Learning a Local Hamiltonian from Local Measurements

Eyal Bairey,[1] Itai Arad,[1] and Netanel H. Lindner[1]
[1] *Physics Department, Technion, 3200003, Haifa, Israel*

Recovering an unknown Hamiltonian from measurements is an increasingly important task for certification of noisy quantum devices and simulators. Recent works have succeeded in recovering the Hamiltonian of an isolated quantum system with local interactions from long-ranged correlators of a single eigenstate. Here, we show that such Hamiltonians can be recovered from local observables alone, using computational and measurement resources scaling linearly with the system size. In fact, to recover the Hamiltonian acting on each finite spatial domain, only observables within that domain are required. The observables can be measured in a Gibbs state as well as a single eigenstate; furthermore, they can be measured in a state evolved by the Hamiltonian for a long time, allowing to recover a large family of time-dependent Hamiltonians. We derive an estimate for the statistical recovery error due to approximation of expectation values using a finite number of samples, which agrees well with numerical simulations.

Introduction. Contemporary condensed matter physics has witnessed great advancements in tools developed to obtain the state of a system given its Hamiltonian. As quantum devices are being rapidly developed, the converse task of recovering the Hamiltonian of a many-body system from measured observables is becoming increasingly important. In particular, it is a necessary step for certifying quantum simulators and devices containing many qubits. As these expand beyond the power of classical devices [1], there is a growing need to certify them using only a polynomial amount of classical computational resources as well as quantum measurements.

Various methods have been suggested for recovering a Hamiltonian based on its dynamics [2–8] or Gibbs state [9–11]. The system-size scaling of the recovery efficiency can be improved using a trusted quantum simulator [12–16], manipulations of the investigated system [17], or accurate measurements of short-time dynamics [18, 19].

Here, we suggest a framework for recovering a generic local Hamiltonian using only polynomial time and measurements. Inspired by the recently introduced method for recovering a local Hamiltonian from measurements on a single eigenstate [20–22], our framework offers four main contributions. First, we generalize to mixed states such as Gibbs states $\rho = \frac{1}{Z}e^{-\beta H}$, treating any state which commutes with the Hamiltonian at the same footing as an eigenstate. Second, our method can be applied to dynamics of arbitrary low-energy initial states time-evolved by the Hamiltonian. Third, it allows to recover time-dependent Hamiltonians if the functional form of their time-dependence is known. Finally, in the case of short-range interactions, we can infer the Hamiltonian of a local patch L based only on local measurements inside L. This implies that a short-ranged Hamiltonian on a large system can be obtained with a number of measurements and computation time *linear* in system size.

Problem setting. We wish to recover the Hamiltonian acting on a region L by measuring observables only in L. We would first like to make these notions precise.

We consider a Hamiltonian H on a finite lattice Λ in d dimensions:

$$H = \sum_i h_i. \tag{1}$$

We assume that H is k-local, such that each h_i acts non-trivially on no more than k spatially contiguous sites (i.e., contained within a ball of diameter k). We focus on a specific subset of sites $L \subseteq \Lambda$. We define its interior $L_0 \subseteq L$ as the sites that are not conected by H to sites outside L (Fig. 1). We denote by H_L the subset of h_i terms in H that act non-trivially on L_0.

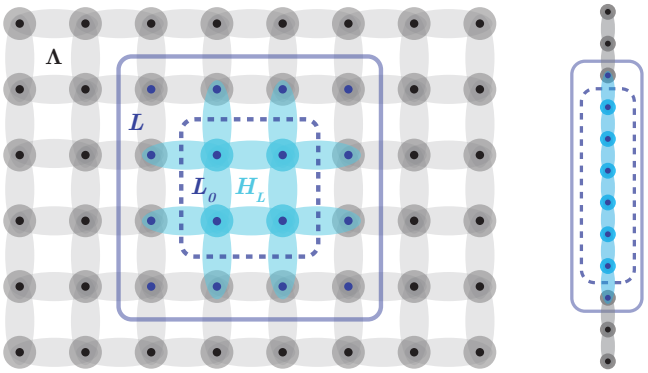

FIG. 1. Recovery from local measurements. Left: our method recovers H_L (light blue), using only measurements of observables residing in L (solid blue line), a sub-region of the whole system Λ. The interior $L_0 \subseteq L$ (dashed blue line) consists of sites interacting only within L. Right: our simulations are performed on $|\Lambda| = 12$ chains, recovering H_L on the 8 middle spins.

We call any state ρ that is stationary under H a *steady state* (taking $\hbar = 1$):

$$i\partial_t \rho = [H, \rho] = 0, \tag{2}$$

In particular, ρ can be any eigenstate as well as a Gibbs state. Our goal is to recover H_L from a steady state of H, based only on measurements in L.

Algorithm. To recover H_L, we identify a set of local constraints on H_L obeyed by any steady state ρ of H.

Since ρ is stationary under H, so is the expectation value $\langle A\rangle \stackrel{\text{def}}{=} \mathrm{Tr}\,(\rho A)$ of any operator A in the state ρ, so that $\partial_t\langle A\rangle = -\langle i[A,H]\rangle = 0$. If A is supported only on L_0, this constraint becomes

$$\langle i[A, H_L]\rangle = 0, \tag{3}$$

since A trivially commutes with $H - H_L$.

The k-local operators acting on L_0 form a linear space. We choose a basis $\{S_m\}_{m=1}^M$ for this space of operators, where M is its dimension. When we expand H_L in this basis,

$$H_L = \sum_{m=1}^{M} c_m S_m, \tag{4}$$

the constraint (3) becomes a linear homogeneous constraint on the vector $\vec{c} = (c_1, c_2, \ldots, c_M)$:

$$\sum_{m=1}^{M} c_m \langle i[A, S_m]\rangle = 0. \tag{5}$$

Using a set of operators $\{A_n\}_{n=1}^N$, each supported on L_0, we obtain a set of N linear constraints:

$$\forall n : \sum_{m=1}^{M} c_m \langle i\,[A_n, S_m]\rangle = 0, \tag{6}$$

which is equivalent to the $N \times M$ real linear equation

$$K\vec{c} = 0, \qquad K_{n,m} \stackrel{\text{def}}{=} \langle i[A_n, S_m]\rangle. \tag{7}$$

The number M of basis elements S_m that span H_L is linear in the subsystem's volume $|L|$. In contrast, the maximal number of constraints scales like the number of linearly independent observables A_n in L_0, which grows exponentially with $|L_0|$. Thus, for a sufficiently large but constant region L (depending on k but not on $|\Lambda|$), we can always have more equations than unknowns, i.e., $N > M$. As argued in Ref. [20], we expect these equations to be generally independent, thereby providing a unique solution $\vec{c}$ up to an overall scale.

Given a region L whose Hamiltonian we wish to learn, our method is therefore as follows:

1. Identify a set of terms $\{S_m\}_{m=1}^M$ spanning the space of possible H_L's.

2. Construct a constraint matrix $K_{N\times M}$ by measuring $\langle i[A_n, S_m]\rangle$ with respect to a set of constraints $\{A_n\}_{n=1}^N$ supported on L_0.

3. Estimate $H_L \propto \sum_{m=1}^M c_m S_m$, with $\vec{c}$ the lowest right-singular vector of K.

The lowest right-singular vector of K is the numerical solution to Eq. (7), the vector that minimizes $\|Kc\|$. Namely, it is the ground-state of the *correlation matrix*,

$$\mathcal{M} = K^T K. \tag{8}$$

Extension to a dynamical setting. So far, we have described how to recover a time-independent H from measurements of its steady state. However, many experimental settings do not have access to an exact steady state of H. Instead, we now describe how to obtain an approximate steady-state from an arbitrary initial state by evolving it with H for long times.

In the dynamical approach, we repeatedly initialize our system in some state $\rho(0)$. We let it evolve for a random time distributed uniformly in $0 \le t' \le t$, before measuring an operator A. The average outcome of these measurements is given by $\mathrm{Tr}\,(\rho_{avg}A)$, where $\rho_{avg} = \frac{1}{t}\int_{t'=0}^{t}\rho(t')dt'$. For a time-independent H, this time-averaged density matrix approaches a steady state in trace norm, since by integrating (2), we obtain:

$$\|[\rho_{avg}, H]\|_1 = \frac{1}{t}\|\rho(t) - \rho(0)\|_1 \le \frac{2}{t}. \tag{9}$$

This allows to recover a time-independent H from a constraint matrix K of time-averaged observables.

The dynamical approach can be extended to time-dependent Hamiltonians of the form:

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{V}f(t), \tag{10}$$

where $f(t)$ is a known function. Similarly to Eq. (9), now the time-averaged commutator $\frac{1}{t}\int_0^t[\rho(t'), \hat{H}(t')]dt'$ must decay with time. Therefore, we estimate the coefficients of $\hat{H}^{(0)}$, $\hat{V}$ as the lowest singular vector of an extended constraint matrix $K_{N\times 2M}$ composed of time-averaged as well as time-modulated measurements (see Appendix A):

$$\begin{aligned}\forall m \le M:\ K_{n,m} &= \frac{1}{t}\int_0^t \langle[A_n, S_m]\rangle dt' \\ K_{n,m+M} &= \frac{1}{t}\int_0^t \langle[A_n, S_m]\rangle f(t')dt'\end{aligned} \tag{11}$$

a. Sample complexity. The complexity of our method depends on the number of observables we need to measure and on the accuracy to which we need to measure each of them. Experimentally, each observable $\langle i[A_n, S_m]\rangle$ can only be measured to finite accuracy due to statistical uncertainty in estimating it using a finite number of samples n_s. We quantify the resulting error in the reconstruction process by the l_2 distance between the normalized recovered and true coefficient vectors [23],

$$\Delta = \|\hat{c}_{true} - \hat{c}_{recovered}\|_2\,, \tag{12}$$

where $\hat{c} \stackrel{\text{def}}{=} \frac{\vec{c}}{\|\vec{c}\|}$.

Following Ref. [20], we analyze the reconstruction error using a simple perturbation theory on the correlation matrix $\mathcal{M}$. We model the error in each entry $K_{n,m}$ obtained by n_s samples as an independent Gaussian with

zero mean and standard deviation $\epsilon \approx n_s^{-1/2}$. To lowest order in ϵ, we estimate the expected error:

$$\mathbb{E}(\Delta) \approx \epsilon \sqrt{\sum_{i>0} \frac{1}{\lambda_i}} \stackrel{\text{def}}{=} \Delta^{est}, \tag{13}$$

where λ_i are the eigenvalues of $\mathcal{M}$ (see Appendix B).

To open a gap in $\mathcal{M}$ between λ_0 and λ_1 and recover a unique Hamiltonian, at least as many constraints N as unknowns M are required. This means measuring $\mathcal{O}(|L|)$ operators $i[A_n, S_m]$, since each constraint A_n commutes with all but a constant number of candidate Hamiltonian terms S_m. Moreover, the Hamiltonian can be reconstructed in linear time in $|L|$ and a linear number of measurements by breaking down L into smaller sub-regions and reconstructing the Hamiltonian on each of them separately. For translationally invariant Hamiltonians a single sub-region is sufficient, with only a constant number of operators to be measured.

Minimizing the support of the measured operators is advantageous for some experimental settings, in which correlations involving multiple sites are hard to measure. Suppose, for example, we wish to recover a generic 2-local H. To obtain more equations than unknowns, we need constraints A_n that act on at least 2 sites. This corresponds to 3-local measurements $i[A_n, S_m]$. Luckily, measurements of only 2-local observables can suffice if a few different steady states are available. These may be Gibbs states at different temperatures, or time-averaged evolutions of different initial conditions. In this setting, each steady state can provide an independent set of constraints. For a 2-local Hamiltonian in one dimension, single-site A_n operators and 5 different steady states can provide sufficient constraints to open a gap in $\mathcal{M}$. More generally, access to multiple steady states allows to recover a k-local H using only k-local measurements.

b. Numerical simulations. To demonstrate the performance of our method, we numerically simulated random one-dimensional spin $\frac{1}{2}$ chains. We considered Hamiltonians consisting of all possible 2-local terms, acting on single spins and nearest neighbors:

$$H = \sum_{l=1}^{|\Lambda|} \sum_{\alpha=1}^{3} c_{l\alpha} \sigma_l^\alpha + \sum_{l=1}^{|\Lambda|-1} \sum_{\alpha=1}^{3} \sum_{\beta=1}^{3} c_{l\alpha\beta} \sigma_l^\alpha \sigma_{l+1}^\beta. \tag{14}$$

In each simulation, we generated a random 2-local Hamiltonian H (Eq. (14)) on $|\Lambda| = 12$ sites by sampling the vector of all coefficients $\vec{c}$ from a Gaussian distribution with zero mean and unit standard deviation, setting the energy scale for what follows. We numerically calculated the ground-state of H, and then recovered H_L from the ground-state in steps. In each step we added one row to the constraint matrix K by choosing a constraint operator A_n and estimating $\{\langle i[A_n, S_m]\rangle\}_{m=1}^M$. Here, A_n is an operator supported on the 6 middle sites L_0, and $\{S_m\}_{m=1}^M$ is the subset of terms in Eq. (14) acting on L_0. To measure the robustness of the reconstruction, we added to the constraint matrix K a noise matrix of independent Gaussian entries with zero mean and standard deviation $\epsilon = 10^{-12}$.

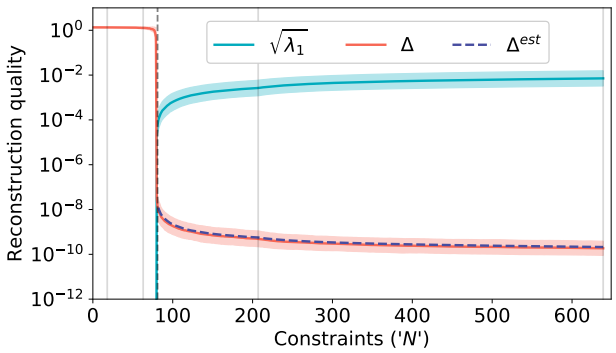

FIG. 2. Quality of Hamiltonian reconstruction as a function of the number of measured constraints N. We generated ground-states of random spin chains (Eq. (14)), and measured local observables $i[A_n, S_m]$ on the 8 middle spins L to recover H_L. When the number N of constraints A_n exceeded the number M of possible Hamiltonian terms S_m (dashed vertical line), $\mathcal{M}$ opened a gap (light blue; $\lambda_0 = 0$ here). The reconstruction error [red, see Eq. (12)] was solely due to the addition of a small Gaussian noise with standard deviation $\epsilon = 10^{-12}$ to each measurement. THe error closely followed an estimate obtained from the spectrum of K [dashed purple, see Eq. (13)]. We used all k-local constraints A_n up to $k = 4$ in an increasing order of support size k. The solid vertical lines denote the transition to $k = 2, 3, 4$ respectively, and within each k we chose the constraints in random order. Results were averaged over 200 random Hamiltonians; the means and standard deviations were calculated after taking the log.

As expected, once sufficiently many constraints had been measured, our procedure recovered the Hamiltonian to high accuracy (Fig. 2). As soon as $N = M - 1$, the correlation matrix $\mathcal{M}$ opened a gap, allowing to recover the coefficient vector $\hat{c}$ given by the ground-state of $\mathcal{M}$. As more constraints were added, the gap gradually grew. The reconstruction error decreased correspondingly, showing excellent agreement with our estimate (13). We also ran simulations on random XY chains to reach larger system sizes ($|\Lambda = 100|$). The gap of the correlation matrix seemed insensitive to the size of the sub-system for the range we examined $7 \leq |L| \leq 13$ (Fig. S1 in Appendix C).

Reconstruction from Gibbs states. Next, we reconstructed H_L for random spin chains from measurements of their Gibbs states. We sampled 200 random Hamiltonians (14) on $|\Lambda| = 12$ sites and generated Gibbs states $\frac{1}{Z} e^{-\beta H}$ for varying $\beta \in [0.01, 1]$. We then measured a fixed number of observables, corresponding to all 4-local constraints A_n supported on the 6 middle spins L_0. We added a small noise ($\epsilon = 10^{-12}$) to each measurement.

Our results show that the reconstruction error increases with temperature (Fig. 3, left). As the system approaches a fully mixed state, the commutator $[H, \rho]$ ap-

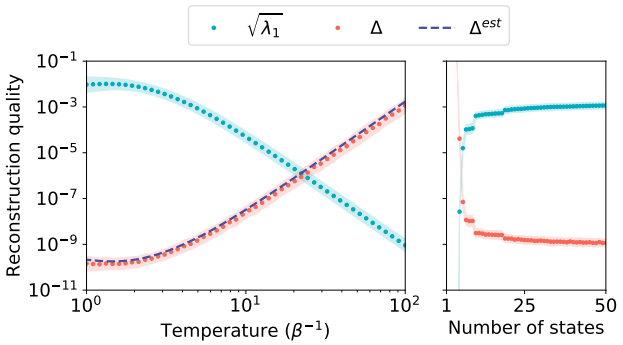

FIG. 3. Quality of Hamiltonian reconstruction from Gibbs states $\rho = \frac{1}{Z}e^{-\beta H}$. We reconstructed H_L on the $L = 8$ middle spins of random spin chains (14) of length $|\Lambda| = 12$ (see Fig. 1). Left: as a function of temperature $T = \beta^{-1}$, using 4-local constraints A_n. The gap of the correlation matrix $\mathcal{M}$ decreased with temperature (light blue). Correspondingly, the reconstruction error (red) due to a small measurement uncertainty ($\epsilon = 10^{-12}$) increased according to the estimate (13) obtained from the spectrum of K (dashed purple). Right: reconstruction with 2-local measurements only, using single-site constraints A_n and multiple Gibbs states of different temperatures. We generated Gibbs states at temperatures in the range $\beta^{-1} = [10^0, 10^2]$, chosen with uniform spacings (in log space) which decreased with the number of states. A few different states sufficed; additional states improved the reconstruction quality. Results were averaged over 200 randomizations.

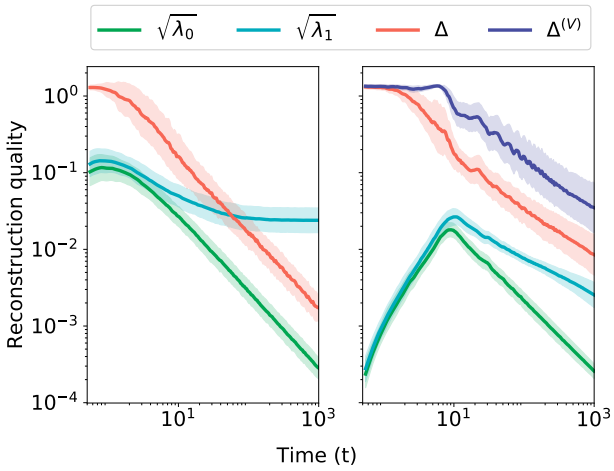

FIG. 4. Reconstruction from dynamics, as a function of time, of the final Hamiltonian following a quench at $t = 0$. Left: reconstruction of a time-independent Hamiltonian from $\rho_{avg}(t) = \frac{1}{t}\int_0^t \rho(t')dt'$. While the first excited eigenvalue λ_1 of the correlation matrix $\mathcal{M}$ saturated (light blue), its lowest eigenvalue λ_0 decayed with time (green), leading to a decrease in the reconstruction error Δ (red). Right: reconstruction of a time-oscillating Hamiltonian $\hat{H}(t) = \hat{H}^{(0)} + J\hat{V}\cos\omega t$. Here λ_1 decreased with time due to heating, leading to a larger reconstruction error Δ compared to the time-independent case (red for $\hat{H}^{(0)}$, purple for $\hat{V}$). Results were averaged over 50 randomizations.

proaches zero for *every* H, which implies that many different H are becoming compatible with ρ. Correspondingly, the elements of the constraint matrix K shrink, and so does its gap. At low temperatures, the reconstruction quality was similar to that of ground-states. By combining measurements performed at different temperatures, we were able to recover H using only 2-local measurements (Fig. 3, right).

Reconstruction from dynamics. To demonstrate Hamiltonian recovery from the dynamics of an initial state, we simulated a quench protocol. We generated two random Hamiltonians $\hat{H}^{(0)}$, $\hat{H}^{(1)}$ on $|\Lambda| = 12$ sites from the ensemble (14). We initialized our system in the ground-state of $\hat{H}^{(0)} + \hat{H}^{(1)}$, and evolved it by $\hat{H}^{(0)}$ alone. This initialization yielded states whose energy with respect to the final Hamiltonian was not too high. We then attempted at different times t to recover $H_L^{(0)}$ on the 8 middle spins using 4-local constraints A_n. We did this by constructing a constraint matrix K_t from time-averaged values of $\langle i[A_n, S_m]\rangle$, sampled at equally spaced intervals $dt = 0.05$ up to time t.

After a transient period, the first excited eigenvalue λ_1 of the correlation matrix $\mathcal{M}$ saturated (Fig. 4). The lowest eigenvalue λ_0 continued to decay, opening a gap which widened with time. This decay fits to the power law $\sqrt{\lambda_0} \propto \frac{1}{t}$, reflecting the expected decay rate of the commutator with the true Hamiltonian from Eq. (9). Here the finite value of λ_0 played the role of noise, leading to reconstrcution error. As λ_0 decreased, the Hamiltonian was reconstructed to better and better accuracy.

Recovery of time-dependent Hamiltonians. We repeated the quench experiments with a final Hamiltonian which depends on time, focusing on a periodic drive with a single frequency: $f(t) = \cos\omega t$. We initialized our system in the ground-state of $\hat{H}^{(0)} + \hat{H}^{(1)}$ and evolved it in time with

$$\hat{H}(t) = \hat{H}^{(0)} + J\hat{V}\cos\omega t, \tag{15}$$

taking $J = 0.5$ and $\omega = 0.05$. We sampled all three terms $\hat{H}^{(0)}, \hat{H}^{(1)}$ and $\hat{V}$ identically using the form given by Eq. (14). We then constructed at different times an extended constraint matrix $K_{N\times 2M}$ from time-averaged as well as time-modulated observables [see Eq. (11)].

As in the time-independent case, λ_0 decayed with time (Fig. 4, right). However, λ_1 decayed too, indicating a small or vanishing gap $\lambda_1 - \lambda_0$ for long times, corresponding to high temperatures (see Fig. 3). Recovery in this case is therefore possible when the system does not heat too quickly, i.e. when λ_1 decays slower than λ_0, which depends on the driving amplitude J and frequency ω (see Appendix D).

Discussion. We suggest a framework for inferring local Hamiltonians. Our framework generalizes the recently-introduced correlation matrix formalism [20–22], applying to Gibbs states and dynamics as well as eigen-

states. Importantly, it allows to recover short-ranged Hamiltonians using measurements as well as computational resources scaling linearly with system size.

We point out that even when the available measurements do not provide sufficient constraints to open a gap in $\mathcal{M}$ and yield a unique Hamiltonian H, our method recovers a linear subspace containing H. This can be combined with additional knowledge, e.g. to verify the accuracy of a prior guess for H or to improve such a guess.

Most of our formalism applies equally well to long-ranged Hamiltonians, in which interactions can involve any arbitrary set of k spins. Our algorithm must then be applied to the whole system Λ at once rather than locally. Still, the number of possible Hamiltonian terms scales polynomially with system size, as $|\Lambda|^k$.

Note that when we enforce stationarity of all possible observables A_n on the full system $L = \Lambda$, our correlation matrix takes the appealing form $\mathcal{M}_{ij}^{\Lambda} = \mathrm{Tr}\left([\rho, S_i]^\dagger [\rho, S_j]\right)$, coinciding with the correlation matrix defined in Ref. [20] (up to a scalar; see Appendix E). If we suffice with the full set of observables A_n on the interior L_0 of a subsystem, Eq. (7) is equivalent to the operator identity $\mathrm{Tr}_{\partial L}[\rho_L, H_L] = 0$. Here, ρ_L is the reduced density matrix on L and $\mathrm{Tr}_{\partial L}$ is a partial trace on the boundary spins $\partial L \stackrel{\mathrm{def}}{=} L - L_0$ included in L but not in its interior (see Ref. [24]). We also note that adding constraints and Hamiltonian terms acting on ∂L converts our algorithm to a method for finding the entanglement Hamiltonian on L (similar to [25]).

ACKNOWLEDGMENTS

We thank Miklos Santha and Anupam Prakash for illuminating discussions, and Renan Gross for critical comments on the manuscript. E. B. and N. L. acknowledge financial support from the European Research Council (ERC) under the European Union Horizon 2020 Research and Innovation Programme (Grant Agreement No. 639172). I. A. acknowledges the support of the Israel Science Foundation (ISF) under the Individual Research Grant 1778/17. N. L. acknowledges support from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (No. FP7/2007–2013) under REA Grant Agreement No. 631696 and from the Israeli Center of Research Excellence (I-CORE) "Circle of Light.".

[1] J. Preskill, (2018), arXiv:1801.00862.
[2] D. Burgarth, K. Maruyama, and F. Nori, Physical Review A - Atomic, Molecular, and Optical Physics **79** (2009), 10.1103/PhysRevA.79.020305, arXiv:0810.2866.
[3] C. Di Franco, M. Paternostro, and M. S. Kim, Physical Review Letters **102** (2009), 10.1103/PhysRevLett.102.187203, arXiv:0812.3510.
[4] J. Zhang and M. Sarovar, Phys. Rev. Lett. **113**, 80401 (2014).
[5] L. E. De Clercq, R. Oswald, C. Flühmann, B. Keitch, D. Kienzler, H. Y. Lo, M. Marinelli, D. Nadlinger, V. Negnevitsky, and J. P. Home, Nature Communications **7** (2016), 10.1038/ncomms11218, arXiv:arXiv:1011.1669v3.
[6] A. Sone and P. Cappellaro, Physical Review A **95** (2017), 10.1103/PhysRevA.95.022335, arXiv:1609.09446.
[7] A. Sone and P. Cappellaro, Physical Review A **96** (2017), 10.1103/PhysRevA.96.062334, arXiv:1702.03280.
[8] Y. Wang, D. Dong, B. Qi, J. Zhang, I. R. Petersen, and H. Yonezawa, IEEE Transactions on Automatic Control **63**, 1388 (2018), arXiv:1610.08841.
[9] K. Rudinger and R. Joynt, Physical Review A - Atomic, Molecular, and Optical Physics **92** (2015), 10.1103/PhysRevA.92.052322, arXiv:arXiv:1410.3029v1.
[10] M. Kieferová and N. Wiebe, Physical Review A **96** (2017), 10.1103/PhysRevA.96.062327, arXiv:1612.05204.
[11] H. J. Kappen, (2018), arXiv:1803.11278.
[12] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, New Journal of Physics **14** (2012), 10.1088/1367-2630/14/10/103013, arXiv:arXiv:1207.1655v1.
[13] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Physical Review Letters **112** (2014), 10.1103/PhysRevLett.112.190501, arXiv:1309.0876.
[14] N. Wiebe, C. Granade, C. Ferrie, and D. Cory, Physical Review A - Atomic, Molecular, and Optical Physics **89** (2014), 10.1103/PhysRevA.89.042314, arXiv:1311.5269.
[15] N. Wiebe, C. Granade, and D. G. Cory, New Journal of Physics **17** (2015), 10.1088/1367-2630/17/2/022005, arXiv:1409.1524.
[16] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'brien, J. G. Rarity, A. Laing, and M. G. Thompson, Nature Physics **13**, 551 (2017), arXiv:1703.05402.
[17] S. T. Wang, D. L. Deng, and L. M. Duan, New Journal of Physics **17** (2015), 10.1088/1367-2630/17/9/093017, arXiv:1505.00665.
[18] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, Physical Review A - Atomic, Molecular, and Optical Physics **84** (2011), 10.1103/PhysRevA.84.012107, arXiv:1002.1330.
[19] M. P. Da Silva, O. Landon-Cardinal, and D. Poulin, Physical Review Letters **107** (2011), 10.1103/PhysRevLett.107.210404, arXiv:1104.3835.
[20] X.-L. Qi and D. Ranard, (2017), arXiv:1712.01850.
[21] E. Chertkov and B. K. Clark, (2018), arXiv:1802.01590v2, arXiv:1802.01590.
[22] M. Greiter, V. Schnells, and R. Thomale, (2018), arXiv:1802.07827.
[23] Equivalently, $\Delta = 2|\sin\frac{\theta}{2}|$, where θ is the angle between the two vectors [20].
[24] A. Anshu, I. Arad, and A. Jain, Physical Review B **94** (2016), 10.1103/PhysRevB.94.195143, arXiv:1603.06049.
[25] W. Zhu, Z. Huang, and Y.-c. He, (2018), arXiv:1806.08060.

[26] I. Peschel, Journal of Physics A: Mathematical and General **36**, L205 (2003), arXiv:0212631 [cond-mat].
[27] S. A. Cheong and C. L. Henley, Physical Review B - Condensed Matter and Materials Physics **69** (2004), 10.1103/PhysRevB.69.075111, arXiv:0206196 [cond-mat].
[28] I. Peschel and V. Eisler, Journal of Physics A: Mathematical and Theoretical **42** (2009), 10.1088/1751-8113/42/50/504003, arXiv:arXiv:0906.1663v3.
[29] S. Sachdev, *Cambridge University Press* (2011) p. 517, arXiv:9811058 [cond-mat].

Appendix A: Recovering time-dependent Hamiltonians

Suppose we wish to recover a time-dependent Hamiltonian of the form:

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{V} f(t), \tag{S1}$$

where $f(t)$ is a known function and $\hat{H}, \hat{V}$ are the operators we wish to learn. For any operator A, Schrodinger's equation now reads:

$$i\partial_t \langle A \rangle = \langle [A, \hat{H}^{(0)}] \rangle + \langle [A, f(t)\hat{V}] \rangle. \tag{S2}$$

Integrating the above equation and expanding in local operators: $\hat{H}^{(0)} = \sum c_m^{(0)} h_m$, $\hat{V} = \sum c_m^{(v)} h_m$, we obtain:

$$\left| \sum_j \frac{c_j^{(0)}}{t} \int_0^t \langle [A, h_m] \rangle dt' + \sum_j \frac{c_j^{(v)}}{t} \int_0^t \langle [A, h_m] \rangle f(t') dt' \right| \leq \frac{2\,\|A\|}{t}. \tag{S3}$$

Minimizing the LHS of (S3) with respect to a set of operators $\{A_n\}_{n=1}^N$ amounts to finding the lowest right-singular vector of the extended constraint matrix $K_{N\times 2M}$, defined as

$$\forall m \leq M: \; K_{n,m} = \frac{1}{t} \int_0^t \langle [A_n, h_m] \rangle dt'$$

$$K_{n,m+M} = \frac{1}{t} \int_0^t \langle [A_n, h_m] \rangle f(t') dt'.$$

Appendix B: Error estimation

Experimentally, each element of the constraint matrix $K_{n,m} = \langle [A_n, S_m] \rangle$ can only be estimated using a finite number of samples n_s. Therefore, the measured empirical constraint matrix $\hat{K}$ deviates from the true one K by a noise matrix. We would like to estimate the error in the recovered Hamiltonian due to this noise.

We study the effect of the noise by treating it as a perturbation. We assume that the correct K has a one-dimensional kernel; namely, we were given a state ρ for which there is only one local Hamiltonian $\vec{c}$ (up to an overall scalar) that satisfies $K\vec{c} = 0$. After many measurements, we can use the central limit theorem to model the noise as a Gaussian matrix:

$$\hat{K} - K \approx \epsilon \hat{E}, \tag{S1}$$

where each entry $\hat{E}_{nm}$ of $\hat{E}$ is an independent random variable with zero mean and unit standard deviation. It is scaled by a small parameter ϵ which decays as $n_s^{-1/2}$.

We wish to estimate the distance between the true and recovered Hamiltonians. This distance is given by $\| |c_0'\rangle - |c_0\rangle \|$, where $|c_0\rangle$ and $|c_0'\rangle$ are the ground-states of the clean $K^T K$ and its noisy estimate $(K+\epsilon\hat{E})^T(K+\epsilon\hat{E})$. We treat $K^T K$ as an unperturbed Hamiltonian, and $\epsilon(\hat{E}^T K + K^T \hat{E})$ as a perturbation to first order in ϵ. We obtain:

$$|c_0'\rangle - |c_0\rangle = \epsilon \sum_{i>0} |c_i\rangle \frac{\langle c_i | \hat{E}^T K + K^T \hat{E} | c_0 \rangle}{\lambda_i - \lambda_0} + \mathcal{O}(\epsilon^2) \tag{S2}$$

$$= \epsilon \sum_{i>0} |c_i\rangle \frac{\langle c_i | K^T \hat{E} | c_0 \rangle}{\lambda_i} + \mathcal{O}(\epsilon^2), \tag{S3}$$

where $|c_i\rangle$ are the eigenstates of $K^T K$ and λ_i the corresponding eigenvalues in increasing order. $\lambda_0 = 0$ since we assumed that an exact reconstruction exists, which also implies that K annihilates $|c_0\rangle$:

$$K^T K |c_0\rangle = 0 \Rightarrow \|K|c_0\rangle\|^2 = \langle c_0 | K^T K | c_0 \rangle = 0. \tag{S4}$$

Similarly, $\|K|c_i\rangle\|^2 = \lambda_i$, so we can write:

$$K|c_i\rangle = \sqrt{\lambda_i} |\tilde{c}_i\rangle \tag{S5}$$

For some unit vector $|\tilde{c}_i\rangle$. Using this we obtain:

$$|c_0'\rangle - |c_0\rangle = \epsilon \sum_{i>0} \frac{1}{\sqrt{\lambda_i}} |c_i\rangle \langle \tilde{c}_i | \hat{E} | c_0 \rangle + \mathcal{O}(\epsilon^2), \tag{S6}$$

and therefore,

$$\| |c_0'\rangle - |c_0\rangle \| = \epsilon \sqrt{\sum_{i>0} \frac{1}{\lambda_i} |\langle \tilde{c}_i | \hat{E} | c_0 \rangle|^2} + \mathcal{O}(\epsilon^2). \tag{S7}$$

We can now average over the noise $\hat{E}$ by invoking Jensen's inequality, together with the concavity of the square root function:

$$\begin{aligned} \mathbb{E} \| |c_0'\rangle - |c_0\rangle \| &= \epsilon \mathbb{E} \sqrt{\sum_{i>0} \frac{1}{\lambda_i} |\langle \tilde{c}_i | \hat{E} | c_0 \rangle|^2} + \mathcal{O}(\epsilon^2) \\ &\leq \epsilon \sqrt{\sum_{i>0} \frac{1}{\lambda_i} \mathbb{E} |\langle \tilde{c}_i | \hat{E} | c_0 \rangle|^2} + \mathcal{O}(\epsilon^2) \\ &= \epsilon \sqrt{\sum_{i>0} \frac{1}{\lambda_i}} + \mathcal{O}(\epsilon^2). \end{aligned} \tag{S8}$$

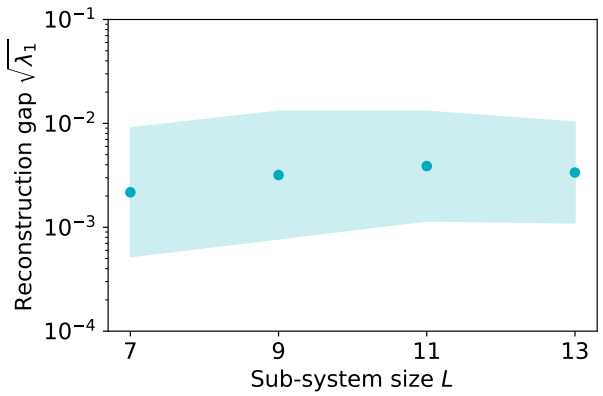

FIG. S1. Recovering H_L for various sub-system sizes $|L|$ of long XY chains with $|\Lambda = 100|$ sites. The reconstruction gap λ_1, which quantifies the robustness of the reconstruction procedure to noise, seems insensitive to sub-system size for $7 \leq |L| \leq 13$. Results are averaged over 50 randomizations, with mean (dark circles) and standard deviation (light shading) were calculated after taking the log.

In the last equality we used the identity $\mathbb{E}|\langle\tilde{c}_i|\hat{E}|c_0\rangle|^2 = 1$, which follows from:

$$\begin{aligned}
\mathbb{E}|\langle\tilde{c}_i|\hat{E}|c_0\rangle|^2 &= \mathbb{E}\left|\sum_{n=1}^{N}\sum_{m=1}^{M}\tilde{c}_i^n c_0^m \hat{E}_{nm}\right|^2 \\
&= \sum_{n,\nu=1}^{N}\sum_{m,\mu=1}^{M}\tilde{c}_i^n\tilde{c}_i^\nu c_0^m c_0^\mu \mathbb{E}\left(\hat{E}_{nm}\hat{E}_{\nu\mu}\right) \\
&= \sum_{n,\nu=1}^{N}\sum_{m,\mu=1}^{M}\tilde{c}_i^n\tilde{c}_i^\nu c_0^m c_0^\mu \delta_{n,\nu}\delta_{m,\mu} \\
&= \sum_{a=1}^{N}\sum_{b=1}^{M}\left|\tilde{c}_i^a\right|^2\left|c_0^b\right|^2 \\
&= \langle\tilde{c}_i|\tilde{c}_i\rangle\langle c_0|c_0\rangle = 1.
\end{aligned} \tag{S9}$$

Appendix C: Scaling of λ_1 with sub-system size

We simulated larger systems of an integrable model to find how the required number of measurements of each observable scales with $|L|$. We examined random XY chains, for which the Hamiltonian is given by:

$$H_{XY} = \frac{1}{2}\sum_{l=1}^{\Lambda}\left[2g_l\sigma_l^z + (1+\gamma_l)\sigma_l^x\sigma_{l+1}^x + (1-\gamma_l)\sigma_l^y\sigma_{l+1}^y\right]. \tag{S1}$$

Using the methods described in [26–29], we constructed reduced density matrices for ground-states on $\Lambda = 100$ sites. We considered sub-regions consisting of $|L| = 7, 9, 11, 13$ sites, and calculated the gaps of the correlation matrices $\mathcal{M}$ constructed with all the 4-local constraints A_n supported on the corresponding interior regions L_0. The gap of the correlation matrix seems insensitive to sub-system size for the sizes we examined (Fig. S1).

Appendix D: Recovery for different driving parameters

Recovery of a time-dependent $H(t)$ is only possible when heating is sufficiently slow. As time progresses, the time-averaged commutator $[\rho(t), H(t)]$ decays, as quantified by λ_0, the lowest right-singular value of the extended constraint matrix $K_{N\times 2M}$. However, since energy is not conserved for such a system, ρ_{avg} could heat up to an infinite-temperature fully-mixed state, which trivially commutes with any Hamiltonian. One measure for this process is the next singular value λ_1, quantifying how well any competing Hamiltonian would commute with ρ on average. Recovery is therefore possible whenever λ_0 decays faster than λ_1, such that the solution does not mix the true Hamiltonian much with any competitor.

Indeed, we find that stronger or slightly faster driving leads to a more rapid decay of the reconstruction gap (Fig. S2). This agrees with our expectation that a larger driving amplitude should lead to faster energy absorption from the drive; we expect the same from a slightly higher driving frequency within the low-frequency regime we study.

Appendix E: Relation to previously-defined correlation matrix

If we wish to recover the Hamiltonian on the full system Λ by enforcing stationarity of all possible constraints A_n, the correlation matrix $\mathcal{M}$ takes the following form, as defined in [20] (up to a multiplicative scalar):

$$\begin{aligned}
M_{ij} &= \sum_n K_{ni}K_{nj} \\
&= \sum_n \langle i\left[h_i, A_n\right]\rangle\langle i\left[h_j, A_n\right]\rangle \\
&= \sum_n \overline{\mathrm{Tr}\left(\rho\left[h_i, A_n\right]\right)}\,\mathrm{Tr}\left(\rho\left[h_j, A_n\right]\right) \\
&= \sum_n \overline{\mathrm{Tr}\left(A_n\left[h_i, \rho\right]\right)}\,\mathrm{Tr}\left(A_n\left[h_j, \rho\right]\right) \\
&= 2^\Lambda\,\mathrm{Tr}\left(\left[h_i,\rho\right]^\dagger\left[h_j,\rho\right]\right)
\end{aligned} \tag{S1}$$

Where we used the identity $\mathrm{Tr}\left(A\left[B,C\right]\right) = \mathrm{Tr}\left(C\left[A,B\right]\right)$ which follows from the cyclic property of the trace, as well as $[A,B] = -[B,A]$. Finally, the last equality follows from the generalized Parseval identity for the Hilbert-Schmidt inner product, namely:

$$\langle v, w\rangle = \sum_i \overline{\langle v, e_i\rangle}\,\langle w, e_i\rangle\,, \tag{S2}$$

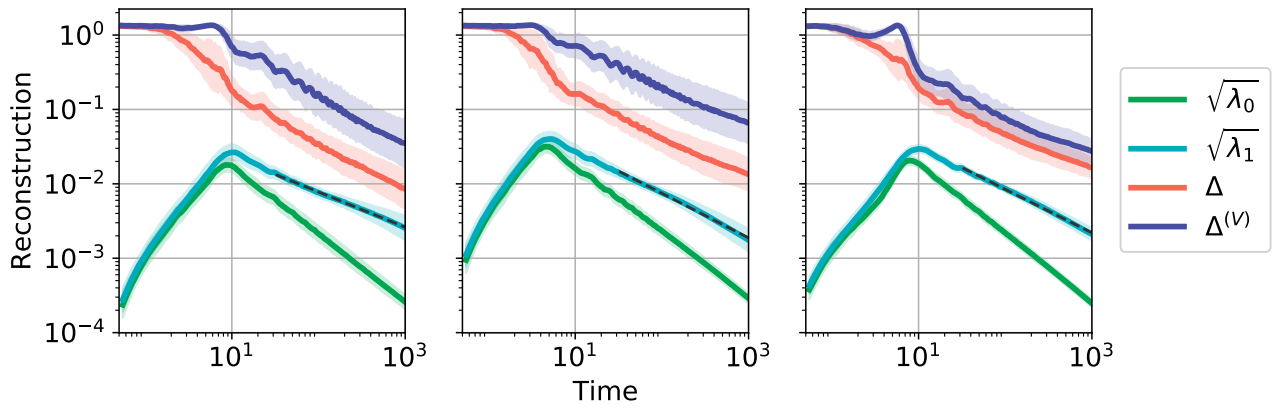

FIG. S2. Recovery of a time-dependent Hamiltonian as a function of driving amplitude and frequency. Left: right panel of Fig. 4 in the main text (driving frequency $\omega = 0.05$, amplitude $J = 0.5$). Middle: double frequency ($\omega = 0.1$, $J = 0.5$). Right: double amplitude ($\omega = 0.05$, $J = 1$). The power α of the long-time decay of the first excited right-singular value $\sqrt{\lambda_1} \sim t^{-\alpha}$ is larger in the central and right panels compared to the left panel (-0.599 ± 0.004 and -0.584 ± 0.003 compared to -0.473 ± 0.002), indicating more noisy recovery. Indeed, we expect faster heating to arise from a stronger driving amplitude, as well as from a higher driving frequency within this low-frequency regime.

taking the operators A_n to be an orthogonal basis for the operators on Λ. We normalize A_n to unity in operator norm (largest eigenvalue) rather than Hilbert-Schmidt norm, since measurements in the lab yield ± 1 outcomes; this is the origin of the 2^Λ factor.