}h^{(1)}\boldsymbol{K}^{(1)} - \boldsymbol{K}^{(1)\dagger}h^{(1)}\boldsymbol{K}^{(0)} - \boldsymbol{K}^{(0)\dagger}h^{(2)}\boldsymbol{K}^{(0)}\\ &= G - \sum_{k,j=1}^{r} h_{jk}^{(0)}L_k^\dagger L_j - \sum_{k=1}^{r}(h_{0k}^{(1)}L_k + h_{k0}^{(1)}L_k^\dagger) - h_{00}^{(2)}I,\end{aligned} \tag{63}$$

which can be set to zero if and only if G is a linear combination of $I, L_k, L_k^\dagger$ and $L_k^\dagger L_j$ $(0 \leq k, j \leq r)$.

In addition,

$$\begin{aligned}\beta^{(3)} &= -\boldsymbol{K}^{(1)\dagger}h^{(1)}\boldsymbol{K}^{(1)} - \boldsymbol{K}^{(0)\dagger}h^{(2)}\boldsymbol{K}^{(1)} - \boldsymbol{K}^{(1)\dagger}h^{(2)}\boldsymbol{K}^{(0)} - \boldsymbol{K}^{(0)\dagger}h^{(3)}\boldsymbol{K}^{(0)}\\ &= -\sum_{k,j=1}^{r} h_{jk}^{(1)}L_k^\dagger L_j - \sum_{k=1}^{r}(h_{0k}^{(2)}L_k + h_{k0}^{(2)}L_k^\dagger) - h_{00}^{(3)}I = 0\end{aligned} \tag{64}$$

can be satisfied by setting the above parameters (which do not appear in the expressions for $\alpha^{(0,1)}$ and $\beta^{(0,1,2)}$) all to zero (other terms in $\beta^{(3)}$ are zero because of the constraints on $h^{(0)}$ and $h^{(1)}$ in Eq. (61) and Eq. (62)). Therefore, when G is a linear combination of $I, L_k, L_k^\dagger$ and $L_k^\dagger L_j$, there exists an h such that $\alpha_{dt} = O(dt)$ and $\beta_{dt} = O(dt^2)$ for the quantum channel $\tilde{\mathcal{E}}_{dt}$; therefore the QFI obeys

$$\mathcal{F}(\rho(t)) \leq 4\frac{t}{dt}\|\alpha_{dt}\| + 4\left(\frac{t}{dt}\right)^2\|\beta_{dt}\|\left(\|\beta_{dt}\| + 2\sqrt{\|\alpha_{dt}\|}\right) = 4\|\alpha^{(2)}\|t + O(\sqrt{dt}), \tag{65}$$

in which $\alpha^{(2)} = (h^{(1)}\boldsymbol{K}^{(0)} + h^{(0)}\boldsymbol{K}^{(1)})^\dagger(h^{(1)}\boldsymbol{K}^{(0)} + h^{(0)}\boldsymbol{K}^{(1)})$ under the constraint $\beta^{(2)} = 0$.