

Quantum Error-Corrected Non-Markovian Metrology

Zachary Mann^{1,2,*}, Ningping Cao^{3,1,4}, Raymond Laflamme^{1,2,3}, and Sisi Zhou^{3,2,1,5,†}

¹*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

²*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

³*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

⁴*Digital Technologies National Research Council Collaboration Center, East Campus 5, Waterloo, Ontario N2L 3G1, Canada*

⁵*Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*



(Received 2 April 2025; revised 10 June 2025; accepted 1 July 2025; published 5 August 2025)

Quantum metrology aims to maximize measurement precision on quantum systems, with a wide range of applications in quantum sensing. Achieving the Heisenberg limit (HL)—the fundamental precision bound set by quantum mechanics—is often hindered by noise-induced decoherence, which typically reduces achievable precision to the standard quantum limit (SQL). While quantum error correction (QEC) can recover the HL under Markovian noise, its applicability to non-Markovian noise remains less explored. In this work, we analyze a hidden Markov model (HMM) in which a quantum probe, coupled to an inaccessible environment, undergoes joint evolution described by Lindbladian dynamics, with the inaccessible degrees of freedom serving as a memory. We derive generalized Knill-Laflamme conditions for the HMM and establish three types of sufficient conditions for achieving the HL under non-Markovian noise using QEC. Additionally, we demonstrate the attainability of the SQL when these sufficient conditions are violated, by analytical solutions for special cases and numerical methods for general scenarios. Our results not only extend prior QEC frameworks for metrology but also provide new insights into precision limits under realistic noise conditions.

DOI: [10.1103/wfyl-wtz3](https://doi.org/10.1103/wfyl-wtz3)

I. INTRODUCTION

Quantum metrology aims to estimate the parameters of quantum systems to the highest attainable precision. Experiments usually consist of preparing a suitable probe state that then interacts with the system, before finally measuring the probe state to obtain information on the parameter. Applications of quantum metrology include gravitational-wave detection [1], high-precision atomic clocks [2], and interferometry [3]. Other promising platforms are nitrogen-vacancy centers in diamond for magnetic [4,5] and electric [6] field sensing, and superconducting technologies for magnetic flux sensing [7,8].

An important result in quantum metrology is the quantum Cramér-Rao bound [9], the quantum generalization

of the classical bound of the same name. For a parameter ω that we wish to estimate, it bounds the variance for any unbiased estimator of ω in terms of the number of experiments M and a quantity $\mathcal{F}$ called the quantum Fisher information (QFI), which is a function of the final state of the probes before measurement:

$$(\Delta\hat{\omega})^2 \geq \frac{1}{M\mathcal{F}}. \quad (1)$$

This bound is saturated in the limit of many experiments ($M \gg 1$) by maximum-likelihood estimators [10]. In this limit, the problem of minimizing the variance of parameter estimation is therefore equivalent to maximizing the quantum Fisher information. If we consider the task of estimating a parameter of a Hamiltonian, then the laws of quantum mechanics place a fundamental bound on the scaling of the QFI. This fundamental limit is called the Heisenberg limit (HL) and states that the QFI scaling can be no greater than a constant times T^2 , where T is the exposure time to the Hamiltonian. Equivalently, the scaling must be no greater than a constant times N^2 , where N is the number of probes used in a given experiment. A typical strategy for obtaining HL scaling in the number of probes is entangling N probes in a GHZ state. In an experimental

*Contact author: zmann@uwaterloo.ca

†Contact author: sisi.zhou26@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

setting, decoherence due to noise often imposes a tighter bound on the scaling of the QFI. In such cases, the best achievable scaling is linear in N and can be achieved with N independent probes. This limit is called the standard quantum limit (SQL).

One proposal for surpassing the SQL despite the presence of noise is to use quantum error correction (QEC) [11–19]. QEC is a protocol that protects quantum information from noise by encoding the information into many physical degrees of freedom, building in redundancy. This redundancy allows for the detection and correction of unwanted operations, or errors, on our state due to interaction with the environment. The principal application of QEC is quantum computation, where computations are performed on logical degrees of freedom encoded in a large number of physical degrees of freedom. It has been shown that QEC can be used to obtain HL scaling in the presence of noise so long as the noise obeys certain assumptions [16,17]. One of the principal assumptions is that the noise has to be Markovian. In this work, we generalize the results of Ref. [16] and derive conditions for achieving the HL using QEC in the presence of non-Markovian noise.

There are many different ways of quantifying Markovianity in a quantum system. Examples of indicators of quantum non-Markovianity are noncompletely positive (CP) divisibility [20–22], information backflow [22–24], and the tensor-product structure of the process tensor [25,26]. Non-Markovian metrology has been studied in many different contexts. It has been studied with non-Markovian bosonic environments [27,28], where the non-Markovianity and the ability to perform fast measurement allow QFI scaling as $N^{3/2}$, called the Zeno limit. It has further been studied for Ramsey spectroscopy [29,30], for sensing magnetic fields [31], for nonequilibrium environments [32], and for optimizing the QFI using process tensors and quantum combs [33–36]. Quantum error correction for non-Markovian noise, modeled by quantum combs, has previously been explored in Ref. [37].

Non-Markovian noise is an important source of decoherence in many experimental quantum sensing platforms. In superconducting circuits, a major source of decoherence is coupling between qubits and two-level quantum states in the environment. These two-level systems are coherent with non-negligible coherence times [38–40], allowing them to serve as a memory mediating temporal correlations in the noise. Further, in many practical cases, few two-level systems are required to accurately model the noise [41–43]. Some theoretical models attribute the ubiquitous $1/f$ noise spectrum in solid-state systems to the presence of two-level systems [44,45].

In, e.g., nitrogen-vacancy centers in diamond, a major source of decoherence is coupling between the electron of the center and the surrounding carbon or nitrogen nuclear spins. The surrounding spins are coherent [46], allowing

them to serve as a memory. The noise in nitrogen-vacancy center experiments can also be accurately modeled using a relatively small number of these environmental spins [47,48].

A. Setup

We consider a quantum system consisting of a probe and inaccessible degrees of freedom that we call the environment. Let $\mathcal{H}_P$ denote the d_P -dimensional Hilbert space of the probe and let $\mathcal{H}_E$ represent the d_E -dimensional Hilbert space of the environment. Further, we suppose that we have access to noiseless auxiliary degrees of freedom. Let $\mathcal{H}_A$ denote the Hilbert space of the auxiliary system. The total state of the quantum system is denoted by ρ , where $\rho \in \mathcal{H}_E \otimes \mathcal{H}_P \otimes \mathcal{H}_A$. In this work, we will consider the case of finite-dimensional systems $d_E, d_P < \infty$. The state of the probe-and-auxiliary system is given by the reduced density matrix $\rho_{PA} = \text{Tr}_E(\rho)$, obtained by tracing out the environment. We suppose that the state ρ evolves in time according to a Lindblad master equation of the form

$$\begin{aligned} \frac{d\rho}{dt} = & -i[H_{EP}, \rho] - i[\omega\tilde{G}, \rho] - i[\omega\tilde{H}_E, \rho] \\ & + \sum_{k=1}^r \left(L_k \rho L_k^\dagger - \frac{1}{2} \{L_k^\dagger L_k, \rho\} \right). \end{aligned} \quad (2)$$

We suppose that $\tilde{G}$ is of the form $\mathbb{1}_E \otimes G \otimes \mathbb{1}_A$; in other words, G only acts on the probe degrees of freedom. Similarly, we suppose that $\tilde{H}_E$ is of the form $H_E \otimes \mathbb{1}_P \otimes \mathbb{1}_A$ and only acts on the environment. Further, we assume that G is nontrivial; i.e., G is not proportional to $\mathbb{1}_P$. We have that G , H_E , and H_{EP} are Hermitian operators. The r different L_k operators are called the jump operators, also known as Lindblad operators. H_{EP} and the L_k act solely on the probe and environment degrees of freedom, acting trivially on the auxiliary system.

The real parameter ω is what we wish to estimate. Both $\omega\tilde{G}$ and ωG will be referred to as the signal Hamiltonian acting on the probe, or simply the signal on the probe. The Hamiltonian ωH_E corresponds to the evolution of the environment dependent on the parameter ω or the signal on the environment. In this model, we suppose that the signal acts independently on probe and environment degrees of freedom.

The H_{EP} and L_k terms correspond to the undesired noise affecting the system. The effect of H_{EP} is referred to as the unitary coupling. The jump operators correspond to the dissipative portion of the noise. These operators only act nontrivially on the environment and probe degrees of freedom and act trivially on the auxiliary degrees of freedom. In this model, we suppose that the noise does not depend on the signal parameter ω .

The model is Markovian when considering the system and environment together, as it evolves according to a

Lindblad master equation. However, only the probe and auxiliary degrees of freedom are accessible experimentally. The time evolution of their reduced state exhibits non-Markovian behavior. For this reason, we refer to this model as a *hidden Markov model* (HMM) [25,49]. It is similar to the *hidden quantum Markov model* studied in the literature [50], the principal difference here being that we consider the reduced density matrix at different times rather than measurement. When we only have access to the dynamics of the probe-and-auxiliary system, the model displays non-Markovianity for all of the aforementioned indicators. In particular, the local dynamics may not be CP divisible and the probe may experience information backflow from the environment.

For each experiment, the environment, probe, and auxiliary system will be reset. As mentioned above, we will denote the number of experiments by M . For a given experiment, we may perform arbitrary control operations on the probe-and-auxiliary system, while leaving the environment untouched. The degrees of freedom of the environment can therefore mediate temporal correlations in the noise perceived by the probe. We will denote by N the number of times we directly measure the probe with a full-rank projective measurement. We will denote by t the exposure time of each probe to the Hamiltonian between two measurements. During this time, we allow the application of nondestructive control operations to the probe-and-auxiliary system. The total experiment time is given by MT , where $T = Nt$. Note that we can also defer all measurements to the end of the evolution by shelving the probe state and swapping in a fresh probe from the auxiliary degrees of freedom. The goal is to obtain conditions for HL scaling in T . We illustrate these different timescales in Fig. 1.

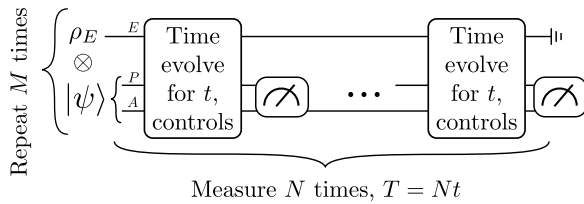


FIG. 1. An illustration of the different timescales in a sensing experiment. In a given experiment, the probe-and-auxiliary system is measured N times, each capturing the signal and the noise for a duration t . At this timescale, the environment is untouched. We repeat such an experiment M times, where at this timescale ($T = Nt$) the environment is reset. Note that the QEC procedure is applied fast and frequently, viewed as quantum controls during the time- t evolution. Although it may involve additional usage of auxiliary systems and measurements, they are not depicted explicitly here for simplicity. The auxiliary system included in the figure only represents noiseless auxiliary system that serve as part of the quantum error-correcting code (QECC).

Finally, we assume access to an unlimited supply of noiseless auxiliary degrees of freedom, enabling the generation of perfect copies of any desired probe state. In experiment, this assumption has been well approximated by considering an electron state in a nitrogen-vacancy sensor as the probe and by using addressable ^{13}C carbon nuclear spins, which are more robust to noise and have longer coherence times, as the auxiliary system [18].

Additionally, we assume that all control operations on both the probes and auxiliary system can be performed with arbitrary speed and precision. For error correction to be feasible, the coherence times of our probe must greatly exceed the duration of control operations. As for the non-Markovianity, it has been shown that for certain superconducting systems interacting with environmental two-level states, the coherence times of the two-level states are of the same order as those of the qubit [38–40,51,52]. This would allow for non-Markovian effects to develop across many rounds of control operations.

B. Summary of results and outline

Consider the HMM. Let $\{|\phi_i\rangle\}_{i=1}^{d_E}$ be an orthonormal basis for $\mathcal{H}_E$. Let $\mathcal{S}^\circ$ be the linear span of the operators $\mathbb{1}_P$, $\langle\phi_i|L_k|\phi_i\rangle$, $\langle\phi_i|L_k^\dagger|\phi_i\rangle$, and $\langle\phi_i|L_k^\dagger|\phi_i\rangle\langle\phi_m|L_j|\phi_m\rangle$ ($\forall k, j \in \{1, 2, \dots, r\}$, $\forall i, m \in \{1, 2, \dots, d_E\}$). Let $\mathcal{S}$ be the linear span of $\mathcal{S}^\circ$, where we add the operators $\langle\phi_i|H_{EP}|\phi_m\rangle$, $\forall i, m \in \{1, \dots, d_E\}$. Let $\mathcal{S}_{EP}$ be the linear span of $\mathbb{1}_{EP}$, L_k , $L_k^\dagger$, $L_k^\dagger L_j$ $\forall k, j$. Finally, we call an interaction diagonal if it is of the form presented in Definition 2. Using these definitions, we summarize the results of the paper in Table I.

The paper is organized as follows. In Sec. II, we consider quantum error correction for the HMM by adapting the Knill-Laflamme (KL) conditions.

In Sec. III, the attainability of the HL using QEC is explored. The results of this section are presented in Fig. 2. In Sec. III A, the QEC conditions of Sec. II are used to generalize the HNLS conditions of Ref. [16] to the HMM. The derived “Hamiltonian not in extended span” (HNES) conditions (Theorem 1), are sufficient conditions for achieving HL scaling in the presence of non-Markovian noise. The rightmost (red) section of Fig. 2 corresponds to when

TABLE I. Summary of results.

Theorem	Conditions	QFI scaling
1	$G \notin \mathcal{S}$	$\mathcal{F} \propto T^2$
2	$\tilde{H}_E = 0, L_k = 0 \forall k$	$\limsup_{t \rightarrow \infty} \mathcal{F}/T^2 > 0$
3	Diagonal interaction, $G \notin \mathcal{S}^\circ$	$\limsup_{t \rightarrow \infty} \mathcal{F}/T^2 > 0$
4	$\tilde{H}_E + \tilde{G} \in \mathcal{S}_{EP}$	$\mathcal{F} = O(T)$
5	Diagonal interaction	$\mathcal{F} = \Omega(T)$

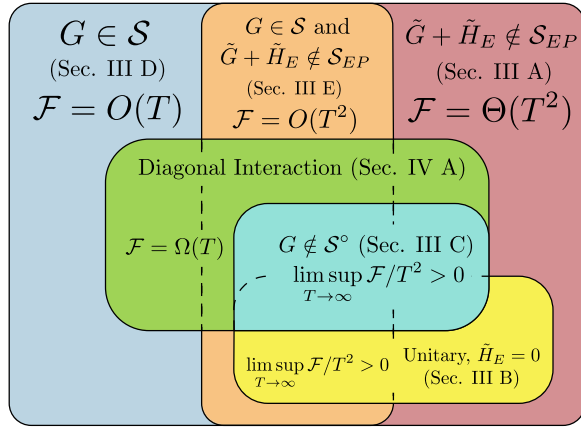


FIG. 2. The best achievable asymptotic scaling of the quantum Fisher information (QFI) for different overlaps of the signal and the noise in the hidden Markov model (HMM). The Heisenberg-limit (HL) scaling in the rightmost region is guaranteed by the “Hamiltonian not in extended span” (HNES) conditions $G \notin \mathcal{S}$ (see Sec. III A). The standard quantum limit (SQL) bound in the leftmost region is guaranteed by the “Hamiltonian not in Lindblad span” (HNLS) conditions (see Sec. III D). Scaling for the middle section depends on the trajectory of the evolution (see Sec. III E). For the diagonal interaction, the scaling is lower bounded by the SQL (see Sec. IV A), or if the “Hamiltonian not in extended Lindblad span” (HNELS) are satisfied (Sec. III C), then the scaling is HL up to an envelope function, as in the unitary case (see Sec. III B). A list of acronyms is presented in Appendix F.

these conditions are satisfied. The error-correction protocol considered in this section is designed to eliminate all noise acting on the probe. As a result, the information about the parameter contained in the environment cannot be exploited for potential gains in sensing.

In Sec. III B, we consider a special case of the model for which the overall dynamics are unitary and the signal only affects the probe ($\tilde{H}_E = 0$) and show that HL scaling can be achieved up to an almost periodic envelope function. This special case corresponds to the yellow bubble of Fig. 2.

In Sec. III C, we consider a special case of the model for which operators partially commute. In this case, a weaker set of conditions than HNES, which we call the “Hamiltonian not in extended Lindblad span” (HNELS) conditions, guarantee that HL scaling can be achieved with QEC up to an almost periodic envelope function, similarly to the unitary case. This special case corresponds to the teal bubble in the middle of Fig. 2. In both the unitary and HNELS cases, non-Markovianity plays a crucial role. A weaker QEC code is employed—one that does not correct the unitary coupling noise term, which can temporarily decohere the probe. However, since this coupling is unitary, we are guaranteed that information will flow back into the system from the environment, allowing for HL scaling.

In Sec. III D, the HNLS conditions of Ref. [16] are applied to rule out the HL in certain cases. This is depicted in the leftmost (blue) section of Fig. 2.

In Sec. III E, an example is presented to prove that the HNLS conditions are sufficient but not necessary. This example illustrates how, in certain cases, the initial environment state must be taken into account in order to derive conditions for HL versus SQL scaling. This corresponds to the middle (orange) section of Fig. 2.

The attainability of SQL scaling in T is considered in Sec. IV. That section is organized as follows. In Sec. IV A, we consider models in which the operators partially commute and show that a combination of QEC and a prepare-and-measure strategy achieves the SQL in T . This is depicted by the green bubble in the center of Fig. 2. Section IV B considers a model in which the operators commute completely, in which case a prepare-and-measure strategy without QEC achieves SQL scaling in T . Once again, we ask the question of whether the SQL is achievable despite the presence of noise. We do not investigate the potential advantages of non-Markovianity for such a strategy. An example of this is the Zeno limit [27,28], where the non-Markovianity allows for $N^{3/2}$ QFI scaling, between SQL and HL.

In Sec. IV C, we consider a prepare-and-measure strategy without QEC for more general models, which do not necessarily have commuting operators. The Fisher information (FI) is computed numerically, providing evidence for generic SQL scaling in the case in which $\tilde{H}_E = 0$. In the more general case of $\tilde{H}_E \neq 0$, the numerical results are inconclusive. Numerical results are presented for a Heisenberg-model-type interaction and for randomly generated master equations on two qubits.

II. QEC FOR HMMs

A. Non-Markovian QEC

We begin by considering QEC for HMMs. We generalize the standard definition of a quantum error-correcting code to one that allows for errors in both the probe and the environment, while preserving the logical information in the probe through local operations only. This can be viewed as a subsystem code where we treat the environment as the gauge degrees of freedom and the probe as the logical subsystem. The difference is that we restrict operations to act solely on the logical subsystem. We prove the following lemma to be the generalized KL conditions for the HMM.

Before deriving error-correction conditions for the HMM, we must first clarify what exactly is meant by an error-correcting code. We do so in the following definition.

Definition 1. Let U be an isometry from $\mathbb{C}^k$ to $\mathbb{C}^{d_{PA}}$ for some integers $k \leq d_{PA}$. Let $\mathcal{E}$ be a set of linear operators $\{E : \mathbb{C}^{d_{PA}d_E} \rightarrow \mathbb{C}^{d'_{PA}d'_E}\}$, where d_E and d'_E are the

dimensions of the environment before and after application of a linear operator. We will say that $(U, \mathcal{E})$ is a quantum error-correcting code (QECC) for $\mathcal{E}$ if there exists a completely positive trace-preserving (CPTP) map $D : \mathcal{B}(\mathbb{C}^{d_{\text{PA}}}) \rightarrow \mathcal{B}(\mathbb{C}^k)$ such that $\forall |\psi\rangle \in \mathbb{C}^k, \forall |\phi\rangle \in \mathbb{C}^{d_E}$ and $\forall E \in \mathcal{E}$, we have that

$$(D \otimes \mathbb{1}_{d_E}) (E(U \otimes \mathbb{1}_{d_E}) |\psi\rangle \langle \psi| \otimes |\phi\rangle \langle \phi| (U^\dagger \otimes \mathbb{1}_{d_E}) E^\dagger) = c_{E,\psi,\phi} |\psi\rangle \langle \psi| \otimes \alpha_{E,\psi,\phi}, \quad (3)$$

where $\alpha_{E,\psi,\phi}$ is a density matrix in $\mathbb{C}^{d_E} \times \mathbb{C}^{d_E}$.

If an error E satisfies Eq. (3), we say that the error is correctable. Here, $\mathbb{C}$ denotes the complex numbers and $\mathcal{B}(\mathbb{C}^d)$ denotes the set of linear operators on the vector space $\mathbb{C}^d$.

QECCs are usually characterized using the KL conditions [53], which are both necessary and sufficient. Lemma 1 presents similar conditions for QECCs for HMMs.

Lemma 1 (KL Conditions for the HMM). Let U be an isometry from $\mathbb{C}^k$ to $\mathcal{H}_{\text{PA}} = \mathcal{H}_P \otimes \mathcal{H}_A$, of dimension d_{PA} , and let $\mathcal{E}$ be a set of linear operators mapping $\mathcal{H}_E \otimes \mathcal{H}_{\text{PA}}$ to $\mathcal{H}_{E'} \otimes \mathcal{H}_{\text{PA}'}$. Let $P = UU^\dagger$ be the code-space projector and let $\{|i\rangle_E\}_{i=1}^{d_E}$ ($\{|i\rangle_{E'}\}_{i=1}^{d_{E'}}$) be an orthonormal basis for $\mathcal{H}_E$ ($\mathcal{H}_{E'}$). We have that $(U, \mathcal{E})$ is a QECC if and only if we have

$$P |l\rangle_E E_i^\dagger |m\rangle_{E'} \langle n|_{E'} E_j |k\rangle_E P = c_{ij}^{lmnk} P, \quad (4)$$

for all $E_i, E_j \in \mathcal{E}$, where the c_{ij}^{lmnk} are scalar constants.

The proof of Lemma 1 can be found in Appendix A. The key insight of the proof is that by defining the set of extended errors,

$$\tilde{E}_{i,n,k} = \langle n|_{E'} E_i |k\rangle_E, \quad (5)$$

the non-Markovian KL conditions reduce to the standard KL conditions on this extended error set. This allows us to reuse all of the error-correcting machinery, including encoding maps and recovery operations, from the standard theory.

Consider the case in which $d_{\text{PA}} = d'_{\text{PA}}$ and $d_E = d'_E$, which captures our model. The KL conditions and Definition 1 ensure that the state is always a product state on $\mathcal{H}_E \otimes \mathcal{H}_{\text{PA}}$. Therefore, the time evolution of the reduced state on $\mathcal{H}_{\text{PA}}$ between two rounds of error correction can always be represented by a CPTP map that depends on the residual states on $\mathcal{H}_E$. The KL conditions ensure that the Kraus operators of this map will be correctable for any choice of the state on $\mathcal{H}_E$. Further, implementing the error-correction protocol requires no knowledge of the environment state.

Note that this error-correcting condition for a QECC is perhaps too strong. In principle, the capabilities to correct errors induced by any state on $\mathcal{H}_E$ are not necessary; one needs only to correct the errors induced by the states on $\mathcal{H}_E$ occurring throughout the evolution of the system. This distinction will be important when we consider quantum metrology. We obtain sufficient but not necessary conditions for the HL for this reason.

Finally, an important remark is that even when perfect QEC fails, the recovery operation still brings the state back to the code space. This will be important for the derivations of Secs. III C and IV A, where even if QEC does not completely correct the noise, it tailors it to a specific form and restricts it to the code space.

Other non-Markovian quantum error-correction conditions exist in the literature. In Ref. [37], the authors derive necessary and sufficient conditions for quantum error correction for quantum combs, which they call the *strategic code formalism*. The model that they consider is very general and encapsulates the HMM defined in Sec. IA. However, since we are considering a simpler model, we can derive conditions using the operators of the master equation directly, rather than considering different contractions of the quantum comb with the QEC control operations.

B. A small example

We will now consider an example of quantum error correction in a spin system to highlight the differences between the Markovian and non-Markovian cases. We will consider a system in which $\mathcal{H}_P$ consists of an ensemble of individually addressable spin-1/2 degrees of freedom (or qubits) and the environment $\mathcal{H}_E$ consists of inaccessible spin-1 degrees of freedom. For simplicity, we will assume that each spin-1/2 particle interacts with its own bath of spin-1 particles. In the non-Markovian case, we will assume that this bath consists of a single coherent spin. We will suppose that the coupling Hamiltonian between a spin-1/2 particle and the i th spin-1 particle in the system is given by

$$H_{\text{EP},i} = \tau_i (S_z^{(i)} \otimes Z + S_x^{(i)} \otimes X),$$

where Z and X are the Pauli matrices acting on the spin-1/2 particle and $S_x^{(i)}, S_z^{(i)}$ are the spin-1 operators acting on the i th spin-1. τ_i is the coupling strength. We have the following:

$$S_z^{(i)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad S_x^{(i)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let $|-1\rangle_E, |0\rangle_E$, and $|1\rangle_E$ be the $-1, 0$, and 1 eigenstates of S_z , respectively. We will suppose that each spin-1 has a

Hamiltonian term of the form

$$H_{E,i} = \Delta (S_z^{(i)})^2,$$

where Δ is a positive strength constant. The total Hamiltonian for a spin-1/2 and its bath is given by

$$H_{\text{tot.}} = \sum_i (H_{E,i} + H_{\text{EP},i}).$$

We first consider the Markovian regime. In this regime, we assume that each spin-1/2 interacts with many spin-1 degrees of freedom. Following the prescription of Ref. [20], we make two approximations. We first suppose that the spin-1/2 degrees of freedom couple weakly to the bath of environmental spins, such that the bath is weakly affected by the interaction ($\Delta \gg \tau_i$). Further, we suppose that there is a dissipative master equation acting on the environment such that the spin-1 degrees of freedom are driven to the $|0\rangle_E$ steady state. A Markovian master equation can then be derived for the spin-1/2 degrees of freedom. This is done in Appendix B. This master equation generates a noise channel with Kraus operators proportional to $\mathbb{1}$ and X . Therefore, a repetition code [54,55] that only corrects X errors is sufficient to protect the logical information.

We next consider the non-Markovian regime. In this regime, we suppose that the spin-1/2 degrees of freedom each interact with a single coherent spin-1 degree of freedom. For simplicity, we suppose that there is no dissipation. In the non-Markovian regime, since the environment state evolves in time, we must consider the set of extended errors

$$K_{ij} = \langle i|_E e^{-itH_{\text{tot.}}} |j\rangle_E,$$

where $|i\rangle_E, |j\rangle_E \in \{|-1\rangle_E, |0\rangle_E, |1\rangle_E\}$. Since we consider arbitrarily fast control, we can consider the errors to first order in t . We obtain the operators

$$\begin{aligned} K_{0,0} &= \mathbb{1}, \quad K_{j,j} = \mathbb{1} - it\Delta\mathbb{1} - jt\tau Z, \\ K_{-1,0} &= K_{0,-1} = K_{1,0} = K_{0,1} = \frac{-i\tau t}{\sqrt{2}} X. \end{aligned}$$

We see that to protect the logical information, a repetition code no longer suffices, due to the presence of Z -rotation errors. Because of the additional error operators introduced by the non-Markovianity, a more capable code is required.

C. Computation versus metrology

There are fundamental differences between the tasks of quantum sensing and quantum computation that must be taken into account when designing a QECC. The principal difference relates to the logical operators, which are

the operators that act nontrivially on the code words without being detected. In quantum computation, the goal is to design codes with logical operators that do not resemble the noise. Logical operators are engineered so that the errors induced by the environment have a very low probability of corrupting the logical information. There is much freedom in engineering these logical operators, allowing for the design of complex codes with very high code distance.

In quantum sensing, we wish to do the opposite. We wish to design codes for which the interaction with the environment due to the signal is projected to a nontrivial logical operator. Contrary to the case of computation, these interactions are not freely chosen. There is a lot less freedom when engineering a QEC protocol for quantum sensing. A trade-off between preserving signal and eliminating noise must be achieved.

As previously shown, the non-Markovian quantum error-correction conditions for the HMM* can be rephrased as the standard KL conditions acting on an extended error set defined by Eq. (A4). In other words, the non-Markovianity of the system introduces a larger variety of errors. Consider the previous example of each qubit interacting with a spin-1 particle. The non-Markovianity changes the form of the errors to either X or Z operators rather than just X operators.

Due to the restricted design freedom, this problem is much more harmful in quantum sensing. Consider the same example, in which we wish to use the spin-1/2 particles to detect a magnetic field. Suppose that the magnetic field acts independently on each spin as a Z operator. We therefore design a code for which Z is a logical operator. In the Markovian case, the errors are of the form X and are therefore distinguishable from Z , and so a repetition code does the job. In the non-Markovian case, we also have Z errors. These errors have the same form as the magnetic field, and so the logical information will always be corrupted no matter which types of QECC are used.

To show why this may not be an issue for computation, we consider the example of a stabilizer code. Stabilizer codes have three important parameters: the number of physical qubits n , the number of logical qubits k encoded in the physical qubits, and finally the distance d . A distance- d stabilizer code can correct error operators that can be written as products of $\lfloor \frac{d-1}{2} \rfloor$ arbitrary single-qubit errors [56]. For unbiased noise, the non-Markovianity will therefore be problematic for quantum error correction if the memory effects in the system cause the errors to be high weight. When this is not the case and when the non-Markovian noise is also local, it has been shown that a fault-tolerance threshold is still possible [57,58]. In the previous example, the non-Markovianity does not affect the weight of the errors.

For the experimental platforms and corresponding noise sources previously mentioned, this is unlikely to be a

problem. For example, in superconducting qubits, one expects the interaction with the two-level chip defects to be relatively local. Even though the memory and coherence of the defect may change the form of the error, it will not affect the weight of the error. Major sources of error in superconducting experiments have instead been crosstalk or high-weight errors due to cosmic rays [59,60].

For quantum sensing and metrology, however, non-Markovianity has the potential to be much more destructive, as we will see below.

III. ACHIEVING THE HL

A. The HNLS conditions for achievability of the HL

In Ref. [16], the authors present sufficient and necessary conditions, called “Hamiltonian not in Lindblad span” (HNLS), for achieving the HL asymptotically. When these conditions are met, the probe and a noiseless auxiliary system can be encoded into a specialized error-correcting code. For this code, all of the noise becomes correctable errors and a portion of the signal remains as a logical operation. This logical component of the signal leads to HL scaling within the code space. However, the HNLS conditions were originally derived for Markovian noise, where no environmental degrees of freedom are present, and the master equation applies solely to the probe. Further, these conditions are necessary to achieve the HL, as noncorrectable errors inevitably lead to complete decoherence of the system and the signal over long timescales.

We generalize these conditions to the HMM. The HNLS conditions were derived under the assumption that error correction can be performed perfectly and arbitrarily fast. We will keep the same assumptions.

Theorem 1 (“Hamiltonian Not in Extended Span” (HNES) Conditions). Let $\{|\phi_i\rangle\}_{i=1}^{d_E}$ be an orthonormal basis for $\mathcal{H}_E$. Let $\mathcal{S}$ be the linear span of the operators $\mathbb{1}_P$, $\langle\phi_i|H_{EP}|\phi_m\rangle$, $\langle\phi_i|L_k|\phi_m\rangle$, $\langle\phi_i|L_k^\dagger|\phi_m\rangle$ and $\langle\phi_i|L_k^\dagger|\phi_m\rangle\langle\phi_n|L_j|\phi_l\rangle$ ($\forall k, j \in \{1, \dots, r\}$, where r is the number of jump operators; $i, m, n, l \in \{1, \dots, d_E\}$).

If $G \notin \mathcal{S}$, then HL scaling can be achieved for the parameter ω . The condition $G \notin \mathcal{S}$ will be referred to as the “Hamiltonian not in extended span” (HNES) conditions. The extended span refers to the possible errors that can act on the reduced state of the probe for arbitrary choices of environment state.

Proof. The proof generalizes the HNLS conditions for Heisenberg-limit scaling from [16], with modifications to accommodate the HMM. It will therefore be very similar to the original proof, with some slight changes. The evolution of the state ρ for a duration Δt is given by the first-order

expansion of the master equation

$$\begin{aligned} \Phi_{\Delta t}(\rho) = & \rho + \left(-i[H_{EP}, \rho] - i[\omega\tilde{G}, \rho] - i[\omega\tilde{H}_E, \rho]\right) \Delta t \\ & + \Delta t \sum_{k=1}^r \left(L_k \rho L_k^\dagger - \frac{1}{2} \{L_k^\dagger L_k, \rho\}\right) + O(\Delta t^2). \end{aligned} \quad (6)$$

To reduce clutter, we use the following notation:

$$A^{(ij)} := \langle\phi_i|A|\phi_j\rangle, \quad (7)$$

where A is a linear operator on $\mathcal{H}_E \otimes \mathcal{H}_P$. This leaves us with an operator acting on P . Let $\mathcal{S}_H$ denote the subspace of Hermitian matrices in $\mathcal{S}$, which is spanned by $L_k^{(ij)} + L_k^{(ij)\dagger}$, $i(L_k^{(ij)} - L_k^{(ij)\dagger})$, $(L_k^\dagger L_m)^{(ij)} + \text{h.c.}$, $i((L_k^\dagger L_m)^{(ij)} - \text{h.c.})$, $i(H_{EP}^{(ij)} - H_{EP}^{(ij)\dagger})$, and $H_{EP}^{(ij)} + \text{h.c.}$ We remark that $\mathcal{S}_H \subseteq \mathcal{S}$. G is Hermitian and can therefore be decomposed into $G = G_{\parallel} + G_{\perp}$, where $G_{\parallel} \in \mathcal{S}_H$ and $G_{\perp} \perp \mathcal{S}_H$. The assumption that $G \notin \mathcal{S}$ implies that $G_{\perp}$ is nonzero. $G_{\perp}$ must have trace zero in order to be orthogonal to $\mathbb{1}_P$ under the trace inner product. The spectral decomposition for $G_{\perp}$ is therefore

$$G_{\perp} = \frac{1}{2} (\text{tr}|G_{\perp}|) (\rho_0 - \rho_1),$$

where $|G_{\perp}| := \sqrt{G_{\perp}^2}$ and ρ_0 and ρ_1 are density matrices. We consider $|C_0\rangle$ and $|C_1\rangle$ to be normalized purifications of ρ_0 and ρ_1 on $\mathcal{H}_P \otimes \mathcal{H}_A$, where $\mathcal{H}_A$ consists of auxiliary degrees of freedom that we append to our system. We further construct $|C_0\rangle$ and $|C_1\rangle$ to have orthonormal supports on $\mathcal{H}_A$. To be explicit, assuming that $\rho_0 = \sum_{i=1}^{d_1} (\rho_0)_{ii} |i\rangle_P \langle i|_P$ and $\rho_1 = \sum_{i=d_1+1}^{d_P} (\rho_1)_{ii} |i\rangle_P \langle i|_P$ are in the diagonalized forms,

$$\begin{aligned} |C_0\rangle &= \sum_{i=1}^{d_1} \sqrt{(\rho_0)_{ii}} |i\rangle_P |i\rangle_A, \\ |C_1\rangle &= \sum_{i=d_1+1}^{d_P} \sqrt{(\rho_1)_{ii}} |i\rangle_P |i\rangle_A, \end{aligned}$$

where $\{|i\rangle_P\}$ and $\{|i\rangle_A\}$ are orthonormal bases of the probe and the auxiliary system, respectively, where the auxiliary system is assumed to be as large as the probe. The supports of ρ_0 and ρ_1 are $\text{span}\{|i\rangle_P\}_{i=1}^{d_1}$ and $\text{span}\{|i\rangle_P\}_{i=d_1+1}^{d_P}$, respectively. This will be the code space, with the code-space

projector given by

$$P = |C_0\rangle\langle C_0| + |C_1\rangle\langle C_1|. \quad (8)$$

The orthogonality of the states on the auxiliary system space implies that

$$\langle C_0|B \otimes \mathbb{1}_A|C_1\rangle = \langle C_1|B \otimes \mathbb{1}_A|C_0\rangle = 0, \quad (9)$$

where B is any linear operator on $\mathcal{H}_P$. We also use the relation

$$\text{tr}(|C_0\rangle\langle C_0| - |C_1\rangle\langle C_1|)(B \otimes \mathbb{1}_A) = \frac{2 \text{tr}(G_\perp B)}{\text{tr}|G_\perp|}. \quad (10)$$

Although a smaller auxiliary system, such as a qubit, is enough to guarantee Eq. (9), it would not be enough to guarantee Eq. (10), which usually requires d_A dimensions.

Equation (10) is zero for any $B \in \mathcal{S}$, allowing us to write

$$\langle C_0|B \otimes \mathbb{1}_A|C_0\rangle = \langle C_1|B \otimes \mathbb{1}_A|C_1\rangle.$$

These equations imply that

$$\begin{aligned} PL_k^{(ij)}P &= \lambda_k^{ij}P, \\ P\left(L_k^\dagger L_m\right)^{(ij)}P &= \mu_{km}^{ij}P, \end{aligned}$$

where this equation is due to the fact that $\left(L_k^\dagger L_m\right)^{(ij)}$ is in the linear span of $L_k^{(ij)\dagger}L_{k'}^{(i'j')}$. By assumption on the span $\mathcal{S}$, we also have that

$$PL_k^{(ij)}L_m^{(i'j')}P = \zeta_{km}^{ij i'j'}P. \quad (11)$$

Further,

$$PH_{\text{EP}}^{(ij)}P = \chi^{ij}P. \quad (12)$$

We see from Eq. (10) that

$$PGP \neq \text{constant} \times P. \quad (13)$$

Let $\rho = \rho_E \otimes \rho_{\text{PA}}$ be a state in the code space, i.e., $P\rho_{\text{PA}}P = \rho_{\text{PA}}$. Let $\rho_E = \sum_{i=0}^{d_E} \alpha_i |\phi_i\rangle\langle\phi_i|$. We can take ρ_E to be diagonal, without loss of generality, since the choice of basis $\{|\phi_i\rangle\}_{i=1}^{d_E}$ is arbitrary. We will use P and $\mathbb{1}_E \otimes P$ interchangeably—it will be clear from the context which

operator we are referring to:

$$\begin{aligned} P\Phi_{\Delta t}(\rho)P &= \rho - i\omega \left[P\tilde{G}P, \rho \right] \Delta t - i\omega [H_E, \rho_E] \otimes \rho_{\text{PA}} \Delta t \\ &\quad + \left(\sum_i \alpha_i \sum_{k=1}^r A_{ik} \otimes \rho_{\text{PA}} \right) \Delta t + O(\Delta t^2), \end{aligned}$$

where

$$\begin{aligned} A_{ik} &:= \sum_{i'j'} \lambda_k^{i'i} \left(\lambda_k^{j'i} \right)^* |\phi_{i'}\rangle\langle\phi_{j'}| \\ &\quad + \sum_{i'} \left((\mu_{kk}^{i'i} - i\chi^{i'i}) |\phi_{i'}\rangle\langle\phi_i| + \text{h.c.} \right). \end{aligned}$$

We also consider the projector $P_\perp = \mathbb{1} - P$. We have that

$$\begin{aligned} P_\perp \Phi_{\Delta t}(\rho)P_\perp &= \Delta t \sum_i \alpha_i \sum_{k=1}^r B_{ik} (|\phi_i\rangle\langle\phi_i| \otimes \rho_{\text{PA}}) B_{ik}^\dagger + O(\Delta t^2), \end{aligned} \quad (14)$$

where $B_{ik} = \sum_j |\phi_j\rangle\langle\phi_i| \otimes (L_k^{(j,i)} - \lambda_k^{ji} \mathbb{1}_{\text{PA}})$. After measurement and neglecting Δt^2 terms, the state is given by

$$\begin{aligned} P\Phi_{\Delta t}(\rho)P + P_\perp \Phi_{\Delta t}(\rho)P_\perp &= \rho - i\omega \left[P\tilde{G}P, \rho \right] \Delta t - i\omega [H_E, \rho_E] \otimes \rho_{\text{PA}} \Delta t \\ &\quad + \left(\sum_i \alpha_i \sum_{k=1}^r A_{ik} \otimes \rho_{\text{PA}} \right) \Delta t \\ &\quad + \Delta t \sum_i \alpha_i \sum_{k=1}^r B_{ik} (|\phi_i\rangle\langle\phi_i| \otimes \rho_{\text{PA}}) B_{ik}^\dagger. \end{aligned}$$

Except for the final term, all noise contributions can be effectively suppressed by fast repeated measurements, which project the system onto the code space—an instance of the quantum Zeno effect. However, the operators B_{ik} in the final term act on the probe, introducing dephasing that cannot be mitigated by measurement alone. Hence, measurement is not enough and the recovery operation is required here as well.

Equation (11) is exactly that of the generalized KL conditions for the HMM in Lemma 1, where we take the errors to be the linear operators acting on the state in Eq. (14). Therefore, there exists a recovery operation $\mathcal{R}$ such that

$$\mathcal{R}(P_\perp \Phi_{\Delta t}(\rho)P_\perp) = \Delta t \sigma_E \otimes \rho_{\text{PA}} + O(\Delta t^2), \quad (15)$$

where $\Delta t \sigma_E$ is a residual operator supported on the environment degrees of freedom after applying the recovery operation to Eq. (15).

Combining Eqs. (16) and (15), we obtain the final state after QEC:

$$\begin{aligned} & P\Phi_{\Delta t}(\rho)P + \mathcal{R}(P_{\perp}\Phi_{\Delta t}(\rho)P_{\perp}) \\ &= \rho - i\omega \left[P\tilde{G}P, \rho \right] \Delta t - i\omega [H_E, \rho_E] \otimes \rho_{PA} \Delta t \\ &+ \Delta t \sigma'_E \otimes \rho_{PA} + O(\Delta t^2), \end{aligned} \quad (16)$$

where $\sigma'_E = \sigma_E + \sum_i \alpha_i \sum_{k=1}^r A_{ik}$. $\sigma'_E \otimes \rho_{PA}$ can also be written as

$$\begin{aligned} & \sigma'_E \otimes \rho_{PA} \\ &= \mathcal{D} \left(-i[H_{EP}, \rho] + \sum_{k=1}^r \left(L_k \rho L_k^\dagger - \frac{1}{2} \{ L_k^\dagger L_k, \rho \} \right) \right), \end{aligned} \quad (17)$$

where $\mathcal{D}(\cdot) = P \cdot P + \mathcal{R}(P_{\perp} \cdot P_{\perp})$. Since the argument of the right-hand side of Eq. (17) has trace zero, then σ'_E must have trace zero due to the fact that $\mathcal{D}$ is a CPTP map and that $\text{Tr}(\rho_{PA}) = 1$. We trace out the environment to obtain the reduced dynamics on the probe, which are unitary,

$$\begin{aligned} & \text{tr}_E[\mathcal{R}(P_{\perp}\Phi_{\Delta t}(\rho)P_{\perp}) + P\Phi_{\Delta t}(\rho)P] \\ &= \rho_{PA} - i\omega [P(G \otimes \mathbb{1}_A)P, \rho_{PA}] \Delta t. \end{aligned} \quad (18)$$

We also neglect the second-order terms in Δt . From here on out, the proof is equivalent to the one presented in Ref. [16]. We summarize the final steps. Let λ_0 and λ_1 be the eigenvalues of $P(G \otimes \mathbb{1}_A)P$ associated with the eigenstates $|C_0\rangle$ and $|C_1\rangle$, respectively. We can see that the code words are eigenstates of the effective Hamiltonian from Eq. (9). By choosing an initial state of

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|C_0\rangle + |C_1\rangle)$$

on the probe-and-auxiliary system, one can show that the QFI scales as

$$\mathcal{F}(\rho_{PA}(t)) = t^2 (\lambda_0 - \lambda_1)^2.$$

Equation (13) guarantees that this QFI is nonzero for $t > 0$. We therefore have that if the HNES conditions are satisfied, the QFI scales as a constant multiplied by t^2 and so we obtain the HL in T . ■

Note that the above QEC protocol to achieve the HL assumes that the control operations implementing the protocol are performed fast and frequently throughout the state evolution. It corresponds to the situation in Fig. 1, in which the probe is only measured at the end of the evolution where $N = 1$ and $T = t$.

Further, although the above protocol yields HL scaling, it is not optimal in general. In particular, the constant factors in the expression of the QFI may be suboptimal. To achieve better scaling, the code may be optimized in the same manner as in Ref. [17].

Finally, from the point of view of information backflow, this error-correction protocol eliminates all non-Markovianity in the system. Since the QEC protocol ensures that the system and environment are always in a product state, information backflow from the environment is not possible. Although this prevents potential correlations in the system from introducing harmful effects, it also prevents any potential sensing advantages due to the non-Markovianity, such as, e.g., using the environment as a memory protected from decoherence [40].

B. Unitary evolution

We now consider the special case of when the evolution of the system is unitary. We will also consider the additional assumption of the signal acting only on the probe, $\tilde{H}_E = 0$. In this scenario, the master equation governing the evolution of the HMM [Eq. (2)] reduces to

$$\frac{d\rho}{dt} = -i[H_{EP}, \rho] - i[\omega\tilde{G}, \rho]. \quad (19)$$

When the HNES conditions are satisfied, the QFI scales as t^2 . For unitary evolution with the signal acting only on the probe, we can guarantee that the QFI scales as a constant times t^2 multiplied by some periodic envelope function, even when the HNES conditions are not satisfied. Once again, let $N = 1$. In particular, we can show that $\limsup_{t \rightarrow \infty} \mathcal{F}/t^2 > 0$ when $d_E < \infty$. We can therefore estimate the parameter with t^{-2} precision as long as we make sure to measure the system at the times when the value of the envelope is above some nonzero constant.

Theorem 2. Suppose that the evolution of the system is unitary and that the signal does not act on the environment; in other words, the evolution of the HMM is generated by Eq. (19). Suppose that $\mathcal{H}_E$ has finite dimension, i.e., $d_E < \infty$. Under these assumptions, there is a QEC protocol that achieves $\limsup_{t \rightarrow \infty} \mathcal{F}/t^2 > 0$. More precisely, it is possible to achieve a QFI that scales as t^2 multiplied by an almost periodic envelope function. By measuring at correct times, we recover the HL.

Proof. We begin by defining a QEC code in the same manner as in the proof of Theorem 1, with the extended span $\mathcal{S}$ replaced by $\text{span}\{\mathbb{1}\}$. $|C_0\rangle$ and $|C_1\rangle$ in this case correspond to the positive and negative parts of $G - \frac{\text{tr}(G)}{d_P} \mathbb{1}_P$ (the eigenstates with eigenvalues zero can be assigned in either the positive or negative parts without affecting the calculations), the supports of which are denoted by projectors P_0 and P_1 . $P_0 + P_1 = \mathbb{1}_P = \mathbb{1}_A$. To guarantee HL

scaling with this QEC code, it is not necessary to act with the recovery operation. Fast and accurate syndrome measurements will be sufficient. The QEC serves to transform the noise and enforce commutation with the signal, facilitating the QFI computation.

Note that since our only controls consist only of rapid measurement, the protocol can be seen as an application of the quantum Zeno effect. By measuring fast and frequently, we freeze the dynamics relating to any noise that is orthogonal to the nontrivial component of the signal.

For an initial state in the code space, the resulting dynamics are governed by the master equation

$$\begin{aligned} \frac{d\rho}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{P\Phi_{\Delta t}(\rho)P + P_{\perp}\Phi_{\Delta t}(\rho)P_{\perp}}{\Delta t} \\ &= -i[PH_{\text{EP}}P, \rho] - i[\omega P\tilde{G}P, \rho], \end{aligned} \quad (20)$$

where P is the projection operator onto the code space. Expanding the evolution to first order, we obtain

$$\begin{aligned} P\Phi_{\Delta t}(\rho)P + P_{\perp}\Phi_{\Delta t}(\rho)P_{\perp} \\ = \rho - i[P(\omega\tilde{G} + H_{\text{EP}})P, \rho]\Delta t + O(\Delta t^2). \end{aligned}$$

We neglect the second-order terms in Δt . This master equation corresponds to unitary evolution generated by the following effective Hamiltonian:

$$H_{\text{eff}} = \omega P\tilde{G}P + PH_{\text{EP}}P.$$

First, we show that the two terms of the Hamiltonian commute. Since, from Eq. (9), $PB_{\text{EP}}P = P_0B_{\text{EP}}P_0 + P_1B_{\text{EP}}P_1$ for any operator B_{EP} acting on the probe and the environment, we have

$$\begin{aligned} [PH_{\text{EP}}P, P\tilde{G}P] \\ = [P_0H_{\text{EP}}P_0, P_0\tilde{G}P_0] + [P_1H_{\text{EP}}P_1, P_1\tilde{G}P_1], \end{aligned} \quad (21)$$

where $P_0 = |C_0\rangle\langle C_0|$ and $P_1 = |C_1\rangle\langle C_1|$. Since $\tilde{G}$ only has support on the probe degrees of freedom, we have that $[\langle C_i|\tilde{G}|C_i\rangle, \langle C_j|H_{\text{EP}}|C_j\rangle] = 0 \forall i, j$. Therefore, the terms cancel and we obtain

$$[PH_{\text{EP}}P, P\tilde{G}P] = 0.$$

We will now show that the HL is achieved. We prepare the following state on the probe and auxiliary system:

$$|\psi_{\text{PA}}\rangle = \frac{1}{\sqrt{2}}(|C_0\rangle + |C_1\rangle),$$

where $|C_0\rangle$ and $|C_1\rangle$ are the code words of the error-correcting code. Let $|\psi_E\rangle$ be an arbitrary state on the environment degrees of freedom. We can take $|\psi_E\rangle$ to be pure,

without loss of generality, by considering a purification of the environment. The total state of the system is

$$|\Psi_{\text{PAE}}\rangle = |\psi_E\rangle \frac{1}{\sqrt{2}}(|C_0\rangle + |C_1\rangle). \quad (22)$$

We can rewrite this state as a sum over the joint eigenstates of $PH_{\text{EP}}P \otimes \mathbb{1}_A$ and $\mathbb{1}_E \otimes P\tilde{G}P \otimes \mathbb{1}_A$:

$$|\Psi_{\text{EPA}}\rangle = \sum_{k=1}^{d_E} c_k^0 |\psi_k^0\rangle |C_0\rangle + \sum_{l=1}^{d_E} c_l^1 |\psi_l^1\rangle |C_1\rangle, \quad (23)$$

where the joint eigenstates of these operators are necessarily product states between the code words and states $|\psi_i^j\rangle$ (defined below) on the environment. Below, we will drop the tensor products with identity on the environment and auxiliary system from the notation to reduce clutter; it will be clear from the context which operators act on which subsystems. To prove that Eq. (23) follows, we know from Eq. (9) that

$$\langle C_0|PH_{\text{EP}}P|C_1\rangle = \langle C_0|P\tilde{G}P|C_1\rangle = 0. \quad (24)$$

This allows us to write the terms of the effective Hamiltonian in the following form:

$$\begin{aligned} PH_{\text{EP}}P &= H_{\text{EP}}^0 \otimes |C_0\rangle\langle C_0| + H_{\text{EP}}^1 \otimes |C_1\rangle\langle C_1|, \\ P\tilde{G}P &= \lambda_0 \mathbb{1} \otimes |C_0\rangle\langle C_0| + \lambda_1 \mathbb{1} \otimes |C_1\rangle\langle C_1|, \end{aligned}$$

where λ_0 and λ_1 are the (nontrivial) eigenvalues of $P\tilde{G}P$ as defined in the previous section and H_{EP}^0 and H_{EP}^1 are operators acting on the environment. From this form, it is clear that the joint eigenvectors of $PH_{\text{EP}}P$ and $P\tilde{G}P$ are the tensor products between the eigenvectors of the blocks H_{EP}^i and the projectors $|C_i\rangle\langle C_i|$.

We now consider the time evolution for our chosen probe state. Since the terms of the Hamiltonian commute, we can independently consider their generated time evolution. We begin with the signal

$$\begin{aligned} e^{-i\omega P\tilde{G}Pt} |\Psi_{\text{EPA}}\rangle &= \sum_{k=1}^{d_E} c_k^0 e^{-i\omega\lambda_0 t} |\psi_k^0\rangle |C_0\rangle \\ &\quad + \sum_{l=1}^{d_E} c_l^1 e^{-i\omega\lambda_1 t} |\psi_l^1\rangle |C_1\rangle. \end{aligned}$$

For the second term, we obtain

$$\begin{aligned} |\Psi_{\text{EPA}}(t)\rangle &= e^{-itPH_{\text{EP}}P} e^{-itP\omega\tilde{G}P} |\Psi_{\text{EPA}}\rangle \\ &= \sum_{k=1}^{d_E} c_k^0 e^{-it\omega\lambda_0} e^{-it\psi_k^0} |\psi_k^0\rangle |C_0\rangle \\ &\quad + \sum_{l=1}^{d_E} c_l^1 e^{-it\omega\lambda_1} e^{-it\psi_l^1} |\psi_l^1\rangle |C_1\rangle, \end{aligned}$$

where ψ_i^j is the eigenvalue of $PH_{\text{EP}}P$ associated with the eigenstate $|\psi_i^j\rangle |C_i\rangle$. We trace out the environment to obtain the reduced state of the probe and auxiliary system:

$$\rho_{\text{PA}} = \begin{pmatrix} \frac{1}{2} & e^{-i\omega t(\lambda_0 - \lambda_1)} \alpha(t) \\ e^{i\omega t(\lambda_0 - \lambda_1)} \alpha(t)^* & \frac{1}{2} \end{pmatrix}.$$

We obtain a 2×2 matrix as the state is contained in the two-dimensional code space, which is spanned by the states $|C_0\rangle$ and $|C_1\rangle$. We know that the diagonals are $\frac{1}{2}$ since the initial state had amplitudes of $\frac{1}{\sqrt{2}}$ in $|C_0\rangle$ and $|C_1\rangle$, which are both eigenstates of $P\tilde{G}P$ and $PH_{\text{EP}}P$. Therefore, the probabilities of measurement in that basis do not change. The value of α is given by

$$\alpha = \sum_{k=1}^{d_E} \sum_{l=1}^{d_E} c_k^0 c_l^{1*} e^{-it(\psi_k^0 - \psi_l^1)} \langle \psi_l^1 | \psi_k^0 \rangle.$$

Let $\Delta\lambda = \lambda_0 - \lambda_1$. The QFI of a mixed state can be calculated using the following equation from Ref. [61]:

$$\mathcal{F}(\rho) = 2 \sum_{i,j} \frac{|\langle \gamma_i | \partial_\omega \rho | \gamma_j \rangle|^2}{\gamma_i + \gamma_j}, \quad (25)$$

where γ_i and $|\gamma_i\rangle$ are the eigenvalue and eigenvector pairs of ρ and $\partial_\omega \rho$ is the derivative of ρ with respect to the parameter ω . The sum is taken over pairs i, j such that $\gamma_i + \gamma_j > 0$.

One can show, by direct calculation, that the eigenvalues of ρ_{PA} are given by

$$\gamma_\pm = \frac{1}{2} \pm |\alpha(t)|. \quad (26)$$

Since the reduced density matrix has positive eigenvalues between zero and one, we can deduce that $|\alpha(t)| \in [-\frac{1}{2}, \frac{1}{2}]$. We rewrite $\alpha(t)$ in the polar form as $\alpha(t) = |\alpha(t)| e^{-i\phi_\alpha(t)}$. We will drop the explicit t dependence from

the notation to reduce clutter. Further, the eigenvectors of ρ_{PA} are given by

$$|\gamma_\pm\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm e^{i\omega t \Delta\lambda - i\phi_\alpha} \end{pmatrix}. \quad (27)$$

Inserting these values into Eq. (25) gives

$$\mathcal{F}(\rho_{\text{PA}}) = 4t^2 (\Delta\lambda)^2 |\alpha(t)|^2.$$

To show the HL, we need to show that

$$\limsup_{t \rightarrow \infty} |\alpha| = \text{constant}.$$

We already know that $\limsup_{t \rightarrow \infty} |\alpha| \leq \frac{1}{2}$ from the earlier-derived condition. We can rewrite α by letting $\bar{c}_m = c_k^0 c_l^{1*} \langle \psi_l^1 | \psi_k^0 \rangle$ and $\phi_m = \psi_k^0 - \psi_l^1$, where $m = (k, l)$ runs over the indices of l and k . The equation for $\alpha(t)$ becomes

$$\alpha(t) = \sum_m \bar{c}_m e^{-it\phi_m}, \quad (28)$$

where m runs over d_E^2 different values. It is important to note that the $\bar{c}_m$ do not depend on t . For $t = 0$, we have $\rho_{\text{PA}} = |\psi_{\text{PA}}\rangle \langle \psi_{\text{PA}}|$, and so $\alpha(0) = \frac{1}{2}$. Our goal is to prove

$$\limsup_{t \rightarrow \infty} |\alpha(t)| = \frac{1}{2}. \quad (29)$$

This condition is very similar to quantum state revival in a closed quantum system, which is guaranteed by the Poincaré recurrence theorem [62–64]. One can prove it by factoring the time evolution to separate signal and noise, and considering the recurrence times of the noise Hamiltonian. In fact, the revival of the QFI is directly implied from the theory of almost periodicity [64,65], which we argue below. For completeness, we further provide an explicit proof of Eq. (29) in Appendix C.

We remark that $\alpha(t)$ is a uniformly convergent Fourier series, as it contains finite d_E^2 terms, indicating its almost periodicity [63–65]. Then, $\alpha(t)$, as an almost periodic function, satisfies $\forall \epsilon > 0$ and $\forall t_0$, there exists an infinite many values of τ such that

$$|\alpha(t_0 + \tau) - \alpha(t_0)| < \epsilon. \quad (30)$$

The HL is achievable so long as we measure the system at proper times at which $\alpha(t)$ is above a certain positive constant. Further, these values of τ are well spread out between $-\infty$ and $+\infty$ such that one cannot construct arbitrarily long intervals between two values. The almost periodicity implies that if we pick a random t from a sufficiently large interval, there is a nonzero probability that $|\alpha(t)|$ is above a positive constant. Therefore, even

if we have no knowledge on the environmental state, we can still find a proper t that achieves the HL by picking randomly. ■

Similar to Sec. III A, the above protocol achieves $\limsup_{T \rightarrow 0} \mathcal{F}/T^2 > 0$ taking $T = t$ and $N = 1$ in Fig. 1, where QEC is performed fast and frequently and no intermediate measurements on the probe system are performed.

As mentioned previously, measuring at random times will yield HL scaling; however, the constant factors in the scaling will likely be suboptimal. In practice, it is very difficult to compute the exact revival times for the system, even with knowledge of the dynamics. Instead, one can probe the revival times experimentally to calibrate the sensing protocol. This can be done with, e.g., a Ramsey experiment. If the revival times are too long, then it may be better to consider the HNES conditions and eliminate all interactions, if possible.

It is worth noting that for systems with small environment dimensions and simple couplings, such as the previously discussed superconducting qubits interacting with two-level systems and nitrogen-vacancy centers interacting with nearby spins, the envelope functions are relatively simple. Further, the revivals have a short period [38,39,46,66], making them easy to probe.

C. Diagonal interactions and the HNELS conditions

We begin by defining the diagonal interaction. This is a special case of the model in which the components of operators supported on the environment degrees of freedom all commute pairwise.

Definition 2 (Diagonal Interaction). We call an HMM in Eq. (2) an HMM with diagonal interactions if and only if there exists an orthonormal basis $\{|\phi_i\rangle\}_{i=1}^{d_E}$ in $\mathcal{H}_E$ such that

$$H_E = \sum_i h_E^i |\phi_i\rangle \langle \phi_i| \otimes \mathbb{1}_{PA}, \quad (31)$$

$$H_{EP} = \sum_i |\phi_i\rangle \langle \phi_i| \otimes H_{EP}^{(i,i)}, \quad (32)$$

$$L_k = \sum_i |\phi_i\rangle \langle \phi_i| \otimes L_k^{(i,i)}, \quad \forall k. \quad (33)$$

For HMMs with diagonal interactions, the quantum dynamics can be simplified through the following lemma.

Lemma 2. Consider a quantum state ρ as an input state to an HMM with diagonal interactions. The output state

after evolution of time t is $\Phi_t(\rho)$, where $\Phi_t(\cdot)$ denotes the evolution channel after time t . We have

$$\text{tr}_E(\Phi_t(\rho)) = \text{tr}_E(\Phi_t((\mathcal{D}_E \otimes \mathbb{1}_{PA})(\rho))), \quad (34)$$

where $\mathcal{D}_E$ is the completely dephasing channel on $\mathcal{H}_E$ such that $\mathcal{D}_E(\cdot) = \sum_k |\phi_k\rangle \langle \phi_k| (\cdot) |\phi_k\rangle \langle \phi_k|$. In addition, Eq. (34) still holds when quantum controls are applied on the probe-and-auxiliary system during the evolution.

Proof. To see that Eq. (34) holds, we observe that it is sufficient to prove the commuting relation between $\mathcal{D}_E$ and Φ_t , because

$$\text{tr}_E(\Phi_t(\rho)) = \text{tr}_E((\mathcal{D}_E \otimes \mathbb{1}_{PA})(\Phi_t(\rho))). \quad (35)$$

The fact that $\mathcal{D}_E \otimes \mathbb{1}_{PA}$ commutes with Φ_t can be proven through noting that $\mathcal{D}_E \otimes \mathbb{1}_{PA}$ commutes with the differential in Eq. (2) for diagonal interactions. Finally, $\mathcal{D}_E \otimes \mathbb{1}_{PA}$ also commutes with any quantum operations acting on the probe-and-auxiliary system because they act on different systems. ■

Lemma 2 implies that for diagonal interactions, it is always sufficient to consider input states that are diagonal in the environmental basis $\{|\phi_i\rangle\}_{i=1}^d$ and the off-diagonal terms will not contribute to the final measurement outcomes at the probe and auxiliary system. Below, we will assume the quantum states under consideration $\rho \in \mathcal{H}_E \otimes \mathcal{H}_P \otimes \mathcal{H}_A$ are diagonal in $\mathcal{H}_E$, i.e.,

$$\rho = \sum_i \alpha_i |\phi_i\rangle \langle \phi_i| \otimes \rho^{(i,i)}, \quad (36)$$

where the $\rho^{(i,i)}$ are density operators on $\mathcal{H}_{PA}$ and the α_i are non-negative constants summing up to one. As we are considering HMMs with diagonal interactions, only $\rho^{(i,i)}$ will evolve over time under the master equation and it is sufficient to track the evolution of $\rho^{(i,i)}$.

Further, when studying diagonal interactions, we can ignore the contribution from $\tilde{H}_E$ as it will not make a difference to the QFI.

Lemma 3. Consider a quantum state ρ in the form of Eq. (36) as the input state to an HMM with diagonal interactions. We have that

$$\text{tr}_E(\Phi_t(\rho)) = \text{tr}_E(\hat{\Phi}_t(\rho)), \quad (37)$$

where $\hat{\Phi}_t(\cdot)$ denotes the evolution channel after time t , where we ignore the contributions of $\tilde{H}_E$. In other words, it is the time evolution under the equation

$$\begin{aligned} \frac{d\rho}{dt} = & -i[H_{EP}, \rho] - i[\omega \tilde{G}, \rho] \\ & + \sum_{k=1}^r \left(L_k \rho L_k^\dagger - \frac{1}{2} \{L_k^\dagger L_k, \rho\} \right). \end{aligned} \quad (38)$$

Proof. First note that $\tilde{H}_E$ commutes with all other operators of the master equation. This allows us to factor the evolution as follows:

$$\Phi_t(\cdot) = \hat{\Phi}_t \left(e^{-i\tilde{H}_E}(\cdot) e^{i\tilde{H}_E} \right) = e^{-i\tilde{H}_E} \hat{\Phi}_t(\cdot) e^{-i\tilde{H}_E}.$$

For a ρ in the form of Eq. (36), time evolution from H_E has no effect,

$$\begin{aligned} e^{-i\tilde{H}_E} \rho e^{i\tilde{H}_E} &= \sum_i \alpha_i e^{-i\tilde{H}_E} |\phi_i\rangle \langle \phi_i| e^{i\tilde{H}_E} \otimes \rho_{PA} \\ &= \sum_i \alpha_i e^{-i\tilde{H}_E} |\phi_i\rangle \langle \phi_i| e^{i\tilde{H}_E} \otimes \rho_{PA} \\ &= \sum_i \alpha_i |\phi_i\rangle \langle \phi_i| \otimes \rho_{PA} = \rho, \end{aligned}$$

and so we see that the evolution due to $\tilde{H}_E$ has no effect and only the evolution due to $\hat{\Phi}_t$ must be considered. ■

When the interaction is diagonal, with less restrictive conditions than the HNES conditions, we can guarantee that the QFI scales as a constant times t^2 multiplied by some periodic envelope function and that $\limsup_{t \rightarrow \infty} \mathcal{F}/t^2 > 0$ when $d_E < \infty$. This is the same scaling as in the case of unitary evolution seen in Sec. III B and guaranteed by Theorem 2.

Theorem 3 (“Hamiltonian Not in Extended Lindblad Span” (HNELS) Conditions). Consider the case of a diagonal interaction. Let $\{|\phi_i\rangle\}_{i=1}^{d_E}$ be the orthonormal basis for $\mathcal{H}_E$ from Definition 2. Let $\mathcal{H}_E$ have finite dimension, i.e., $d_E < \infty$. Let $\mathcal{S}^\circ$ be the linear span of the operators $\mathbb{1}_P$, $\langle \phi_i | L_k | \phi_i \rangle$, $\langle \phi_i | L_k^\dagger | \phi_i \rangle$, and $\langle \phi_i | L_k^\dagger | \phi_i \rangle \langle \phi_m | L_j | \phi_m \rangle$ ($\forall k, j \in \{1, 2, \dots, r\}, \forall i, m \in \{1, 2, \dots, d_E\}$). $\mathcal{S}^\circ$ is similar to $\mathcal{S}$, except that we exclude the contributions from H_{EP} and only consider diagonal elements.

If $G \notin \mathcal{S}^\circ$, then $\limsup_{t \rightarrow \infty} \mathcal{F}/t^2 > 0$. Further, it is possible to achieve a QFI that scales as t^2 multiplied by a periodic envelope function. We will refer to the containment $G \notin \mathcal{S}^\circ$ as the “Hamiltonian not in extended Lindblad span” (HNELS) conditions.

Proof. We consider ρ in diagonal form [Eq. (36)] since we trace out the environment at the end and may thus apply Lemmas 2 and 3. For a ρ of this form, the master equation governing the HMM takes the form

$$\begin{aligned} \frac{d\rho}{dt} &= \sum_i \alpha_i |\phi_i\rangle \langle \phi_i| \otimes \left(-i\omega [G, \rho^{(i,i)}] - i[H_{EP}^{(i,i)}, \rho^{(i,i)}] \right. \\ &\quad \left. + \sum_{j=1}^r \left(L_j^{(i,i)} \rho^{(i,i)} (L_j^{(i,i)})^\dagger - \frac{1}{2} \left\{ (L_k^{(i,i)})^\dagger L_k^{(i,i)}, \rho^{(i,i)} \right\} \right) \right). \end{aligned}$$

We employ the same construction for a QEC code as in Sec. III A, except that we swap $\mathcal{S}$ for $\mathcal{S}^\circ$. For an initial state in the code space, the resulting dynamics are governed by the master equation

$$\begin{aligned} \frac{d\rho}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{P\Phi_{\Delta t}(\rho)P + \mathcal{R}(P_\perp \Phi_{\Delta t}(\rho)P_\perp)}{\Delta t} \\ &= \sum_i \alpha_i |\phi_i\rangle \langle \phi_i| \otimes \left(-i\omega [PGP, \rho^{(i,i)}] \right. \\ &\quad \left. - i[P H_{EP}^{(i,i)} P, \rho^{(i,i)}] + \sum_{j=1}^r \left(P L_j^{(i,i)} \rho^{(i,i)} (L_j^{(i,i)})^\dagger P \right. \right. \\ &\quad \left. \left. + \mathcal{R} \left(P_\perp L_j^{(i,i)} \rho^{(i,i)} (L_j^{(i,i)})^\dagger P_\perp \right) \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \left\{ P (L_k^{(i,i)})^\dagger L_k^{(i,i)} P, \rho^{(i,i)} \right\} \right) \right), \end{aligned} \quad (39)$$

where P is the projection operator onto the code space. Since $G \notin \mathcal{S}^\circ$, we have that PGP acts nontrivially in the code space. Further, by design of the QEC code, the last three terms of Eq. (39) give $\rho^{(i,i)}$ up to a scalar. This scalar must be zero since the jump-operator portion of the master equation has trace zero and the channel $P \cdot P + \mathcal{R}(P \cdot P)$ is trace preserving. Therefore, the evolution is given by

$$\frac{d\rho}{dt} = \sum_i \alpha_i |\phi_i\rangle \langle \phi_i| \otimes -i \left[P(\omega G + H_{EP}^{(i,i)}) P, \rho^{(i,i)} \right],$$

which can be rewritten as

$$\frac{d\rho}{dt} = -i \left[P \left(\omega \tilde{G} + \sum_i |\phi_i\rangle \langle \phi_i| \otimes H_{EP}^{(i,i)} \right) P, \rho \right]. \quad (40)$$

We see from Eq. (40) that after quantum error correction, for a diagonal interaction, the overall evolution of the system and environment is unitary. Application of Theorem 2 therefore yields the desired QFI scaling.

Note that in this proof, a first code has been used to obtain unitary dynamics on the full system. In order to prove the scaling of Theorem 2, the probe-and-auxiliary system have been encoded in a QECC that imposes commutation between the signal and coupling Hamiltonian. It is not necessary here to concatenate the two codes, as the first code used in the proof already has all the properties necessary to prove the HL scaling. ■

D. The HNLS conditions for when the HL is unattainable

Using the HNLS conditions of Ref. [16] on the full system, we can obtain conditions for when the HL is unattainable.

Theorem 4 (Unattainability of the HL by Violating the HNLS Conditions [16]). Let $\mathcal{S}_{\text{EP}}$ be the linear span of $\mathbb{1}_{\text{EP}}$, $L_k, L_k^\dagger, L_k^\dagger L_j \forall k, j$. If $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \in \mathcal{S}_{\text{EP}}$, then the SQL is the best achievable scaling. In other words, $\mathcal{F} = O(T)$.

The proof is as follows. The HNLS conditions of Ref. [16] tell us that if $\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P \in \mathcal{S}_{\text{EP}}$, then the SQL is the best achievable asymptotic scaling with controls that act on the full system, i.e., the probe, the auxiliary system, and the environment. In our analysis, we have restricted ourselves to control operations acting only on the probe-and-auxiliary system, which is a strict subset of the set of full system controls. Therefore, the SQL is necessarily the best achievable asymptotic scaling for the local controls.

Further, we can relate these conditions to the previously derived HNES conditions (Theorem 1). Consider

$$\begin{aligned} \mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P &\in \mathcal{S}_{\text{EP}} \\ \Rightarrow \mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P &= \sum_{A \in \mathcal{S}_{\text{EP}}} c_A A, \end{aligned}$$

where $c_A \in \mathbb{C}$. We simply trace out the environment to obtain the following:

$$G = \frac{1}{d_E} \sum_i \sum_{A \in \mathcal{S}_{\text{EP}}} c_A \langle \phi_i | A | \phi_i \rangle - \text{tr}(H_E) \mathbb{1}_P.$$

From this equation, we see that $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \in \mathcal{S}_{\text{EP}}$ implies $G \in \mathcal{S}$. We can obtain a simple yet useful result by taking the contrapositive to this statement. If $G \notin \mathcal{S}$, then $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \notin \mathcal{S}_{\text{EP}}$. These containment conditions and their relations to the scaling of the QFI are presented in Fig. 2.

E. Dependence of the QFI on the initial environment state

We have so far shown that when the HNES conditions are satisfied ($G \in \mathcal{S}$), the HL is achievable using QEC. We have further shown that $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \in \mathcal{S}_{\text{EP}}$ implies that the HL is not attainable, using the HNLS conditions. This leaves one scenario to study, the case in which $G \in \mathcal{S}$ but $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \notin \mathcal{S}_{\text{EP}}$.

When $G \in \mathcal{S}$ but $(\mathbb{1}_E \otimes G + H_E \otimes \mathbb{1}_P) \notin \mathcal{S}_{\text{EP}}$, the master equation [Eq. (2)] does not fully determine the scaling of the QFI. We must also take into account the initial state of the environment and the trajectory of the evolution to find the scaling. We show this with the following example, for which we can be limited by either HL or SQL depending on the initial environment state.

We wish to sense the parameter ω with signal Hamiltonian $\tilde{G} = \mathbb{1}_E \otimes Z$. We consider an environment and probe

that are both qubits, so that $\mathcal{H}_E \otimes \mathcal{H}_P = \mathbb{C}^2 \otimes \mathbb{C}^2$. Further, we assume no unitary coupling between the qubits, i.e., $H_{\text{EP}} = 0$, and no signal on the environment $H_E = 0$. We consider a model with the following two jump operators:

$$L_1 = |0\rangle\langle 0| \otimes \mathbb{1}, \quad L_2 = |1\rangle\langle 1| \otimes Z. \quad (41)$$

Here, Z refers to the Pauli σ_Z matrix. We first consider the situation in which the environment is initially in the $|0\rangle$ state, such that the initial density matrix is given by $\rho_{\text{EP}}(0) = |0\rangle\langle 0| \otimes \rho_{P,0}$, where $\rho_{P,0}$ is the initial probe state. One can see that throughout the evolution, the state will always remain in a product state. The environment remains unchanged and stays in the $|0\rangle\langle 0|$ state. We can therefore trace out the environment and look only at the dynamics on the probe. The reduced density matrix on the probe evolves under the equation

$$\frac{d\rho_{\text{PA}}}{dt} = -i\omega [G, \rho_{\text{PA}}]. \quad (42)$$

For this choice of initial state, the probe evolution is entirely generated by the signal Hamiltonian and the evolution is Markovian and unitary. This guarantees the achievability of the HL.

Similarly, consider the case in which the initial environment state is $|1\rangle$. Once again, the state always remains in a product state throughout the evolution and, in particular, the environment always remains in the $|1\rangle\langle 1|$ state and does not change. We can trace out the environment and look at the evolution of the probe. In this case, the probe evolves under the following master equation:

$$\frac{d\rho_{\text{PA}}}{dt} = -i\omega [G, \rho_{\text{PA}}] + Z\rho_{\text{PA}}Z - \rho_{\text{PA}}. \quad (43)$$

The reduced dynamics are Markovian and therefore the results of Ref. [16] can be applied. For these dynamics, the HL cannot be achieved using any quantum controls due to violation of the HNLS conditions. In this case, the SQL is the best achievable scaling since the signal Hamiltonian and the jump operator are both given by the same matrix.

IV. ACHIEVING THE SQL

We now consider the attainability of the SQL in the case in which the HNES conditions are violated. If the HNES conditions are violated but the HNLS conditions on the full system are satisfied, then, as seen in Sec. III E, the SQL can bound the QFI from above for certain initial states of the environment. Further, if the HNLS conditions are not satisfied, then we are also guaranteed an SQL upper bound (Sec. III D). The question remains of whether linear scaling in N is always achievable; in other words, if these bounds can always be saturated.

For the Markovian models considered in Ref. [16], the SQL can be achieved using a prepare-and-measure strategy. It consists of preparing the probe in a state $|\psi\rangle$, letting it evolve for a constant time t without applying any control operations, and then either measuring the state or shelving the state to replace it with a fresh probe. For this strategy in the Markovian regime, the system is completely reset after each measurement and the SQL in N is achievable due to the additivity of the QFI for product states. For the non-Markovian model, measurements on the probe do not reset the environment. This can lead to correlations between different probes, mediated by the environment, which in general can lead to higher or lower Fisher information for the distribution of measurement outcomes [67].

In this section, we investigate whether linear scaling in N is always achievable for the QFI. We first investigate this claim with an exactly solvable model and then treat the general case with numerical methods. We numerically compute the FI for a finite number of probes and then linearly extrapolate SQL scaling. The special case in which the probe and environment undergo unitary evolution has been considered in Ref. [68], where the authors show that a prepare-and-measure strategy achieves the SQL in N . We conjecture that the SQL remains achievable with this strategy when jump operators are added to the equation and when the signal acts only on the probe. We leave a proof of this claim for the general non-Markovian setting as an open problem.

A. Achieving the SQL for a diagonal interaction

We will consider a diagonal interaction as defined by Definition 2, as we did when considering the HNLS conditions of Sec. III C. We simplify the dynamics further by applying QEC on the probe-and-auxiliary system to impose commutation of interactions on the probe degrees of freedom as well, allowing us to directly compute a QFI lower bound in a basis in which all operators are simultaneously diagonal. To obtain the QFI lower bound, we will measure the probe at each time step and calculate the classical Fisher information (FI), which we denote by $\mathcal{I}(\omega)$, of the probability distribution for the measurement outcomes. As mentioned above, we will employ the prepare-and-measure strategy. The FI of the joint probability distribution is a lower bound for the QFI of the quantum state of all output probes [10]. Consequently, observing SQL scaling in the FI ensures a QFI scaling no less than SQL. In our case, the joint probability distribution for the measurement outcomes is a mixture of binomial distributions. The FI of such a distribution scales linearly in N asymptotically. We shall use these arguments to prove the following theorem.

Theorem 5. Consider the case of a diagonal interaction (Definition 2). With an interaction of this form, SQL scaling of the QFI is always achievable in N using a

combination of a prepare-and-measure strategy and QEC, for suitable choices of t .

Proof. The proof is as follows. Here, we define the QEC code in the same manner as in the proof of Theorem 2, with $\mathcal{S}$, $|C_0\rangle$, and $|C_1\rangle$ defined in the same way. Contrary to the unitary case, in this section, we will need to perform the recovery operation during the QEC protocol. Although the code here cannot correct any nontrivial errors, it is sufficient to transform the dynamics on the probe system to commuting operations, with the recovery operation being

$$\begin{aligned} \mathcal{R}(\cdot) = & |C_0\rangle \langle C_0| \text{tr}(\mathbb{1}_P \otimes (P_0)_A(\cdot)) \\ & + |C_1\rangle \langle C_1| \text{tr}(\mathbb{1}_P \otimes (P_1)_A(\cdot)). \end{aligned} \quad (44)$$

For an initial state in the code space, the dynamics are governed by the following equation:

$$\begin{aligned} \frac{d\rho}{dt} = & \lim_{\Delta t \rightarrow 0} \frac{P\Phi_{\Delta t}(\rho)P + \mathcal{R}(P_{\perp}\Phi_{\Delta t}(\rho)P_{\perp})}{\Delta t} \\ = & -i[PH_{\text{EP}}P, \rho] - i[\omega P\tilde{G}P, \rho] + \sum_{i=1}^r \left(PL_i\rho L_iP \right. \\ & \left. + \mathcal{R}(P_{\perp}L_i\rho L_i^{\dagger}P_{\perp}) - \frac{1}{2}\{PL_i^{\dagger}L_iP, \rho\} \right), \end{aligned}$$

where P is the projection operator onto the code space. Application of Lemma 3 allows us to ignore the signal on the environment. Assuming an initial state in the form of Eq. (36) and using Eq. (44), we have

$$\begin{aligned} \frac{d\rho^{(i,i)}}{dt} = & -i[\omega P(G \otimes \mathbb{1}_A)P, \rho^{(i,i)}] \\ & - i\left[PH_{\text{EP}}^{(i,i)}P + \frac{\text{Im}(a_i)}{2}\bar{Z}, \rho^{(i,i)}\right] + \frac{\text{Re}(a_i)}{2}(\bar{Z}\rho^{(i,i)}\bar{Z} - \rho^{(i,i)}), \end{aligned}$$

where the logical operator $\bar{Z} = |C_0\rangle \langle C_0| - |C_1\rangle \langle C_1|$ and

$$\begin{aligned} a_i = & \sum_k \frac{1}{2} \left(\langle C_0| \left(L_k^{\dagger} L_k \right)^{(i,i)} |C_0\rangle + \langle C_1| \left(L_k^{\dagger} L_k \right)^{(i,i)} |C_1\rangle \right) \\ & - \langle C_0| L_k^{(i,i)} |C_0\rangle \langle C_0| L_k^{\dagger (i,i)} |C_0\rangle. \end{aligned} \quad (45)$$

It is clear that $\rho^{(i,i)}$ are evolving under Pauli-Z rotations and Markovian dephasing noise, and the strengths of the Pauli-Z rotation and the dephasing noise are different with different i .

We consider the prepare-and-measure strategy. At each time step, we prepare the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|C_0\rangle + |C_1\rangle)$. For the measurement basis, we choose

$$|\pm_i\rangle = \frac{1}{\sqrt{2}}(|C_0\rangle \pm i|C_1\rangle),$$

which we associate with the outcomes 1 and -1 , respectively. Let X_n denote the random variable associated with

the measurement at the n th time step. We will denote by b_n the outcome of the measurement for the n th step. Consider a given measurement sequence with outcomes $b_1, \dots, b_N$, the probability of this event occurring is given by the joint distribution $P(X_N = b_N, \dots, X_1 = b_1)$.

We will first consider the time evolution between two arbitrarily chosen steps of the prepare-and-measure protocol. As in the proof of Theorem 3, we will factor the evolution into two steps. The first step will be evolution under the master equation

$$\begin{aligned} \frac{d\rho^{(i,i)}}{dt} = & -i \left[PH_{\text{EP}}^{(i,i)} P + \frac{\text{Im}(a_i)}{2} \bar{Z}, \rho^{(i,i)} \right] \\ & + \frac{\text{Re}(a_i)}{2} (\bar{Z} \rho^{(i,i)} \bar{Z} - \rho^{(i,i)}), \end{aligned}$$

which corresponds to the evolution due only to the noise. We can integrate the equation to obtain the CPTP map Λ_t corresponding to the time evolution for a duration t :

$$\Lambda_t(|\psi\rangle\langle\psi|) = A_0|\psi\rangle\langle\psi|A_0^\dagger + A_1|\psi\rangle\langle\psi|A_1^\dagger, \quad (46)$$

where

$$A_0 = \sqrt{\frac{1 + \exp(-\text{Re}(\Gamma_i)t)}{2}} \exp\left(-\frac{i}{2}\text{Im}(\Gamma_i)t\bar{Z}\right),$$

and where

$$A_1 = \sqrt{\frac{1 - \exp(-\text{Re}(\Gamma_i)t)}{2}} \bar{Z} \exp\left(-\frac{i}{2}\text{Im}(\Gamma_i)t\bar{Z}\right).$$

Here, $\Gamma_i = i \text{tr}(\bar{Z} P H_{\text{EP}}^{(i,i)} P) + a_i$. We have taken our initial probe state to be $|\psi\rangle$. Further, we know that for $t = 0$, Λ_t leaves the state unchanged and acts as the identity channel, as $A_0(t = 0) = \mathbb{1}_A$ and $A_1(t = 0) = 0$. After evolving under this channel, the next step is the unitary evolution due to the signal.

$$\begin{aligned} & e^{-i\omega t P(G \otimes \mathbb{1}_A)P} \Lambda_t(|\psi\rangle\langle\psi|) e^{i\omega t P(G \otimes \mathbb{1}_A)P} \\ &= \sum_{k=0}^1 e^{-i\omega t P(G \otimes \mathbb{1}_A)P} A_k |\psi\rangle\langle\psi| A_k^\dagger e^{i\omega t P(G \otimes \mathbb{1}_A)P}. \end{aligned}$$

The final step is to measure in the chosen basis. The probability of obtaining an outcome of ± 1 is given by

$$\begin{aligned} & P(X_n = b_n | \rho_E = |\phi_i\rangle\langle\phi_i|) \\ &= \langle \pm_i | e^{-i\omega t P(G \otimes \mathbb{1}_A)P} \Lambda_t(|\psi\rangle\langle\psi|) e^{i\omega t P(G \otimes \mathbb{1}_A)P} | \pm_i \rangle. \end{aligned}$$

First, we see that the probability of obtaining $b_n = 1$ is independent of n due to the fact that for an initial environment state of $|\phi_i\rangle\langle\phi_i|$, the environment remains

invariant and does not mediate correlations between measurements. Let $p_i := P(X_n = 1 | \rho_E = |\phi_i\rangle\langle\phi_i|)$ and $q_i = 1 - p_i$ be the probability of obtaining $b_n = -1$. The probability of obtaining $+1$ is given by

$$p_i = \frac{1}{2} + \frac{1}{2} \exp(-\text{Re}(\Gamma_i)t) \sin((\text{Im}(\Gamma_i) + \omega(\lambda_0 - \lambda_1))t), \quad (47)$$

where λ_0 and λ_1 are the eigenvalues of $P(G \otimes \mathbb{1}_A)P$ associated with $|C_0\rangle, |C_1\rangle$. Further, $\lambda_0 \neq \lambda_1$, as we assume that G acts nontrivially on the probe. The derivative of the probability with respect to the sensing parameter ω is given by

$$\begin{aligned} \partial_\omega p_i = & \frac{1}{2} \exp(-\text{Re}(\Gamma_i)t) (\lambda_0 - \lambda_1)t \\ & \times \cos((\text{Im}(\Gamma_i) + \omega(\lambda_0 - \lambda_1))t). \end{aligned} \quad (48)$$

Since the measurement outcomes are independent, the joint probability can be expressed as the product of the marginals

$$\begin{aligned} & P(X_N = b_N, \dots, X_1 = b_1 | \rho_E = |\phi_i\rangle\langle\phi_i|) \\ &= \prod_{n=1}^N p_i^{(1+b_n)/2} (1 - p_i)^{(1-b_n)/2}. \end{aligned} \quad (49)$$

For an arbitrary diagonal environment state, such as in Eq. (36), we sum the joint probabilities [Eq. (49)] for the different environment basis states,

$$\begin{aligned} & P(X_N = b_N, \dots, X_1 = b_1) \\ &= \sum_{i=1}^{d_E} \alpha_i \prod_{n=1}^N p_i^{(1+b_n)/2} (1 - p_i)^{(1-b_n)/2}. \end{aligned}$$

As mentioned previously, the order of the outcomes does not matter and we recognize a sum of binomial distributions, also called a binomial mixture. For a measurement sequence $b_1, \dots, b_N$, let N_1 count the $+1$ outcomes in the sequence. The associated probability distribution for N_1 is

$$P(N_1 = n_1) = \sum_{i=1}^{d_E} \alpha_i \binom{N}{n_1} p_i^{n_1} (1 - p_i)^{N-n_1}. \quad (50)$$

From this point on, we will assume, without loss of generality, that all the p_i are distinct when we evaluate them at the true value of ω and where a specific value of t has already been fixed. We may do this since similar values of p can be combined into single terms of the sum. Suppose, e.g., that $p_2 = p_3$. Then, we can set $\alpha_2 \mapsto \alpha_2 + \alpha_3$ and remove the index 3 from the sum.

To show that this strategy achieves the SQL, one must show that the FI of this binomial mixture scales linearly in N asymptotically. We show this in Appendix D. ■

B. Two-qubit dephasing model

We will now consider a model that has a diagonal interaction but for which we do not need to apply QEC. Further, for this model we will not measure or compute FI. Rather, we will shelve the probe states and replace them with new ones. For certain control parameters, the QFI of this state will be computed exactly to show SQL scaling.

We now consider a model in which $\tilde{G}$, H_{EP} , $\tilde{H}_E$, and L_k all commute pairwise. We model both the environment and the probe by qubits $\mathcal{H}_P \simeq \mathcal{H}_E \simeq \mathbb{C}^2$. Let $G = Z$ and $H_{EP} = Z \otimes Z$. Let there be a single jump operator $L = \mathbb{1}_E \otimes Z$. Without loss of generality, we set $\tilde{H}_E = 0$, as the commutation of the model allows for the application of Lemma 3. This model corresponds to pure phase noise on both the system and the environment, as the interaction is completely diagonal on both the probe and the environment.

We can see that since $\tilde{G} = L$, the conditions of Theorem 4 are satisfied and so the SQL is the best attainable scaling. Further, the results from Theorem 5 of the previous section guarantee that the SQL is achievable with a prepare-and-measure strategy. This example, however, is particular since we can show that the QFI itself scales linearly in N , rather than a FI lower bound. Further, QEC will not be used in this example to tailor the noise.

We will prepare N probe states in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Since all the operators commute, we can factor out the time evolution into three steps: evolution by the channel corresponding to L , evolution H_{EP} , and, finally, evolution by G .

Let ρ_E be the initial state of the environment. We begin with the evolution from the jump operators,

$$\frac{d\rho}{dt} = (\mathbb{1}_E \otimes Z) \rho (\mathbb{1}_E \otimes Z) - \rho.$$

One can show that evolving under this equation for a duration t is equivalent to applying the following dephasing channel on the probe:

$$\mathcal{D}_t(\rho_{PA}) = \left(\frac{1 + e^{-2t}}{2} \right) \rho_{PA} + \left(\frac{1 - e^{-2t}}{2} \right) Z \rho_{PA} Z.$$

For $\rho_{PA} = |+\rangle\langle+|$, this channel is equivalent to a depolarizing channel:

$$\mathcal{D}_t(|+\rangle\langle+|) = e^{-2t} |+\rangle\langle+| + (1 - e^{-2t}) \frac{\mathbb{1}_P}{2}.$$

Next, we apply the time evolution due to H_{EP} to the state. The unitary operator for this time evolution is

$$\begin{aligned} U_{EP}(t) &= e^{-itZ \otimes Z} \\ &= |0\rangle\langle 0| \otimes e^{-itZ} + |1\rangle\langle 1| \otimes e^{itZ}. \end{aligned}$$

Application of this unitary yields

$$\begin{aligned} U_{EP}(t) [\rho_E \otimes \mathcal{D}_t(|+\rangle\langle+|)] U_{EP}(t)^\dagger \\ = \sum_{i,j=0}^1 |i\rangle\langle i| \rho_E |j\rangle\langle j| \otimes e^{-i(-1)^i t Z} \mathcal{D}_t(|\psi\rangle\langle\psi|) e^{i(-1)^j t Z}. \end{aligned}$$

Finally, we consider the time evolution due to the signal. This corresponds to unitary evolution on the probe generated by G . The state of the environment and the first probe after time t is given by

$$\begin{aligned} \sum_{i,j=0}^1 |i\rangle\langle i| \rho_E |j\rangle\langle j| \\ \otimes e^{-i\omega t Z} e^{-i(-1)^i t Z} \mathcal{D}_t(|\psi\rangle\langle\psi|) e^{i(-1)^j t Z} e^{i\omega t Z}. \end{aligned}$$

Since we are considering a prepare-and-measure strategy, after this time t the probe is swapped out for a new probe and the process is repeated. We repeat the process for N probes. One can show that the final output state ρ_{EP}^f for the environment and the N probes is given by

$$\begin{aligned} \rho_{EP}^f(t) &= \sum_{i,j=0}^1 |i\rangle\langle i| \rho_E |j\rangle\langle j| \\ &\otimes \left[e^{-i\omega t Z} e^{-i(-1)^i t Z} \mathcal{D}_t(|\psi\rangle\langle\psi|) e^{i(-1)^j t Z} e^{i\omega t Z} \right]^{\otimes N}. \end{aligned}$$

Since we have access only to the states of the probe, we trace out the environment to obtain the final probe state ρ_{PA}^f ,

$$\begin{aligned} \rho_{PA}^f(t) &= \sum_{i=0}^1 \langle i| \rho_E |i\rangle \left[e^{-2t} e^{-i\omega t Z} e^{-i(-1)^i t Z} \right. \\ &\cdot |+\rangle\langle+| e^{i(-1)^i t Z} e^{i\omega t Z} + (1 - e^{-2t}) \frac{\mathbb{1}_P}{2} \left. \right]^{\otimes N}. \end{aligned} \quad (51)$$

If we fix the time $t = \frac{\pi}{2}$, then this density matrix is easily diagonalizable. Indeed, the state reduces to

$$\rho_{PA}^f\left(\frac{\pi}{2}\right) = \left[e^{-\pi} e^{-i\omega \frac{\pi}{2} Z} |-\rangle\langle-| e^{i\omega \frac{\pi}{2} Z} + (1 - e^{-\pi}) \frac{\mathbb{1}_P}{2} \right]^{\otimes N}.$$

Since the QFI of a product state is the sum of the QFIs of the factors of the product [61], then the FI of the final state

$\mathcal{F}\left(\rho_{\text{PA}}^f\left(\frac{\pi}{2}\right)\right)$ is given by

$$\mathcal{F}\left(\rho_{\text{PA}}^f\left(\frac{\pi}{2}\right)\right) = N\mathcal{F}(\rho_1), \quad (52)$$

where

$$\rho_1 = e^{-\pi} e^{-i\omega\frac{\pi}{2}Z} |-\rangle \langle -| e^{i\omega\frac{\pi}{2}Z} + (1 - e^{-\pi}) \frac{\mathbb{1}_P}{2}. \quad (53)$$

So long as $\mathcal{F}(\rho_1)$ is independent of N and greater than zero, then the SQL is achieved.

We recall that the QFI of a mixed state can be calculated using Eq. (25). One can show that ρ_1 has eigenvalues $(e^{-2t} + (1 - e^{-2t})/2)$ and $(1 - e^{-2t})/2$ with eigenvectors $e^{-i\omega\frac{\pi}{2}Z} |-\rangle$ and $e^{-i\omega\frac{\pi}{2}Z} |+\rangle$, respectively. This yields a QFI of

$$\mathcal{F}(\rho_1) = e^{-2\pi} \pi^2, \quad (54)$$

which is both independent of N and greater than zero. Therefore, for the diagonal interaction, the SQL is attainable with a prepare-and-measure strategy without using QEC to tailor the noise, as in the Markovian case.

C. Numerical sampling of the classical FI

In the previous subsections, we have considered models with restricted structure, particularly diagonal interactions. We have shown SQL scaling in those specific cases. We now consider more generic models and present numerical evidence that the SQL is still generally achievable with prepare-and-measure strategies, even without QEC, mirroring the Markovian case [16,17].

Consider the joint quantum state of N probes. If these probes are qudits, then the Hilbert space scales as d_P^N and so the computation of the QFI using exact diagonalisation and Eq. (25) quickly becomes intractable as N grows. To simplify the problem, we instead measure the probes at each time step and calculate the FI.

For N rounds of prepare-and-measure, we can associate the measurements with random variables $\{X_1, \dots, X_N\}$, forming a stochastic process. These random variables can be correlated, with the correlations mediated by the environment. It has been shown that if the process is stationary and its Markov order, which quantifies the memory length, is finite, then the classical FI will scale linearly in N [67]. It is the minimum l such that

$$P(X_N | X_{N-1}, \dots, X_1) = P(X_N | X_{N-1}, \dots, X_{N-l}).$$

The HMM under consideration does not generally possess these two properties.

We compute the classical FI using Eq. (D1). Since the same probe state is prepared at every time step, the i th measurement outcome depends solely on the residual state

of the environment after the previous measurement. Consequently, it suffices to track only the $d_E \times d_E$ density matrix of the environment and its derivative throughout the time evolution. This approach enables the computation of the probabilities and their derivatives for the measurement outcomes $\{x_1, \dots, x_N\}$ using constant memory, thereby avoiding the exponential memory requirements associated with storing the full density matrix for N probes. We trade the memory overhead for exponential time complexity. If each random variable X_i takes on n_{out} possible different values, then there are $(n_{\text{out}})^N$ different measurement outcomes, which must all be calculated separately to compute the FI exactly.

To get around this exponential scaling, we use random sampling to approximate the FI. Our methods are inspired by previous works using Monte Carlo to calculate the FI [69,70]. We consider the following unbiased estimator for the FI:

$$\hat{I}(\omega) = \frac{1}{S} \sum_{s=1}^S \left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^2, \quad (55)$$

where $\{x_1^{(s)}, \dots, x_N^{(s)}\}$ is a sample from $\{X_1, \dots, X_N\}$ and the s index differentiates the samples. The variable S is the total number of samples. Each sample is obtained by computing the time evolution. At each measurement step, we randomly decide the outcome. To do so, we randomly sample from a Bernoulli distribution with parameter $p = P(X_n = x_n | X_{n-1} = x_{n-1}, \dots)$, where $P(X_n = x_n | X_{n-1} = x_{n-1}, \dots)$ is given by the trace of the non-normalized environment density matrix after projecting the probe into the state associated with a $+1$ outcome.

This method provides an approximate value of the FI. For a fixed N , the variance of this estimator scales as $O(S^{-1})$ and so convergence is guaranteed for a large number of samples. The constant hidden by the big O may depend on N and t . We *a priori* do not know this functional dependence. We can, however, numerically and heuristically estimate the variance to provide error bars for Figs. 3 and 4. The details of the variance estimation are provided in Appendix E.

1. Heisenberg interaction

We first consider the prepare-and-measure strategy for a model in which the operators of the master equation do not commute. We once again model the probes and environment as qubits $\mathcal{H}_P \simeq \mathcal{H}_E \simeq \mathbb{C}^2$. We consider a signal given by $G = Z$ and a single jump operator $L = \mathbb{1}_E \otimes Z$. We consider a Heisenberg-type interaction for the coupling Hamiltonian,

$$H_{\text{EP}} = \gamma (X \otimes X + Z \otimes Z + Y \otimes Y), \quad (56)$$

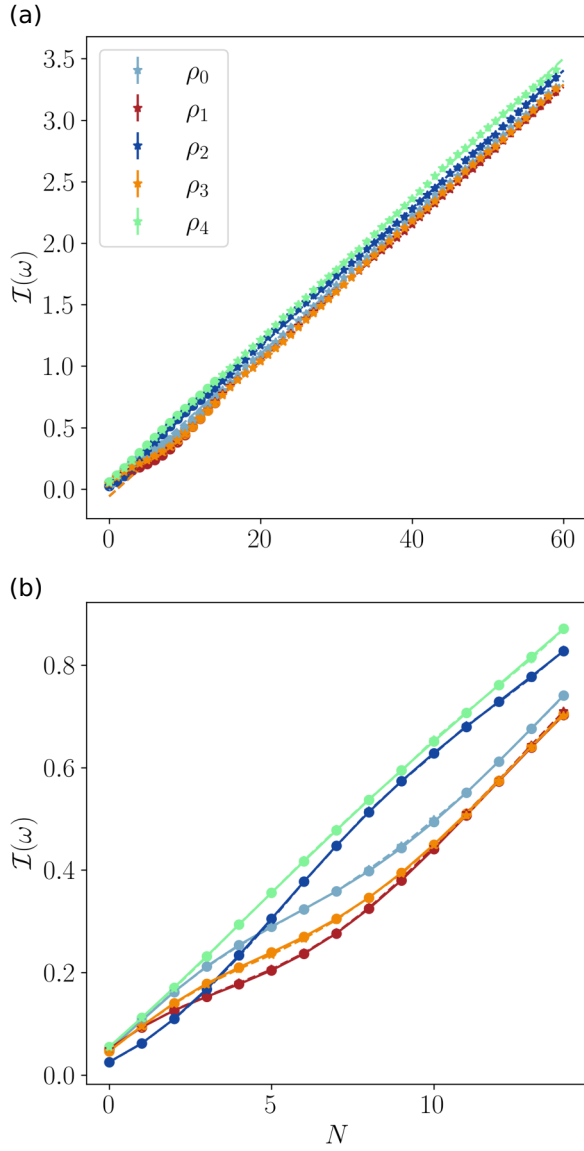


FIG. 3. The FI ($\mathcal{I}(\omega)$) for the Heisenberg model with $H_E = 0$. The solid dots have been calculated exactly and the stars have been calculated by random sampling. The different colors correspond to different initial environment states ρ_i sampled from Haar-random states. (a) The FI for system sizes up to $N = 60$. The dashed lines show linear regression on the data. (b) A comparison of exact computation and Monte Carlo sampling FI values up to $N = 15$. The lines serve to guide the eye. The dashed and solid lines interpolate the random-sampling and exact FI values, respectively.

where γ is the coupling strength. We will consider two cases of the model. In the first case, there is no signal acting on the environment and $H_E = 0$. In the second case, there is a signal acting on the environment and $H_E = Z$.

Here, we fix the measurement basis to be the eigenstates of the Pauli- X operator and we prepare each probe

TABLE II. Linear regressions for the Fisher information (FI) as a function of the number of probes for the Heisenberg-type model with $H_E = 0$. Associated curves are shown in Fig. 3.

i	Slope	r_c coefficient
0	0.05755739026318657	0.9987758048158253
1	0.057886505716965755	0.997690310855459
2	0.05267854285583423	0.9980036537841903
3	0.05734837817591377	0.9990203746173694
4	0.0643685351834392	0.9956504727114459

in the state $|+_i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$. We fix the parameters to be $\gamma = 1$, $\omega = 0$, and $t = 0.25$. The initial state of the environment is obtained by sampling pure states from the Haar measure. The same five random states are used for both the $H_E = 0$ and $H_E = Z$ models. The FI is then calculated using 25 000 samples.

In Fig. 3, we show the classical FI as a function of the number of prepare-and-measure rounds for the model in which $H_E = 0$. The FI curves are calculated for five pure Haar-random initial environment states $\{\rho_0, \dots, \rho_4\}$. For $N \leq 15$, the FI is computed both exactly and by random sampling. At larger N , the values are exclusively from random sampling.

In Fig. 3(a), we show the FI values up to $N = 60$ as well as linear regression curves for the data. For each initial environment state ρ_i , a linear regression is performed on the values of the FI to show the linear (SQL) scaling. The results of the linear regressions are presented in Table II. The linear regression gives us the slope of the linear relation along with the Pearson correlation coefficient r_c , which measures the linear correlation between N and the FI. A higher r_c indicates a stronger correlation, with $r_c \in [-1, 1]$. The r_c coefficients of Table II are all very close to 1, which strongly indicates that the relationship between the FI and N is linear. Further, the values for the slopes are all above zero. Therefore, we see that the SQL is achieved up to depths of 60. We conjecture that this trend holds in the limit of $N \rightarrow \infty$.

In Fig. 3(b), we show the FI values up to $N = 15$ from both exact and random-sampling methods. The curves in this figure serve only to guide the eye. This figure shows the agreement between the sampling results and exact computations.

In Fig. 4 and Table III, we show the same information but for the $H_E = Z$ model, where the signal also acts on the environment. Similarly to Fig. 3(b), Fig. 4(b) shows excellent agreement between the exact and sampled FI. In Fig. 4(a), the presence of signal on the environment seems to lower the values of the FI. Further, the scaling of the FI does not seem to be linear. This is reflected by the values of r_c obtained for the model in Table III, which are much lower than those of Table II. In particular, the values of r_c

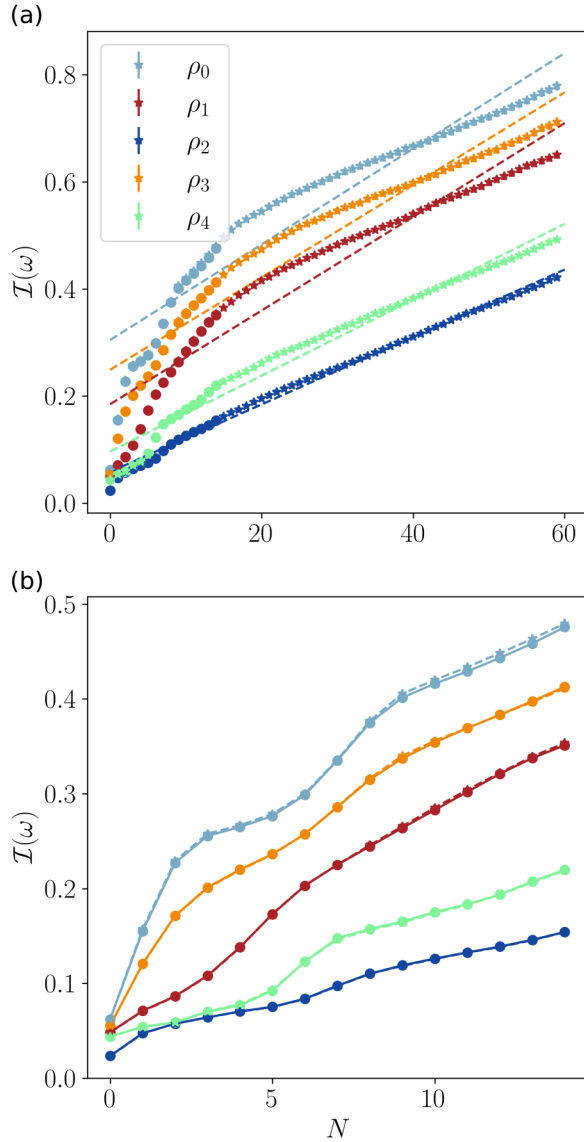


FIG. 4. The FI ($\mathcal{I}(\omega)$) for the Heisenberg model with $H_E = Z$. The solid dots have been calculated exactly and the stars have been calculated by random sampling. The different colors correspond to different initial environment states ρ_i sampled from Haar-random states. (a) The FI for system sizes up to $N = 60$. The dashed lines show linear regression on the data. (b) A comparison of exact computation and Monte Carlo sampling FI values up to $N = 15$. The lines serve to guide the eye. The dashed and solid lines interpolate the random-sampling and exact FI values, respectively.

are not high enough for us to confidently conclude that the SQL is achieved.

2. Random master equations

To provide further evidence that the SQL is always achievable, we numerically study the prepare-and-measure strategy for a model in which the noise operators of the

TABLE III. Linear regressions for the FI as a function of the number of probes for the Heisenberg-type model with $H_E = Z$. Associated curves are shown in Fig. 4.

i	Slope	r_c coefficient
0	0.00893164136192775	0.937906503836237
1	0.008752382336452712	0.9550822417044361
2	0.006298895354174299	0.9968012417020177
3	0.00863620471038918	0.9498835201202465
4	0.007083260388947953	0.9857235070906991

master equation are randomly generated. We again model the probe and environment by qubits, $\mathcal{H}_P \simeq \mathcal{H}_E \simeq \mathbb{C}^2$. The protocol does not require any auxiliary degrees of freedom. We once again consider the signal to be $G = Z$ and suppose that the parameter we are sensing is $\omega = 0$. For each round of the protocol, we prepare the probe in the state $|+\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ and allow the system to evolve for a time $t = 0.1$, before finally measuring in the eigenbasis of the X operator. As in the Heisenberg case, we consider two cases of the model, $H_E = 0$ and $H_E = Z$.

We consider a model with three jump operators. The jump operators are obtained by sampling random matrices from the Ginibre ensemble [71,72]. For the coupling Hamiltonian H_{EP} , we sample a random matrix A from the Ginibre ensemble. We then set $H_{EP} = \frac{1}{2}(A + A^\dagger)$. Finally, the initial environment state is taken to be a Haar-random pure state.

For a given master equation, we compute the FI as a function of the number of rounds N from $N = 1$ to $N = 30$. The values of the FI are calculated using the random-sampling method of Sec. IV C with 25 000 samples. A linear fit is then applied to the resulting curve, yielding a slope and an r_c coefficient. These quantities have been computed for 1050 different random master equations. The same 1050 master equations are used for the $H_E = 0$ and $H_E = Z$ models.

The distributions of the r_c values and slopes for the $H_E = 0$ models are presented in Figs. 5(a) and 5(b) respectively. First, we see from Fig. 5(a) that all of the obtained r_c values are very close to 1, with the lowest being 0.9820 rounded to the fourth digit. There is therefore a strong linear relationship between the FI and N for all of the random models studied. Therefore, it appears that linear scaling of the FI up to small fluctuations is a general property with this choice of strategy. Further, the minimum slope obtained is approximately 1.48466×10^{-5} . Although the minimal slope is quite small, it is still nonzero and of a few orders of magnitude above machine precision. All of the randomly generated models therefore have approximately linear FI scaling with a nonzero slope. This is strong evidence that the prepare-and-measure strategy can generically achieve the SQL. It is important to note that since we have chosen the same values of t , the same measurement basis, and the same initial state for all of the random-noise

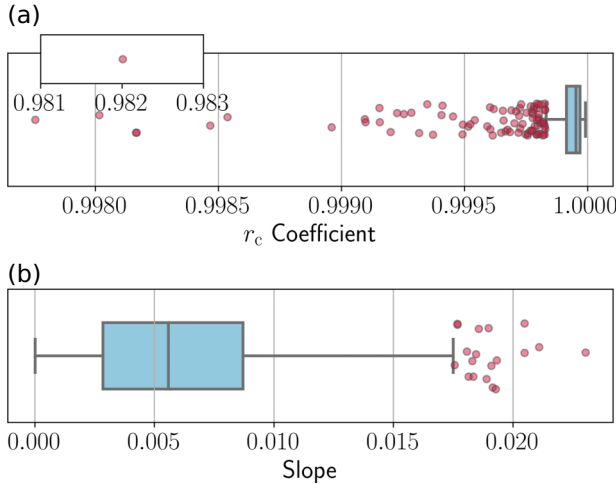


FIG. 5. The statistics for the linear fits to the FI of 1050 random models with $H_E = 0$. The FI of each data point is obtained with a prepare-and-measure strategy with noise from a random master equation. The boxes extend from the first to the third quartile, split by the median. The whiskers extend from the quartiles to 1.5 times the interquartile range. (a) The slopes of the linear fits. The subplot in the top-left corner shows an isolated outlier and has been separated for aesthetic purposes. (b) The Pearson coefficients r_c of the linear fits.

models considered, the values of the FI obtained may be far from optimal, even among prepare-and-measure strategies. We can therefore expect an increase in performance by fine tuning these parameters for each random model considered.

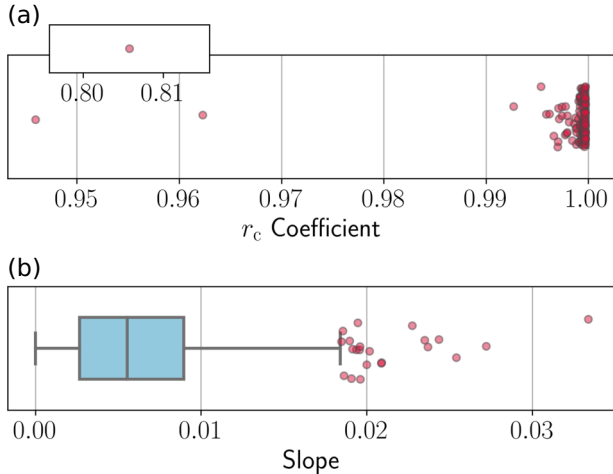


FIG. 6. The statistics for the linear fits to the FI of 1050 random models with $H_E = Z$. The FI of each data point is obtained with a prepare-and-measure strategy with noise from a random master equation. (a) The slopes of the linear fits. The subplot in the top-left corner shows an isolated outlier, and has been separated for aesthetic purposes. (b) The Pearson coefficients r_c of the linear fits.

In Fig. 6, we present the same data for the $H_E = Z$ case of the random models. The minimum r_c value among the models is 0.8058 rounded to the fourth digit and the minimum slope is 3.2925×10^{-6} . We see that the majority of the r_c values of Fig. 6(a) are clustered between 0.99 and 1, indicating that for the majority of the random models, the relationship is strongly linear and therefore the scaling is probably SQL. However, there are three notable outliers. It is therefore unclear whether the SQL is always achieved in the presence of an environment signal, $H_E = Z$.

V. CONCLUSIONS

In this work, we have considered the attainability of HL scaling in quantum metrology using QEC when the probe is subject to non-Markovian noise. We have restricted our investigation to an HMM, which displays non-Markovianity for many choices of metric. We have presented the HNES conditions, which guarantee the HL using QEC, and generalized the HNLS conditions of Ref. [16] to the non-Markovian regime. Contrary to the Markovian conditions, the HNES conditions are sufficient but not necessary. Through an example, we have demonstrated that deriving a general necessary condition requires taking into account the initial state of the environment and may potentially require tracking the trajectory of the system. Thus, deriving such conditions is challenging. We have further presented alternative conditions for obtaining the HL in certain scenarios in which the HNES conditions may be violated. The trade-off is that achieving the HL in these scenarios may require a relatively strict measurement schedule. First, we have shown that when there are no dissipative terms in the master equation, i.e., when the evolution of the entire system is unitary, and when the signal acts only on the probe, then one can always achieve HL scaling of the QFI. Contrary to the HNES conditions, which guarantee exact quadratic scaling of the QFI, unitary evolution only guarantees quadratic scaling up to an almost periodic envelope function. In order to obtain the HL, one must therefore measure the system and specific recurrences times. Second, we have shown that for a specific class of interactions, which we call diagonal interactions, the assumptions of the HNES conditions may be weakened. We have presented the weaker HNELS conditions, which guarantee HL scaling of the QFI. Once again, the HL scaling is up to a periodic envelope function and so measurements must be well timed.

We have shown that the original HNLS conditions can be applied to the full system of probe and environment to guarantee that the SQL is the best achievable limit in certain cases. A summary of all the cases for the scaling of the QFI is presented in Fig. 2.

Finally, we have investigated whether SQL scaling within a given experiment is always possible when we do not reset the state of the environment. We have shown that

it is always achievable in a special case in which the model has a diagonal interaction. We have presented numerical evidence of SQL scaling for a specific Heisenberg model in which the operators do not commute and for a large number of randomly generated models. We have considered a finite-size scaling argument on the FI, where the FI is calculated exactly for a few measurements and then via random sampling for a large number of measurements. When the signal acts on only the probe, there is strong evidence that the SQL is always achievable (see Fig. 3). When the signal also acts on the environment, the numerics are inconclusive due to the presence of important outliers in the data (see Fig. 4).

Many open problems remain. First, there is the problem of deriving sufficient and necessary conditions for HL scaling. Second, a proof for the general attainability of the SQL in a given experiment is still lacking. Lastly, our work assumes a signal that acts unitarily and independently on the probe and environment. This setup is not completely general and leaves open the problem of QFI scaling when the dissipative terms as well as the unitary coupling depend on the parameter to be estimated.

ACKNOWLEDGMENTS

We would like to thank Senrui Chen, Matthew Dusch- enes, Noah Gorgichuk, Liang Jiang, Gideon Lee, Benjamin MacLellan, Akimasa Miyake, Arsalan Motamedi, Andy Schang, Ruhi Shah, Shravan Shravan, and Tyler Thurtell for helpful discussions. S.Z. and N.C. acknowledge funding provided by Perimeter Institute for Theoretical Physics, a research institute supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities. S.Z. acknowledges the support from the National Research Council of Canada (Grant No. AQC-217-1). N.C. acknowledges the support from the Applied Quantum Computing Challenge Program at the National Research Council of Canada. R.L. would like to thank Mike and Ophelia Lazaridis for funding.

DATA AVAILABILITY

The data that support the findings of this paper are openly available [73].

APPENDIX A: KL CONDITIONS FOR THE HMM

We will follow the definitions and proofs of Refs. [56, 74]. We can modify Definition 1 to obtain an equivalent picture that includes the syndrome information, in which case we say that $(U, \mathcal{E})$ is a QECC if there exists an isometry V such that

$$(V \otimes \mathbb{1}_{d'})E(U \otimes \mathbb{1}_d)|\psi\rangle \otimes |\phi\rangle = |\psi\rangle \otimes \sqrt{c_{E,\psi,\phi}}|S_{E,\psi,\phi}\rangle,$$

where here the state S contains both the syndrome and the environment degrees of freedom. This picture is obtained by taking the isometry V to be the Stinespring dilation of the channel D . The states remain pure in this picture since the trace is not taken. Let d_S be the dimension necessary for the dilation, such that $|S_{E,\psi,\phi}\rangle \in \mathbb{C}^{d_S}$.

As for standard error-correcting codes, if $(U, \mathcal{E})$ is a QECC, then $(U, \text{span}(\mathcal{E}))$ is a QECC. The proof is the same as in the Markovian case. We follow the proof given in ref. [56, 74]. Consider $F, E \in \mathcal{E}$. As seen previously, there exists V such that $\forall |\psi\rangle \in \mathbb{C}^k$ and $\forall |\phi\rangle \in \mathbb{C}^{d_E}$, we have that

$$\begin{aligned} (V \otimes \mathbb{1}_{d'_E})E(U \otimes \mathbb{1}_{d_E})|\psi\rangle|\phi\rangle &= \sqrt{c_{E,\psi,\phi}}|\psi\rangle|S_{E,\psi,\phi}\rangle, \\ (V \otimes \mathbb{1}_{d'_E})F(U \otimes \mathbb{1}_{d_E})|\psi\rangle|\phi\rangle &= \sqrt{c_{F,\psi,\phi}}|\psi\rangle|S_{F,\psi,\phi}\rangle. \end{aligned}$$

By linearity, we have

$$\begin{aligned} (V \otimes \mathbb{1}_{d'_E})(aE + bF)(U \otimes \mathbb{1}_{d_E})|\psi\rangle|\phi\rangle \\ = |\psi\rangle(a\sqrt{c_{E,\psi,\phi}}|S_{E,\psi,\phi}\rangle + b\sqrt{c_{F,\psi,\phi}}|S_{F,\psi,\phi}\rangle) \end{aligned}$$

and so we see that $aE + bF$, where $a, b \in \mathbb{R}$ is also a correctable error, completing the proof. Using these properties, we can now prove Lemma 1.

Proof of Lemma 1. The proof is similar to those of the standard KL conditions. We begin by showing that the conditions are necessary. Let $|\psi\rangle$ and $|\phi\rangle$ be linearly independent states of $\mathbb{C}^k$. Let E_i and E_j be errors in $\mathcal{E}$. We consider orthogonal basis states $|\chi_1\rangle, |\chi_2\rangle \in \mathbb{C}^{d_E}$. Since we have a QECC, then there exists an isometry V such that

$$\begin{aligned} (V \otimes \mathbb{1}_{d'_E})E_i(U \otimes \mathbb{1}_{d_E})|\psi\rangle|\chi_1\rangle \\ = \sqrt{c_{E_i,\psi,\chi_1}}|\psi\rangle|S_{E_i,\psi,\chi_1}\rangle, \end{aligned} \quad (\text{A1})$$

$$\begin{aligned} (V \otimes \mathbb{1}_{d'_E})E_j(U \otimes \mathbb{1}_d)|\phi\rangle|\chi_2\rangle \\ = \sqrt{c_{E_j,\phi,\chi_2}}|\phi\rangle|S_{E_j,\phi,\chi_2}\rangle. \end{aligned} \quad (\text{A2})$$

We now consider a state of the form

$$|\eta\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle,$$

where $|\psi^\perp\rangle$ is a vector that is orthogonal to $|\psi\rangle$. We have that

$$\begin{aligned} (V \otimes \mathbb{1}_{d'_E})E_i(U \otimes \mathbb{1}_{d_E})|\eta\rangle|\chi_1\rangle \\ = \sqrt{c_{E_i,\eta,\chi_1}}|\eta\rangle|S_{E_i,\eta,\chi_1}\rangle. \end{aligned} \quad (\text{A3})$$

We now project with $|\psi\rangle\langle\psi|$ on both Eqs. (A1) and (A3) and then divide the second equation by α . The left-hand

sides become equal, and we equate them to obtain

$$\sqrt{c_{E_i, \psi, \chi_1}} |S_{E_i, \psi, \chi_1}\rangle = \sqrt{c_{E_i, \eta, \chi_1}} |S_{E_i, \eta, \chi_1}\rangle \quad \forall \eta.$$

We can therefore conclude that the residual syndrome and environment state do not depend on the choice of initial state, i.e.,

$$c_{E_i, \psi, \chi_1} \equiv c_{E_i, \chi_1}.$$

Knowing this, we now take the inner product between Eqs. (A1) and (A3) with the projector $|\xi_l\rangle_S \langle \xi_k|_S$ sandwiched between the bra and ket, where $|\xi_l\rangle_S$ and $|\xi_k\rangle_S$ are orthogonal basis elements of $\mathbb{C}^{d_S}$. We also relabel χ_1 to χ_m and χ_2 to χ_n . This gives us the new equation

$$\begin{aligned} & \langle \phi | \langle \chi_m | (U^\dagger \otimes \mathbb{1}_{d_E}) E_j^\dagger (V^\dagger \otimes \mathbb{1}_{d'_E}) \\ & \cdot (\mathbb{1}_k \otimes |\xi_l\rangle_S \langle \xi_k|_S) (V \otimes \mathbb{1}_{d'_E}) E_i (U \otimes \mathbb{1}_{d_E}) |\psi\rangle | \chi_n \rangle \\ & = \sqrt{c_{E_i, \chi_n} c_{E_j, \chi_m}} \langle S_{E_j, \chi_m} | |\xi_l\rangle \langle \xi_k | S_{E_i, \chi_n} \rangle \langle \phi \rangle \psi. \end{aligned}$$

By defining

$$\begin{aligned} c_{ij}^{mnlk} &= \sqrt{c_{E_i, \chi_m} c_{E_j, \chi_n}} \langle S_{E_j, \chi_n} | |\xi_l\rangle \\ & \times \langle \xi_k | S_{E_i, \chi_m} \rangle \langle \phi \rangle \psi, \end{aligned}$$

the equation is reduced to

$$P \langle \chi_m | E_i^\dagger | \chi_l \rangle \langle \chi_k | E_j | \chi_n \rangle P = c_{ij}^{mnlk} P.$$

This is exactly Eq. (4) for the KL conditions, thus completing the proof that the conditions are necessary for a QECC.

We now show that the conditions are sufficient. By defining

$$\tilde{E}_{i,n,k} = \langle n |_{E'} E_i | k \rangle_E, \quad (\text{A4})$$

one can see that the generalized conditions are equivalent to the standard KL conditions of Ref. [53], where we consider the extended error set $\{\tilde{E}_{i,n,k}\}$ indexed by i , n , and k acting solely on the probe.

Let $|\chi\rangle \in \mathbb{C}^{d_E}$ be the environment state and $|\psi\rangle \in \mathbb{C}^{d_P}$ be a state in the code space. We expand the environment state over an orthonormal basis to obtain

$$|\chi\rangle = \sum_{l=0}^{d_E} c_l |l\rangle_E.$$

Let $E_i \in \mathcal{E}$ be an error. We then have that

$$\begin{aligned} E_i |\psi\rangle | \chi \rangle &= \left(\sum_{n=0}^{d'_E} \sum_{k=0}^{d_E} \langle n |_{E'} E_i | k \rangle_E \otimes | n \rangle_{E'} \langle k |_E \right) \left(|\psi\rangle \otimes \sum_{l=0}^{d_E} c_l |l\rangle_E \right) \\ &= \sum_{n=0}^{d'_E} \left[\left(\sum_{k=0}^{d_E} c_k \tilde{E}_{i,n,k} |\psi\rangle \right) \otimes | n \rangle_{E'} \right]. \end{aligned}$$

Since we suppose that the standard KL conditions hold for $\tilde{E}_{i,n,k}$ on the probe, then by Ref. [53] there exists a CPTP map D that acts solely on the probe such that

$$\begin{aligned} & (D \otimes \mathbb{1}_{d'_E}) \left(\sum_{n,n'=0}^{d'_E} \left[\left(\sum_{k,k'=0}^{d_E} c_k c_{k'}^* \tilde{E}_{i,n,k} |\psi\rangle \langle \psi | \tilde{E}_{i,n',k'}^\dagger \right) \right. \right. \\ & \left. \left. \otimes | n \rangle_{E'} \langle n' |_{E'} \right] \right) = |\psi\rangle \langle \psi | \otimes \sum_{n,n'=0}^{d'_E} \alpha_{nn'} | n \rangle_{E'} \langle n' |_{E'}, \end{aligned}$$

where the $\alpha_{nn'}$ are the coefficients of the residual non-normalized state in the environment. This holds since the D map can correct for linear combinations of the errors $\tilde{E}_{i,n,k}$. We need only compose with a partial trace map on the environment degrees of freedom to obtain the exact definition for a QECC. The measurement and recovery operators for the HMM are obtained from the code for the reduced errors $\tilde{E}_{i,n,k}$. ■

APPENDIX B: EXAMPLE NOISE IN THE MARKOV APPROXIMATION

We will consider N environmental spin-1 degrees of freedom. We will consider a single spin-1/2, as we will assume that the noise on each spin-1/2 is uncorrelated. The total Hamiltonian of the system can be written as

$$H = \sum_{i=1}^N H_{E,i} + \sum_{i=1}^N H_{EP,i},$$

where

$$H_{EP,i} = \tau_i (S_x^{(i)} \otimes X + S_z^{(i)} \otimes Z).$$

Here, the superscript is used to index the spins on which the operator acts. We will suppose, without loss of generality, that the total Hamiltonian has no nontrivial terms acting on the probe. If this were not the case, then these terms could be eliminated by the fast and accurate universal control. This spin-1 Hamiltonian is inspired by the Hamiltonian of nitrogen-vacancy centers in diamond [75,76]. The full

interaction Hamiltonian can be written

$$H_I = \sum_{i=1}^N \sum_{A \in \{X, Z\}} \tau_i S_A^{(i)} \otimes A.$$

We set

$$H_{E,i} = \Delta (S_z^{(i)})^2, \quad (\text{B1})$$

where Δ is the zero-field splitting. τ_i is the strength of the coupling between the environmental spin i and the spin-1/2, which we assume to be small. In the Markovian regime, we will assume that $\Delta \gg \tau_i$, such that the effect of the coupling on the environment state can be neglected. We see that $|0\rangle_E$ is the ground state of the zero-field splitting Hamiltonian given by Eq. (B1). Following the prescription of Ref. [23], we can obtain a Markovian master equation acting on the spin-1/2 by making the Born-Markov approximation. In this approximation, we assume that the state can be written as

$$\rho(t) \approx \rho_P(t) \otimes (|0\rangle_E \langle 0|_E)^{\otimes N}. \quad (\text{B2})$$

In other words, the environment spins are approximated to always be in the state $(|0\rangle_E \langle 0|_E)^{\otimes N}$ due to their weak coupling and fast decoherence. The master equation acting on the probe is then given by Ref. [23, Eq. (3.143)]:

$$\frac{d\rho_P}{dt} = \sum_{A, Q \in \{X, Z\}} \gamma_{AQ} \left(A \rho_P Q - \frac{1}{2} \{QA, \rho_P\} \right). \quad (\text{B3})$$

We have neglected the Lamb-shift Hamiltonian as it can be offset by our controls, leaving only the dissipative terms. The decay rates γ_{AQ} are given by Ref. [20, Eq. (3.139)]:

$$\gamma_{AQ} = \sum_{i,j=1}^N \int_{\mathbb{R}} dt e^{i\omega t} \text{Tr} \left[e^{itH_{E,i}} S_A^{(i)} e^{-itH_{E,i}} S_Q^{(j)} (|0\rangle_E \langle 0|_E)^{\otimes N} \right].$$

In particular, we see that $\gamma_{AQ} = 0$ if A or Q is equal to Z , and so, Eq. (B2) reduces to

$$\frac{d\rho_P}{dt} = \gamma_{XX} (X \rho_P X - \rho).$$

We recognize the equation for dephasing along the X axis. The corresponding channel has Kraus operators proportional to $\mathbb{1}$ and X .

APPENDIX C: DIRECT PROOF OF THE LIMIT SUPERIOR FOR THE UNITARY CASE

It is sufficient to show for any $t_0 > 0$ and $\epsilon > 0$, there exists $t > t_0$ such that $|\alpha(t) - \alpha(0)| = |\alpha(t) - \frac{1}{2}| < \epsilon$.

First, due to the continuity and periodicity of $e^{i\phi}$ as a function of ϕ , there is a constant δ that depends on ϵ such that $|\sum_m \bar{c}_m e^{-it\phi_m} - \sum_m \bar{c}_m| < \epsilon$ whenever $|t\phi_m - 2k_m\pi| \leq \delta$ for some set of integers $\{k_m\}$ and all m . Without loss of generality, we also assume that δ is small enough such that $2\pi \lceil 1/(4\pi\delta) \rceil > t_0$ and $4\pi\delta^2 < \delta$. Second, according to Dirichlet's approximation theorem [77], there exist an integer $1 \leq q \leq \lceil 1/\delta^2 \rceil^{d_E^2}$ and a set of integers $\{q_m\}$ such that $|\phi_m - q_m/q| \leq \frac{1}{\lceil 1/\delta^2 \rceil q} \leq \delta^2/q$. Finally, let $t = 2\pi \lceil 1/(4\pi\delta) \rceil q$. We have

$$\left| t\phi_m - t\frac{q_m}{q} \right| \leq \left\lceil \frac{1}{4\pi\delta} \right\rceil q 2\pi \frac{\delta^2}{q} \leq \frac{\delta + 4\pi\delta^2}{2} < \delta. \quad (\text{C1})$$

The above provides a construction of t that satisfies $t = 2\pi \lceil 1/(4\pi\delta) \rceil q > t_0$ and $|t\phi_m - t\frac{q_m}{q}| < \delta$ as required, which implies $|\sum_m \bar{c}_m e^{-it\phi_m} - \sum_m \bar{c}_m| < \epsilon$. Therefore, we have $\limsup_{t \rightarrow \infty} |\alpha(t)| = \frac{1}{2}$, leading to a QFI scaling of

$$\limsup_{t \rightarrow \infty} \frac{\mathcal{F}(\rho_{\text{PA}})}{t^2} = (\Delta\lambda)^2. \quad (\text{C2})$$

APPENDIX D: PROOF OF LINEAR FI SCALING FOR THE BINOMIAL MIXTURE

The FI, which we denote by $\mathcal{I}$, can be calculated using the following formula:

$$\begin{aligned} \mathcal{I}(\omega) &= \sum_{b_1, \dots, b_N} \left(\frac{\partial_\omega P(b_1, \dots, b_N)}{P(b_1, \dots, b_N)} \right)^2 P(b_1, \dots, b_N) \\ &= \mathbb{E} \left[\left(\frac{\partial_\omega P(b_1, \dots, b_N)}{P(b_1, \dots, b_N)} \right)^2 \right]. \end{aligned} \quad (\text{D1})$$

We first consider the numerator of Eq. (D1) and take the derivative of Eq. (50),

$$\begin{aligned} \partial_\omega P(N_1 = n_1) &= \sum_{i=1}^{d_E} \alpha_i \binom{N}{n_1} p_i^{n_1} (1 - p_i)^{N-n_1} \frac{n_1 - Np_i}{p_i(1 - p_i)} (\partial_\omega p_i). \end{aligned} \quad (\text{D2})$$

We will choose $t > 0$ such that $p_i \neq 0$ or 1. We can see from Eq. (47) that for $t = 0$, $p_i = 0.5$. Further, the probability depends continuously on the variable t . There therefore exists a choice $t > 0$ such that p_i is not too far from $\frac{1}{2}$, for all i . Similarly, we will choose t such that $\partial_\omega p_i \neq 0$. We can see from Eq. (48) that for $t > 0$ but small, the derivative is

given by

$$\partial_{\omega} p_i = \frac{1}{2} \exp(-\text{Re}(\Gamma_i)t) (\lambda_0 - \lambda_1)t + O(t^2),$$

and so for a choice of small t , the derivative does not vanish. Below, we assume that t is small enough so that for all i , $0 < p_i < 1$, and $\partial_{\omega} p_i \neq 0$.

The full FI can be written as

$$\mathcal{I}(\omega) = \sum_{n_1=0}^N \frac{\left(\sum_{i=1}^{d_E} \alpha_i (\partial_{\omega} p_i) p_i^{n_1} (1-p_i)^{N-n_1} \frac{n_1 - N p_i}{p_i(1-p_i)} \right)^2}{\left(\sum_{i=1}^{d_E} \alpha_i p_i^{n_1} (1-p_i)^{N-n_1} \right)^2} \cdot \left(\binom{N}{n_1} \sum_{i=1}^{d_E} \alpha_i p_i^{n_1} (1-p_i)^{N-n_1} \right). \quad (\text{D3})$$

To show the asymptotic linear scaling of the FI, we will need to prove three important bounds. Let $i_* = \text{argmax} \{p_i \mid \alpha_i > 0\}$ and let $p_* = p_{i_*}$. Such a p_i always exists since we have assumed $\sum_i \alpha_i = 1$ and so there must be at least one nonzero α_i . To prove that the SQL is always achievable, we will need the following three bounds. The proofs of the bounds can be found below.

Lemma D1. $\forall \eta_1 > 0, \epsilon_1 > 0$, and $\forall p_i \neq p_*$, $\exists N_1^{(1)}$ such that $\forall N > N_1^{(1)}$, and $\forall n_1$ satisfying

$$\frac{n_1}{N} > \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_1,$$

we have that

$$\left| \alpha_* p_*^{n_1} (1-p_*)^{N-n_1} \cdot \frac{n_1 - N p_*}{p_* (1-p_*)} (\partial_{\omega} p_*) \right| > \eta_1 \left| \alpha_i p_i^{n_1} (1-p_i)^{N-n_1} \cdot \frac{n_1 - N p_i}{p_i (1-p_i)} (\partial_{\omega} p_i) \right|. \quad (\text{D4})$$

Note that the right-hand side of this equation satisfies

$$\frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_1 > p_i.$$

Lemma D2. $\forall \eta_2 > 0, \epsilon_2 > 0$, and $\forall p_i \neq p_*$, $\exists N_1^{(2)}$ such that $\forall N > N_1^{(2)}$, and $\forall n_1$ satisfying

$$\frac{n_1}{N} > \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_2,$$

we have that

$$\alpha_* p_*^{n_1} (1-p_*)^{N-n_1} > \eta_2 \alpha_i p_i^{n_1} (1-p_i)^{N-n_1}. \quad (\text{D5})$$

Lemma D3. $\forall 1 > \eta_3 > 0, \epsilon_3 > 0$, $\exists N_1^{(3)}$ such that $\forall N > N_1^{(3)}$, $\exists n_1^{(3)} > (p_* - \epsilon_3)N$ such that $n_1 > n_1^{(3)}$ implies that

$$\sum_{n_1 > n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - N p_*}{p_* (1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}} > \eta_3 \sum_{n_1=1}^N \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - N p_*}{p_* (1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}}. \quad (\text{D6})$$

We can combine these three bounds to show linear scaling. Let $p_2 = \max \{p_i \mid \alpha_i \neq 0, p_i < p_*\}$ be the second-largest p_i . If such a p_i does not exist, then there is only one binomial distribution in the mixture. For that particular case, we know that the FI of a binomial distribution scales linearly in N , completing the proof that the SQL is achievable. Indeed, the FI of a binomial distribution with parameter p_* is given by

$$\mathcal{I}(\omega) = \frac{N}{p_*(1-p_*)} (\partial_{\omega} p_*)^2. \quad (\text{D7})$$

We now consider the case in which p_2 is well defined. We choose $\epsilon_1, \epsilon_2, \epsilon_3$ (as in Lemmas D1, D2, and D3, respectively) small enough such that

$$\frac{\log\left(\frac{1-p_2}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_2)}{p_2(1-p_*)}\right)} + \max\{\epsilon_1, \epsilon_2\} < p_* - \epsilon_3.$$

We show that this inequality holds in Lemma D4. Then, for any $\eta_1, \eta_2 > d_E$ and $1 > \eta_3 > 0$, we know that for any N large enough, there exists a corresponding n_1^* such that

$$p_2 < \frac{\log\left(\frac{1-p_2}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_2)}{p_2(1-p_*)}\right)} + \max\{\epsilon_1, \epsilon_2\} < \frac{n_1^*}{N} < p_* - \epsilon_3,$$

and such that Eqs. (D4), (D5), and (D6) hold. Application of these bounds allows us to derive the following inequalities for the FI.

We first remove a few terms of the sum of Eq. (D3). Since all the terms are positive, this leads to a lower bound:

$$\mathcal{I}(\omega) \geq \sum_{n_1 > n_1^*} \binom{N}{n_1} \frac{\left(\sum_{i=1}^{d_E} \alpha_i (\partial_{\omega} p_i) p_i^{n_1} (1-p_i)^{N-n_1} \frac{n_1 - N p_i}{p_i(1-p_i)} \right)^2}{\left(\sum_{i=1}^{d_E} \alpha_i p_i^{n_1} (1-p_i)^{N-n_1} \right)^2}.$$

Next, we apply Lemma D1 to lower bound the numerator:

$$\mathcal{I}(\omega) \geq \sum_{n_1 > n_1^*} \binom{N}{n_1} \cdot \frac{\left((1 - d_E \eta_1^{-1}) \alpha_*(\partial_\omega p_*) p_*^{n_1} (1 - p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)} \right)^2}{\sum_{i=1}^{d_E} \alpha_i p_i^{n_1} (1 - p_i)^{N-n_1}},$$

where we have upper bounded each of the terms associated with p_i with $p_i \neq p_*$ by η_1^{-1} times the p_* term. We give each of these terms a sign opposite to that of the p_* term, so that we obtain a lower bound. We now apply Lemma D2 to bound the denominator:

$$\mathcal{I}(\omega) \geq \sum_{n_1 > n_1^*} \binom{N}{n_1} \cdot \frac{\left((1 - d_E \eta_1^{-1}) \alpha_*(\partial_\omega p_*) p_*^{n_1} (1 - p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)} \right)^2}{(1 + d_E \eta_2^{-1}) \alpha_* p_*^{n_1} (1 - p_*)^{N-n_1}},$$

where we similarly bound the denominator by bounding all the non- p_* terms by η_2^{-1} times the p_* term and assigning them the same sign to obtain a lower bound. We can rewrite this bound as

$$\mathcal{I}(\omega) \geq \frac{(1 - d_E \eta_1^{-1})^2}{1 + d_E \eta_2^{-1}} \eta_3 \cdot \sum_{n_1 > n_1^*} \binom{N}{n_1} \frac{\left(\alpha_*(\partial_\omega p_*) p_*^{n_1} (1 - p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)} \right)^2}{\alpha_* p_*^{n_1} (1 - p_*)^{N-n_1}}.$$

Finally, we apply Lemma D3 to the inequality to obtain

$$\mathcal{I}(\omega) \geq \frac{(1 - d_E \eta_1^{-1})^2}{1 + d_E \eta_2^{-1}} \eta_3 \cdot \sum_{n_1=0}^N \binom{N}{n_1} \frac{\left(\alpha_*(\partial_\omega p_*) p_*^{n_1} (1 - p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)} \right)^2}{\alpha_* p_*^{n_1} (1 - p_*)^{N-n_1}}.$$

We recognize the formula for the FI of a binomial distribution and can therefore substitute the known result, Eq. (D7), to obtain a final lower bound for the FI of the binomial mixture:

$$\mathcal{I}(\omega) \geq \eta_3 \frac{(1 - d_E \eta_1)^2}{1 + d_E \eta_2} \alpha_*^2 (\partial_\omega p_*)^2 \frac{N}{p_*(1 - p_*)}.$$

We see that the final lower bound scales linearly in N , thus completing the proof that the SQL is achievable.

1. Proof of Lemma D1

We will show that $\forall \eta_1 > 0, \epsilon_1 > 0$, and $\forall p_i \neq p_*$, $\exists N_1^{(1)}$ such that $\forall N > N_1^{(1)}$, and $\forall n_1$ satisfying

$$\frac{n_1}{N} > \frac{\log \left(\frac{1-p_i}{1-p_*} \right)}{\log \left(\frac{p_*(1-p_i)}{p_i(1-p_*)} \right)} + \epsilon_1, \quad (\text{D8})$$

we have that

$$\left| \alpha_* p_*^{n_1} (1 - p_*)^{N-n_1} \cdot \frac{n_1 - N p_*}{p_*(1 - p_*)} (\partial_\omega p_*) \right| > \eta_1 \left| \alpha_i p_i^{n_1} (1 - p_i)^{N-n_1} \cdot \frac{n_1 - N p_i}{p_i(1 - p_i)} (\partial_\omega p_i) \right|. \quad (\text{D9})$$

Proof. We can rearrange the terms of Eq. (D9) to obtain the following equation:

$$\left| \frac{p_*(1 - p_i)}{p_i(1 - p_*)} \right|^{n_1} \left| \frac{1 - p_*}{1 - p_i} \right|^N \left| \frac{n_1 - N p_*}{n_1 - N p_i} \right| > \eta_1 \frac{\alpha_i p_*(1 - p_*)}{\alpha_* p_i(1 - p_i)} \left| \frac{\partial_\omega p_i}{\partial_\omega p_*} \right|.$$

We will denote the right-hand side of this equation by $\hat{\eta}$. We remark $\hat{\eta}$ depends on neither N nor n_1 . Taking the natural log of both sides, we obtain

$$n_1 \log \left| \frac{p_*(1 - p_i)}{p_i(1 - p_*)} \right| + N \log \left| \frac{1 - p_*}{1 - p_i} \right| + \log \left| \frac{n_1 - N p_*}{n_1 - N p_i} \right| > \log \hat{\eta}.$$

First, we let $n_1 = N k_1$. We divide the previous equation by N and shuffle terms around to obtain

$$k_1 > \left(\log \left| \frac{p_*(1 - p_i)}{p_i(1 - p_*)} \right| \right)^{-1} \cdot \left(\log \left| \frac{1 - p_i}{1 - p_*} \right| + \frac{1}{N} \log \left| \frac{k_1 - p_i}{k_1 - p_*} \right| + \frac{\log \hat{\eta}}{N} \right).$$

There exists an $N_1^{(1)}$ such that $N > N_1^{(1)}$ implies that the last two terms when expanding the right-hand side of the equation are smaller than ϵ_1 . This yields the inequality

$$k_1 > \frac{\log \left(\frac{1-p_i}{1-p_*} \right)}{\log \left(\frac{p_*(1-p_i)}{p_i(1-p_*)} \right)} + \epsilon_1$$

and so the condition for Eq. (D9) to hold is that

$$k_1 = \frac{n_1}{N} > \frac{n_1^{(1)}}{N}, \quad N > N_1^{(1)},$$

where we let

$$\frac{n_1^{(1)}}{N} = \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_1. \quad (\text{D10})$$

■

In the proofs of the main text, we have used the following property of $n_1^{(1)}$.

Lemma D4. For $p_i < p_*$ and $0 < p_i, p_* < 1$, we have

$$\frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} < p_*.$$

Proof. We can rewrite this equation as

$$\log\left(\frac{1-p_i}{1-p_*}\right) - p_* \log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right) < 0.$$

Taking the exponentials of both sides, we obtain the inequality

$$\exp\left(\log\left(\frac{1-p_i}{1-p_*}\right) - p_* \log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)\right) < 1. \quad (\text{D11})$$

We will now prove this inequality. First, we remark that for $p_i = p_*$, we have that

$$\exp\left(\log\left(\frac{1-p_*}{1-p_*}\right) - p_* \log\left(\frac{p_*(1-p_*)}{p_*(1-p_*)}\right)\right) = 1.$$

We now consider the derivative of the left-hand side with respect to p_* :

$$\begin{aligned} \frac{\partial}{\partial p_*} \exp\left(\log\left(\frac{1-p_i}{1-p_*}\right) - p_* \log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)\right) = \\ - \left(\frac{1-p_i}{1-p_*}\right)^{1-p_*} \left(\frac{p_i}{p_*}\right)^{p_*} \left[\log\left(\frac{p_*}{p_i}\right) + \log\left(\frac{1-p_i}{1-p_*}\right)\right]. \end{aligned}$$

We see that the derivative is strictly negative for all $p_i < p_*$ such that $0 < p_i, p_* < 1$. In other words, for $p_* > p_i$, as we increase p_* , the function defined by the left-hand side of Eq. (D11) decreases from the value of 1 at $p_i = p_*$. Therefore, the inequality of Eq. (D11) must hold. ■

Equivalently, by the density of real numbers, this lemma states that there exists an $\epsilon > 0$ such that

$$\frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} < p_* - \epsilon.$$

2. Proof of Lemma D2

We now consider the second bound. We show that $\forall \eta_2 > 0, \forall p_i \neq p_*, \exists N_1^{(2)}$ such that $\forall N > N_1^{(2)}, \exists n_1^{(2)}$ such that $n_1 > n_1^{(2)}$ implies

$$\alpha_* p_*^{n_1} (1-p_*)^{N-n_1} > \eta_2 \alpha_i p_i^{n_1} (1-p_i)^{N-n_1}.$$

Proof. We can rearrange the equation by following the same steps as in the previous derivation, to obtain

$$k_1 > \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \frac{\log \eta_2}{N}.$$

For all $\epsilon_2 > 0$, there exists $N_1^{(2)}$ such that for $N > N_1^{(2)}$, the last term is smaller than ϵ_2 . This gives us

$$k_1 > \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_2. \quad (\text{D12})$$

Once again, the choice of $N_1^{(2)}$ is somewhat arbitrary, as we can choose to suppress ϵ_2 more or less in exchange for lowering the value of $n_1^{(2)}$, which we define as

$$\frac{n_1^{(2)}}{N} = \frac{\log\left(\frac{1-p_i}{1-p_*}\right)}{\log\left(\frac{p_*(1-p_i)}{p_i(1-p_*)}\right)} + \epsilon_2.$$

Once again, we must choose $N_1^{(2)}$ big enough such that $n_1^{(2)} < N$ and such that there exist integer values of n_1 such that $n_1^{(2)} < n_1 < N$. ■

3. Proof of Lemma D3

Finally, we prove a final bound. We will show that $\forall 1 > \eta_3 > 0, \epsilon_3 > 0, \exists N_1^{(3)}$ such that $\forall N > N_1^{(3)}, \exists n_1^{(3)} > (p_* - \epsilon_3)N$ such that $n_1 > n_1^{(3)}$ implies that

$$\begin{aligned} \sum_{n_1 > n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)}\right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}} \\ > \eta_3 \sum_{n_1=1}^N \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - N p_*}{p_*(1-p_*)}\right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}}. \quad (\text{D13}) \end{aligned}$$

Proof. We can decompose the right-hand side (RHS) into partial sums,

$$\begin{aligned} \text{RHS} &= \eta_3 \sum_{n_1 \leq n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - Np_*}{p_*(1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}} \\ &\quad + \eta_3 \sum_{n_1 > n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - Np_*}{p_*(1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}}. \end{aligned}$$

We can bound the first term,

$$\begin{aligned} &\eta_3 \sum_{n_1 \leq n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - Np_*}{p_*(1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}} \\ &= \eta_3 \sum_{n_1 \leq n_1^{(3)}} \binom{N}{n_1} p_*^{n_1} (1-p_*)^{N-n_1} \left(\frac{n_1 - Np_*}{p_*(1-p_*)} \right)^2 \\ &\leq \eta_3 \frac{N^2}{(p_*(1-p_*))^2} \sum_{n_1 \leq n_1^{(3)}} \binom{N}{n_1} p_*^{n_1} (1-p_*)^{N-n_1}. \end{aligned}$$

This final sum can now be bounded from above using the Chernoff bound [78],

$$\sum_{n_1 \leq n_1^{(3)}} \binom{N}{n_1} p_*^{n_1} (1-p_*)^{N-n_1} \leq \exp \left(-ND \left(\left\| \frac{n_1^{(3)}}{N} \right\| p_* \right) \right),$$

where $D \left(\frac{n_1^{(3)}}{N} \| p_* \right)$ is the Kullback-Leibler divergence for a binomial distribution. For any $\epsilon_3 > 0$, the Kullback-Leibler divergence is positive when $\frac{n_1^{(3)}}{N} < p_* - \epsilon_3$. For a positive Kullback-Leibler divergence, the partial sum is exponentially suppressed in the limit of large N . In other words, $\forall \epsilon_3 > 0, \exists N_1^{(3)}$ such that $N > N_1^{(3)}$ implies that

$$\eta_3 \frac{N^2}{p_*(1-p_*)} \sum_{n_1 \leq n_1^{(3)}} p_*^{n_1} (1-p_*)^{N-n_1} < \epsilon_3$$

and so

$$\text{RHS} < \eta_3 \sum_{n_1 > n_1^{(3)}} \binom{N}{n_1} \frac{\left(p_*^{n_1} (1-p_*)^{N-n_1} \frac{n_1 - Np_*}{p_*(1-p_*)} \right)^2}{p_*^{n_1} (1-p_*)^{N-n_1}} + \epsilon_3.$$

Taking $\epsilon_3 < 1 - \eta_3$ proves the inequality of Eq. (D6). ■

APPENDIX E: ESTIMATE FOR THE VARIANCE OF THE FI ESTIMATOR

We calculate the variance of Eq. (55):

$$\begin{aligned} \text{Var} [\hat{\mathcal{I}}(\omega)] &= \text{Var} \left[\frac{1}{S} \sum_{s=1}^S \left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^2 \right] \\ &= \frac{1}{S^2} \sum_{s=1}^S \text{Var} \left[\left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^2 \right]. \end{aligned} \quad (\text{E1})$$

There are no covariance terms in Eq. (E1), since all of our samples are uncorrelated, as they are produced via direct sampling of the probability distribution. Using the definition of the variance, we obtain

$$\begin{aligned} \text{Var} [\hat{\mathcal{I}}(\omega)] &= \frac{1}{S^2} \sum_{s=1}^S \text{Var} \left[\left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^2 \right] \\ &= \frac{1}{S^2} \sum_{s=1}^S \left(\mathbb{E} \left[\left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^4 \right] \right. \\ &\quad \left. - \mathbb{E} \left[\left(\frac{\partial_\omega P(x_1^{(s)}, \dots, x_N^{(s)})}{P(x_1^{(s)}, \dots, x_N^{(s)})} \right)^2 \right]^2 \right). \end{aligned}$$

The expectation value is independent of the sample. Further, we recognize the definition of the FI in the second expectation value. The equation then becomes

$$\text{Var} [\hat{\mathcal{I}}(\omega)] = \frac{1}{S} \mathbb{E} \left[\left(\frac{\partial_\omega P(x_1, \dots, x_N)}{P(x_1, \dots, x_N)} \right)^4 \right] - \frac{\mathcal{I}(\omega)^2}{S}.$$

We can upper bound the variance:

$$\text{Var} [\hat{\mathcal{I}}(\omega)] \leq \frac{1}{S} \mathbb{E} \left[\left(\frac{\partial_\omega P(x_1, \dots, x_N)}{P(x_1, \dots, x_N)} \right)^4 \right]. \quad (\text{E2})$$

The right-hand side of this equation can be estimated with

$$\frac{1}{S} \sum_{s=1}^S \left(\frac{\partial_\omega P(x_1, \dots, x_N)}{P(x_1, \dots, x_N)} \right)^4. \quad (\text{E3})$$

We therefore take Eq. (E3) to be a biased estimator of our variance, allowing us to draw error bars in Fig. 3. Although the estimator is biased, it converges to a value that is higher than the true value. The small error bars on Fig. 3 are therefore a good indicator of the convergence of the FI estimator. This method remains a heuristic, as although Eq. (E3) converges for high-enough S , it is unclear how different values of N affect the convergence and whether the 25 000 samples are sufficient.

APPENDIX F: LIST OF ACRONYMS

CPTP	Completely positive trace-preserving
FI	Fisher information, or classical Fisher information
HL	Heisenberg limit
HMM	Hidden Markov model
HNLS	Hamiltonian not in Lindblad span
HNES	Hamiltonian not in extended span
HNELS	Hamiltonian not in extended Lindblad span
KL	Knill-Laflamme
QEC	Quantum error correction
QFI	Quantum Fisher information
SQL	Standard quantum limit

- [1] R. Schnabel, N. Mavalvala, D. E. McClelland, and P. K. Lam, Quantum metrology for gravitational wave astronomy, *Nat. Commun.* **1**, 121 (2010).
- [2] H. Katori, Optical lattice clocks and quantum metrology, *Nat. Photonics* **5**, 203 (2011).
- [3] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photonics* **5**, 222 (2011).
- [4] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin, High-sensitivity diamond magnetometer with nanoscale resolution, *Nat. Phys.* **4**, 810 (2008).
- [5] G. Balasubramanian, I. Y. Chan, R. Kolesov, M. Al-Hmoud, J. Tisler, C. Shin, C. Kim, A. Wojcik, P. R. Hemmer, A. Krueger, T. Hanke, A. Leitenstorfer, R. Bratschkitsch, F. Jelezko, and J. Wrachtrup, Nanoscale imaging magnetometry with diamond spins under ambient conditions, *Nature* **455**, 648 (2008).
- [6] F. Dolde, H. Fedder, M. W. Doherty, T. Nöbauer, F. Rempp, G. Balasubramanian, T. Wolf, F. Reinhard, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup, Electric-field sensing using single diamond spins, *Nat. Phys.* **7**, 459 (2011).
- [7] R. Kleiner, D. Koelle, F. Ludwig, and J. Clarke, Superconducting quantum interference devices: State of the art and applications, *Proc. IEEE* **92**, 1534 (2004).
- [8] S. Danilin, N. Nugent, and M. Weides, Quantum sensing with tunable superconducting qubits: Optimization and speed-up, *New J. Phys.* **26**, 103029 (2024).
- [9] C. W. Helstrom, Quantum detection and estimation theory, *J. Stat. Phys.* **1**, 231 (1969).
- [10] M. G. A. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inf.* **07**, 125 (2009).
- [11] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Increasing sensing resolution with error correction, *Phys. Rev. Lett.* **112**, 150801 (2014).
- [12] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Quantum error correction for metrology, *Phys. Rev. Lett.* **112**, 150802 (2014).
- [13] W. Dür, M. Skotiniotis, F. Froewis, and B. Kraus, Improved quantum metrology using quantum error correction, *Phys. Rev. Lett.* **112**, 080801 (2014).
- [14] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, *Quantum* **1**, 27 (2017).
- [15] R. Demkowicz-Dobrzański, J. Czapkowski, and P. Sekatski, Adaptive quantum metrology under general Markovian noise, *Phys. Rev. X* **7**, 041009 (2017).
- [16] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, *Nat. Commun.* **9**, 78 (2018).
- [17] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, *PRX Quantum* **2**, 010343 (2021).
- [18] T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A. Sushkov, M. Lukin, A. Retzker, B. Naydenov, L. McGuinness, and F. Jelezko, Quantum metrology enhanced by repetitive quantum error correction, *Phys. Rev. Lett.* **116**, 230502 (2016).
- [19] Y. Matsuzaki and S. Benjamin, Magnetic-field sensing with quantum error detection under the effect of energy relaxation, *Phys. Rev. A* **95**, 032303 (2017).
- [20] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2007).
- [21] Á. Rivas, S. F. Huelga, and M. B. Plenio, Entanglement and non-Markovianity of quantum evolutions, *Phys. Rev. Lett.* **105**, 050403 (2010).
- [22] D. Chruściński, Dynamical maps beyond Markovian regime, *Physics Reports* **992**, 1 (2022).
- [23] H.-P. Breuer, E.-M. Laine, and J. Piilo, Measure for the degree of non-Markovian behavior of quantum processes in open systems, *Phys. Rev. Lett.* **103**, 210401 (2009).
- [24] X.-M. Lu, X. Wang, and C. P. Sun, Quantum Fisher information flow and non-Markovian processes of open systems, *Phys. Rev. A* **82**, 042103 (2010).
- [25] S. Milz and K. Modi, Quantum stochastic processes and quantum non-Markovian phenomena, *PRX Quantum* **2**, 030201 (2021).
- [26] S. Milz, D. Egloff, P. Taranto, T. Theurer, M. B. Plenio, A. Smirne, and S. F. Huelga, When Is a non-Markovian quantum process classical? *Phys. Rev. X* **10**, 041049 (2020).
- [27] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum metrology in non-Markovian environments, *Phys. Rev. Lett.* **109**, 233601 (2012).
- [28] K. Macieszczak, Zeno limit in frequency estimation with non-Markovian environments, *Phys. Rev. A* **92**, 010102 (2015).
- [29] F. Riberi, L. M. Norris, F. Beaudoin, and L. Viola, Frequency estimation under non-Markovian spatially correlated quantum noise, *New J. Phys.* **24**, 103011 (2022).
- [30] D. Xie and A. M. Wang, Quantum metrology in correlated environments, *Phys. Lett. A* **378**, 2079 (2014).
- [31] 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).
- [32] W. W. Cheng, Nonequilibrium effects on the precision of parameter estimation, *Ann. Phys.* **535**, 2300180 (2023).
- [33] A. Altherr and Y. Yang, Quantum metrology for non-Markovian processes, *Phys. Rev. Lett.* **127**, 060501 (2021).
- [34] Y. Yang, Memory effects in quantum metrology, *Phys. Rev. Lett.* **123**, 110501 (2019).

- [35] S. Kurdzialek, P. Dulian, J. Majsak, S. Chakraborty, and R. Demkowicz-Dobrzański, Quantum metrology using quantum combs and tensor network formalism, *New J. Phys.* **27**, 013019 (2025).
- [36] S. Kurdzialek, F. Albarelli, and R. Demkowicz-Dobrzański, Universal bounds in quantum metrology in presence of correlated noise, [arXiv:2410.01881](https://arxiv.org/abs/2410.01881).
- [37] A. Tanggara, M. Gu, and K. Bharti, Strategic code: A unified spatio-temporal framework for quantum error-correction, [arXiv:2405.17567](https://arxiv.org/abs/2405.17567).
- [38] M. Neeley, M. Ansmann, R. C. Bialczak, M. Hofheinz, N. Katz, E. Lucero, A. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis, Process tomography of quantum memory in a Josephson-phase qubit coupled to a two-level state, *Nat. Phys.* **4**, 523 (2008).
- [39] Y. Shalibo, Y. Rofe, D. Shwa, F. Zeides, M. Neeley, J. M. Martinis, and N. Katz, Lifetime and coherence of two-level defects in a Josephson junction, *Phys. Rev. Lett.* **105**, 177001 (2010).
- [40] A. M. Zagoskin, S. Ashhab, J. R. Johansson, and F. Nori, Quantum two-level systems in Josephson junctions as naturally formed qubits, *Phys. Rev. Lett.* **97**, 077001 (2006).
- [41] A. Agarwal, L. P. Lindoy, D. Lall, F. Jamet, and I. Rungger, Modelling non-Markovian noise in driven superconducting qubits, *Quantum Sci. Technol.* **9**, 035017 (2024).
- [42] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu, Decoherence in Josephson qubits from dielectric loss, *Phys. Rev. Lett.* **95**, 210503 (2005).
- [43] J. H. Cole, C. Müller, P. Bushev, G. J. Grabovskij, J. Lisenfeld, A. Lukashenko, A. V. Ustinov, and A. Shnirman, Quantitative evaluation of defect-models in superconducting phase qubits, *Appl. Phys. Lett.* **97**, 252501 (2010).
- [44] J. Burnett, L. Faoro, I. Wisby, V. L. Gurtovoi, A. V. Chernykh, G. M. Mikhailov, V. A. Tulín, R. Shaikhaidarov, V. Antonov, P. J. Meeson, A. Y. Tzalenchuk, and T. Lindström, Evidence for interacting two-level systems from the $1/f$ noise of a superconducting resonator, *Nat. Commun.* **5**, 4119 (2014).
- [45] D. L. Mickelsen, H. M. Carruzzo, and C. C. Yu, Interacting two-level systems as a source of $1/f$ charge noise in quantum dot qubits, *Phys. Rev. B* **108**, 195307 (2023).
- [46] L. Childress, M. V. Gurudev Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin, Coherent dynamics of coupled electron and nuclear spin qubits in diamond, *Science* **314**, 281 (2006).
- [47] J. R. Maze, J. M. Taylor, and M. D. Lukin, Electron spin decoherence of single nitrogen-vacancy defects in diamond, *Phys. Rev. B* **78**, 094303 (2008).
- [48] Z.-H. Wang and S. Takahashi, Spin decoherence and electron spin bath noise of a nitrogen-vacancy center in diamond, *Phys. Rev. B* **87**, 115122 (2013).
- [49] L. Rabiner and B. Juang, An introduction to hidden Markov models, *IEEE ASSP Mag.* **3**, 4 (1986).
- [50] A. Monras, A. Beige, and K. Wiesner, Hidden quantum Markov models and non-adaptive read-out of many-body states, *Appl. Math. Comput. Sci.* **3**, 93 (2011).
- [51] G. Sun, Z. Zhou, B. Mao, X. Wen, P. Wu, and S. Han, Entanglement dynamics of a superconducting phase qubit coupled to a two-level system, *Phys. Rev. B* **86**, 064502 (2012).
- [52] C. Müller, J. H. Cole, and J. Lisenfeld, Towards understanding two-level-systems in amorphous solids: Insights from quantum circuits, *Rep. Prog. Phys.* **82**, 124501 (2019).
- [53] E. Knill, R. Laflamme, and L. Viola, Theory of quantum error correction for general noise, *Phys. Rev. Lett.* **84**, 2525 (2000).
- [54] A. Peres, Reversible logic and quantum computers, *Phys. Rev. A* **32**, 3266 (1985).
- [55] V. V. Albert and P. Faist, Quantum repetition code, in *The Error Correction Zoo* (2022), https://errorcorrectionzoo.org/c/quantum_repetition.
- [56] D. Gottesman, *Surviving as a quantum computer in a classical world (2024 draft)* (UMD, College Park, Maryland, 2024), <https://www.cs.umd.edu/~dgottesman/>.
- [57] B. M. Terhal and G. Burkard, Fault-tolerant quantum computation for local non-Markovian noise, *Phys. Rev. A* **71**, 012336 (2005).
- [58] P. Aliferis, D. Gottesman, and J. Preskill, Quantum accuracy threshold for concatenated distance-3 codes, *Quantum Info. Comput.* **6**, 97 (2006).
- [59] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush, D. Bacon, J. C. Bardin, J. Basso, A. Bengtsson, S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, *Nature* **614**, 676 (2023).
- [60] M. McEwen, L. Faoro, K. Arya, A. Dunsworth, T. Huang, S. Kim, B. Burkett, A. Fowler, F. Arute, J. C. Bardin, A. Bengtsson, A. Bilmes, B. B. Buckley, N. Bushnell, Z. Chen, R. Collins, S. Demura, A. R. Derk, C. Erickson, M. Giustina *et al.*, Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits, *Nat. Phys.* **18**, 107 (2022).
- [61] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, *J. Phys. A: Math. Theor.* **53**, 023001 (2019).
- [62] P. Bocchieri and A. Loinger, Quantum recurrence theorem, *Phys. Rev.* **107**, 337 (1957).
- [63] L. S. Schulman, Note on the quantum recurrence theorem, *Phys. Rev. A* **18**, 2379 (1978).
- [64] I. C. Percival, Almost periodicity and the quantal H theorem, *J. Math. Phys.* **2**, 235 (1961).
- [65] C. Corduneanu, *Almost Periodic Oscillations and Waves* (Springer, New York, 2009).
- [66] Y.-N. Lu, Y.-R. Zhang, G.-Q. Liu, F. Nori, H. Fan, and X.-Y. Pan, Observing information backflow from controllable non-Markovian multichannels in diamond, *Phys. Rev. Lett.* **124**, 210502 (2020).
- [67] M. Radaelli, G. T. Landi, K. Modi, and F. C. Binder, Fisher information of correlated stochastic processes, *New J. Phys.* **25**, 053037 (2023).
- [68] M. Guta and J. Kiukas, Equivalence classes and local asymptotic normality in system identification for quantum Markov chains, *Commun. Math. Phys.* **335**, 1397 (2015).
- [69] J. C. Spall, Monte Carlo computation of the Fisher information matrix in nonstandard settings, *J. Comput. Graph. Stat.* **14**, 889 (2005).

- [70] J. C. Spall, in *2008 American Control Conference (IEEE, Seattle, Washington, 2008)*, pp. 2395–2400.
- [71] J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices, *J. Math. Phys.* **6**, 440 (1965).
- [72] T. Can, Random Lindblad dynamics, *J. Phys. A: Math. Theor.* **52**, 485302 (2019).
- [73] Z. Mann, S. Zhou, N. Cao, and R. Laflamme, zach-et-m/nm-qec-metrology: paper code, 2025.
- [74] D. W. Leung and M. Vasmer, Quantum error correction and fault tolerance lecture notes (University of Waterloo, Waterloo, Ontario, 2024), <https://www.math.uwaterloo.ca/~wcleung/qic890-w2024.html>.
- [75] M. Liddy, *Optimal Control Theory Techniques for Nitrogen Vacancy Ensembles* (University of Waterloo, 2022).
- [76] K. Rama Koteswara Rao and D. Suter, Level anti-crossings of a nitrogen-vacancy center in diamond: Decoherence-free subspaces and 3D sensors of microwave magnetic fields, *New J. Phys.* **22**, 103065 (2020).
- [77] Editorial Board, Encyclopedia of Mathematics, http://encyclopediaofmath.org/index.php?title=Editorial_Board&oldid=41109, “Dirichlet theorem”.
- [78] R. Arratia and L. Gordon, Tutorial on large deviations for the binomial distribution, *Bull. Math. Biol.* **51**, 125 (1989).