npj | quantum information

Article

Published in partnership with The University of New South Wales

https://doi.org/10.1038/s41534-025-01044-7

Lindblad-like quantum tomography for non-Markovian quantum dynamical maps

S. Varona[1,2] ✉, M. Müller[1,3] & A. Bermudez[2]

We introduce Lindblad-like quantum tomography (LℓQT) as a quantum characterization technique of time-correlated noise in quantum information processors. This approach enables the estimation of time-local master equations, including their possible negative decay rates, by maximizing a likelihood function subject to dynamical constraints. We discuss LℓQT for the dephasing dynamics of single qubits in detail, which allows for a neat understanding of the importance of including multiple snapshots of the quantum evolution in the likelihood function, and how these need to be distributed in time depending on the noise characteristics. By a detailed comparative study employing both frequentist and Bayesian approaches, we assess the accuracy and precision of LℓQT of a dephasing quantum dynamical map that goes beyond the Lindblad limit, focusing on two different microscopic noise models that can be realised in either trapped-ion or superconducting-circuit architectures. We explore the optimization of the distribution of measurement times to minimize the estimation errors, assessing the superiority of each learning scheme conditioned on the degree of non-Markovinity of the noise, and setting the stage for future experimental designs of non-Markovian quantum tomography.

The field of quantum information processing has witnessed a remarkable progress in the last years[1–5], as evidenced by the key advances reported in[6–18]. This progress lays the groundwork for the eventual demonstration of practical quantum advantage in real-world applications[19]. Central to these advancements and future breakthroughs is the exceptional level of isolation and control achieved over quantum information processors (QIPs), enabling the application and integration of various strategies to fight against the accumulation of errors during quantum computations. These strategies can be implemented either during the processing of quantum information or, alternatively, post-measurement, falling into three distinct categories: quantum error suppression (QES)[20–23], quantum error mitigation (QEM)[24–27], and quantum error correction (QEC)[28–30].

The development and optimization of these techniques for specific architectures greatly benefits from a comprehensive understanding of the underlying sources of noise, including a thorough quantum characterization, verification and validation (QCVV) of the noise models[31–33]. By addressing the noise characteristics, researchers can tailor their strategies to suit the specific requirements of different platforms, thereby enhancing the reliability and performance of QIPs. For instance, the presence of spatial and temporal noise correlations is a critical consideration in some techniques of QES, such as decoherence-free subspaces[34–38] and dynamical decoupling[39–44], respectively. Likewise, in the context of QEC, the presence of spatial[45–47] and temporal[48–50] noise correlations must be carefully accounted for when considering fault-tolerant quantum computation beyond the idealized regime of independent and identically distributed errors. In this work, we will focus on the characterization of qubit dynamics under temporally correlated noise. This can actually lead to a non-Markovian quantum evolution, which will require reconsidering some of the established characterization tools for the dynamics of Markovian open quantum systems. Before delving into more specific details about the characterization of non-Markovian quantum evolutions, we note that the degree of non-Markovianity[51–53] can play a role in the effectiveness of QES[54,55] and QEM[56,57] techniques.

Within the set of QCVV techniques[31–33], quantum process tomography (QPT) aims at characterizing the most generic type of process that can account for the evolution of a quantum system, solely constrained by the laws of quantum mechanics[58–62]. For a specific evolution time, after preparing and measuring the system in an informationally complete setting, one can make an estimate of the completely-positive trace-preserving (CPTP) quantum channel[63,64] that determines a snapshot of the quantum evolution. Therefore, QPT has been applied for the characterization of quantum gates in various experimental QIPs[65–73], typically restricted to small number of qubits. In addition to the inherent complexity of QPT as the number of qubits increases, this characterization must be repeated for each instant of time of interest, in order to obtain a coarse-grained reconstruction of the full, i.e. a one-parameter family of CPTP channels[74–76] that governs the

[1]Institute for Quantum Information, RWTH Aachen University, Aachen, Germany. [2]Instituto de Física Teórica, UAM-CSIC, Universidad Autónoma de Madrid, Cantoblanco, Madrid, Spain. [3]Peter Grünberg Institute, Theoretical Nanoelectronics, Forschungszentrum Jülich, Jülich, Germany.
✉e-mail: svarona@physik.rwth-aachen.de

time evolution of the quantum system. Although repeating QPT can allow to characterise non-Markovian evolutions, the associated overhead can limit the precision in architectures where the number of measurements shots cannot be sufficiently large[77].

A strategy to overcome this limitation is to focus on the estimation of the generators of the noisy dynamics, rather than on the various coarse-grained snapshots. For time-homogeneous quantum dynamical maps, which form a semigroup, the time evolution can be described by the exponential of a Lindblad super-operator[74,78,79]. Although one may expect that the generators of this type of Markovian noise can be obtained by simple algebraic manipulations of a single QPT snapshot at any arbitrary time, this approach can lead to inconsistencies[80–83] due to the branches of the complex logarithm. Therefore, alternative QCVV techniques are required. One possibility is use the Lindbladian generators for a parametrization of the quantum evolution, which can then be inferred using different learning strategies[65,81,82,84–90]. Lindblad learning aims to estimate the Hamiltonian under which the system evolves and, additionally, the jump operators and dissipation rates that govern the non-unitary part of the dynamics of the noisy QIP. However, it is important to note that Lindblad learning is based on the Lindblad master equation, valid only for Markovian system-environment interactions, i.e., memoryless interactions in which information flows from the system to the environment but never flows back.

However, noise in real QIPs does not always fall in this category, and temporal correlations and even non-Markovianity can play an important role, as alluded above in connection to QES, QEM and QEC. Hence, it would be desirable to extend the Lindblad learning to encompass non-Markovian noise scenarios, such that the quantum dynamical maps are no longer a semigroup, nor can they be divided into the composition of sequential CP channels at any intermediate time[75,76]. Several characterization techniques for related problems have been developed in recent years. Quantum noise spectroscopy protocols[91,92] have evolved into characterization tools of non-Markovian systems, as in ref. 93, where the authors develop a filter-function formalism based on frames. A general non-Markovian quantum process framework was proposed by ref. 94 allowing to characterize non-Markovian quantum processes using process tensor tomography (PTT). This later led to proposals of quantum non-Markovian process tomography[95]. Reference[96] introduced into this PTT frameworks the possibility of memory effects within the system control itself and from undesirable interactions between the control and the environment. This contrasts ref. 94 where perfect gates were assumed. Reference[97] characterizes certain type of non-Markovian errors in quantum gates and[98] proposes also characterization of quantum gates but with a compressed approach that reduces the resources needed in comparison to, for instance, ref. 96.

In general, any quantum evolution that results from the coupling of a quantum system to a larger environment, or to a set of noisy controls modeled by stochastic processes, can be expressed in terms of a time-local master equation by using a time-convolutionless formulation[99] of the Nakajima-Zwanzig integro-differential equation[74,100,101]. These time-local master equations generalize the aforementioned Lindblad master equation[78,79], and can be expressed in a canonical form that connects directly with the degree of non-Markovianity[102]. In essence, the characterization of these time-local master equations would require a time-dependent parametrization of the Hamiltonian, jump operators and dissipation rates, which can then be incorporated into a maximum-likelihood estimation that parallels the Markovian Lindblad limit[87–90]. In this work, we call this QCVV technique *Lindblad-like quantum tomography* (LℓQT), and develop it in the simplest possible scenario: the dephasing dynamics of a single qubit. We present a detailed comparative study of this QCVV technique, considering both a frequentist and a Bayesian approach for the statistical inference. We consider minimal dephasing models, both semi-classical and fully quantum-mechanical, in which the temporal correlations and degree of non-Markovianity can be independently controlled. By making a careful connection to the the theory of asymptotic inference and Bayesian estimation, we quantify both the accuracy and precision of LℓQT. We discuss how the amount of temporal correlations and the degree of non-Markovianity can play a key role in deciding which of the two approaches is preferable when learning the non-Lindblad qubit dephasing.

This article is organized as follows. In subsection "Learning time-local master equations by maximum-likelihood estimation" we review the techniques of Lindblad quantum tomography and present LℓQT, a generalization of Lindblad learning that allows us to characterize non-Markovian noise. In subsections "Frequentist approach to non-Markovian inference" and "Bayesian approach to non-Markovian inference" two approaches to LℓQT are presented. Subsection "Lindblad-like quantum tomography for non-Markovian dephasing" presents LℓQT applied to dephasing noise. The frequentist and Bayesian approaches are later compared in a performance analysis in subsections "Markovian semi-classical dephasing" and "Non-Markovian quantum dephasing", where we also study how measurement times should be selected to reduce the number of necessary measurements and the error in the estimation of noise parameters.

Results

Learning time-local master equations by maximum-likelihood estimation

The Lindblad master equation generalizes the Schrödinger equation to open and noisy quantum systems[74,78,79], and describes the non-unitary time evolution of the density matrix of the system, defined as a positive-definite unit-trace linear operator $\rho \in \mathsf{D}(\mathscr{H}_S) \subset \mathsf{L}(\mathscr{H}_S)$ in a Hilbert space of dimension $d = \dim \mathscr{H}_S$[64]. This master equation can be written in terms of an infinitesimal generator $\mathrm{d}\rho/\mathrm{d}t = \mathscr{L}_{H,G}(\rho)$, namely

$$\mathscr{L}_{H,G}(\rho) = -\mathrm{i}\big[H,\rho\big] + \sum_{\alpha,\beta=1}^{d^2-1} G_{\alpha\beta}\left(E_\alpha \rho E_\beta^\dagger - \frac{1}{2}\left\{E_\beta^\dagger E_\alpha, \rho\right\}\right), \qquad (1)$$

where the Hamiltonian $H \in \mathsf{Herm}(\mathscr{H}_S)$ is a Hermitian operator, and we have introduced the so-called dissipation Lindblad matrix, a positive semidefinite matrix $G \in \mathsf{Pos}(\mathbb{C}^{d^2-1})$. Here, $\mathscr{B} = \{E_0 = \mathbb{1}_d, E_\alpha : \alpha \in \{1,\cdots,d^2-1\}\}$ forms an operator basis $\mathsf{L}(\mathscr{H}_S) = \mathrm{span}\{\mathscr{B}\}$ and, together with the Lindblad matrix, determines the dissipative non-unitary dynamics of the system. Diagonalizing the Lindblad matrix, we obtain

$$\mathscr{L}_{H,G}(\rho) = -\mathrm{i}\big[H,\rho\big] + \sum_{n=1}^{d^2-1} \gamma_n\left(L_n \rho L_n^\dagger - \frac{1}{2}\left\{L_n^\dagger L_n, \rho\right\}\right), \qquad (2)$$

where $\{L_n : n \in \{1,\cdots,d^2-1\}\} \in \mathsf{L}(\mathscr{H}_S)$ are the jump operators responsible of generating the different noise processes with dissipative decay rates $\gamma_n \in \mathbb{R}^+$. The goal of Lindblad learning is to estimate $\mathscr{L}_{H,G}$ or, equivalently the system Hamiltonian and the dissipation rates and jump operators $\{H, \gamma_n, L_n\}$, using a finite number of measurements[65,81,82,84–89,103]. In particular, our work starts from a maximum-likelihood approach[104] to Lindbladian quantum tomography (LQT)[87,89,90].

In the general case, LQT involves preparing an informationally complete set of initial states $s \in \mathbb{S}_0$, allowing the system to evolve over a set of times $i \in \mathbb{I}_t$, and performing measurements in different basis $b \in \mathbb{M}_b$, with corresponding outcomes $m_b \in \mathbb{M}_{m_b}$ (or more generally using a POVM). These independent configurations, consisting of initial states, evolution times, and measurement outcomes, provide the necessary data to estimate the Hamiltonian, dissipation rates, and jump operators, ensuring a complete reconstruction of the Lindbladian dynamics. LQT makes use of a total of N_{shot} measurement shots, also known as trials in the context of statistics, which will be distributed among the different initial states, instants of time and measurement basis $N_{\mathrm{shot}} = \sum_{s,i,b} N_{s,i,b}$. Therefore, the total number of measurement shots is a function of number of initial states, measurement basis, measurement outcomes and measurement times, expressed as $N_{\mathrm{shot}}(|\mathbb{S}_0|, |\mathbb{M}_b|, |\mathbb{M}_{m_b}|, |\mathbb{I}_t|)$, where $|\mathbb{A}|$ denotes the cardinality of the set $\mathbb{A}$, i.e., the number of elements in the set. In the experiment, one would

count the number of times N_{s,i,b,m_b} that the m_b outcome is obtained for each of the configurations, such that $N_{s,i,b} = \sum_{m_b} N_{s,i,b,m_b}$. This provides a data set $\mathbb{D} = \{N_{s,i,b,m_b}\}$ that can be understood as a random sample of the corresponding random variable $\tilde{N}_{s,i,b,m_b}$ obtained from N_{shot} experimental measurements. We use tildes to refer to stochastic variables. In this case, the probability of obtaining the outcome m_b when measuring in basis b is given by $p_{s,i,b}(m_b) = \text{Tr}\left\{M_\mu \mathscr{E}_{t_i,t_0}(\rho_{0,s})\right\}$, where M_μ represents the corresponding projector of the combination (b, m_b). $\mathscr{E}_{t,t_0} \in \mathsf{C}(\mathscr{H}_S)$ is a one-parameter family of completely-positive trace-preserving (CPTP) channels[64] describing the actual time evolution of the noisy quantum system. Each measurement configuration gives rise to an independent binomial distribution with $N_{s,i,b}$ trials.

The data set $\mathbb{D}$ can be used to estimate the Hamiltonian H and Lindblad matrix G by maximizing the likelihood function, which is defined as the probability distribution of the combination of all the independent binomials $p_{s,i,b}(m_b)$. These binomials can be approximated by the Lindbladian of Eq. (1), namely $p_{s,i,b}(m_b) \mapsto p^{\text{L}}_{s,i,b}(m_b) = \text{Tr}\{M_\mu e^{(t_i-t_0)\mathscr{L}_{H,G}}(\rho_{0,s})\}$. The larger this likelihood function is for a given pair H, G, the better the Lindbladian description approximates the observed data[104]. Taking the negative logarithm of the likelihood function we obtain a cost function

$$\mathsf{C}_{\text{L}}(H, G) = -\sum_{s,i,b,m_b} N_{s,i,b,m_b} \log p^{\text{L}}_{s,i,b}(m_b), \tag{3}$$

with the minimum located in the same place as the maximum of the original likelihood. The optimization problem is thus converted into a non-linear minimization of this Lindbladian cost function. By minimizing this non-convex estimator or, instead, a convex approximation based on linearization and/or compressed sensing[90], LQT provides an estimate of the generators $\hat{H}, \hat{G}$ that yield the best match with the observed data, where we will use hats to refer to estimated quantities. We note that this minimization is subject to constraints on Hamiltonian hermicity and Lindblad matrix semidefinite positiveness.

In this work, we describe our first steps in the development of a learning procedure for non-Markovian quantum dynamical maps that supersedes the above LQT. In particular, we consider the statistical inference of the generators of quantum dynamical maps that need not fulfill CP-divisibility, the property of a quantum dynamical map where the evolution between any two intermediate times can be described by a completely positive map[51,105]. These maps $\rho(t) = \mathscr{E}^{\text{TL}}_{t,t_0}(\rho_0)$ no longer have the Lindbladian generator of Eq. (1), but are instead governed by a *time-local master equation* that, when expressed in a canonical form[102], reads $\dot{\rho} = -\text{i}[H(t), \rho] + \mathscr{D}_{\text{TL}}(\rho)$ with

$$\mathscr{D}_{\text{TL}}(\rho) = \sum_n \gamma_n(t)\left(L_n(t)\rho L_n^\dagger(t) - \frac{1}{2}\{L_n^\dagger(t)L_n(t), \rho\}\right). \tag{4}$$

Here, the Hamiltonian $H(t)$, as well as the dissipative rates and jump operators $\gamma_n(t), L_n(t)$, can be time dependent. It is important to note that the 'rates' are no longer required to be positive semidefinite. The possibility of encountering negative rates is directly linked with the non-Markovianity of the quantum evolution[51,52,75]. This time-local master equation can always be written in the form of Eq. (1) by letting $H \mapsto H(t)$ and $G \mapsto G(t)$, such that the corresponding quantum dynamical map will depend on the history of the time-dependent generators $\mathscr{E}^{\text{TL}}_{t,t_0} = \mathscr{E}^{\text{TL}}_{t,t_0}(\{H(t'), G(t')\})$ for all $t' \in [t, t_0]$. Here, the t' argument reflects the fact that now the dynamical map between initial time t_0 and time t depends on all intermediate times. We thus formulate a *Lindblad-like quantum tomography* (LℓQT) by upgrading the LQT cost function in Eq. (3) to a time-local one that can encompass non-Markovian effects

$$\mathsf{C}_{\text{TL}}(\{H(t'), G(t')\}) = -\sum_{s,i,b,m_b} N_{s,i,b,m_b} \log p^{\text{TL}}_{s,i,b}(m_b), \tag{5}$$

where the theoretical probabilities are calculated following

$$p^{\text{TL}}_{s,i,b}(m_b) = \text{Tr}\left\{M_{b,m_b}\mathscr{E}^{\text{TL}}_{t_i,t_0}(\rho_{0,s})\right\}. \tag{6}$$

The above cost function must be minimized subject to dynamical constraints

$$\left(\hat{H}(t), \hat{G}(t)\right) = \text{argmin}\left\{\mathsf{C}_{\text{TL}}(\{H(t'), G(t')\})\right\}, \tag{7}$$

subject to $H(t') \in \mathsf{Herm}(\mathscr{H}_S), G(t') \in \mathsf{Herm}(\mathbb{C}^{d^2-1})$. Therefore, we see that in addition to the time dependences, the dissipation matrix is no longer required to be semi-positive definite, but only Hermitian, and can thus support negative decay rates and incorporate non-Markovian effects.

The crucial property that differentiates LQT from other learning approaches such as quantum process tomography[58–62] is that the estimator includes $|\mathbb{I}_t|$ different instants of time $\{t_i, i \in \mathbb{I}_t\} \subset T$, instead of focusing on a single snapshot of the quantum dynamical map. Although we have shown in ref. 90 that, in certain regimes, an accurate LQT can be obtained by focusing on a single snapshot for Lindbladian evolution, this will not be the case for time-correlated and non-Markovian quantum evolutions. In this case, it will be crucial to include the information of various snapshots into the cost function. In fact, we address in this work how many snapshots would be required, and which particular instants of time would be optimal in order to learn about the memory effects of a time-correlated or a non-Markovian noisy quantum evolution. We note that a black-box approach to LℓQT is a very complicated problem, as the parameters of the Hamiltonian and dissipation matrix can have any arbitrary time dependence. In order to progress further, we instead look into physically-motivated models for LℓQT, allowing us to restrict the search space, and start by focusing on a simple and, yet, very relevant setting: a single qubit subject to time-correlated dephasing noise, which can result in non-Markovian quantum dynamics. Our techniques and conclusions may be useful when generalising to more complicated non-Markovian dynamics, aiming at the characterization of non-Markovian noise in gate sets of QIPs to optimise a tomographic analysis[77].

As stated in Eq. (7) we can follow a frequentist approach similarly to the one in Markovian Lindblad quantum tomography[87,89,90], but now taking into account the time-dependence of the decay rates, which will require an a priori selection of the evolution times at which the system is probed. In light of the LℓQT estimation problem of Eq. (7), we do not need to consider arbitrary time-dependent functions for the Hamiltonian and dissipation matrix, but can actually find an effective parametrization that reduces drastically the search space. This means that the minimization in Eq. (7) can be over the parameters $\boldsymbol{\theta}$ of the Hamiltonian and dissipation matrix. Although several intermediate times can be taken, a single time instant actually suffices for an accurate learning in LQT[90], which makes it simpler and computationally lighter. As the noise becomes time correlated, reconstructing the more complex time-dependent dynamics via LℓQT, demands multiple time steps to achieve high accuracies. LQT has in principle the same number of parameters $d^2(d^2-1)$ to be learnt as quantum process tomography, and to do so requires at least $d^2(d^2-1)$ measurement configurations. For the general case, where arbitrary time-dependent rates need to be estimated, a naive blind-search approach would require repeated LQT at many time points, leading to high sample complexity. In contrast, LℓQT's use of a microscopic parameterization reduces the need for extensive measurements by imposing physical constraints, which partially addresses the sample complexity challenge and makes it a more scalable method. Thus in parameterized LℓQT we will need at least as many measurement configurations as the number of parameters to estimate.

Alternatively to this frequentist approach, we can also consider a Bayesian approach that exploits a physically-motivated prior knowledge about these noise parameters, which can be represented by a certain

probability distribution. After performing measurements on the system at certain instants of time and certain measurement basis, we update this probability distribution with the new acquired information, a process that is repeated until reaching a target accuracy for the parameter estimation. The Bayesian approach has the advantage of choosing, at each step, the most convenient subsequent time and measurement basis at which to measure by maximizing the information one would gain. In principle, this can lead to a reduction in the number of measurements required to reach a certain accuracy with respect to those required by the frequentist approach. In practice, however, the non-Markovianity of the quantum evolution can modify this argument, as our probabilistic account of the model parameters can be affected by the actual time correlations of the underlying random process.

Frequentist approach to non-Markovian inference

In this section, we focus on the frequentist approach, which builds on the relative frequencies of observed outcomes. The LℓQT cost function of Eq. (5) can be rewritten as follows

$$\mathsf{C}_{\mathrm{TL}}(\boldsymbol{\theta}) = -\sum_{s,i,b} N_{s,i,b} \sum_{m_b} \tilde{f}_{s,i,b}(m_b\,|\boldsymbol{\theta}_\star) \log p^{\mathrm{TL}}_{s,i,b}(m_b|\boldsymbol{\theta}), \tag{8}$$

where $\tilde{f}_{s,i,b}(m_b\,|\boldsymbol{\theta}_\star) = N_{s,i,b,m_b}/N_{s,i,b}$ is the ratio of the number of m_b measurement outcomes observed N_{s,i,b,m_b} to the total of $N_{s,i,b}$ shots collected at the instant of time i and initial state s when measuring the system with the POVM element b. Our notation remarks that these relative frequencies carry information about the real noise parameters $\boldsymbol{\theta}_\star$ we aim at estimating. In addition, the estimator depends on $p_{s,i,b}(m_b|\boldsymbol{\theta})$ shown in Eq. (6), where we make explicit the dependence on the parametrized noise.

The minimization in Eq. (7) can explicitely written as a minimization over the noise parameters. As a consequence, the frequentist approach can be recast as a statistical problem of parameter point estimation[104], namely

$$\hat{\boldsymbol{\theta}}_{\mathrm{F}} = \mathsf{argmin}_{\boldsymbol{\theta}}\{\mathsf{C}_{\mathrm{TL}}(\boldsymbol{\theta}) :\ \boldsymbol{\theta} \in \Theta = \mathbb{R}^n\}. \tag{9}$$

Instead of the general LℓQT learning over $d^2(d^2-1)$ parameters, which increase exponentially $d = 2^{n_q}$ with the number of qubits n_q and can arbitrarily change in time, our procedure revolves around the estimation of n noise parameters, which are independent of the system size and the evolution times. On the other hand, the imprecision of our estimates will indeed depend on our choice of the evolution times, forcing us to go beyond the LQT single-time estimator[90], and actually measure at optimal times, initial states and measurement elements for which the estimation imprecision can be minimized. Let us note that the conditions to minimise this cost function are the same as those that minimize the Kullback-Leibler divergence[106,107], which is the following relative entropy

$$\mathsf{D}_{\mathrm{KL}}(\boldsymbol{p}_1||\boldsymbol{p}_2) = -\sum_k p_{1,k} \log\left(\frac{p_{2,k}}{p_{1,k}}\right), \tag{10}$$

between the experimental $p_{1,k} \in \{\tilde{f}_{s,i,b}(m_b\,|\boldsymbol{\theta}_\star)\}$ and parameterized theoretical $p_{2,k} \in \{p^{\mathrm{TL}}_{s,i,b}(m_b\,|\boldsymbol{\theta})\}$ probability distributions, provided one considers variations with respect to the estimation parameters $\boldsymbol{\theta}$.

The above estimator depends on the data set $\mathbb{D}$, and is thus also a stochastic variable, which will be characterized by its mean $\mathbb{E}[\hat{\boldsymbol{\theta}}_{\mathrm{F}}]$ and its moments, such as the covariance matrix

$$[\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{\mathrm{F}})]_{n_1,n_2} = \mathbb{E}\Big[\Big(\hat{\theta}_{n_1} - \mathbb{E}[\hat{\theta}_{n_1}]\Big)\Big(\hat{\theta}_{n_2} - \mathbb{E}[\hat{\theta}_{n_2}]\Big)\Big]. \tag{11}$$

We note that the expectation values are taken with respect to the probability distributions for the measurements of $p_{s,i,b}(m_b)$ which, implicitly, also have the stochastic average over the random dephasing noise in the semi-classical model, or a partial trace over the environment in the quantum-mechanical one. The nice property of the maximum-likelihood estimator is that it is asymptotically unbiased, such that $\mathsf{B}_{\boldsymbol{\theta}_\star}(\boldsymbol{\theta}_{\mathrm{F}}) = \mathbb{E}[\hat{\boldsymbol{\theta}}_{\mathrm{F}}] - \boldsymbol{\theta}_\star \to \boldsymbol{0}$ for a sufficiently large N_{shot}. Moreover, its asymptotic covariance matrix saturates the Cramér-Rao bound[108] relating the estimation precision to the Fisher information matrix, which quantifies the amount of information in $\mathbb{D}$ about the unknown parameters. If we momentarily assume that the measurements occur at a single instant of time $t = t_i$, a single initial state and a single measurement basis with outcomes that are independent and identically distributed, the covariance matrix becomes $\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{\mathrm{F}}) \approx (N_{s,i,b} I_{F,s,i,b}(\boldsymbol{\theta}_\star))^{-1}$, where

$$[I_{F,s,i,b}(\boldsymbol{\theta}_\star)]_{n_1,n_2} = \mathbb{E}\left[\frac{\partial \log\left(p^{\mathrm{TL}}_{s,i,b}(m_b|\boldsymbol{\theta})\right)}{\partial\theta_{n_1}} \frac{\partial \log\left(p^{\mathrm{TL}}_{s,i,b}(m_b|\boldsymbol{\theta})\right)}{\partial\theta_{n_2}}\bigg|_{\boldsymbol{\theta}_\star}\right]. \tag{12}$$

In this work, we deal with the more general case in which we measure at several times, therefore the random variables are not identically distributed, and the total number of shots need not be the same for different measurement configurations, i.e., instants of time, initial states and measurement basis. In this case, we must take a linear combination of the Fisher information matrices of each measurement configuration weighted by the proportion of measurements taken at each one of these configurations[109]. We thus obtain the asymptotic covariance matrix

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{\mathrm{F}}) \approx \Sigma_{\hat{\boldsymbol{\theta}}} \equiv \left[N_{\mathrm{shot}} \sum_{s,i,b} \frac{N_{s,i,b}}{N_{\mathrm{shot}}} I_{F,s,i,b}(\boldsymbol{\theta}_\star)\right]^{-1}, \tag{13}$$

such that, the more the Ramsey estimator varies under changes of the noise parameters, the bigger the amplification of the noise parameter is and, thus, the smaller the imprecision one can achieve. As we can see, the imprecision of the estimate will scale with $1/\sqrt{N_{\mathrm{shot}}}$, such that the Ramsey estimator is asymptotically consistent in the $N_{\mathrm{shot}} \to \infty$ asymptotic limit[104]. The asymptotic statistics of the maximum-likelihood estimator is further explained in the Methods section.

We also note that, in this limit, the observed relative frequencies will be normally distributed, such that one can consider minimizing a weighted least-squares cost function. For instance, assuming a binary case where $m_b \in \{0, 1\}$, corresponding to a qubit measurement, we have

$$\mathsf{C}_{\mathrm{TL}}(\boldsymbol{\theta}) \approx \sum_{s,i,b} \frac{\left(\tilde{f}_{s,i,b}(0|\boldsymbol{\theta}_\star) - p^{\mathrm{TL}}_{s,i,b}(0|\boldsymbol{\theta})\right)^2}{\tilde{\sigma}^2_{\tilde{f}_{s,i,b}}}, \tag{14}$$

where $\tilde{f}_{s,i,b}(0\,|\boldsymbol{\theta}_\star) = N_{s,i,b,0}/N_{s,i,b}$ is the ratio of the number of outcomes observed $N_{s,i,b,0}$ to the total of shots $N_{s,i,b}$ collected at initial state s, the instant of time t_i and measurement element b. Here, $\tilde{\sigma}^2_{\tilde{f}_{s,i,b}}$ is the variance of these measured samples. These $\tilde{f}_{s,i,b}(0\,|\boldsymbol{\theta}_\star)$ follow a binomial proportion distribution with mean given by $p^{\mathrm{TL}}_{s,i,b}(0|\boldsymbol{\theta}_\star)$ and, in the limit of a large $N_{s,i,b}$, they can be approximated by a normal distribution with mean $p^{\mathrm{TL}}_{s,i,b}(0|\boldsymbol{\theta}_\star)$ and variance $\sigma^2_{\tilde{f}_{s,i,b}}$ according to the central limit theorem. The expected variance of the measurement configuration (s, i, b) is

$$\sigma^2_{\tilde{f}_{s,i,b}} = \frac{p^{\mathrm{TL}}_{s,i,b}(0\,|\boldsymbol{\theta})\left(1 - p^{\mathrm{TL}}_{s,i,b}(0\,|\boldsymbol{\theta})\right)}{N_{s,i,b}}, \tag{15}$$

since we are sampling from a binomial distribution. This approximation allows us to use a simpler weighted least-squares algorithm, such as the trust-region reflective algorithm implemented in SciPy, where we can optionally set some bounds for the parameters to be estimated. The computational complexity of the trust-region reflective algorithm is dominated in each iteration by solving the trust-region subproblem, which

is typically $O(n^3)$, with n the number of parameters. The number of iterations depends on the complexity of the cost function, the accuracy of the initial guess, and the desired level of convergence. However, once the algorithm gets close to the solution, the remaining convergence is rapid, showing a quadratic rate, meaning the distance to the solution decreases at a rate proportional to the square of the distance at the previous step.

Once the properties of the Ramsey frequentist estimate $\hat{\boldsymbol{\theta}}_{\mathrm{F}}$ have been discussed, we can search for the optimal measurement times, initial states and measurement elements that would lead to a Ramsey estimator with the lowest possible imprecision for a given finite N_{shot}. Depending on which parameter we are interested in, we may be interested in minimizing a particular component of the covariance matrix $[\Sigma_{\hat{\boldsymbol{\theta}}}]_{n_1,n_1}$ in Eq. (13) or, alternatively, minimize its determinant as a whole. In the asymptotic limit in which $\hat{\boldsymbol{\theta}}_{\mathrm{F}}$ follows a multivariate normal distribution, $\det\Sigma_{\hat{\boldsymbol{\theta}}}$ is proportional to the volume enclosed by the covariance elliptical region, so it is a good measure of the dispersion of the distribution, and a good way to quantify the accuracy of the estimation. Sometimes, we will also use $\det\Sigma_{\theta}^{1/2n}$ instead, which can be more easily compared to the individual standard deviations $[\Sigma_{\hat{\boldsymbol{\theta}}}]^{1/2}_{n_1,n_1}$, as both quantities scale with $1/\sqrt{N_{\mathrm{shot}}}$. Minimizing $\det\Sigma_{\hat{\boldsymbol{\theta}}}$ has also the advantage that the optimal measurement configurations obtained are independent of the parameters we want to determine, assuming the different parametrizations have the same number of parameters and that a coordinate transformation exists between parametrizations. In this case, the determinants of the covariance matrices are related by $\det\Sigma_{\theta} = \det\Sigma_{\theta'}\det J^2$, with J the Jacobian of the coordinate transformation between both parametrizations, which does not depend on the measurement configurations and therefore will have no influence in the minimization. Before presenting these results, we discuss an approach based on Bayesian inference[104].

Bayesian approach to non-Markovian inference

Rather than considering the relative frequencies as approximations of the underlying probability distribution with a certain fixed value of $\boldsymbol{\theta}_\star$, the idea of Bayesian inference is to quantify statistically our knowledge about the noise parameters, and how this knowledge gets updated as we collect more information via measurements. Hence, the noise parameters become continuous stochastic variables themselves $\boldsymbol{\theta}_\star \mapsto \tilde{\boldsymbol{\theta}}$ that take values $\boldsymbol{\theta} \in \Theta$ according to a prior probability density function (PDF) $\pi_0(\boldsymbol{\theta})$. This probability distribution quantifies our uncertainty about the noise parameters before making any measurement $\mathbb{D}_0 = \emptyset$. At each $\ell > 0$ Bayesian step, we measure the system enlarging the data set sequentially $\mathbb{D}_{\ell-1} \mapsto \mathbb{D}_\ell = \mathbb{D}_{\ell-1} \cup \delta\mathbb{D}_\ell$, where $\delta\mathbb{D}_\ell \subset \mathbb{D} = \{N_{i,s,b,m_b}\}$ contains a number of measurement outcomes $|\delta N_\ell|$ that is a fraction of the total N_{shot}. These outcomes will be labeled as $\delta\mathbb{D}_\ell = \{N_{i_\ell,s_\ell,b_\ell,m_{b_\ell}}\}$. The measurements in this data set are again binary Bernoulli trials, and can be described by a joint distribution of independent binomials $p_{\tilde{N}}(\delta\mathbb{D}_\ell)$, defined in analogy to the likelihood function, but only extended to the configurations measured in the particular Bayesian step. The prior $(\ell-1)$-th probability distribution is then updated by using Bayes' rule based on the parametrized probability distributions $p_\ell^{\mathrm{TL}}(\delta\mathbb{D}_\ell|\boldsymbol{\theta})$ being understood as probabilities conditioned on our statistical knowledge of the noise parameters

$$\pi_\ell(\boldsymbol{\theta}|\mathbb{D}_\ell) = \frac{1}{p_{\ell-1}^{\mathrm{TL}}(\delta\mathbb{D}_\ell)} p_\ell^{\mathrm{TL}}(\delta\mathbb{D}_\ell|\boldsymbol{\theta})\pi_{\ell-1}(\boldsymbol{\theta}). \tag{16}$$

Here, $p_{\ell-1}^{\mathrm{TL}}(\delta\mathbb{D}_\ell) = \int_\Theta \mathrm{d}^n\theta\, p_\ell^{\mathrm{TL}}(\delta\mathbb{D}_\ell|\boldsymbol{\theta})\pi_{\ell-1}(\boldsymbol{\theta})$ is a normalization constant required to interpret $\pi_\ell(\boldsymbol{\theta}|\mathbb{D}_\ell) \mapsto \pi_\ell(\boldsymbol{\theta})$ as a probability distribution describing our updated knowledge about the noise, which will be used as the subsequent prior.

One of the main differences with respect to the frequentist approach is that we have, at each step, a probability distribution from which one can obtain a Bayesian estimate

$$\hat{\boldsymbol{\theta}}_{\mathrm{B}} = \mathbb{E}_{\boldsymbol{\theta}}[\boldsymbol{\theta}] = \int_\Theta \mathrm{d}^n\theta\, \pi_\ell(\boldsymbol{\theta})\boldsymbol{\theta}. \tag{17}$$

We note that this Bayesian estimator $\hat{\boldsymbol{\theta}}_{\mathrm{B}}$ minimizes the Bayesian risk associated to a squared error loss function over all possible estimators $\hat{\tau}$, $\hat{\boldsymbol{\theta}}_{\mathrm{B}} = \mathrm{argmin}\{\mathsf{R}_{\mathrm{TL}}(\hat{\tau})\}$[104] with

$$\mathsf{R}_{\mathrm{TL}}(\hat{\tau}) = \sum_{i_\ell,s_\ell,b_\ell,m_{b_\ell}\in\delta\mathbb{D}_\ell} \int_\Theta \mathrm{d}^n\theta\, \pi_\ell(\boldsymbol{\theta}) p_\ell^{\mathrm{TL}}\big(\delta\mathbb{D}_\ell|\boldsymbol{\theta}\big)(\hat{\tau}-\boldsymbol{\theta})^2. \tag{18}$$

In addition to the expectation value in Eq. (17), since we have the updated probability distributions, we can quantify how our uncertainty about the noise parameters changes via the associated covariance matrix $[\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{\mathrm{B}})]_{n_1,n_2}$ or any other statistics, regardless of the size of the Bayesian data set $\mathbb{D}_\ell$. This differs from our previous arguments for covariance of the maximum likelihood estimate $\hat{\boldsymbol{\theta}}_{\mathrm{F}}$ in Eq. (13), which require working in the asymptotic regime $N_{\mathrm{shot}} \to \infty$. In experimental situations in which this regime cannot be reached, we note that one could use Monte Carlo sampling techniques to estimate precision of $\hat{\boldsymbol{\theta}}_{\mathrm{F}}$[77,110,111], although these deal with the estimator based on the full likelihood function.

Another crucial difference of the Bayesian approach is that, instead of choosing a predefined set of evolution times, initial states and measurement basis, we can find the optimal time at which we should measure to maximize the information gain at each Bayesian step. For each update, we thus solve for

$$(i_\ell, s_\ell, b_\ell) = \mathtt{argmax}_{i,s,b}\big\{\mathbb{E}\big[\mathsf{D}_{\mathrm{KL}}\big(\pi_\ell(\boldsymbol{\theta}|\mathbb{D}_\ell)\,||\,\pi_\ell(\boldsymbol{\theta})\big)\big]\big\}, \tag{19}$$

where $\pi_\ell(\boldsymbol{\theta}|\mathbb{D}_\ell)$ is the posterior probability in Eq. (16) corresponding to the measurement results one would obtain by measuring at a time t_i, initial state s and measurement basis b, and enlarging the data set as $\mathbb{D}_{\ell'} = \mathbb{D}_\ell \cup \delta\mathbb{D}_{\ell'}$. In the expression above, we are making use of the Kullback-Leibler divergence of Eq. (10) between the posterior and the prior, searching for a measurement configuration that maximizes the relative entropy between the prior and any of the possible posteriors, such that one gains the maximum amount of information at each Bayesian step. Therefore, not only the data set is enlarged sequentially $\mathbb{D}_\ell \supset \mathbb{D}_{\ell-1}$, but also the specific times are chosen adaptively. In light of the fixed set of measurement times in the frequentist approach $\{t_i : i \in \mathbb{I}_t\}$, we note that the total set of updated times after N_{B} Bayesian steps $\{t_\ell, \ell \in \{0,\cdots,N_{\mathrm{B}}\}\}$ can be very different, and that is the reason why we use a different notation. Computing Eq. (19) can be quite time-consuming, especially when dealing with a large number of parameters. In practice, long computation times may lead to a reduction in the frequency of experimental shots, which is undesirable. To avoid this, we can take tens or hundreds of shots at each step before computing again the optimal measurement configuration. This will not change the results significantly, since a Bayesian update of a single shot does not change the prior much and the optimal measurement time of next step remains very similar to the previous one. We note that the Bayesian approach is ultimately related to the maximum-likelihood estimation in the asymptotic limit $N_{\mathrm{shot}} \to \infty$. When the variance of the prior is small, the maxima of the Kullback-Leibler divergence in Eq. (10) between prior and posterior are localized at the optimal measurement configurations obtained by minimizing the asymptotic covariance matrix of the maximum likelihood-estimation in Eq. (13). Additionally, if we make a single Bayesian update in which we take a very big number of shots at several measurement configurations, the likelihood function relating the posterior and the prior will be a joint distribution of independent binomials. Given the big number of shots taken in this single step, the posterior distribution will be mainly shaped by the likelihood function, which contains most of the information about the parameters. The position of the maximum of this PDF will be located at the most likely value of the parameters, and therefore it is in

agreement with the maximum-likelihood estimator that takes measurements at those measurement configurations. The Bayesian approach has the advantage that, at each step, we measure at the most convenient configuration that maximizes the expected information gain, and this can lead to an overall reduction of the number of measurements needed. Also, as noted above, the final estimate relies on a probability distribution, so that we can immediately derive confidence intervals without requiring any asymptotic limit.

For the Bayesian approach, where we use sequential Monte Carlo[112] to implement it numerically, the computational complexity is dominated by the computation of the best measurement configuration. At each step in the Bayesian approach, we need to compute the expected posterior of the measurement configurations, and then select the measurement configuration maximizing the Kullback-Leibler divergence between posterior and prior. The complexity of this calculation depends on the number of particles in the particle filter, N_{filter}, the number of initial states, the number of measurement bases, the number of measurement outcomes and the number of measurement times considered. Therefore, we have a complexity $O(N_{\text{filter}} \times |\mathbb{S}_0| \times |\mathbb{M}_b| \times |\mathbb{M}_{m_b}| \times |\mathbb{I}_t|)$. At each one of these $N_{\text{filter}} \times |\mathbb{S}_0| \times |\mathbb{M}_b| \times |\mathbb{M}_{m_b}| \times |\mathbb{I}_t|$ operations we need to evaluate the likelihood function $p_\ell^{\text{TL}}(\delta\mathbb{D}_\ell|\boldsymbol{\theta})$ to compute the posterior weight. Therefore if the posterior function is costly to compute we could think of precomputing it on a grid and then make linear interpolation to reduce the computational cost, as the authors of ref. 113 do. N_{filter} may need to increase exponentially with the number of parameters, which makes the frequentist approach more suitable in the case of large number of parameters. However, this can be moderated by the specific structure of the model and how concentrated the posterior is.

Lindblad-like quantum tomography for non-Markovian dephasing

Let us formulate the LℓQT for the time-local master equation of Eq. (4) for the dephasing of a single qubit, including temporal correlations. As discussed in the Methods section, either in a semi-classical or quantum-mechanical model of pure dephasing, the qubit dynamics can be described by the time-local master equation in Eq. (4) with a simple Hamiltonian $H(t) = \frac{1}{2}\omega_0\sigma_z$ and a single jump operator $L(t) = \sigma_z$, both of which are time independent

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{\mathrm{i}}{2}\left[\omega_0\sigma_z, \rho\right] + \gamma(t)\left(\sigma_z\rho\sigma_z - \rho\right). \tag{20}$$

On the contrary, the decay rate is time-dependent and contains memory information about the noise fluctuations

$$\gamma(t) = \frac{1}{4}\int_0^t \mathrm{d}t'(C(t,t') + C(t',t)). \tag{21}$$

In a semi-classical model, $C(t,t') = \mathbb{E}[\delta\tilde{\omega}(t)\delta\tilde{\omega}(t')]$ is the auto-correlation function of a stochastic process $\delta\tilde{\omega}(t)$ representing frequency fluctuations. This master equation is the result of averaging over the stochastic process $\rho(t) = \mathbb{E}[\tilde{\rho}(t)]$, and is valid for a random process in a second-order cumulant expansion known as the fast-fluctuation expansion or, alternatively, for a Gaussian random process with arbitrary correlation times τ_c, as discussed in the Methods section. Alternatively, for a fully quantum-mechanical dephasing model, $C(t,t') = \frac{1}{2}\mathrm{Tr}_B\{B_I(t)B_I(t')\rho_B^{\text{ss}}\}$ is the auto-correlation function on the stationary state of the environment/bath ρ_B^{ss}, which induces fluctuations in the qubit frequency via the bath operators $B_I(t)$. In this case, the time-local master equation is the result of tracing over the bath degrees of freedom $\rho(t) = \mathrm{Tr}_B\{\rho_{\text{SB}}(t)\}$, and is valid in a second-order cumulant expansion also discussed in the Methods section. Assuming wide-sense stationarity, i.e., the mean of the stochastic process is constant and the auto-correlation function only depends on $t' - t$, we can introduce the power spectral density (PSD) of the noise

$$C(t-t') = \int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\,S(\omega)\,\mathrm{e}^{\mathrm{i}\omega(t-t')}. \tag{22}$$

Therefore, the time-dependent rate becomes

$$\gamma(t) = \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}\omega\,S(\omega)f_\gamma(\omega,t), \tag{23}$$

where we have introduced the following modulation function

$$f_\gamma(\omega,t) = \frac{t}{2\pi}\,\mathrm{sinc}\,(\omega t). \tag{24}$$

It will be useful to define the symmetrized auto-correlation function and the symmetrized PSD as $\bar{C}(t,t') = \frac{1}{2}(C(t,t') + C(t',t))$ and $\bar{S}(\omega) = \frac{1}{2}(S(\omega) + S(-\omega))$, since for dephasing noise only the symmetric part of the auto-correlation function and the symmetric part of the PSD influence the time evolution, as shown in the Methods section.

As discussed in the previous sections, we focus on an effective parametrization of the Hamiltonian and dissipation matrix that reduces the search space. Considering the expression in Eq. (23), the dephasing rate will depend on a certain number n of real-valued noise parameters $\boldsymbol{\theta}_\star \in \Theta = \mathbb{R}^n$ via the PSD $S_{\boldsymbol{\theta}_\star}(\omega)$, where we have made explicit its parametrization. Since the Hamiltonian of the qubit is very simple, and only depends on the transition frequency ω_0 that is known with a high precision using spectroscopic methods, it need not be included in the learning. Technically, this means that we can assume that the driving used to initialize and measure the qubit is resonant with the transition (see Fig. 1), and work in the rotating frame presented in the Methods section. Likewise, we have only one possible jump operator $L = \sigma_z$, such that the learning can focus directly on the estimation $\hat{\boldsymbol{\theta}}$ of the actual noise parameters $\boldsymbol{\theta}_\star$, which translate into the estimation of the dephasing rate $\hat{\gamma}(t)$ via Eq. (23). Therefore, LℓQT becomes a non-linear minimization problem for the cost function in Eq. (5). Before giving more details on this problem, we discuss relevant properties of the dephasing quantum dynamical map.

For a white-noise noise model with a vanishing correlation time $\tau_c = 0$, one has a flat PSD $S(\omega) = c$ and a constant dephasing rate $\gamma(t) = c/2$. This leads to an exponential decay of the coherences $p_{+,i,x}(m_x) = \frac{1}{2}\left(1 + (-1)^{m_x}\mathrm{e}^{-t_i/T_2}\right)$, where $m_x \in \{0, 1\}$ correspond to the projective measurements on $|+\rangle, |-\rangle$, respectively, and we have introduced a decoherence time $T_2 = 2/c$. If the real system is affected by white dephasing noise, there is thus a single noise parameter to learn $\theta_\star = c_\star$ or, alternatively, the real decoherence time $\theta_\star = T_{2\star}$. We note that this procedure is in complete agreement with the LQT based on the corresponding Lindblad master equation in Eq. (2). On the other hand, for a time-correlated dephasing noise with a structured PSD, the coherence decay will generally differ from the above exponential law, with the exception of the long-time regime $t_i \gg \tau_c$, where $p_{+,i,x}(m_x) \approx \frac{1}{2}\left(1 + (-1)^{m_x}\mathrm{e}^{-t_i/T_2}\right)$ and one finds an effective decoherence time controlled by the static part of the PSD $T_2 = 2/S(0)$. As t_i increases towards τ_c, the decay will no longer be a time-homogeneous exponential, which can actually be a consequence of (but not a prerequisite for) a non-Markovian quantum evolution. In this more general situation, we will have more noise parameters $\boldsymbol{\theta}_\star$ to learn.

Let us now connect to the formalism of filter functions[40,41,43,44,114–117], which appears naturally when considering the time evolution of the coherences at any instant of time. This follows from the exact solution of the time-local master equation in the rotating frame, which reads

$$p_{+,i,x}^{\text{TL}}(m_x) = \frac{1}{2}\left(1 + (-1)^{m_x}\mathrm{e}^{-\Gamma(t_i)}\right) =: p_i^{\text{TL}}(m_x), \tag{25}$$

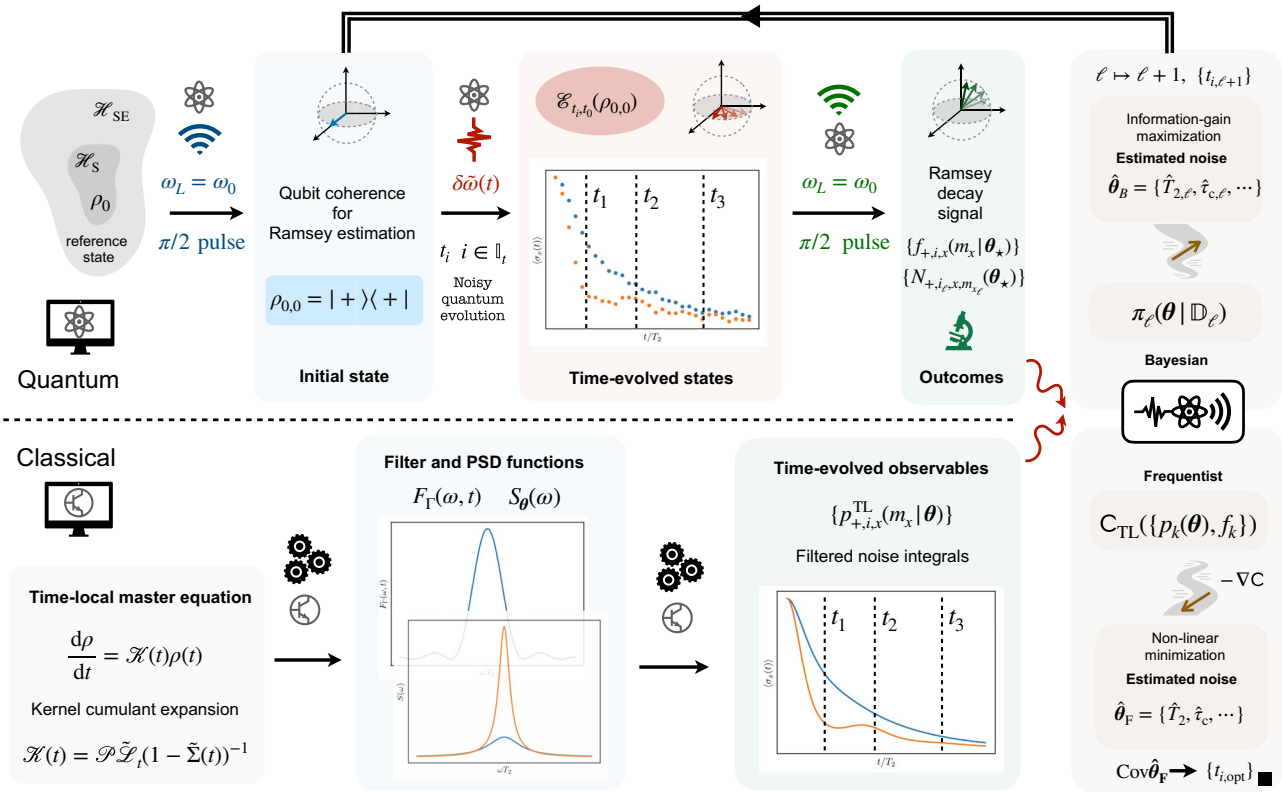

Fig. 1 | Scheme for LℓQT protocol of a dynamical dephasing map. In the case of pure dephasing, the LℓQT scheme reduces to a problem of Ramsey estimation. In the QIP, the qubit is initialized in a single state in the equator of Bloch's sphere $\rho_{0,0} = |+\rangle\langle+|$ by a resonant driving inducing a $\pi/2$ pulse on a reference state $\rho_0 = |0\rangle\langle 0|$. The qubit then evolves for different times $t_i : i \in \mathbb{I}_t$, after which it is subjected to another resonant $\pi/2$ pulse, which effectively rotates to the Pauli measurement basis x. As a consequence of the device dephasing noise with actual noise $\boldsymbol{\theta}_\star$ parameters, the $m_x \in \{0, 1\}$ outcomes are collected as relative frequencies $\{f_{+,i,x}(m_x|\boldsymbol{\theta}_\star)\}$ which will decay in time. This data can be collected by distributing the measurement shots among pre-fixed evolution times, connecting to the frequentist estimation protocol. Alternatively, in a Bayesian approach, one collects the data by successive ℓ-update steps $\delta\mathbb{D}_\ell = \{N_{+,i_\ell,x,m_{x,\ell}}(\boldsymbol{\theta}_\star)\}$. In the lower half, we show the classical part of the estimation, which starts by solving the time-local master equation for a particular parametrized power spectral density $S_{\boldsymbol{\theta}}(\omega)$ and filter function. This leads to the theoretically predicted probabilities $\{p^{\mathrm{TL}}_{+,i}(m_x|\boldsymbol{\theta})\}$. These probabilities and the relative frequencies are either fed into a log-likelihood cost function $\mathscr{C}_{\mathrm{TL}}(\boldsymbol{\theta})$ that must be minimized in order to obtain the frequentist estimate $\boldsymbol{\theta}_{\mathrm{F}}$, or used in a Bayesian step to update our prior knowledge of the noise parameters $\pi_\ell(\boldsymbol{\theta}|\mathbb{D}_\ell)$, converging to the final Bayesian estimate $\boldsymbol{\theta}_{\mathrm{B}}$. In the frequentist approach, by calculating the covariance matrix, we can estimate the precision and find the optimal measurement times, which will depend on how many noise parameters we aim at estimating, and changes with their specific values. In the Bayesian approach, knowledge about the noise parameters is updated sequentially by selecting evolution times that are expected to yield the greatest information gain at each step, thereby implementing feedback and adaptive data acquisition from the quantum device.

where we have introduced the time integral of the decay rates

$$\Gamma(t) = 2\int_0^t \mathrm{d}t'\gamma(t'), \tag{26}$$

and simplified the notation by omitting the initial state and the measurement basis, as they will be unique for the estimation of the dephasing map. Using the Fourier transform in Eq. (22), this integral can be rewritten in terms of the noise PSD as

$$\Gamma(t) = \int_{-\infty}^{\infty} \mathrm{d}\omega\, S(\omega)F_\Gamma(\omega,t),\;\; F_\Gamma(\omega,t) = \int_0^t \mathrm{d}t' f_\gamma(\omega,t'), \tag{27}$$

where we have introduced a filter function that reads

$$F_\Gamma(\omega,t) = \frac{t}{2}\eta_{\frac{2}{t}}(\omega). \tag{28}$$

Note that, by making use of the nascent Dirac delta

$$\eta_\epsilon(x) = \frac{\epsilon}{\pi x^2}\sin^2\left(\frac{x}{\epsilon}\right), \tag{29}$$

where $\eta_\epsilon(x) \to \delta(x)$ as $\epsilon \to 0^+$, and $\int_{-\infty}^{\infty}\mathrm{d}x\,\eta_\epsilon(x) = 1$, one sees that in the long-time limit $\epsilon = 2/t \to 0^+$, $\tilde{f}_\Gamma(\omega,t) \approx \frac{t}{2}\delta(\omega)$ becomes a Dirac delta distribution, such that $\Gamma(t) \approx tS(0)/2$. This agrees with the above coarse-grained prediction for a decoherence time $T_2 = 2/S(0)$. Therefore, physically, the conditions for the long-time limit to be accurate is that $t \gg \tau_c$.

In this article, we are not interested in this Markovian limit, as we aim at estimating the time-local master equation that depends on the full decay rate $\gamma(t)$, including situations in which non-Markovianity becomes manifest. To quantify this, we note that the dephasing quantum dynamical map

$$\mathscr{E}^{\mathrm{TL}}_{t,0}(\rho_0) = (1-p(t))\rho_0 + p(t)\sigma_z\rho_0\sigma_z, \tag{30}$$

is non-Markovian when it is not CP-divisible. In Eq. (30), we have introduced the following time-dependent probability for the occurrence of phase-flip errors

$$p(t) = \frac{1}{2}\left(1 - \mathrm{e}^{-\Gamma(t)}\right). \tag{31}$$

Following[105], the degree of non-Markovianity of the quantum evolution can be obtained by integrating over all times for which the rate of the time-local

master equation is negative

$$\mathcal{N}_{\rm CP} = \int \mathrm{d}t \big(|\gamma(t)| - \gamma(t)\big). \tag{32}$$

An alternatively measure of non-Markovianity is based on the trace distance of two arbitrary initial states[52,118], which will decrease with time when there is a flow of information from the system into the noisy environment. When this information flows back, the trace distance increases, and the qubit can recohere for a finite lapse of time, such that one gets a non-Markovian quantum evolution. The instantaneous variation of trace distance is given by $\sigma(t) = \frac{\mathrm{d}}{\mathrm{d}t} D(\mathscr{E}^{\rm TL}_{t,0}(\rho_0), \mathscr{E}^{\rm TL}_{t,0}(\rho_0'))$, with D being the trace distance[1], and a positive $\sigma(t)$ is thus a measure of non-Markovianity, which can be expressed in terms of the time intervals in which the phase-flip error probability decreases infinitesimally with time

$$\mathcal{N}_{\rm TD} = \max_{\rho_0,\rho_0'} \int_{\sigma(t)>0} \mathrm{d}t\, \sigma(t) = -\int_{\gamma(t)<0} \mathrm{d}t\, \dot{p}(t). \tag{33}$$

Here, we have rewritten this measure in terms of the error probability of the phase-flip channel of Eq. (31), which changes infinitesimally with $\dot{p}(t) = \gamma(t)\mathrm{e}^{-\Gamma(t)}$. Hence, the non-Markovianity condition translates into a dynamical situation in which phase-flip errors do not increase monotonically during the whole evolution. When the dephasing rate attains negative values, the phase-flip error probability can decrease, such that the qubit momentarily recoheres (see the orange lines in Fig. 1). In this simple case of pure dephasing noise, we see that both measures of non-Markovianity in Eqs. (32) and (33) depend on the rate $\gamma(t)$ attaining negative values, and both are equal to 0 if $\gamma(t)$ is always positive. We define and detect non-Markovianity in this case using Eqs. (32) and (33). However, it is important to note that these measures provide sufficient, but not necessary, conditions for detecting all possible forms of non-Markovian behavior. For instance, certain non-Markovian dynamics may not yield an increase in trace distance.

Once these additional properties of the dephasing quantum dynamical map have been discussed, we can move back to the estimation LℓQT protocol of Eq. (7), and how it can be simplified even further. As noted above, in contrast to the informationally-complete set of initial states an measurements that must be considered for the general cost function of LℓQT in Eq. (5), we can work with a smaller number of configurations by noting that all information of the dephasing map can be extracted by preparing a single initial state $\rho_0 = |+\rangle\langle+|$, and measuring in a single basis $M_{x,\pm}$ (see Fig. 1). Indeed, this combination of initial state and measurement basis yields the expected value $p^{\rm TL}_{+,i,x}(m_x) = \frac{1}{2}(1 + (-1)^{m_x}\mathrm{e}^{-\Gamma(t_i)})$. Similarly, if we chose $\rho_0 = |+\mathrm{i}\rangle\langle+\mathrm{i}|$ and $M_{y,\pm}$, we would obtain $p^{\rm TL}_{+\mathrm{i},i,y}(m_y) = \frac{1}{2}(1 + (-1)^{m_y}\mathrm{e}^{-\Gamma(t_i)})$, which yields the same information. The rest of combinations of initial states and measurement bases yield constant values and do not provide any information about the decay rates.

As advanced in the introduction, we would like to know how many snapshots $|\mathbb{I}_t|$ are required, at which the system is measured after evolving for $\{t_i : i \in \mathbb{I}_t\}$ and, moreover, which are the optimal times of those snapshots in terms of the specific details of the noise. We take here two different routes: the frequentist and the Bayesian approach. In the frequentist case, for pure dephasing we have the cost function

$$\mathsf{C}^{\rm pd}_{\rm TL}(\boldsymbol{\theta}) = -\sum_i N_i \sum_{m_x} \tilde{f}_i(m_x\,|\boldsymbol{\theta}_\star) \log p^{\rm TL}_i(m_x|\boldsymbol{\theta}), \tag{34}$$

which is greatly simplified with respect to the general case in Eq. (8), as we only have a single initial state and a single measurement basis. Since we are monitoring the coherence of the qubit, this cost function corresponds to a *Ramsey-type estimator*, where $\tilde{f}_i(0\,|\boldsymbol{\theta}_\star) = N_{i,0}/N_i$ ($\tilde{f}_i(1\,|\boldsymbol{\theta}_\star) = N_{i,1}/N_i$) is the ratio of the number of outcomes observed $N_{i,0}$ ($N_{i,1} = N_i - N_{i,0}$) to the total of N_i shots collected at the instant of time t_i, when measuring the system with the POVM element $M_{x,0}$ ($M_{x,1}$). Our notation remarks that these relative frequencies carry information about the real noise parameters $\boldsymbol{\theta}_\star$ we aim at estimating. In addition, the estimator depends on $p_i(m_x|\boldsymbol{\theta})$ shown in Eqs. (25)–(26), which stand for the probabilities obtained by solving the time-local dephasing master equation of Eq. (20), where we make explicit the dependence on the parametrized noise. Minimizing $\det \Sigma_{\boldsymbol{\theta}}$ we can determine the optimal measurement times of the estimator. For the Bayesian approach, at each $\ell > 0$ Bayesian step, we measure the system enlarging the data set sequentially $\mathbb{D}_{\ell-1} \mapsto \mathbb{D}_\ell = \mathbb{D}_{\ell-1} \cup \delta\mathbb{D}_\ell$, where $\delta\mathbb{D}_\ell \subset \mathbb{D} = \{N_{i,+,x,m_x}\}$ contains a number of measurement outcomes $|\delta N_\ell|$ that is a fraction of the total $N_{\rm shot}$. These outcomes will be labelled as $\delta\mathbb{D}_\ell = \{N_{i_\ell,+,x,m_{x_\ell}}\}$. At each step, we maximize the information gain of Eq. (19) to determine the optimal measurement time. As shown in Fig. 1 the scheme for LℓQT in the case of pure dephasing can be divided into the following steps:

a) Quantum experiment:

1. Initialize the reference state $\rho_0 = |0\rangle\langle 0|$ (see Fig. 1).
2. Prepare $\rho_{0,0} = |+\rangle\langle+|$ by a resonant $\pi/2$ pulse.
3. Let the qubit evolve under the pure dephasing for different evolution times t_i.
4. Apply a second resonant $\pi/2$ pulse, rotating the qubit to the Pauli-X measurement basis.
5. Collect the outcomes $m_x \in \{0, 1\}$ and compute the frequencies $\{f_{+,i,m_x}\}$.

b) Classical processing and frequentist approach:

1. Solve the time-local master equation with a parametrized power spectral density $S_{\boldsymbol{\theta}}(\omega)$ and a filter function, yielding the predicted probabilities $\{p_{+,i,m_x}(\boldsymbol{\theta})\}$.
2. From the quantum experiment we have a set of frequencies $\{f_{+,i,m_x}\}$ at different times t_i. The predicted probabilities and measured relative frequencies are used in a log-likelihood cost function, which is minimized to estimate noise parameters $\boldsymbol{\theta}$.
3. By calculating the covariance matrix, the precision of the estimation is determined.

c) Classical processing and Bayesian approach:

1. Solve the time-local master equation with a parametrized power spectral density $S_{\boldsymbol{\theta}}(\omega)$ and a filter function, yielding the predicted probabilities $\{p_{+,i,m_x}(\boldsymbol{\theta})\}$.
2. We have some prior knowledge of the parameters $\boldsymbol{\theta}$, which is represented by a prior probability distribution.
3. At each step, evolution time t_i is chosen to maximize the information gain of the parameters. A new quantum experiment is performed with this new evolution time.
4. After each measurement or set of measurements the relative frequencies are used to update the prior distribution of the parameters.

We present below a detailed comparison of the two approaches, frequentist and Bayesian, determining the regimes in which each of them is better than the other. For the Bayesian protocol design, we have used the Python package Qinfer[119], which numerically implements the operations needed by using a sequential Monte Carlo algorithm for the updates.

Let us study some dephasing dynamics in which we can apply the two approaches we have just introduced, and make a comparative study of their performance when learning parametrized dephasing maps with time-correlated noise. We study two different dephasing models: a Markovian semi-classical dephasing and a non-Markovian quantum dephasing. These models are selected for several reasons. The semi-classical model is a well-established generalization of purely Lindbladian evolution, and in fact, contains the Lindbladian regime as a limiting case. This allows us to connect with previous results from LQT and Ramsey interferometry in the context of quantum sensing and quantum clocks. On the other hand, the non-Markovian case represents the simplest possible generalization of the Markovian scenario by introducing a shifted Lorentzian PSD, making it an ideal testbed for LℓQT. Furthermore, it has experimental relevance, as we can use a laser-cooled vibrational mode in trapped ions to obtain a fully-

tunable implementation of this non-Markovian noise. These two dephasing models offer both classical and quantum noise characterization through the symmetry of the PSD, and they provide a versatile platform to demonstrate LℓQT.

Regarding the choice between frequentist and Bayesian inference, the majority of quantum technologies traditionally use frequentist approaches. However, Bayesian inference is becoming increasingly relevant, particularly in situations where the number of shots is limited and clock cycle times are large, such as in trapped-ion platforms. As we show below, Bayesian methods offer advantages in scenarios where physical priors are available and can provide better estimates in non-Markovian cases, where the dynamics are more complex.

For each one of the cases presented below, we have a PSD with some real parameters $\boldsymbol{\theta}_\star$ from which we can determine the time-dependent measurement probabilities $p_i^{\rm TL}(m_x|\boldsymbol{\theta}_\star)$. From these probabilities, we numerically take the necessary samples to simulate the experiment by using the SciPy implementation of a binomial random variable[120].

Markovian semi-classical dephasing

We now apply both estimation techniques for the LℓQT of a dephasing quantum dynamical map that goes beyond the Markovian Lindblad assumptions. In particular, we consider a time-correlated frequency noise $\delta\tilde{\omega}(t)$ that is described by an Ornstein-Uhlenbeck (OU) random process[121,122]. This process has an underlying multi-variate Gaussian joint PDF, and incorporates a correlation time $\tau_{\rm c}>0$ above which the correlations between consecutive values of the process become very small. In fact, beyond the relaxation window $t,t'>\tau_{\rm c}$, the correlations show an exponential decay

$$C(t-t')=\frac{c\tau_{\rm c}}{2}\,{\rm e}^{-\frac{|t-t'|}{\tau_{\rm c}}}, \tag{35}$$

where $c>0$ is a so-called diffusion constant. Since this correlations only depend on the time differences, the process is wide-sense stationary. Moreover, on the basis of its Gaussian joint PDF, it can be shown that the process is indeed strictly stationary. Being Gaussian, all the information is thus contained in its two-point functions or, alternatively, in its PSD

$$S(\omega)=\frac{c\tau_{\rm c}^2}{1+(\omega\tau_{\rm c})^2}, \tag{36}$$

which has a Lorentzian shape. This Gaussian process then leads to an exact time-local master equation for the dephasing of the qubit in Eq. (20) with a time-dependent decay rate

$$\gamma(t)=\frac{1}{4}S(0)\Big(1-{\rm e}^{-\frac{t}{\tau_{\rm c}}}\Big). \tag{37}$$

In the long-time limit $t\gg\tau_{\rm c}$, one recovers a constant decay rate $\gamma(t)\approx S(0)/4=c\tau_{\rm c}^2/4$, which connects to our previous discussion of the effective exponential decay of the Ramsey signal $p_i^{\rm TL}(m_x)\approx\big(1+(-1)^{m_x}{\rm e}^{-t_i/T_2}\big)/2$ and the decoherence time $T_2=2/c\tau_{\rm c}^2$. On the other hand, for shorter time scales, we see the effects of the noise memory through a time-inhomogeneous evolution of the coherences that goes beyond a Lindbladian description. The time-dependent decay rate is always positive, such that the two measures of non-Markovianity in Eqs. (32)-(33) vanish exactly. The pure dephasing quantum dynamical map of a qubit subjected to OU frequency noise is thus Markovian albeit not Lindbladian.

From the perspective of LℓQT, we have two parameters to learn $\boldsymbol{\theta}=(c,\tau_{\rm c})\in\Theta=\mathbb{R}_+^2$, which fully parametrize the PSD, the decay rate or, alternatively, the Ramsey attenuation factor

$$\Gamma(t)=\frac{1}{2}S(0)\Big(t-\tau_{\rm c}\Big(1-{\rm e}^{-\frac{t}{\tau_{\rm c}}}\Big)\Big). \tag{38}$$

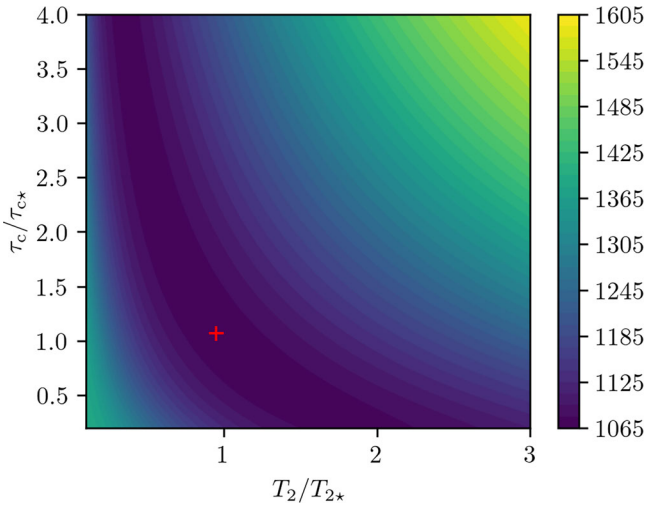

Fig. 2 | Gradient descent for OU dephasing LℓQT. Contour plot of the Ramsey cost function $\mathsf{C}_{\rm TL}^{\rm pd}(\boldsymbol{\theta})$ of Eq. (34) as a function of the noise parameters $(T_2,\tau_{\rm c})$, where we recall that $T_2=2/c\tau_{\rm c}^2$ plays the role of an effective decoherence time in the long-time limit. We choose two evolution times and, for illustration purposes, fix them at $t_1=\tau_{\rm c\star}$, $t_2=2\tau_{\rm c\star}$, fixing the OU diffusion constant such that $T_{2\star}/\tau_{\rm c\star}=1$. We distribute $N_{\rm shot}=2\times10^3$ shots equally per time step. The gradient descent of this convex problem converges towards the global minimum, and is marked with a red cross at $T_2/T_{2\star}\approx0.95$ and $\tau_{\rm c}/\tau_{\rm c\star}\approx1.07$, lying close to the real noise parameters $\boldsymbol{\theta}\mapsto\boldsymbol{\theta}_\star$.

a. Frequentist Ramsey estimators. We can now evaluate the LℓQT cost function $\mathsf{C}_{\rm TL}^{\rm pd}(\boldsymbol{\theta})$ in Eq. (34) by substituting the attenuation factor in Eq. (38) in the likelihood function $p_i^{\rm TL}(m_x|\boldsymbol{\theta})=\big(1+(-1)^{m_x}{\rm e}^{-\Gamma(t_i)}\big)/2$ after a certain set of evolution times $\{t_i,i\in\mathbb{I}_t\}$, and the relative frequencies for the measurement outcomes $\tilde{f}_i(m_x|\boldsymbol{\theta}_\star)$. For a Lindbladian dynamics, a single measuring time t_1 would suffice for the estimation[90], which can actually be solved for analytically in the present pure dephasing context, as discussed in Sec. I of the Supplementary Material. For the OU dephasing, this is no longer the case, and we actually need at least two times t_1,t_2. In order to assess the performance of the frequentist minimization problem in Eq. (9) under shot noise, we numerically generate the relative frequencies $\tilde{f}_1(m_x|\boldsymbol{\theta}_\star),\tilde{f}_2(m_x|\boldsymbol{\theta}_\star)$ at two instants of time by sampling the probability distribution with the actual OU parameters $\boldsymbol{\theta}_\star=(c_\star,\tau_{\rm c\star})$ a number of times $N_{\rm shot}=N_1+N_2$. In the following, rather than learning $(c_\star,\tau_{\rm c\star})$, we will focus on two noise parameters with units of time $(T_{2\star},\tau_{\rm c\star})$, where we recall that $T_{2\star}=2/c_\star\tau_{\rm c\star}^2$ is an effective decoherence time in the long-time limit. In Fig. 2, we present a contour plot of this two-time cost function $\mathsf{C}_{\rm TL}^{\rm pd}(\boldsymbol{\theta})$, which is actually convex and allows for a neat visualization of its global minimum. We also depict with a red cross the result of a gradient-descent minimization, where one can see that the estimates $\hat{\boldsymbol{\theta}}_{\rm F}=(\hat{T}_2,\hat{\tau}_{\rm c})$ are close to the real noise parameters. The imprecision of the estimate is a result of the shot noise, which we now quantify.

In order to find the optimal evolution times t_1, t_2 that maximize the precision of our estimates, we can minimize the covariance in Eq. (13) which, in turn, requires maximizing the Fisher information matrix in Eq. (12). By Taylor expanding the cost function, we can actually find a linear relation between the estimate difference $\delta\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}_{\rm F}-\boldsymbol{\theta}_\star$, and the differences between the parametrized probabilities and the relative frequencies $\delta\tilde{f}_i=p_i^{\rm TL}(m_x|\hat{\boldsymbol{\theta}})-\tilde{f}_i(m_x|\boldsymbol{\theta}_\star)$, namely

$$\begin{pmatrix}\delta\hat{T}_2\\ \delta\hat{\tau}_{\rm c}\end{pmatrix}=\frac{1}{\mathscr{N}}\begin{pmatrix}-\frac{\Gamma'_{\tau_{\rm c}}(t_2)({\rm e}^{2\Gamma(t_1)}-1)}{\sinh\Gamma(t_1)} & \frac{\Gamma'_{\tau_{\rm c}}(t_1)({\rm e}^{2\Gamma(t_2)}-1)}{\sinh\Gamma(t_2)}\\ \frac{\Gamma'_{T_2}(t_2)({\rm e}^{2\Gamma(t_1)}-1)}{\sinh\Gamma(t_1)} & -\frac{\Gamma'_{T_2}(t_1)({\rm e}^{2\Gamma(t_2)}-1)}{\sinh\Gamma(t_2)}\end{pmatrix}\begin{pmatrix}\delta\tilde{f}_1\\ \delta\tilde{f}_2\end{pmatrix}, \tag{39}$$

where $\mathscr{N}=\Gamma'_{T_2}(t_1)\Gamma'_{\tau_{\rm c}}(t_2)-\Gamma'_{\tau_{\rm c}}(t_1)\Gamma'_{T_2}(t_2)$, and we have introduced a shorthand notation for the partial derivatives $\Gamma'_\tau(t_i)=\partial\Gamma(t_i)/\partial\tau_{\rm c}$,

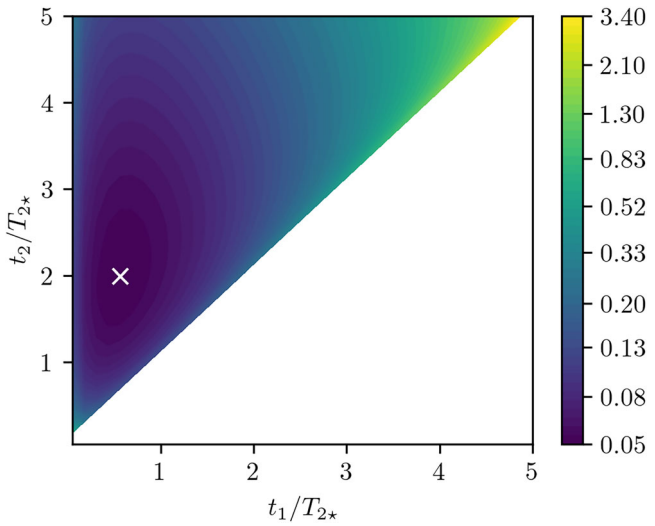

Fig. 3 | Asymptotic covariance matrix for OU dephasing. We represent $\det\left(\Sigma_{\hat{\boldsymbol{\theta}}}(t_1,t_2)\right)^{1/2}$ as a function of the t_1 and t_2 measurement times selected. For $t_1 = t_2$ the determinant diverges, since it is not possible to determine the parameters by just measuring at a single time. The true parameters are $\tau_{c\star} = T_{2\star}/2$ and the number of shots is $N_1 = N_2 = 5 \cdot 10^3$. The optimal times that minimize the determinant are indicated with a white cross, $t_1 \approx 0.56T_{2\star}$, $t_2 \approx 1.99T_{2\star}$. The determinant increases rapidly when moving away from this minimum. Note the color bar scale is logarithmic. The plot is symmetric with respect to the line $t_1 = t_2$, therefore we just represent the part $t_2 > t_1$.

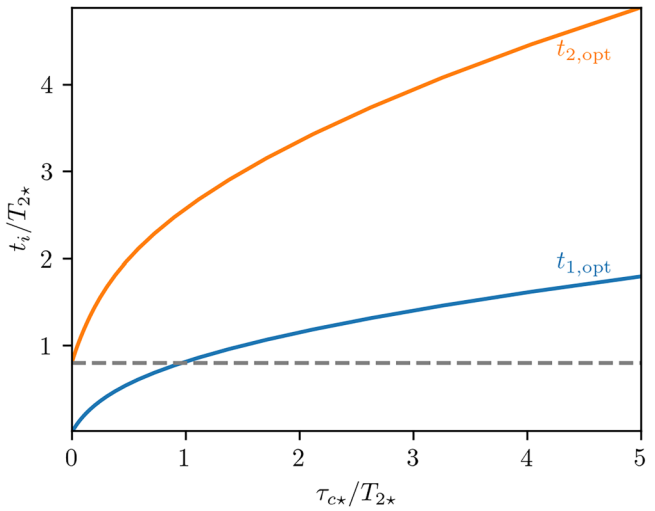

Fig. 4 | Optimal times for OU dephasing LℓQT. We represent $(t_{1,\mathrm{opt}}, t_{2,\mathrm{opt}}) = \mathrm{argmin}\{\det(\Sigma(t_1,t_2))\}$ as a function of the ratio of the real noise parameters $\tau_{c\star}/T_{2\star}$. In the regime where $\tau_{c\star} \ll T_{2\star}$, we find that $t_2 \approx 0.8T_{2\star}$ and $t_1 \approx 0$. The optimal time of this measurement lies very close to the Lindbladian result $t = 0.797T_{2\star}$ (gray dashed line) found when there is no time correlation and we have a single parameter T_c, which is an exact analytical result discussed in Sec. I of the Supplementary Material.

$\Gamma'_{T_2}(t_i) = \partial\Gamma(t_i)/\partial T_2$. The first thing one notices is that, fixing $t_2 = t_1$, the factor $\mathcal{N} = 0$, and the difference between the estimation and the true value of the parameters diverges, signaling the fact that one cannot learn two noise parameters using a Ramsey estimator with a single instant of time. The second result one finds is that, in the asymptotic limit $N_i \to \infty$, the estimate differences will follow a bi-variate normal distribution. This follows from the fact that Eq. (39) is a linear combination of the differences between the finite frequencies and the binomial probabilities, which are known to follow a normal distribution $N\left(\boldsymbol{0}, \mathrm{diag}\left(\sigma^2_{f_1}, \sigma^2_{f_2}\right)\right)$ with binomial variances $\sigma^2_{f_i}$ defined in Eq. (15). Therefore, the frequentist estimates will also be normally distributed $\delta\hat{\boldsymbol{\theta}} \sim N\left(\boldsymbol{0}, \Sigma_{\hat{\boldsymbol{\theta}}}\right)$ according to

$$\Sigma_{\hat{\boldsymbol{\theta}}}(t_1,t_2) = M_{\hat{\boldsymbol{\theta}}}(t_1,t_2)\begin{pmatrix}\sigma^2_{f_1} & 0\\ 0 & \sigma^2_{f_2}\end{pmatrix}M^{\mathrm{T}}_{\hat{\boldsymbol{\theta}}}(t_1,t_2), \tag{40}$$

where $M_{\hat{\boldsymbol{\theta}}}(t_1,t_2)$ is the matrix in Eq. (39). A more detailed derivation of the relationship between shot noise and the uncertainty in estimation, as well as the asymptotic covariance matrix, is provided in the Methods section. From this perspective, the aforementioned divergence for $t_1 = t_2$ is a consequence of the singular nature of this matrix, which cannot be thus inverted.

A measure of the imprecision of the estimation is then obtained from $\det(\Sigma_{\hat{\boldsymbol{\theta}}}(t_1,t_2)) = (\det M_{\hat{\boldsymbol{\theta}}})^2\sigma^2_{f_1}\sigma^2_{f_2} \propto 1/N_1N_2 = 1/N_1(N_{\mathrm{shot}} - N_1)$ which, in this bi-variate case, can be related to the area enclosed by a covariance ellipse. We thus clearly see that the maximum precision will be obtained when $N_1 = N_2 = N_{\mathrm{shot}}/2$. Turning to the optimal measurement times, we can now numerically minimize the determinant of the asymptotic covariance matrix

$$\{t_{i,\mathrm{opt}}\} = \mathrm{argmin}\left\{\det\Sigma_{\hat{\boldsymbol{\theta}}}\left(\{t_i\}\right)\right\}, \tag{41}$$

finding the two optimal values at which the signal shows the highest sensitivity to changes in the OU noise (see Fig. 3). The optimal times obtained in this way are depicted in Fig. 4 as a function of the noise correlation time. In the limit where this correlation time is much smaller than the effective decoherence time $\tau_{c\star} \ll T_{2\star}$, the signal only carries important information about the noise for times that are much larger than the correlation time. Hence, we are in the long-time limit where the decay rate is constant $\gamma(t) \approx 1/2T_2$ and one expects to find agreement with a purely Lindbladian dephasing noise. As discussed in Sec. I of the Supplementary Material, the LQT for pure dephasing requires a single measuring time, and can be analytically found by minimizing the standard deviation of the estimated noise parameter. This solution yields an optimal time $t_{\mathrm{opt}} = 0.797T_{2\star}$, which is actually very close to the intercept of the curve of $t_{2,\mathrm{opt}}$ shown in Fig. 4. In this long-time regime, we find $t_{1,\mathrm{opt}} \approx 0$ indicating that measurements at time $t_{2,\mathrm{opt}}$ will be mainly used to determine parameter T_2, while those at $t_{1,\mathrm{opt}} \approx 0$ contribute to determine the much smaller τ_c.

In the more general case in which the time correlation of the OU noise yields important memory effects, the measurement times have to be adapted to specific optimal values, which are in general larger than the purely Lindbladian limit as shown in Fig. 4. Since these optimal times depend on the parameters we aim at learning, it is not straightforward to devise a practical strategy to minimise the imprecision of the frequentist estimates. In the pure Lindbladian case, one may foresee that the experimentalist will have an accurate prior knowledge of the T_2 time, such that the measurements can all be implemented close to the predicted optimal time. On the other hand, for the OU noise, one has the additional noise correlation time τ_c, which is related to deviations from the time-homogeneous exponential decay of the coherences and is not typically characterised experimentally. The frequentist procedure to operate at the optimal regime of estimation would then need to distribute the total N_{shot} in smaller groups that are applied in sequence, each time shifting the measurement times to try to get to the optimal point. One can foresee that this procedure will not be optimal, as one will loose many measurements along the way and, moreover, not scalable to other situations in which one aims at learning more noise parameters also optimally.

b. Bayesian Ramsey estimators. Let us now describe how a Bayesian inference for OU dephasing LℓQT would proceed, which will provide an experimental procedure to operate at the optimal estimation times. We start by commenting on the fully-uncorrelated Lindbladian limit discussed in Sec. I of the Supplementary Material, where the optimization of the

measurement time for each Bayesian step in Eq. (19) can also be solved analytically. Considering that the prior probability distribution $\pi_\ell(\gamma)$ for our knowledge about the decay rate at the ℓ-th step is Gaussian, we can focus on how its mean and variance change as one takes the next Bayesian step. In Sec. I of the Supplementary Material, we show that, minimizing the Bayesian variance of the next step, one finds optimal measurement times that agree with the above frequentist prediction, albeit for the knowledge of the decay rate that we actually have at each particular step $t_\ell = 0.797/2\hat{\gamma}_\ell$ or, alternatively, of the decoherence time $T_{2,\ell}$. This result is very encouraging, as the experimentalist may only have a crude guess of this value, but it gets automatically updated towards the optimal regime. This motivates an extension to time-correlated dephasing such as the OU noise.

For the OU dephasing, we have two parameters to learn, and we can maximize the Kullback-Leibler divergence of Eq. (19) to obtain the subsequent optimal time t_ℓ for the next Ramsey measurement(s), and the corresponding extension of the data set $\mathbb{D}_{\ell-1} \mapsto \delta\mathbb{D}_\ell = \mathbb{D}_{\ell-1} \cup \delta\mathbb{D}_{\ell-1}$. We then proceed by measuring at this time, updating the prior, and starting the optimization step all over again to finally find the estimates in Eq. (17) $\hat{\boldsymbol{\theta}}_{\rm B} = (\hat{\tau}_{{\rm c},\ell}, \hat{T}_{2,\ell})$. As shown in Fig. 5, as one collects more and more data, the Bayesian measurement times cluster at two single times, and tend to alternate between them. Remarkably, these times are the optimal $t_{1,\rm opt}$ and $t_{2,\rm opt}$ predictions of the frequentist approach shown in Fig. 4. We can see how the Bayesian approach automatically finds the optimal measurement setting to learn a time-correlated dephasing noise.

Let us now present a detailed comparison of the precision of the frequentist and Bayesian approaches. For the frequentist approach, we can obtain the expected covariance of the estimator $\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{\rm F})$ by performing several runs, and computing the covariance matrix of the results. We emphasise that this is not the asymptotic $\Sigma_{\hat{\boldsymbol{\theta}}}$ discussed previously, and does not require a very large number of measurement shots. For the Bayesian approach we obtain a posterior probability distribution after ℓ steps, $\pi_\ell(\boldsymbol{\theta})$, and we can directly compute the covariance matrix Σ_ℓ of this posterior distribution. A good measure of the uncertainty of each one of the approaches can be obtained by taking the determinant of the corresponding covariance matrix. For a Gaussian distribution, this quantity $\det\Sigma$ gives us the elliptical area associated to the bi-variate Gaussian covariance, and we can define an average radius $\bar{R} = \sqrt{\det\Sigma}/\pi$. In order for this to scale as $1/\sqrt{N_{\rm shot}}$ and that it has the same units as the standard deviation, we will use the square root of this radius $\det\Sigma^{1/4}$. Therefore, we will compare $\det\Sigma^{1/4}$ for both frequentist and Bayesian approaches by taking the ratio of the determinants,

$$r_{\rm OU} = \sqrt{[4]}{\frac{\det\left(\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{F,\rm opt})\right)}{\det\Sigma_{\ell_{\rm max}}}}, \tag{42}$$

with $\hat{\boldsymbol{\theta}}_{F,\rm opt}$ the frequentist estimator taking measurements at optimal times and considering for both approaches the same number of total measurements $N_{\rm shot}$. This ratio is represented in Fig. 6 as a function of $\tau_{{\rm c},\star}/T_{2\star}$. As we can see, the Bayesian approach is better for small number of measurements, since it has some prior knowledge of the parameters. As we make more measurements and the frequentist estimator keeps measuring at optimal times we get to the opposite situation. Finally, in the limit of big $N_{\rm shot}$ the Bayesian approach takes also most measurements at these optimal times and the ratio saturates to $r_{\rm OU} \approx 1$, indicating that both approaches offer a similar precision. Let us emphasize, however, that the frequentist approach will not operate in practice at the optimal times, as these depend on the noise parameters one aims at estimating. It is also interesting to note that, as the correlation time of the noise increases, the region where the Bayesian strategy overcomes the frequentist one grows also in terms of the required number of shots. In regimes in which the time correlations are much larger than the effective decoherence time, the Bayesian approach will always be preferred unless one can perform a prohibitively-large number of measurements.

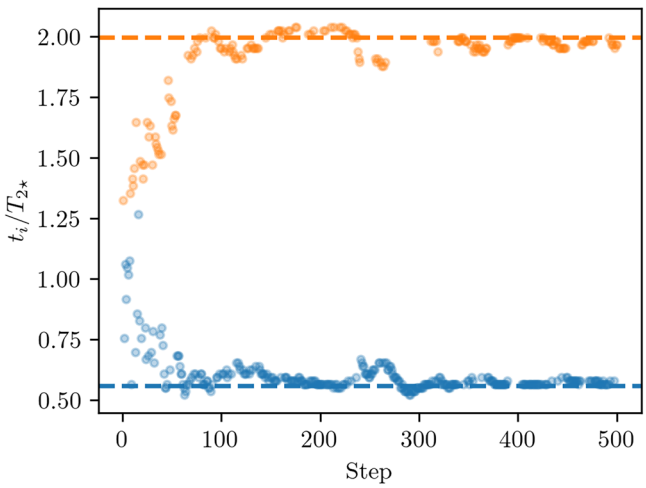

Fig. 5 | Bayesian optimal measurement time at each step. We consider OU noise and set $\tau_{{\rm c}\star} = T_{2\star}/2$. The prior of the parameters is a continuous uniform distribution with $T_2 \in [T_{2\star}/3, 3T_{2\star}]$, $\tau \in [\tau_{{\rm c}\star}/3, 3\tau_{{\rm c}\star}]$. At each Bayesian step, we increase the data set with $|\delta\mathbb{D}_\ell| = 50$ new outcomes, which are used to compute the Bayesian update. The total number of measurements after 500 steps is $500 \times 50 = 25000$. In the Bayesian steps at the beginning, we see that the evolution times lie around $t \sim T_{2\star}$. Later on, the algorithm tends to alternate between two times, which correspond to the optimal times $t_{1,\rm opt} \approx 0.56T_{2\star}$, $t_{2,\rm opt} \approx 1.99T_{2\star}$ in the case of $\tau_{{\rm c}\star} = T_{2\star}/2$. As the number of steps increases and and the posterior gets closer to the true values of $\tau_{{\rm c}\star}$ and $T_{2\star}$, the Bayesian algorithm tends to select measurement times which are closer to the optimal times.

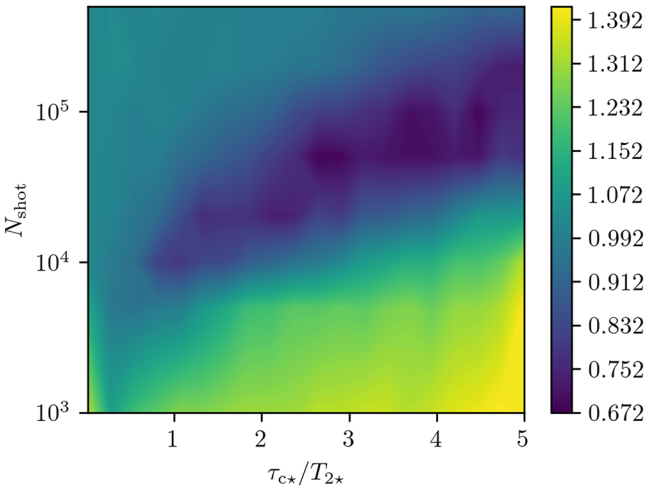

Fig. 6 | Comparison of frequentist and Bayesian approaches for OU noise. The ratio of Eq. (42) computed for different parameter values $\tau_{{\rm c}\star}$ and total number of shots $N_{\rm shot}$. 5000 frequentist runs were done to estimate $\det\mathrm{Cov}(\hat{\boldsymbol{\theta}}_{F,\rm opt})$, while 200 Bayesian runs were done to estimate $\det\Sigma_{\ell_{\rm max}}$ for each point $(\tau_{{\rm c},\star}, N_{\rm shot})$. In the Bayesian approach 100 shots were taken at each step and the prior is a continuous uniform distribution with $T_2 \in [T_{2\star}/3, 3T_{2\star}]$, $\tau \in [\tau_{{\rm c}\star}/3, 3\tau_{{\rm c}\star}]$. Analogously, for the least-squares minimization algorithm used in the frequentist approach (trust-region reflective algorithm), we set the same parameter bounds as the ones of the uniform distribution.

Non-Markovian quantum dephasing

Let us now move on to the discussion of LℓQT for a non-Markovian dephasing dynamics. In the previous section, we have shown that a semi-classical dephasing with OU noise, an archetype for time-correlated Gaussian random processes, yields a dephasing map that, although departing from the time-homogeneous Lindbladian case, does not fall under the class of non-Markovian quantum dynamical maps. We have shown how

both the frequentist and Bayesian approaches can learn the time-local master equation, which is parametrized in terms of an effective decoherence time T_2 and a correlation time τ_c. In this section, we focus on a quantum-mechanical dephasing noise that can actually lead to non-Markovianity in the qubit evolution, and see how the degree of non-Markovianity affects the precision of both the frequentist and LℓQT. We consider an apparently mild modification of the noise PSD with respect to the OU case in Eq. (36). In particular, we use

$$S(\omega) = 4g_{\bar{n}}^2 \frac{\kappa}{(\omega + \Delta_c)^2 + (\kappa/2)^2}, \tag{43}$$

which is a Lorentzian of width κ centered around $-\Delta_c$, and reaching a maximum of $16g_{\bar{n}}^2/\kappa$. We note that for $\Delta_c = 0$, we recover the previous OU case (36) with $\tau_c = 2/\kappa$ and $c = 4g_{\bar{n}}^2\kappa$. On the other hand, for $\Delta_c \neq 0$, this PSD is not an even function $S(\omega) \neq S(-\omega)$, and the associated frequency noise cannot arise from a semi-classical stochastic model[123]. Instead, this particular PSD can be deduced from a quantum-mechanical dephasing model as discussed in the Methods section, and applied to a qubit coupled to a dissipative bosonic mode.In the context of superconducting circuits[124], Δ_c is the detuning of a bosonic microwave resonator with respect to the frequency of an external driving, which is considered to be resonant with the qubit, such that $\Delta_c = \omega_c - \omega_0$. In addition, g is a qubit-resonator cross-Kerr coupling that leads to a bosonic enhancement $g_{\bar{n}}^2 = g^2\bar{n}$, where $\bar{n}$ is the average bosonic occupation of the driven resonator, and κ is the rate of spontaneous emission/loss of photons into the electromagnetic environment. We note that a similar dynamics can be engineered in a two-ion crystal, in analogy to[125], such that one of the ions encodes the qubit in a pair of ground state/metastable levels, while the other one is continuously Doppler cooled via a laser that is red-detuned with respect to a dipole-allowed transition. This laser then drives the carrier and motional sidebands, and effectively laser cools the common vibrational modes, one of which will play the role of the above dissipative bosonic mode, such that the above ω_c will now be its vibrational frequency. The role of the above κ is then played by the laser cooling rate, and the phonon population in the steady state $\bar{n}$ depends on the difference of laser cooling and heating processes[126], which can be controlled by the Rabi frequency and detuning of the laser that drives the dipole-allowed transition. The dissipative phonons will then act as an effective Lorentzian bath for the qubits[125,127,128]. We consider the qubit to be subjected to a far-detuned sideband coupling, which induces a second-order cross-Kerr coupling of strength g describing a phonon-dependent ac-Stark shift on the qubit levels.

In any of the two architectures discussed, when the coupling between the bosonic mode and the qubit is weaker than the dissipative rate $g_{\bar{n}} \ll \kappa$, one can truncate the cumulant expansion of a time-convolutionless master equation of the qubit at second order such that, after tracing over the driven-dissipative mode in its stationary state, one arrives at a time-local dephasing master equation of the form given in Eq. (20). This master equation will be controlled by an auto-correlation function for the bath operator $B(t) = 2g(a^\dagger a - \bar{n})$, following the notation used below Eq. (21) and in the Methods section. In particular, making use of the quantum regression theorem[129], this auto-correlation can be expressed as

$$C(t - t') = 4g_{\bar{n}}^2 e^{-\frac{\kappa}{2}|t-t'|} e^{-i\Delta_c(t-t')}, \tag{44}$$

which coincides with the OU auto-correlation function in Eq. (35) when $\Delta_c = 0$. Being wide-sense stationary, one can Fourier transform this function as shown in Eq. (22), leading to the displaced Lorentzian PSD in Eq. (43). Following Eq. (23), one can obtain the following time-dependent decay rate

$$\gamma(t) = \frac{1}{4}S(0)\left(1 - e^{-\frac{\kappa}{2}t}\left(\cos\Delta_c t + \frac{2\Delta_c}{\kappa}\sin\Delta_c t\right)\right). \tag{45}$$

In comparison to Eq. (37), this decay rate presents additional oscillatory terms for $\Delta_c \neq 0$ that will play an important role for the non-Markovianity of the quantum dynamical map. The attenuation factor that controls the decay of the coherence is

$$\Gamma(t) = \frac{S(0)}{2}\left(t + \frac{2}{\kappa}\frac{(2\Delta_c/\kappa)^2 - 1}{(2\Delta_c/\kappa)^2 + 1} + \frac{2}{\kappa}e^{-\frac{\kappa}{2}t}\cos\varphi(t)\right), \tag{46}$$

where we have introduced $\varphi(t) = \Delta_c t - 2\arctan\frac{2\Delta_c}{\kappa}$.

Recalling that we can assign a correlation time to this noise by $\tau_c = 2/\kappa$, one would expect to recover a Markovian Lindbladian description for $t \gg \tau_c$. In this limit, the term linear in t in Eq. (46) is the dominant one, which leads to a time-homogeneous exponential decay of the Ramsey probabilities $p_i^{\mathrm{TL}}(m_x|\boldsymbol{\theta}) \approx \left(1 + (-1)^{m_x}e^{-t_i/T_2}\right)/2$ with an associated decoherence time $T_2 = 2/S(0) = ((\kappa/2)^2 + \Delta_c^2)/2g_{\bar{n}}^2\kappa$. As in the OU case, for shorter times, the memory effects will start playing a bigger role in the qubit dynamics, such that the Ramsey decay is no longer a time-homogeneous exponential. Moreover, in this particular case, these memory effects can give rise to a non-Markovianity that can be understood as a backflow of information from the environment into the system. According to Eq. (32) or (33), non-Markovianity occurs when the decay rate takes negative values $\gamma(t) < 0$. This can only happen if the second contribution in Eq. (45) dominates over the first one, which cannot happen if $\Delta_c = 0$. Since this contribution is suppressed by $\mathbf{e}^{-\kappa t/2}$, we will need the frequency of the oscillations Δ_c to be sufficiently large in comparison to κ, such that one can get a non-vanishing degree of non-Markovianity. In Fig. 7, we depict the measure of non-Markovianity of Eq. (32) as a function of Δ_c and $\tau_c = 2/\kappa$. We see that the parameter regime $\Delta_c \gtrsim 1.82\kappa$ (white dashed line) is where $\gamma(t)$ can become negative at some time during the evolution, leading to larger non-Markovianity as both Δ_c and τ_c are further increased.

Let us now discuss the statistical inference for the LℓQT of this non-Markovian dephasing map, and compare the frequentist and Bayesian approaches to the statistical estimation.

a. Frequentist Ramsey estimators. As in the OU case, we need to minimize $\mathsf{C}_{\mathrm{TL}}^{\mathrm{pd}}(\boldsymbol{\theta})$ in Eq. (34), where the likelihood function $p_i^{\mathrm{TL}}(m_x|\boldsymbol{\theta}) = \left(1 + (-1)^{m_x}e^{-\Gamma(t_i)}\right)/2$ now depends on the new attenuation factor in Eq. (46). Since there are three noise parameters $\boldsymbol{\theta}_\star = (g_{\bar{n}\star}^2, \kappa_\star, \Delta_{c\star}) \in \Theta = \mathbb{R}_+^3$, we shall at least need to measure at three different times. To assess the performance of the frequentist approach under shot noise, we numerically generate the relative frequencies $\tilde{f}_1(m_x|\boldsymbol{\theta}_\star), \tilde{f}_2(m_x|\boldsymbol{\theta}_\star), \tilde{f}_3(m_x|\boldsymbol{\theta}_\star)$ at these

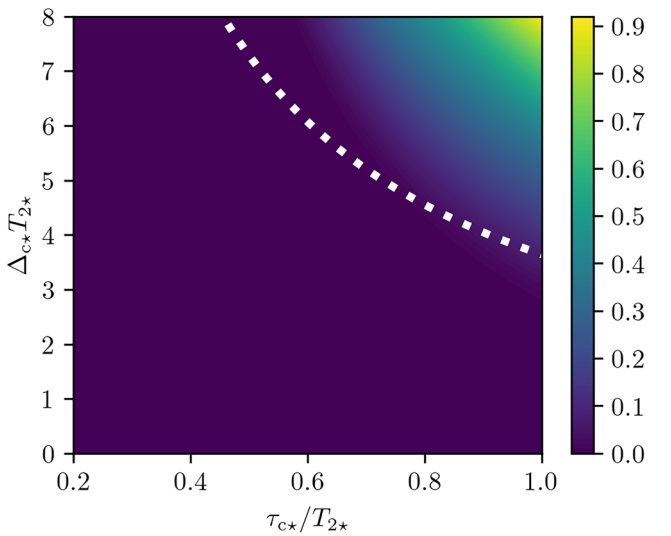

Fig. 7 | Non-Markovianity measure for the quantum dephasing map. Non-Markovianity measure of Eq. (32) for different values of $\tau_c = 2/\kappa$ and Δ_c. The dashed line signals the limit between the Markovian and the non-Markovian regions and corresponds to the equation $\Delta_c \approx 1.82\kappa$. The non-Markovian region is located in the upper right corner.

times by sampling the probability distribution with the real noise parameters $\boldsymbol{\theta}_\star$ a number of times $N_{\text{shot}} = N_1 + N_2 + N_3$. In order to find the three optimal times, we minimize the determinant of the covariance matrix in Eq. (41), which will have an similar expression as Eq. (40), but now expressed in terms of 3×3 matrices for the underlying trivariate normal distribution. Since we are optimizing t_1, t_2 and t_3, one may wonder if we also could improve the estimation by redistributing the total number of measurements differently at each of these times. However, according to our prescription in which the imprecision is quantified by the long-run Gaussian PDF, which defines an elliptical volume in this case, one gets $\det \Sigma_{\hat{\boldsymbol{\theta}}}(t_{1,\text{opt}}, t_{2,\text{opt}}, t_{3,\text{opt}}) \propto 1/N_1N_2N_3$. Therefore, the number of measurements must be equally distributed between the three different times $N_1 = N_2 = N_3 = N_{\text{shot}}/3$.

The minimization in Eq. (41) then yields the optimal times $\{t_{1,\text{opt}}, t_{2,\text{opt}}, t_{3,\text{opt}}\}$ shown in Fig. 8, which have been represented as a function of $\tau_{c\star}/T_{2\star}$ for a fixed value of $\Delta_{c\star}$. The non-zero value of the later is responsible for the fact that, for $\tau_{c\star} \gtrsim 0.73T_{2\star}$, we cross the red-dashed line and enter a non-Markovian regime in which the effective dephasing dynamics is no longer CP-divisible. In this non-Markovian regime, the optimal times tend to be close to the local maxima of $\tilde{f}_i(m_x|\boldsymbol{\theta}_\star)$, providing a large amount of information about the dephasing noise. Conversely, deep in the Markovian regime $\tau_{c\star} \ll T_{2\star}$, it suffices to measure at $t_{3,\text{opt}} \approx 0.797T_{2\star}$ in order to determine $T_{2\star}$, in agreement with the analytical Lindblad result, while $t_{1,\text{opt}}$ and $t_{2,\text{opt}}$ tend to zero and would be used to determine the two other noise parameters.

b. Bayesian Ramsey estimators. Let us now move to the Bayesian approach, where we have some prior knowledge as shown in Eq. (16) of the parameters $\boldsymbol{\theta} = (g_{\vec{n}}^2, \kappa, \Delta_c)$ that gets updated at each Bayesian step t_ℓ by enlarging the data set with $\delta\mathbb{D}_\ell$. Minimizing the relative entropy between the prior and the posterior in Eq. (19), we obtain the subsequent evolution time t_ℓ, and update our knowledge about the noise parameters in the best possible way. As shown in Fig. 9, the Bayesian procedure starts by using evolution update times that are scattered in a broad range of values. However, as the number of iterations increases and we gain more knowledge, they tend to cluster around three well-defined times. In fact, as shown by the corresponding dashed lines, these times coincide with the optimal measurement times of the frequentist approach in Fig. 8. Therefore, if we start a Bayesian experiment with some prior, and we let the experiment run for sufficiently long number of steps, we will learn the optimal times automatically.

In order to compare the precision of the frequentist and Bayesian approaches, we proceed in analogy to the OU noise by looking for a parameter that captures the relative precision of the two approaches in Eq. (42). We now have to consider that the determinant of the trivariate Gaussian covariance $\det \Sigma$ is a volume, and we can define an average radius as $\bar{R} = \sqrt{[3]}\det \Sigma/(4\pi/3)$. We can quantify the precision by taking the square root of this radius, which scales like a standard deviation. Altogether, the relative precision of the frequentist and Bayesian approaches is defined by the ratio

$$r_{\text{NM}} = \sqrt[6]{\frac{\det\left(\text{Cov}(\hat{\boldsymbol{\theta}}_{\text{F,opt}})\right)}{\det \Sigma_{\ell_{\max}}}}. \tag{47}$$

Here, $\hat{\boldsymbol{\theta}}_{F,\text{opt}}$ is the frequentist estimator taking measurements at the optimal times, and we consider the same number of total measurements N_{shot} for both approaches. This ratio is represented in Fig. 10 for $N_{\text{shot}} = 2 \times 10^4$, which is still far from the asymptotic regime of large N_{shot} where one would obtain $r_{\text{NM}} \approx 1$ in similarity to the results found for the OU noise in Fig. 6. Instead of exploring how this ratio changes with the number of measurement shots, we are here interested in understanding how the degree of non-Markovianity can affect the performance of the two estimation strategies. We thus set $N_{\text{shot}} = 2 \times 10^4$, since the variance is already good enough but the cost in terms of number of measurements is still not too big, and plot the precision ratio as a function of the real noise parameters. As we can see in Fig. 10, the blue region represents a regime in which the frequentist approach is slightly better than the Bayesian one, and coincides with the regime of Markovian dephasing that is delimited by the dashed white line. The continuous white line marks the ratio contour line with $r_{\text{NM}} = 1$, and thus delimits the part of the blue region in which the frequentist approach with optimal times is preferable. In the green and yellow areas, which coincide with the non-Markovian regime, the Bayesian approach becomes preferable and the advantage can actually be quite significant. As we go

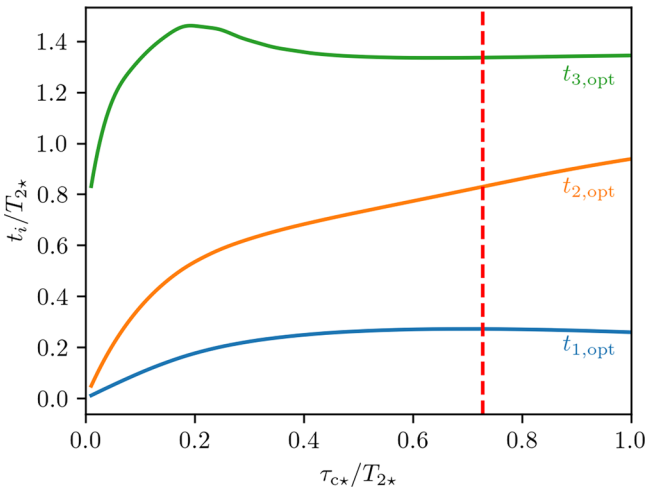

Fig. 8 | Optimal times for quantum dephasing LℓQT. We represent the three optimal times $(t_{1,\text{opt}}, t_{2,\text{opt}}, t_{3,\text{opt}}) = \text{argmin}\{\det(\Sigma_{\hat{\boldsymbol{\theta}}}(t_1, t_2, t_3))\}$ as a function of the ratio of the real noise parameters $\tau_{c\star}/T_{2\star}$, setting $\Delta_{c\star} = 5/T_{2\star}$. The vertical red dashed line separates the Markovian regime (left) from the non-Markovian one (right). On the Markovian regime, as $\tau_{c\star} \ll T_{2\star}$, we see that the t_1, t_2 in blue and orange tend to zero, while t_3 in green tends to the optimal value of the pure Lindbladian case $t_3 \approx 0.797T_{2\star}$. On the other hand, as $\tau_{c\star}$ increases and memory effects become more relevant, the three optimal times start to depart from each other. In the non-Markovian regime, the optimal time $t_{2,\text{opt}}$ is closer to $t_{3,\text{opt}}$ than it is to $t_{1,\text{opt}}$. Moreover, this optimal $t_{1,\text{opt}}$ time starts to decrease as $\tau_{c\star}$ goes deep into the non-Markovian regime.

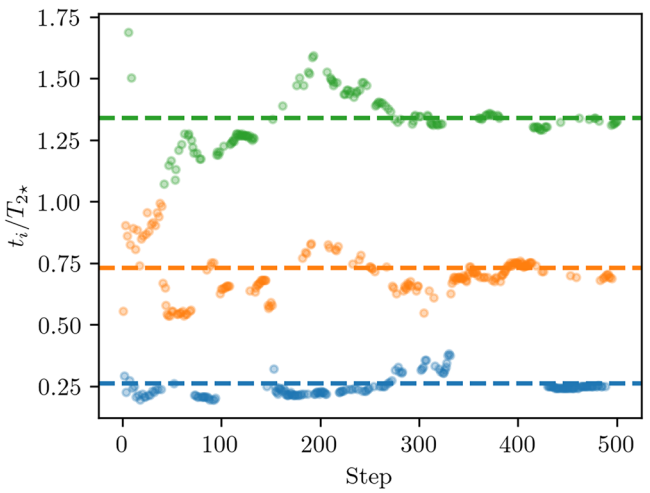

Fig. 9 | Bayesian optimal measurement time at each step. We consider the non-Markovian quantum dephasing with noise parameters $\tau_{c,\star} = T_{2\star}/2$ and $\Delta_{c\star} = 5/T_{2\star}$. We consider as initial prior a continuous uniform distribution with $g_{\vec{n}}^2 \in [g_{\vec{n}\star}^2/3, 3g_{\vec{n}\star}^2]$, $\kappa \in [3\kappa_\star, \kappa_\star/3]$, $\Delta_c \in [3\Delta_{c\star}, \Delta_{c\star}/3]$. At each Bayesian step, we increase the data set with $|\delta\mathbb{D}_\ell| = 50$ new outcomes. As the Bayesian protocol proceeds, we see that the Bayesian update times cluster in three groups of data, which lie around the optimal times $t_{1,\text{opt}}$, $t_{2,\text{opt}}$, $t_{3,\text{opt}}$ of the frequentist approach, which are represented by dashed lines.

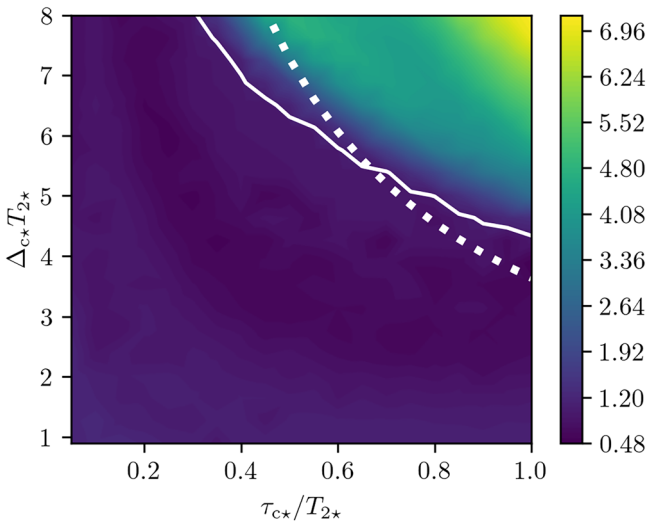

Fig. 10 | Ratio between determinants of covariance matrices of frequentist and Bayesian approaches. The ratio of Eq. (47) computed for different parameter values $\tau_{c\star}$ and $\Delta_{c\star}$, and a fixed number of total measurements $N_{\text{shot}} = 2 \cdot 10^4$. The dashed white line separates the Markovian from the non-Markovian regime, with the non-Markovian region located in the upper right corner. The continuous white line indicates the level where the ratio equals 1. 5000 frequentist runs were done to estimate $\det \text{Cov}(\hat{\boldsymbol{\theta}}_{\text{F,opt}})$, while 30 Bayesian runs were done to estimate $\det \Sigma_{\ell_{\max}}$ for each point $(\tau_{c\star}, \Delta_{c\star})$. In the Bayesian approach 100 shots were taken at each step and the prior is a continuous uniform distribution with $\boldsymbol{\theta} \in [\boldsymbol{\theta}_\star/3, 3\boldsymbol{\theta}_\star]$. Analogously, for the least-squares minimization used in the frequentist approach (trust-region reflective algorithm), we set the same parameter bounds as the ones of the uniform distribution.

deeper in the non-Markovian regime, the oscillating term in Eq. (46) becomes bigger and the decay of the coherence exhibits an increasing number of local maxima. These local maxima represent times that provide a significant amount of information in terms of the Kullback-Leibler divergence of Eq. (19). Therefore, the presence of more local maxima in the non-Markovian regime makes it easier for the Bayesian method to, even if the prior information is minimal, select a time as useful as the asymptotically optimal times.

It is also useful to quantify the advantage of one estimator with respect to the other in terms of number of measurements that is required to reach a target precision, which amounts to reaching the same value of the above covariance determinant. The values shown for the ratio of the determinants in Fig. 10 can be converted into the ratio of number of measurements by assuming that both covariance determinants in Eq. (47) scale as $1/\sqrt{N_{\text{shot}}}$ even in the non-asymptotic regime. Although there can be corrections to this scaling, this can give us in most cases an idea of the proportion of measurement shots one can save by using the best estimator. With this assumption, we get $N_{\text{shot,F}}/N_{\text{shot,B}} \approx r_{\text{NM}}^2$, with r_{NM} the ratio of the determinants of two different approaches. Thus, if for instance $r_{\text{NM}} = 1.5$, we obtain that $N_{\text{shot,F}}/N_{\text{shot,B}} \approx 2.25$. Therefore, we get more than a 2× reduction in the number of measurements with the Bayesian approach in comparison to the frequentist one to reach the same precision. Going back to the values of Fig. 10, we see that the Bayesian approach can result in a considerable improvement as one goes deep in the non-Markovian limit. Before closing this subsection, it is worth recalling once more that, in practice, the frequentist estimation will never be performed at the three optimal times, and the advantage of the Bayesian approach can be even bigger. In Sec. II of the Supplementary Material, we present a detailed comparison of these estimators with another one in which the shots are evenly distributed between the measurements after evolution times that cover uniformly the whole time interval T. As discussed in Sec. II of the Supplementary Material, the Bayesian approach is preferable for most of the parameter values, and can again show a big advantage as one enters into the non-Markovian regime.

Discussion

We have presented LℓQT, a new tool designed to characterize non-Markovian dephasing noise in QIPs, building upon the established framework of Lindblad quantum tomography. LℓQT extends the applicability of Lindblad learning to scenarios where temporal correlations and non-Markovian dynamics play a significant role. In particular, it allows us to extend the characterization of the generators of quantum dynamical maps that go beyond the time-homogeneous Lindblad limit, which connect to a time-local master equation that can display negative decay rates in certain time intervals and, thus, strictly non-Markovian quantum evolutions. Through a detailed comparative study, both frequentist and Bayesian approaches to LℓQT are presented, offering insights into the accuracy and precision of noise estimation under different conditions.

By focusing on the time-correlated dephasing quantum dynamical map of a single qubit, we show that LℓQT can be formally expressed as a parameter estimation process, which simplifies the most general learning scheme to a single initial state and a single measurement basis. In particular, the problem reduces to a time-correlated Ramsey estimator for a parametrized decay rate, which depends on the noise parameters via a filtered power spectral density of the noise. In the frequentist approach, the focus lies on optimizing measurement times to reduce the number of necessary measurements while minimizing error in parameter estimation. By leveraging statistical inference techniques, the frequentist approach provides valuable insights into the efficiency and effectiveness of LℓQT, particularly in scenarios with varying degrees of temporal correlations and non-Markovianity. Conversely, the Bayesian approach offers a more dynamic and adaptive framework, allowing for the incorporation of prior knowledge and iterative updates to refine noise estimates over time.

We have compared the performance of both approaches for two different dephasing quantum dynamical maps, either for a semi-classical or for a quantum-mechanical noise model. In both cases, the microscopically-motivated parametrization allows one to interpolate between a fully Markovian Lindblad limit, for which we derive analytical solutions for the optimal estimation, and a time-correlated and even non-Markovian regime which require a different distribution of the optimal and Bayesian measuring times. Interestingly, in the quantum-mechanical dephasing model, which can be obtained from a microscopic model of a qubit that is coupled to a dissipative bosonic mode in both superconducting-circuit and trapped-ion architectures, the best of the two approaches depends on whether we are in the Markovian or the non-Markovian regime. The Bayesian approach yields much better results in the non-Markovian regime, showing that it is able to automatically adapt to the particularities of the non-Markovian evolution to make much better estimations with a limited number of shots. Moreover, we also compare to more standard schemes considered in the context of Lindbladian quantum tomography, in which the measurements are distributed uniformly (see Supplementary Material). In this case we show an advantage of our schemes that again becomes more appreciable in the non-Markovian regime for the Bayesian approach. However, we also find that for the frequentist approach, the optimal distribution of measurements does not always outperform the uniform distribution when the number of measurments is small and the asymptotic regime has not yet been reached. These results may suggest that the observed patterns extend to other similar cases, but further exploration would be required to confirm this.

The semi-classical dephasing model we have discussed has also been considered in the context of central spin models, where the central spin/qubit is dipolarly-coupled to an ensemble of environmental spins that play the role of a mesoscopic bath. For the case of NV centers in diamond[130], the bath corresponds to nearby substitutional nitrogen atoms, also known as P1 centers, which have a much smaller splitting than the qubit. Due to the mismatch, energy exchange processes can be neglected, and only longitudinal ones survive leading to a time-correlated dephasing. For these systems, assuming incoherent dynamics of different bath spins and negligible back action is a good approximation, such that the above semi-classical model with an Ornstein-Uhlenbeck process turns out to be a reasonably-

good approximation. More generally, if the bath consists of weakly-interacting nuclear spins that get dipolarly-coupled to the qubit, bath correlations might be relevant. Although this is not generally valid, when the leading bath correlations can be reduced to pairs of spins, a similar Gaussian semi-classical model can still be used[131], albeit considering other PSDs. In more general situations, one should consider larger groups of spins, and explore other spin-bath dephasing models. Although this goes beyond the scope of this work, we believe that incorporating a detailed microscopically-motivated parametrization of this type of non-Markovian dephasing should allow for a more efficient parametrization of the noise that could be incorporated in a LℓQT that is similar in spirit to the one considered in this work.

Future research shall explore the extension of these non-Markovian characterization techniques to larger quantum systems, combining the effect of spatial and temporal correlations. More importantly, our work sets the stage to generalize to more complex situations beyond pure dephasing, specially focusing on scalability and robustness, and eventually targeting the noise in full universal gate sets of QIPs.

Methods

Time-local master equation for pure dephasing

For the sake of completeness, we present here a derivation of the time-local master equation in Eq. (20) for a qubit subjected to time-correlated dephasing noise, both in a semi-classical and a fully quantum-mechanical model. This serves to introduce well-known concepts and set our notation following[77].

a. Semi-classical time-correlated dephasing.- The qubit evolves under a stochastic rotating-frame Hamiltonian

$$\tilde{H}(t) = \frac{1}{2}\delta\tilde{\omega}(t)\sigma_z, \tag{48}$$

where $\delta\tilde{\omega}(t)$ is the detuning of the qubit with respect to the frequency of a driving used in the initialization/measurement stages with respect to, and we have set $\hbar = 1$. We use a tilde to highlight the random nature of $\delta\tilde{\omega}(t)$, which is modeled as a stochastic process with zero mean $\mathbb{E}[\delta\tilde{\omega}(t)] = 0$, thus assuming that the driving frequency is resonant with the qubit transition on average. We recall that the averages are taken with respect to the underlying joint PDF of the process for any finite set of times $p_{\delta\tilde{\omega}}(\boldsymbol{\delta\omega}) = p_{t_1,t_2,\cdots,t_n}(\delta\omega_1, \delta\omega_2, \cdots, \delta\omega_n)$, $\forall n \in \mathbb{Z}^+ : \delta\omega_n = \delta\omega(t_n) \in \mathbb{R}$, $\{t_i\}_{i=1}^n \in T$, which fulfills the conditions $p_{\delta\tilde{\omega}}(\boldsymbol{\delta\omega}) \geq 0$ and $\int \prod_n \mathrm{d}\delta\omega_n p_{\delta\tilde{\omega}}(\boldsymbol{\delta\omega}) = 1$[121,132]. Physically, these stochastic fluctuations can either stem from frequency/phase noise of the drive, or from additional external fields that shift the energy of the qubit. For each individual trajectory of the noise $\delta\tilde{\omega}(t)$, the evolution of an initial qubit state $\rho_0 = |\psi_0\rangle\langle\psi_0|$ in the rotating frame is purely unitary but random, giving rise to $\tilde{\rho}(t)$, and expectation values will thus depend on stochastic averages $\mathbb{E}$, leading to a completely-positive trace-preserving (CPTP) map after averaging $\rho(t) = \mathbb{E}[\tilde{\rho}(t)] = \mathscr{E}_{t,t_0}(\rho_0)$[1,64,133,134]. The corresponding stochastic differential equations are

$$\frac{\mathrm{d}\tilde{\rho}}{\mathrm{d}t} = \tilde{\mathscr{L}}_t(\tilde{\rho}(t)), \tilde{\mathscr{L}}_t(\bullet) = -\mathrm{i}[\tilde{H}(t), \bullet], \tag{49}$$

where one sees that the noise $\delta\tilde{\omega}(t)$ thus enters multiplicatively. Using the Nakajima-Zwanzig[74] projection operators $\mathscr{P} = \mathbb{E}$, and $\mathscr{Q} = 1 - \mathscr{P}$, we can find differential equations for the averaged density matrix using

$$\rho(t) = \mathscr{P}(\tilde{\rho}(t)) : \ \mathscr{Q}(\rho_0) = 0 = \mathscr{P}(\tilde{\mathscr{L}}_t). \tag{50}$$

In fact, this averaged evolution can be written as a time-local master equation[99,135,136], namely

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \mathscr{K}(t)\rho(t), \tag{51}$$

where $\mathscr{K}(t)$ is the so-called time-convolutionless kernel that encapsulates the effects that the finite memory of the time-correlated noise has on the qubit. In particular, this kernel can be expressed as follows

$$\mathscr{K}(t) = \mathscr{P}\tilde{\mathscr{L}}_t(1 - \tilde{\Sigma}(t))^{-1}, \tag{52}$$

where we have used a super-operator playing the role of a 'self-energy', which can be expanded as

$$\tilde{\Sigma}(t) = \sum_m \alpha^m \tilde{\Sigma}_m(t), \mathscr{K}(t) = \sum_n \mathscr{P}\tilde{\mathscr{L}}_t \left(\sum_m \alpha^m \tilde{\Sigma}_m(t)\right)^n. \tag{53}$$

In this way, the kernel is organised in a power series of a microscopic coupling α that characterizes the order of magnitude of the coupling of the system to the external noise, and Eq. (50) can be used to show that only even terms contribute

$$\mathscr{K}(t) = \sum_n \mathscr{K}_{2n}(t). \tag{54}$$

This series agrees with the Kubo and Van Kampen cumulant expansion[137,138], and one finds that the n-th order term can be expressed in terms of $n-1$ nested time-ordered integrals[74], being the lowest-order contribution $\mathscr{K}_2(t) = \int_0^t \mathrm{d}t' \mathscr{P}(\tilde{\mathscr{L}}_t \tilde{\mathscr{L}}_{t'})$. This term is controlled by the auto-correlation of the stochastic process, leading to

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{4}\int_0^t \mathrm{d}t' (C(t,t') + C(t',t))\left(\sigma^z\rho(t)\sigma^z - \rho(t)\right), \tag{55}$$

which, for wide-sense stationary processes, can be expressed in terms of the PSD of the stochastic process

$$C(t-t') = \mathbb{E}[\delta\tilde{\omega}(t)\delta\tilde{\omega}(t')] = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} S(\omega)\, \mathrm{e}^{\mathrm{i}\omega(t-t')}. \tag{56}$$

We note that for any wide-sense stationary classical noise, the PSD is even $S(\omega) = S(-\omega)$[123], and $C(t,t') = C(|t-t'|) = C(t',t)$ such that the symmetrized autocorrelation function and the symmetrized PSD introduced below Eq. (24) already contain all of the required information for a second-order approximation. The truncation at this order is justified by first noting that the autocorrelation is typically concentrated within $|t-t'| \leq \tau_c$, where τ_c is a characteristic correlation time. Due to the cluster property[135,139], one finds that $\mathscr{K}_2(t) \sim \alpha\zeta$ with $\alpha = \sqrt{C(0)}$ and a small parameter

$$\zeta = \alpha\tau_c = \sqrt{\frac{\tau_c}{T_2}}, \tag{57}$$

where we have defined a characteristic time as $T_2 = 2/S(0)$. The cluster property for the higher n-th order contributions, which have $(n-1)$ nested integrals, states that the corresponding kernels scale with $\mathscr{K}_n(t) \sim \alpha\zeta^{n-1}$, justifying a low-order truncation whenever the condition $\zeta \ll 1$ is met. This is known as a fast-fluctuation expansion and, back from the rotating frame, yields the time-local master equation in Eq. (20).

Let us note that the above truncation rests on the importance of the memory effects $\tau_c$within the T_2 time. As discussed in more detail in the "Results", this T_2 time controls the time scale for the decay of coherences $\langle\sigma_x(t)\rangle \approx \mathbf{e}^{-t/T_2}$ in a long-time Lindbladian limit $t \gg \tau_c$. However, for shorter times, the structure of the noise can actually lead to deviations from this limit, leading to a coherence decay that is not exponential. As emphasized in the main text, this is not an univocal signal of non-Markovianity for the qubit evolution. We note that there is an exception to the $\tau_c \ll T_2$ requirement for Gaussian random processes, which are defined by a joint PDF that is a multivariate normal distribution for any set of times. In this

case, the time-local master equation in Eq. (20) is actually an exact result, independently of the value of ζ. In fact, the higher-order contributions to the kernel[74] vanish identically $\mathscr{K}_n(t) = 0, \forall n > 2$, due to Isserlis' theorem, most commonly referred to as Wick's theorem in the context of physics $\mathbb{E}[\delta\tilde{\omega}(t_1)\delta\tilde{\omega}(t_2)\cdots\delta\tilde{\omega}(t_n)] = \sum_{\sigma\in S_n}\mathbb{E}[\delta\tilde{\omega}(t_{\sigma(1)})\delta\tilde{\omega}(t_{\sigma(2)})]\cdots\mathbb{E}[\delta\tilde{\omega}(t_{\sigma(n-1)})\delta\tilde{\omega}(t_{\sigma(n)})]$, where S_n is the group of all possible permutations of n elements, e.g. $\sigma(1, 2, \cdots, n-1, n) = (n, 1, 2, \cdots, n-1)$.

b. Quantum-mechanical time-correlated dephasing.- We consider a single qubit coupled to an environment, and evolving $\mathrm{d}\rho_{\mathrm{SB}}/\mathrm{d}t = \mathscr{L}_t(\rho_{\mathrm{SB}})$ under the following Liouvillian

$$\mathscr{L}_t = \mathscr{L}_{\mathrm{SB}} + \mathscr{L}_{\mathrm{B}}, \mathscr{L}_{\mathrm{SB}}(\bullet) = -\frac{\mathrm{i}}{2}\left[\left(\omega_0 + B(t)\right)\sigma^z, \bullet\right], \tag{58}$$

where $B(t)$ is an environment/bath operator that introduces fluctuations on the qubit frequency, and $\mathscr{L}_{\mathrm{B}}(\bullet)$ is the Liouvillian of the bath. In the standard description of quantum master equations, the environment is macroscopically large and subject to a purely-unitary evolution $\mathscr{L}_{\mathrm{B}}(\bullet) = -\mathbf{i}[H_{\mathrm{B}}, \bullet]$. When the system-environment coupling is weak, one can assume that the environment remains unaltered, such that the evolution takes place on the qubit but there is no back action $\rho_{\mathrm{SB}} = \rho(t)\otimes\rho_B$. A Born-Markov approximation then yields a non-unitary master equation for the qubit[129]. This can be expressed as a time convolutionless master equation as the one discussed in the previous subsection in Eq. (51), also truncated at second order, where $\mathscr{P}$ is now a super-operator tracing over the bath degrees of freedom $\mathscr{P}(\rho_{\mathrm{SB}}) = \mathrm{Tr}_B(\rho_{\mathrm{SB}}) = \rho$[74]. The non-unitary evolution of the qubit results from the large number of degrees of freedom in the environment, such that the purity of the state can only decrease with no recurrences.

Let us note, however, that the conditions under which these assumptions are made can be more general, and the degrees of freedom playing the role of an environment need not be macroscopically large. The crucial requirement is that the time with which the effective environment reaches its steady state $\mathscr{L}_{\mathrm{B}}(\rho_{\mathrm{B}}^{\mathrm{ss}}) = 0$ must be much shorter than the timescale of interest in which the system evolves $\rho(t)$. In the present context, this is the case of a single bosonic mode that exchanges energy with a larger electromagnetic bath with a certain rate κ. The bath Liouvillian reads

$$\mathscr{L}_{\mathrm{B}}(\bullet) = -\mathbf{i}[H_{\mathrm{B}}(t), \bullet] + \kappa\left(a\rho a^\dagger - \frac{1}{2}\{a^\dagger a, \bullet\}\right), \tag{59}$$

where $H_{\mathrm{B}}(t)$ is the bosonic mode Hamiltonian, which can include external drivings, and $a^\dagger$, a are the bosonic creation and annihilation operators, respectively. The condition for this single driven-dissipative mode to act as an environment is that κ must be much larger than the coupling strength inside $B(t)$. In the context of the superconducting circuits discussed in the main text, κ is the rate of photon loss in a resonator, and H_{B} must contain a linear resonant microwave driving of the resonator that controls the non-zero number of photons in the steady state[124]. For trapped ions, κ will be the rate of sympathetic cooling of a vibrational model in a two-ion crystal, which will also be supplemented with a smaller heating rate[126]. The difference of these two rates controls the population of phonons in the steady state, and can be controlled by an external laser.

We now move to the interaction picture with respect to the bare system Liouvillian $\rho_I(t) = \mathbf{e}^{t\mathscr{L}_{\mathrm{S}}}(\rho(t))$ with $\mathscr{L}_{\mathrm{S}}(\bullet) = -\mathbf{i}[\frac{1}{2}\omega_0\sigma_z, \bullet]$, and the bare bath Liouvillian, i.e., $B_I(t) = \mathbf{e}^{t\mathscr{L}_{\mathrm{B}}}(B(t))$. The key step is that, due to the fast decay of the bath, for the timescales of interest $t \gg 1/\kappa$, one can assume that $\rho_{\mathrm{SB}}(t) = \rho(t)\otimes\rho_B^{\mathrm{ss}}$, the second-order time-convolutionless master equation can be expressed as in Eq. (55) with

$$C(t, t') = \mathbb{E}[B_I(t)B_I(t')] = \mathrm{Tr}_B\{B_I(t)B_I(t')\rho_{\mathrm{B}}^{\mathrm{ss}}\}. \tag{60}$$

Here, we have assumed that $\mathscr{P}(B_I(t)) = 0$, and we note that $B(t)$ need not commute with itself at different times. Once more, if these quantum-mechanical auto-correlation functions are wide-sense stationary, $C(t, t') = C(t - t')$. We note that, in contrast to the semi-classical case where $S(\omega) = S(-\omega)$, this is not necessarily the case in the quantum-mechanical case $S(\omega) \neq S(-\omega)$[123]. However, in the case of pure dephasing, the time evolution in Eq. (55) only depends on the symmetrized auto-correlation function $\bar{C}(t, t') = \frac{1}{2}(C(t, t') + C(t', t))$ and therefore only the symmetric part of the auto-correlation function influences the time evolution

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{2}\int_0^t \mathrm{d}t'\bar{C}(t, t')\left(\sigma^z\rho(t)\sigma^z - \rho(t)\right). \tag{61}$$

Moving back to the Schrödinger picture, we obtain the master equation in Eq. (20), which will only depend on the symmetrized noise PSD defined below Eq. (24).

Error analysis and asymptotic statistics

We show here the asymptotic normality of the maximum-likelihood estimator in Eq. (9) when N_{shot} tends to infinity and how this is related to the Fisher information in Eq. (13). The asymptotic covariance matrix is used in the main text to obtain the optimal times of the frequentist estimator and can be used to obtain approximate confidence intervals of estimations. To simplify the notation and reduce complexity, we consider the pure dephasing case here, where the cost function is simplified. However, the conclusions remain the same, and the general case can be recovered by substituting $i \to i, s, b$ and $m_x \to m_b$. In this case, instead of considering only measurement points in time, we account for measurement configurations, including evolution times, initial states, measurement bases, and measurement outcomes.

Our estimation $\hat{\boldsymbol{\theta}}_F$ is affected by the limited number of measurement shots N_i at each point in time t_i, which will cause $\tilde{f}_i$ to behave as a normal random variable and will produce a random error in the final estimation of the parameters. The maximum-likelihood cost function reads

$$\mathsf{C}_{\mathrm{TL}}^{\mathrm{pd}}(\boldsymbol{\theta}) = -\sum_{i,m_x} N_i\tilde{f}_{i,m_x}\log p_i(m_x|\boldsymbol{\theta}), \tag{62}$$

where $\tilde{f}_{i,m_x}$ denotes $\tilde{f}_i(m_x|\boldsymbol{\theta}_\star)$ and $p_i(m_x|\boldsymbol{\theta})$ denotes $p_i^{\mathrm{TL}}(m_x|\boldsymbol{\theta})$. The minimum of the cost function satisfies

$$\partial_{\boldsymbol{\theta}}\mathsf{C}_{\mathrm{TL}}^{\mathrm{pd}} = -\sum_{i,m_x} N_i\tilde{f}_{i,m_x}\frac{\partial_{\boldsymbol{\theta}}p_i(m_x|\boldsymbol{\theta})}{p_i(m_x|\boldsymbol{\theta})} = 0. \tag{63}$$

Since $\tilde{f}_{i,m_x} = p_i(m_x|\boldsymbol{\theta}_\star) + \Delta\tilde{f}_{i,m_x}$, with $\Delta\tilde{f}_{i,m_x}$ small, the minimum of the cost function is slightly displaced from the true minimum $\boldsymbol{\theta}_\star$ to $\boldsymbol{\theta}_\star + \Delta\boldsymbol{\theta}$. Taylor expanding Eq. (63) around $\boldsymbol{\theta}_\star$ to first order we have

$$\begin{aligned}\partial_{\theta_j}\mathsf{C}_{\mathrm{TL}}^{\mathrm{pd}} \approx -\sum_{i,m_x} N_i\left[p_i(m_x|\boldsymbol{\theta}) + \Delta\tilde{f}_{i,m_x}\right]\left[\frac{\partial_{\theta_j}p_i(m_x|\boldsymbol{\theta})}{p_i(m_x|\boldsymbol{\theta})}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_\star}\right.\\ \left.+\sum_k\partial_{\theta_k}\frac{\partial_{\theta_j}p_i(m_x|\boldsymbol{\theta})}{p_i(m_x|\boldsymbol{\theta})}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_\star}\Delta\theta_k\right] = 0.\end{aligned} \tag{64}$$

Keeping only first-order terms in $\Delta\boldsymbol{\theta}$ and $\Delta\tilde{f}_{i,m_x}$ and simplifying we obtain

$$\sum_{i,j}\Delta\boldsymbol{\theta}_j N_i[I_i(\boldsymbol{\theta}_\star)]_{jk} = \sum_{i,m_x} N_i\Delta\tilde{f}_{i,m_x}\frac{\partial_{\theta_k}p_i(m_x|\boldsymbol{\theta})}{p_i(m_x|\boldsymbol{\theta})}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_\star}, \tag{65}$$

where we have defined the matrix

$$\begin{aligned}&[I_i(\boldsymbol{\theta}_\star)]_{jk} =\\ &\sum_{m_x} p_i(m_x|\boldsymbol{\theta}_\star)\partial_{\theta_j}\log p_i(m_x|\boldsymbol{\theta})\partial_{\theta_k}\log p_i(m_x|\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_\star}.\end{aligned} \tag{66}$$

Note that this is the Fisher information given in Eq. (12). Taking into account that $\Delta\tilde{f}_{i,1} = -\Delta\tilde{f}_{i,0}$ and $\tilde{f}_{i,1} = 1 - \tilde{f}_{i,0}$, and defining the matrices

$$J_{jk} = \sum_i N_i [I_i(\boldsymbol{\theta}_\star)]_{jk}, \qquad F_{ji} = N_i \frac{\partial_{\theta_j} p_i(0|\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_\star}}{p_i(0|\boldsymbol{\theta}_\star)(1 - p_i(0|\boldsymbol{\theta}_\star))}, \tag{67}$$

we arrive at the expression

$$\Delta\theta_k = \sum_{i,j} J^{-1}_{jk} F_{ji} \Delta\tilde{f}_{i,0}, \tag{68}$$

which relates differences between the expected and measured values $\Delta\tilde{f}_{i,m_x} = \tilde{f}_{i,m_x} - p_i(m_x|\boldsymbol{\theta}_\star)$ to differences between the estimated and the true parameters $\Delta\theta_k$. When N_i is sufficiently large, $\tilde{f}_{i,0}$ behaves as the normal distribution $N\left[\mu = p_i(0|\boldsymbol{\theta}_\star), \sigma^2_{\tilde{f}} = p_i(0|\boldsymbol{\theta}_\star)p_i(1|\boldsymbol{\theta}_\star)/N_i\right]$. Thus, we have

$$\Delta\tilde{f}_{i,0} \sim N\left[\mu = 0, \sigma^2_{\tilde{f}_i} = p_i(0|\boldsymbol{\theta}_\star)p_i(1|\boldsymbol{\theta}_\star)/N_i\right]. \tag{69}$$

After the linear transformation of Eq. (68), we obtain that $\Delta\boldsymbol{\theta}$ behaves as

$$\Delta\boldsymbol{\theta} \sim N\left[\mu = 0,\ \Sigma_{\hat{\boldsymbol{\theta}}} = J^{-1}\right], \tag{70}$$

where we have used that $J^{-1}F\mathrm{diag}(\sigma^2_{\tilde{f}})F^T(J^{-1})^T = J^{-1}$. Thus, the shot noise produces a normal random error with covariance matrix $\Sigma_{\boldsymbol{\theta}} = J^{-1}$ which scales as $1/N_{\mathrm{shot}}$, with $N_{\mathrm{shot}} = \sum_i N_i$, which is the result presented in Eq. (13).

Data availability

The following GitHub repository contains the code to reproduce our findings: https://github.com/varona/llqt.

Code availability

The code used to conduct the analyses and generate the results presented in this work is openly accessible at https://github.com/varona/llqt.

Received: 29 May 2024; Accepted: 11 May 2025;
Published online: 07 June 2025

References

1. Nielsen, M. A. & Chuang, I. L.*Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
2. Bruzewicz, C. D., Chiaverini, J., McConnell, R. & Sage, J. M. Trapped-ion quantum computing: Progress and challenges. *Appl. Phys. Rev.* **6**, 021314 (2019).
3. Morgado, M. & Whitlock, S. Quantum simulation and computing with Rydberg-interacting qubits. *AVS Quantum Sci.* **3**, 023501 (2021).
4. Wendin, G. Quantum information processing with superconducting circuits: a review. *Rep. Prog. Phys.* **80**, 106001 (2017).
5. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
6. Zhong, H.-S. et al. Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
7. Pino, J. M. et al. Demonstration of the trapped-ion quantum CCD computer architecture. *Nature* **592**, 209–213 (2021).
8. Egan, L. et al. Fault-tolerant control of an error-corrected qubit. *Nature* **598**, 281–286 (2021).
9. Postler, L. et al. Demonstration of fault-tolerant universal quantum gate operations. *Nature* **605**, 675–680 (2022).
10. Krinner, S. et al. Realizing repeated quantum error correction in a distance-three surface code. *Nature* **605**, 669–674 (2022).
11. Zhao, Y. et al. Realization of an error-correcting surface code with superconducting qubits. *Phys. Rev. Lett.* **129**, 030501 (2022).
12. Acharya, R. et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature* **614**, 676–681 (2023).
13. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023).
14. Bluvstein, D. et al. Logical quantum processor based on reconfigurable atom arrays. *Nature* **626**, 58–65 (2024).
15. Gupta, R. S. et al. Encoding a magic state with beyond break-even fidelity. *Nature* **625**, 259–263 (2024).
16. Wang, Y. et al. Fault-tolerant one-bit addition with the smallest interesting colour code 2309.09893 (2023).
17. Yamamoto, K., Duffield, S., Kikuchi, Y. & Muñoz Ramo, D. Demonstrating Bayesian quantum phase estimation with quantum error detection. *Physical Review Research* **6** https://doi.org/10.1103/PhysRevResearch.6.013221 (2024).
18. Ryan-Anderson, C. et al. Implementing fault-tolerant entangling gates on the five-qubit code and the color code https://arxiv.org/abs/2208.01863 (2022).
19. Dalzell, A. M. et al. Quantum algorithms: A survey of applications and end-to-end complexities 2310.03011 (2023).
20. Viola, L., Knill, E. & Lloyd, S. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.* **82**, 2417–2421 (1999).
21. Khodjasteh, K. & Lidar, D. A. Fault-tolerant quantum dynamical decoupling. *Phys. Rev. Lett.* **95**, 180501 (2005).
22. Khodjasteh, K. & Viola, L. Dynamically error-corrected gates for universal quantum computation. *Phys. Rev. Lett.* **102**, 080501 (2009).
23. Lidar, D. A.*Review of Decoherence-Free Subspaces, Noiseless Subsystems, and Dynamical Decoupling*, 295–354 (John Wiley & Sons, Ltd, 2014). https://doi.org/10.1002/9781118742631.ch11.
24. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
25. Nation, P. D., Kang, H., Sundaresan, N. & Gambetta, J. M. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum* **2**, 040326 (2021).
26. van den Berg, E., Minev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors. *Nat. Phys.* **19**, 1116–1121 (2023).
27. Cai, Z. et al. Quantum error mitigation. *Rev. Mod. Phys.* **95**, 045005 (2023).
28. Calderbank, A. R. & Shor, P. W. Good quantum error-correcting codes exist. *Phys. Rev. A* **54**, 1098–1105 (1996).
29. Steane, A. M. Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77**, 793–797 (1996).
30. Terhal, B. M. Quantum error correction for quantum memories. *Rev. Mod. Phys.* **87**, 307–346 (2015).
31. Eisert, J. et al. Quantum certification and benchmarking. *Nat. Rev. Phys.* **2**, 382–390 (2020).
32. Kliesch, M. & Roth, I. Theory of quantum system certification. *PRX Quantum* **2**, 010201 (2021).
33. Gebhart, V. et al. Learning quantum systems. *Nat. Rev. Phys.* **5**, 141–156 (2023).
34. Zanardi, P. & Rasetti, M. Noiseless quantum codes. *Phys. Rev. Lett.* **79**, 3306–3309 (1997).
35. Lidar, D. A., Chuang, I. L. & Whaley, K. B. Decoherence-free subspaces for quantum computation. *Phys. Rev. Lett.* **81**, 2594–2597 (1998).
36. Knill, E., Laflamme, R. & Viola, L. Theory of quantum error correction for general noise. *Phys. Rev. Lett.* **84**, 2525–2528 (2000).
37. Kempe, J., Bacon, D., Lidar, D. A. & Whaley, K. B. Theory of decoherence-free fault-tolerant universal quantum computation. *Phys. Rev. A* **63**, 042307 (2001).
38. Lidar, D. A. & Birgitta Whaley, K.*Decoherence-Free Subspaces and Subsystems*, 83-120 (Springer Berlin Heidelberg, 2003). https://doi.org/10.1007/3-540-44874-8_5.

39. Viola, L. & Lloyd, S. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A* **58**, 2733–2744 (1998).
40. Kofman, A. G. & Kurizki, G. Unified theory of dynamically suppressed qubit decoherence in thermal baths. *Phys. Rev. Lett.* **93**, 130406 (2004).
41. Gordon, G., Erez, N. & Kurizki, G. Universal dynamical decoherence control of noisy single- and multi-qubit systems. *J. Phys. B: At., Mol. Optical Phys.* **40**, S75 (2007).
42. Uhrig, G. S. Keeping a quantum bit alive by optimized π-pulse sequences. *Phys. Rev. Lett.* **98**, 100504 (2007).
43. Cywiński, L., Lutchyn, R. M., Nave, C. P. & Das Sarma, S. How to enhance dephasing time in superconducting qubits. *Phys. Rev. B* **77**, 174509 (2008).
44. Biercuk, M. J., Doherty, A. C. & Uys, H. Dynamical decoupling sequence construction as a filter-design problem. *J. Phys. B: At., Mol. Optical Phys.* **44**, 154002 (2011).
45. Clemens, J. P., Siddiqui, S. & Gea-Banacloche, J. Quantum error correction against correlated noise. *Phys. Rev. A* **69**, 062313 (2004).
46. Klesse, R. & Frank, S. Quantum error correction in spatially correlated quantum noise. *Phys. Rev. Lett.* **95**, 230503 (2005).
47. Aharonov, D., Kitaev, A. & Preskill, J. Fault-tolerant quantum computation with long-range correlated noise. *Phys. Rev. Lett.* **96**, 050504 (2006).
48. Terhal, B. M. & Burkard, G. Fault-tolerant quantum computation for local non-Markovian noise. *Phys. Rev. A* **71**, 012336 (2005).
49. Aliferis, P., Gottesman, D. & Preskill, J. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Info Comput.* **6**, 97–165 (2006).
50. Ng, H. K. & Preskill, J. Fault-tolerant quantum computation versus Gaussian noise. *Phys. Rev. A* **79**, 032318 (2009).
51. Ángel, R., Huelga, S. F. & Plenio, M. B. Quantum non-Markovianity: characterization, quantification and detection. *Rep. Prog. Phys.* **77**, 094001 (2014).
52. Breuer, H.-P., Laine, E.-M., Piilo, J. & Vacchini, B. Colloquium: Non-Markovian dynamics in open quantum systems. *Rev. Mod. Phys.* **88**, 021002 (2016).
53. Li, L., Hall, M. J. & Wiseman, H. M. Concepts of quantum non-Markovianity: A hierarchy. *Phys. Rep.* **759**, 1–51 (2018).
54. Addis, C., Ciccarello, F., Cascio, M., Palma, G. M. & Maniscalco, S. Dynamical decoupling efficiency versus quantum non-Markovianity. *N. J. Phys.* **17**, 123004 (2015).
55. Berk, G. D., Milz, S., Pollock, F. A. & Modi, K. Extracting quantum dynamical resources: consumption of non-Markovianity for noise reduction. *npj Quantum Inf.* **9**, 104 (2023).
56. Hakoshima, H., Matsuzaki, Y. & Endo, S. Relationship between costs for quantum error mitigation and non-Markovian measures. *Phys. Rev. A* **103**, 012611 (2021).
57. Donvil, B. I. C., Lechler, R., Ankerhold, J. & Muratore-Ginanneschi, P. Quantum trajectory approach to error mitigation 2305.19874 (2023).
58. Chuang, I. L. & Nielsen, M. A. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.* **44**, 2455–2467 (1997).
59. Poyatos, J. F., Cirac, J. I. & Zoller, P. Complete characterization of a quantum process: The two-bit quantum gate. *Phys. Rev. Lett.* **78**, 390–393 (1997).
60. Fiurášek, J. & Hradil, Zcv Maximum-likelihood estimation of quantum processes. *Phys. Rev. A* **63**, 020101 (2001).
61. Sacchi, M. F. Maximum-likelihood reconstruction of completely positive maps. *Phys. Rev. A* **63**, 054104 (2001).
62. Ježek, M., Fiurášek, J. & Hradil, Zcv Quantum inference of states and processes. *Phys. Rev. A* **68**, 012305 (2003).
63. Nielsen, M. A. & Chuang, I. L.*Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
64. Watrous, J.*The Theory of Quantum Information* (Cambridge University Press, 2018).
65. Childs, A. M., Chuang, I. L. & Leung, D. W. Realization of quantum process tomography in NMR. *Phys. Rev. A* **64**, 012314 (2001).
66. Weinstein, Y. S. et al. Quantum process tomography of the quantum Fourier transform. *J. Chem. Phys.* **121**, 6117–6133 (2004).
67. Mitchell, M. W., Ellenor, C. W., Schneider, S. & Steinberg, A. M. Diagnosis, prescription, and prognosis of a Bell-state filter by quantum process tomography. *Phys. Rev. Lett.* **91**, 120402 (2003).
68. O'Brien, J. L. et al. Quantum process tomography of a controlled-not gate. *Phys. Rev. Lett.* **93**, 080502 (2004).
69. Riebe, M. et al. Process tomography of ion trap quantum gates. *Phys. Rev. Lett.* **97**, 220407 (2006).
70. Home, J. P. et al. Complete methods set for scalable ion trap quantum information processing. *Science* **325**, 1227–1230 (2009).
71. Monz, T. et al. Realization of the quantum Toffoli gate with trapped ions. *Phys. Rev. Lett.* **102**, 040501 (2009).
72. Bialczak, R. C. et al. Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits. *Nat. Phys.* **6**, 409–413 (2010).
73. Poletto, S. et al. Entanglement of two superconducting qubits in a waveguide cavity via monochromatic two-photon excitation. *Phys. Rev. Lett.* **109**, 240505 (2012).
74. Breuer, H. P. & Petruccione, F.*The theory of open quantum systems* (Oxford University Press, Great Clarendon Street, 2002).
75. de Vega, I. & Alonso, D. Dynamics of non-Markovian open quantum systems. *Rev. Mod. Phys.* **89**, 015001 (2017).
76. Chruściński, D. Dynamical maps beyond Markovian regime 2209.14902 (2022).
77. Velázquez, J. M. S. et al. Dynamical quantum maps for single-qubit gates under non-Markovian phase noise 2402.14530 (2024).
78. Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48**, 119–130 (1976).
79. Gorini, V., Kossakowski, A. & Sudarshan, E. C. G. Completely positive dynamical semigroups of n level systems. *J. Math. Phys.* **17**, 821–825 (1976).
80. Wolf, M. M., Eisert, J., Cubitt, T. S. & Cirac, J. I. Assessing non-Markovian quantum dynamics. *Physical Review Letters* **101** https://doi.org/10.1103/PhysRevLett.101.150402 (2008).
81. Boulant, N., Havel, T. F., Pravia, M. A. & Cory, D. G. Robust method for estimating the Lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points. *Phys. Rev. A* **67**, 042322 (2003).
82. Howard, M. et al. Quantum process tomography and Linblad estimation of a solid-state qubit. *N. J. Phys.* **8**, 33 (2006).
83. Onorati, E., Kohler, T. & Cubitt, T. Fitting quantum noise models to tomography data 2103.17243 (2021).
84. Bairey, E., Guo, C., Poletti, D., Lindner, N. H. & Arad, I. Learning the dynamics of open quantum systems from their steady states. *N. J. Phys.* **22**, 032001 (2020).
85. Pastori, L., Olsacher, T., Kokail, C. & Zoller, P. Characterization and verification of Trotterized digital quantum simulation via Hamiltonian and Liouvillian learning. *PRX Quantum* **3**, 030324 (2022).
86. França, D. S., Markovich, L. A., Dobrovitski, V. V., Werner, A. H. & Borregaard, J. Efficient and robust estimation of many-qubit Hamiltonians 2205.09567 (2022).
87. Samach, G. O. et al. Lindblad tomography of a superconducting quantum processor. *Phys. Rev. Appl.* **18**, 064056 (2022).
88. Zhang, H., Pokharel, B., Levenson-Falk, E. & Lidar, D. Predicting non-Markovian superconducting-qubit dynamics from tomographic reconstruction. *Phys. Rev. Appl.* **17**, 054018 (2022).
89. Ben Av, E., Shapira, Y., Akerman, N. & Ozeri, R. Direct reconstruction of the quantum-master-equation dynamics of a trapped-ion qubit. *Phys. Rev. A* **101**, 062305 (2020).
90. Dobrynin, D., Cardarelli, L., Müller, M. & Bermudez, A. Compressed-sensing Lindbladian quantum tomography with trapped ions 2403.07462 (2024).

91. Álvarez, G. A. & Suter, D. Measuring the Spectrum of Colored Noise by Dynamical Decoupling. *Phys. Rev. Lett.* **107**, 230501 (2011).
92. Yuge, T., Sasaki, S. & Hirayama, Y. Measurement of the Noise Spectrum Using a Multiple-Pulse Sequence. *Phys. Rev. Lett.* **107**, 170504 (2011).
93. Chalermpusitarak, T. et al. Frame-based filter-function formalism for quantum characterization and control. *PRX Quantum* **2**, 030315 (2021).
94. Pollock, F. A., Rodríguez-Rosario, C., Frauenheim, T., Paternostro, M. & Modi, K. Non-markovian quantum processes: Complete framework and efficient characterization. *Phys. Rev. A* **97**, 012127 (2018).
95. White, G., Pollock, F., Hollenberg, L., Modi, K. & Hill, C. Non-markovian quantum process tomography. *PRX Quantum* **3**, 020344 (2022).
96. White, G. A. L., Jurcevic, P., Hill, C. D. & Modi, K. Unifying non-Markovian characterisation with an efficient and self-consistent framework 2312.08454 (2023).
97. Wei, K. X., Pritchett, E., Zajac, D. M., McKay, D. C. & Merkel, S. Characterizing non-markovian off-resonant errors in quantum gates. *Phys. Rev. Appl.* **21**, 024018 (2024).
98. Gullans, M., Caranti, M., Mills, A. & Petta, J. Compressed gate characterization for quantum devices with time-correlated noise. *PRX Quantum* **5**, 010306 (2024).
99. Chaturvedi, S. & Shibata, F. Time-convolutionless projection operator formalism for elimination of fast variables. Applications to Brownian motion. *Z. f.ür. Phys. B Condens. Matter* **35**, 297–308 (1979).
100. Nakajima, S. On Quantum Theory of Transport Phenomena: Steady Diffusion. *Prog. Theor. Phys.* **20**, 948–959 (1958).
101. Zwanzig, R. Ensemble Method in the Theory of Irreversibility. *J. Chem. Phys.* **33**, 1338–1341 (2004).
102. Hall, M. J. W., Cresser, J. D., Li, L. & Andersson, E. Canonical form of master equations and characterization of non-Markovianity. *Phys. Rev. A* **89**, 042120 (2014).
103. Bužek, V. Reconstruction of Liouvillian superoperators. *Phys. Rev. A* **58**, 1723–1727 (1998).
104. Rossi, R. J.*Mathematical statistics an introduction to likelihood based inference / by Richard J. Rossi*. (John Wiley & Sons, Hoboken, NJ, 2018), 1st edition. edn.
105. Rivas, A., Huelga, S. F. & Plenio, M. B. Entanglement and non-Markovianity of quantum evolutions. *Phys. Rev. Lett.* **105**, 050403 (2010).
106. Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).
107. Mogilevtsev, D., R^eháček, J. & Hradil, Z. Self-calibration for self-consistent tomography. *N. J. Phys.* **14**, 095001 (2012).
108. James, F.*Statistical Methods in Experimental Physics* (WORLD SCIENTIFIC, 2006), 2nd edn. https://doi.org/10.1142/6096.
109. Hoadley, B. Asymptotic Properties of Maximum Likelihood Estimators for the Independent Not Identically Distributed Case. *Ann. Math. Stat.* **42**, 1977 – 1991 (1971).
110. Blume-Kohout, R. Robust error bars for quantum tomography 1202.5270 (2012).
111. Thinh, L. P., Faist, P., Helsen, J., Elkouss, D. & Wehner, S. Practical and reliable error bars for quantum process tomography. *Phys. Rev. A* **99**, 052311 (2019).
112. Doucet, A., de Freitas, N. & Gordon, N. J. (eds.) *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science (Springer, 2001). https://doi.org/10.1007/978-1-4757-3437-9.
113. Gerster, L. et al. Experimental Bayesian calibration of trapped-ion entangling operations. *PRX Quantum* **3**, 020350 (2022).
114. Kofman, A. G. & Kurizki, G. Acceleration of quantum decay processes by frequent observations. *Nature* **405**, 546–550 (2000).
115. Kofman, A. G. & Kurizki, G. Universal dynamical control of quantum mechanical decay: Modulation of the coupling to the continuum. *Phys. Rev. Lett.* **87**, 270405 (2001).
116. Uhrig, G. S. Exact results on dynamical decoupling by π pulses in quantum information processes. *N. J. Phys.* **10**, 083024 (2008).
117. Almog, I. et al. Direct measurement of the system-environment coupling as a tool for understanding decoherence and dynamical decoupling. *J. Phys. B: At., Mol. Optical Phys.* **44**, 154006 (2011).
118. Rivas, Á., Huelga, S. F. & Plenio, M. B. Quantum non-Markovianity: Characterization, quantification and detection. *Rep. Prog. Phys.* **77**, 094001 (2014).
119. Granade, C. et al. QInfer: Statistical inference software for quantum applications. *Quantum* **1**, 5 (2017).
120. Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).
121. Gardiner, C. W.*Handbook of stochastic methods for physics, chemistry and the natural sciences*, vol. 13 of *Springer Series in Synergetics* (Springer-Verlag, Berlin, 2004), third edn.
122. Gillespie, D. T. The mathematics of Brownian motion and Johnson noise. *Am. J. Phys.* **64**, 225 (1996).
123. Clerk, A. A., Devoret, M. H., Girvin, S. M., Marquardt, F. & Schoelkopf, R. J. Introduction to quantum noise, measurement, and amplification. *Rev. Mod. Phys.* **82**, 1155–1208 (2010).
124. von Lüpke, U. et al. Two-qubit spectroscopy of spatiotemporally correlated quantum noise in superconducting qubits. *PRX Quantum* **1**, 010305 (2020).
125. Bermudez, A., Schaetz, T. & Plenio, M. B. Dissipation-assisted quantum information processing with trapped ions. *Phys. Rev. Lett.* **110**, 110502 (2013).
126. Cirac, J. I., Blatt, R., Zoller, P. & Phillips, W. D. Laser cooling of trapped ions in a standing wave. *Phys. Rev. A* **46**, 2668–2681 (1992).
127. Cormick, C., Bermudez, A., Huelga, S. F. & Plenio, M. B. Dissipative ground-state preparation of a spin chain by a structured environment. *N. J. Phys.* **15**, 073027 (2013).
128. Bermudez, A. & Schaetz, T. Quantum transport of energy in controlled synthetic quantum magnets. *N. J. Phys.* **18**, 083006 (2016).
129. Carmichael, H. J.*An open systems approach to quantum optics* (Springer-Verlag, Berlin Heidelberg, 1993).
130. de Lange, G., Wang, Z. H., Risté, D., Dobrovitski, V. V. & Hanson, R. Universal dynamical decoupling of a single solid-state spin from a spin bath. *Science* **330**, 60–63 (2010).
131. Cywiński, L., Witzel, W. M. & Das Sarma, S. Pure quantum dephasing of a solid-state electron spin qubit in a large nuclear spin bath coupled by long-range hyperfine-mediated interactions. *Phys. Rev. B* **79**, 245314 (2009).
132. Kampen, N. V.*Stochastic processes in physics and chemistry* (North Holland, 2007).
133. Kraus, K. General state changes in quantum theory. *Ann. Phys.* **64**, 311–335 (1971).
134. Choi, M.-D. Completely positive linear maps on complex matrices. *Linear Algebra its Appl.* **10**, 285–290 (1975).
135. Terwiel, R. Projection operator method applied to stochastic linear differential equations. *Physica* **74**, 248–265 (1974).
136. Zoller, P. Theoretical Quantum Optics: A practical guide to stochastic processes & calculus (Master and PhD Course, University of Innsbruck) (2018).
137. Kubo, R. Generalized cumulant expansion method. *J. Phys. Soc. Jpn.* **17**, 1100–1120 (1962).
138. Van Kampen, N. A cumulant expansion for stochastic linear differential equations. i. *Physica* **74**, 215–238 (1974).
139. Fox, R. F. Critique of the generalized cumulant expansion method. *J. Math. Phys.* **17**, 1148–1153 (2008).

Acknowledgements

The project leading to this publication has received funding from the US Army Research Office through Grant No. W911NF-21-1-0007. A.B acknowledges support from PID2021-127726NB- I00 (MCIU/AEI/FEDER,

UE), from the Grant IFT Centro de Excelencia Severo Ochoa CEX2020-001007-S, funded by MCIN/AEI/10.13039/501100011033, from the CSIC Research Platform on Quantum Technologies PTI-001, and from the European Union's Horizon Europe research and innovation programme under grant agreement No 101114305 (-MILLENION-SGA1- EU Project). M.M. furthermore acknowledges support by the European Union's Horizon Europe research and innovation program under Grant Agreement No. 101046968 (BRISQ), the ERC Starting Grant QNets through Grant No. 804247, by the Germany ministry of science and education (BMBF) via the VDI within the project IQuAn, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy -Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1- 390534769. This research is also part of the Munich Quantum Valley (K-8), which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus.

Author contributions

S.V. adapted LℓQT to the frequentist and Bayesian approaches, studying the theoretical characteristics and performing the numerical simulations. A.B. and M.M. conceptualized the initial ideas and theoretical tools, and supervised the theoretical work and the writing of the manuscript. All authors contributed to the review of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-025-01044-7.

Correspondence and requests for materials should be addressed to S. Varona.