

PAPER • OPEN ACCESS

Compressed-sensing Lindbladian quantum tomography with trapped ions

To cite this article: Dmitrii Dobrynin *et al* 2025 *Quantum Sci. Technol.* **10** 045041

View the [article online](#) for updates and enhancements.

You may also like

- [Typical machine learning datasets as low-depth quantum circuits](#)
Florian J Kiwit, Bernhard Jobst, Andre Luckow et al.
- [Adiabatic dynamics of entanglement](#)
Einar Gabbasso and Achim Kempf
- [On the role of symmetry and geometry in global quantum sensing](#)
Julia Boeyens, Jonas Glatthard, Edward Gandar et al.

Modular cryogenics platform adapts to new era of practical quantum computing

ICE

With a cube-based design that fits into a standard rack mount, the **ICE-Q platform** delivers the reliability and scalability needed to exploit quantum systems in real-world operating environments.

[Click to read the article on physicsworld.com](#)



Quantum Science and Technology



PAPER

OPEN ACCESS

RECEIVED
19 February 2025

REVISED
7 August 2025

ACCEPTED FOR PUBLICATION
4 September 2025

PUBLISHED
23 September 2025

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Compressed-sensing Lindbladian quantum tomography with trapped ions

Dmitrii Dobrynin^{1,2}, Lorenzo Cardarelli^{1,2} , Markus Müller^{1,2} and Alejandro Bermudez^{3,*}

¹ Institute for Quantum Information, RWTH Aachen University, D-52056 Aachen, Germany

² Peter Grünberg Institute, Theoretical Nanoelectronics, Forschungszentrum Jülich, D-52425 Jülich, Germany

³ Instituto de Física Teórica, UAM-CSIC, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

* Author to whom any correspondence should be addressed.

E-mail: alejandro.bermudez@csic.es

Keywords: quantum information processors (QIPs), noise benchmarking, maximum likelihood estimation, compressed sensing techniques, Lindbladian quantum tomography (LQT)

Abstract

Characterizing the dynamics of quantum systems is a central task for the development of quantum information processors (QIPs). It serves to benchmark different devices, learn about their specific noise, and plan the next hardware upgrades. However, this task is also very challenging, for it requires a large number of measurements and time-consuming classical processing. Moreover, when interested in the time dependence of the noise, there is an additional overhead since the characterization must be performed repeatedly within the time interval of interest. To overcome this limitation while, at the same time, ordering the learned sources of noise by their relevance, we focus on the inference of the dynamical generators of the noisy dynamics using Lindbladian quantum tomography (LQT). We propose two different improvements of LQT that alleviate previous shortcomings. In the weak-noise regime of current QIPs, we manage to linearize the maximum likelihood estimation of LQT, turning the constrained optimization into a convex problem to reduce the classical computation cost and to improve its robustness. Moreover, by introducing compressed sensing techniques, we reduce the number of required measurements without sacrificing accuracy. To illustrate these improvements, we apply our LQT tools to trapped-ion experiments of single- and two-qubit gates, advancing in this way the previous state of the art.

1. Introduction

The progress on quantum technologies witnessed over the past decades has relied on the development of high-precision techniques to isolate, manipulate, and interrogate quantum systems. When these systems increase in size, as required for instance in quantum-advantage demonstrations based on quantum-information processors (QIPs) [1, 2], the development of efficient calibration and characterization strategies becomes a central task [3, 4]. This requires a mix of tools to estimate quantum states, their time evolution, and the operations used for their measurements, all of which are encompassed within the broad subject of quantum tomography (QT). One may say that QT is both a blessing and a curse, for it provides us with a well-defined route for the detailed learning of quantum systems but, at the same time, it entails a very large complexity. In most cases, this complexity scales with a certain power of the Hilbert space dimension $d = 2^N$, and thus grows exponentially with the number of qubits N . The specific so-called sampling complexity of various QT protocols is discussed in more detail in appendix A, which also serves to set our notation and introduce key concepts in the QT literature. This includes rigorous proofs that incorporate the limited accuracy one can achieve in any realistic system, including noise and the limited number of experimental runs to sample the probability distributions associated to the quantum state.

In light of this exponential complexity, QT in full generality has been mostly limited to small-sized systems, as illustrated in the estimation of quantum states ρ [5–7]. Small-size state-QT has become a

standard practice in various platforms such as photons [8, 9], neutral atoms [10–12], trapped ions [13–15], nuclear magnetic resonance [16, 17] and superconducting circuits [18, 19]. In an effort to minimize the complexity, alternative QT schemes have been proposed, exploiting symmetry [20–22] and entanglement [23–25] arguments to constrain the possible states in order to reduce the QT costs. A reasonable constraint in this respect is that states produced in high-fidelity QIPs are close to ideal pure states, and thus correspond to low-rank density matrices ρ . This can be exploited via compressed sensing (CS) techniques, originally developed for the recovery of sparse signals by random sampling with a rate that is smaller than the one expected from the signal bandwidth [26, 27]. Even if the resources of CS-QT still scale exponentially with the number of qubits N [28–32] (see appendix A), the overall reduction can be important in practice, and has allowed to push state-QT to larger registers [PhysRevLett.113.040503], [33] including the experimental CS-QT of 7-qubit states [34].

QT requirements become even more demanding when one is not only interested in states, but also in their dynamics $\rho(t_0) \mapsto \rho(t) = \mathcal{E}_{t,t_0}(\rho_0)$ [35–39]. As discussed in appendix A, even for a single snapshot of this dynamics at $t \in T = [t_0, t_f]$, the complexity of this so-called process-QT presents an even faster exponential scaling with N . Therefore, most experiments of full process-QT have also been limited to small-sized systems, such as two-qubit entangling gates in nuclear magnetic resonance [40, 41], photons [42, 43], trapped ions [44–46], and superconducting circuits [47, 48]. Paralleling our discussion of possible strategies to reduce the cost of state-QT, one can restrict either the snapshot $\mathcal{E}_{t,t_0}$ to specific quantum channels of Pauli type [49, 50], or apply compressed-sensing techniques assuming the channel has a reduced Kraus rank $r_\kappa \ll d^2 = 4^N$ [29, 51]. Both of these techniques still have a complexity that scales exponentially with N (see appendix A), although they can lead to a practical overall improvement. Note that in the context of high-fidelity QIPs, low-rank channels lie very close to a specific target unitary operation, i.e. a quantum gate, the knowledge of which can be exploited to define a basis and reduce the resource scaling of process CS-QT to a polynomial one [52, 53].

Having discussed this, we can now delve into the central theme of the current work. If one is interested in learning the noisy real-time dynamics of the QIP to estimate which is the optimal time duration of a gate, i.e. the evolution time for which errors are minimized, the above process-QT of $\mathcal{E}_{t,t_0}$ require repeating the whole procedure over and over again for each evolution time one is interested in $t \in T$. Since the dynamics of closed quantum systems must be generated by an underlying Hamiltonian $\mathcal{E}_{t,t_0}(\rho) = U(t, t_0)\rho U^\dagger(t, t_0)$, e.g. $U(t, t_0) = \exp(-i(t - t_0)H)$ if the Hamiltonian is constant, one may sidestep this repetition overhead by focusing on the estimation of the Hamiltonian H [54–57]. Using the N -qubit Pauli basis $E_\alpha \in \mathcal{B}_P$ (A2), the Hamiltonian can be expressed in terms of $d^2 - 1$ real numbers

$$H = \sum_{\alpha=1}^{d^2-1} c_\alpha E_\alpha, \quad (1)$$

which can be grouped in a Hamiltonian vector $\mathbf{c} \in \mathbb{R}^{d^2-1}$, where we have excluded a trivial overall shift of the energies. Since $d^2 = 4^N$, one may naively expect to face similar exponential scalings of the complexity. However, when restricting the type of possible Hamiltonians using microscopic information [58–60], or exploiting entanglement arguments [61], one can again reduce the complexity to a polynomial scaling. If such detailed prior knowledge is not available, one can still exploit rather general constraints on the locality of the interactions in physical systems, and develop QT schemes that employ polynomial resources to estimate the Hamiltonian [62–66] (see appendix A for the description of rigorous scalings).

From the perspective of CS, Hamiltonian tomography makes a drastic assumption by considering dynamical quantum maps $\mathcal{E}_{t,t_0}$ with rank $r_\kappa = 1$. In the context of QIPs, this amounts to limiting the possible noise in gates to a systematic mis-calibration or drift leading to coherent errors. This type of errors does certainly not exhaust all important sources of noise in experiments, and one thus needs to go beyond this limit. For more generic errors, the dynamical quantum map has a larger rank and generally lacks an inverse, falling into the class of completely-positive trace-preserving (CPTP) linear super-operators [67, 68]. In particular, there is a type of CPTP maps called Markovian, which can be divided at any intermediate time $t' \in [t_0, t]$ as $\mathcal{E}_{t,t_0} = \mathcal{E}_{t,t'} \circ \mathcal{E}_{t',t_0}$, such that $\mathcal{E}_{t,t'}$ is also a physical CPTP map [69, 70]. For the time-homogeneous CPTP maps of interest in our work, this divisibility follows from the existence of a Liouvillian generator [71], generalizing the case of the Hamiltonian to a so-called Lindbladian $\mathcal{E}_{t,t_0}(\rho) = \mathcal{E}_{t-t_0}(\rho) = \exp\{(t - t_0)\mathcal{L}\}(\rho)$ [72, 73], acting on physical states as

$$\mathcal{L}(\rho) = -i \left[\sum_{\alpha} c_\alpha E_\alpha, \rho \right] + \sum_{\alpha, \beta} \frac{G_{\alpha\beta}}{2} \left(E_\alpha \rho E_\beta^\dagger - \{E_\beta^\dagger E_\alpha, \rho\} \right). \quad (2)$$

Here, we have introduced the dissipative Lindblad matrix $G \in \text{Pos}(\mathbb{C}^{d^2-1})$, which must be positive semidefinite to guarantee that the dynamical quantum map is indeed a one-parameter family of CPTP channels [74]. The goal of *Lindbladian* QT (LQT) is to estimate the elements of the Lindblad matrix, possibly in conjunction with those of the Hamiltonian, from measurement data. As noted in appendix A, LQT has in principle the same number of parameters $d^2(d^2 - 1)$ to be learnt as full process-QT, with the advantage that it need not be repeated over time. Moreover, the knowledge of the Lindbladian can give more physical insight onto the error sources and how to optimize their suppression, as compared to QT of the full process matrix.

One may naively expect that estimating $\mathcal{E}_{t-t_0}$ via process-QT allows one to find the Lindbladian by taking the matrix logarithm, such that LQT would reduce to standard process-QT. However, as noted in [71, 75–77], this requires searching through the complex logarithm branches, of which there is an infinite number, and can lead to inconsistencies in the presence of errors. LQT thus requires an independent tomographic strategy, the origin of which may be traced back to [78], which put forth a linear inversion method similar to standard process-QT [35, 36]. In the work [40], a microscopically-motivated parametrization of the dynamical quantum map was used for a more accurate extraction of the Lindblad generators from the linear inversion. However, in the presence of errors, the approximate inversion process can lead to unphysical estimates leading to generators that do not yield a CPTP map. In [75, 76], a three-step procedure for Lindblad learning was proposed, which starts with a possible unphysical estimate, and then applies a non-linear least-squares fit with an added penalty for unphysical generators, followed by a final filtering step. Paralleling the advances in Hamiltonian learning, recent works have also exploited the locality of interactions to improve LQT [79–81], although we note that there are no rigorous proofs of the sampling complexity to our knowledge.

In this work, we unveil two directions of improvement for LQT by constraining the estimation in a way that the associated dynamical map corresponds to an admissible physical process. The constraints are imposed via a maximum-likelihood (ML) philosophy, as first considered in the context of state-QT [6] and process-QT [37–39]. For LQT [82, 83], ML estimation goes along similar lines but, instead of using a generic CPTP dynamical quantum map $\rho(t) = \mathcal{E}_{t,t_0}(\rho_0)$, it parametrizes the time evolution using the Lindbladian $\rho(t) = e^{(t-t_0)\mathcal{L}}(\rho_0)$ (2), and incorporates it in a likelihood estimator. The estimation proceeds by a non-linear optimization subject to constraints on the Hermitian nature of the Hamiltonian H , and the positive semidefiniteness of the Lindblad matrix G . We also note that an alternative LQT strategy has been presented in [84], which minimizes a least-square estimator. These LQT methods have been demonstrated in experiments with superconducting circuits [82, 84] and trapped ions [83].

In this work, we partake in the development of ML-LQT, presenting advances that are then applied to trapped-ion QIPs. First, we show that a linearization procedure in the regime of low-error gates attained by modern QIPs transforms the ML-LQT estimator into a convex one which, in turn, allows for a more efficient optimization. To make further improvements, we combine this linearized ML-LQT with CS, exploiting the fact that the noise is not completely unstructured, but is instead controlled by a reduced number of leading noise sources that will depend on the specific QIP. The corresponding Lindblad matrix will be controlled by a few leading generators, which directly translates into its sparseness in that particular basis. We take advantage of this feature by developing an accurate CS estimation of the Lindbladian with a reduced number of measurements, demonstrating that informational completeness in CS-LQT is not necessary. We also introduce statistical convergence criteria to avoid over-fitting the measurement data. Ultimately, we test the proposed LQT methods on experimental data from trapped-ion QIPs. This broadens the previous experimental implementations [83] by considering two-qubit gates, preparing the ground for more efficient LQT schemes for multi-qubit dynamics that will be explored in the future.

This article is organized as follows. In section 2, we describe the ML estimation for LQT. Section 3 contains our results for the linearization of ML-LQT for high-fidelity QIPs, together with two different algorithms for the conjugate gradient (CG) descent. We present a comparison of the linearized ML-LQT to the standard full ML-LQT, and identify in which regimes either of the two linearized algorithms has a better performance. In section 4, we introduce a compressed-sensing technique to improve the linearized LQT for situations in which the noise is structured and sparse, and one can estimate its generators with non informationally-complete datasets. These improvements of LQT by linearization and CS are applied to trapped-ion single- and two-qubit gates in section 5, advancing LQT to trapped-ion experiments with real not-injected noise. We present our conclusions and results in section 6. We include a more detailed discussion of the sampling complexity of QT in appendix A. Appendix B contains a detailed derivation of the linearisation in the context of LQT, whereas appendix C describe the details of two algorithms for linearized ML-LQT.

2. ML LQT

In this section, we present a more detailed account of ML-LQT, starting from the theory of Markovian quantum master equations [74, 85, 86]. As noted in the introduction, for the dynamical quantum map $\mathcal{E}_{t,t_0} = e^{(t-t_0)\mathcal{L}}$ to represent a physically-admissible quantum evolution, the generator (2) must be expressed in terms of a positive semidefinite Lindblad matrix G . Its diagonalization $G\mathbf{u}_n = \gamma_n\mathbf{u}_n$ yields non-negative eigenvalues $\gamma_n \geq 0$ and associated eigenvectors $\mathbf{u}_n = \sum_{\alpha} u_{\alpha,n}\mathbf{e}_{\alpha}$. This allows us to rewrite the Liouvillian (2) in terms of the Lindblad-type master equation

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_{n=1}^{d^2-1} \gamma_n \left(L_n \rho L_n^\dagger - \frac{1}{2} \{L_n^\dagger L_n, \rho\} \right), \quad (3)$$

where $L_n = \sum_{\alpha} u_{\alpha,n} E_{\alpha}$ are the so-called jump operators, and γ_n play the role of the corresponding decay rates [74]. In the case of atomic qubits, these jump operators are not arbitrary but typically correspond to leading noise sources, such as the amplitude-damping or dephasing processes [35]. Estimating the decay rates and jump operators, especially when there are several noise sources present, provides more direct insight than the full dynamical quantum maps, as one can order the error generators L_n in decreasing order of γ_n . Moreover, the knowledge of γ_n, L_n allows one to assess which ‘software’ and/or ‘hardware’ improvements of the QIP would be most effective in combating the actual noise.

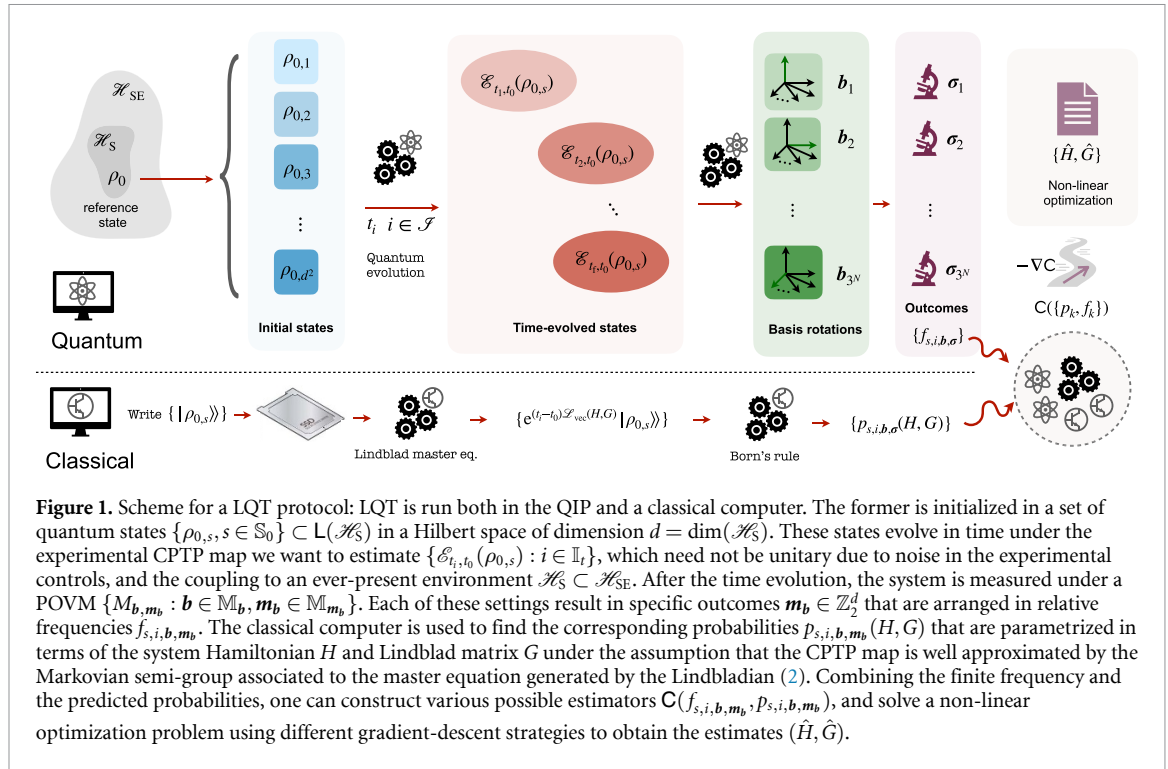
In order to learn the Lindbladian, ML estimation introduces a cost function or estimator that quantifies the differences between the observed and estimated probability distributions, such as the negative log-likelihood estimator

$$\mathbf{C}(p_1, p_2) = - \sum_k p_{1,k} \log(p_{2,k}). \quad (4)$$

Here, the (multi)index k spans the sample space of the probability distribution. In the context of process-QT, p_1 and p_2 correspond to the observed and predicted probability distributions for the measurements performed on the time-evolved state $\rho(t) = \mathcal{E}_{t,t_0}(\rho_0)$, respectively. The observed probabilities $p_{1,k}$ are approximated by the relative measurement frequencies f_k , the components of which are given by the specific observed outcomes for each of the LQT configurations composed of the initialization, evolution, and measurement steps. Going beyond state QT and process QT, these relative frequencies are no longer a vector (A5) nor a matrix (A11), but instead a tensor $f_{s,i,\mu}$ with indexes $k = (s, i, \mu)$. In this work, we consider the positive operator-valued measure (POVM) elements $\{M_{\mu} : \mu \in \mathbb{M}_f\}$ according to local Pauli measurements (A4), which are applied to the time-evolved states from a set (A8) of initial states $\{\rho_{0,s} : s \in \mathbb{S}_0\}$ after a set of evolution times of interest $\{t_i : i \in \mathbb{I}_t\} \subset T$. We note that, in contrast to process-QT which typically focuses on a single channel at a single snapshot, the estimator (4) does in general include various evolution times [82]. As discussed in more detail below, we remark that these times need not densely cover the interval of interest $|\mathbb{I}_t| \gg 1$. In fact, for the Markovian evolutions hereby studied, it will suffice to use a single snapshot $|\mathbb{I}_t| = 1$ to learn the Lindblad generators, whereas more snapshots will be required in situations in which the noise is time-correlated [87]. Once we obtain an estimate of the generators, it is possible to integrate the corresponding Lindblad master equation to infer the dynamics at any desired time $t \in T$.

The estimator in equation (4) also depends on the predicted probabilities p_2 , which are derived from the solution of the Lindblad equation (3) when considering the same set of POVM elements, initial states, and probing times. Hence, $p_{2,k}(\mathbf{c}, G) = \text{Tr}\{M_{\mu} e^{(t_i-t_0)\mathcal{L}(\mathbf{c}, G)}(\rho_{0,s})\}$, such that the log-likelihood estimator (4) is implicitly parametrized in terms of the Hamiltonian and Lindblad matrix (2). Due to the time-homogeneous character of the Markovian dynamical map, we can set $t_0 = 0$ without loss of generality, since there are no memory effects in the quantum evolution that depend on the specific initial time. From this perspective, the estimator (4) is related to the likelihood function $\mathbf{L}(\{f_k, p_{2,k}\})$ for the joint multinomial probability distribution with which the outcomes would be observed with relative frequencies $\{f_k\}$, assuming an underlying statistical model that is parametrized by the Hamiltonian and Lindblad matrix $\{p_{2,k}(\mathbf{c}, G)\}$. The maximum of this likelihood function gives the model for which the observed outcomes are most probable, and corresponds a minimum after taking the negative logarithm and rescaling the result $\mathbf{C}(\{f_k, p_{2,k}(\mathbf{c}, G)\}) \propto -\log(\mathbf{L}(\{f_k, p_{2,k}\}))$ as in equation (4).

The ML-LQT protocol is schematized in figure 1, where we show different triples $(\rho_{0,s}, t_i, M_{\mu})$ that will be referred to as configurations. Starting from a single reference state, a set of initial states $\{\rho_{0,s}, s \in \mathbb{S}_0\}$ is prepared by acting with local single-qubit gates, after which the system evolves under the Lindbladian we aim at estimating for different times $\{t_i, i \in \mathbb{I}_t\}$, and is finally measured according to a POVM $\{M_{\mu}, \mu \in \mathbb{M}_f\}$. We



assume for now on that one performs local Pauli measurements (A4), such that $\mu = (b, m_b) \in \mathbb{M}_f = \mathbb{M}_b \times \mathbb{M}_{m_b}$, where the vector b specifies the $\{x, y, z\}$ basis for the measurement of each qubit, while the vector $m_b \in \{+, -\}^d$ determines the corresponding possible binary outcomes. Hence, the number of independent configurations (A10) is larger than the number of parameters to be estimated per time step $n_{\text{conf},i} = 3^N d^2 (d-1) \geq d^2 (d^2 - 1)$, such that the LQT scheme is informationally complete (IC). The log-likelihood estimator reads

$$\mathbf{C}_{\text{full}}(\mathbf{c}, G) = - \sum_{s,i,\mu} f_{s,i,\mu} \log \left(\text{Tr} \left\{ M_{\mu} e^{(t_i-t_0)\mathcal{L}(\mathbf{c}, G)} (\rho_{0,s}) \right\} \right), \quad (5)$$

where the relative-frequency tensor $f_{s,i,\mu} = N_{s,i,b,m_b} / N_{s,i,b}$ is an approximation to the observed probability, calculated by the ratio of the number of observed m_b -outcomes N_{s,i,b,m_b} with respect to the number of measurements $N_{s,i,b} = \sum_{m_b} N_{s,i,b,m_b}$ for a particular initial state, evolution time and measurement basis. Once the estimator is defined, its minimization subject to a positive semidefinite constraint provides the full ML estimate of the Lindbladian

$$\begin{aligned} (\hat{\mathbf{c}}_{\text{full}}, \hat{G}_{\text{full}}) &= \text{argmin} \{ \mathbf{C}_{\text{full}}(\mathbf{c}, G) \} \\ &\times \text{subject to } \mathbf{c} \in \mathbb{R}^{d^2-1}, G \in \text{Pos}(\mathbb{C}^{d^2-1}). \end{aligned} \quad (6)$$

We note that the positive semidefinite constraint can be imposed by using a Cholesky decomposition $G = L_G L_G^\dagger$ in terms of a lower triangular matrix L_G with real and positive diagonal entries, paralleling some of the approaches for state-QT [76, 88]. As a result, we can perform unconstrained numerical optimization using general purpose optimization libraries, such as the C++ ALGLIB [89] for nonlinear CG descent. For the descent iterations, we need to numerically approximate the gradient, which requires solving the Lindblad master equation (2) for sets of neighboring $\mathbf{c}, G$ parameters, which we accomplish by vectorizing the density matrix $\rho \rightarrow |\rho\rangle\rangle \in \mathbb{C}^{d^2}$, and calculating matrix exponentials of the corresponding Lindbladian. We will refer to this method as full ML-LQT [82, 83] to distinguish it from the optimized routines for linearized and compressed-sensing ML-LQT that have been devised in this work.

3. Linearized LQT in high-fidelity QIPs

Let us start by describing the first improvement for Lindblad tomography proposed in this work: linearization. The full ML-LQT (6) is a non-convex optimization problem that can present multiple local minima and saddle points, such that convergence to a global minimum is not guaranteed. As a consequence,

one may end up in a local minimum yielding a biased estimate of the Hamiltonian, dissipation rates and/or jump operators. To overcome these limitations, we propose to use a linearization that can be applied in high-fidelity QIPs in which the coherent part of the time evolution is known $U(t, t_0) = \exp(-i(t - t_0)H)$, such that H can be excluded from the learning process. We can then restrict the estimation procedure to the effect of the Lindblad matrix G during a certain time scale Δt that is small even if the effect of the coherent evolution is not. This is the case of gates in current QIPs, which are mildly affected by weak external noise.

With these assumptions, we can apply a linearization procedure based on error process matrices [90] to the current ML-LQT, and show that the constrained minimization problem (6) turns into a convex one. In this regime, as described in detail in appendix B, the parametrized probabilities can be written as linear functions of the Lindblad matrix

$$p_{s,i,\mu} = p_{s,i,\mu}^u + \sum_{\alpha,\beta} \Phi_{s,i,\mu}^{\alpha\beta} G_{\alpha\beta}, \quad (7)$$

where the contribution from the ideal gate unitary reads

$$p_{s,i,\mu}^u = \text{Tr} \{ M_\mu U(t_i, t_0) \rho_{0,s} U^\dagger(t_i, t_0) \}. \quad (8)$$

The weak Markovian noise linear in G that we aim at learning is fully contained in the second term, which depends on the following set of matrices $\{\Phi_{s,i,\mu}\}$ for each configuration, which are expressed similarly to equation (8), namely

$$\Phi_{s,i,\mu}^{\alpha\beta} = \text{Tr} \{ M_\mu U(t_i, t_0) \delta \rho_s^{\alpha\beta}(t_i, t_0) U^\dagger(t_i, t_0) \}. \quad (9)$$

The specific effect of the noise is thus contained in

$$\delta \rho_s^{\alpha\beta}(t_i, t_0) = \sum_{\alpha\beta} \int_{t_0}^{t_i} dt' [\mathbb{W}^\dagger(t') \mathbb{B}^{\alpha\beta} \mathbb{W}(t')]_{\alpha\beta} E_\alpha \rho_{0,s} E_\beta^\dagger, \quad (10)$$

which depends on the matrices $\{\mathbb{B}^{\alpha\beta}\}$ and $\mathbb{W}(t')$ defined in equations (B4) and (B10). We refer the reader to the appendix B for the full derivation, which starts with a Lindblad representation in an operator basis different from equation (2) in order to prepare the ground for the next compressed-sensing improvement. The linearized log-likelihood estimator obtained after this weak-noise approximation reads as follows

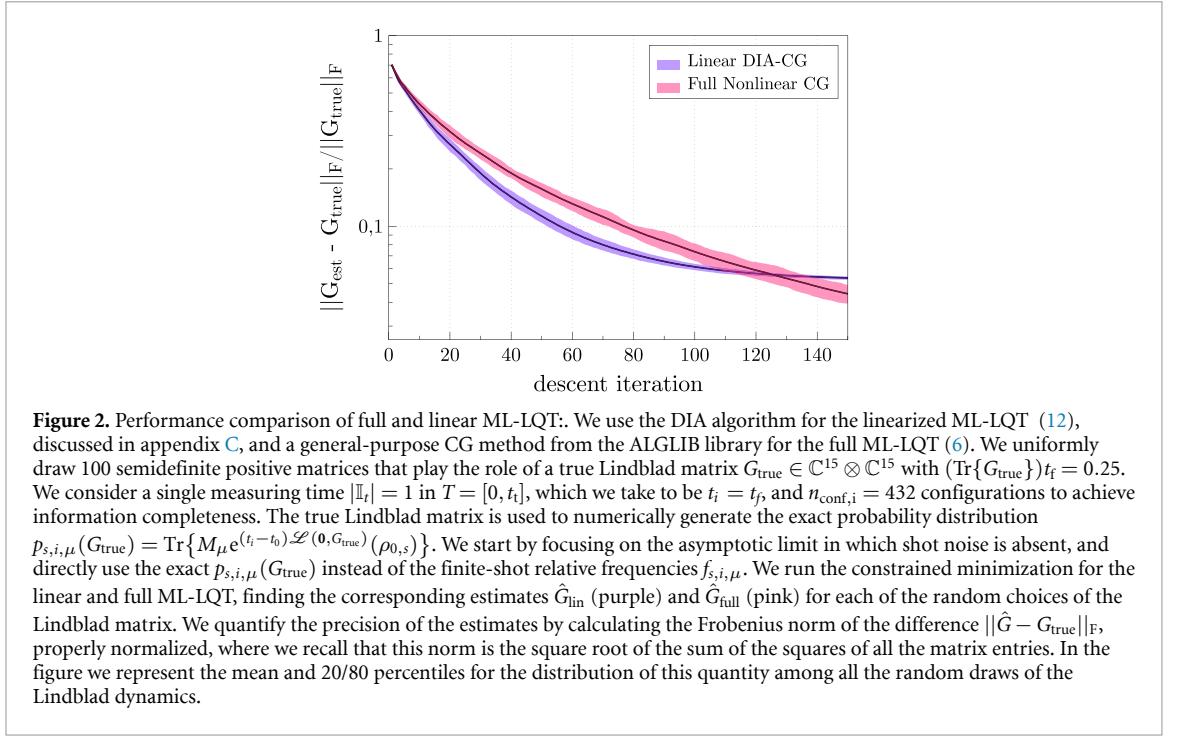
$$\mathbf{C}_{\text{lin}}(G) = - \sum_{s,i,\mu} f_{s,i,\mu} \log \left(p_{s,i,\mu}^u + \sum_{p,q} \Phi_{s,i,\mu}^{pq} G_{pq} \right). \quad (11)$$

Hereafter, the method based on equation (11) will be referred to as linear ML-LQT. Let us comment on the relevance of this linearization for the optimization. In the full ML-LQT (5), when considering a complete knowledge of the Hamiltonian, one must numerically integrate the full Lindblad master equation for each time t_i , given an initial guess of the Lindblad matrix $\hat{G}_{\text{full}}$, and update it according to a gradient-descent method in the search for the optimal solution to the non-convex minimization of equation (6). In the linearized estimator (11), the evaluation of equation (9) still requires a numerical integration at each t_i , but this is highly simplified as we only need to exponentiate the Hamiltonian for an infinitesimal time step once, and not repeat the exponentiation of the vectorized Lindbladian for each updated value of $\hat{G}_{\text{lin}}$. Moreover, the important advantage of the linearized ML-LQT is that the optimization problem

$$\begin{aligned} \hat{G}_{\text{lin}} = \text{argmin} \{ \mathbf{C}_{\text{lin}}(G) \} \\ \times \text{subject to } G \in \text{Pos}(\mathbb{C}^{d^2-1}), \end{aligned} \quad (12)$$

has become convex, as the estimator is now a linear function of G , and the positive semidefinite constraint draws a convex cone over the Lindblad matrices. Linearized ML-LQT is thus guaranteed to have a unique solution $\hat{G}_{\text{lin}}$ in contrast to $\hat{G}_{\text{full}}$, which can have a more complex minima landscape.

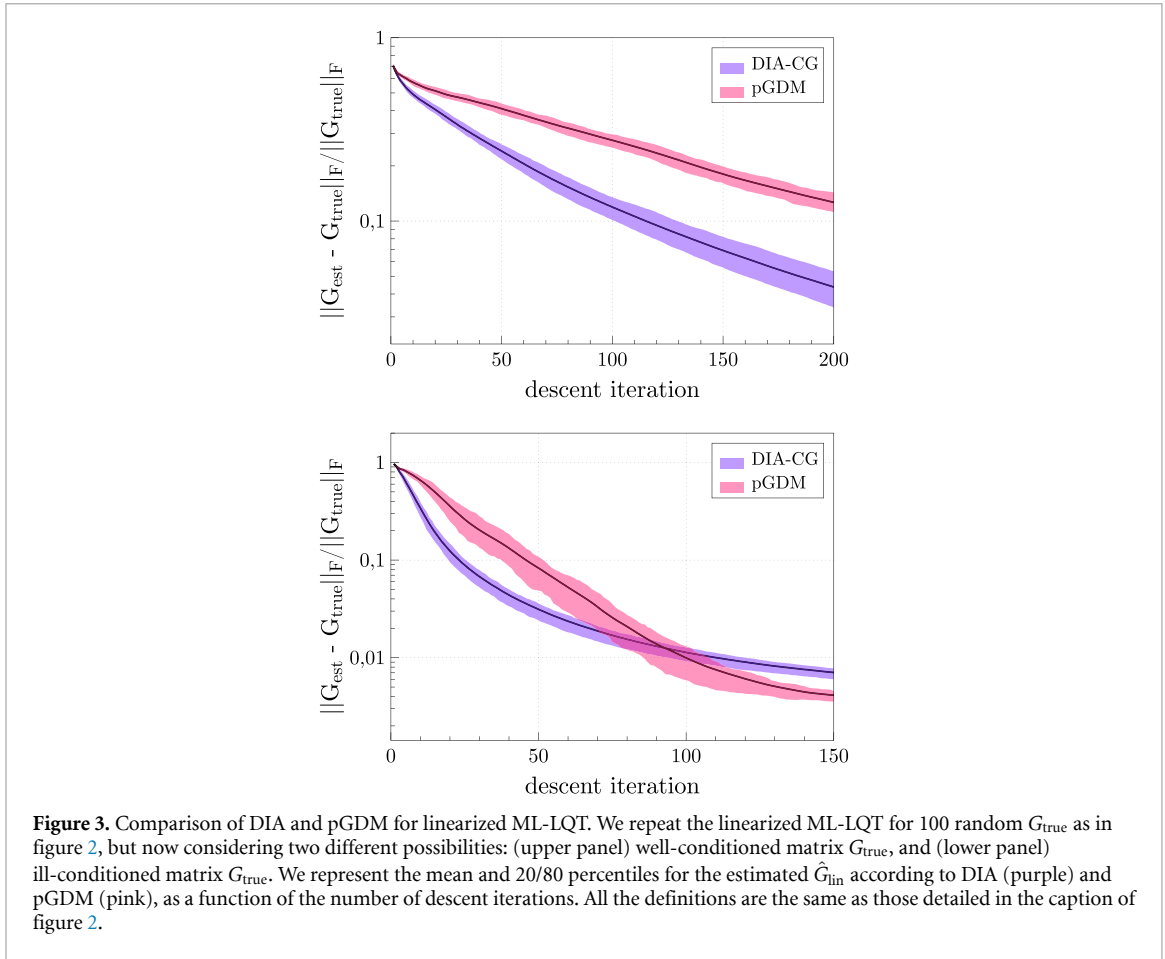
In appendix C, we present two efficient methods for the convex minimization of the linearized ML-LQT estimator (11), namely the diluted iterative algorithm (DIA) and the projected gradient descent with momentum (pGDM). Both of these methods have been considered in the context of state-QT [88, 91] but, to our knowledge, not for LQT. We start by comparing the generic CG methods for the full ML-LQT (6) to the DIA approach for the linearized ML-LQT (12), which makes use of the analytical expression for the gradient of the linear estimator $\mathbf{C}_{\text{lin}}$ (11), together with a line search and a CG descent, to converge towards the unique global minimum $\hat{G}_{\text{lin}}^{\text{DIA}}$. We consider a two-qubit system $N = 2, d = 4$, subject to a small Markovian



noise corresponding to a random Lindblad matrix G_{true} obtained by sampling a uniform random unitary in the Haar measure sense [92]. This unitary is applied to an arbitrary state $|\psi\rangle \in \mathbb{C}^{15} \otimes \mathbb{C}^{15}$, after which one traces over one of the subsystems. As a result, we get a random positive semidefinite matrix $G_{\text{true}} \in \text{Pos}(\mathbb{C}^{15})$ of trace one, sampled uniformly according to the Hilbert–Schmidt distance [93]. To change the scale of G_{true} , we multiply the result by any desired prefactor, which is here set to $(\text{Tr}\{G_{\text{true}}\})_{t_f} = 0.25$. The Hamiltonian H_{true} for each generated G_{true} is chosen to produce a $\pi/2$ single qubit rotation along a random axis for each run. Once a true Lindbladian is randomly chosen, we simulate numerically the dynamics without any approximations, and obtain the exact POVM probabilities $f_{s,i,\mu} \mapsto p_{s,i,\mu}$, focusing on a single snapshot $t_i = t_f$. By using the exact probabilities, we momentarily dispense with the effects of shot noise in order to compare the convergence of the full and linearized ML-LQT approaches. In sections below, we will go beyond these approximations and also consider real experimental data with shot noise and other SPAM errors.

In figure 2, we show how both the full-CG (6) (pink) and linear DIA (12) (purple) minimizations yield a value of the estimated Lindblad matrix $\hat{G}$ that gets closer to G_{true} with each descent iteration. One can see how, on average, the convergence of the linear DIA requires less descent iterations than the full ML-LQT, which is further supported by the guaranteed convergence of DIA to the global minimum ensured by the convexity of the linearized estimator. Moreover, the calculation of the gradient for the full ML-LQT requires that, at each step, one must solve the full Lindblad master equation with various values of the estimated matrix G , instead of using the simpler linearized expressions for the gradient. Therefore, in addition to the increased number of descent steps in the full ML-LQT, each of them requires more classical computational resources. This results in the limited precision of the full nonlinear CG together with a significant runtime slowdown of optimization compared to linear DIA; we observed on average $\approx 50\times$ and $\approx 900\times$ difference of optimization time-per-iteration for 1 and 2 qubit estimations respectively, testing on a single-core CPU. As shown in the figure, the full-GC algorithm shows a slower convergence and a wider confidence interval. On the contrary, the linear DIA shows a much smoother convergence towards the true Lindblad matrix, even after a small number of iterations, as it learns any random Lindbladian with a similar accuracy showcasing its higher robustness. After a number of descent steps, the DIA error begins to saturate due to the limitations of the inherent linear approximation. In contrast, the full-CG minimization continues to improve, provided that one allows for a sufficiently large number of descent iterations. As expected, the linearized ML-LQT (12) is a leading-order approximation, and one should not aim at accuracies for which other higher-order corrections may become dominant.

In summary, the full ML-LQT can reach arbitrary precision in this idealized shot-noise-free situation, at the expense of a much higher-cost in post-processing times and less robustness due to the lack of convexity. We note that this can be a limitation if the ML-LQT is used as a real-time diagnosis tool to calibrate experimental devices, especially if the noise changes during the post-processing time. In a more realistic situation in which shot noise and other SPAM errors would be present, the accuracy of the



Lindbladian reconstruction cannot reach arbitrary precision, as we will see in more detail below. In the weak noise limit, the higher-order corrections to the already-small effect of the noise may lead to even smaller contributions that get completely overshadowed by the shot noise, unless one repeats the experiment an extremely large number of times, which can be prohibitive in QIP platforms with relatively long cycle times, such as trapped ions. Therefore, in such a situation, aiming at a higher accuracy by switching to the full ML-LQT might actually be counterproductive.

Once the advantages of the linearized vs full ML-LQT have been identified, we can compare two different optimization routines for the linearized case. We have so far explored the convergence of the DIA, which deals with the positive-semidefinite constraint of G explicitly by using a Cholesky decomposition. As discussed in appendix C.2, an alternative approach based on pGDM allows for updates on the estimated $\hat{G}$ which do not satisfy the constraint, but instead project back to the set of physically-admissible Lindbladians in a subsequent step. These methods have been proven to be more efficient in the context of QT of states with a high purity (low rank) [91], as they lie very close to the boundary of the space of physically-admissible states. In this case, imposing the positive semidefinite and unit trace constraints at all steps can dramatically slow down the convergence. In the case of linearized ML-LQT, the Lindblad matrix G has no trace constraints as in state-QT, so it is not *a priori* clear if similar benefits will be found when running a pGDM algorithm for the estimation of a low-rank Lindblad matrix. In order to make a quantitative comparison of the convergence of the two linearized ML-LQT algorithms, we also take random samples of low-rank matrices G_{true} by considering Haar uniformly random projectors.

Once more, we focus on the $N = 2, d = 4$ case, using randomly generated two-qubit Lindblad matrices G_{true} , and setting their scale so that with $(\text{Tr}\{G_{\text{true}}\})_{t_f} = 0.01$. The comparison for these linear estimators is shown in figure 3, where we recall that we are still focusing on the idealized shot-noise-free regime. In the upper panel, when G_{true} is sampled from the Hilbert–Schmidt uniform set and thus has many similar eigenvalues, DIA (purple) exhibits faster convergence than the pGDM (pink). However, when G_{true} is a random projector, we observe in the lower panel that pGDM takes over, converging to closer estimates $\hat{G}_{\text{lin}}^{\text{pGDM}}$ than the DIA $\hat{G}_{\text{lin}}^{\text{DIA}}$ for a sufficient number of descent iterations. For a specific range of descent iterations, the pGDM method shows a much quicker descent for low-rank Lindblad matrices than for uniform ones, although we also note that there is a larger variance depending on the specific random

Lindbladian one is aiming to learn, as compared to the DIA. In summary, we can conclude that, if one suspects that an experiment will be affected by a leading noise source such that G_{true} will only have few substantial eigenvalues, pGDM methods are favored with respect to DIA. In the following section, we will discuss how the general linearized strategy, regardless of the DIA or pGDM approach used for the gradient descent, can actually be improved even further in situations where the noise is really structured by using compressed-sensing techniques.

4. Compressed sensing LQT with structured noise

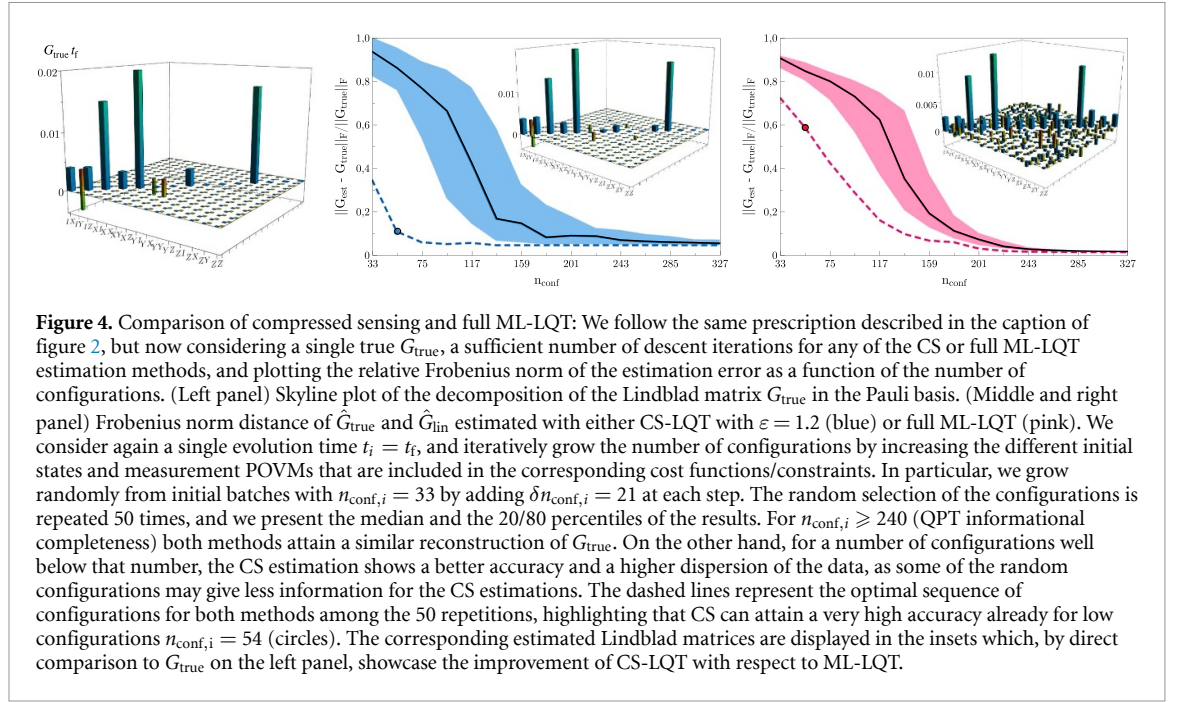
The number of parameters that must be estimated in LQT $d^2(d^2 - 1)$ scales exponentially with the qubit number $d = 2^N$. As noted by the end of last section, however, the Lindblad matrix may have a much smaller number of leading jump operators (3) that describe the main sources of noise. As discussed in appendix A, the operator-sum representations of the dynamical maps (A12) associated to such a reduced Lindbladian would thus have a low Kraus rank, which can be exploited by CS [29, 51, 52, 94, 95]. In fact, for process-QT [52, 53], when one knows the specific unitary operator that lies close to the actual time evolution, CS can be optimized to reduce the number of configurations to a polynomial one by exploiting the sparseness of the process matrix in a certain basis. To our knowledge, the application of compressed-sensing techniques to ML-LQT has not been considered yet. In this section, we explore how this prior on the noise structure can also be exploited for our linearized ML-LQT, allowing us to obtain a faithful estimation of the Lindblad matrix with a much lower number of measurements.

CS originates in the classical theory of signal processing, where the signal is described by a certain sparse vector that can be recovered from an under-sampled set of measurements [26, 27]. A standard approach to maximize sparsity is to minimize the 1-norm of the signal vector, defined as the sum of the absolute values of its components. The incorporation of CS to process-QT allows for the estimation of the process matrix using informationally incomplete data sets by defining a CS estimator that resembles this norm applied to the process matrix [96]. In the present work, we use such a CS estimator to learn the Lindblad matrix in the linear regime of equation (7). This procedure maximizes the sparsity of the Lindblad matrix subject to a constraint with which the Lindblad model must be able to reproduce the observed measurement outcomes with a certain error ε with respect to a least-square measure of the distance between the theoretical and measured probability distributions. This can be formalized through the following constrained convex minimization problem

$$\begin{aligned} \hat{G}_{\text{lin}}^{\text{CS}} = \underset{\alpha, \beta}{\text{argmin}} \left\{ C_{\text{CS}}(G) = \sum_{\alpha, \beta} (|\text{Re}(G_{\alpha\beta})| + |\text{Im}(G_{\alpha\beta})|) \right\} \\ \times \text{subject to } \|f - p^u - \sum_{\alpha, \beta} G_{\alpha\beta} \Phi^{\alpha\beta}\|_2 < \sqrt{n_{\text{conf}}} \varepsilon. \end{aligned} \quad (13)$$

Here, $f, p^u \in \mathbb{R}^{n_{\text{conf}}}$ are, respectively, the observed relative frequencies and the contribution of the ideal unitary evolution to the estimated probabilities (8), both of which have been vectorized in the space of configurations. The total number of configurations per time step is $n_{\text{conf},i} = 3^N d^2(d-1)$, although we remark that the CS method will use a smaller number as we are interested in reducing the sampling complexity. Likewise, we have introduced a set of vectors $\Phi^{\alpha\beta} \in \mathbb{C}^{n_{\text{conf}}}$ by reshaping the linearized contribution of the Markovian dynamics (9) to the estimated probabilities (7). Let us remark that the distance between the measured and estimated probabilities is no longer present in the cost function, but instead appears in the constraint. While the previous ML estimation aims at inferring the Lindblad model with which one could predict the observed outcomes with a higher likelihood, CS attempts to find the sparsest Lindblad model at the expense of obtaining a lower probability to reconstruct the observations. In fact, the above ε acts as a trade-off parameter, as it balances the tendency to fit the experimental data minimizing the least-square distance or maximizing the sparseness of $\hat{G}$.

Although, as advanced in the introduction and explained further in appendix A, CS for process-QT can get considerable improvements even without prior knowledge of the sparsifying basis [29, 51], one can get further improvements towards a polynomial scaling when this is known [52, 53]. This basis dependence becomes very transparent at the level of the Markovian jump operators (3). To quantify these possible improvements, we start by considering LQT for a 2-qubit system as before, but considering a restricted set of Markovian noise jump operators $L_n \in \{L_{\text{deph},1}, L_{\text{deph},2}, L_{\text{damp},1}, L_{\text{damp},2}, L_{\text{bf}}\}$ instead of randomly sampling an unstructured Lindblad matrix. We include single-qubit dephasing with rates $\gamma_{\text{deph},1} \neq \gamma_{\text{deph},2}$ and jump operators $L_{\text{deph},1} = \sigma_z \otimes \mathbb{1}_2$, $L_{\text{deph},2} = \mathbb{1}_2 \otimes \sigma_z$, single-qubit amplitude damping with rates $\gamma_{\text{damp},1} \neq \gamma_{\text{damp},2}$ and operators $L_{\text{damp},1} = \sigma_- \otimes \mathbb{1}_2$, $L_{\text{damp},2} = \mathbb{1}_2 \otimes \sigma_-$ and, finally, correlated bit-flip errors $L_{\text{bf}} = \sigma_x \otimes \sigma_x$ with



a rate γ_{bf} . All these decay rates and associated jump operators will define the true Lindblad matrix G_{true} we aim at learning.

In light of these jump operators and in order to exploit the sparseness of the Lindbladian, we use the Pauli basis (A2) in the linearization. To obtain an estimate $\hat{G}_{\text{lin}}^{\text{CS}}$ via equation (13), we need to calculate the linearized contributions to the predicted probabilities (7). In the present case, we set $H = 0$ such that the ideal evolution is the identity and we are simply estimating the qubit decoherence as a quantum memory. Hence, the vector $\mathbf{p}^{\text{u}}$ (8) can be readily obtained. Finally, the set of vectors $\Phi^{\alpha\beta}$ (9) require a numerical evaluation of equation (10) which, ultimately, requires using $b_{\alpha}^{\text{p}} = \delta_{p,\alpha}$ in equations (B4)–(B5) for the Pauli-basis choice. Once all these quantities are at our disposal, we employ the general augmented Lagrangian algorithm from the C++ ALGLIB library [97] for the constrained minimization (13) choosing a target ε parameter.

The main motivation for turning to CS-LQT techniques is that one can obtain accurate estimations using a smaller set of configurations. We recall that, for the 2-qubit case, the total number of configurations in the ML-LQT is $n_{\text{conf},i} = 3^N d^2 (d - 1) = 432$ per time step, while informational completeness requires $d^2 (d^2 - 1)^2 = 240$ linearly-independent configurations. In order to address the accuracy of the CS-LQT in under-sampled situations, we randomly draw a small number of triples $(\rho_{s,0}, t_i, M_{\mu})$ from all possible configurations, and gradually increase the set by subsequently including $\delta n_{\text{conf},i}$ additional random configurations until all of the $n_{\text{conf},i} = 432$ configurations are incorporated, considering again a single $|\mathbb{I}_t| = 1$ snapshot at $t = t_f$. In this informationally-complete regime, the CS-LQT should give similar estimates to the linearized ML-LQT.

For each CS estimation (13), which uses a specific choice of configurations for which the measurement data would be collected, we calculate the Frobenius norm of the difference between the estimated $\hat{G}_{\text{lin}}^{\text{CS}}$ and the true sparse Lindblad matrix G_{true} , which is depicted in the left panel of figure 4. The central and right panels illustrate the result of our numerical simulations, where the true G_{true} is used to generate the $f_{s,i,\mu} \mapsto p_{s,i,\mu}(G_{\text{true}})$ probability distribution, dispensing again with the effects of shot noise and SPAM errors for the moment. We depict the normalized Frobenius distance between the estimated and true Lindblad matrices as a function of the number of configurations being included in the linear CS-LQT (central panel) and ML-LQT (right panel) estimators. The solid lines represent the mean, while the shaded areas are the 20/80 confidence intervals, showing how the accuracy of the estimation grows for different random ways in which the number of configurations is increased. In addition, the dashed lines represent the optimal configuration choice for this specific G_{true} among the 50 repetitions, with a filled circle that indicates the Frobenius distance for $n_{\text{conf},i} = 54$, being the corresponding estimated $\hat{G}_{\text{lin}}^{\text{CS}}$, $\hat{G}_{\text{lin}}^{\text{ML-LQT}}$ represented in the corresponding insets. The comparison of these dashed lines shows that the CS-LQT strategy can reach a much more accurate estimate for an under-sampled data set than the linear ML-LQT approach. As we can directly see in the insets, the CS-LQT estimated Lindblad matrix resembles more closely the true one (left panel), whereas the linear ML-LQT estimation is corrupted by many small non-zero coefficients, lowering considerably the overall estimation accuracy. The superiority of CS-LQT is a consequence of the sparseness

of the true Lindblad matrix G_{true} , and the underlying sequence of configurations underlying the optimal CS-LQT estimation are those that providing more relevant information about the noise, determined by the Lindblad matrix eigenbasis. The identification of this eigenbasis can be based on microscopic knowledge about the QIP, or from the knowledge acquired by running full ML-LQT for small system sizes, and then extrapolating the conclusions about the noise structure to larger systems.

In summary, ML-LQT chooses to retain the small components of the estimated Lindblad matrix in order to reach a higher likelihood between the measured and estimated probabilities. However, this leads to a larger error with respect to CS-LQT in undersampled regimes, as the latter neglects almost all of those very small components by minimizing the sparseness-based cost function (13). The CS estimate captures the most prominent noise components at this level, providing more than a ten-fold reduction in the number of shots with respect to the ML-LQT with all 432 configurations. Therefore, if a large number of measurements are inaccessible experimentally, or one scales to larger system sizes where the shots need to be distributed among more configurations, CS-LQT will offer a superior solution to ML-LQT if one has some prior knowledge about the structure and sparseness of the noise.

5. Single- and two-qubit trapped-ion LQT

Trapped-ion systems have played a key role in the development of quantum computers. In the seminal work [98], it was proposed that ion crystals can function as registers for the realization of a quantum computer. In this setup, quantum information is encoded in the electronic levels of the ions and manipulated using a universal gate set that involves additional lasers to excite their collective vibrations. This proposal, which was first implemented in [99], paved the way for extensive experimental and theoretical work, establishing trapped ions as one of the leading platforms in the pursuit of constructing fault-tolerant quantum computers [100, 101].

At present, trapped-ion QIPs have served to experimentally realize various noisy intermediate-scale quantum (NISQ) algorithms over the years [102–108]. Moreover, there are ongoing efforts in developing trapped-ion quantum error correction (QEC) [109–123], and developing a detailed microscopic noise modeling to assess the QEC performance under realistic experimental conditions [124–128]. The success of NISQ and QEC endeavours heavily relies on the high-fidelity universal gate set native to the trapped-ion platform [129–134]. Trapped ions have also pioneered several QT experiments for the characterization of entangled states [14, 15, 25, 135, 136], as well as single- and multi-qubit gates [44–46, 137–141]. In fact, many pioneering schemes for QT, such as randomized benchmarking [142, 143] and extensions thereof [144], or gate set tomography [145], were first implemented with trapped ions [130, 133, 146–152]. In the context of LQT, there has only been one previous experiment to the best of our knowledge, which focused on the spontaneous emission of a single trapped-ion qubit under different engineered decay channels [83]. However, LQT has not been applied to learn the Lindblad generators of real non-injected noise in high-fidelity gates including, in particular, LQT for noisy two-qubit gates. The goal of this section is to fill in this gap using data from the experiments in [152–154].

In order to apply the linearized ML and CS tools for LQT in real trapped-ion systems, we first need to reconsider the above strategies for a finite number of measurements N_{shots} , such that $p_{s,i,\mu} \approx f_{s,i,\mu}$. As discussed below equation (5) and in appendix A, one records the number of times N_{s,i,b,m_b} that the m_b -outcome is observed for each initial state $\rho_{0,s}$, evolution time $t_i \in T$, and measurement basis b . We consider that N_{shots} are equally distributed among each setting $N_{\text{shots}} = 3^N d^2 |\mathbb{I}_t| \times N_{\text{sc}}$, such that $N_{\text{sc}} = \sum_{m_b} N_{s,i,b,m_b}$ is the same $\forall s, i, b$, and we can obtain the relative frequency by simply taking a ratio $f_{s,i,b,m_b} = N_{s,i,b,m_b} / N_{\text{sc}}$. The ML-LQT and CS-LQT strategies presented in the previous sections will not only be limited by the effects discussed previously, such as the accuracy of the non-linear minimization or the number of configurations included in the estimator, but also by stochastic errors associated to this shot noise. In addition, SPAM errors in the state-preparation and measurement will also affect the inference.

Another point that has not been addressed in detail yet is that, when learning from real data, we do not know the true Lindblad matrix, and thus cannot provide estimates for the accuracy of our estimates, or find how many descent iterations and how many configurations would be required to reach a specific target. We now describe a simple way in which the readout data is not only used to extract the model parameters by ML or CS, but also for error analysis and hypothesis testing. In principle, if one could collect large numbers of data, it would be possible to use the multinomial distribution of the likelihood function to perform a statistical analysis of errors and confidence intervals for the estimation [155]. In this section, we follow a much simpler strategy that could actually be performed during the gradient descent or as one increases the number of configurations, providing us with a simple criterion of convergence: the Pearson χ^2_P test [156]. This test assesses the goodness of a fit between a joint multinomial distribution with probabilities parametrized by the estimated Lindbladian $\hat{c}, \hat{G}$, and the finite-frequency approximation to the distribution

measured in the experiment. The reduced Pearson χ^2_P statistics reads

$$\chi^2_P = \frac{N_{\text{shots}}}{(d-1)} \sum_{s,i,b,m_b} \frac{\left(p_{s,i,b,m_b}(\hat{\mathbf{c}}, \hat{\mathbf{G}}) - f_{s,i,b,m_b}\right)^2}{p_{s,i,b,m_b}(\hat{\mathbf{c}}, \hat{\mathbf{G}})}. \quad (14)$$

This quantity converges to the reduced χ^2 of a Gaussian distribution with variance $\sigma^2_\chi = 2/(d-1)$ in the large- N_{shots} limit, and can be used as a simple indicator of goodness of fit: $\chi^2_P \gg 1$ when the parametrized probabilities are not an adequate model to reproduce the data, and $\chi^2_P \ll 1$ when there is overfitting, and the theoretically parametrized probabilities are improperly accounting for the noise, or the underlying error bars would be over-estimated. A heuristic converge criterion is to stop when the Pearson χ^2_P falls below unity $\chi^2_{P,c} = 1$. An important note on the applicability of the method is that the number of counts in any of the ‘bins’ N_{scf,s,i,b,m_b} should not be too small [156], which might be the case for high-fidelity gates. We can now obtain estimates for the discrepancy between observed outcomes and the values expected under the Markovian Lindblad model as one performs subsequent descents in the non-linear minimizations, or as one includes more configurations in the estimators towards informational completeness. In this way, we can decide when to stop the LQT learning without knowing the true Lindbladian.

5.1. LQT for single-qubit gates

Let us now consider these aspects for the LQT of trapped-ion quantum gates. We start from the simple single-qubit rotations, focusing on a $\pi/2$ rotation around the x axis

$$R_X(\Theta(t_f, t_0)) = e^{-i\frac{\Theta(t_f, t_0)}{2}\sigma_x}, \quad \Theta(t, t_0) = \int_{t_0}^{t_f} dt' \Omega(t'), \quad (15)$$

where $t_f - t_0$ is the gate time, and $\Omega(t')$ defines a possibly-modulated Rabi frequency for the specific qubit transition. The ideal target unitary $U(t_f, t_0) = (\mathbb{I}_2 - i\sigma_x)/\sqrt{2}$ is obtained by setting $\Theta(t_f, t_0) = \pi/2$ in the expression above. As trapped-ion QIPs routinely achieve very high fidelities for single-qubit gates [101], we can start by assuming a perfect knowledge of this unitary, and focusing the LQT on the estimation of the unknown Lindbladian matrix G_{true} . We use the measurement data from the $^{40}\text{Ca}^+$ experiments by P. Schindler *et al* [153, 154], in which the configurations include $|\mathbb{S}_0| = 4$ initial states, a single time at the end of the gate $|\mathbb{I}_t| = 1$, and $|\mathbb{M}_b| = 3$ for the measurements in the $b \in \{x, y, z\}$ single-qubit basis, each of which has a binary Pauli outcome $m_b \in \{-1, +1\}$. Therefore, the total number of independent configurations is $n_{\text{conf}} = 12$, among which the total number of measurements $N_{\text{shot}} = 1.2 \cdot 10^5$ were distributed uniformly. Note that the number of configurations equals the required number of Lindbladian parameters $d^2(d^2 - 1) = 12$ if one were to learn both the Hamiltonian and the Lindblad matrix. To gauge the relevance of the shot noise, we start by running our previous LQT algorithms on numerically simulated data, where we sample N_{shots} times from the ideal probability distribution $p_{s,i,b,\sigma}^u$ (8). We thus generate N_{shots} uniform random numbers and numerically simulate the Bernoulli trials, collecting the number of obtained outcomes $N_{s,i,b,\sigma}$ for a fixed number of shots $N_{sc} = 10^4$ per initial state, evolution time and measurement basis. In the present case, we have a single evolution time $t_i = t_f$. In this way, we simulate the effects of shot noise in the finite frequencies $f_{s,i,b,\sigma}^u$. As there are no other sources of noise, one expects that running the LQT algorithms will provide a structureless Lindblad matrix $\hat{G}_u$, leading to estimated jump operators $\{\hat{J}_n^u\}_{n=1}^3$ (3) that have no preferred direction in the qubit’s Bloch sphere. On the other hand, the eigenvalues γ_n^{sn} can be used to quantify the level of the contribution of the shot noise to the decay rates $\bar{\gamma}_{\text{sn}} = \max\{\gamma_n^u : n \in \{1, 2, 3\}\}$. For $N_{\text{shots}} = 1.2 \cdot 10^5$ shots, we obtain $\bar{\gamma}_{\text{sn}} \approx 0.8 \cdot 10^{-3}$. This sets a lower bound below which the estimated decay rates, which will be derived using the real experimental data below, would be dominated by the shot noise and thus be inconclusive.

Let us now analyze the real experimental data which, in addition to shot noise, will also be afflicted by SPAM errors. In the left panel of figure 5, we display the results of the estimated Lindblad matrix $\hat{G}_{\text{lin}}^{\text{DIA}}$ obtained by solving the linearized ML-LQT (12) for the $\pi/2$ pulse using the DIA for gradient descent. On the right panel, we display the Pearson reduced χ^2_P -test, showing that we fall below the critical $\chi^2_{P,c} = 1$ very fast as the number of descent steps is increased, reaching a minimal value of $\chi^2_{P,\text{min}} = 0.695$ that supports a reasonably good fit. By diagonalizing the estimated Lindblad matrix $\hat{G}_{\text{lin}}^{\text{DIA}}$, we obtain a single leading decay rate $\hat{\gamma}_{\text{max}} = 5.6 \cdot 10^{-3} > \bar{\gamma}_{\text{sn}}$, while the other two are below the 10^{-12} level. Considering that this leading decay rate is almost $10\times$ bigger than the estimated shot noise, we can conclude that our DIA ML-LQT is not limited by shot noise for the number of measurement $N_{\text{shots}} = 1.2 \cdot 10^5$ performed in the experiment. We can thus discern the underlying structure of the sources of noise from the LQT. In fact, the jump operator corresponding to the maximum decay rate is found to have the following decomposition

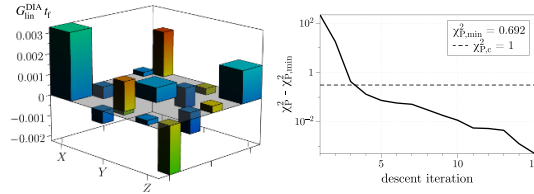


Figure 5. linearized ML-LQT for trapped-ion single-qubit gates: We consider the informationally complete setting with $n_{\text{config},i} = 12$ configurations, each measured projectively with $N_{s,i,b} = 10^4$ shots, and solve the linearized ML-LQT (12) with our DIA. (Left panel) Skyline plot of the estimated Lindblad matrix $\hat{G}_{\text{lin}}^{\text{DIA}}$, showing predominant contributions for elements involving the X and Z Pauli matrices. (Right panel) Pearson χ_P^2 as a function of the gradient descent iteration. After just a few interactions, starting from a diagonal G with $\text{Tr}\{G\}_{t_f} = 0.25$, we already obtain a value well below the $\chi_{P,c}^2 = 1$, signaling a reasonable convergence.

$L_{\text{max}} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z$ with $|n_x| = 0.77$, $|n_y| = 0.36$, $|n_z| = 0.53$, which shows that the main error generator of single-qubit gates is biased towards the gate rotation axis.

We note that, for the single-qubit case, we have also run the CS-LQT (13), but the estimates are similar to those of linear ML-LQT, as the size of the Lindblad matrix is already small and there is no big sparsity that can be exploited. We will show below that CS becomes more useful for two qubits.

5.2. LQT for two-qubit Mølmer–Sørensen (MS) gates

Let us now turn to the LQT of entangling trapped-ion gates. In particular, we focus on the two-qubit MS gate [157, 158]. This gate exploits the quantized vibrations of the ion crystal, i.e. phonons, to generate entanglement between a pair of qubits that belong to the same ion crystal. This gate has become a workhorse in trapped-ion quantum computing, as it allows to achieve high-fidelity gates even in the presence of thermal fluctuations, i.e. without requiring perfectly groundstate-cooled vibrational modes. For two qubits, this unitary gate can be interpreted as the coherent evolution under an Ising-type Hamiltonian

$$R_{XX}(\Theta(t_f, t_0)) = e^{-i \frac{\Theta(t_f, t_0)}{2} \sigma_x \otimes \sigma_x}. \quad (16)$$

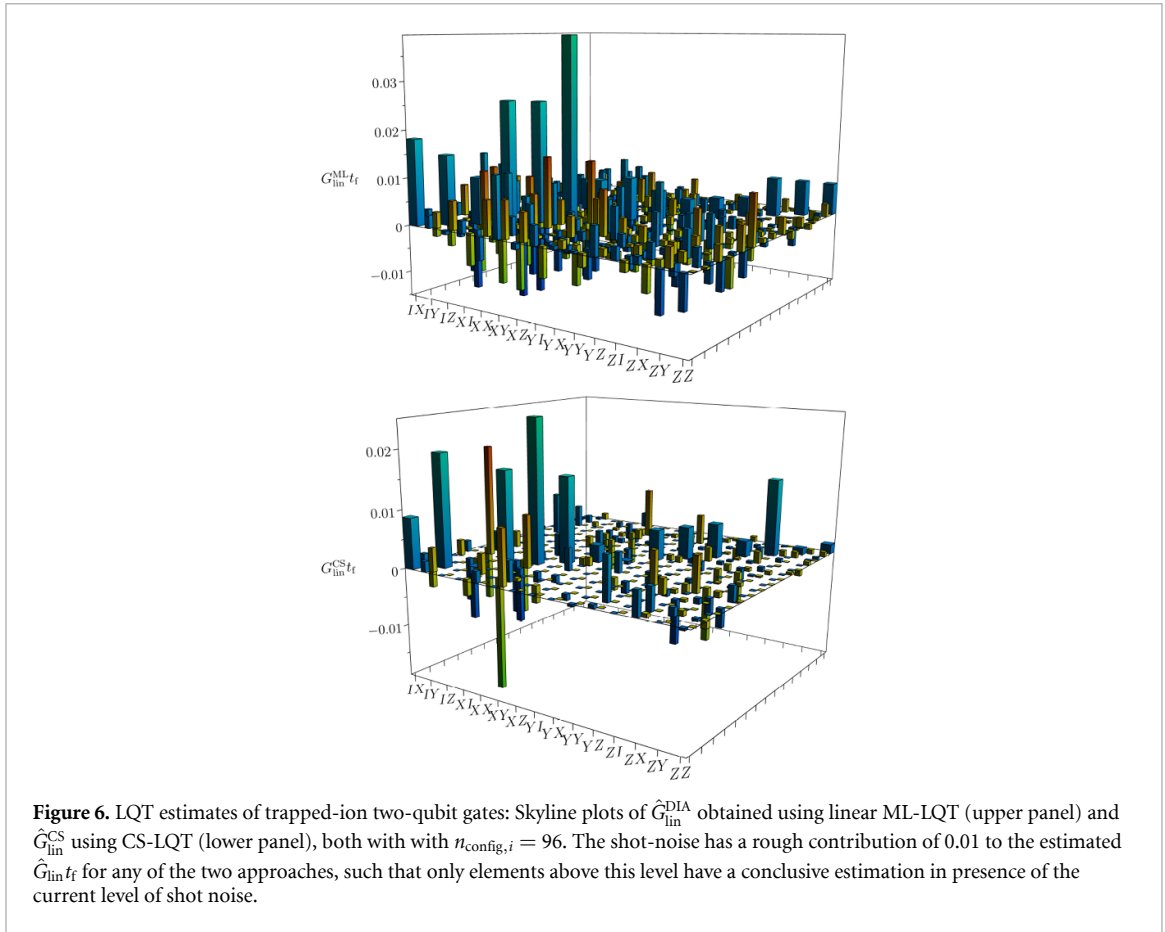
Here, $\Theta(t_f, t_0)$ is the pulse area that depends on the gate time $t - t_0$, laser parameters such as intensity and frequency, as well as the frequencies and Lamb–Dicke parameters associated to the vibrational modes that mediate the interactions [159]. In particular, assuming that the MS gate is obtained by a pair of beams of opposite detuning $\pm \mu_L$ with respect to the qubit transitions, driving the sidebands for the longitudinal center-of-mass mode of frequency ν_z near-resonantly for a two-ion crystal, one finds

$$\Theta(t_f, t_0) = \frac{\eta_{\text{LD}}^2 \nu_z}{8\pi^2} \frac{\mu_L - \nu_z}{\mu_L + \nu_z} \int_{t_0}^{t_f} dt_1' \int_{t_0}^{t_f} dt_2' \Omega(t_1') \Omega(t_2'), \quad (17)$$

where η_{LD} is the Lamb–Dicke parameter, and $\Omega(t)$ is the Rabi frequency. Here, we have assumed that the gate time fulfills $t_f - t_0 = 2\pi/(\mu_L - \nu_z)$, i.e. a single-loop MS gate is realized, and that the laser intensities are calibrated such that $\Theta(t_f, t_0) = \pi/2$, and the target MS gate is $U(t_f, t_0) = (\mathbb{1}_4 - i\sigma_x \otimes \sigma_x)/\sqrt{2}$.

We use the experimental data gathered on the $^{40}\text{Ca}^+$ setup described in [152], which includes $|\mathbb{S}_0| = 16$ initial states, a single time at the end of the gate $|\mathbb{I}_t| = 1$, and $|\mathbb{M}_b| = 9$ for the measurements in the two-qubit basis $\mathbf{b} \in \{xx, xy, xz, \dots, zz\}$, each leading to $|\mathbb{M}_{m_b}| = 3$ independent outcomes, e.g. $\mathbf{m}_b \in \{(+1, +1), (+1, -1), (-1, +1)\}$. Therefore, the total number of independent configurations is $n_{\text{conf}} = 432$, for which the data set contains a total number of measurement outcomes $N_{\text{shots}} = 1.44 \cdot 10^5$ this time. In contrast to the single-qubit case, the configuration set now exceeds the $d^2(d^2 - 1) = 240$ real parameters that are required to determine the Lindbladian, as we already discussed in the previous numerical simulations.

As in the single-qubit case, we can first simulate numerically the ideal MS gate subject to shot noise by sampling N_{shots} times from the ideal probability distribution $p_{s,i,\mu}^u$ (8). Since the number of shots per configuration $N_{\text{sc}} = 10^3$ is smaller than in the single-qubit case $N_{\text{sc}} = 10^4$, one can expect the effect of shot noise to be larger. In fact, we find that the shot-noise threshold for the decay rates is set at $\bar{\gamma}_{\text{sn}} = 1.7 \cdot 10^{-2}$, which is larger than before. As a direct consequence of shot noise, we find that the goodness of fit for the ML-LQT algorithm is worse in this case, reaching chi-square values that still do not reach the convergence criterion $\chi_{P,\text{min}}^2 = 2.1 > 1 = \chi_{P,c}^2$. We depict the estimated Lindblad matrices $\hat{G}_{\text{lin}}^{\text{DIA}}$, $\hat{G}_{\text{lin}}^{\text{CS}}$ for both the DIA for linear ML-LQT and the CS-LQT in figure 6. The skyline plot for the linear DIA displays a fluctuating and non-sparse landscape (upper panel), with most of the peaks below the percent level being associated to the structureless shot noise. It should be noted that the biggest decay rate that can be extracted in this way is



$\hat{\gamma}_{\text{max}} = 6.7 \cdot 10^{-2}$, which is less than four times larger than the shot noise threshold, and thus more sensitive to this noise than the trapped-ion single-qubit case, which had a larger number of shots per configuration. The lower panel of figure 6 displays the CS-LQT estimate for the Lindblad matrix, which is also affected by the same level of shot noise, but has a sparser structure so that most of the small peaks associated to shot noise no longer appear.

In addition to the undesired effect of shot noise, one should also mention that there might be errors in the MS gate that go beyond the LQT assumptions, including SPAM errors of the unitaries used for state preparation and readout, as well as time-dependent and non-Markovian effects that can be expected when considering the dynamical phonons of the two-ion crystal as a small dynamic and thermally-fluctuating environment. In order to discern among these possibilities, insight can be gained by moving to the CS techniques for LQT. After determining $\hat{G}$ using the ML-LQT method, we can diagonalize it and derive a new basis set $\mathcal{A}$ in which the sparsity of the transformed G_{true} matrix would increase, which raises the possibility that CS techniques may actually provide an advantage for LQT. Assuming that the experimental noise only changes mildly in the time between the ML-LQT and the CS-LQT, we can repeat the Lindblad learning using CS in the sparse basis set $\mathcal{A}$, which would point to a route to make the most of the available number of shots by distributing them among the configurations that carry more information about the relevant noise sources. Since the experimental data [152] is already fixed, we will analyze the prospect of this idea using the same data set, but arranging the configurations so that the ones with more information come first to optimally benefit from CS.

In figure 7 we show the results of this analysis. We start by running the full ML-LQT $\hat{G}_{\text{full}}$ considering the IC set of configurations with $n_{\text{config},i} = 432$. With this, we can plot the respective normalized Frobenius distances with respect to the $\hat{G}_{\text{lin}}^{\text{DIA}}$ or $\hat{G}_{\text{lin}}^{\text{CS}}$ estimates, as a function of the number of configurations considered, bearing in mind that there will be an intrinsic error associated to the underlying linear approximation. The pink line represents the accuracy of the linear DIA approx, which lies above the CS-LQT estimate (blue) for small numbers of configurations. We find that only after informational completeness is attained $n_{\text{config},i} = 240$, does the linear DIA approach become preferable with respect to the CS-LQT. Finally, we use the information of $\hat{G}_{\text{full}}$ to learn about the sparsifying basis $\mathcal{A}$, and adapt the CS-LQT approach by incorporating this information in the linearization (9) via a new set of matrices $\{\mathbb{B}^{pq}\}$ (B4). In this way, the Lindblad matrix has a higher sparseness, and the CS sensing can actually capture more accurately the

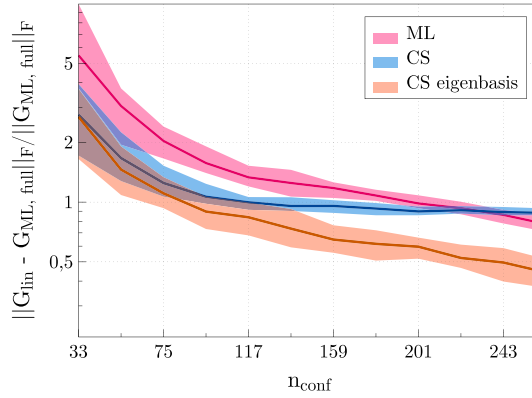


Figure 7. LQT performance for trapped-ion two-qubit gates: We apply the full ML-LQT (6) for $n_{\text{config},i} = 432$ and a total of $N_{\text{shots}} = 1.44 \cdot 10^5$ measurements. The estimated $\hat{G}_{\text{full}}$ is used to calculate the Frobenius distances with respect to the linear estimates based on DIA or CS learning $\hat{G}_{\text{lin}}^{\text{DIA}}$ (pink), $\hat{G}_{\text{lin}}^{\text{CS}}$ (blue), restricted to a smaller number of configurations $n_{\text{config},i} < 432$. The specific configurations are selected randomly, and grown in batches of $\delta n_{\text{config},i} = 21$ following the approach of figure 4. Below information completeness at $n_{\text{config},i} = 240$, one can see how the CS-LQT approach presents a clear benefit with respect to the linear ML-LQT. The solid lines represent the median, and the shaded areas the 20/80 percentiles. The orange solid line (shaded area) represents the mean (percentiles) of the CS-LQT using the $\mathcal{A}$ eigenbasis.

relevant noise sources even for a small number of configurations well-below informational completeness. The results displayed in orange in figure 7 clearly show that the advantage of this idea for the CS-LQT since, when working in the $\mathcal{A}$ basis, we find that the estimation $\hat{G}_{\text{lin}}^{\text{CS}}|_{\mathcal{A}}$ performs, on average, considerably better than the previous approaches, always for configuration numbers that are lower than the informationally-completeness threshold $n_{\text{conf}} = 240$.

6. Conclusions and outlook

In this work, we have introduced and demonstrated two improvements for LQT that advance the characterization of current high-fidelity QIPs. First of all, we have shown how one can linearize the ML estimation of the Lindbladian in the limit of weak noise, which turns the previous non-convex minimizations of LQT [82–84] into a convex problem. We have presented a pair of descent algorithms for linearized ML-LQT that build on the analytical expressions of the gradients of the linear ML estimator, and we have shown that both of these algorithms simplify considerably the classical computational cost of previous full non-convex ML-LQT. We have presented a numerical analysis to show in which situations each of the proposed linear schemes should be preferable, which depends on the rank of the Lindblad matrix one wants to estimate.

This last observation led us to develop a CS approach for LQT, exploiting the structure of the noise, in particular the existence of leading error sources, to estimate the Lindbladian with a smaller number of configurations, such that one could make the most of the available number of measurement shots. We have shown that CS-LQT can yield important advantages in small QIPs that are limited by shot noise, which could also be extended to larger systems in order to partially minimize the exponential scaling in required resources, provided the eigenbasis for CS can be extrapolated from the one learnt for the smaller system sizes. We have applied our improved LQT toolbox to experimental trapped-ion data for single- and two-qubit gates considering, for the first time, real non-injected noise. We have shown that our LQT techniques allow us to extract the leading sources of noise, identifying the main Lindblad jump operators responsible for the incoherent noise. For two-qubit entangling gates, we have shown that CS combined with a method to obtain knowledge about the sparsifying basis can be the key to obtain accurate LQT estimates using a fairly small set of measurement shots.

As an outlook, we believe that it would be interesting to develop LQT further in three directions. On the one hand, it would be desirable to develop rigorous estimates about the sampling complexity, both for the ML-LQT and CS-LQT strategies discussed in this work, putting them at the same level of other QT protocols for state and process estimation in which the resource scalings are clear, and optimal solutions have been devised. Another interesting avenue for further research is the generalization of LQT to account for non-Markovian dynamics in open quantum systems. Here, the possibility to include various intermediate times in the evolution $t_i \in T$ will play a crucial role to capture time correlations of the noise. Searching for optimal schemes depending on the structure of the colored noise is an interesting open question. Finally, we would like to mention a promising direction for future research in LQFT, exploring alternative compressive

tomography techniques [160] that go beyond traditional CS assumptions by incorporating rigorous certification of informational completeness through semidefinite programming. Extending these approaches to encompass LQT can lead to an autonomous verification when the acquired measurement data suffice reconstruct the Lindbladian without requiring *a priori* assumptions on the low rank or the structure of the leading noise generators. Moreover, adaptive strategies that dynamically optimize the choice of the measurement basis along the tomographic procedure are also very interesting, promising further reductions and automation on the required sampling resources. Crucially, these methods do not require prior knowledge of the noise structure or its representation basis, enabling more robust and flexible tomography protocols applicable to complex noise.

Data availability statement

The experimental data used in some parts of the manuscript to run the compressed sensing LQT was shared by the Innsbruck trapped-ion group. This data can be shared upon reasonable request, provided we convey the request to the Innsbruck team. All other data used for numerical simulations can be shared upon reasonable request. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

We would like to thank the Innsbruck trapped-ion group for providing us with the experimental data [152–154] used in section 5. The project leading to this publication has received funding from the US Army Research Office through Grant No. W911NF-21-1-0007. M M and L C acknowledge support by the ERC Starting Grant QNets Grant Number 804247, the EU H2020-FETFLAG-2018-03 under Grant Agreement number 820495, by the Germany ministry of science and education (BMBF) via the VDI within the project IQuAn, by the Deutsche Forschungsgemeinschaft through Grant No. 449905436, and and under Germany's Excellence Strategy 'Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1' 390534769. 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. A B and M M acknowledge support from the European Union's Horizon Europe research and innovation programme under Grant No. 101114305 ('MILLENION-SGA1' EU Project). The authors gratefully acknowledge the computing time provided to them at the NHR Center NHR4CES at RWTH Aachen University (Project Number p0020074). This is funded by the Federal Ministry of Education and Research, and the state governments participating on the basis of the resolutions of the GWK for national high performance computing at universities (www.nhr-verein.de/unsere-partner).

Appendix A. Complexity of QT

This appendix serves to set our notation following that of [155], and to review key QT results for the characterization of QIPs. We will devote special attention to understanding the complexity of the various QT strategies, which will serve to frame more precisely our work on Lindbladian tomography.

In the context of QT, one typically starts from the characterization of quantum states in systems with a single register of N qubits, which can be repeatedly prepared in a specific state ρ . This state, which we aim at estimating from measurement data, is mathematically described by a positive semidefinite operator of unit trace $\rho \in \mathcal{D}(\mathcal{H}_S)$ [68], where $\mathcal{H}_S = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ is the $d = 2^N$ dimensional N -qubit Hilbert space. By repeated preparation, we have n copies of the state at our disposal, which quantify our resources by fixing the number of samples $N_{\text{shots}} = n$ of an underlying probability distribution from which we aim at inferring ρ . These copies are thus measured sequentially for QT according to a POVM [35], which is defined by a collection of $|\mathbb{M}_f|$ POVM elements associated to the measurement outcomes $\{M_\mu : \mu \in \mathbb{M}_f\}$. Each of these POVM elements corresponds to a positive semidefinite operator $M_\mu \in \text{Pos}(\mathcal{H}_S)$ that acts on an individual copy of the state ρ , and the set is constrained to resolve the identity $\sum_\mu M_\mu = \mathbb{1}_d$ [68]. Mathematically, the POVM can be understood as a mapping from state space to a probability space in which the probability vector $\mathbf{p}$ encodes the full measurement statistics. According to Born's rule, this mapping reads

$$\rho \mapsto \mathbf{p} = \sum_\mu \text{Tr}\{M_\mu \rho\} \mathbf{e}_\mu \in \mathbb{R}^{|\mathbb{M}_f|}, \quad (\text{A1})$$

where $\{\mathbf{e}_\mu : \mu \in \mathbb{M}_f\}$ are the standard unit vectors. The components of $\mathbf{p}$ fulfill $p_\mu \geq 0$ and $\sum_\mu p_\mu = 1$, which makes the formal connection to a probability distribution. A measurement is said to be IC if it allows one to invert the above mapping, and use the measured probabilities $\mathbf{p}$ to recover the $d^2 - 1$ real parameters that characterize a generic quantum state $\rho \in \mathcal{D}(\mathcal{H}_S)$ [68].

An IC-POVM must thus contain $|\mathbb{M}_f|_{\text{ind}} \geq d^2$ linearly-independent elements, exceeding by one the required number of parameters to estimate ρ due to the constraint $\sum_\mu M_\mu = \mathbb{1}_d$ [161]. A standard IC-POVM follows from the Pauli basis, where we recall that the operators $\{E_\alpha : \alpha \in \{0, \dots, d^2 - 1\}\}$ form an orthogonal basis of the space of linear operators $\mathcal{L}(\mathcal{H}_S)$ if $\text{Tr}\{E_\alpha^\dagger E_\beta\} = d\delta_{\alpha,\beta}$, where we consider the Hilbert–Schmidt scalar product. By taking tensor products of Pauli matrices, we can define the unnormalized Pauli basis as

$$E_\alpha \in \mathcal{B}_P = \{\mathbb{1}_2, \sigma_x, \sigma_y, \sigma_z\}^{\otimes N}, \quad (\text{A2})$$

leading to Hermitian and involutory basis operators $E_\alpha^\dagger = E_\alpha, E_\alpha^2 = \mathbb{1}_d$, which thus have $m_\alpha = \pm 1$ eigenvalues. We can define a POVM with elements proportional to the orthogonal projectors onto the corresponding eigenspaces

$$\mu = (\alpha, m_\alpha), \quad M_{\alpha, m_\alpha} = \frac{1}{d^2 - 1} \left(\frac{\mathbb{1}_d + m_\alpha E_\alpha}{2} \right), \quad (\text{A3})$$

where we excluded $E_0 = \mathbb{1}_2 \otimes \dots \otimes \mathbb{1}_2 = \mathbb{1}_d$ for $\alpha = 0$. The IC-POVM thus contains $|\mathbb{M}_f| = |\mathbb{M}_\alpha \times \mathbb{M}_\sigma| = 2(d^2 - 1)$ elements, and one typically speaks of $d^2 - 1$ measurement settings with 2 possible outcomes each. We note that only $|\mathbb{M}_{\text{ind}}| = d^2$ POVM elements are linearly independent due to the constraint $(d^2 - 1) \sum_{m_\alpha} M_{\alpha, m_\alpha} = \mathbb{1}_d, \forall \alpha$.

This measurement setup would allow one to infer the non-trivial expectation values $\langle E_\alpha \rangle = 2(d^2 - 1) \text{Tr}\{M_{\alpha, +} \rho\} - 1$ that are required for the estimation of the state ρ [67]. However, except for the single-qubit case, most of these POVM elements correspond to joint operators not native in most QIPs, e.g. $M_\mu \propto (\mathbb{1}_4 + \sigma^x \otimes \sigma^x)$ is a projector onto the even-parity subspace of $N = 2$ qubits, which typically requires entangling CNOT gates preceding an ancilla-qubit projective measurement [67]. To avoid introducing ancillas, which would also need their own characterization, and the additional complexity of working with sequences of entangling gates, we focus on a different IC-POVM with local Pauli projectors

$$\mu = (\mathbf{b}, \mathbf{m}_b), \quad M_{\mathbf{b}, \mathbf{m}_b} = \frac{1}{3^N} P_{b_1, m_{b_1}} \otimes \dots \otimes P_{b_N, m_{b_N}}. \quad (\text{A4})$$

Here, $b_j \in \{x, y, z\}$ indicates the Pauli basis of the j th qubit projector, and $m_{b_j} = \pm 1$ the corresponding eigenvalue, e.g. $P_{x, \pm} = (\mathbb{1}_2 \pm \sigma_x)/2$. Altogether, we have 3^N measurement bases with d possible outcomes each. Therefore, the POVM has $|\mathbb{M}_f| = |\mathbb{M}_b \times \mathbb{M}_{\mathbf{m}_b}| = 3^N d$ elements, although not all of them are independent, as the set of projectors with a fixed basis resolves the identity $\sum_{\mathbf{m}_b} P_{(b_1, m_{b_1})} \otimes \dots \otimes P_{(b_N, m_{b_N})} = \mathbb{1}_d$. Accordingly, we have $|\mathbb{M}_{\text{ind}}| = 3^N(d - 1) + 1 \geq d^2$ independent POVM elements and, thus, an IC-POVM.

According to the QT scenario described above, one would ideally need to prepare the same state ρ repeatedly and measure it a number of times that scales with either 4^N times for global Pauli measurements (A3) or 6^N times for local ones (A4). In practice, however, any of these resource counts is an idealization, as various sources of noise and errors occur in any experiment, rendering the measurements imperfect. At the very least, one is always confronted with projection/shot noise due to the finite number of measurement shots N_b per measurement basis [162], which can only provide us with a relative-frequency approximation of the probability vector

$$\mathbf{f} = \sum_{\mathbf{b}, \mathbf{m}_b} \frac{N_{\mathbf{b}, \mathbf{m}_b}}{N_{\text{shots}}} \mathbf{e}_{\mathbf{b}, \mathbf{m}_b} \approx \mathbf{p}, \quad \sum_{\mathbf{b}, \mathbf{m}_b} f_{\mathbf{b}, \mathbf{m}_b} = 1. \quad (\text{A5})$$

Here, $N_{\mathbf{b}, \mathbf{m}_b}$ stands for the number of observed $\mathbf{m}_b$ outcomes associated to the measurement basis $\mathbf{b}$, such that $N_b = \sum_{\mathbf{m}_b} N_{\mathbf{b}, \mathbf{m}_b}$ and $N_{\text{shots}} = \sum_{\mathbf{b}} N_b$ is the total number of shots in the experiment and, thus, the number of copies of the state ρ that the register must be sequentially prepared into $N_{\text{shots}} = n$. Hence, a more realistic description should account for errors in the inversion of equation (A1), which can lead to the estimation of unphysical states when dealing with overcomplete POVMs, require one to move to ML methods.

In general, QT can only aim at estimating the state ρ by an approximate $\hat{\rho}$ with a certain error $\varepsilon > 0$. This error can be quantified by the trace-distance of the corresponding states $\varepsilon = \|\rho - \hat{\rho}\|_1 = \text{Tr}\{\sqrt{(\rho - \hat{\rho})^2}\}$, where one uses the Schatten 1-norm [68]. Therefore, a more meaningful question regarding the complexity of state-QT is to quantify the resources required to reach a desired target error ε . In the present context, this

should be expressed in terms of the required number of copies of the state at our disposal, which also correspond to the total number of measurement shots that will be performed N_{shots} , and will be larger than the previous idealized scalings. As shown in [30–32, 163], the optimal strategy using a fixed sequence of POVM elements employs randomized measurements, which are obtained by acting with N_{shots} uniformly-sampled random unitaries $U \in \mathcal{U}(\mathcal{H}_S)$ prior to a sequence of specific projective measurements of the resulting states on the computational basis. By finding both upper and lower bounds on the resources, these works show that this scheme is optimal and requires $N_{\text{shots}} \propto 2^{3N}/\varepsilon^2$ copies of the state [30, 31, 163]. Note that a better scaling $N_{\text{shots}} \propto 2^{2N}/\varepsilon^2$ can be achieved if one has access to a multiple-copy register $\rho^{\otimes N_{\text{shots}}}$ and can perform collective entangled measurements [31, 164]. Finally, the sampling complexity can change if the number of outcomes of each of the projective measurements is independent of the system size. An important example is that of the binary-outcome Pauli measurements in equation (A3), where the scaling is instead $N_{\text{shots}} \propto 2^{4N}/\varepsilon^2$ and is also optimal [165].

Due to this exponential sampling complexity, the resources for the standard approach to state-QT become prohibitive already for intermediate-sized systems. As discussed in the introduction, there has been a considerable effort in devising alternative schemes with a lower cost by e.g. restricting the set of possible states to an ansatz with a smaller number of parameters. This can be motivated by a specific symmetry, such as permutation symmetry [20–22], which leads to a polynomial scaling with N_{shots} . Instead, one can restrict the set of possible states according to their entanglement content, and perform efficient QT within the set of area-law states employing polynomial resources [23–25]. A different assumption is that of low-rank QT which, although allowing for a smaller gain, is less restrictive and, arguably, common to the majority of states created in recent experiments. With the low errors achieved by current QIPs, these states are close to ideal pure states $\rho \approx |\psi\rangle\langle\psi|$ and, thus, have a low rank $r \ll 2^N$. Building on ideas of CS to recover a large sparse vector or matrix by randomly sampling a much smaller number of its elements [26, 27], CS QT has been shown to require a number of copies N_{shots} that scales with $\mathcal{O}(Nr^2 2^{2N}/\varepsilon^2)$ when based on Pauli measurements [28, 29]. This scaling can be improved further to $\mathcal{O}(r^2 2^{2N}/\varepsilon^2)$ by considering collective measurements on the multiple-copy register $\rho^{\otimes N_{\text{shots}}}$ [30, 31]. Although these sampling complexities still scale exponentially with the number of qubits N_{shots} , the improvement is considerable in comparison to the previous scaling $N_{\text{shots}} \propto 2^{4N}/\varepsilon^2$ in standard state-QT.

Let us now discuss the complexity of QT for the time evolution of a quantum system, commonly referred to as quantum process tomography [35–39]. This task is also of primary importance in QIPs for which the precise characterization of a universal gate set [35] allows one to identify, model, and possibly amend the errors in a quantum computation. Knowing the precise error model of the native gate set is important for an accurate estimation of the error threshold in fault-tolerant quantum computation [166–170]. In an idealized error-free situation, the exponential scaling of resources follows directly from the description of any admissible quantum evolution by a dynamical quantum map $\mathcal{E}_{t,t_0} \in \mathcal{C}(\mathcal{H}_S) \forall t \in T = [t_0, t_f]$. Each snapshot of the dynamical quantum map belongs to the set of completely-positive and trace-preserving channels $\mathcal{C}(\mathcal{H}_S)$ acting on the space of linear operators $\mathcal{L}(\mathcal{H}_S)$ [68]. In the following discussion, we will only refer to the tomography of a quantum channel, and hence consider a single snapshot at $t \in T$. This channel admits a representation in terms of a process χ matrix

$$\rho_0 \mapsto \rho(t) = \mathcal{E}_{t,t_0}(\rho_0) = \sum_{\alpha,\beta} \chi_{\alpha\beta}(t, t_0) E_{\alpha} \rho_0 E_{\beta}^{\dagger}. \quad (\text{A6})$$

For equation (A6) to represent an admissible physical process, the process matrix $\chi(t, t_0)$ must be semidefinite positive and subjected to a so-called [35] trace constraint

$$\chi_{\alpha\beta}(t, t_0) \in \text{Pos}(\mathbb{C}^{d^2}), \quad \sum_{\alpha,\beta} \chi_{\alpha\beta}(t, t_0) E_{\beta}^{\dagger} E_{\alpha} = \mathbb{1}_d. \quad (\text{A7})$$

Accordingly, $d^2(d^2 - 1)$ real parameters per snapshot are required to describe the time evolution $\rho(t) \in \mathcal{D}(\mathcal{H}_S)$, and one will similarly to state-QT face exponential scalings [171].

In order to determine the time evolution of $\chi(t, t_0)$, the standard process-QT [35, 36] requires the preparation of $|\mathbb{S}_0| = d^2$ linearly-independent initial states $\{\rho_{0,s} : s \in \mathbb{S}_0\}$. A typical choice is to consider all possible tensor products of four states

$$\rho_{0,s} \in \{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |+\rangle\langle +| + |i\rangle\langle i| + |i\rangle\langle +| + |+\rangle\langle i|\}^{\otimes N}, \quad (\text{A8})$$

where $\{|0\rangle, |1\rangle\}$ are the qubit computational basis, and $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|i\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$ are two other cardinal states on the qubit Bloch sphere. Additionally, one must probe the system with an IC-POVM measurement at various instants of time, each of which is described by $|\mathbb{M}_{\text{ind}}| \geq d^2$

linearly-independent elements $\{M_\mu : \mu \in \mathbb{M}_f\}$, e.g. the IC-POVM in equation (A4). One thus gets a formal mapping between the process matrix and a probability matrix

$$\chi(t, t_0) \mapsto [p(t, t_0)]_{\mu s} = \text{Tr}\{M_\mu \mathcal{E}_{t, t_0}(\rho_{0, s})\}, \quad (\text{A9})$$

the columns of which correspond to the probability vectors for the measurement statistics. To recover the full dynamical quantum map, this equation must be inverted [35, 36] at each instant of time $\{t_i : i \in \mathbb{I}_t\} \subset T$, which has an additional overhead increasing the QT complexity even further. The independent triples $(\rho_{0, s}, t_i, M_\mu)$ form our LQT configurations. Considering the set of initial states (A8), and the independent POVM projectors used in our work (A4), we thus have

$$n_{\text{conf}, i} = 3^N d^2 (d - 1) \quad (\text{A10})$$

configurations per time step. Each of this requires preparing a single copy of the time-evolved states, such that the resources in terms of the number of configurations is $n_{\text{config}} = \sum_i n_{\text{conf}, i}$. In QPT, one typically focuses on reconstructing the quantum channel for a single snapshot $t_i \in T$.

In a realistic scenario, one must again consider the estimation error ε , which will require repeating the measurements a total number of times N_{shots} that is much larger than the above configurations. Due to the finite number of repetitions, shot noise only provides us with an approximation to the above probability matrix which, using equation (A4), reads

$$f(t_i, t_0) = \sum_s \sum_{\mathbf{b}, \mathbf{m}_b} \frac{N_{s, i, \mathbf{b}, \mathbf{m}_b}}{N_{s, i, \mathbf{b}}} \mathbf{e}_{\mathbf{b}, \mathbf{m}_b} \otimes \mathbf{e}_s \approx p(t, t_0). \quad (\text{A11})$$

Here, $N_{s, i, \mathbf{b}, \mathbf{m}_b}$ stands for the number of observed $\mathbf{m}_b$ -outcomes associated to each measurement basis $\mathbf{b}$, and now also to each initial state $\rho_{0, s}$ for the specific snapshot $t_i \in T$. Hence, $N_{s, i, \mathbf{b}} = \sum_{\mathbf{m}_b} N_{s, i, \mathbf{b}, \mathbf{m}_b}$ is the number of shots per initialization and measurement basis, and $N_{\text{shots}} = \sum_{s, i, \mathbf{b}} N_{s, i, \mathbf{b}}$ is the total number of shots performed for the process-QT.

In general, the rigorous proofs that underlie our previous discussion on the sampling complexity of state-QT cannot be directly ported to process-QT [51], except in some particular cases. Paralleling our description of state-QT, one can restrict the processes $\mathcal{E}_{t, t_0}$ to specific types, in search for more efficient strategies for process-QT. For instance, if the process is restricted to the family of Pauli channels, one can devise almost optimal QT strategies, and even provide mathematical proofs addressing the sample complexity [49, 172]. In this type of channels, one considers the above Pauli basis $E_\alpha \in \mathcal{B}_P$ (A2) as the orthonormal operator basis in equation (A6), and the channel ansatz is restricted by imposing that the process matrix must be diagonal $\chi_{\alpha\beta}(t, t_0) = p_{P, \alpha}(t - t_0) \delta_{\alpha, \beta}$, where $\delta_{\alpha, \beta}$ is the Kronecker delta. Then, the positive semidefinite constraint simply requires the positivity of $p_{P, \alpha} \geq 0$, and the trace constraint becomes $\sum_\alpha p_{P, \alpha} = 1$, such that the channel is fully determined by a probability vector $\mathbf{p}_P(t - t_0)$ with components known as the Pauli error rates. Process-QT then requires estimating the vector of Pauli error rates by an approximate one $\hat{\mathbf{p}}(t - t_0)$ with a given 1-norm distance $\varepsilon = \|\mathbf{p}_P(t - t_0) - \hat{\mathbf{p}}(t - t_0)\|_1 = \sum_\alpha |p_{P, \alpha}(t - t_0) - \hat{p}_{P, \alpha}(t - t_0)|$.

In a similar spirit to the previous QT scenario, one may consider a quantum register of N_{shots} qubits that can be initialized in different states, then evolved under the channel that we aim at estimating, which may also be interleaved with a fixed sequence of unitary gates. Finally, each of the N_{shots} resulting states can be measured individually with a fixed sequence of POVM elements. In this case, we do not have exact copies of the same state as in state-QT, which are probed by distributing the finite number of N_{shots} experimental shots among the smaller number of measurement settings, but rather a collection of configurations involving different time-evolved states measured in different settings, among which we distribute the total number of shots N_{shots} . As discussed in [49, 50], when the channel to be estimated is a snapshot $\mathcal{E}_{t, t_0}$ of the Pauli type, and one uses intermediate random Pauli gates to average over the noise, the resources scale with $N_{\text{shots}} = \mathcal{O}(N2^{3N}/\varepsilon^2)$ which, up to a logarithmic correction, is similar to the optimal scaling for the QT of states [30, 31, 163]. Indeed, a lower bound presented in [50] shows that, up to these logarithmic corrections, this Pauli-channel tomography is already optimal. It is interesting to note that performing Pauli tomography can actually lead to quantum advantage, as the complexity using a quantum processor can lead to an exponential reduction [173].

Similarly to our discussion of quantum state tomography, we now discuss a less restrictive assumption that allows applying CS for the QT of channels [29, 51, 52, 94, 95]. The process-matrix description (A6) is equivalent to the Kraus operator-sum representation [174, 175], namely

$$\mathcal{E}_{t, t_0}(\rho_0) = \sum_n K_n(t, t_0) \rho_0 K_n^\dagger(t, t_0), \quad (\text{A12})$$

where the Kraus operators are constrained to resolve the identity in order to describe a CPTP map. Indeed, this constraint can be obtained by constructing the Kraus operators as follows

$$\sum_n K_n^\dagger(t, t_0) K_n(t, t_0) = \mathbb{1}_d, \quad K_n(t, t_0) = \sqrt{\chi_n} \sum_\alpha v_{\alpha, n} E_\alpha, \quad (\text{A13})$$

where χ_n and $\mathbf{v}_n = \sum_\alpha v_{\alpha, n} \mathbf{e}_\alpha$ are the eigenvalues and eigenvectors of the process matrix, $\chi(t, t_0) \mathbf{v}_n = \chi_n \mathbf{v}_n$, respectively. For some quantum channels, it may turn out that the process matrix has zero eigenvalues, such that the sum in the Kraus decomposition (A12) is truncated

$$\mathcal{E}_{t, t_0}(\rho_0) \approx \sum_{n=1}^{r_\kappa} K_n(t, t_0) \rho_0 K_n^\dagger(t, t_0), \quad (\text{A14})$$

terminating at a certain Kraus rank $r_\kappa \leq d^2 = 4^N$. This is the case, for instance, of pure-unitary dynamics in which $r_\kappa = 1$.

Just as the QT of most states produced by current QIPs can be improved by assuming low-rank states, the gates that these processors employ to produce such states are very close to unitaries, and thus have a low Kraus rank $r_\kappa \ll 4^N$. As discussed in [29, 51], CS of low-rank channels using a fixed sequence of state preparation, evolution, and Pauli measurements, allow one to derive upper bounds on the required resources $N_{\text{shots}} \leq \mathcal{O}(N2^{5N}/\varepsilon^2)$. Although, to our knowledge, there are no lower bounds to precisely determine the sampling complexity and argue about optimality, the exponential scaling in resources is clear. At this point, it is important to note that, if one could further restrict the low-rank channel by knowing which specific unitary operator the time evolution is close to, which is typically the case of high-fidelity QIPs, one could use a specific basis where the process matrix is sparse, and obtain further improvements for QT. In fact, if one has this prior information, although there are still no strict upper and lower bounds for the sample complexity, the number of QT configurations has been shown to scale polynomially [52, 53].

When the dynamics is purely unitary, instead of estimating the quantum channels for various snapshots to reconstruct a coarse-grained version of the dynamical quantum map, one may directly target the system Hamiltonian H generating such a time evolution. Indeed, for dynamical quantum maps (A12) with Kraus rank $r_\kappa = 1$, the constraint (A13) can be fulfilled by considering a single time-ordered exponential with a time-dependent Hamiltonian $K_1(t, t_0) = \mathcal{T} \left\{ \exp \left(-i \int_{t_0}^t dt' H(t') \right) \right\}$, the estimation of which can be less resource intensive, specially when the Hamiltonian is constant (1). The idea is that the typical Hamiltonians describing physical QIPs do not actually require $d^2 - 1$ parameters when one chooses an appropriate operator basis, such as $E_\alpha \in \mathcal{B}_p$ (A2). Taking into account the tensor-product character of this basis, and the typical locality of interactions, one can restrict oneself to

$$H \approx \sum_{\alpha=1}^m c_\alpha E_\alpha. \quad (\text{A15})$$

Here, the Hamiltonian is parametrized by a smaller set of real coefficients $\{c_\alpha : \alpha \in \{1, \dots, m\}\}$ with $m = \mathcal{O}(\text{poly}(N)) \ll 4^N$ [4], which clearly resembles the situation that motivated the low-rank truncation of the quantum channel (A14). Remarkably, if the initial state commutes with the Hamiltonian, the estimation of the Hamiltonian only requires solving a linear system of equations for the vector of couplings $\mathbf{c}$, which lies in the kernel of an observable correlation matrix C . This matrix consists of equal-time two-point functions in the operator basis $C_{\alpha\beta} = \text{Tr}\{i[E_\alpha, E_\beta]\rho\}$ [62–66, 176]. Allowing for adaptive sequences where the initial state, control fields, and measurements, can be varied according to a Bayesian update, one can find an estimate $\hat{\mathbf{c}}$ considering a fidelity error $F(\hat{\mathbf{c}}, \mathbf{c}) = |\hat{\mathbf{c}}^\dagger \mathbf{c}|^2 / (\|\hat{\mathbf{c}}\|_2^2 \|\mathbf{c}\|_2^2) \leq 1 - \varepsilon$ [176]. In the work [176], the authors showed that, up to logarithmic corrections, the complexity of this Hamiltonian learning is polynomial $N_{\text{shots}} = \mathcal{O}(N^3 k^{3D} / (\varepsilon \Delta)^{3/2})$, where a k -local Hamiltonian in D dimensions has at most $m = \mathcal{O}(Nk^{3D})$ terms, and the correlation matrix is assumed to have an energy gap Δ .

As emphasized in the introduction, in order to account for realistic errors in QIPs, one should upgrade closed-system Hamiltonian QT to open quantum systems that do not evolve unitarily. The theory of open quantum systems aims at describing the dynamics of the system after tracing over an ever-present environment $\rho = \text{Tr}_{\text{E}}\{|\Psi_{\text{SE}}\rangle\langle\Psi_{\text{SE}}|\} \in \mathcal{D}(\mathcal{H}_{\text{S}})$ with $|\Psi_{\text{SE}}\rangle \in \mathcal{H}_{\text{SE}}$, which can be generally described by an exact integro-differential equation known as the Nakajima–Zwanzig equation [74, 85, 86]. When the coupling to the environment is weak, and the timescale of interest is much larger than the environmental correlation time τ_c , one can approximate it by a much simpler Markovian master equation. In more detail, τ_c sets a characteristic timescale for the decay of the environmental correlations and, when the time-scale of interest is much larger $\Delta t \gg \tau_c$, one can coarse-grain to derive a time-local differential equation for the quantum system that is not affected by environmental memory effects. In this Born–Markov limit, this equation has a

Lindblad form (2), such that $d\rho/dt = \mathcal{L}(\rho)$ is fully determined by the Hamiltonian H and Lindblad matrix G . The goal of Lindbladian QT is to estimate both the Hamiltonian H and Lindblad matrix G (2). Let us note that, by choosing the Pauli basis with $E_0 = \mathbb{I}_d$, one can incorporate some of the Lindbladian contributions in the Hamiltonian part $H \mapsto \tilde{H} = H + \sum_{\alpha} \text{Im}\{G_{0\alpha}\}E_{\alpha}$, such that the fully incoherent part of the dynamics is encapsulated in a $(d^2 - 1) \times (d^2 - 1)$ positive semidefinite matrix $G \rightarrow \tilde{G} \in \text{Pos}(\mathbb{C}^{d^2-1})$. In the main text, we avoid the tildes to simplify notation. Hence, in addition to the $d^2 - 1$ real parameters for the Hamiltonian $\{\tilde{c}_{\alpha} = c_{\alpha} + \text{Im}\{G_{0\alpha}\}\}$, we require an additional set of $(d^2 - 1)^2$ real parameters for $\{G_{\alpha\beta}\}$. Altogether this yields $d^2(d^2 - 1)$, which is the same parameter count one finds for the full process matrix.

Appendix B. Linearizing LQT

In this appendix, we elaborate on the linearization of the Markovian evolution operator $\exp((t - t_0)\mathcal{L})$, following an approach based on error process matrices [90]. The main idea in [90] is to factor out the target unitary $U(t, t_0)$ in the estimation of the process matrix $\chi(t, t_0)$ of a noisy gate (A6), and to only learn the error process matrix $\chi^{\text{err}}(t, t_0)$. As discussed in the main text, in the limit of weak noise, this factoring can have important practical consequences in ML-LQT and CS-LQT, rendering the constrained optimization convex and the Lindbladian estimation much more efficient. We assume knowledge of the coherent part of the evolution $U(t, t_0) = \exp(-i(t - t_0)H)$, which narrows the learning task to the sole estimation of the Lindblad matrix G . The state of the system at a time $t_{i+1} = t_i + \Delta t$ can be derived from the Lindblad master equation, which, in the weak-noise limit, can be approximated by a Suzuki-Trotter expansion. Moreover, the dissipative evolution operator now admits a linear approximation $\exp(\mathcal{L}_G \Delta t) \approx (\mathbb{I} + \mathcal{L}_G \Delta t)$ even when the coherent part is not small, where $\mathbb{I}(\rho) = \rho$ is the identity channel. Preparing the ground for the linearized compressed-sensing formulation, we will allow the Lindbladian to be defined as

$$\mathcal{L}_G(\rho) = \sum_{p,q} G_{pq} \left(B_p \rho B_q^{\dagger} - \frac{1}{2} \{ B_q^{\dagger} B_p, \rho \} \right), \quad (\text{B1})$$

where we consider any operator basis $\mathcal{L}(\mathcal{H}_S) = \text{span}\{\mathcal{B}\}$ with $\mathcal{B} = \{B_p : p \in \{0, \dots, d^2 - 1\}\}$, and only impose that $\text{Tr}\{B_p\} = 0$. To simplify the presentation, we keep the same notation of the G matrix as in the Pauli-basis equation (2), although the main idea is that the matrix will be different and sparser if we allow the $B_p = \sum_{\alpha} b_{\alpha}^p E_{\alpha}$ to be certain linear combinations of Pauli operators. A clear and simple example is that of a single-qubit spontaneous emission where using $B_p = (\sigma_x - i\sigma_y)/2 = \sigma^{-}$ can increase sparseness.

The equation for time evolution from the sth initial state can then be expanded as

$$\begin{aligned} \rho_s(t_{i+1}) &\approx U(t_{i+1}, t_i) \rho_s(t_i) U^{\dagger}(t_{i+1}, t_i) + U(t_{i+1}, t_i) \sum_{p,q} G_{p,q} \Delta t \sum_{\alpha,\beta} b_{\alpha}^p b_{\beta}^{q*} \\ &\times \left(E_{\alpha} \rho_s(t_i) E_{\beta}^{\dagger} - \frac{1}{2} \{ E_{\beta}^{\dagger} E_{\alpha}, \rho_s(t_i) \} \right) U^{\dagger}(t_{i+1}, t_i). \end{aligned} \quad (\text{B2})$$

This leading-order evolution can be rewritten compactly as

$$\rho_s(t_{i+1}) \approx \sum_{\alpha,\beta} \left(\chi_{\alpha\beta}^I + \sum_{\gamma,\delta} \mathbb{B}_{\alpha\beta}^{pq} G_{pq} \Delta t \right) U_{\Delta t} E_{\alpha} \rho_s(t_i) E_{\beta}^{\dagger} U_{\Delta t}^{\dagger}, \quad (\text{B3})$$

where we have introduced a short-hand notation $U_{\Delta t} = U(t_{i+1}, t_i)$, and the following quantities following from an expansion in the Pauli basis and standard algebra

$$\chi_{\alpha\beta}^I = \delta_{\alpha,0} \delta_{\beta,0}, \quad \mathbb{B}_{\alpha\beta}^{pq} = b_{\alpha}^p b_{\beta}^q - \frac{1}{2} \left(c_{\beta}^{pq} \delta_{\alpha,0} + \delta_{0,\beta} c_{\alpha}^{pq*} \right), \quad (\text{B4})$$

which also depend on

$$c_{\alpha}^{pq} = \sum_{\gamma,\delta} b_{\gamma}^p b_{\delta}^{q*} \text{Tr} \left\{ E_{\alpha}^{\dagger} E_{\delta} E_{\gamma} \right\}. \quad (\text{B5})$$

We can now make use of equation (A6) to identify a process matrix for the infinitesimal dynamical quantum map in the weak-noise limit. The only difference is that the state resulting from this map is then acted with the unitary as follows

$$\rho_s(t_{i+1}) \approx U_{\Delta t} \left(\sum_{\alpha,\beta} \chi_{\alpha\beta}^{\text{err}}(t_{i+1}, t_i) E_{\alpha} \rho_s(t_i) E_{\beta}^{\dagger} \right) U_{\Delta t}^{\dagger}. \quad (\text{B6})$$

Since we have factored out the effect of the ideal target unitary $U(t_{i+1}, t_i)$ from deviations caused by the dissipation, this process matrix actually contains information about the noise and is known as the infinitesimal error matrix [90, 177], namely

$$\chi_{\alpha,\beta}^{\text{err}}(t_{i+1}, t_i) = \chi_{\alpha,\beta}^I + \sum_{\gamma,\delta} \mathbb{B}_{\alpha\beta}^{pq} G_{pq} \Delta t. \quad (\text{B7})$$

The novelty in our work with respect to [90] is that, as noted above, we allow for a more general general basis, and will consider this linearization in the context of ML estimation.

The contribution from the first term χ^I actually leads to the ideal unitary, whereas the second term $\Delta\chi^{\text{err}}(t_{i+1}, t_i)$ is responsible for the small errors in the weak-noise limit. The full dynamical quantum map for the time evolution $t \in T$ can be obtained by composing the above infinitesimal maps for small time periods Δt to obtain the following temporal sequence of M channels with $\Delta t = (t_f - t_0)/M$ and $t_i = t_0 + (\Delta t)i$. This can be schematically depicted as follows

$$- \chi^{\text{err}}(t_1, t_0) - U(t_1, t_0) - \chi^{\text{err}}(t_2, t_1) - U(t_2, t_1) - \dots - \chi^{\text{err}}(t_M, t_{M-1}) - U(t_M, t_{M-1}) -, \quad (\text{B8})$$

As we can see, the dynamical quantum map is described by the composition of, first, the infinitesimal error channel, followed by the infinitesimal unitary evolution, which is then repeated sequentially. In most theoretical treatments of noisy circuits, however, one models a faulty gate by the action of the full ideal unitary composed with the full error channel. To make connection with these studies, we can swap the order of any unitary V and $\tilde{\chi}^{\text{err}}$ by a simple unitary transformation

$$- V - \chi^{\text{err}} - = - \chi_V^{\text{err}} - V -$$

where we have introduced the following basis transformation for the error process matrix

$$\chi^{\text{err}} = \mathbb{W}^\dagger \chi_{U_{\Delta t}}^{\text{err}} \mathbb{W}, \quad \mathbb{W}_{\alpha\beta} = \text{Tr} \left\{ E_\alpha^\dagger U_{\Delta t} E_\beta U_{\Delta t}^\dagger \right\}. \quad (\text{B9})$$

With this transformation, we can swap all $\chi^{\text{err}}(t_{i+1}, t_i)$ in equation (B8) to the left, such that the dynamical quantum map is described by

$$- \chi^{\text{err}}(t_1, t_0) - \chi_{U(t_1, t_0)}^{\text{err}}(t_2, t_1) - \chi_{U(t_2, t_1)U(t_1, t_0)}^{\text{err}}(t_3, t_2) - \dots - U(t_1, t_0) - U(t_2, t_1) - U(t_3, t_2) - \dots$$

In the first order approximation [90], the composition of a sequence of error process matrices can be written as a sum. Using the group composition for the unitary evolution and letting $\Delta t \rightarrow 0$, we can convert the sums into time integrals, such that

$$\begin{aligned} \chi_{\alpha\beta}^{\text{err}}(t, t_0) &\approx \chi_{\alpha\beta}^I + \sum_{pq} G_{pq} \int_{t_0}^t dt' [\mathbb{W}^\dagger(t') \mathbb{B}^{pq} \mathbb{W}(t')]_{\alpha\beta}, \\ \mathbb{W}_{\alpha\beta}(t') &= \text{Tr} \left\{ E_\alpha^\dagger U(t_0 + t', t_0) E_\beta U^\dagger(t_0 + t', t_0) \right\}, \end{aligned} \quad (\text{B10})$$

making it a linear function of the Lindblad matrix G . As a result, the density matrix for the evolution of an initial state ρ_s also becomes a simple linear function of G .

Once we have obtained the full time evolution in terms of the weak-noise error process matrix, we can apply it to the ML-LQT (12) or CS-LQT (13) discussed in the main text. In both cases, we need to calculate the predicted measurement probabilities for a triple $(\rho_{0,s}, t_i, M_\mu)$. Using the above expression, we find

$$p_{i,s,\mu} = p_{i,s,\mu}^u + \sum_{\alpha,\beta} \Phi_{i,s,\mu}^{\alpha\beta} G_{\alpha\beta} \quad (\text{B11})$$

with the definitions introduced in equations (8) and (9) of the main text.

Since we have assumed to know the gate unitaries by design, and the measurement configurations are experimentally predefined, the p^u probabilities and the Φ matrix should be calculated only once. Afterwards, we choose an estimator for $G \mapsto \hat{G}$, which will be obtained by fitting the modeled values of $p_{s,i,\mu}$ to observed frequencies $f_{s,i,\mu}$. For instance, the log-likelihood cost function is now explicitly written as

$$\mathcal{C}_{\text{lin}}(G) = - \sum_{i,s,\mu} f_{i,s,\mu} \log \left(p_{i,s,\mu}^u + \sum_{\alpha,\beta} \Phi_{i,s,\mu}^{\alpha\beta} G_{\alpha\beta} \right), \quad (\text{B12})$$

which is the solution presented in the main text (11) for linearized ML-LQT, and also appears in the CS-LQT constraint (13) after a least-squares approximation.

Appendix C. Convex gradient descent for linear LQT

In this appendix, we adapt some of the efficient methods for convex quantum state tomography to our linearized ML-LQT. In particular, we consider the DIA [88] and the pGDM [91] in the context of Lindbladian tomography.

C.1. DIA

Let us consider the DIA [88] for our convex problem of linearized ML-LQT, which we rewrite here for convenience

$$\begin{aligned} \text{minimize } \mathbf{C}_{\text{lin}}(G) &= - \sum_k f_k \log(p_k^u + \text{Tr}\{\Phi_k^T G\}) \\ \text{subject to } G &= L_G L_G^\dagger, \end{aligned} \quad (\text{C1})$$

where we recall that the multi-index $k = (s, i, \mu)$ contains all information about the initial state, evolution time, and POVM element, L_G is a lower-triangular matrix, and we have simplified the notation further by using matrix products. In contrast to process-QT, G is not constrained further to have unit trace. The DIA algorithm starts by deriving a closed analytical expression for the gradient by varying $\mathbf{C}$ with respect to G , which, for our linearized ML-LQT estimator, gives

$$\delta \mathbf{C}_{\text{lin}}(G) = - \sum_k \frac{f_k}{p_k^u + \text{tr}\{\Phi_k^T G\}} \text{Tr}\{\Phi_k^T \delta G\} = \text{Tr}\{R \delta G\}, \quad (\text{C2})$$

where we have introduced the matrices

$$R = - \sum_k (f_k \Phi_k^T) (p_k^u + \text{Tr}\{\Phi_k^T G\})^{-1}, \quad (\text{C3})$$

$$\delta G = \delta L_G L_G^\dagger + L_G \delta L_G^\dagger. \quad (\text{C4})$$

We note that R is hermitian, since Φ_a is hermitian. Rewriting the variation in a ‘vector-of-matrices’ form we get

$$\delta \mathbf{C}_{\text{lin}} = \text{Tr}\left\{ \left(L_G^\dagger R, R L_G \right) \cdot \left(\delta L_G, \delta L_G^\dagger \right)^T \right\} = \langle \mathbf{g}_C, \delta \mathbf{z} \rangle, \quad (\text{C5})$$

with vectors defined as follows

$$\mathbf{g}_{\text{DIA}} = (R L, L^\dagger R)^T, \quad \mathbf{z} = (L_G, L_G^\dagger)^T. \quad (\text{C6})$$

We note that the vector $\mathbf{g}_C$ is the gradient of the estimator $\mathbf{C}_{\text{lin}}$ with respect to $L_G, L_G^\dagger$ and, thus, the minimum is found when

$$\mathbf{g}_{\text{DIA}} = \left(\frac{\partial \mathbf{C}_{\text{lin}}}{\partial L_G}, \frac{\partial \mathbf{C}_{\text{lin}}}{\partial L_G^\dagger} \right)^T = \mathbf{0} \quad \Rightarrow \quad R L_G = L_G^\dagger R = 0. \quad (\text{C7})$$

By setting $\delta \mathbf{z} = -\eta \mathbf{g}_{\text{DIA}}$, the estimator $\mathbf{C}_{\text{lin}}$ follows the gradient descent to the global minimum. One can perform different standard line-search techniques to choose an optimal value of η , which decreases $\mathbf{C}_{\text{lin}}$ as fast as possible at each step. For a chosen η_n , the update of the G matrix is given by

$$G_{n+1} = \left(C_n + \delta L_{G,n} \right) \left(C_n^\dagger + \delta L_{G,n}^\dagger \right) = (1 - \eta_n R_n) G_n (1 - \eta_n R_n). \quad (\text{C8})$$

The gradient descent is known to be a sub-optimal optimization algorithm since it possesses no information about the previous directions of the descent, and thus can exhibit zig-zag trajectories for problems with large condition numbers. One of the most efficient deterministic algorithms circumventing this deficiency is CG [178]. Using the recommendation from [179], we employ the Polak–Ribière type of nonlinear CG for the solution of equation (C1).

Require: $G_{\text{init}} = G_0 \rightarrow \mathbf{z}_0 = (L_{G,0}, L_{G,0}^\dagger), \xi \in (0, 1), \varepsilon;$

```

1:  $n = 0$ , compute  $R_0$ ;
2:  $\mathbf{g}_{\text{DIA},0} = (R_0 L_{G,0}, L_{G,0}^\dagger R_0), \mathbf{h}_0 = -\mathbf{g}_{\text{DIA},0};$ 
3: while  $|\text{Tr}\{R_n L_{G,n}\}| > \varepsilon$  do
4:    $n = n + 1;$ 
5:   line search of  $\eta_n$  for  $\mathbf{z}_n = \mathbf{z}_{n-1} + \eta_n \mathbf{h}_{n-1};$ 
6:   compute  $R_n(\mathbf{z}_n);$ 
7:    $\mathbf{g}_{\text{DIA},n} = (R_n L_{G,n}, L_{G,n}^\dagger R_n);$ 
8:   set  $\gamma_n = \max \left\{ \frac{\langle \mathbf{g}_{\text{DIA},n}, \mathbf{g}_{\text{C},n} - \xi \mathbf{g}_{\text{DIA},n-1} \rangle}{\langle \mathbf{g}_{\text{DIA},n-1}, \mathbf{g}_{\text{DIA},n-1} \rangle}, 0 \right\}$ 
9:    $\mathbf{h}_n = -\mathbf{g}_{\text{DIA},n} + \gamma_n \mathbf{h}_{n-1};$ 
10: return  $\hat{\mathbf{z}} = (L_{G,n}, L_{G,n}^\dagger) \rightarrow \hat{G} = L_{G,n} L_{G,n}^\dagger.$ 

```

The line search for η_n at step 5 can be implemented by choosing arbitrary η', η'' , and calculating $\mathbf{z}'_n = \mathbf{z}_{n-1} + \eta'_n \mathbf{h}_{n-1}$ and $\mathbf{z}''_n = \mathbf{z}_{n-1} + \eta''_n \mathbf{h}_{n-1}$. Then, using the three values $\mathbf{C}_{\text{lin}}(\mathbf{z}_{n-1}) = \mathbf{C}_{\text{lin}}(\eta = 0)$, $\mathbf{C}_{\text{lin}}(\mathbf{z}'_n) = \mathbf{C}_{\text{lin}}(\eta = \eta'_n)$, $\mathbf{C}_{\text{lin}}(\mathbf{z}''_n) = \mathbf{C}_{\text{lin}}(\eta = \eta''_n)$, one interpolates a quadratic polynomial $\mathbf{C}_{\text{lin}}(\eta) = a\eta^2 + b\eta + c$, and uses its minimum as the η_n . One should be careful though not to use too large values of η', η'' , because it could result in breaking of the first order approximation in $\|G\|\Delta t \ll 1$ underlying the linearization. We also note that step 8 introduces the so-called the Polak–Ribière factor for the CG descent. In this work we chose $\eta'' = 2\eta'$ and optimized the value of η' for each figure featuring the DIA algorithm from the range (0.1, 1). ξ was fixed at 0.5.

C.2. pGDM

In this part of the appendix, we adapt a different type of methods to deal with the positive-semidefinite constraint in the linearized ML-LQT. In state-QT, the use of a Choleski decomposition to explicitly deal with the constraints of the density matrix has been shown to be responsible for a convergence slowdown of QT algorithms [91, 180, 181], particularly when the state ρ to be estimated approaches a pure state and thus has a very small rank. These works use a method to speed up the estimation by exploiting a projected descent, such that the density matrix is not constrained to be positive-semidefinite along the descent (it is just required to remain hermitian), but it is instead projected back to the physical space after each ML descent iteration. A suggested algorithm in [91] is the pGDM, which showed superior convergence for QT of high-purity and thus low-rank states.

The pGDM for the linearized ML-LQT can be presented as follows. We denote by $\bar{G}$ and G physical positive-semidefinite and possibly unphysical matrices, respectively. Since we do not use the Cholesky decomposition, the gradient of the linear estimator is found by varying with respect to G directly

$$\delta \mathbf{C}_{\text{lin}}(G) = \text{Tr}\{R \delta G\} = \langle g_{\text{pGDM}}, \delta G \rangle, \quad (\text{C9})$$

such that $g_{\text{pGDM}} = R$. The syntax of the algorithm is much simpler in this case.

Require: $G_{\text{init}}, 0.9 < \gamma < 1, \eta > 0, \varepsilon;$

```

1:  $n = 0, \bar{G}_0 = P\{G_{\text{init}}\}, M_0 = 0;$ 
2: while  $|\mathbf{C}_{\text{lin}}(\bar{G}_n) - \mathbf{C}_{\text{lin}}(\bar{G}_{n-1})| > \varepsilon$  do
3:    $n = n + 1;$ 
4:    $M_n = \gamma M_{n-1} - \eta g_{\text{pGDM}}(\bar{G}_{n-1});$ 
5:    $\bar{G}_n = P\{\bar{G}_{n-1} + M_n\};$ 
6: return  $\bar{G}_n.$ 

```

We note that, in step 4, we introduce a so-called momentum increase/change, which is controlled by the ‘friction’ constant γ and the accumulation constant η . As in the DIA of the previous subsection, one should be careful not to set too large values of η that would imply a breakdown of the first-order approximation underlying the linearization. We note that, in contrast to DIA, the gradient descent step η is not optimized by using a line search, nor do we apply a CG strategy. Instead, both η and the descent directions are held fixed. In step 5 of the algorithm, the projection P is done by diagonalizing the corresponding matrix and setting to zero the contribution from all negative eigenvalues. We also note that one could pre-calculate the new momentum update at the position based on the previously accumulated inertia, namely use $g_{\text{pGDM}}(\bar{G}_{n-1} + \gamma M_{n-1})$ instead of $g_{\text{pGDM}}(\bar{G}_{n-1})$. This is called the Nesterov update [182], and it has been shown to be beneficial in many non-linear minimization problems. Additionally, when $\mathbf{C}_{\text{lin}}(\bar{G}_n)$ increases

above a threshold, one can reset the descent momentum to zero [183]. In this work we found the values of $\gamma = 0.99$, $\eta = 3 \times 10^{-4}$, and no Nesterov update to achieve the fastest convergence to the minimum.

ORCID iDs

Lorenzo Cardarelli  0000-0001-5010-5167

Alejandro Bermudez  0000-0002-7331-1139

References

- [1] Aaronson S and Chen L 2017 Complexity-theoretic foundations of quantum supremacy experiments *Proc. 32nd Computational Complexity Conf. (CCC'17) (Schloss Dagstuhl–Leibniz-Zentrum Fuer Informatik, Dagstuhl (DEU))*
- [2] Harrow A W and Montanaro A 2017 Quantum computational supremacy *Nature* **549** 203
- [3] Eisert J, Hangleiter D, Walk N, Roth I, Markham D, Parekh R, Chabaud U and Kashefi E 2020 Quantum certification and benchmarking *Nat. Rev. Phys.* **2** 382
- [4] Carrasco J, Elben A, Kokail C, Kraus B and Zoller P 2021 Theoretical and experimental perspectives of quantum verification *PRX Quantum* **2** 010102
- [5] Vogel K and Risken H 1989 Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase *Phys. Rev. A* **40** 2847
- [6] Hradil Z 1997 Quantum-state estimation *Phys. Rev. A* **55** R1561
- [7] M Paris and J Řeháček eds 2004 *Quantum State Estimation* (Springer)
- [8] Smithey D T, Beck M, Raymer M G and Faridani A 1993 Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum *Phys. Rev. Lett.* **70** 1244
- [9] James D F V, Kwiat P G, Munro W J and White A G 2001 Measurement of qubits *Phys. Rev. A* **64** 052312
- [10] Janicke U and Wilkens M 1995 Tomography of atom beams *J. Mod. Opt.* **42** 2183
- [11] Kurtsiefer C, Pfau T and Mlynek J 1997 Measurement of the Wigner function of an ensemble of helium atoms *Nature* **386** 150
- [12] Klose G, Smith G and Jessen P S 2001 Measuring the quantum state of a large angular momentum *Phys. Rev. Lett.* **86** 4721
- [13] Poyatos J F, Walser R, Cirac J I, Zoller P and Blatt R 1996 Motion tomography of a single trapped ion *Phys. Rev. A* **53** R1966
- [14] Leibfried D, Meekhof D M, King B E, Monroe C, Itano W M and Wineland D J 1996 Experimental determination of the motional quantum state of a trapped atom *Phys. Rev. Lett.* **77** 4281
- [15] Roos C F, Lancaster G P T, Riebe M, Häffner H, Hänsel W, Gulde S, Becher C, Eschner J, Schmidt-Kaler F and Blatt R 2004 Bell states of atoms with ultralong lifetimes and their tomographic state analysis *Phys. Rev. Lett.* **92** 220402
- [16] Chuang I L, Gershenfeld N, Kubinec M G and Leung D W 1998 Bulk quantum computation with nuclear magnetic resonance: theory and experiment *Proc. R. Soc. A* **454** 447
- [17] Chuang I L, Gershenfeld N and Kubinec M 1998 Experimental implementation of fast quantum searching *Phys. Rev. Lett.* **80** 3408
- [18] Steffen M, Ansmann M, Bialczak R C, Katz N, Lucero E, McDermott R, Neeley M, Weig E M, Cleland A N and Martinis J M 2006 Measurement of the entanglement of two superconducting qubits via state tomography *Science* **313** 1423
- [19] Filipp S *et al* 2009 Two-qubit state tomography using a joint dispersive readout *Phys. Rev. Lett.* **102** 200402
- [20] D'Ariano G M, Maccone L and Painsi M 2003 Spin tomography *J. Opt. B: Quantum Semiclass. Opt.* **5** 77
- [21] Tóth G, Wieczorek W, Gross D, Krischek R, Schwemmer C and Weinfurter H 2010 Permutationally invariant quantum tomography *Phys. Rev. Lett.* **105** 250403
- [22] Moroder T, Hyllus P, Tóth G, Schwemmer C, Niggebaum A, Gaile S, Gühne O and Weinfurter H 2012 Permutationally invariant state reconstruction *New J. Phys.* **14** 105001
- [23] Cramer M, Plenio M B, Flammia S T, Somma R, Gross D, Bartlett S D, Landon-Cardinal O, Poulin D and Liu Y-K 2010 Efficient quantum state tomography *Nat. Commun.* **1** 149
- [24] Baumgratz T, Gross D, Cramer M and Plenio M B 2013 Scalable reconstruction of density matrices *Phys. Rev. Lett.* **111** 020401
- [25] Lanyon B P *et al* 2017 Efficient tomography of a quantum many-body system *Nat. Phys.* **13** 1158
- [26] Donoho D 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289
- [27] Candès E J, Romberg J K and Tao T 2006 Stable signal recovery from incomplete and inaccurate measurements *Commun. Pure Appl. Math.* **59** 1207
- [28] Gross D, Liu Y-K, Flammia S T, Becker S and Eisert J 2010 Quantum state tomography via compressed sensing *Phys. Rev. Lett.* **105** 150401
- [29] Flammia S T, Gross D, Liu Y-K and Eisert J 2012 Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators *New J. Phys.* **14** 095022
- [30] Kueng R, Rauhut H and Terstiege U 2017 Low rank matrix recovery from rank one measurements *Appl. Comput. Harmon. Anal.* **42** 88
- [31] Haah J, Harrow A W, Ji Z, Wu X and Yu N 2017 Sample-optimal tomography of quantum states *IEEE Trans. Inf. Theory* **63** 5628
- [32] Note that a better scaling $n \propto 2^2 N / \varepsilon^2$ can be achieved if one has access to the n -copy register $\rho^{\otimes n}$ and can perform collective entangled measurements [31, 164]. In the case of rank- r states, the scaling with such collective measurements becomes $O(r2^N / \varepsilon^2)$ with additional logarithmic corrections [31].
- [33] Schwemmer C, Tóth G, Niggebaum A, Moroder T, Gross D, Gühne O and Weinfurter H 2014 Experimental comparison of efficient tomography schemes for a six-qubit state *Phys. Rev. Lett.* **113** 040503
- [34] Riofrío C A, Gross D, Flammia S T, Monz T, Nigg D, Blatt R and Eisert J 2017 Experimental quantum compressed sensing for a seven-qubit system *Nat. Commun.* **8** 15305
- [35] Chuang I L and Nielsen M A 1997 Prescription for experimental determination of the dynamics of a quantum black box *J. Mod. Opt.* **44** 2455
- [36] Poyatos J F, Cirac J I and Zoller P 1997 Complete characterization of a quantum process: the two-bit quantum gate *Phys. Rev. Lett.* **78** 390
- [37] Fiurášek J and Hradil Z *c v* 2001 Maximum-likelihood estimation of quantum processes *Phys. Rev. A* **63** 020101
- [38] Sacchi M F 2001 Maximum-likelihood reconstruction of completely positive maps *Phys. Rev. A* **63** 054104
- [39] Ježek M, Fiurášek J and Hradil Z *c v* 2003 Quantum inference of states and processes *Phys. Rev. A* **68** 012305

- [40] Childs A M, Chuang I L and Leung D W 2001 Realization of quantum process tomography in NMR *Phys. Rev. A* **64** 012314
- [41] Weinstein Y S, Havel T F, Emerson J, Boulant N, Saraceno M, Lloyd S and Cory D G 2004 Quantum process tomography of the quantum Fourier transform *J. Chem. Phys.* **121** 6117
- [42] Mitchell M W, Ellenor C W, Schneider S and Steinberg A M 2003 Diagnosis, prescription and prognosis of a Bell-state filter by quantum process tomography *Phys. Rev. Lett.* **91** 120402
- [43] O'Brien J L, Pryde G J, Gilchrist A, James D F V, Langford N K, Ralph T C and White A G 2004 Quantum process tomography of a controlled-not gate *Phys. Rev. Lett.* **93** 080502
- [44] Riebe M, Kim K, Schindler P, Monz T, Schmidt P O, Körber T K, Hänsel W, Häffner H, Roos C F and Blatt R 2006 Process tomography of ion trap quantum gates *Phys. Rev. Lett.* **97** 220407
- [45] Home J P, Hanneke D, Jost J D, Amini J M, Leibfried D and Wineland D J 2009 Complete methods set for scalable ion trap quantum information processing *Science* **325** 1227
- [46] Monz T, Kim K, Hänsel W, Riebe M, Villar A S, Schindler P, Chwalla M, Hennrich M and Blatt R 2009 Realization of the quantum Toffoli gate with trapped ions *Phys. Rev. Lett.* **102** 040501
- [47] Bialczak R C *et al* 2010 Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits *Nat. Phys.* **6** 409
- [48] Poletto S *et al* 2012 Entanglement of two superconducting qubits in a waveguide cavity via monochromatic two-photon excitation *Phys. Rev. Lett.* **109** 240505
- [49] Flammia S T and Wallman J J 2020 Efficient estimation of Pauli channels *ACM Trans. Quantum Comput.* **1** 1–32
- [50] Fawzi O, Oufkir A and França D S 2023 Lower bounds on learning Pauli channels (arXiv:2301.09192 [quant-ph])
- [51] Kliesch M, Kueng R, Eisert J and Gross D 2019 Guaranteed recovery of quantum processes from few measurements *Quantum* **3** 171
- [52] Shabani A, Kosut R L, Mohseni M, Rabitz H, Broome M A, Almeida M P, Fedrizzi A and White A G 2011 Efficient measurement of quantum dynamics via compressive sensing *Phys. Rev. Lett.* **106** 100401
- [53] Rodionov A V, Veitia A, Barends R, Kelly J, Sank D, Wenner J, Martinis J M, Kosut R L and Korotkov A N 2014 Compressed sensing quantum process tomography for superconducting quantum gates *Phys. Rev. B* **90** 144504
- [54] Schirmer S G, Kolli A and Oi D K L 2004 Experimental Hamiltonian identification for controlled two-level systems *Phys. Rev. A* **69** 050306
- [55] Cole J H, Schirmer S G, Greentree A D, Wellard C J, Oi D K L and Hollenberg L C L 2005 Identifying an experimental two-state Hamiltonian to arbitrary accuracy *Phys. Rev. A* **71** 062312
- [56] Devitt S J, Cole J H and Hollenberg L C L 2006 Scheme for direct measurement of a general two-qubit Hamiltonian *Phys. Rev. A* **73** 052317
- [57] de Clercq L E, Oswald R, Flühmann C, Keitch B, Kienzler D, Lo H-Y, Marinelli M, Nadlinger D, Negnevitsky V and Home J P 2016 Estimation of a general time-dependent Hamiltonian for a single qubit *Nat. Commun.* **7** 11218
- [58] Di Franco C, Paternostro M and Kim M S 2009 Hamiltonian tomography in an access-limited setting without state initialization *Phys. Rev. Lett.* **102** 187203
- [59] Burgarth D and Maruyama K 2009 Indirect Hamiltonian identification through a small gateway *New J. Phys.* **11** 103019
- [60] Zhang J and Sarovar M 2014 Quantum Hamiltonian identification from measurement time traces *Phys. Rev. Lett.* **113** 080401
- [61] Holzäpfel M, Baumgratz T, Cramer M and Plenio M B 2015 Scalable reconstruction of unitary processes and Hamiltonians *Phys. Rev. A* **91** 042129
- [62] da Silva M P, Landon-Cardinal O and Poulin D 2011 Practical characterization of quantum devices without tomography *Phys. Rev. Lett.* **107** 210404
- [63] Chertkov E and Clark B K 2018 Computational inverse method for constructing spaces of quantum models from wave functions *Phys. Rev. X* **8** 031029
- [64] Qi X-L and Ranard D 2019 Determining a local Hamiltonian from a single eigenstate *Quantum* **3** 159
- [65] Bairey E, Arad I and Lindner N H 2019 Learning a local Hamiltonian from local measurements *Phys. Rev. Lett.* **122** 020504
- [66] Li Z, Zou L and Hsieh T H 2020 Hamiltonian tomography via quantum quench *Phys. Rev. Lett.* **124** 160502
- [67] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge University Press)
- [68] Watrous J 2018 *The Theory of Quantum Information* (Cambridge University Press)
- [69] Rivas A, Huelga S F and Plenio M B 2010 Entanglement and non-Markovianity of quantum evolutions *Phys. Rev. Lett.* **105** 050403
- [70] Ángel Rivas S F H and Plenio M B 2014 Quantum non-Markovianity: characterization, quantification and detection *Rep. Prog. Phys.* **77** 094001
- [71] Wolf M M, Eisert J, Cubitt T S and Cirac J I 2008 Assessing non-Markovian quantum dynamics *Phys. Rev. Lett.* **101** 150402
- [72] Lindblad G 1976 On the generators of quantum dynamical semigroups *Commun. Math. Phys.* **48** 119
- [73] Gorini V, Kossakowski A and Sudarshan E C G 1976 Completely positive dynamical semigroups of n -level systems *J. Math. Phys.* **17** 821
- [74] Breuer H P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford University Press)
- [75] Boulant N, Havel T F, Pravia M A and Cory D G 2003 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
- [76] Howard M, Twamley J, Wittmann C, Gaebel T, Jelezko F and Wrachtrup J 2006 Quantum process tomography and Lindblad estimation of a solid-state qubit *New J. Phys.* **8** 33
- [77] Onorati E, Kohler T and Cubitt T 2021 Fitting quantum noise models to tomography data (arXiv:2103.17243 [quant-ph])
- [78] Bužek V 1998 Reconstruction of Liouvillian superoperators *Phys. Rev. A* **58** 1723
- [79] Bairey E, Guo C, Poletti D, Lindner N H and Arad I 2020 Learning the dynamics of open quantum systems from their steady states *New J. Phys.* **22** 032001
- [80] Pastori L, Olsacher T, Kokail C and Zoller P 2022 Characterization and verification of Trotterized digital quantum simulation via Hamiltonian and Liouvillian learning *PRX Quantum* **3** 030324
- [81] França D S, Markovich L A, Dobrovitski V V, Werner A H and Borregaard J 2022 Efficient and robust estimation of many-qubit Hamiltonians (arXiv:2205.09567 [quant-ph])
- [82] Samach G O *et al* 2022 Lindblad tomography of a superconducting quantum processor *Phys. Rev. Appl.* **18** 064056
- [83] Av Ben E, Shapira Y, Akerman N and Ozeri R 2020 Direct reconstruction of the quantum-master-equation dynamics of a trapped-ion qubit *Phys. Rev. A* **101** 062305
- [84] Zhang H, Pokharel B, Levenson-Falk E and Lidar D 2022 Predicting non-Markovian superconducting-qubit dynamics from tomographic reconstruction *Phys. Rev. Appl.* **17** 054018

- [85] Nakajima S 1958 On quantum theory of transport phenomena: steady diffusion *Prog. Theor. Phys.* **20** 948
- [86] Zwanziġ R W 1960 Ensemble method in the theory of irreversibility *J. Chem. Phys.* **33** 1338
- [87] Varona S, Müller M and Bermudez A 2025 Lindblad-like quantum tomography for non-Markovian dynamical dephasing maps *Npj Quantum Inf.* **11** 96 (available at: www.nature.com/articles/s41534-025-01044-7)
- [88] Reháček J, Hradil Z, Knill E and Lvovsky A I 2007 Diluted maximum-likelihood algorithm for quantum tomography *Phys. Rev. A* **75** 042108
- [89] Bochkhanov S Alglib (199–2024) (available at: www.alglib.net)
- [90] Korotkov A N 2013 Error matrices in quantum process tomography (arXiv:1309.6405)
- [91] Bolduc E, Knee G C, Gauger E M and Leach J 2017 Projected gradient descent algorithms for quantum state tomography *npj Quantum Inf.* **3** 44
- [92] Mezzadri F 2007 How to generate random matrices from the classical compact groups (arXiv:math-ph/0609050 [math-ph])
- [93] Życzkowski K and Sommers H-J 2001 Induced measures in the space of mixed quantum states *J. Phys. A: Math. Gen.* **34** 7111
- [94] Kosut R L 2009 Quantum process tomography via l1-norm minimization (arXiv:0812.4323 [quant-ph])
- [95] Baldwin C H, Kalev A and Deutsch I H 2014 Quantum process tomography of unitary and near-unitary maps *Phys. Rev. A* **90** 012110
- [96] Kosut R L 2008 Quantum process tomography via l1-norm minimization (arXiv:0812.4323)
- [97] Bochkhanov S Alglib (available at: <http://alglib.net>)
- [98] Cirac J I and Zoller P 1995 Quantum computations with cold trapped ions *Phys. Rev. Lett.* **74** 4091
- [99] Schmidt-Kaler F *et al* 2003 Realization of the Cirac–Zoller controlled-not quantum gate *Nature* **422** 408
- [100] Häffner H, Roos C and Blatt R 2008 Quantum computing with trapped ions *Phys. Rep.* **469** 155
- [101] Bruzewicz C D, Chiaverini J, McConnell R and Sage J M 2019 Trapped-ion quantum computing: progress and challenges *Appl. Phys. Rev.* **6** 021314
- [102] Gulde S, Riebe M, Lancaster G P T, Becher C, Eschner J, Häffner H, Schmidt-Kaler F, Chuang I L and Blatt R 2003 Implementation of the Deutsch–Jozsa algorithm on an ion-trap quantum computer *Nature* **421** 48
- [103] Barrett M D *et al* 2004 Deterministic quantum teleportation of atomic qubits *Nature* **429** 737
- [104] Riebe M *et al* 2004 Deterministic quantum teleportation with atoms *Nature* **429** 734
- [105] Chiaverini J *et al* 2005 Implementation of the semiclassical quantum Fourier transform in a scalable system *Science* **308** 997
- [106] Monz T, Nigg D, Martinez E A, Brandl M F, Schindler P, Rines R, Wang S X, Chuang I L and Blatt R 2016 Realization of a scalable Shor algorithm *Science* **351** 1068
- [107] Figgatt C, Maslov D, Landsman K A, Linke N M, Debnath S and Monroe C 2017 Complete 3-qubit Grover search on a programmable quantum computer *Nat. Commun.* **8** 1918
- [108] Wan Y *et al* 2019 Quantum gate teleportation between separated qubits in a trapped-ion processor *Science* **364** 875
- [109] Chiaverini J *et al* 2004 Realization of quantum error correction *Nature* **432** 602
- [110] Schindler P, Barreiro J T, Monz T, Nebendahl V, Nigg D, Chwalla M, Hennrich M and Blatt R 2011 Experimental repetitive quantum error correction *Science* **332** 1059
- [111] Nigg D, Müller M, Martinez E A, Schindler P, Hennrich M, Monz T, Martin-Delgado M A and Blatt R 2014 Quantum computations on a topologically encoded qubit *Science* **345** 302
- [112] Linke N M, Gutierrez M, Landsman K A, Figgatt C, Debnath S, Brown K R and Monroe C 2017 Fault-tolerant quantum error detection *Sci. Adv.* **3** 10
- [113] Negnevitsky V, Marinelli M, Mehta K K, Lo H-Y, Flühmann C and Home J P 2018 Repeated multi-qubit readout and feedback with a mixed-species trapped-ion register *Nature* **563** 527
- [114] Flühmann C, Nguyen T L, Marinelli M, Negnevitsky V, Mehta K and Home J P 2019 Encoding a qubit in a trapped-ion mechanical oscillator *Nature* **566** 513
- [115] Stricker R *et al* 2020 Experimental deterministic correction of qubit loss *Nature* **585** 207
- [116] de Neeve B, Nguyen T L, Behrle T and Home J 2020 Error correction of a logical grid state qubit by dissipative pumping (arXiv:2010.09681)
- [117] Erhard A *et al* 2021 Entangling logical qubits with lattice surgery *Nature* **589** 220
- [118] Egan L *et al* 2021 Fault-tolerant control of an error-corrected qubit *Nature* **598** 281
- [119] Debroy D M, Egan L, Noel C, Risinger A, Zhu D, Biswas D, Cetina M, Monroe C and Brown K R 2021 Optimizing stabilizer parities for improved logical qubit memories *Phys. Rev. Lett.* **127** 240501
- [120] Ryan-Anderson C *et al* 2021 Realization of real-time fault-tolerant quantum error correction *Phys. Rev. X* **11** 041058
- [121] Hilder J, Pijn D, Onishchenko O, Stahl A, Orth M, Lekitsch B, Rodriguez-Blanco A, Müller M, Schmidt-Kaler F and Poschinger U G 2022 Fault-tolerant parity readout on a shuttling-based trapped-ion quantum computer *Phys. Rev. X* **12** 011032
- [122] Postler L *et al* 2022 Demonstration of fault-tolerant universal quantum gate operations *Nature* **605** 675
- [123] Ryan-Anderson C *et al* 2022 Implementing fault-tolerant entangling gates on the five-qubit code and the color code (arXiv:2208.01863)
- [124] Bermudez A *et al* 2017 Assessing the progress of trapped-ion processors towards fault-tolerant quantum computation *Phys. Rev. X* **7** 041061
- [125] Gutiérrez M, Müller M and Bermúdez A 2019 Transversality and lattice surgery: exploring realistic routes toward coupled logical qubits with trapped-ion quantum processors *Phys. Rev. A* **99** 022330
- [126] Bermudez A, Xu X, Gutiérrez M, Benjamin S C and Müller M 2019 Fault-tolerant protection of near-term trapped-ion topological qubits under realistic noise sources *Phys. Rev. A* **100** 062307
- [127] Parrado-Rodríguez P, Ryan-Anderson C, Bermudez A and Müller M 2021 Crosstalk suppression for fault-tolerant quantum error correction with trapped ions *Quantum* **5** 487
- [128] Rodriguez-Blanco A, Shahandeh F and Bermudez A 2022 Witnessing entanglement in trapped-ion quantum error correction under realistic noise (arXiv:2212.07479 [quant-ph])
- [129] Harty D T C *et al* 2014 High-fidelity preparation, gates, memory and readout of a trapped-ion quantum bit *Phys. Rev. Lett.* **113** 220501
- [130] Ballance C J, Harty T P, Linke N M, Sepiol M A and Lucas D M 2016 High-fidelity quantum logic gates using trapped-ion hyperfine qubits *Phys. Rev. Lett.* **117** 060504
- [131] Gaebler J P *et al* 2016 High-fidelity universal gate set for $^9\text{Be}^+$ ion qubits *Phys. Rev. Lett.* **117** 060505
- [132] Harty T P, Sepiol M A, Allcock D T C, Ballance C J, Tarlton J E and Lucas D M 2016 High-fidelity trapped-ion quantum logic using near-field microwaves *Phys. Rev. Lett.* **117** 140501

- [133] Erhard A, Wallman J J, Postler L, Meth M, Stricker R, Martinez E A, Schindler P, Monz T, Emerson J and Blatt R 2019 Characterizing large-scale quantum computers via cycle benchmarking *Nat. Commun.* **10** 5347
- [134] Zarantonello G, Hahn H, Morgner J, Schulte M, Bautista-Salvador A, Werner R F, Hammerer K and Ospelkaus C 2019 Robust and resource-efficient microwave near-field entangling $^9\text{Be}^+$ gate *Phys. Rev. Lett.* **123** 260503
- [135] Häffner H et al 2005 Scalable multiparticle entanglement of trapped ions *Nature* **438** 643
- [136] Stricker R, Meth M, Postler L, Edmunds C, Ferrie C, Blatt R, Schindler P, Monz T, Kueng R and Ringbauer M 2022 Experimental single-setting quantum state tomography *PRX Quantum* **3** 040310
- [137] Riebe M, Chwalla M, Benhelm J, Häffner H, Hänsel W, Roos C F and Blatt R 2007 Quantum teleportation with atoms: quantum process tomography *New J. Phys.* **9** 211
- [138] Monz T et al 2009 Realization of universal ion-trap quantum computation with decoherence-free qubits *Phys. Rev. Lett.* **103** 200503
- [139] Wang S X, Labaziewicz J, Ge Y, Shewmon R and Chuang I L 2010 Demonstration of a quantum logic gate in a cryogenic surface-electrode ion trap *Phys. Rev. A* **81** 062332
- [140] Navon N, Akerman N, Kotler S, Glickman Y and Ozeri R 2014 Quantum process tomography of a Mølmer-Sørensen interaction *Phys. Rev. A* **90** 010103
- [141] Tinkey H N, Meier A M, Clark C R, Seck C M and Brown K R 2021 Quantum process tomography of a Mølmer-Sørensen gate via a global beam *Quantum Sci. Technol.* **6** 034013
- [142] Emerson J, Alicki R and Życzkowski K 2005 Scalable noise estimation with random unitary operators *J. Opt. B: Quantum Semiclass. Opt.* **7** S347
- [143] Magesan E, Gambetta J M and Emerson J 2011 Scalable and robust randomized benchmarking of quantum processes *Phys. Rev. Lett.* **106** 180504
- [144] Wallman J J and Emerson J 2016 Noise tailoring for scalable quantum computation via randomized compiling *Phys. Rev. A* **94** 052325
- [145] Blume-Kohout R, Gamble J K, Nielsen E, Mizrahi J, Sterk J D and Maunz P 2013 Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit (arXiv:1310.4492 [quant-ph])
- [146] Knill E, Leibfried D, Reichle R, Britton J, Blakestad R B, Jost J D, Langer C, Ozeri R, Seidelin S and Wineland D J 2008 Randomized benchmarking of quantum gates *Phys. Rev. A* **77** 012307
- [147] Brown K R, Wilson A C, Colombe Y, Ospelkaus C, Meier A M, Knill E, Leibfried D and Wineland D J 2011 Single-qubit-gate error below 10^{-4} in a trapped ion *Phys. Rev. A* **84** 030303
- [148] Gaebler J P et al 2012 Randomized benchmarking of multiqubit gates *Phys. Rev. Lett.* **108** 260503
- [149] Blume-Kohout R, Gamble J K, Nielsen E, Rudinger K, Mizrahi J, Fortier K and Maunz P 2017 Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography *Nat. Commun.* **8** 14485
- [150] Mavadia S, Edmunds C L, Hempel C, Ball H, Roy F, Stace T M and Biercuk M J 2018 Experimental quantum verification in the presence of temporally correlated noise *npj Quantum Inf.* **4** 7
- [151] Wright K et al 2019 Benchmarking an 11-qubit quantum computer *Nat. Commun.* **10** 5464
- [152] Pogorelov I et al 2021 Compact ion-trap quantum computing demonstrator *PRX Quantum* **2** 020343
- [153] Schindler P, Müller M, Nigg D, Barreiro J T, Martinez E A, Hennrich M, Monz T, Diehl S, Zoller P and Blatt R 2013 Quantum simulation of dynamical maps with trapped ions *Nat. Phys.* **9** 361
- [154] Schindler P et al 2013 A quantum information processor with trapped ions *New J. Phys.* **15** 123012
- [155] Velázquez J M S, Steiner A, Freund R, Guevara-Bertsch M, Marciniak C D, Monz T and Bermudez A 2024 Dynamical quantum maps for single-qubit gates under non-Markovian phase noise (arXiv:2402.14530 [quant-ph])
- [156] Taylor J 1982 *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (University Science Books)
- [157] Sørensen A and Mølmer K 1999 Quantum computation with ions in thermal motion *Phys. Rev. Lett.* **82** 1971
- [158] Sørensen A and Mølmer K 2000 Entanglement and quantum computation with ions in thermal motion *Phys. Rev. A* **62** 022311
- [159] Cai Z, Luan C Y, Ou L, Tu H, Yin Z, Zhang J N and Kim K 2023 Entangling gates for trapped-ion quantum computation and quantum simulation *J. Korean Phys. Soc.* **82** 882
- [160] Teo Y S and Sánchez-Soto L L 2021 Modern compressive tomography for quantum information science *Int. J. Quantum Inf.* **19** 2140003
- [161] Flammia S T, Silberfarb A and Caves C M 2005 Minimal informationally complete measurements for pure states *Found. Phys.* **35** 1985
- [162] Itano W M, Bergquist J C, Bollinger J J, Gilligan J M, Heinzen D J, Moore F L, Raizen M G and Wineland D J 1993 Quantum projection noise: population fluctuations in two-level systems *Phys. Rev. A* **47** 3554
- [163] Chen S, Huang B, Li J, Liu A and Sellke M 2022 Tight bounds for state tomography with incoherent measurements (arXiv:2206.05265 [quant-ph])
- [164] O'Donnell R and Wright J 2016 Efficient quantum tomography *Proc. 48th Annual ACM Symp. on Theory of Computing (STOC '16)* (Association for Computing Machinery) pp 899–912
- [165] Angus L 2021 Learning quantum states without entangled measurements *Master's Thesis* (available at: <http://hdl.handle.net/10012/17663>)
- [166] Dennis E, Kitaev A, Landahl A and Preskill J 2002 Topological quantum memory *J. Math. Phys.* **43** 4452
- [167] Raussendorf R, Harrington J and Goyal K 2007 Topological fault-tolerance in cluster state quantum computation *New J. Phys.* **9** 199
- [168] Fowler A G, Stephens A M and Groszkowski P 2009 High-threshold universal quantum computation on the surface code *Phys. Rev. A* **80** 052312
- [169] Landahl A J, Anderson J T and Rice P R 2011 Fault-tolerant quantum computing with color codes (arXiv:1108.5738 [quant-ph])
- [170] Katzgraber H G, Bombin H and Martin-Delgado M A 2009 Error threshold for color codes and random three-body Ising models *Phys. Rev. Lett.* **103** 090501
- [171] Mohseni M, Rezakhani A T and Lidar D A 2008 Quantum-process tomography: Resource analysis of different strategies *Phys. Rev. A* **77** 032322
- [172] Harper R, Flammia S T and Wallman J J 2020 Efficient learning of quantum noise *Nat. Phys.* **16** 1184
- [173] Chen S, Zhou S, Seif A and Jiang L 2022 Quantum advantages for Pauli channel estimation *Phys. Rev. A* **105** 032435
- [174] Hellwig K E and Kraus K 1970 Operations and measurements. II *Commun. Math. Phys.* **16** 142
- [175] Kraus K 1971 General state changes in quantum theory *Ann. Phys., NY* **64** 311
- [176] Evans T J, Harper R and Flammia S T 2019 Scalable Bayesian Hamiltonian learning (arXiv:1912.07636 [quant-ph])

- [177] Dewes A, Ong F R, Schmitt V, Lauro R, Boulant N, Bertet P, Vion D and Esteve D 2012 Characterization of a two-transmon processor with individual single-shot qubit readout *Phys. Rev. Lett.* **108** 057002
- [178] Shewchuk J R 1994 An introduction to the conjugate gradient method without the agonizing pain *Technical Report*
- [179] Teo Y S 2013 Numerical estimation schemes for quantum tomography (arXiv:[1302.3399](#))
- [180] Gonçalves D, Gomes-Ruggiero M and Lavor C 2016 A projected gradient method for optimization over density matrices *Optim. Methods Softw.* **31** 328
- [181] Shang J, Zhang Z and Ng H K 2017 Superfast maximum-likelihood reconstruction for quantum tomography *Phys. Rev. A* **95** 062336
- [182] Nesterov Y 2004 *Introductory Lectures on Convex Optimization: A Basic Course (Applied Optimization)* (Springer)
- [183] O'Donoghue B and Candes E 2012 Adaptive restart for accelerated gradient schemes (arXiv:[1204.3982](#))