

PAPER • OPEN ACCESS

Hamiltonian and Liouvillian learning in weakly-dissipative quantum many-body systems

To cite this article: Tobias Olsacher *et al* 2025 *Quantum Sci. Technol.* **10** 015065

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

You may also like

- [An operational definition of quantum information scrambling](#)
Gabriele Lo Monaco, Luca Innocenti, Dario Cilluffo et al.
- [Pseudomode treatment of strong-coupling quantum thermodynamics](#)
Francesco Albarelli, Bassano Vacchini and Andrea Smirne
- [Quantum state tomography based on infidelity estimation](#)
Yong Wang, Lijun Liu, Tong Dou et al.

**ECONOMIST
IMPACT**

4th annual Commercialising Quantum Global

May 13th-14th 2025 • London

1,000 attendees
100 speakers
60 content rich sessions
30 case studies

Register

Quantum Science and Technology



PAPER

OPEN ACCESS

RECEIVED
4 November 2024

REVISED
5 December 2024

ACCEPTED FOR PUBLICATION
13 December 2024

PUBLISHED
8 January 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.



Hamiltonian and Liouvillian learning in weakly-dissipative quantum many-body systems

Tobias Olsacher^{1,2,7} , Tristan Kraft^{1,3,4,7,*} , Christian Kokail^{1,2,5,6}, Barbara Kraus^{1,3,4} and Peter Zoller^{1,2}

¹ Institute for Theoretical Physics, University of Innsbruck, Technikerstraße 21A, 6020 Innsbruck, Austria

² Institute for Quantum Optics and Quantum Information of the Austrian Academy of Sciences, Innsbruck, Austria

³ TUM School of Natural Sciences, Technical University of Munich, James-Frank-Str. 1, D-85748 Garching, Germany

⁴ Munich Center for Quantum Science and Technology (MCQST), Schellingstraße 4, D-80799 München, Germany

⁵ ITAMP, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, United States of America

⁶ Department of Physics, Harvard University, Cambridge, MA 02138, United States of America

⁷ These two authors contributed equally.

* Author to whom any correspondence should be addressed.

E-mail: tristan.kraft@tum.de

Keywords: quantum simulation, Hamiltonian learning, Liouvillian learning, verification

Abstract

We discuss Hamiltonian and Liouvillian learning for analog quantum simulation from non-equilibrium quench dynamics in the limit of weakly dissipative many-body systems. We present and compare various methods and strategies to learn the operator content of the Hamiltonian and the Lindblad operators of the Liouvillian. We compare different ansätze based on an experimentally accessible ‘learning error’ which we consider as a function of the number of runs of the experiment. Initially, the learning error decreases with the inverse square root of the number of runs, as the error in the reconstructed parameters is dominated by shot noise. Eventually the learning error remains constant, allowing us to recognize missing ansatz terms. A central aspect of our approaches is to (re-)parametrize ansätze by introducing and varying the dependencies between parameters. This allows us to identify the relevant parameters of the system, thereby reducing the complexity of the learning task. Importantly, this (re-)parametrization relies solely on classical post-processing, which is compelling given the finite amount of data available from experiments. We illustrate and compare our methods with two experimentally relevant spin models.

1. Introduction

Controllable quantum many-body systems, when scaled to a large number of particles, hold the potential to function as quantum computers or quantum simulators, addressing computational problems that are considered intractable for classical computers [1]. Remarkable progress has been reported recently in building quantum simulators, as programmable special-purpose quantum devices, to solve quantum many-body problems efficiently, which finds applications in condensed matter [2], high-energy physics [3], and quantum chemistry [4], in both equilibrium and non-equilibrium dynamics. Quantum simulation can be realized as analog or digital quantum simulators. In analog simulation, a target Hamiltonian finds a natural implementation on a quantum device, exemplified by ultracold bosonic and fermionic atoms in optical lattices as Hubbard models [5–8], or spin models with trapped ions [9–12], Rydberg tweezer arrays [13–16], and superconducting qubits [17–19]. The unique feature of analog quantum simulators is the scalability to large particle numbers. In contrast, digital quantum simulation [20] represents the time evolution of a given many-body Hamiltonian using a freely programmable sequence of Trotter steps implemented via single and multi-qubit entangling quantum gates.

An outstanding challenge in quantum simulation is the ability to predict properties of many-body observables with controlled error while scaling to a regime of potential quantum advantage [21–24]. Given the increase of complexity of these systems, methods to characterize, and thus verify, the proper functioning

of quantum simulators are required [25–27]. This includes verification, that the correct many-body Hamiltonians are being implemented, and a complete characterization of (weak) decoherence due to unwanted couplings to an environment or fluctuating external fields. In the present paper, we approach this goal by studying Hamiltonian and Liouvillian learning for analog quantum simulators. Previous works have studied various scenarios for Hamiltonian and Liouvillian learning [28–40], for instance, by comparing with a trusted simulator [28], or based on the preparation of steady states [31–35]. Alternative approaches are based on dynamics in (long-time) quenches [29, 34–37], the estimation of the time-derivatives of few-qubit observables from short-time evolution [38, 39], or require additional resources such as intermediate gates [41, 42].

Here, we will be interested in Hamiltonian and Liouvillian learning from dynamics in long-time quenches, with only few experimental requirements such as the preparation of product states, and product measurements. We assume that the dynamics of the experimental quantum simulator is described by a master equation, where the Hamiltonian acts as a generator of the coherent many-body dynamics, while Lindbladian terms model the noise. The goal of Hamiltonian and Liouvillian learning is to learn the *operator structure*, reminiscent of a *principal component analysis* [43], of both the many-body Hamiltonian, as one-, two- or few-body interaction terms including their couplings, and the quantum jump operators in the dissipative Liouvillian, representing local or non-local (global) quantum and classical noise. The scalability and efficiency of Hamiltonian and Liouvillian learning are related to the assumption that physical Hamiltonians and Liouvillians will only involve few-body interactions and quantum jump operators, leading to a polynomial scaling of the number of terms to be learned with system size.

Our study below discusses and compares various scenarios and strategies of Hamiltonian and Liouvillian learning, which we illustrate by simulating learning protocols for various model cases. Our work is motivated by present trapped-ion experiments, where quantum simulators realize 1D interacting weakly dissipative spin-1/2 chains. This allows quench experiments to be performed, and we will be interested in learning the Hamiltonian and Liouvillian from experimental quench data observed at various quench times. Learning the Hamiltonian and Liouvillian requires many experimental runs. In each run projective measurements of spins are performed in various bases, allowing to measure multi-spin correlation functions up to shot noise. In addition, learning protocols will prepare many initial states, which in our case can be pure or mixed, thus resulting in stability against state-preparation errors. A central aspect of our study below will thus be an investigation of the experimentally measurable learning error of Hamiltonian and Liouvillian, and its scaling with the number of measurement runs.

The paper is structured as follows. Section 2 outlines the specific scenario we are examining and the core theoretical concepts involved. Section 2.1 establishes some constraints on the system’s Hamiltonian and Liouvillian, measurable through simple quench experiments. We then present the main equations that enable us to infer the Hamiltonian and Liouvillian from experimental data in sections 2.2 and 2.3. The first approach is based on the equation of motion of the expectation value of general observables, known as Ehrenfest’s theorem, while the second approach is based on generalized energy conservation. In section 3, we compare the two methods and showcase learning protocols for various model scenarios through numerical simulations. We compare different ansätze for the operator content, using an experimentally measurable quantity, which we identify as a learning error. These ansätze are derived by re-parametrizing our ansatz, typically involving data recycling and classical post-processing. The learning process can be divided into two phases as a function of experimental runs N_{runs} : in the early phase, the learning error is dominated by shot-noise and decays $\sim N_{\text{runs}}^{-1/2}$. In the later phase, systematic errors become dominant due to missing terms in the ansatz, or an insufficient ansatz. This leads to a constant learning error independent of the number of measurements, indicating the need to extend our ansatz.

2. Hamiltonian and Liouvillian learning of many-body systems

2.1. Background

We consider analog quantum simulation in a regime, where the engineered quantum many-body system of interest is weakly coupled to a decohering environment. We assume that the system dynamics is described by a master equation with Lindblad form [44],

$$\frac{d}{dt}\varrho = -i[H, \varrho] + \frac{1}{2} \sum_k \gamma_k \left([l_k \varrho, l_k^\dagger] + [l_k, \varrho l_k^\dagger] \right) \equiv \mathcal{H}(\varrho) + \mathcal{L}(\varrho), \quad (1)$$

comprising a coherent term $\mathcal{H}(\varrho) = -i[H, \varrho]$ with many-body *Hamiltonian* H and a dissipative term \mathcal{L} , referred to as the *Liouvillian*. The Lindblad quantum jump operators l_k describe dissipative processes coupling the system to an environment. Here, $\gamma_k \geq 0$ is the physical domain for the corresponding damping

rates, where the dynamics is described by a completely positive and trace-preserving map [45]⁷. Throughout this work, we assume that the Hamiltonian and the Liouvillian are time-independent. In typical experimental settings, the Hamiltonian terms are significantly larger than dissipative processes $\|H\| \gg \gamma_k$.

Below we will be interested in analog quantum simulation of spin-1/2 systems, as implemented with trapped ions [9–12], Rydberg tweezer arrays [14–16], or superconducting circuits [17, 19]. An illustrative example of a two-local spin-Hamiltonian in one spatial dimension, which we have in mind, is given by

$$H = \sum_{i=1}^{N-1} J_{i,i+1}^z \sigma_i^z \sigma_{i+1}^z + \sum_{i=1}^{N-2} J_{i,i+2}^z \sigma_i^z \sigma_{i+2}^z + B_x \sum_{i=1}^N \sigma_i^x + B_z \sum_{i=1}^N \sigma_i^z, \quad (2)$$

which describes a next-nearest-neighbor Ising model of N spins with longitudinal- and transverse fields.

The Liouvillian, \mathcal{L} , in equation (1) is defined by its Lindblad, or quantum jump operators l_k . Examples of Lindblad operators that typically appear in experiments include spontaneous emission, described by local Lindblad operators $l_k = \sigma_k^-$, or local dephasing, represented by $l_k = \sigma_k^z$. Besides local dissipation, we will be interested in identifying the presence of collective dissipative effects, for instance, in the form of collective dephasing caused by globally fluctuating laser fields or effective magnetic fields, leading to a collective Lindblad operator $l = \sum_k \sigma_k^z$.

Before we proceed with the discussion, we need to establish certain conditions on the Hamiltonian and Liouvillian of the system. For a time-independent observable, O , Ehrenfest's theorem, in the context of equation (1), states that

$$\frac{d}{dt} \langle O \rangle = \langle -i[O, H] \rangle + \frac{1}{2} \sum_k \gamma_k \langle l_k^\dagger [O, l_k] + [l_k^\dagger, O] l_k \rangle, \quad (3)$$

where $\langle X \rangle = \text{tr}[X \varrho]$, for any state ϱ . In its integral form, this generalized Ehrenfest theorem yields the following condition for the Hamiltonian and Liouvillian

$$\langle O \rangle_T - \langle O \rangle_0 = -i \int_0^T \langle [O, H] \rangle dt + \frac{1}{2} \sum_k \gamma_k \int_0^T \langle l_k^\dagger [O, l_k] + [l_k^\dagger, O] l_k \rangle_t dt, \quad (4a)$$

where $\langle X \rangle_t = \text{tr}[X \exp\{t(\mathcal{H} + \mathcal{L})\} \varrho(0)]$. This condition simplifies if we consider $O = H$, which leads to

$$\langle H \rangle_T - \langle H \rangle_0 = \frac{1}{2} \sum_k \gamma_k \int_0^T \langle l_k^\dagger [H, l_k] + [l_k^\dagger, H] l_k \rangle_t dt. \quad (4b)$$

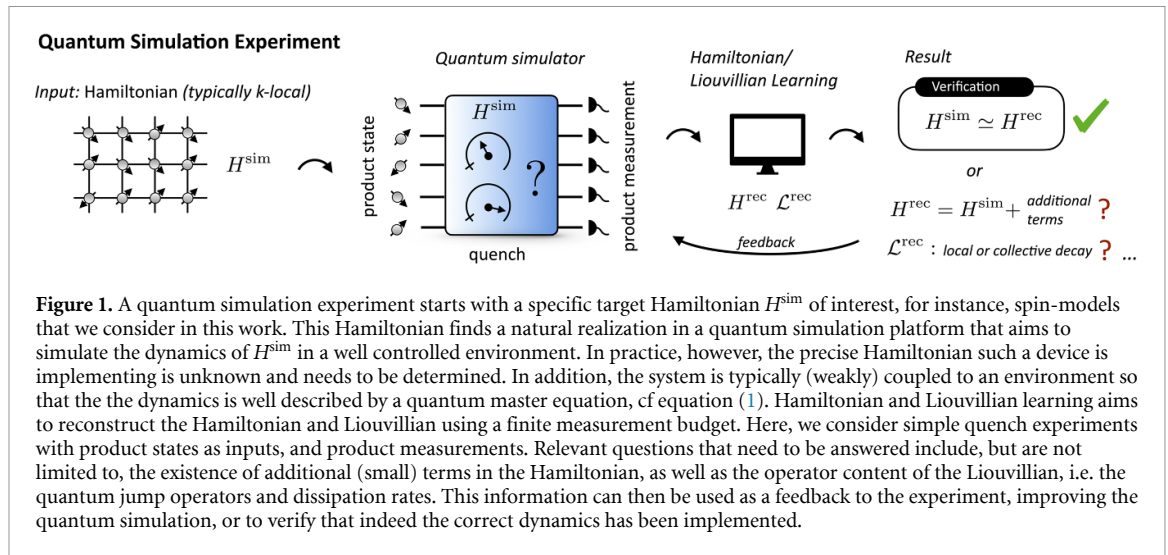
It describes generalized energy conservation, including the loss of total energy of the system during a quench of duration T . If $\gamma_k = 0$ for all k , this equation indicates the conservation of energy. In our Hamiltonian and Liouvillian learning protocol, equations (4a) and (4b) will play a crucial role. We will generalize the protocol of [36] for Hamiltonian learning in the absence of dissipation, and present it in a form particularly suited for learning H and \mathcal{L} in the limit of weak dissipation.

As a final remark, let us elaborate on the experimental procedure that we are considering which can be used to probe the conditions in equations (4a) and (4b) (see also figure 1 for an illustration). Starting from a product state, $|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_N\rangle$, which can be experimentally prepared with high fidelity, one evolves the state under the Hamiltonian and Liouvillian for some time t . The resulting state, $\varrho_t = \exp[t(\mathcal{H} + \mathcal{L})] \varrho_0$, is measured in a product basis, for instance, in the Pauli basis. Clearly, the limiting quantity, here, and in the following, will be the total number of runs of quench experiments, which we will denote by N_{runs} . Given, however, that many-body Hamiltonians typically consist of a few quasi-local operators, many of the required measurements can be carried out simultaneously in a single run and using classical post-processing, or via the randomized measurement toolbox [46].

2.2. Hamiltonian and Liouvillian learning from the generalized Ehrenfest theorem

Often, in an experimental setting, the detailed structure of the Hamiltonian and the Liouvillian are unknown. We present here a method to learn the operator content of H and \mathcal{L} , and the corresponding parameters from experimental data. For instance, in the context of the spin system in and below

⁷ The Hamiltonian, H , the dissipation rates, γ_k , and the Lindblad operators, l_k , are uniquely determined by the dynamics, if one requires the following: (i) H is traceless, (ii) the l_k are traceless and orthonormal, i.e. $\text{tr}(l_k) = 0$ and $\text{tr}(l_m^\dagger l_n) = \delta_{mn}$, and (iii) the γ_k are not degenerate [45].



equation (2), identifying the operator content means identifying the Pauli operators that appear in the decomposition of H and the Lindblad operators l_k .

We start by choosing an ansatz for the operator content for the Hamiltonian and Liouvillian. Specifically, as an ansatz for H we choose

$$A(\mathbf{c}) = \sum_{j=1}^n c_j h_j, \quad (5)$$

with parameters $\mathbf{c} = (c_1, \dots, c_n)$, and h_j traceless and hermitian for all j . As an example, one could choose the h_j to be few-body Pauli operators. As an ansatz for the Liouvillian we choose

$$\mathcal{D}(\mathbf{d}) = \frac{1}{2} \sum_{k=1}^m d_k \left([a_k \varrho, a_k^\dagger] + [a_k, \varrho a_k^\dagger] \right), \quad (6)$$

with Lindblad operators $\{a_k\}$, and parameterized by the corresponding non-negative dissipation rates, $\mathbf{d} = (d_1, d_2, \dots, d_m)$. Here, a_k is intended to be an ansatz operator for a single Lindblad operator, l_k . However, this does not have to be the case as we will discuss in section 2.3.

Consider the equation of motion of a general observable, O (not necessarily commuting with H), given by the generalized Ehrenfest theorem in integral form in equation (4a). Inserting the ansatz, $A(\mathbf{c})$ for H , and $\mathcal{D}(\mathbf{d})$ for the Liouvillian, and imposing the resulting constraint for a set of observables $\{O_i\}$, with $i = 1, \dots, p$, one obtains a set of linear equations for \mathbf{c} and \mathbf{d} . These can be written as a simple matrix equation⁸;

$$K_H \mathbf{c} + K_D \mathbf{d} = \mathbf{b}. \quad (7)$$

Here, K_H is a $p \times n$ matrices, and K_D is a $p \times m$ matrix, with entries defined by

$$(K_H)_{ij} = -i \int_0^T \langle [O_i, h_j] \rangle_t dt, \quad (8)$$

and

$$(K_D)_{ik} = \frac{1}{2} \int_0^T \langle a_k^\dagger [O_i, a_k] + \text{h.c.} \rangle_t dt, \quad (9)$$

and the vector \mathbf{b} is defined by $b_i = \langle O_i \rangle_T - \langle O_i \rangle_0$. As these equations hold for any density matrix, ϱ , our protocol, similar to the Hamiltonian learning protocol in [36], is resistant to state-preparation errors. Moreover, although we have imposed the above equations for multiple observables, $\{O_i\}$, and a single input state, ϱ , one may also consider multiple input states. Note also, that similar constraints have also been used in [33] to learn Liouvillians from their steady-states.

⁸ To be more precise, we insert the adjoint \mathcal{D}^\dagger of the ansatz for the Liouvillian into equation (7). The adjoint is defined by $\text{tr}[X\mathcal{D}(Y)] = \text{tr}[\mathcal{D}^\dagger(X)Y]$ for all test operators X, Y . We note, that both superoperators contain the same Lindblad operators and dissipation rates.

In the discussion above, we have implicitly assumed that our ansatz is chosen such as to contain the Hamiltonian and Liouvillian, i.e. there exist vectors \mathbf{c}^H , and \mathbf{d}^γ , such that $A(\mathbf{c}^H) = H$, and $\mathcal{D}(\mathbf{d}^\gamma) = \mathcal{L}$. However, in practice we do not expect this to be the case, and we will call such an ansatz *insufficient*⁹. Nevertheless, we can determine the parameters of our ansatz minimizing the violation of equation (7) by minimizing the squared residuals;

$$(\mathbf{c}^{\text{rec}}, \mathbf{d}^{\text{rec}}) = \underset{(\mathbf{c}, \mathbf{d}) : \mathbf{d} \geq 0}{\text{argmin}} \|K_H \mathbf{c} + K_D \mathbf{d} - \mathbf{b}\|, \quad (10)$$

where $\|\mathbf{x}\|$ denotes the 2-norm. In case of an insufficient ansatz, the value of the linear optimization problem, $\min_{(\mathbf{c}, \mathbf{d})} \|K_H \mathbf{c} + K_D \mathbf{d} - \mathbf{b}\|$, will be strictly bounded away from zero.

2.3. Hamiltonian and Liouvillian learning from generalized energy conservation

Instead of inserting our ansatz into equation (4a), we may also insert the ansatz into the much simpler generalized energy conservation condition in equation (4b). Imposing the resulting constraint for a *set* of initial (product) states $\{\varrho_i\}$, where $i = 1, \dots, p$, leads to the following simple matrix equation;

$$[M_H + M_D(\mathbf{d})] \mathbf{c} = 0. \quad (11)$$

The matrices M_H and $M_D(\mathbf{d})$ are $p \times n$ matrices defined by

$$(M_H)_{ij} = \langle h_j \rangle_{i,0} - \langle h_j \rangle_{i,T}, \quad (12)$$

and $M_D(\mathbf{d}) = \frac{1}{2} \sum_k d_k M^{(k)}$, with

$$(M^{(k)})_{ij} = \int_0^T \langle a_k^\dagger [h_j, a_k] + [a_k^\dagger, h_j] a_k \rangle_{i,t} dt, \quad (13)$$

and $\langle X \rangle_{i,t} = \text{tr}[X \varrho_i(t)]$. In contrast to equation (7), where we can impose constraints for multiple input states and multiple observables, the observables appearing in equation (11) are fixed by the ansatz. Therefore to obtain more constraints, one needs to consider more input states. Moreover, as these equations hold for any state, ϱ , the reconstruction is resistant to state preparation errors.

Similar as above, in the case of an insufficient ansatz that cannot fulfill equation (11), one can reconstruct those parameters minimizing the violation of generalized energy conservation;

$$(\mathbf{c}^{\text{rec}}, \mathbf{d}^{\text{rec}}) = \underset{(\mathbf{c}, \mathbf{d}) : \|\mathbf{c}\|=1, \mathbf{d} \geq 0}{\text{argmin}} \|[M_H + M_D(\mathbf{d})] \mathbf{c}\|. \quad (14)$$

Without dissipation, where $\mathbf{d} = 0$, this is a linear optimization problem, while in the presence of dissipation, i.e. $\mathbf{d} > 0$, equation (14) involves non-linear optimization over \mathbf{c} and \mathbf{d} . Note, that for a fixed \mathbf{d} , the minimization over \mathbf{c} results in the smallest singular value, λ_1 , of the constraint matrix $M_H + M_D(\mathbf{d})$. Computing the minimum of the cost function in equation (14) then corresponds to finding non-negative dissipation rates \mathbf{d}^{rec} for which the smallest singular value, λ_1 , of the constraint matrix $M_H + M_D(\mathbf{d})$ attains its minimum. Numerically we find \mathbf{d}^{rec} using the ‘Dividing Rectangles’ (DIRECT) algorithm implemented in SciPy with default settings¹⁰. The corresponding \mathbf{c}^{rec} is then the right-singular vector corresponding to λ_1 . In the following we will denote the singular values of the constraint matrix by $\lambda_n \geq \dots \geq \lambda_2 \geq \lambda_1 \geq 0$.

2.3.1. Learning the overall scale, and conserved quantities

The generalized energy conservation conditions in equation (4b) define the Hamiltonian only up to a scalar factor since they are linear in H . Additionally, these conditions hold for any operator Q with $[Q, H] = 0$, i.e. for all *conserved quantities* of H . An ansatz, $A(\mathbf{c})$, then may contain the Hamiltonian, H , and other conserved quantities of H which admit a decomposition in the form of $A(\mathbf{c})$. For instance, if we choose our ansatz to be k -local, it contains at most k -local conserved quantities, such as the total magnetization, $Q \sim \sum_k \sigma_k^z$, which is a sum of local operators.

So let us assume that an ansatz contains two linearly independent conserved quantities H , and Q , corresponding to two linearly independent vectors \mathbf{c}^H , and \mathbf{c}^Q (the generalization is straightforward). Then,

⁹ As an example, the Hamiltonian may not only contain operators h_j , but also additional operators h'_j not contained in our ansatz. On the level of equation (11) this would result in a truncation of the matrix M which then contains fewer columns than would be required in order to reconstruct \mathbf{c}^H .

¹⁰ The lower bound for each d_k is set to $d_k^{\text{min}} = 0$, and the upper bound is set proportional to the maximum dissipation rate of the model $d_k^{\text{max}} = k\gamma_{\text{max}} = k\max(\{\gamma_k\})$. We check convergence for different values of k and different numbers of cost-function evaluations.

by singular value decomposition, one obtains two right singular vectors $\{\mathbf{v}, \mathbf{v}_\perp\}$, which both satisfy equation (11), and degenerate singular values $\lambda_1 = \lambda_2 = 0$. However, the right singular vectors do not necessarily have to correspond to \mathbf{c}^H and \mathbf{c}_\perp^H . Thus, in the presence of conserved quantities, naively solving equation (14) would in general reconstruct the parameters of a linear combination of conserved quantities.

To distinguish \mathbf{c}^H from other conserved quantities, we impose additional constraints that can only be fulfilled by H , but not by other conserved quantities, including scalar multiples of H . Such a constraint can be obtained from equation (4a) for a generic observable, O , not commuting with H . We spell out the modified equations in appendix A. The optimization problem in equation (14) is then modified to

$$(\mathbf{c}^{\text{rec}}, \mathbf{d}^{\text{rec}}) = \arg \min_{(\mathbf{c}, \mathbf{d}): \mathbf{d} \geq 0} \left\| \begin{pmatrix} M(\mathbf{d}) \\ \xi M^{\text{add}} \end{pmatrix} \mathbf{c} - \begin{pmatrix} 0 \\ \xi \mathbf{b}(\mathbf{d}) \end{pmatrix} \right\|, \quad (15)$$

where $M(\mathbf{d}) = M_H + M_D(\mathbf{d})$ is as in equation (14), and M^{add} and $\mathbf{b}(\mathbf{d})$ contain additional constraints. The *auxiliary parameter* ξ controls a penalty that is added for violating the condition determining H . Choosing this parameter large enough allows us to reconstruct a $\mathbf{c}^{\text{rec}} \approx \mathbf{c}^H$, with the correct overall scale, as a unique solution of equation (15). We emphasize, that by choosing the appropriate additional constraints, equation (15) can also be used to learn other conserved quantities of H .

2.3.2. Learning the full Liouvillian

Since we have dissipative dynamics, one expects that accurately learning the Hamiltonian also requires learning all the individual dissipation rates. This is, however, not necessarily the case. In some cases, different dissipative processes give the same contribution to equation (4b), and therefore, cannot be distinguished by these conditions. More specifically, when inserting an ansatz into this condition, the resulting constraint matrices $\{M^{(k)}\}$ in equation (13), where k labels the different Lindblad operators, become linearly dependent. Then, the decomposition $M_D(\mathbf{d}) = \sum_k d_k M^{(k)}$ is not unique. This dependence leads to symmetries in the cost function in equation (14), such that the Hamiltonian H can be learned exactly, for different dissipation rates \mathbf{d}^{rec} .

In order to learn the full Liouvillian and resolve individual dissipation rates, one needs to add additional constraints in the same way as one does for excluding conserved quantities and for learning the overall scale. This can then be phrased as an optimization problem of the form of equation (15), which we explain in detail in appendix A. This is a unique feature of equation (11). Below, we will compare Hamiltonian and Liouvillian learning based on equation (7), which, generically, does not admit such symmetries, and equation (11), including additional constraints, where we will show how to choose these additional constraints to learn the full Liouvillian.

3. Numerical case studies

In this section we want to illustrate the Hamiltonian and Liouvillian learning methods developed in the previous sections in the context of several experimentally relevant spin-Hamiltonians, more specifically the Ising model in equation (2), and a long-range Hamiltonian with algebraically decaying interactions that is naturally found in trapped ion experiments [9]. We will study the procedure of learning these Hamiltonians, in the presence of weak dissipation, using only a finite amount of numerically simulated measurement data. Moreover, we will aim to learn the Liouvillian, and in particular distinguish local from collective dissipation. As a preparation for the following discussion let us discuss some details of the learning procedure.

3.1. Preliminaries

To learn the Hamiltonian and Liouvillian of a many-body system, the constraint matrices in equation (7), or equation (11), are estimated from a finite amount of experimental data. Therefore, we have to solve the corresponding optimization problem using their noisy estimates, e.g. $\tilde{K}_H = K_H + E$, where E is an error matrix with random entries. Crucially, E is experimentally inaccessible. Therefore, as an experimentally accessible figure of merit to assess the progress of our learning procedure, one may consider the minimum of the objective function in equation (10), or equation (14), given by the smallest singular value λ_1 . In early stages of the learning, where shot-noise, represented by the matrix E , is the dominant source of error, one expects this minimum to decrease $\sim N_{\text{runs}}^{-1/2}$ as we increase N_{runs} , see appendix B for details. In this regime, we will consider the reconstructed parameters \mathbf{c}^{rec} and \mathbf{d}^{rec} , for which we compute error bars using bootstrapping methods. This allows us to identify dominant terms in the Hamiltonian and Liouvillian, thus learning the dominant operator structure. For a sufficient ansatz, this error will asymptotically converge to zero, while for an insufficient ansatz the error will be strictly bounded away from zero, i.e. it will *plateau* in later stages of the learning, where missing ansatz terms are the dominant source of error. Therefore, when observing a plateau we need to extend our ansatz by additional terms.

A fundamental tension in Hamiltonian and Liouvillian learning is the tradeoff between the number of parameters of an ansatz, and the number of measurements that are required to estimate these parameters from experimental data. While ansätze comprising many parameters allow us to learn many aspects of the dynamics, they require a large measurement budget to reduce shot noise. In contrast, in the case of a limited measurement budget, it is quite useful to reduce the number of parameters to just a few relevant ones. This is also physically justified as many experimentally relevant Hamiltonians can be effectively characterized by only a few parameters, which may not even increase with system size. Illustrative examples include translationally invariant systems, or systems exhibiting algebraically decaying interactions. While ansätze with few parameters only require a small measurement budget, they might be insufficient in the limit of an infinite number of measurements. However, as long as these insufficiencies are small, learning will be limited by shot noise. Therefore, our goal in the following is to find an ansatz comprising only few parameters and the dominant terms of the Hamiltonian and Liouvillian, which do not show any insufficiencies given our limited measurement budget.

To find this ansatz, we compare different ansätze by varying their operator content as well as their parametrization, i.e. the dependencies between parameters. We emphasize that *reparametrization* can be done solely by classical post-processing of the data. More specifically, by transforming the vector of parameters, \mathbf{c} , via a *parametrization matrix* G to a new vector of parameters, $\mathbf{c}_G = G^T \mathbf{c}$, where G encodes dependencies between parameters in \mathbf{c} . This leads to a transformation of the learning equations in equations (7) and (11) (see appendix C for technical details). The nature of this reparametrization is best understood by considering the following simple example. An ansatz $A = \sum_{k=1}^N c_k \sigma_k^z$ with N independent parameters is mapped to a translation invariant ansatz via $G^T = (1, 1, \dots, 1)/\sqrt{N}$. Indeed, $A(G^T \mathbf{c}) = \bar{c} \sum_{k=1}^N \sigma_k^z$ is translation invariant, with the overall scale $\bar{c} = 1/\sqrt{N} \sum_{k=1}^N c_k$ as the only free parameter.

We will demonstrate in the following numerical simulations that the ratio of the smallest two singular values of the constraint matrices, λ_1/λ_2 , serves as a quantifier for the learning error of an ansatz and a useful figure of merit to compare different ansätze. This is motivated by the fact that the smallest singular value, λ_1 , quantifies the learning error, e.g. the violation of generalized energy conservation in equation (4b), and that a larger gap $\delta = \lambda_2$ tightens the upper bound on $|\sin(\theta)|$, where $\theta = \angle(\mathbf{c}^H, \mathbf{c}^{\text{rec}})$ is the reconstruction error (see appendix B for details). As we will see below, the gap is typically larger for ansätze with fewer parameters. Therefore, when comparing different ansätze, we choose the one which minimizes λ_1/λ_2 . This results in an ansatz containing only few parameters and small missing terms, i.e. with an early shot-noise scaling, and a low plateau (see figure 2 for an illustration).

As a final remark, let us note that in case the operators O_i , h_j and a_k are Pauli operators, the expectation values in equations (9) and (13) evaluate to $\alpha \langle h_j \rangle$, where $\alpha = 0$ if $[l_k, h_j] = 0$, and $\alpha = -4$ otherwise, and respectively for O_i . Therefore, in case the operators $\{O_i\}$, $\{h_j\}$ are few-body operators these expectation values as well as the commutator $[O_i, h_j]$ remain few-body Paulis. Thus, many of the operators commute and can thus be measured jointly.

Estimating the integrals in equation (A3) requires measuring time-resolved expectation values of the corresponding operators over the duration of the quench, see figure 2. Here, we use the composite Simpson's rule¹¹ which approximates an integral as a series of parabolic segments between N_t equally spaced points $t_m = m \cdot \Delta t$, with $m \in 0, \dots, N_t$ and distance $\Delta t = T/N_t$. In the following, we choose a sufficiently small time step Δt to ensure that errors arising from discretizing the integral can be disregarded in comparison to shot-noise errors¹².

3.2. Learning from generalized energy conservation

To begin, we illustrate Hamiltonian and Liouvillian learning from generalized energy conservation based on equation (11). We assume that the analog quantum simulator is governed by a master equation of the following form

$$\frac{d}{dt} \varrho = -i[H, \varrho] + \sum_{\eta \in \{+, -, z\}} \frac{\gamma_\eta}{2} \sum_{k=1}^N \left(\left[\sigma_k^\eta \varrho, \sigma_k^{\eta, \dagger} \right] + \text{h.c.} \right). \quad (16)$$

Here, H is the Ising Hamiltonian of equation (2) and the Lindblad operators σ_+ , σ_- , and σ_z represent spontaneous absorption, spontaneous emission, and local dephasing, respectively. In this model case, the

¹¹ It reads $\int_0^T \langle P \rangle_t dt \approx \frac{1}{3} \Delta t \left[\langle P \rangle_0 + 4 \sum_{k=1}^{N_t/2} \langle P \rangle_{t_{2k-1}} + 2 \sum_{k=1}^{N_t/2-1} \langle P \rangle_{t_{2k}} + \langle P \rangle_{T} \right]$.

¹² Indeed, for fixed time T one obtains $\int_0^T \langle h \rangle_t dt = I(\Delta t) + K \Delta t^4$, where $I(\Delta t)$ is the integral approximation, and K is a constant. When expressed in terms of N_t this reads $I(N_t) + \tilde{K}/N_t^4$. However, as $I(N_t)$ can only be estimated from data, one obtains $I(N_t) = \tilde{I}(N_t) + \epsilon$, only up to a statistical error ϵ . As $I(N_t)$ is linear, the variance of ϵ roughly scales like $1/(N_t N_s)$, where N_s is the number of shots per time-point.

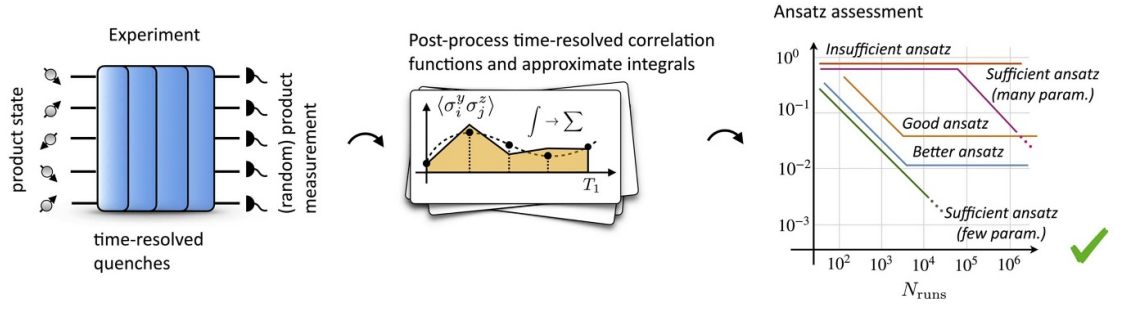


Figure 2. Hamiltonian and Liouvillian learning begins by choosing an ansatz for the operator structure of the Hamiltonian and Liouvillian. Then, one measures time-resolved expectation values of correlation functions, and estimates of their integrals over a period of total time T . Here, we choose to measure the correlation functions at discrete points and estimate the integral using Simpson's rule. The learning progress is assessed by considering the ratio λ_1/λ_2 , which can be directly obtained from experimental data, as a function of N_{runs} . This allows us to compare different ansätze, and to notice when terms in the ansatz are missing, changing the ansatz involves reparametrization, which typically can be done by classical post-processing.

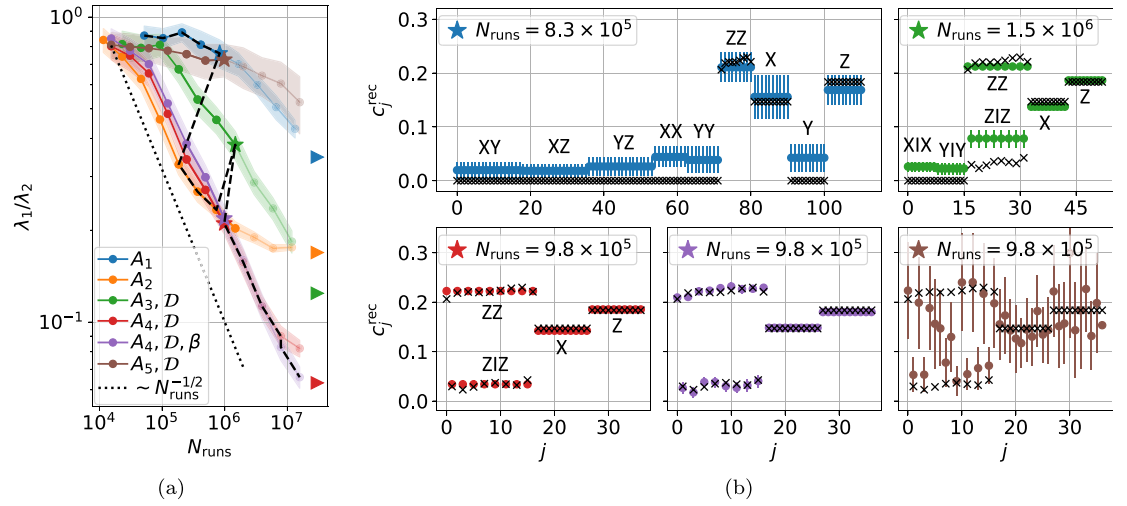


Figure 3. Hamiltonian learning in presence of weak dissipation. (a) The learning error λ_1/λ_2 as a function of the number of runs, N_{runs} , of the simulator for different ansätze described in the main text. The colored triangles indicate the asymptotic values for $N_{\text{runs}} \rightarrow \infty$. The dashed line follows the learning procedure described in the main text. (b) Snapshots of the (normalized) reconstructed Hamiltonian parameters, c_j^{rec} , for some of the ansätze described in the main text. The capital letters denote groups of operators in the ansatz, and within each group the site indices of the coefficients are sorted increasingly from left to right, e.g. the coefficients in group XY correspond to the operators $\sigma_k^x \sigma_{k+1}^y + \sigma_k^y \sigma_{k+1}^x$ in equation (17), and similarly for all other groups of operators. Black crosses indicate the true parameters of the model Hamiltonian as defined below. The point at which these snapshots are taken in terms of N_{runs} is also indicated by asterisks in the left panel. All error bars are computed via bootstrapping as explained in appendix D, using 80 samples. *Model parameters:* The Hamiltonian parameters in equation (16) are chosen as $B_x = 4/5B_z$. Moreover, the coupling strengths are described by 4-th order polynomials of the form $J_{i,i+1}^x = B_z \sum_{l=1}^5 a_l \cdot x_{i,i+1}^{l-1}$, and $J_{i,i+2}^z = B_z \sum_{l=1}^5 b_l \cdot x_{i,i+2}^{l-1}$ where $x_{i,j} = [(i+j) - (N+1)]/N$. We choose the following coefficients: $\mathbf{a} = (6/5, 1/20, 1/5, 0, -2/5)$ and $\mathbf{b} = (1/5, 1/20, -2/5, 0, 4/5)$. The dissipation rates are chosen such that $(\gamma_+, \gamma_-, \gamma_z) = (1, 1.5, 2) \times 10^{-2} B^z \ll \|H\|$. To estimate the integrals in equation (13), we use $B_z \Delta t = 1/64$, i.e. 64 equidistant times, where at each time we spend only $1/64$ of the measurements per basis, compared to the quench times $B_z T \in [0, 0.5, 1]$. Expectation values at $t = 0$ are calculated exactly from the initial state.

Hamiltonian couplings and dissipation rates are chosen such that the dominant terms are the nearest-neighbor couplings and fields of the Ising Hamiltonian, while the sub-dominant terms are next-nearest-neighbor couplings in the Hamiltonian as well as the dissipative processes. Moreover, we included small spatial variations in the couplings of the Hamiltonian. We summarize the choice of parameters in the caption of figure 3.

In the following, our goal will be to learn the above-described model Hamiltonian and Liouvillian from simulated quench experiments. We will limit the total number of simulated runs $N_{\text{runs}} \approx 10^6$, which is a reasonable limit for experiments with trapped ions. We will illustrate how c^{rec} , including its error bars that we obtain via bootstrapping, and the learning error λ_1/λ_2 can be used in this scenario to identify the operator content and relevant parameters of the model Hamiltonian and Liouvillian.

1. Identifying the dominant terms of the Hamiltonian. In a first step, we seek to identify the dominant terms in the Hamiltonian H , assuming (prior knowledge) that dissipation is typically weak compared to the

Hamiltonian. Therefore, in this first step we will only include an ansatz for the Hamiltonian and no ansatz for the Liouvillian. As, interactions are typically finite range, we generically expect the dominant terms to be nearest-neighbor terms. Moreover, in quantum simulation of condensed matter models one expects more or less homogeneous couplings (translational invariance). Therefore, we start with the following ansatz for the Hamiltonian

$$A_1 = c_{xx} \sum_{k=1}^{N-1} \sigma_k^x \sigma_{k+1}^x + c_{xy} \sum_{k=1}^{N-1} (\sigma_k^x \sigma_{k+1}^y + \sigma_k^y \sigma_{k+1}^x) + \cdots + c_x \sum_{k=1}^N \sigma_k^x + c_y \sum_{k=1}^N \sigma_k^y + c_z \sum_{k=1}^N \sigma_k^z, \quad (17)$$

comprising all nearest-neighbor interactions, i.e. $\{xx, yy, zz, xy, xz, yz\}$, with spatially homogeneous coefficients. As commuting operators can be measured jointly, the operators in A_1 can be measured by classically post-processing data obtained from measuring the following nine product operators, independently of the system size: $\sigma_x^{\otimes N}$, $\sigma_y^{\otimes N}$, $\sigma_z^{\otimes N}$, and six operators of the form $\sigma_a \otimes \sigma_b \otimes \sigma_a \otimes \sigma_b \otimes \cdots$, for all combinations $a \neq b$ of distinct Pauli operators.

In figure 3(a) we plot λ_1/λ_2 for the ansatz A_1 as function of N_{runs} (blue line). At small N_{runs} one expects to observe a decrease of λ_1/λ_2 with N_{runs} indicating the errors are dominated by shot noise. As the ansatz A_1 is insufficient, i.e. misses terms present in equation (2), one expects a plateau in λ_1/λ_2 for large N_{runs} . This plateau appears beyond our maximum measurement budget at $N_{\text{runs}} \approx 10^7$ experimental runs, which is indicated by a blue triangle on the y -axis in figure 3(a). Nevertheless, at $N_{\text{runs}} \approx 10^6$ the reconstructed parameters \mathbf{c}^{rec} in figure 3(b) (blue data) identify the dominant terms in the Hamiltonian, i.e. nearest-neighbor $\sigma_z \otimes \sigma_z$ interactions and fields in σ_x and σ_z direction.

At this point, we cannot rule out the presence of the other terms in the operator ansatz A_1 . However, as the values of the non-dominant terms are much smaller than the ones of the dominant terms (see figure 3(b), blue data), we remove all non-dominant terms from the ansatz in the next step. It should be noted that, in case those terms were present in the Hamiltonian, we would encounter a plateau in λ_1/λ_2 at a later stage of the learning process. We will later see that this is, however, not the case (as dissipative terms are still missing in our ansatz). On the other hand, if the dominant terms identified from \mathbf{c}^{rec} of A_1 constitute a good approximation to the Hamiltonian, we expect λ_1/λ_2 to decrease. This is indeed the case as the reparametrized ansatz,

$$A_2 = c_{zz} \sum_{k=1}^{N-1} \sigma_k^z \sigma_{k+1}^z + c_x \sum_{k=1}^N \sigma_k^x + c_z \sum_{k=1}^N \sigma_k^z \quad (18)$$

leads to a much smaller value of λ_1/λ_2 . We emphasize that learning with the ansatz A_2 only requires measurements of $\sigma_x^{\otimes N}$ and $\sigma_z^{\otimes N}$. Therefore, some of the measurements performed for A_1 can be recycled for learning the parameters of A_2 . By further measuring, a plateau in λ_1/λ_2 starts to emerge at $N_{\text{runs}} \approx 3 \times 10^5$, as shown in figure 3(a) (orange line). At this point we have identified all the dominant terms in H . To continue the learning process we need to extend the operator ansatz. Therefore, in the next step we want to learn subdominant terms of H , as well as the Liouvillian.

2. Learning the sub-dominant terms of the Hamiltonian, and learning the Liouvillian. One expects, *a priori*, that when learning smaller terms of the Hamiltonian, weak dissipative effects become relevant. Therefore, we now include an ansatz for the Liouvillian

$$\mathcal{D}(\varrho) = \sum_{\eta \in \{+, -, z\}} \frac{d_\eta}{2} \sum_{k=1}^N \left(\left[\sigma_k^\eta \varrho, \sigma_k^{\eta, \dagger} \right] + \text{h.c.} \right), \quad (19)$$

and extend our ansatz for the Hamiltonian by next-nearest neighbor couplings. As before, we choose our ansatz to be spatially homogeneous, which leads to

$$A_3 = A_2 + c_{x1x} \sum_{k=1}^{N-2} \sigma_k^x \sigma_{k+2}^x + c_{y1y} \sum_{k=1}^{N-2} \sigma_k^y \sigma_{k+2}^y + c_{z1z} \sum_{k=1}^{N-2} \sigma_k^z \sigma_{k+2}^z. \quad (20)$$

This ansatz for the Hamiltonian and Liouvillian requires measurements of the form $\sigma_x^{\otimes N}$, $\sigma_y^{\otimes N}$ and $\sigma_z^{\otimes N}$, which means that all of the data taken for A_2 can be reused for learning the parameters of A_3 . Note that including dissipation requires the same measurement bases, but at various times to estimate the integrals in equation (13).

An ansatz for the dissipation and next-nearest-neighbor terms now leads to a lower plateau of λ_1/λ_2 compared to A_2 . However, this plateau appears for $N_{\text{runs}} \geq 10^7$, see figure 3(a) (green line), which exceeds

our assumed measurement budget. Nevertheless, the reconstructed parameters of the Hamiltonian, \mathbf{c}^{rec} , at $N_{\text{runs}} \approx 10^6$, suggest that the next-nearest-neighbor $\sigma_z \otimes \sigma_z$ couplings are larger compared to other next-nearest-neighbor terms, see figure fig:example1(b) (green data). Therefore, we reparametrize our ansatz for the Hamiltonian to

$$A_4 = A_2 + c_{z1z} \sum_{k=1}^{N-2} \sigma_k^z \sigma_{k+2}^z. \quad (21)$$

Indeed, this leads to a significantly smaller ratio λ_1/λ_2 compared to A_2 and A_3 , see figure 3(a) (red line). Moreover, the small error bars in the reconstructed parameters in figure 3(b) (red data) show that $N_{\text{runs}} = 10^6$ is already sufficient for learning A_4 . Note that at this point, we have identified all dominant and subdominant terms. However, so far we have not learned a possible spatial structure of the couplings in H .

3. Learning the spatial variations of the couplings. As can be seen in figure 3(a) (red line), the plateau in A_4 is reached at $N_{\text{runs}} \approx 10^7$, which again exceeds our available measurement budget. Therefore, we are not allowed to conclude that ansatz A_4 is insufficient to describe the Hamiltonian in equation (2) by only considering the ratio λ_1/λ_2 . However, at this point, one can choose to test other ansätze and compare their corresponding reconstructed parameters, or to see if smaller values of λ_1/λ_2 are achieved. As one typically expects small spatial variations in the couplings of H , and A_4 was already a good approximation of H , we want to test for spatial variations on top of the spatially homogeneous ansatz A_4 . To this end, we simply treat the parametrization of the ansatz A_4 as a regularization of the cost function, see appendix C for technical details. That is, we start with the optimization problem in equation (14), and add a penalty term with control parameter β for deviating from a given parametrization. Choosing $\beta \gg 1$ imposes the exact parametrization of A_4 , and we then successively decrease β until one observes the emergence of spatial variations, and the corresponding error bars, see figure 3(b) (purple data). This ensures that our learning process is dominantly limited by shot noise, and not by having too few parameters in the ansatz.

With a larger measurement budget one could then successively increase N_{runs} while decreasing β , until the desired accuracy is reached, which is reminiscent of Bayesian learning [31, 43]. Therefore, we conclude that we have indeed successfully determined the operator content and parameters of the Hamiltonian and Liouvillian in our simulated experiment.

As a final step, we may compare the reconstructed parameters of our learning procedure to the ones obtained from a ‘naive’ ansatz of the form

$$A_5 = \sum_{k=1}^{N-1} c_{zz}^{(k)} \sigma_k^z \sigma_{k+1}^z + \sum_{k=1}^{N-2} c_{z1z}^{(k)} \sigma_k^z \sigma_{k+2}^z + \sum_{k=1}^N c_x^{(k)} \sigma_k^x + \sum_{k=1}^N c_z^{(k)} \sigma_k^z,$$

that has a total of $4N - 3$ parameters. One notices that our strategy leads to a much more accurate reconstruction, cf figure 3(b) (brown data).

3.3. Learning from generalized Ehrenfest theorem

In this section, we will learn the Hamiltonian and Liouvillian of a model system from the generalized Ehrenfest theorem in equation (7). As this method requires the estimation of expectation values of many observables, $\{O_i\}$, we will use the randomized measurement toolbox [46] to obtain the necessary data. This will also allow us to demonstrate Hamiltonian learning in large systems of tens of qubits, and provide evidence that given a fixed number of shots the reconstruction error is independent of the system size, if the number of Hamiltonian parameters can be kept constant.

We consider a model system involving long-range spin–spin interactions, as realized in trapped-ion setups. The effective Hamiltonian is given by [9]

$$H_{XY} = \sum_{i < j} J_{ij} \left(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y \right) + \sum_{i=1}^N B_i \sigma_i^z, \quad (22)$$

with interaction strengths J_{ij} and site-dependent magnetic field B_i . In an idealized description, it is often assumed that

$$J_{ij} = \frac{J_0}{|p(i) - p(j)|^\alpha}, \quad B_i = B_z, \quad (23)$$

where $p(i)$ is the position of the i th ion and with tunable $0 \leq \alpha \leq 3$. To add features to the model, we assume that the ions are randomly shifted from their equilibrium positions, i.e. $p(i) = i + r_i$, where each r_i is

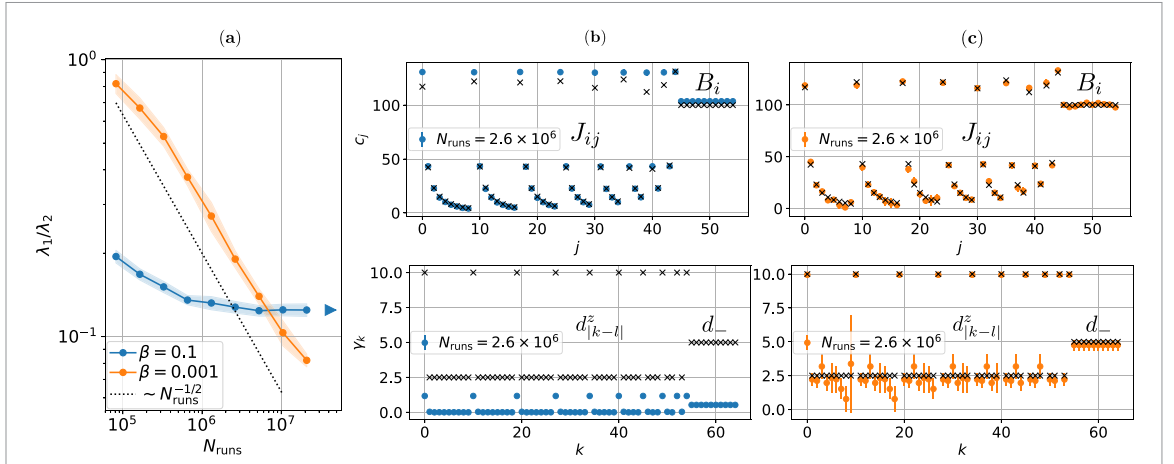


Figure 4. Learning collective dissipation using random measurements. (a) The ratio λ_1/λ_2 of the matrix $(M, -b)$ for different regularization parameters β . (b), (c) Snapshots of the learned Hamiltonian parameters, c_j^{rec} , corresponding to the parameters J_{kl} and B_i defined in equation (22). The required measurement budget is also indicated in the left panel. Black crosses correspond to the parameters of the model Hamiltonian specified below. All error bars are computed via bootstrapping as explained in appendix D, using 40 samples. The time-traces in equation (13) are evaluated using $B_z \Delta_T = 1/20$, i.e. using 20 time-steps, using the same amount of random measurements at each time up to a time of $B_z T = 0.5$. Expectation values at $t = 0$ are calculated exactly from the initial state. *Model parameters:* For the Hamiltonian we choose parameters as specified in equation (23) with $B_z = 100$, $J_0 = 6/5B_z$ and $\alpha = 1.5$. Moreover, in the Liouvillian, we choose $\gamma_- = B_z/20$ in equation (24), as well as $\Gamma_{kl}^z = \gamma_k^z \delta_{kl} + \Gamma_0$ with $\Gamma_0 = B_z/40$ and $\gamma_k^z = 3\Gamma_0$ in equation (25).

sampled uniformly from $[-0.05, 0.05]$. Models with random ion positions are studied, for example, in the context of topological defects in the Frenkel–Kontorova model [47, 48].

To account for spontaneous emission of trapped ions, we add the spatially homogeneous Lindblad terms

$$\mathcal{L}^-(\varrho) = \frac{\gamma_-}{2} \sum_{k=1}^N \left(\left[\sigma_k^- \varrho, \sigma_k^{-\dagger} \right] + \left[\sigma_k^-, \varrho \sigma_k^{-\dagger} \right] \right). \quad (24)$$

Moreover, we consider the presence of dephasing terms originating from the presence of a fluctuating, classical magnetic field, that leads to shifts of the energy levels of the k th ion proportional to $B_k^{\text{cl}}(t)$. It can be shown that in the white-noise limit, i.e. with correlation function $\langle B_k^{\text{cl}}(t) B_l^{\text{cl}}(t') \rangle = \Gamma_{kl}^z \delta(t - t')$, this leads to Lindblad terms of the form

$$\mathcal{L}^z(\varrho) = \sum_{k,l=1}^N \Gamma_{kl}^z \left\{ \sigma_k^z \varrho \sigma_l^z - \frac{1}{2} \left(\sigma_k^z \sigma_j^z \varrho + \varrho \sigma_k^z \sigma_j^z \right) \right\}. \quad (25)$$

In particular, uncorrelated magnetic field fluctuations lead to a diagonal matrix $\Gamma_{kl}^z = \gamma_k^z \delta_{kl}$, whereas global fluctuations, with $B_k^{\text{cl}}(t) = B^{\text{cl}}(t)$ for all k , lead to a constant matrix $\Gamma_{kl}^z = \Gamma_0$. In the present example we will, for simplicity, only consider uncorrelated and global fluctuations, which lead to a matrix

$$\Gamma_{kl}^z = \gamma_k^z \delta_{kl} + \Gamma_0, \quad (26)$$

(see also figure 4 for the specific choices of parameters). We note, that in experiments with, for instance, long ion strings, or local magnetic fields one expects a more complicated structure of the matrix Γ_{kl}^z .

1. Hamiltonian and Liouvillian learning using randomized measurements. In section 3.2, we already illustrated learning of the operator content, and the relevant parameters of the Hamiltonian from minimal assumptions. Here, we want to focus on distinguishing local from collective dephasing, and analyse the scaling of the reconstruction error. Therefore, we start with a parametrized ansatz, that already incorporates the algebraically decaying spin–spin interactions in H_{XY} , that one would expect to find in a trapped-ion experiment, as well as a constant magnetic field. We choose as an ansatz for the Hamiltonian

$$A(c) = \sum_{i < j} \frac{c_0}{|i - j|^\alpha} \left(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y \right) + b \sum_k \sigma_k^z, \quad (27)$$

which depends non-linearly on the parameter α , and hence, will require non-linear minimization of λ_1 over α as explained in appendix C. As an ansatz for the Liouvillian we choose

$$\mathcal{D}(\varrho) = \frac{d_-}{2} \sum_{k=1}^N \left(\left[\sigma_k^- \varrho, \sigma_k^{-\dagger} \right] + \text{h.c.} \right) + \sum_{k,l=1}^N d_{|k-l|}^z \mathcal{D}_{kl}^z(\varrho), \quad (28)$$

that includes single-spin spontaneous decay with jump operator σ_k^- , and single-spin as well as multi-spin collective dephasing \mathcal{D}_{kl}^z with operator structure as described in equation (25).

To obtain constraint operators, $\{O_i\}$, we proceed as follows; we choose a set of random Pauli measurements, $\{P_\alpha\} = \{\sigma^{\alpha_1} \otimes \dots \otimes \sigma^{\alpha_N}\}$, with $\alpha_i \in \{x, y, z\}$, which we measure at consecutive times from $t = 0$ up to $t = T$, similar as before (see figure 4 for details). From this set of Pauli measurements we extract all one- and two-qubit operators that are contained in $\{P_\alpha\}$, as well as their commutators with the ansatz operators and jump operators (in case those are contained). For these operators and commutators we extract expectation values to obtain the data necessary to estimate the constraint matrix in equation (7). Choosing the set $\{P_\alpha\}$ large enough ensures that the necessary operators are contained in this set.

The results are presented in figure 4. In figure 4(a) we show the ratio of the two smallest singular values of the matrix $(K_H, K_D, -\mathbf{b})$, which can be obtained by rewriting equation (7), where K_H, K_D , and \mathbf{b} are the constraint matrices and vector respectively. The quantity λ_1/λ_2 shows a similar scaling behavior in early and late stages of the learning procedure as the analog quantity in section 3.2. Moreover, it correctly recognizes the insufficiency of the translation invariant ansatz ($\beta = 0.1$, blue), as well as the larger number of parameters of the almost unparametrized ansatz ($\beta = 0.001$, orange). In figure 4(b) we show the learned Hamiltonian parameters (top) and dissipation rates (bottom) for both ansätze, and a total measurement budget of $N_{\text{runs}} = 2.6 \times 10^6$, in comparison to the true values (black). For the translation invariant ansatz (blue) one notices that while the reconstructed parameters of the Hamiltonian roughly resemble the true parameters, the dissipation rates turn out wrong. We have numerically verified that this artefact does not disappear for larger measurement budget. However, by lowering the penalty β for deviating from translation inference the solution \mathbf{c}^{rec} correctly accounts for disorder, and the dissipation rates are learned accurately.

2. Comparison between learning from Ehrenfest constraints and energy conservation. To make a fair comparison between Hamiltonian and Liouvillian learning from energy conservation and the Ehrenfest theorem, we have learned the same model system using equation (11). This, however, also requires the measurement of additional constraints to learn the overall scale, and to distinguish local from collective dephasing due to symmetries of H_{XY} . The results are discussed in appendix F. One observes, that in a system with conserved quantity, learning from the Ehrenfest theorem in equation (7) requires less measurements, when compared to learning from generalized energy conservation in equation (11). For instance, verifying the presence of collective dephasing using the Ehrenfest theorem requires $N_{\text{runs}} \approx 10^6$ (cf figure 4), compared to $N_{\text{runs}} \approx 10^7$ when using energy conservation (cf figure 9 in appendix F).

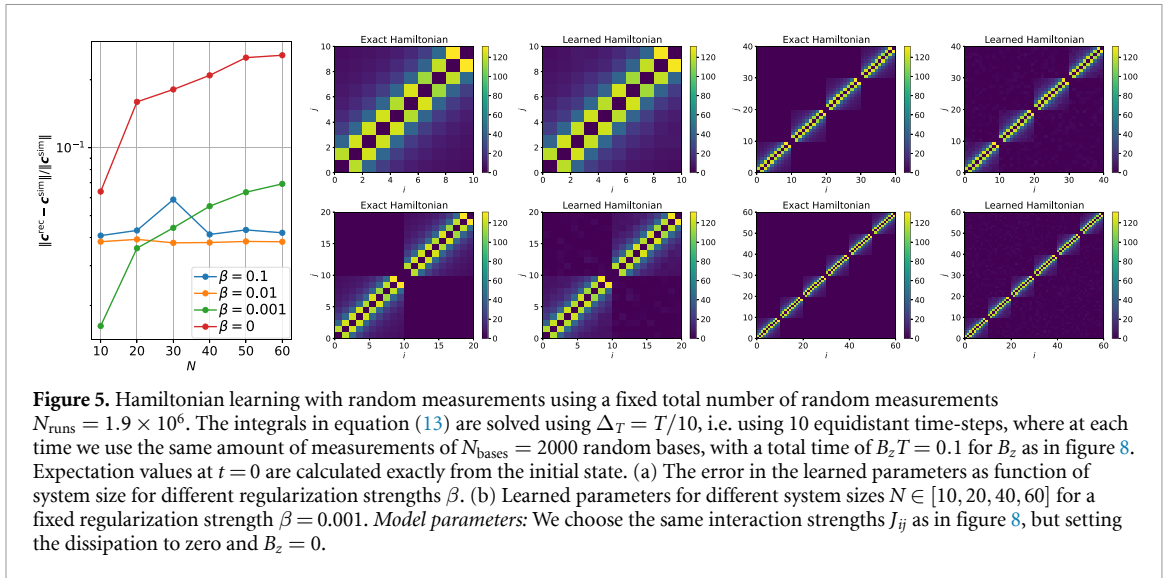
3. Scaling to larger system sizes. We now want to study the scaling behavior of the reconstruction error of the Hamiltonian parameters, by varying system size from $N = 10$ up to $N = 60$ spins, for a Hamiltonian that is dominantly described by a constant number of parameters. In order to simplify the presentation and to reduce the classical computation time needed for simulating the measurements we make the following modifications to our model system: (i) we set the magnetic field and the dissipation rates to zero. (ii) We consider a Hamiltonian of the form

$$H = \sum_{k=1}^{N_{\text{sub}}} \sum_{i \neq j=1}^N J_{ij} \left(\sigma_{(kN+i)}^x \sigma_{(kN+j)}^x + \sigma_{(kN+i)}^y \sigma_{(kN+j)}^y \right), \quad (29)$$

consisting of N_{sub} non-interacting subsystems of $N = 10$ spins each, which makes it easy to simulate on a classical computer.

To learn the Hamiltonian in equation (29) for the different system sizes, we want to compare learning without parametrization, where the number of parameters in the ansatz grows with systems size, to learning using a parametrized ansatz where the number of parameters is kept constant with system size. Therefore, we choose an ansatz

$$A(\mathbf{c}) = \sum_{k=1}^{N_{\text{sub}}} \sum_{i \neq j=1}^{10} \frac{c_0}{|i-j|^\alpha} \left(\sigma_{i+k}^x \sigma_{j+k}^x + \sigma_{i+k}^y \sigma_{j+k}^y \right) + \sum_{k \neq l=1}^{N_{\text{sub}}} \sum_{i \neq j=1}^{10} c_1 \left(\sigma_{i+k}^x \sigma_{j+l}^x + \sigma_{i+k}^y \sigma_{j+l}^y \right), \quad (30)$$



with a constant number of parameters, and that depends non-linearly on the parameter α similar to equation (27). As in the previous example the parametrization of this ansatz is not strictly enforced, but depends on the chosen regularization strength β (see appendix C for technical details). We emphasize that for $\beta = 0$ we are learning from an unparametrized ansatz, with an independent parameter for each operator in the ansatz. In this case, our reconstructed Hamiltonian does not have the subsystem-structure of the true Hamiltonian in equation (29) inherently built-in; rather, this structure is learned by including all possible inter-subsystem interaction terms.

We learn the Hamiltonian in equation (29) for $N = 10, \dots, 60$ systems. We choose a total of 2000 random Pauli measurements for each of a total of 11 time-steps up to $T = 1$, leading to a total measurement budget of $N_{\text{runs}} = 1.5 \times 10^6$. For the constraint matrix M and vector b in equation (10) we choose the constraints to be all possible one- and two-qubit constraint operators, and evaluate the corresponding constraint matrix M and vector b . We then determine the learned coefficients c^{rec} from equation (10).

The result is shown in figure 5. In figure 5(a) we show the relative error of the learned parameters defined by

$$\Delta c = \|c^{\text{rec}} - c^{\text{sim}}\| / \|c^{\text{sim}}\| \quad (31)$$

as a function of the system size N . We find that for weakly, or unparametrized ansätze, for which the number of learned parameters grows with system size, the reconstruction error grows only moderately. This is likely explained by the fact that in larger systems a single round of measurements extracts more information as more correlation functions can be measured in a single shot. Moreover, for parametrized ansätze with a system-size independent number of parameters the reconstruction error also remains independent of the system-size.

4. Conclusion and outlook

In this study, we have devised and compared different methods for learning the Hamiltonian and Liouvillian in the analog quantum simulation of many-body systems. Our work applies to a scenario where one has direct access to the quantum device, however, in the literature, other scenarios have been considered, where one does not have direct access to the quantum device, see, e.g. [25, 49]. Our method is applicable in a regime of experimental relevance where the dissipation is weak compared to the coherent evolution. Our protocols are based on quench experiments, where initial product states evolve under coherent and dissipative dynamics, and the resulting state is measured in a product basis. Hamiltonian and Liouvillian learning can be understood as a sample efficient process tomography of quantum simulators. The learning methods begin with an ansatz for the operator structure of the Hamiltonian and Liouvillian. The quality of this ansatz can be monitored by measuring the learning error. Our strategy encompasses the reparametrization of the ansatz. This typically allows data recycling from previous measurement runs, but requires additional classical

post-processing. This approach enables us to identify step by step first the dominant operator content of the Hamiltonian and Liouvillian, and successively sub-dominant terms within a limited measurement budget. A distinctive feature of learning from generalized energy conservation is that we can learn the Hamiltonian without the necessity to learn the entire Liouvillian, thus reducing the number of parameters to be learned. However, we demonstrated that additional constraints can be employed to learn the entire Liouvillian and to ascertain the overall scale of the Hamiltonian. Furthermore, these additional constraints can facilitate the learning of the Hamiltonian even when there are conserved quantities as operators commuting with the system Hamiltonian. We demonstrate how randomized measurements [46] can be utilized to learn the Hamiltonian and Liouvillian from the generalized Ehrenfest theorem, as this requires the estimation of many different correlation functions. Moreover, this allowed us to numerically demonstrate Hamiltonian learning for large systems consisting of tens of spins. While the focus of the present paper has been on spin models, the central ideas of Hamiltonian and Liouvillian learning also carry over to Bose and Fermi Hubbard models.

Extensions of the present work should consider scenarios where the experimental Hamiltonian (and Liouvillian) involves a large number of small terms, which, e.g. emerge as corrections in effective many-body spin models in a low-energy description. Such a formulation might involve a statistical description as learning of an ensemble of Hamiltonians. Along similar lines, Hamiltonian learning might also account for slow drifts of experimental Hamiltonians and Liouvillians. Considering alternative viewpoints, Hamiltonian and Liouvillian learning can also be phrased in the language of Bayesian inference, similar to [31], establishing an interesting link between techniques of parameter estimation in multi-parameter quantum metrology, and optimal sensing with finite measurement budgets. Finally, exploring alternative routes should include using (short-range) entangled states as inputs, which may be easily prepared in experiments.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

We would like to thank Manoj K Joshi, Benoit Vermersch, and Christian Roos for interesting discussions and valuable feedback on the manuscript. TK would like to thank H Chau Nguyen for helpful discussions. This research is supported by the U.S. Air Force Office of Scientific Research (AFOSR) via IOE Grant No. FA9550-19-1-7044 LASCeM, by the European Union's Horizon Europe programmes HORIZON-CL4-2022-QUANTUM-02-SGA via the Project 101113690 (PASQuanS2.1) and HORIZON-CL4-2021-DIGITAL-EMERGING-02-10 under Grant Agreement No. 101080085 (QCFD), by the Austrian Science Fund (FWF) through the grants SFB BeyondC (Grant No. F7107- N38) and P 32273-N27 (Stand-Alone Project), by the Simons Collaboration on Ultra-Quantum Matter, which is a grant from the Simons Foundation (651440, P.Z.), and by the Institut für Quanteninformation. Innsbruck theory is a member of the NSF Quantum Leap Challenge Institute Q-Sense. TK and BK acknowledge funding from the BMW endowment fund. Our research is part of the Munich Quantum Valley, which is supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus. C.K. acknowledges support from the NSF through a grant for ITAMP at Harvard University. The computational results presented here have been achieved (in part) using the LEO HPC infrastructure of the University of Innsbruck.

Appendix A. Additional constraints

We consider the equation of motion of a general observable, O , not commuting with H , as given in equation (4a). Inserting the ansatz $A(\mathbf{c})$ for H and $\mathcal{D}(\mathbf{d})$ for the Liouvillian one obtains again a simple matrix equation;

$$M^{\text{add}} \mathbf{c} = \mathbf{b}(\mathbf{d}), \quad (\text{A1})$$

where

$$M_{ij}^{\text{add}} = \int_0^T \langle -i [O, h_j] \rangle_{i,t} dt, \quad (\text{A2})$$

and

$$b_i(\mathbf{d}) = \langle O \rangle_{i,T} - \langle O \rangle_{i,0} - \frac{1}{2} \sum_k d_k \int_0^T \langle a_k^\dagger [O, a_k] + \text{h.c.} \rangle_{i,t} dt. \quad (\text{A3})$$

Note, that similar constraints have also been used in [33] to learn Liouvillians from their steady-states. Compared to the conditions in equation (11), the constraints above contain additional integrals in equation (A2) that need to be estimated from experimental data. To obtain the reconstructed parameters we solve the combined system of equations including the additional constraints, i.e.

$$\mathbf{c}_0^{\text{rec}}(\xi) = \arg \min_{\mathbf{c}} \left[\min_{\mathbf{d} \geq 0} \left\| \begin{pmatrix} M(\mathbf{d}) \\ \xi M^{\text{add}} \end{pmatrix} \mathbf{c} - \begin{pmatrix} 0 \\ \xi \mathbf{b}(\mathbf{d}) \end{pmatrix} \right\| \right], \quad (\text{A4})$$

where the parameter ξ controls the relative weight between the constraints defined by $M(\mathbf{d})$ and M^{add} . In analogy to equation (14), solving equation (A4) requires a simultaneous minimization over \mathbf{c} and \mathbf{d} , while ξ serves as a ‘hyper-parameter’ of the optimization. For $\xi = 0$ the vector $\mathbf{c}_0^{\text{rec}}(0)$ may be a linear combination of the Hamiltonian and additional, linearly independent, conserved quantities due to the degeneracy of the spectrum of $M(\mathbf{d})$, as discussed in the main text. Then, by choosing the value of ξ large enough, one removes components of conserved quantities from $\mathbf{c}_0^{\text{rec}}$. In a similar way, one can choose additional constraints such that the solution for the dissipation rates \mathbf{d} becomes unique.

Note that the norm of $\mathbf{c}_0^{\text{rec}}$ in equation (A4) depends on ξ and becomes exact only in the limit $\xi \rightarrow \infty$. A finite ξ typically leads to a smaller value for the overall scale of the Hamiltonian. This is because the homogeneous part of equation (A4) is perfectly solved for $\mathbf{c} = 0$. Nevertheless, one obtains a $\mathbf{c}_0^{\text{rec}} \propto \mathbf{c}^H$. Then, the correct overall scale s , defined by $\mathbf{c}^H = s \cdot \mathbf{c}_0^{\text{rec}}$, can be determined solely via the additional constraints

$$\sum_j M_{ij}^{\text{add}} (s \mathbf{c}_0^{\text{rec}})_j = b_i(\mathbf{d}^{\text{rec}}). \quad (\text{A5})$$

where \mathbf{d}^{rec} are the optimal dissipation rates determined from equation (A4). Then, averaging over all additional constraints

$$s = \frac{1}{p} \sum_{i=1}^p \frac{b_i(\mathbf{d}^{\text{rec}})}{(M^{\text{add}} \mathbf{c}_0^{\text{rec}})_i} \quad (\text{A6})$$

yields the correct overall scale.

Appendix B. Effect of shot noise

Here, we discuss in more detail the role of the two smallest singular values of the constraint matrix. As an example we will consider equation (11), but similar arguments also apply to equation (7).

In case the ansatz for the dynamics is sufficient, the only source of error is shot noise. Therefore, with a finite measurement budget one obtains a noisy estimate, $\tilde{M}(\mathbf{d}) = M(\mathbf{d}) + E(\mathbf{d})$, of the true constraint matrix $M(\mathbf{d})$, with an additive error matrix $E(\mathbf{d})$. One can establish the following bound on the perturbed singular value λ_1

$$\lambda_1 [\tilde{M}(\mathbf{d}^{\text{rec}})] \leq \lambda_1 [\tilde{M}(\mathbf{d}^\gamma)] \leq \|E(\mathbf{d}^\gamma)\|, \quad (\text{B1})$$

where the first inequality holds as \mathbf{d}^{rec} is the minimum of the cost-function in equation (14), and the second inequality follows from Weyl’s inequality and the fact that $\lambda_1[M(\mathbf{d}^\gamma)] = 0$ for a sufficient ansatz. Therefore, in early stages of the learning procedure, one expects λ_1 to decrease $\sim N_{\text{runs}}^{-1/2}$ as we increase the number of runs, N_{runs} , in the experiment. As the ansatz is sufficient $\lambda_1[\tilde{M}(\mathbf{d}^{\text{rec}})]$ will be zero in the absence of shot noise. However, if an ansatz is insufficient, λ_1 is strictly bounded away from zero, even in the absence of shot noise. Thus, for N_{runs} sufficiently large, λ_1 will reach a plateau. Here, the dominant source of error will be systematic errors due to missing terms in our ansatz.

In case of a sufficient ansatz and without degeneracy of the smallest singular values, one can understand the role of the singular value λ_2 . To this end, one considers the angle $\theta = \angle(\mathbf{c}^H, \mathbf{c}^{\text{rec}})$. It is a well known result in singular subspace perturbation theory, that the stability of a singular vector under perturbation depends on the gap between the corresponding singular value and the remainder of the spectrum, which is known as

the Davis–Kahan–Wedin $\sin(\theta)$ -theorem [50, 51]. For the case we consider this theorem establishes the following upper bound on the angle θ [51]

$$|\sin(\theta)| \leq \frac{\|E(\mathbf{d}^{\text{rec}}) + R(\Delta\mathbf{d})\|}{\delta}, \quad (\text{B2})$$

where $R(\Delta\mathbf{d}) = \sum_k \Delta d_k M_D^{(k)}$. Here, $\delta = \lambda_2 - \lambda_1^{\text{exact}}$, and $\lambda_1^{\text{exact}} = 0$ for a sufficient ansatz. In case the elements of E are i.i.d. Gaussian random variables one can establish average bounds on $|\sin(\theta)|$ [36]. However, we emphasize that the noise will in general be correlated due to the fact that we perform many of the required measurements in parallel. Moreover, the error matrix $E(\mathbf{d})$ is experimentally inaccessible. Nevertheless, as we demonstrate in section 3, the ratio λ_1/λ_2 , which can be directly computed from experimental data, serves as a quantity to assess an ansatz.

Appendix C. (Re-)Parameterization

We want to discuss here in more detail the different possibilities to reparametrize an ansatz. In general, there are two possibilities, but we will show that both can be understood in terms of regularization of the cost function in equation (14). Note, that the same holds for equation (10).

In the first reparametrization the vector of parameters, \mathbf{c} , is mapped via a *parametrization matrix* G to a new vector of parameters, $\mathbf{c}_G = G^T \mathbf{c}$. Here, the matrix G is a $n \times \tilde{n}$ matrix, with $n \geq \tilde{n}$, encoding the dependencies between the n parameters in \mathbf{c} , and the \tilde{n} parameters in \mathbf{c}_G . Moreover, one requires that $G = (\mathbf{g}_1, \dots, \mathbf{g}_{\tilde{n}})$, with orthonormal columns $\langle \mathbf{g}_i | \mathbf{g}_j \rangle = \delta_{ij}$. One can easily verify that $G^T G = \mathbb{1}_{\tilde{n} \times \tilde{n}}$, and $GG^T = \sum_{i=1}^{\tilde{n}} |\mathbf{g}_i\rangle\langle \mathbf{g}_i|$, which is a projector onto the support of G^T , and thus G is an isometry. In particular cases G can also depend on non-linear parameters, i.e.

$$G = G(\boldsymbol{\alpha}) = (\mathbf{g}_1(\boldsymbol{\alpha}), \dots, \mathbf{g}_{\tilde{n}}(\boldsymbol{\alpha})). \quad (\text{C1})$$

In either case, the new ansatz is given by $A_G(\mathbf{c}_G) = A(G\mathbf{c}_G)$. This also transforms the constraint matrices via

$$M_G^H = M^H G, \quad M_G^D(\mathbf{d}) = M^D(\mathbf{d}) G, \quad (\text{C2})$$

where the columns of the new constraint matrices are obtained as linear combinations of the columns of the old constraint matrices. The parametrized reconstructed parameters can be obtained as solutions of

$$\mathbf{c}_G^{\text{rec}} = \arg \min_{\mathbf{c}_G, \|\mathbf{c}_G\|=1} \left[\min_{\mathbf{d} \geq 0} \|[M_G^H + M_G^D(\mathbf{d})] \mathbf{c}_G\| \right], \quad (\text{C3})$$

and similarly for equation (A4). In the case where G depends on non-linear parameters $\boldsymbol{\alpha}$, the above optimization also includes a minimization over $\boldsymbol{\alpha}$. Numerically, the optimal $\boldsymbol{\alpha}$ can be found similarly to the optimal dissipation rates \mathbf{d}^{rec} in equation (14), using the DIRECT algorithm in SciPy.

We wish to emphasize that by using this way of parametrizing an ansatz, one obtains reconstructed parameters, where the dependencies, encoded in the matrix G , are exactly fulfilled. Examples for such a reparametrization include, for instance, disregarding operators from the operator content of $A(\mathbf{c})$. This can be understood as a reparametrization, where $(\mathbf{g}_i)_k = 0$, for all $i \in [0, \tilde{n}]$, and $k \in [0, n]$ for which we want to remove the corresponding operator h_k from the operator content of $A(\mathbf{c})$.

Instead of imposing an exact parametrization on an ansatz, one may only impose it approximately. In practice, this may be very useful as parametrizations are almost never exactly fulfilled, but only to a very good approximation. To this end, one adds a penalty term to the cost function in equation (14), which acts as a regularizing term, giving preference to solutions approximately admitting a certain parametrization. In case of a parametrization G , as defined above, the corresponding optimization problem reads

$$\mathbf{c}^{\text{rec}}(\beta) = \arg \min_{\mathbf{c}, \|\mathbf{c}\|=1} \left[\min_{\mathbf{d} \geq 0} \|[M_H + M_D(\mathbf{d})]\mathbf{c}\| + \beta \|\mathbb{1} - GG^T\| \mathbf{c} \right], \quad (\text{C4})$$

where $\beta \in [0, \infty)$ is the regularization strength. For $\beta = 0$ this corresponds to the original unconstrained (i.e. unparametrized) problem in equation (14). For non-zero β the last term adds a penalty, whenever \mathbf{c} has a component outside of the range of GG^T , where the parametrization implied by G is exactly fulfilled. The larger β , the more \mathbf{c}^{rec} is constrained by the parametrization G . In the limit $\beta \rightarrow \infty$ the parametrization G is fulfilled exactly, and the optimization problem corresponds to the one in equation (C3). This can be seen as follows; the objective function in equation (C3) can be rewritten as $\|[M_H + M_D(\mathbf{d})]GG^T \mathbf{c}\|$, where the feasible region consists of all \mathbf{c} for which $\mathbf{c} = GG^T \mathbf{c}$. This in turn can be written as an unconstrained problem

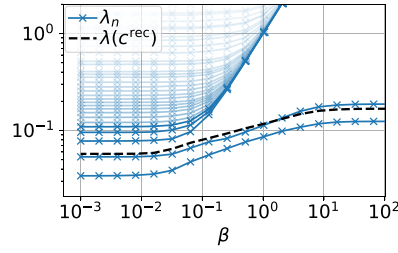


Figure 6. Singular value spectrum of $M_G(\beta)$ defined in equation (C6) as a function of the regularization strength β for the model introduced in section 3.3 in the main text. Increasing β opens a gap between the image of the parametrization matrix G introduced in appendix C and the rest of the spectrum. Within this subspace, $\lambda(c^{\text{rec}})$, defined in equation (C7), follows the best approximation to the Hamiltonian of the system (dashed line).

in equation (C4), where $\beta \rightarrow \infty$. Note, that solving the minimization problem in equation (C4) is equivalent to solving the linear problem

$$c^{\text{rec}}(\beta) = \arg \min_{c, \|c\|=1} \left[\min_{d \geq 0} \left\| \begin{pmatrix} M_H + M_D(d) \\ \beta(\mathbb{1} - GG^T) \end{pmatrix} c \right\| \right]. \quad (\text{C5})$$

Finally, one notices that given that $GG^T c^H = c^H$, i.e. the parametrization does not render an ansatz insufficient, the bound in equation (B2) can only tighten. To see this, one observes that $\lambda_2[M_G(d)] \geq \lambda_2[M(d)]$, and $\|E(d)G\| \leq \|E(d)\| \|G\| = \|E(d)\|$ for any isometry G . Therefore, reducing the number of parameters in general increases the gap $\delta = \lambda_2$ of the constraint matrix $M(d)$.

We now want to study the effect of β on the singular value spectrum of the matrix

$$M_G(\beta) = \begin{pmatrix} M_H + M_D(d^{\text{rec}}) \\ \beta(\mathbb{1} - GG^T) \end{pmatrix}, \quad (\text{C6})$$

for the model system in section 3.3 (see figure 6(a)). As expected, in the limit $\beta \rightarrow 0$, the spectrum converges to the spectrum of the unparametrized constraint matrix $M(d^{\text{rec}})$. Here, the gap between the lowest singular values and the rest of the spectrum becomes very small, resulting in an unstable solution. Then, when increasing β , all singular values, whose corresponding right-singular vectors are incompatible with the regularization increase with β , opening a gap to the subspace spanned by the regularization, i.e. the image of G , which in the case of the ansatz in equation (27) is two-dimensional. Note, that the Hamiltonian in equation (22) only approximately lies in the image of G . Therefore, also the lowest singular values initially increase with β , until they reach a constant value, that corresponds to an exactly enforced parametrization. Along this path, we can monitor the minimum of the cost function in equation (C5) defined via

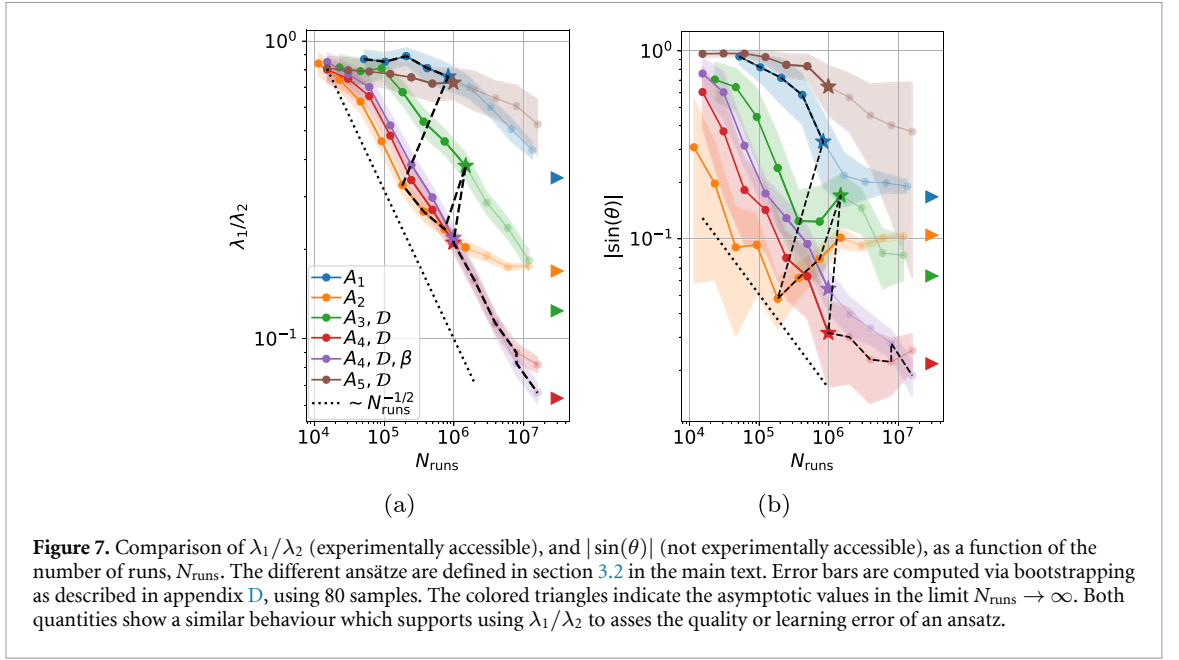
$$\lambda(c^{\text{rec}}) = \left\| \begin{pmatrix} M_H + M_D(d^{\text{rec}}) \\ \beta(\mathbb{1} - GG^T) \end{pmatrix} c^{\text{rec}} \right\|, \quad (\text{C7})$$

which follows the singular value that corresponds to the best approximation of the Hamiltonian.

Appendix D. Bootstrapping

To obtain error bars from data we use the Bootstrapping method, which we will introduce below. Assume we are given a single realization x_1, \dots, x_n of a set of independent and identically distributed random variables X_1, \dots, X_n with unknown distribution function. We are interested in estimating the variance of a given function $T(X_1, \dots, X_n)$. To do so we draw n times with replacement from x_1, \dots, x_n , yielding a sample x_1^*, \dots, x_n^* , and then evaluate $t = T(x_1^*, \dots, x_n^*)$. We repeat this procedure r times obtaining $t^{(1)}, \dots, t^{(r)}$, and then estimate the variance of T from the sample variance of $t^{(1)}, \dots, t^{(r)}$.

In our case the X_1, \dots, X_n are the individual measurements in a given basis for fixed initial state and simulation time. One can also think of X_1, \dots, X_n to be individual measurements of a given observable for fixed initial state and simulation time, in the case where the measurements of different observables are independent. Then the quantities of interest $T(X_1, \dots, X_n)$ are, e.g. the learned parameters c^{rec} , or the ratio of singular values λ_1/λ_2 . The number of samples r is chosen the minimum possible integer, such that the error bars do not significantly change anymore when further increasing r .



Appendix E. Comparison between λ_1/λ_2 and $|\sin(\theta)|$

In figure 7 we provide a comparison between the experimentally measurable quantity λ_1/λ_2 , which we use to assess the quality of an ansatz, and the experimentally non-accessible angle between the reconstructed parameters, \mathbf{c}^{rec} , and the Hamiltonian parameters, \mathbf{c}^H , and a function of the number of runs, N_{runs} . One observes that both quantities show a very similar behaviour over almost the entire range of N_{runs} considered here. In particular, in the limit $N_{\text{runs}} \rightarrow \infty$ a larger value of λ_1/λ_2 corresponds to larger $|\sin(\theta)|$, and vice versa. This supports our idea to use λ_1/λ_2 as an experimentally accessible quantity to assess the quality and learning error of a given ansatz.

Appendix F. Learning from generalized energy conservation and additional constraints

In the following we compare the learning based on the generalized Ehrenfest theorem, presented in section 3.3 in the main text, to the learning based on energy conservation. We consider the same model systems comprising the long-range Hamiltonian H_{XY} as well as local and collective dissipation, introduced in section 3.3 in the main text.

We will use equation (27) as an ansatz for the Hamiltonian. As we also want to learn the dissipative processes, we will start with an ansatz for the Liouvillian, comprising single-qubit Lindblad operators that typically appear in the context of trapped-ion experiments, i.e.

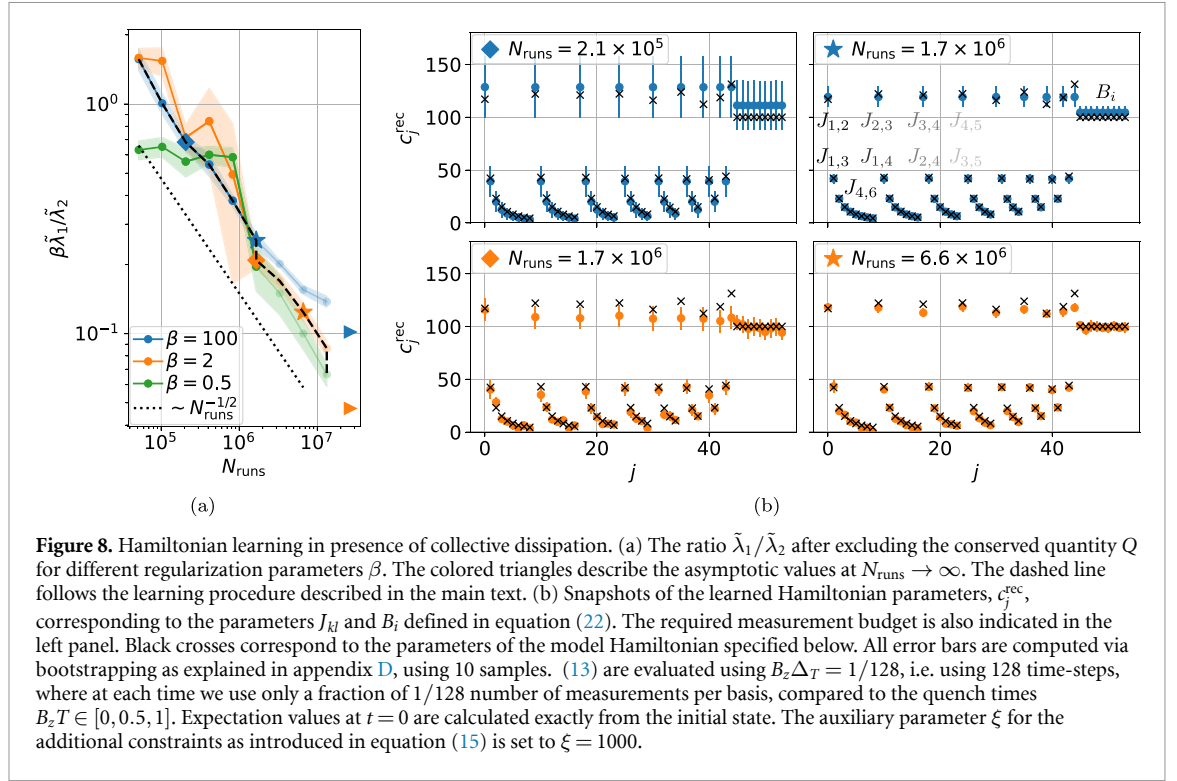
$$\mathcal{D}^{\text{loc}}(\varrho) = \sum_{\eta \in \{+, -, z\}} \frac{d_\eta}{2} \sum_{k=1}^N \left(\left[\sigma_k^\eta \varrho, \sigma_k^{\eta, \dagger} \right] + \text{h.c.} \right). \quad (\text{F1})$$

Note that this ansatz does not include collective dissipation as given by the off-diagonal elements of Γ_{kl}^z in the model Liouvillian. Nevertheless, we will find, that the ansatz in equation (F1) is sufficient for learning the Hamiltonian, and we will discuss in detail why this is the case.

Firstly, one notices that the ansatz A contains the total magnetization defined by

$$Q = B \sum_k \sigma_k^z, \quad (\text{F2})$$

which commutes with H_{XY} for any $B \in \mathbb{R}$. Therefore, Q is a conserved quantity of \mathcal{L}_H . This leads to a degeneracy of the singular value spectrum of the constraint matrix $M(\mathbf{d})$, i.e. we observe, that $\lambda_1 \approx \lambda_2$. In particular, this means that the reconstructed parameters \mathbf{c}^{rec} will be a linear combination of the parameters of H_{XY} and Q . To reconstruct the parameters of the Hamiltonian, including its overall-scale, we need to exclude Q , as well as scalar multiples of the Hamiltonian, $\nu \cdot H_{XY}$, as possible solutions for \mathbf{c}^{rec} . This is achieved by imposing additional constraints as discussed in section 2.3. To this end, we choose a set of operators, $\mathcal{O} = \{\sigma_1^x, \sigma_1^y, \sigma_1^z\}$, where some operators do not commute with Q , while others do not commute with H_{XY} .



Here, one could in principle also choose random operators. This leads to a combined linear optimization as in equation (15) [see also appendix A for the detailed structure of the additional constraints]. Moreover, one can define a projected constraint matrix with non-degenerate spectrum $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots$, where $c^{\text{rec}} \approx c^H$ is the right-singular vector corresponding to $\tilde{\lambda}_1$. Then, $\tilde{\lambda}_1 / \tilde{\lambda}_2$ becomes the analog of λ_1 / λ_2 ¹³.

In figure 8(a) we monitor $\tilde{\lambda}_1 / \tilde{\lambda}_2$ as a function of N_{runs} for the ansatz A, and for different values of the regularization parameter β . One starts with a large value of the regularization parameter, here, $\beta = 100$, which strongly imposes the parametrization of A , leading to reconstructed parameters without spatial disorder, see figure 8(b) for $N_{\text{runs}} = 2.1 \times 10^5$ (blue data). Then further increasing N_{runs} reduces the size of the error bars in c^{rec} , as is also shown in figure 8(b) for $N_{\text{runs}} = 1.7 \times 10^6$ (blue data). However, an ansatz with $\beta = 100$ cannot account for disorder in the coupling terms in H_{XY} , which is shown by the plateau in $\tilde{\lambda}_1 / \tilde{\lambda}_2$, here, starting at around $N_{\text{runs}} \approx 10^7$ (blue line), which, again, is above our measurement budget. Nevertheless, we decrease the regularization parameter β , which leads to larger error bars, but also allows to learn some of the spatial disorder in H_{XY} , as is shown in figure 8(b) for $\beta = 2$ and $N_{\text{runs}} = 1.7 \times 10^6$ (orange data). Note that already before reaching a plateau for $\beta = 2$ in figure 8(a) (orange line) the error bars of c^{rec} become very small, which can be seen in figure 8(b) for $\beta = 2$ and $N_{\text{runs}} = 6.6 \times 10^6$ (orange data). This suggest to further reduce β . Ultimately, this process of subsequently reducing β and increasing N_{runs} allows us to learn H_{XY} up to statistical errors, and leads to $\tilde{\lambda}_1 / \tilde{\lambda}_2 \sim N_{\text{runs}}^{-1/2}$ for the entire range of N_{runs} .

So far we have used an ansatz for the Liouvillian which does not include the collective dephasing terms in equation (25). Nevertheless, we can learn the Hamiltonian to a high accuracy, which constitutes a useful feature of our approach. To see why this is the case, we evaluate the matrix $M_D(\mathbf{d})$ by summing equation (13) over all Lindblad operators, yielding

$$[M_D(\mathbf{d})]_{\kappa, (i,j)} = \frac{1}{2} \sum_{k,l} d_{kl} [M^{(k,l)}]_{\kappa, (i,j)} = \frac{1}{2} \int_0^T \sum_{k,l} d_{kl} \langle \sigma_k^z [h_{ij}, \sigma_l^z] + \text{h.c.} \rangle_{\kappa, t} dt, \quad (\text{F3})$$

where κ labels input states. Here, we only include the interaction terms, $h_{ij} = \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y$, of the Hamiltonian H_{XY} , since the contribution of the collective dissipation in equation (25) vanishes for the field

¹³ In the case of two linearly independent conserved quantities our figure of merit $\lambda_1 / \lambda_2 \approx 1$ independent of N_{runs} . Nevertheless, we can define a new matrix $\tilde{M}(\mathbf{d})$, by projecting the kernel of $M(\mathbf{d})$, spanned by all vectors \mathbf{c} that belong to conserved quantities of H , onto the vector of reconstructed parameters of the Hamiltonian c^{rec} . The spectrum $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots$ of $\tilde{M}(\mathbf{d})$ is gapped, and $\tilde{\lambda}_1 / \tilde{\lambda}_2$ becomes the analogue of λ_1 / λ_2 . For more than two linearly independent conserved quantities one proceeds similarly.

terms σ_i^z . Considering the sum over operators appearing in the expectation value in equation (F3), we can split this sum into

$$\begin{aligned} & \sum_{k=1}^N d_{kk} (\sigma_k^z [h_{ij}, \sigma_k^z] + \text{h.c.}) + \sum_{k \neq l}^N d_{kl} (\sigma_k^z [h_{ij}, \sigma_l^z] + \text{h.c.}) \\ &= -4 \sum_{k=1}^N d_{kk} (\delta_{ki} + \delta_{kj}) h_{ij} + 4 \sum_{k \neq l}^N d_{kl} (\delta_{ki} \delta_{lj} + \delta_{kj} \delta_{li}) h_{ij}, \end{aligned}$$

which vanishes for $d_{kl} = d_0$. Therefore, the constraint matrix in equation (11) and hence also the reconstructed parameters of the Hamiltonian are not affected by including collective dissipation into our ansatz. However, this will change if we include additional constraints, as we demonstrate in the following.

F.1. Learning collective dissipation

To discover the collective dissipation given by equation (25) we choose the following ansatz for the Liouvillian

$$\mathcal{D}^{\text{col}}(\varrho) = \frac{d_-}{2} \sum_{k=1}^N \left([\sigma_k^- \varrho, \sigma_k^{-\dagger}] + \text{h.c.} \right) + \sum_{k,l=1}^N (d_z \delta_{kl} + d_z^{\text{col}} (1 - \delta_{kl})) \mathcal{D}_{kl}^z(\varrho), \quad (\text{F4})$$

that includes single-qubit spontaneous decay with jump operator σ_k^- , and single-qubit as well as collective dephasing \mathcal{D}_{kl}^z with operator structure as described in equation (25). Learning collective dissipation requires the set of constraint operators to also include two-qubit terms. This is because single-qubit operators are not affected by the multi-qubit Lindblad operators in equation (25). Therefore, we may choose the following set of constraint operators

$$\mathcal{O} = \{ \sigma_1^x, \sigma_1^y, \sigma_1^z, \sigma_2^x, \sigma_2^y, \sigma_2^z, \sigma_1^x \sigma_2^x, \sigma_1^x \sigma_2^y, \sigma_1^y \sigma_2^x, \sigma_1^y \sigma_2^y \}.$$

Note that we included here the local operators despite the fact that they cannot help to learn the collective dissipation. This is due to the facts that their expectation values can be obtained from post-processing the data obtained from measuring the two-body operators and that they can be used to better learn the overall scale and the local dissipation rates, as discussed above. Let us emphasize, that the ratio λ_1/λ_2 is not affected by taking collective dissipation in the ansatz into account. Therefore, we instead use

$$\Delta^{\text{add}} := \|M^{\text{add}} \mathbf{c}^{\text{rec}} - \mathbf{b}(\mathbf{d}^{\text{rec}})\|, \quad (\text{F5})$$

as a measure of how well the additional constraints are fulfilled. In case an ansatz for the Liouvillian is insufficient, not any possible additional constraint can be exactly fulfilled. In such a case, Δ^{add} will be bounded away from zero in the limit $N_{\text{runs}} \rightarrow \infty$. However, in practice, with only a few additional constraints, e.g. the local Pauli operators discussed above, this is not necessarily the case.

In figure 9(a) we monitor Δ^{add} as a function of N_{runs} for the Hamiltonian ansatz in equation (27) with $\beta = 2$, for different ansätze for the Liouvillian. For an ansatz that does not include dissipation (blue line in figure 9(a)) we observe a plateau of Δ^{add} already at very early stages of the learning procedure, here at $N_{\text{runs}} \approx 10^5$. On the other hand, we find that the ansätze containing local (orange line) and collective (green line) dissipation lead to similar values of Δ^{add} in early states of the learning procedure. Only above $N_{\text{runs}} \approx 10^7$ the insufficiency of the local ansatz equation (F1) would become evident from considering Δ^{add} , which is beyond our available measurement budget. Nevertheless, we can study the reconstructed parameters shown in figure 9(b) and their corresponding error bars. For the ansatz with local dissipation (orange data) we find that the learned rates converge to the exact values, with a good reconstruction achieved at around $N_{\text{runs}} \approx 10^6$. Note that here the learned dephasing rate is the sum $d_z + d_z^{\text{col}}$. For the ansatz including collective dissipation (green data) we obtain similarly accurate values for γ_- . However, the local and global dephasing rates, d_z and d_z^{col} respectively, have larger error bars. Only around $N_{\text{runs}} \approx 10^7$ we observe the emergence of non-zero dissipation rates for collective dephasing. This concludes the learning procedure of the Hamiltonian and Liouvillian.

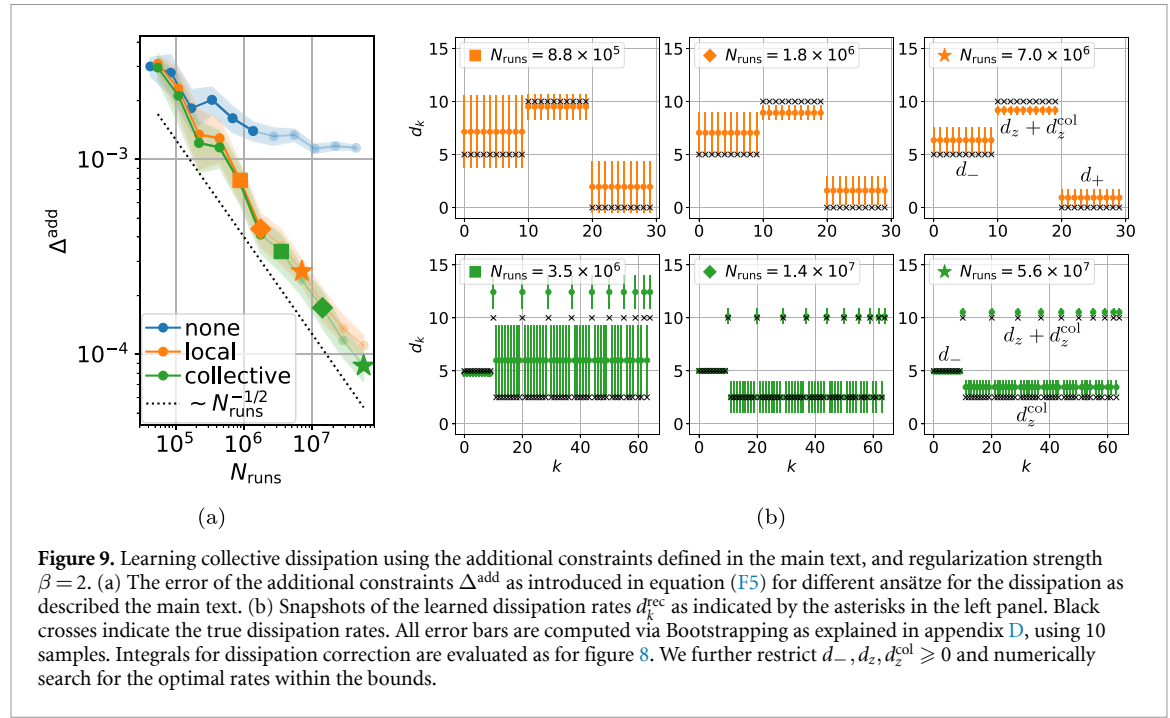


Figure 9. Learning collective dissipation using the additional constraints defined in the main text, and regularization strength $\beta = 2$. (a) The error of the additional constraints Δ^{add} as introduced in equation (F5) for different ansätze for the dissipation as described in the main text. (b) Snapshots of the learned dissipation rates d_k^{rec} as indicated by the asterisks in the left panel. Black crosses indicate the true dissipation rates. All error bars are computed via Bootstrapping as explained in appendix D, using 10 samples. Integrals for dissipation correction are evaluated as for figure 8. We further restrict d_- , d_z , $d_z^{\text{col}} \geq 0$ and numerically search for the optimal rates within the bounds.

ORCID iDs

Tobias Olsacher <https://orcid.org/0000-0002-2307-4657>

Tristan Kraft <https://orcid.org/0000-0002-5268-5377>

Barbara Kraus <https://orcid.org/0000-0001-7246-6385>

Peter Zoller <https://orcid.org/0000-0003-4014-1505>

References

- [1] Altman E *et al* 2021 Quantum simulators: architectures and opportunities *PRX Quantum* **2** 017003
- [2] Tarruell L and Sanchez-Palencia L 2018 Quantum simulation of the Hubbard model with ultracold fermions in optical lattices *C. R. Phys.* **19** 365
- [3] Bauer C W *et al* 2023 Quantum simulation for high-energy physics *PRX Quantum* **4** 027001
- [4] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S C and Yuan X 2020 Quantum computational chemistry *Rev. Mod. Phys.* **92** 015003
- [5] Gross C and Bloch I 2017 Quantum simulations with ultracold atoms in optical lattices *Science* **357** 995
- [6] Sompet P *et al* 2022 Realizing the symmetry-protected Haldane phase in Fermi–Hubbard ladders *Nature* **606** 484
- [7] Léonard J, Kim S, Kwan J, Segura P, Grusdt F, Repellin C, Goldman N and Greiner M 2023 Realization of a fractional quantum Hall state with ultracold atoms *Nature* **619** 495
- [8] Zhang W-Y *et al* 2023 Scalable multipartite entanglement created by spin exchange in an optical lattice *Phys. Rev. Lett.* **131** 073401
- [9] Monroe C *et al* 2021 Programmable quantum simulations of spin systems with trapped ions *Rev. Mod. Phys.* **93** 025001
- [10] Joshi M K, Kokail C, van Bijnen R, Kranzl F, Zache T V, Blatt R, Roos C F and Zoller P 2023 Exploring large-scale entanglement in quantum simulation *Nature* **624** 539
- [11] Kiesenhofer D, Hainzer H, Zhdanov A, Holz P C, Bock M, Ollikainen T and Roos C F 2023 Controlling two-dimensional Coulomb crystals of more than 100 ions in a monolithic radio-frequency trap *PRX Quantum* **4** 020317
- [12] Guo S A *et al* 2024 A site-resolved 2d quantum simulator with hundreds of trapped ions (arXiv:2311.17163 [quant-ph])
- [13] Weimer H, Müller M, Lesanovsky I, Zoller P and Büchler H P 2010 A Rydberg quantum simulator *Nat. Phys.* **6** 382
- [14] Semeghini G *et al* 2021 Probing topological spin liquids on a programmable quantum simulator *Science* **374** 1242
- [15] Ebadi S *et al* 2021 Quantum phases of matter on a 256-atom programmable quantum simulator *Nature* **595** 227
- [16] Chen C *et al* 2023 Continuous symmetry breaking in a two-dimensional Rydberg array *Nature* **616** 691
- [17] Kim Y *et al* 2023 Evidence for the utility of quantum computing before fault tolerance *Nature* **618** 500
- [18] Tindall J, Fishman M, Stoudenmire E M and Sels D 2024 Efficient tensor network simulation of IBM's eagle kicked ising experiment *PRX Quantum* **5** 010308
- [19] Rosenberg E *et al* 2024 Dynamics of magnetization at infinite temperature in a Heisenberg spin chain *Science* **384** 48
- [20] 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
- [21] Trivedi R, Rubio A F and Cirac J I 2022 Quantum advantage and stability to errors in analogue quantum simulators (arXiv:2212.04924 [quant-ph])
- [22] Kashyap V, Styliaris G, Mouradian S, Cirac J I and Trivedi R 2024 Accuracy guarantees and quantum advantage in analogue open quantum simulation with and without noise (arXiv:2404.11081 [quant-ph])
- [23] Cai Y, Tong Y and Preskill J 2023 Stochastic error cancellation in analog quantum simulation (arXiv:2311.14818 [quant-ph])

- [24] Daley A J, Bloch I, Kokail C, Flannigan S, Pearson N, Troyer M and Zoller P 2022 Practical quantum advantage in quantum simulation *Nature* **607** 667
- [25] Hangleiter D, Kliesch M, Schwarz M and Eisert J 2017 Direct certification of a class of quantum simulations *Quantum Sci. Technol.* **2** 015004
- [26] 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
- [27] Carrasco J, Elben A, Kokail C, Kraus B and Zoller P 2021 Theoretical and experimental perspectives of quantum verification *PRX Quantum* **2** 010102
- [28] Wiebe N, Granade C, Ferrie C and Cory D G 2014 Hamiltonian learning and certification using quantum resources *Phys. Rev. Lett.* **112** 190501
- [29] Holzäpfel M, Baumgratz T, Cramer M and Plenio M B 2015 Scalable reconstruction of unitary processes and hamiltonians *Phys. Rev. A* **91** 042129
- [30] Wang S-T, Deng D-L and Duan L-M 2015 Hamiltonian tomography for quantum many-body systems with arbitrary couplings *New J. Phys.* **17** 093017
- [31] Evans T J, Harper R and Flammia S T 2019 Scalable Bayesian Hamiltonian learning (arXiv:1912.07636 [quant-ph])
- [32] Bairey E, Arad I and Lindner N H 2019 Learning a local Hamiltonian from local measurements *Phys. Rev. Lett.* **122** 020504
- [33] 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
- [34] Gu A, Cincio L and Coles P J 2024 Practical Hamiltonian learning with unitary dynamics and Gibbs states *Nat. Commun.* **15** 312
- [35] Haah J, Kothari R and Tang E 2024 Learning quantum Hamiltonians from high-temperature Gibbs states and real-time evolutions *Nat. Phys.* **20** 1027–31
- [36] Li Z, Zou L and Hsieh T H 2020 Hamiltonian tomography via quantum quench *Phys. Rev. Lett.* **124** 160502
- [37] Ott R, Zache T V, Prüfer M, Erne S, Tajik M, Pichler H, Schmiedmayer J and Zoller P 2024 Hamiltonian learning in quantum field theories (arXiv:2401.01308 [cond-mat.quant-gas])
- [38] Zubida A, Yitzhaki E, Lindner N H and Bairey E 2021 Optimal short-time measurements for Hamiltonian learning (arXiv:2108.08824 [quant-ph])
- [39] Stilck França D, Markovich L A, Dobrovitski V V, Werner A H and Borregaard J 2024 Efficient and robust estimation of many-qubit Hamiltonians *Nat. Commun.* **15** 311
- [40] Hangleiter D, Roth I, Fuksa J, Eisert J and Roushan P 2021 Robustly learning the hamiltonian dynamics of a superconducting quantum processor (arXiv:2108.08319 [quant-ph])
- [41] Huang H-Y, Tong Y, Fang D and Su Y 2023 Learning many-body Hamiltonians with Heisenberg-limited scaling *Phys. Rev. Lett.* **130** 200403
- [42] Yu W, Sun J, Han Z and Yuan X 2023 Robust and efficient Hamiltonian learning *Quantum* **7** 1045
- [43] Mehta P, Bukov M, Wang C-H, Day A G, Richardson C, Fisher C K and Schwab D J 2019 A high-bias, low-variance introduction to machine learning for physicists *Phys. Rep.* **810** 1
- [44] Gardiner C and Zoller P 2004 *Quantum noise: a handbook of markovian and non-markovian quantum stochastic methods with applications to quantum optics* (Springer)
- [45] Gorini V, Kossakowski A and Sudarshan E C G 1976 Completely positive dynamical semigroups of n-level systems *J. Math. Phys.* **17** 821
- [46] Elben A, Flammia S T, Huang H-Y, Kueng R, Preskill J, Vermersch B and Zoller P 2022 The randomized measurement toolbox *Nat. Rev. Phys.* **5** 9
- [47] Chelpanova O, Kelly S P, Schmidt-Kaler F, Morigi G and Marino J 2024 Dynamics of quantum discommensurations in the Frenkel-Kontorova chain (arXiv:2401.12614 [cond-mat.stat-mech])
- [48] García-Mata I, Zhirov O V and Shepelyansky D L 2007 Frenkel-Kontorova model with cold trapped ions *Eur. Phys. J. D* **41** 325
- [49] Liu Z, Devulapalli D, Hangleiter D, Liu Y-K, Kollár A J, Gorshkov A V and Childs A M 2024 Efficiently verifiable quantum advantage on near-term analog quantum simulators (arXiv:2403.08195 [quant-ph])
- [50] Davis C and Kahan W M 1970 The rotation of eigenvectors by a perturbation. III *SIAM J. Numer. Anal.* **7** 1
- [51] Wedin P-Å 1972 Perturbation bounds in connection with singular value decomposition *BIT Numer. Math.* **12** 99