Hamiltonian Tomography via Quantum Quench

Zhi Li, 1, 2, 3 Liujun Zou, 3 and Timothy H. Hsieh 3

1 Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States
2 Pittsburgh Quantum Institute, Pittsburgh, Pennsylvania 15260, United States
3 Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

We show that it is possible to uniquely reconstruct a generic many-body local Hamiltonian from a single pair of generic initial and final states related by evolving with the Hamiltonian for any time interval. We then propose a practical version of the protocol involving multiple pairs of such initial and final states. Using the eigenstate thermalization hypothesis, we provide bounds on the protocol’s performance and stability against errors from measurements and in the ansatz of the Hamiltonian. The protocol is efficient (requiring experimental resources scaling polynomially with system size in general and constant with system size given translation symmetry) and thus enables analog and digital quantum simulators to verify implementation of a putative Hamiltonian.

The advent of quantum many-body simulators [1–12] has enabled the exploration of complex quantum dynamics beyond the capabilities of classical computers. Given this significant potential, it is vital to determine accurately the Hamiltonian actually being realized by a simulator. However, Hamiltonian tomography for a generic many-body system is challenging, precisely due to the fact that the complexity of many-body dynamics makes benchmarking with a classical computer difficult. Thus far, most progress has been made in systems with either a special Hamiltonian or a small size [13–28].

In this Letter, we introduce two protocols for Hamiltonian tomography (Fig. 1). The first is more of conceptual interest: we show that it is possible to uniquely reconstruct a generic many-body Hamiltonian with local interactions given only a single pair of generic initial and final states related by time evolving with the Hamiltonian. Our approach is partly motivated by the recent proposals of reconstructing a many-body Hamiltonian from a single eigenstate or steady state [29–33]. As in these studies, our protocol relies on the physical assumption that the Hamiltonian is local, which implies that the number of its parameters scales polynomially with system size. However, in contrast to these works, our approach does not require a steady state and instead relies precisely on how a generic state changes in time.

Despite being conceptually interesting, this first protocol is impractical and we thus propose a second, practical version that uses multiple pairs of initial and final states connected by time evolving with the Hamiltonian. This protocol only requires measuring a set of local observables in each pair of initial and final states. We bound the errors in the reconstructed Hamiltonian due to errors in both the measurements and the ansatz of the Hamiltonian, and the fidelity of the reconstruction can be increased to unity by using more pairs of initial-final states.

We emphasize that our locality assumption does not necessarily have to be spatial locality; we only require that all interactions in the Hamiltonian involve a fixed number of degrees of freedom that is independent of system size. Hence, our protocol applies to all analog and digital quantum simulators, from superconducting circuits to trapped ion platforms. We only require measurement of local observables, which can be done with high precision thanks to recent advances in single-site resolution.

**Tomography from a single quench.**—Given a Hamiltonian $H$ and an initial state $|\psi(0)\rangle$ of a $D$-dimensional Hilbert space, we ask whether one can determine $H$ from only $|\psi(0)\rangle$ and $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$. Without any restrictions, it is impossible to determine $H$ uniquely, since $H$ has $O(D^2)$ parameters while the given wave functions contain $O(D)$ parameters. However, a many-body Hamiltonian with local interactions has number

---

**FIG. 1.** Two protocols for Hamiltonian tomography. (a) From a single pair of initial and final states related by time evolving with the Hamiltonian, $|\psi(t)\rangle = U|\psi(0)\rangle$ with $U = \exp(-iHt)$, a generic Hamiltonian can be uniquely reconstructed. (b) A practical version using $p$ pairs of initial and final states related by time evolving with the Hamiltonian and requiring only local measurements on the initial and final states.
of parameters scaling polynomially with system size, $L$, whereas $D$ is exponential in $L$, and thus reconstructing a local $H$ is possible in principle.

Without loss of generality, we consider a traceless many-body Hamiltonian and decompose it into $n$ local interaction terms:

$$H = \sum_{\alpha=1}^{n} c_{\alpha} \mathcal{O}_{\alpha}, \quad (1)$$

where $\{\mathcal{O}_{\alpha}\}$ are traceless local Hermitian operators, $\{c_{\alpha}\}$ are coupling constants, and $n$ is a polynomial of $L$ due to locality. Our goal is to determine $\{c_{\alpha}\}$ from $|\psi(0)\rangle$ and $|\psi(t)\rangle = e^{-iHT} |\psi(0)\rangle$.

Our approach is based on the simple observation of generalized energy conservation:

$$\langle \psi(t) | H^m | \psi(t) \rangle = \langle \psi(0) | H^m | \psi(0) \rangle \quad (2)$$

for any positive integer $m$. These equations place many constraints on the variables $\{c_{\alpha}\}$. Indeed, substituting Eq. (1) (with $c$ replaced by $x$, a symbol for unknowns) into Eq. (2) yields (repeated indices summed over):

$$M^{(m)}_{\alpha_1 \alpha_2 \cdots \alpha_m} x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_m} = 0, \quad (3)$$

where

$$M^{(m)}_{\alpha_1 \alpha_2 \cdots \alpha_m} = \langle \psi(0) | \mathcal{O}_{\alpha_1} \mathcal{O}_{\alpha_2} \cdots \mathcal{O}_{\alpha_m} | \psi(0) \rangle - \langle \psi(t) | \mathcal{O}_{\alpha_1} \mathcal{O}_{\alpha_2} \cdots \mathcal{O}_{\alpha_m} | \psi(t) \rangle. \quad (4)$$

More explicitly, we have $M^{(1)}_{\alpha_1} x_{\alpha_1} = 0$, $M^{(2)}_{\alpha_1 \alpha_2} x_{\alpha_1} x_{\alpha_2} = 0$, etc., which we refer to as first order, second order, etc. This constitutes a system of polynomial equations for $\{x_{\alpha}\}$ which can be used to determine $\{c_{\alpha}\}$ (and thus $H$). Note that without knowledge of $t$, one can always rescale $H$ and $t$ to leave $Ht$ invariant, so $\{c_{\alpha}\}$ will be determined up to an overall multiplicative factor. Hereafter we omit the “up to a factor” caveat.

Although Eq. (2) is valid for any positive integer $m$, it can be shown [34] that at most $(D - 1)$ of these equations are independent. As long as $D - 1 \geq n$, our procedure uses the first $n$ equations ($m = 1, 2 \cdots n$) to determine $\{c_{\alpha}\}$. We find that if there exists one example of a Hamiltonian, $H_0$, that can be uniquely reconstructed from this procedure, then a generic Hamiltonian can be uniquely reconstructed from this procedure. The proof of our claim is based on the analytic properties of results of polynomial equations, and we leave the technical details to the Supplementary Material [34]. There we also show that given the time $t$ as an additional input, a generic $H$ can be determined completely (including the overall scale).

Our procedure applies to generic Hamiltonians and initial states, and it can be immediately generalized for mixed initial states. However, there will be fine-tuned cases for which it fails. For example, if $|\psi(0)\rangle$ is an eigenstate of $H$, Eq. (2) is trivially satisfied and we cannot obtain $H$ from it. Nevertheless, in this special case one can in principle employ the methods in Refs. [29–33] to determine the Hamiltonian. Furthermore, if $H$ admits a conserved quantity $Q$ that can be decomposed into the same set of local operators, $\{\mathcal{O}_{\alpha}\}$, then Eq. (2) yields any linear combination of $H$ and $Q$.

**Proof of concept.**—In order to demonstrate uniqueness for generic cases, one needs to show the existence of $H_0$ for arbitrary system size, which is a difficult problem. In this Letter, we present three physically relevant checks.

The first is an analytical check on a spin-1/2 chain. We consider a translationally invariant transverse-field Ising model with three random couplings:

$$H = c_1 \sum_i Z_i Z_{i+1} + c_2 \sum_i X_i + c_3 \sum_i Y_i Y_{i+2}, \quad (5)$$

in which $X, Y, Z$ are Pauli operators. We have added a next-nearest-neighbor interaction; otherwise, the problem is trivial and the first order (linear) equation in Eq. (3) is enough to determine $c_1/c_2$. Using a small $t$ approximation, it can be analytically shown [34] that the first and second order equations have two solutions, only one of which satisfies the third order equation, and this unique reconstruction for small $t$ is sufficient to establish unique reconstruction for finite $t$. We emphasize that the small $t$ approximation is not a requirement of our protocol; it is an intermediate step in establishing the uniqueness of reconstruction.

We also numerically check reconstruction using finite time evolution for similar models. For example, we check the Ising model with transverse and longitudinal fields and the Heisenberg model for chains up to length $L = 10$:

$$H = c_1 \sum_{i=1}^{L-1} Z_i Z_{i+1} + c_2 \sum_{i=1}^{L} X_i + c_3 \sum_{i=1}^{L} Z_i, \quad (6)$$

$$H = c_1 \sum_{i=1}^{L-1} X_i X_{i+1} + c_2 \sum_{i=1}^{L-1} Y_i Y_{i+1} + c_3 \sum_{i=1}^{L-1} Z_i Z_{i+1}. \quad (6)$$

We find similar uniqueness results as above.

The third check is closer to a generic Hamiltonian. We consider a transverse field Ising model with random spatially varying couplings $c_{1,i}$ and $c_{2,i}$:

$$H = \sum_{i=1}^{L} c_{1,i} X_i + \sum_{i=1}^{L-1} c_{2,i} Z_i Z_{i+1} \quad (7)$$

Due to the high computational complexity [the number of coefficients in Eq. (3) is exponential in $n$], we only check the case of $L = 4$ (7 local operators in total) for illustration. Using the Gröbner basis [35] technique, we find that $n = 7$ polynomial equations indeed determine $\{c\}$ uniquely.

The above checks provide evidence for the existence of $H_0$ for various classes of Hamiltonians, which strongly
suggests that a generic local many-body Hamiltonian can be uniquely reconstructed from a single pair of initial-final states related via time evolving with the Hamiltonian.

Reconstruction from multiple quenches.—Though the above protocol is conceptually interesting, it is impractical because the experimental and computational complexity are both exponential in \( n \). Hence, below we present a more practical method for Hamiltonian tomography, whose experimental and computational complexity is only a polynomial of \( n \) and whose sensitivity to errors can be controlled.

Since the high complexity of the first approach arises from the higher order equations \([m \geq 2]\) in Eq. (3), we will only keep the linear equation \((m = 1)\). In order to uniquely determine \( \{c_{\alpha}\} \), we need at least \((n - 1)\) linear equations. Hence we use \( p \geq (n - 1) \) pairs of initial and final states related by time evolving with \( H \):

\[
\{ |\psi_i(0)\rangle \rightarrow |\psi_i(t)\rangle = e^{-iHt} |\psi_i(0)\rangle, \quad i = 1, 2 \cdots p \}, \tag{8}
\]

As a result, we have a linear system of equations \( Mx = 0 \), where \( M \) is a \( p \times n \) matrix with entries

\[
M_{\alpha i} = \langle \psi_i(0) | \mathcal{O}_\alpha | \psi_i(0) \rangle - \langle \psi_i(t) | \mathcal{O}_\alpha | \psi_i(t) \rangle. \tag{9}
\]

In principle it is sufficient to use \( p = n - 1 \), and the kernel of \( M \) will generically be one dimensional and equal to \( \{c_{\alpha}\} \). (The uniqueness can be addressed by the same technique as for the single quench protocol.) In contrast to the first approach, the number of measurements required is proportional to the number of coefficients in \( M \), which is \( O(n^2) \).

Inevitably there will be experimental errors in state preparation, time evolution, and measurements, and these will lead to errors in the reconstructed Hamiltonian. We can mitigate these errors by utilizing more than \((n - 1)\) pairs of initial-final states. Using \( p > n - 1 \), we want to find the best fit of \( \{x\} \) for \( Mx = 0 \) when \( M_{\alpha i} \) has random error. We use the standard least squares method for a homogeneous system of equations: the best estimate of \( \{x\} \) is given by the right singular vector of \( M \) with the smallest singular value (or equivalently, the eigenvector of \( M^T M \) with the smallest eigenvalue).

A similar approach was used in Ref. [36] to find local integrals of motion; however, while that approach used different time slices from evolving a fixed initial state, our approach fixes one time and uses different pairs of initial and final states, which leads to more efficient use of experimental resources, as we will discuss below.

Stability against measurement errors.—To quantify the error in the reconstructed Hamiltonian \( H = \sum_\alpha x_\alpha \mathcal{O}_\alpha \), we define \( \theta \) to be the angle between the vector of reconstructed coupling constants \( x \) and the vector of actual coupling constants \( c \), and correspondingly the fidelity \( F = | \cos \theta | \) and reconstruction error \( E = | \sin \theta | \).

Consider an error model in which each matrix element \( M_{\alpha i} \) has an additive error uniformly distributed between \( \pm \epsilon \). Using both standard perturbation theory and non-perturbative results in statistical theory [37], we find that the average reconstruction error is bounded:

\[
E \leq C \sqrt{\frac{n}{p}} \frac{\epsilon}{\lambda} \tag{10}
\]

for \( \epsilon < \lambda \), where \( C \) is a constant and \( \lambda \) is the gap between the two lowest singular values of \( \frac{1}{\sqrt{p}} M \) \( (\lambda \) becomes independent of \( p \) as \( p \to \infty \)) [34].

Therefore, by increasing the number of pairs of initial-final states, \( p \), one can decrease the reconstruction error [38]. As Eq. (10) implies, we would like \( \lambda \) to be as large as possible. The gap is determined by \( M \), which in turn depends on the important choices of time \( t \) and the ensemble of initial states. Intuitively, larger \( t \) is preferable so that the initial and final states are more distinguishable in terms of local observables. In the same vein, the initial states should not be too random; otherwise, the time evolution has little effect on changing local observables (for example, we find that Haar random initial states perform poorly). At the same time, states in the initial ensemble should be distinct enough to provide independent information about \( H \).

We now discuss the dependence of \( \lambda \) on these factors in more details. To understand the singular values of \( \frac{1}{\sqrt{p}} M \), we note that

\[
\left( \frac{1}{p} M^T M \right)_{\alpha \beta} = \langle \psi | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi \rangle \langle \psi | \mathcal{O}_\beta - \mathcal{O}_\beta(t) | \psi \rangle, \tag{11}
\]
where the overline denotes an average over the initial state ensemble. We denote that choosing the initial states from an ensemble of random product states provides a robust (and experimentally practical) scheme. Specifically, for a system of qubits we consider initial states given by

$$|\psi(0)\rangle = \prod_i |\phi_i\rangle,$$  \hspace{1cm} (12)

with $|\phi_i\rangle$ a random state on the Bloch sphere at site $i$. (Alternatively, for each site one can choose $|\phi_i\rangle = |\pm 1\rangle$ in either the $X,Y,$ or $Z$ basis.) Our results hereafter will assume the random Bloch sphere states as the initial states.

We analyze a generic class of Hamiltonians with random onsite and nearest-neighbor interactions in a spin$-1/2$ chain:

$$H = \sum_{i=1}^{L} \sum_{\alpha=1}^{3} c_{ia} \sigma_i^\alpha + \sum_{i=1}^{L-1} \sum_{a,b=1}^{3} c_{iab} \sigma_i^a \sigma_{i+1}^b.$$  \hspace{1cm} (13)

Here $c_{ia}$ and $c_{iab}$ are random variables and $\sigma^{1,2,3} = \{X,Y,Z\}$ are Pauli operators. We will average over random Hamiltonian realizations $c_{ia}, c_{iab} \in (-1,1)$ and error realizations in $(-\epsilon, \epsilon)$. In Fig. 2, we plot the reconstruction error vs. $p$ and confirm the $1/\sqrt{p}$ dependence in Eq. (10). We use the experimentally reasonable timescale $t = 1$ for $L = 8$ and observe that $p = 2n$ number of pairs is sufficient to achieve a high fidelity of 0.98 when $\epsilon = 0.1$.

Figure 2 also shows that our approach performs substantially better than evolving a fixed initial state with different times. This is because in the latter approach, if the times are too large, the final states will be locally thermal and hardly distinguishable. On the other hand, states separated by small time are also hard to distinguish. Using these states in Eq. (9) would result in $M$ and its gap $\lambda$ being smaller. This is also evident when we use our method and choose a small $t$ (see reconstruction fidelity versus $t$ in Fig. 3; the fidelity at $t \approx 0$ is nonzero because the overlap between two random vectors in $\mathbb{R}^n$ is roughly $1/\sqrt{n}$). It is only at later times that $M$ and $\lambda$ become sufficiently nonzero, and the fidelity approaches unity.

To understand the behavior of the fidelity curve in Fig. 3 and in particular its late time value, we have the following result. Consider the correlation function $C_{\alpha\beta}(t) = \frac{1}{D} \text{Tr}[O_{\alpha}(0)O_{\beta}(t)]$ (an $n \times n$ matrix) and let $l_{\text{max}}$ be the maximum size of an operator in $\{O_{\alpha}\}$. For example, the operator $Z_i Z_j$ has size 2 regardless of the distance $|i-j|$; a local Hamiltonian has $l_{\text{max}}$ of $O(1)$. Using Eq. (11) and analyzing operator growth [34], we find

$$\lambda(t) \geq \left(\frac{1}{3}\right)^{l_{\text{max}}/2}s_2(I - C(t)),$$  \hspace{1cm} (14)

where $s_2(I - C(t))$ denotes the 2nd smallest singular value of $I - C(t)$. Due to operator scrambling, we expect $C(t)$ to decay in time; assuming the eigenstate thermalization hypothesis (ETH), matrix elements of $C(\infty)$ are $O\left(\frac{1}{D}\right)$ [39] with fluctuations exponentially small in system size [40]. We use similar techniques to find that $s_2(I - C(\infty)) \geq 1 - O(1/L)$. Eq. (14) thus provides an $O(1)$ lower bound for the gap.

We emphasize that the above lower bound is independent of system size. Indeed, we find in Fig. 3 that the late time reconstruction fidelity is insensitive to system size. However, the time scale at which the maximal fidelity is reached depends on the time scale at which $C(t)$ asymptotes, which is expected to increase at most polynomially with system size. Note that the fidelity timescale also depends on the error magnitude: in the limit of zero error, the timescale for saturation approaches zero. We observe in the numerics that for experimentally reasonable $L = 10$, error $\epsilon = 0.1$, and pairs of states $p = 2n$, an $O(1)$ time is more than sufficient to reach maximal fidelity.

Stability against errors in the ansatz of the Hamiltonian.—In reality, interactions are not strictly compactly supported; there are small nonlocal interactions. This motivates consideration of an enlarged Hamiltonian

$$H = \sum_{\alpha} c_{\alpha} O_{\alpha} + \sum_{\beta} c'_{\beta} O'_{\beta}$$  \hspace{1cm} (15)

in which the support of $O'_{\beta}$ is not necessarily bounded, but we assume $\|c'\| \ll \|c\|$.

Restricted by experimental and computational resources, it is desirable to model only the dominant interactions. Assuming the real Hamiltonian is described
by Eq. (15) and we use our method to reconstruct an “effective” Hamiltonian $H = \sum_{\alpha} x_{\alpha} O_{\alpha}$, how different will $x$ be from $c$?

The matrix in Eq. (9) now consists of two pieces $(M; M')$ from $\{O_\alpha\}, \{\tilde{O}_\beta\}$ respectively. The actual coefficient vector $\begin{pmatrix} c \\ c' \end{pmatrix}$ is the singular vector of $(M; M')$ with singular value 0. By restricting the operator set to $\{O_\alpha\}$, one only measures $M$ and calculates its minimal singular vector $x$.

We find [34] that the reconstruction error between $x, c$ is controlled by

$$E \leq \frac{\frac{1}{2} M'_{p2}}{s_2\left(\frac{1}{2} M\right)} \|c'\|.$$  \hfill (16)

Recall that $s_2\left(\frac{1}{2} M\right)$ denotes the second smallest singular value of $\frac{1}{2} M$. Using ETH, we show that $s_2\left(\frac{1}{2} M\right) \geq \left(\frac{1}{3}\right)^{l_{\max}/2} - O(\frac{1}{M})$, where $l_{\max}$ is the size of the largest operator in $\{O_\alpha\}$ (not $\{\tilde{O}_\beta\}$). Therefore $\left(\frac{1}{3}\right)^{l_{\max}/2}$ is $O(1)$ and does not vanish even though the full Hamiltonian may contain arbitrarily nonlocal operators. We also show that $\left\|\frac{1}{2} M'_{p2}\right\|$ has an $O(1)$ bound. Thus, as long as $\|c'\| \ll \|c\|$, our reconstruction will succeed.

Summary.—We have shown that a single quantum quench is sufficient in principle to reconstruct a generic many-body Hamiltonian with local interactions. We also propose a practical version involving multiple quantum quenches from random initial product states and requiring only measurement of local observables. Using ETH, we analytically bound the reconstruction error arising from measurement errors and ignorance of nonlocal interactions. The efficiency and robustness of our protocol enable quantum simulators to determine precisely the Hamiltonian being implemented.

Acknowledgements.—We thank Ivaro Alhambra, Ehud Altman, Xiaoliang Qi, and Beni Yoshida for helpful discussions, and Gabriela Secara for help with figure design. Z.L. is grateful for the hospitality of the Visiting Graduate Fellowship program at Perimeter Institute where this work was carried out. Z.L. is also supported by the PQI fellowship from Pittsburgh Quantum Institute. L.Z. is supported by the John Bardeen Postdoctoral Fellowship. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

[1] R. Blatt and C. F. Roos, “Quantum simulations with trapped ions,” Nature Physics 8, 277–284 (2012).

[2] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, “Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator,” Nature 551, 601–604 (2017).

[3] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe, “Emergence and frustration of magnetism with variable-range interactions in a quantum simulator,” Science 340, 583–587 (2013).

[4] M. H. Devoret and R. J. Schoelkopf, “Superconducting circuits for quantum information: an outlook,” Science 339, 1169 (2013).

[5] Jay M. Gambetta, Jerry M. Chow, and Matthias Steffen, “Building logical qubits in a superconducting quantum computing system,” npj Quantum Information 3, 2 (2017).

[6] Markus Greiner, Olaf Mandel, Tilman Esslinger, Theodor W. Hänsch, and Immanuel Bloch, “Quantum phase transition from a superfluid to a mott insulator in a gas of ultracold atoms,” Nature 415, 39–44 (2002).

[7] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger, “Many-body physics with ultracold gases,” Rev. Mod. Phys. 80, 885–964 (2008).

[8] Hannes Bernien, Sylvain Schwartz, Alexander Keesling, Harry Levine, Ahmed Omran, Hannes Pichler, Soonwon Choi, Alexander S. Zibrov, Manuel Endres, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin, “Probing many-body dynamics on a 51-atom quantum simulator,” Nature 551, 579–584 (2017).

[9] Robert Jördens, Niels Strohmaier, Kenneth Günter, Henning Moritz, and Tilman Esslinger, “A mott insulator of fermionic atoms in an optical lattice,” Nature 455, 204 EP – (2008).

[10] Christian Gross and Immanuel Bloch, “Quantum simulations with ultracold atoms in optical lattices,” Science 357, 995–1001 (2017).

[11] Anton Mazurenko, Christie S. Chiu, Geoffrey Ji, Maxwell F. Parsons, Márton Kanász-Nagy, Richard Schmidt, Fabian Grusdt, Eugene Demler, Daniel Greif, and Markus Greiner, “A cold-atom fermi–hubbard antiferromagnet,” Nature 545, 462–466 (2017).

[12] B. Chiaro et al., “Growth and preservation of entanglement in a many-body localized system,” (2019), arXiv:1910.06024 [cond-mat.dis-nn].

[13] S. G. Schirmer, A. Kolli, and D. K. L. Oi, “Experimental hamiltonian identification for controlled two-level systems,” Phys. Rev. A 69, 050306 (2004).

[14] Simon J. Devitt, Jared H. Cole, and Lloyd C. L. Hollenberg, “Scheme for direct measurement of a general two-qubit hamiltonian,” Phys. Rev. A 73, 052317 (2006).

[15] Jared H. Cole, Sonia G. Schirmer, Andrew D. Greenree, Cameron J. Wellard, Daniel K. L. Oi, and Lloyd C. L. Hollenberg, “Identifying an experimental two-state hamiltonian to arbitrary accuracy,” Phys. Rev. A 71, 062312 (2005).

[16] Jared H. Cole, Simon J Devitt, and Lloyd C L Hollenberg, “Precision characterization of two-qubit hamiltonians via entanglement mapping,” Journal of Physics A: Mathematical and General 39, 14649–14658 (2006).

[17] C. Di Franco, M. Paternostro, and M. S. Kim, “Hamiltonian tomography in an access-limited setting without state initialization,” Phys. Rev. Lett. 102, 187203 (2009).

[18] Daniel Burgarth and Koji Maruyama, “Indirect hamilt-
nian identification through a small gateway,” New Journal of Physics 11, 103019 (2009).
[19] Daniel Burgarth, Koji Maruyama, and Franco Nori, “Coupling strength estimation for spin chains despite restricted access,” Phys. Rev. A 79, 020305 (2009).
[20] Daniel Burgarth, Koji Maruyama, and Franco Nori, “Indirect quantum tomography of quadratic hamiltonians,” New Journal of Physics 13, 013019 (2011).
[21] Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin, “Practical characterization of quantum devices without tomography,” Phys. Rev. Lett. 107, 210404 (2011).
[22] C. Senko, J. Smith, P. Richerme, A. Lee, W. C. Campbell, and C. Monroe, “Coherent imaging spectroscopy of a quantum many-body spin system,” Science 345, 430–433 (2014).
[23] Jun Zhang and Mohan Sarovar, “Quantum hamiltonian identification from measurement time traces,” Phys. Rev. Lett. 113, 080401 (2014).
[24] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and D. G. Cory, “Hamiltonian learning and certification using quantum resources,” Phys. Rev. Lett. 112, 190501 (2014).
[25] Sheng-Tao Wang, Dong-Ling Deng, and L-M Duan, “Hamiltonian tomography for quantum many-body systems with arbitrary couplings,” New Journal of Physics 17, 093017 (2015).
[26] Ruichao Ma, Clai Owens, Aman LaChapelle, David I. Schuster, and Jonathan Simon, “Hamiltonian tomography of photonic lattices,” Phys. Rev. A 95, 062120 (2017).
[27] Philipp Hauke, Maciej Lewenstein, and André Eckardt, “Tomography of band insulators from quench dynamics,” Phys. Rev. Lett. 113, 045303 (2014).
[28] P. Jurcevic, P. Hauke, C. Maier, C. Hempel, B. P. Lanyon, R. Blatt, and R. Blatt, “Spectroscopy of interacting quasiparticles in trapped ions,” Phys. Rev. Lett. 115, 100501 (2015).
[29] James R. Garrison and Tarun Grover, “Does a single eigenstate encode the full hamiltonian?” Phys. Rev. X 8, 021026 (2018).
[30] Xiao-Liang Qi and Daniel Ranard, “Determining a local Hamiltonian from a single eigenstate,” Quantum 3, 159 (2019).
[31] Eli Chertkov and Bryan K. Clark, “Computational inverse method for constructing spaces of quantum models from wave functions,” Phys. Rev. X 8, 031029 (2018).
[32] Eyal Bairey, Itai Arad, and Netanel H. Lindner, “Learning a local hamiltonian from local measurements,” Phys. Rev. Lett. 122, 020504 (2019).
[33] Tim J. Evans, Robin Harper, and Steven T. Flammia, “Scalable bayesian hamiltonian learning,” (2019), arXiv:1912.07636 [quant-ph].
[34] See the Supplemental Material for detailed derivations of the claims in the main text.
[35] David A. Cox, John B. Little, and Donal O’Shea, Ideals, varieties, and algorithms an introduction to computational algebraic geometry and commutative algebra (Springer Int. Publ., Switzerland, 2015).
[36] Gregory Bentsen, Ionut-Dragos Potirniche, Vir B. Balchandani, Thomas Scaffidi, Xiangyu Cao, Xiao-Liang Qi, Monika Schleier-Smith, and Ehud Altman, “Integrable and chaotic dynamics of spins coupled to an optical cavity,” Phys. Rev. X 9, 041011 (2019).
[37] T. Tony Cai and Anru Zhang, “Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics,” Ann. Statist. 46, 60–89 (2018).
[38] Why not fix $p$ and perform more measurements $k$ for each matrix element $M_{i\alpha}$ since the central limit theorem also guarantees that errors decay as $1/\sqrt{k}$? We find that in the experimentally relevant regime of $p$ marginally greater than $n$, error decreases faster than $1/\sqrt{p}$ (see Fig. 2), and thus our approach is more efficient.
[39] Yichen Huang, Fernando G. S. L. Brandão, and Yong-Liang Zhang, “Finite-size scaling of out-of-time-ordered correlators at late times,” Phys. Rev. Lett. 123, 010601 (2019).
[40] Álvaro M. Alhambra, Jonathon Riddell, and Luis Pedro García-Pintos, “Time evolution of correlation functions in quantum many-body systems,” Phys. Rev. Lett. 124, 110605 (2020).
I. Single quench tomography

A. Dependence in the equal moments conditions

To solve for $H$, we make use of the “equal moments conditions”, which we copy here for convenience:

$$
\langle \psi(t) | H^m | \psi(t) \rangle = \langle \psi(0) | H^m | \psi(0) \rangle, \quad \forall m \in \mathbb{N}.
$$

Here, we show that at most $(D - 1)$ of them are independent, where $D$ is the dimension of the Hilbert space.

We decompose $H$ as $H = \sum_i E_i | i \rangle \langle i |$, then Eq. (1) becomes:

$$
\sum_{i=1}^{D} a_i E_i^m = 0,
$$

where $a_i = \langle \psi(t) | i \rangle^2 - \langle \psi(0) | i \rangle^2$. For simplicity, we first consider the case where $H$ has no degeneracy. In this case, the first $D$ equations ($m = 0, 1, \cdots, D - 1$) are enough to show that $a_i = 0$ ($\forall i$). This is because the above equation can be viewed as a linear equation of $a_i$, whose coefficient matrix is the the Vandermonde matrix and its
determinant is nonzero (provided that $H$ has no degeneracy):
\[ \begin{vmatrix} 1 & E_1 & E_2 & \cdots & E_{D-2} & E_{D-1} \\ 1 & E_2 & E_2 & \cdots & E_{D-2} & E_{D-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & E_D & E_2 & \cdots & E_{D-2} & E_{D-1} \end{vmatrix} = \prod_{1 \leq i \leq j \leq D} (E_j - E_i) \neq 0, \quad (3) \]

Now that we know $a_i = 0$, Eq. (2) holds for all $m \in \mathbb{N}$. Note that the equation with $m = 0$ is redundant: $\sum_i a_i = 0$ automatically due to the normalization condition $\langle \psi(t)| \psi(t) \rangle = \langle \psi(0)| \psi(0) \rangle = 1$. Therefore, at most $(D-1)$ equations in Eq. (1) are independent.

If $H$ has degeneracy, we can simply group $a_i$ together for those $i$'s with the same $E_i$. The same argument shows that at most $(D - 1)$ of them are independent, where $D$ is the number of different eigenvalues of $H$ and $D \leq D$.

In summary, at most $(D - 1)$ equations in Eq. (1) are independent.

**B. Uniqueness of reconstruction**

Recall that we want to determine $\{c\}$ from Eq. (1) with $m = 1, 2, \cdots, n$. More explicitly:
\[
\begin{aligned}
M_{\alpha_1}^{(1)} x_{\alpha_1} &= 0, \\
M_{\alpha_2}^{(2)} x_{\alpha_2} x_{\alpha_2} &= 0, \\
\quad \ldots \ldots \\
M_{\alpha_n}^{(n)} x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_n} &= 0,
\end{aligned}
\]

where
\[ M_{\alpha_1 \alpha_2 \cdots \alpha_m}^{(m)} = \langle \psi(0)| O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_n} \psi(0) \rangle - \langle \psi(t)| O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_n} \psi(t) \rangle. \quad (5) \]

Here we use $\{x\}$ as symbols of unknown variables, to distinguish it from $\{c\}$.

In this section, we prove the following theorem:

**Theorem 1.** Fix the system and fix a basis $\{O_{\alpha}\}$ for local operators. If there exists a Hamiltonian $H_0 = \sum_{\alpha} c_{\alpha}^0 O_{\alpha}$ such that the coupling constants $c_{\alpha}^0$ can be uniquely determined on $\mathbb{C}$ (up to a multiplicative factor) from the above equations, then almost all $H = \sum_{\alpha} c_{\alpha} O_{\alpha}$ can be uniquely determined (up to a factor).

**Proof.** Without loss of generality, we assume $c_{n}^0 \neq 0$. We will show that at most a measure zero subset of $\mathbb{R}^n$ (the space of $c$) may fail to be determined uniquely. Since $\{c_n = 0\}$ has measure zero, in the following we only need to consider the subspace $\mathbb{R}^n \setminus \{c_n = 0\}$.

The polynomial system Eq. (4) contains $n$ equations for $n$ homogeneous variables, it is (formally) over-determined. However, by construction, it always has at least one nonzero solution $x = c$. Having at least one nonzero solution is equivalent to that a polynomial of the coefficients $M$, called Macaulay resultant, vanishes $[1]$:
\[ \text{poly}_1(M) = 0. \quad (6) \]

Here, we need a further result.

**Lemma 2.** When $c_n \neq 0$, then Eq. (4) has solutions other than $x = c$ (including those with $x_n = 0$) if and only if some polynomials (denoted by poly$_2$ collectively) of the coefficients $\{M\}$ vanish simultaneously.

**Proof of Lemma.** Denote the original homogeneous equations as
\[ F^{(m)}(x_1, \cdots, x_n) = \sum_{\alpha_1 \alpha_2 \cdots \alpha_m} M_{\alpha_1 \alpha_2 \cdots \alpha_m} x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_m} = 0, \quad (m = 1, 2, \cdots, n). \quad (7) \]

Under the following transformation:
\[
\begin{aligned}
x_\alpha &\rightarrow y_\alpha + c_\alpha y_n \quad (\alpha = 1, 2, \cdots, n - 1), \\
x_n &\rightarrow c_n y_n,
\end{aligned}
\]

\[ \text{poly}_2(M) = 0. \]
the original polynomial equations become
\[ G^{(m)}(y_1, \ldots, y_n) = F^{(m)}(y_1 + c_1 y_n, \ldots, y_{n-1} + c_{n-1} y_n, c_n y_n) = 0, \]  
(9)
which can be expanded as:
\[ G^{(m)}(y_1, \ldots, y_n) = \sum_{j=0}^{m} G^{(m)}_j(y_1, \ldots, y_{n-1}) y_n^{m-j}, \]
(10)
where \( G^{(m)}_j \) is a homogeneous polynomial of \((y_1, \ldots, y_{n-1})\) with degree \(j\). Solutions of Eq. (7) are in one-to-one correspondence to solutions of Eq. (10), since the transformation Eq. (8) is invertible. The obvious solution \(x = c\) of the original equations now becomes \(y_1 = \cdots = y_{n-1} = 0, \; y_n \neq 0\), which imples:
\[ G^{(m)}_0(y_1, \ldots, y_{n-1}) = 0, \; \forall m. \]
(11)
Therefore, we want to show that the following homogeneous equations
\[ G^{(m)}(y_1, \ldots, y_n) = \sum_{j=1}^{m} G^{(m)}_j(y_1, \ldots, y_{n-1}) y_n^{m-j} \]
(12)
has no solutions other than \([0, 0, \ldots, 0, 1]\) if and only if the coefficients are special in the sense that some polynomial relations of the coefficients are satisfied.

Denote \(\{a\}\) to be the set of coefficients (of \(\{y\}\)) in those \(G^{(m)}(y)\). Denote \(N\) to be the number of coefficients \(\{a\}\). Now let us regard \(G^{(m)}\) as polynomials of both \(a\) and \(y\) (linear in \(a\), homogeneous in \(y\)), \(G^{(m)}(a, y)\). Construct the ideal
\[ J = (G^{(1)}(a, y), G^{(2)}(a, y), \ldots, G^{(N)}(a, y)), \]
(13)
it defines a variety on \(\mathbb{C}^N \times \mathbb{C}^{p-1}\). Consider the ideal \(I\) corresponding to \(\mathbb{C}^N \times \{y_1 = \cdots = y_{n-1} = 0\}\), namely,
\[ I = (y_1, y_2, \cdots, y_{n-1}). \]
(14)
We construct the ideal quotient \[2\) \((J : I^\infty)\). This construction serves to exclude \(y_1 = \cdots = y_n = 0\) (which should be excluded since we are considering homogeneous variables) and the obvious solution \([0, \ldots, 0, 1]\). Then the variety \(\pi((J : J^\infty))\) is exactly the space of those extra solutions \((a, y)\), and the projection of \(\pi((J : J^\infty))\) to \(\mathbb{C}^N\), denoted by \(\pi((J : J^\infty))\), is exactly the space of “special” coefficients.

According to the main theorem of elimination [2], \(\pi((J : J^\infty))\) is Zariski closed (note that to use the main theorem of elimination, it is essential to work on \(\mathbb{C}\) instead of \(\mathbb{R}\), and to project along projective space \(\mathbb{C}^{p-1}\) instead of \(\mathbb{C}^{n-1}\)). In other words, it can be described as the set of common zeros for several polynomials of \(a\). Since each \(a\) is a linear combination of original coefficients \(M\) in Eq. (7), we conclude that Eq. (7) has solutions other than \(x \propto c\) if and only if some polynomial relations (“poly\(_2\)) of the coefficients \(M\) are satisfied, at least when \(c_n \neq 0\).

\[ \square \]

Back to the original problem. The success of uniquely reconstructing \(H_0\) implies:
\[ \text{poly}_2(M(e^0, t, \psi(0))) \neq 0. \]
(15)
According to Eq. (5), each \(M^{(m)}_{a_1a_2\cdots a_n}\) is a real-analytic function of \(\{c_{a_i}\}, t\) and \(|\psi(0)\rangle\). Therefore poly\(_2\)(\(M\)) is also real-analytic in \(\{c_{a_i}\}, t\) and \(|\psi(0)\rangle\). It can be shown [3] that a real-analytic function is nonzero almost everywhere if it is nonzero at one point. Therefore, Eq. (15) implies \(\text{poly}_2(M(c, t, \psi(0))) \neq 0\) for generic \(c\), which implies the success for uniquely reconstructing a generic \(H\).

\[ \square \]

C. Determining the multiplicative factor

Suppose we know \(H\) up to a multiplicative factor, we show that generically this factor can be determined from \(|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle\) and the knowledge of \(|\psi(t)\rangle, |\psi(0)\rangle\) and \(t\).

We will show that, equivalently, \(e^{-iHt} |\psi(0)\rangle\) is generically injective in \(t\) (even modulo the phase). Physically speaking, generically the Poincare recurrence cannot be exact. This means that, generically, given \(|\psi(0)\rangle, e^{-iHt} |\psi(0)\rangle\)
and $e^{-i\alpha H_t}|\psi(0)\rangle$ cannot be identical state, $\forall \alpha \neq 1$. So when $|\psi(t)\rangle$ and $t$ are also given, $H$ can be generically determined without the freedom of multiplying a factor.

We expand $H$ as $H = \sum_i E_i |i\rangle \langle i|$. Then:

$$|\psi(t)\rangle = \sum_i e^{-iE_it} \langle i|\psi(0)\rangle |i\rangle .$$  \hspace{1cm} (16)

If $|\psi(t)\rangle \propto |\psi(t')\rangle$, then

$$e^{-iE_it} = e^{-iE_{i't'}-i\theta}, \text{ for } i \text{ such that } \langle i|\psi(0)\rangle \neq 0,$$

or equivalently,

$$\left( E_i - \frac{\theta}{t - t'} \right) (t - t') = 2\pi p_i, \; p_i \in \mathbb{Z}, \text{ for } i \text{ such that } \langle i|\psi(0)\rangle \neq 0. $$

(18)

This is impossible generically, since generically $\langle i|\psi(0)\rangle \neq 0$ for $\forall i$, and $E_i$ are mutually incommensurate even after shift.

The above analysis also provides us an algorithm to find the scale. Before presenting the algorithm, we note that any such algorithm will require precise information about $H$ and $|\psi\rangle$, and will be numerically unstable due to approximate Poincare recurrence.

Assuming we know $H$ up to multiplicative factor, and $\psi(0), \psi(t)$ up to phases, then Eq. (16) gives us:

$$E_it = i\log \frac{\langle i|\psi(t)\rangle}{\langle i|\psi(0)\rangle} \overset{\text{def}}{=} \alpha_i + 2\pi p_i + \theta, \text{ for } i \text{ such that } \langle i|\psi(0)\rangle \neq 0.$$  \hspace{1cm} (19)

Denote $E_{ij} = E_i - E_j$ (and similarly for $\alpha_{ij}$ and $p_{ij}$). Picking up $i, j, k$ such that $\frac{E_{ij}}{E_{jk}}$ is irrational and $\langle i|\psi(0)\rangle, \langle j|\psi(0)\rangle, \langle k|\psi(0)\rangle \neq 0$, then:

$$p_{ij}E_{jk} - p_{jk}E_{ij} = \frac{E_{ij}\alpha_{jk} - E_{jk}\alpha_{ij}}{2\pi}.$$  \hspace{1cm} (20)

This is a $\mathbb{Z}$–linear combination of two incommensurate numbers $E_{jk}$ and $E_{ij}$, therefore the coefficients can be found uniquely (by exhaustive search). Then $t$ can be determined from:

$$E_{ij}t = \alpha_{ij} + 2\pi p_{ij}.$$  \hspace{1cm} (21)

**D. Analytical checks of uniqueness**

**1. Linear approximation**

According to the Taylor expansion of $e^{-iHt}$, the coefficients in the above equations become:

$$M^{(m)}_{\alpha_1\alpha_2\ldots\alpha_m} = it \langle \psi(0)| [H, O_{\alpha_1} O_{\alpha_2} \cdots O_{\alpha_m}] |\psi(0)\rangle + O(t^2) \overset{\text{def}}{=} \tilde{M}^{(m)}_{\alpha_1\alpha_2\ldots\alpha_m} + O(t^2).$$  \hspace{1cm} (22)

We want to take $\tilde{M}$ as an approximation to $M$, and use this approximation to understand whether the original polynomial system, Eq. (4), has a unique solution.

First, note that Eq. (4) with $M$ substituted by $\tilde{M}$ always has at least one solution, i.e., $x = c$. This can be seen by fixing $x = c$ and then expanding Eq. (4) in $t$. For the uniqueness, the following statement justifies our linear approximation. It tells us that if the linearized equation has a unique solution, then the solution of the original equation is also unique.

**Claim 3.** If there exists a poly$_2$, denoted by $f$, such that $f(\tilde{M}(t)) \neq 0$ as a function of $t$, then $f(M(t)) \neq 0$ for small enough $t$ and also for generic $t$.

**Proof.** We use the following simple observation: under rescaling $M \rightarrow \lambda M$, solutions of Eq. (7) remain unchanged. Therefore, the set of poly$_2$ actually generates a homogeneous ideal and hence $f$ can be assumed to be homogeneous without loss of generality.

Due to the homogeneity of $f$, if we expand $f(M(t))$ in $t$, the term with the lowest degree in $t$ will exactly be $f(\tilde{M})$. Therefore, $f(\tilde{M}) \neq 0$ implies $f(M(t)) \neq 0$ for small enough $t$, which also implies that $f(M(t)) \neq 0$ for generic $t$ due to analyticity. 

\qed
2. Translationally invariant systems

Our formalism simplifies a lot for translationally invariant systems. In this case, the ansatz for the Hamiltonian will be:

\[ H = \sum_{\alpha} c_\alpha \sum_{j=1}^{L} O_{\alpha j}, \]  

where \( L \) is the number of sites, \( O_{\alpha j} \) are local operators at site \( j \). Our equal moments condition has the same form as Eq. (4) with \( O_\alpha \) now equal to \( \sum_{j=1}^{L} O_{\alpha j} \) (\( \alpha \) goes over local operator types):

\[
M^{(m)}_{\alpha_1 \alpha_2 \ldots \alpha_m} = \langle \psi(0) \mid (\sum_{j=1}^{L} O_{\alpha_1 j})(\sum_{j=1}^{L} O_{\alpha_2 j}) \ldots (\sum_{j=1}^{L} O_{\alpha_m j}) \mid \psi(0) \rangle - \langle \psi(t) \mid (\sum_{j=1}^{L} O_{\alpha_1 j})(\sum_{j=1}^{L} O_{\alpha_2 j}) \ldots (\sum_{j=1}^{L} O_{\alpha_m j}) \mid \psi(t) \rangle.
\]

The linear approximation will be determined by:

\[
\tilde{M}^{(m)}_{\alpha_1 \alpha_2 \ldots \alpha_m} = it \langle \psi(0) \mid \left[ H, (\sum_{j=1}^{L} O_{\alpha_1 j})(\sum_{j=1}^{L} O_{\alpha_2 j}) \ldots (\sum_{j=1}^{L} O_{\alpha_m j}) \right] \mid \psi(0) \rangle.
\]

3. A spin chain example

We consider the spin chain example in the main text:

\[ H = c_1 \sum_i Z_i Z_{i+1} + c_2 \sum_i X_i + c_3 \sum_i Y_i Y_{i+2}. \]  

(26)

In this model, there are three types of local operators, so we will try to use three equations to determine \( c \). We will show that the linear approximation equations, with \( \mid \psi(0) \rangle \) taken as a translationally invariant product state \( \mid \psi \rangle = \prod_i \mid n_i \rangle \) (\( \mid n_i \rangle \) is a fixed state on the Bloch sphere), indeed generically have a unique solution.

The linearized equations are as follows:

\[
\sum_{i=1}^{L} c_\delta \langle \psi \mid [O_{\delta i}, O_{\alpha i}] \mid \psi \rangle x_\alpha = 0,
\]

\[
\sum_{i=1}^{L} \sum_{j=1}^{L} c_\delta \langle \psi \mid [O_{\delta i}, O_{\alpha i} O_{\beta j}] \mid \psi \rangle x_{\alpha} x_{\beta} = 0,
\]

\[
\sum_{i=1}^{L} \sum_{j=1}^{L} \sum_{k=1}^{L} c_\delta \langle \psi \mid [O_{\delta i}, O_{\alpha i} O_{\beta j} O_{\gamma k}] \mid \psi \rangle x_{\alpha} x_{\beta} x_{\gamma} = 0.
\]

(27)

In the above equations, \( c_\delta \) are constants, \( x_\alpha, x_\beta, x_\gamma \) are variables, summation over \( \alpha, \beta, \gamma, \delta \) are omitted. We only keep \( O_{\delta i} \) instead of \( \sum_i O_{\delta i} \) due to translational invariance. Here, we will only briefly describe what happens instead of presenting the explicit form of the above equations\(^1\).

For the 1st equation, note that only when local terms \( O_{\delta i} \) and \( O_{\alpha i} \) have overlap can the commutator be nonzero, hence only a few terms contribute.

For the 2nd equation, commutators survive only if \( O_{\delta i} \) has overlap with \( O_{\alpha i} O_{\beta j} \). There are two possibilities:

- When one of \( i \) and \( j \) is far away from 1, say \( j \), then it is just \( [O_{\delta 1}, O_{\alpha i}] O_{\beta j} \). This produces a term proportional to \( L \) as well as an \( O(1) \) “offset” (see below for an explicit example). The term proportional to \( L \) is exactly the one in the 1st equation.
- When \( O_{\delta 1}, O_{\alpha i} \) and \( O_{\beta j} \) are “connected”, we get \( O(1) \) non-vanishing terms.

\(^1\) It is lengthy and not easy to work out by hand. We did it symbolically on computer.
At the end, only $O(1)$ terms coming from the connected case and $O(1)$ terms coming from the offset of the disconnected case contribute.

For example, consider the case such that $\delta \leftrightarrow ZZ, \alpha \leftrightarrow ZZ, \beta \leftrightarrow X$, then

$$
Z_1Z_2, \left( \sum_i Z_iZ_{i+1}\right) \left( \sum_i X_i \right) = \left( \sum_i Z_iZ_{i+1}\right)(iY_1Z_2 + iZ_1Y_2),
$$

and

$$
\langle \psi | Z_1Z_2, \left( \sum_i Z_iZ_{i+1}\right) \left( \sum_i X_i \right) | \psi \rangle = \langle \psi | \left( Z_0Z_1 + Z_1Z_2 + Z_2Z_3 + \sum_{i \neq 0,1,2} Z_iZ_{i+1}\right)(iY_1Z_2 + iZ_1Y_2) | \psi \rangle
$$

$$
= 2(n_x^2n_x + n_x + in_yn_z) + (2L - 6)n_z^2(in_yn_z).
$$

Here $n_x = \langle n | X | n \rangle$, etc. The term proportional to $L$ is given by a corresponding term $\langle \psi | Z_1Z_2, \left( \sum_i X_i \right) | \psi \rangle$ in the first equation.

For the 3rd equation, the situation is similar. The final equation contains $L^2, L^1$ and $L^0$ terms; the $L^2$ term is exactly the 1st equation while the $L^1$ term can be derived from the 1st and 2nd equation. Only the $L^0$ term (from connected correlations and offsets) contribute.

We calculated the solution of the 1st and 2nd equation, and found that there are two solutions up to a factor. This is as expected, since a line and a quadratic curve in two dimensions (two independent unknowns here) generically has two intersections. Moreover, it can be explicitly checked that only one of the solution survives the 3rd equation.

In summary, we have analytically proved that our procedure works for the model in Eq. (26).

II. Multiple quench tomography

A. Homogeneous linear regression

In this section, we consider the following homogeneous linear regression problem. Assume

$$
\sum_\alpha M_{i\alpha}x_\alpha = 0,
$$

for all $i$, but the knowledge of $M$ comes with errors. How should we get a good estimation for the vector $x$?

Let us apply the least square method to estimate $x$, i.e., we will find $x$ which minimizes

$$
S = \sum_i \text{error}_i^2 = \sum_i (\sum_\alpha M_{i\alpha}x_\alpha)^2 = x^TM^TMx.
$$

However, since each equation is homogeneous, a constraint should be imposed to fix the overall magnitude of $x$ (otherwise the least square solution will be $x = 0$). We use the following natural constraint:

$$
\sum_\alpha x_\alpha^2 = x^2 = 1.
$$

This constraint can also be regarded as fixing $\text{Tr}(H^2)$ if $\text{Tr}(O_\alpha O_{\alpha'}) \propto \delta_{\alpha\alpha'}$.

Applying the method of Lagrange multipliers, we get the following equations:

$$
M^TMx - \lambda x = 0
$$

$$
x^T x = 1,
$$

where $\lambda$ is the Lagrange multiplier. Therefore, $x$ is an eigenvector of $M^TM$ with eigenvalue $\lambda$. Plug back into the original penalty function Eq. (31), we get:

$$
S = \lambda x^Tx = \lambda.
$$

Therefore, we should pick the vector $x$ corresponding to the minimal $\lambda$. In order words, the least square solution of the linear regression problem is the singular vector of $M$ corresponding to the minimal singular value $\sqrt{\lambda}$. 
B. Stability against error

In this section we prove that the average reconstruction error caused by errors in $M$ is upper bounded by

$$C \sqrt{\frac{m \epsilon}{p \lambda}},$$

(35)

when $\epsilon$ is small enough. Here $\lambda$ is the gap between the smallest singular value (0) and the 2nd smallest singular value of $\frac{1}{\sqrt{p}} M$. In this subsection, we will assume that $M$ is gapped ($\lambda > 0$), otherwise the above inequality holds trivially.

1. Perturbative analysis

We would like to compare the singular vectors of $M + E$ and $M$, where $E$ is a $p \times n$ matrix or errors. Only in this section, to simplify the computation, we model the error to be i.i.d $N(0, \epsilon)$. Due to properties of Gaussian distributions, elements of $U^\dagger EV$ are also iid $N(0, \epsilon)$ for any unitary matrices $U$ ($p$ by $p$) and $V$ ($n$ by $n$). Therefore, we can assume $M$ is in its singular-value-decomposed form:

$$
\begin{bmatrix}
\sqrt{p} \lambda_1 \\
\sqrt{p} \lambda_2 \\
\sqrt{p} \lambda_{n-1} \\
0 \\
0 \\
0 \\
0
\end{bmatrix}.
$$

(36)

Now we perform standard perturbation theory on $M^T M$. Since

$$\Delta(M^T M) = (M + E)^T (M + E) - M^T M = M^T E + E^T M + E^T E,$$

(37)

Let $|1\rangle$ be the eigenvector of $M^T M$ with eigenvalue 0. Then the perturbed eigenvector is:

$$|\bar{1}\rangle = |1\rangle - \sum_{i=2}^{n} \frac{\langle i | M^T E + E^T M | 1 \rangle |i\rangle}{p \lambda_i^2} + O(\epsilon^2).$$

(38)

Using Eq. (36) and $M |1\rangle = 0$, we have:

$$\langle i | M^T E + E^T M | 1 \rangle = \langle i | M^T E | 1 \rangle = \sqrt{p} \lambda_i \epsilon_i,$$

(39)

where $\epsilon_i = \langle i | E | 1 \rangle$ is a $N(0, \epsilon)$ random variable. Therefore,

$$|\bar{1}\rangle = |1\rangle - \sum_{i=2}^{n} \frac{\epsilon_i}{\sqrt{p} \lambda_i} |i\rangle + O(\epsilon^2),$$

(40)

and the angle between reconstructed and actual couplings is given by

$$|\sin \theta| = \left\| \sum_{i=2}^{n} \frac{\epsilon_i}{\sqrt{p} \lambda_i} |i\rangle \right\| + O(\epsilon^2)$$

$$= \left\| \sum_{i=2}^{n} \frac{\epsilon_i^2}{p \lambda_i} \right\| + O(\epsilon^2).$$

(41)

On average, we will have

$$|\sin \theta| \leq \sqrt{\frac{n \epsilon}{p \lambda}} + O(\epsilon^2),$$

(42)

where $\bar{\lambda}^2$ is the harmonic mean of $\lambda_i^2$, which is no less than the minimal one ($\lambda$). Therefore the expectation value of error $\sin \theta$ is bounded by Eq. (35).
2. Nonperturbative analysis

Strictly speaking, one needs to worry about the higher order terms in the perturbation theory, even if the perturbation theory converges. It is also possible that when $\epsilon$ is some $o(1)$ value (as $n \to \infty$ or $p \to \infty$), the 1st order result is already not a good approximation. Therefore, we need some nonperturbative results to estimate the error in the realistic case, where $\epsilon$ is at constant level (albeit small).

The problem of singular vector perturbation has been actively studied (nonperturbatively) in the mathematical and statistical literature. Here we point out the following result which is relevant to us, see theorem 3 in Ref. 4:

\[
\mathbb{E} \|\sin \Theta(V, V')\|^2 \leq \frac{Cn(p\lambda^2\epsilon^2 + p\epsilon^4)}{p^2\lambda^4} \leq \frac{2Cn}{p} \frac{\epsilon^2}{\lambda^2} \quad \text{(if } \epsilon \leq \lambda). \tag{43}
\]

Here, $C$ is some absolute constant; $V$ and $V'$ are the space spanned by the first $n - 1$ (right) singular vectors of $M$ and $M + E$, respectively, therefore $\|\sin \Theta(V, V')\| = |\sin \theta|$.

C. Analysis of the gap

Recall that

\[ M_{i\alpha} = \langle \psi_i | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi_i \rangle, \tag{44} \]

where $i = 1 \cdots p$ label the initial-final states pair, $\alpha = 1 \cdots n$ label the parameters in $H$. Define a random variable

\[ X = (\langle \psi | \mathcal{O}_1 - \mathcal{O}_1(t) | \psi \rangle, \cdots, \langle \psi | \mathcal{O}_n - \mathcal{O}_n(t) | \psi \rangle). \tag{45} \]

It is a random variable since $|\psi\rangle$ comes from a random ensemble. Then the matrix elements of $\frac{1}{p}M^TM$ are

\[ \frac{1}{p}(M^TM)_{\alpha\beta} = \frac{1}{p} \sum_{i=1}^{p} \langle \psi_i | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi_i \rangle \langle \psi_i | \mathcal{O}_\beta - \mathcal{O}_\beta(t) | \psi_i \rangle = \frac{1}{p} \sum_{i=1}^{p} X_i^T X_i. \tag{46} \]

Therefore, $\frac{1}{p}M^TM$ is the sample covariance matrix (not centered) of the random variable $X$, or in other words the sample average of $X^TX$.

In the following, we will use the actual covariance matrix of $X$ as an estimation of $\frac{1}{p}M^TM$:

\[ \frac{1}{p}M^TM \leftarrow \mathbb{E}[X^TX] = \langle \psi | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi \rangle \langle \psi | \mathcal{O}_\beta - \mathcal{O}_\beta(t) | \psi \rangle, \tag{47} \]

where the overline means ensemble average over initial states $|\psi\rangle$. In other words, we take $p \to \infty$ limit. The behavior at finite $p$ will be discussed in subsection II C 5.

We are free to choose the initial states ensemble. But in any case, when $t = 0$ we have $M = 0$ and the gap of $M/\sqrt{p}$ is $\lambda = 0$. This is why at short times the fidelity is low.

1. Haar random initial states

In this case, a crude estimation will show that the stability is bad. Indeed, both the harmonic mean gap $\bar{\lambda}$ and the minimal gap are bounded by the arithmetic mean of the eigenvalues:

\[ \lambda^2 \leq \bar{\lambda}^2 \leq \frac{1}{n - 1} \text{Tr}(\frac{1}{p}M^TM) = \frac{1}{n - 1} \sum_{\alpha=1}^{n} |\langle \psi | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi \rangle|^2, \tag{48} \]

With Haar integral,

\[ |\langle \psi | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi \rangle|^2 = \int d\psi |\langle \psi | \mathcal{O}_\alpha - \mathcal{O}_\alpha(t) | \psi \rangle|^2 = \frac{1}{D(D + 1)} [\text{Tr}^2(\mathcal{O}_\alpha - \mathcal{O}_\alpha(t)) + \text{Tr}[(\mathcal{O}_\alpha - \mathcal{O}_\alpha(t))^2]] \leq \frac{4}{D + 1}. \tag{49} \]

Here we have used a Haar average formula and the fact that $\mathcal{O}_\alpha$ is traceless. This exponentially small gap makes the reconstruction very sensitive to the error.
2. Constraint on operator growth from energy conservation

Assuming $H$ has generic spectrum, then the late time value of the coefficient of $A$ in the expansion of $B(t)$ is:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{1}{D} \text{Tr}(AB(t)) dt = \lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{1}{D} \sum_{E,E'} \langle E| A| E' \rangle e^{iE't} \langle E'| B| E \rangle e^{-iEt} dt = \frac{1}{D} \sum_E A_E B_E.\quad (50)$$

with $A_E \equiv \langle E| A| E \rangle$ and $B_E \equiv \langle E| B| E \rangle$. For traceless local operators $A, B$, we will derive a formula for this quantity for systems with geometrically local interaction, assuming eigenstate thermalization hypothesis (ETH), following Ref. [5].

ETH tells us that there exist smooth enough functions $f_A$ and $f_B$ such that:

$$|A_E - f_A(E/L)| \leq \frac{1}{\text{poly}(L)},\quad (51)$$

and similarly for $B$, where $L$ is the number of sites (volume), and $\text{poly}(L)$ means some polynomial of $L$ with high enough degree. Moreover, it can be shown that for a trace-less local operator $A$ [5]:

$$f_A(0) = 0, \quad \frac{f'_A(0)}{L} = \frac{\text{Tr}(HA)}{\text{Tr}(H^2)},\quad (52)$$

and similarly for $B$. Therefore,

$$\frac{1}{D} \sum_E A_E B_E = \frac{1}{D} \sum_E f'_A(0)\frac{E}{L} f'_B(0)\frac{E}{L} E + \frac{1}{D} \sum_{E \leq L^{1+\epsilon}} \left( A_E B_E - f'_A(0)\frac{E}{L} f'_B(0)\frac{E}{L} \right) + \frac{1}{D} \sum_{E > L^{1+\epsilon}} \left( A_E B_E - f'_A(0)\frac{E}{L} f'_B(0)\frac{E}{L} \right) = \frac{\text{Tr}(HA)\text{Tr}(HA)}{D\text{Tr}(H^2)} + O\left( \frac{1}{L^2} \right).\quad (53)$$

The fact that the error is $O\left( \frac{1}{L^2} \right)$ can be proved with some technical results in Ref. [5]. The first term here is $O\left( \frac{1}{L^2} \right)$. For example, if $A = \mathcal{O}_\alpha, B = \mathcal{O}_\beta, H = \sum_{\alpha=1}^N c_\alpha \mathcal{O}_\alpha$, and $\text{Tr}(\mathcal{O}_\alpha \mathcal{O}_\beta) \propto \delta_{\alpha,\beta}$, then the first term is just:

$$\sum_{\alpha=1}^N c_\alpha c_\beta.\quad (54)$$

Eq. (53) tells us that if both $A$ and $B$ have overlap with the Hamiltonian, then the weight of $A$ in $B(t)$ will be only polynomially small, $O\left( \frac{1}{L^2} \right)$, instead of exponentially small if one were to naively replace the time evolution operator by Haar random unitary. This can be intuitively understood from energy conservation: an $O(1)$ perturbation of $A$ gives an $O(1)$ perturbation to the conserved energy. This perturbation is then evenly distributed among each $B(t)$, so that the correlation of $A$ and $B(t)$ is of order $O\left( \frac{1}{T} \right)$.

Regarding the deviation between $C(t)$ and the ETH value in Eq. (53), denoted by $\frac{1}{D} \text{Tr}(AB(\infty))$, it is proved in Ref. [6] that:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \left( \frac{1}{D} \text{Tr}(AB(t)) - \frac{1}{D} \text{Tr}(AB(\infty)) \right)^2 dt \leq \frac{1}{D} \|A\|^2 \|B\|^2.\quad (55)$$

We note that the only assumption used to prove this inequality is just a generic energy spectrum, rather than the more stringent ETH [6].
3. Random Product States: Exact Results

Suppose at $t = 0$ all $O_{\alpha}$’s are Pauli strings normalized such that $\text{Tr}(O_{\alpha}(0)O_{\beta}(0)) = D\delta_{\alpha\beta}$. Let us expand $O_{\alpha}(t)$ as

$$O_{\alpha}(t) = \sum_s P_s C_{s,\alpha}(t),$$

(56)

where $P_s$ means Pauli operator with string $s$, e.g., $s = (\cdots, I, I, Z, X, I, I, \cdots)$. Note $s \neq (\cdots, I, I, I, \cdots)$ because $O_{\alpha}$ is taken to be traceless. The coefficients $C_{s,\alpha}(t)$ depends on $s, i, t,$ and $c$. Due to energy conservation, $H(t) = H$. So for each $s$, we have:

$$\sum_\alpha C_{s,\alpha}(t)c_\alpha = \sum_\alpha C_{s,\alpha}(0)c_\alpha = \begin{cases} c_\alpha, & s = O_{\alpha} \text{ for some } \alpha \\ 0, & s \notin \{O\}. \end{cases}$$

(57)

Also note that $\text{Tr}(O_{\alpha}(0)O_{\beta}(0)) = D\delta_{\alpha\beta}$ implies

$$\sum_s C_{s,\alpha}(t)C_{s,\beta}(t) = \delta_{\alpha\beta}.$$  

(58)

Now consider the covariance:

$$\frac{1}{p}(M^TM)_{\alpha\beta} = \langle \psi | O_{\alpha} - O_{\alpha}(t) | \psi \rangle \langle \psi | O_{\beta} - O_{\beta}(t) | \psi \rangle.$$

(59)

We note the following formula for random product states average:

$$\langle \psi | P_s | \psi \rangle \langle \psi | P_s' | \psi \rangle = \delta_{ss'} \left( \frac{1}{3} \right)^{|s|}, \quad (l_s = \text{size of } P_s).$$

(60)

Here $P_s$ means a Pauli string. Alternatively, one can also use a random $XYZ$ ensemble. Namely, each spin can choose from $X = \pm 1, Z = \pm 1, Z = \pm 1$ with equal $(\frac{1}{6})$ probability. The above formula still holds for this ensemble. With this formula, the covariance can be expanded as:

$$\frac{1}{p}(M^TM)_{\alpha\beta} = \left( \frac{1}{3} \right)^{|s|} \delta_{\alpha\beta} - \left( \frac{1}{3} \right)^{|s|} C_{\beta,\alpha}(t) - \left( \frac{1}{3} \right)^{|s|} C_{\alpha,\beta}(t) + \sum_s C_{s,\alpha}(t)C_{s,\beta}(t)\left( \frac{1}{3} \right)^{|s|}$$

$$= \left[ \left( \frac{1}{3} \right)^{|s|} \delta_{\alpha\beta} - \left( \frac{1}{3} \right)^{|s|} C_{\beta,\alpha}(t) - \left( \frac{1}{3} \right)^{|s|} C_{\alpha,\beta}(t) + \sum_{s \in \{O\}} C_{s,\alpha}(t)C_{s,\beta}(t)\left( \frac{1}{3} \right)^{|s|} \right] + \sum_{s \notin \{O\}} C_{s,\alpha}(t)C_{s,\beta}(t)\left( \frac{1}{3} \right)^{|s|}$$

$$= A_{\alpha\beta} + \tilde{A}_{\alpha\beta}.$$  

(61)

Here $\{O\}$ is the set of local operators in the Hamiltonian.

Define three matrices $B, C(t), \tilde{C}(t)$ as:

$$B_{\alpha\beta} = \left( \frac{1}{3} \right)^{|s|} \delta_{\alpha\beta},$$

$$C_{s\alpha}(t) = \text{expansion coefficients in Eq. (56)} \quad (s \in \{O\})$$

$$\tilde{C}_{s\alpha}(t) = \text{expansion coefficients in Eq. (56)} \quad (s \notin \{O\}).$$

(62)

Then

$$A(t) = B - C(t)^T B - BC(t) + C(t)^T BC(t)$$

$$= (I - C(t)^T)B(I - C(t)).$$

(63)

Therefore, $A(t)$ is a positive semi-definite matrix. Similarly, $\tilde{A}(t) = \tilde{C}(t)^T B \tilde{C}(t)$ is also positive semi-definite.

Moreover, $A$ has the same zero mode as $\frac{1}{p}M^TM$. Indeed, since $\frac{1}{p}M^TM = A + \tilde{A}$ with both term positive semi-definite, a zero mode of $\frac{1}{p}M^TM$ must be a zero mode of $A$. Conversely, if $Ax = 0$, then we must have $x = Cx$ since $B$ is positive definite. This also implies $\tilde{C}x = 0$ since Eq. (58) is just

$$C^TC + \tilde{C}^T \tilde{C} = I.$$
We will use the following result\(^2\): if \(X \succeq Y\) as Hermitian operators (in the sense that \(X - Y\) is positive semidefinite), then
\[
\sigma_i(X) \geq \sigma_i(Y)
\]
for all \(i\), where \(\sigma_i(X)\) and \(\sigma_i(Y)\) are the \(i\)-th smallest eigenvalues of \(X\) and \(Y\), respectively. Since \(\tilde{A}\) is positive semi-definite, we have:
\[
A + \tilde{A} \succeq A \geq \left(\frac{1}{3}\right)^{l_{\max}} (I - C^T)(I - C),
\]
where \(l_{\max}\) is maximal length of Pauli operators that appear in \(\{O_\alpha\}\). Therefore
\[
\lambda^2 = \sigma_2(A + \tilde{A}) \geq \sigma_2\left(\left(\frac{1}{3}\right)^{l_{\max}} (I - C^T)(I - C)\right).
\]

Note that we have actually proved the existence of a gap of \(M\), therefore the uniqueness of the reconstruction, without referring to “generic”. This is a consequence of ETH. Indeed, if there are other conserved quantities (and our approach for Hamiltonian tomography will not yield a unique solution), the ETH does not hold anymore.

4. Random Product States: Understanding from ETH

If we use the late time value Eq. (54) predicted by ETH, and ignore the \(O(\frac{1}{L^2})\) term, then:
\[
C \leftarrow \frac{cc^T}{\|c\|^2} = C_0.
\]
The spectrum of \((I - C_0^T)(I - C_0)\) will just be \(0, 1, \cdots, 1\), where the \(n - 1\) eigenvectors with eigenvalue 1 are orthogonal to \(c\), and \(c\) is the last eigenvector with eigenvalue 0. Then we have:
\[
\lambda^2 \geq \left(\frac{1}{3}\right)^{l_{\max}}.
\]
Taking the \(O(\frac{1}{L^2})\) correction and the exponentially small fluctuation Eq. (55) into consideration, we obtain:
\[
\lambda \geq \left(\frac{1}{3}\right)^{l_{\max}/2 - O(\frac{1}{L})}.
\]

As a side note, we may want to take \((I - C_0^T)B(I - C_0)\) as an approximation to \(\frac{1}{p} M^T M\). Recall that \(\frac{1}{p} M^T M = A(t) + \tilde{A}(t)\), this approximation amounts to ignore \(\tilde{A}(t), A(t) - A(\infty)\), and the \(O(\frac{1}{L})\) correction in the end value. It turns out the approximation works quite well.

In Fig. 1, we plot the histogram of eigenvalues of \(\frac{1}{p} M^T M\). In this example, we use random 1-body and 2-body Pauli operators as the basis \(\{O_\alpha\}\), so \(l = 1, 2\). The matrix \((I - C_0^T)B(I - C_0)\) can be diagonalized exactly, and the spectrum is:
\[
0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}.
\]
Here \(\mu_1\) is the square sum of those \(c_\alpha\) corresponding to operators \(O_\alpha\) with length 1. Note that \(\frac{1}{3} - \frac{2}{9}\mu_1^2 \in \left[\frac{1}{9}, \frac{1}{3}\right]\), the gap \(\lambda^2\) is always \(\frac{1}{9}\), as expected. The above spectrum concentrates at \(\frac{1}{3}\) and \(\frac{1}{3}\), with one exception at 0 the other one between \(\frac{1}{9}, \frac{1}{3}\). As can be seen from Fig. 1, this approximation matches the actual spectrum very well.

\(^2\) This result is an easy corollary of Weyl’s inequality, or can be proved by the Courant min-max principle.
5. Random Product States: Effects of finite $t$ and $p$

In the above, we have used Eq. (47) to approximate $\frac{1}{p} M^{T} M$. This approximation is accurate as $p \to \infty$ due to the law of large numbers.

To discuss the accuracy when $p$ is finite, we need to consider the covariance of $X_{\alpha}X_{\beta}$:

$$\text{cov}(X_{\alpha}X_{\beta}, X_{\gamma}X_{\delta}) = \mathbb{E}[X_{\alpha}X_{\beta}X_{\gamma}X_{\delta}] - \mathbb{E}[X_{\alpha}X_{\beta}]\mathbb{E}[X_{\gamma}X_{\delta}].$$

(72)

Obviously $|X_{\alpha}| \leq 2$, since $||O_{\alpha} - O_{\alpha}(t)|| \leq 2$. Therefore $\text{cov}(X_{\alpha}X_{\beta}, X_{\gamma}X_{\delta}) = O(1)$. According to the central limit theorem,

$$\left| \frac{1}{p} (M^{T} M)_{\alpha\beta} - \mathbb{E}[(X^{T}X)]_{\alpha\beta} \right| \leq \frac{O(1)}{\sqrt{p}} \quad (\forall \alpha, \beta),$$

(73)

for typical sample. The maximum difference between the gap of $\frac{1}{p} (M^{T} M)$ ($p$ finite) and the gap of $\mathbb{E}[(X^{T}X)]$ ($p$ infinite) will be $\frac{O(n)}{\sqrt{p}}$. Therefore, as long as $p = \Omega(n^2)$, the approximation Eq. (47) will be reasonable.

This is just a rough (but rigorous) estimation. If $\frac{1}{p} M^{T} M - \mathbb{E}[(X^{T}X)]$ behaves like a random matrix, then typical difference of the gaps will be controlled by $O(\frac{n}{\sqrt{p}})$ and $p = \Omega(n)$ will be enough to make the approximation Eq. (47) valid.

In Fig. 2, we plot the gap vs. $p$, when $L = 8$ and $t = 10$. We find that the gap grows with $p$. This is consistent with our observation that at small $p$, the error decreases faster than $\frac{1}{\sqrt{p}}$. We can also see that the gap at large $p$ is consistent with our prediction Eq. (70).
In Fig. 3, we plot the gap vs. $t$ for different $p$. We see that the gap grows with $t$, consistent with the intuition that a large time interval makes the fidelity better.

### D. Stability against ignorance

Assuming the exact Hamiltonian is:

$$H = \sum_\alpha c_\alpha O_\alpha + \sum_\beta c'_\beta O_\beta,$$

(74)

but one may only pick up part of the local operators $\{O_\alpha\}$ and obtain the following reconstructed Hamiltonian:

$$H = \sum_i x_\alpha O_\alpha.$$

(75)

The reconstruction matrix now splits into two parts:

$$M_{\text{tot}} = (M; M').$$

(76)

The actual coefficient vector $\left(\begin{array}{c} c \\ c' \end{array}\right)$ is the singular vector of $M_{\text{tot}}$ with singular value 0. By restricting the operator set to $\{O_\alpha\}$, one only measures $M$ and calculates its minimal singular vector $x$.

**Theorem 4.** The reconstruction error $\sin \theta$ between $x$ and $c$ is controlled by

$$|\sin \theta| \leq \frac{\|\sqrt{\frac{1}{p}} M' c'\|}{s_2(\sqrt{\frac{1}{p}} M) \|c\|} \leq \frac{\|\sqrt{\frac{1}{p}} M'\|}{s_2(\sqrt{\frac{1}{p}} M) \|c\|} \|c'\| \|c\|,$$

(77)

where $s_2(\sqrt{\frac{1}{p}} M)$ denotes the second smallest singular value of $\sqrt{\frac{1}{p}} M$.

**Proof.** We will use $s_2$ to represent $s_2(\sqrt{\frac{1}{p}} M)$ for notational simplicity. Assuming $s_2$ is nonzero for now, otherwise the above inequality is trivially valid. Because $\left(\begin{array}{c} c \\ c' \end{array}\right)$ is the coupling vector of the full Hamiltonian, $\sqrt{\frac{1}{p}} M c + \sqrt{\frac{1}{p}} M' c' = 0$ and we have:

$$\left\| \sqrt{\frac{1}{p}} M c \right\| = \left\| \sqrt{\frac{1}{p}} M' c' \right\| \overset{\text{def}}{=} \epsilon \|c\|. $$

(78)

This implies $\sqrt{\frac{1}{p}} M$ has at least one singular value in $[0, \epsilon]$. Let us assume $\epsilon < s_2$, otherwise the theorem is again trivially valid. Therefore, $\sqrt{\frac{1}{p}} M$ must have exactly one singular value in $[0, \epsilon]$ and no singular values in $(\epsilon, s_2)$. Denote the singular vector to be $\bar{c}$ (so this is what one will get by our reconstruction method) and decompose $c$ as $c^\parallel + c^\perp$ with respected to $\bar{c}$. Then

$$\epsilon \|c\| = \left\| \sqrt{\frac{1}{p}} M c \right\| \geq s_2 \|c^\perp\|. $$

(79)

Denote the angle between $c$ and $\bar{c}$ to be $\theta$, then

$$\sin \theta = \frac{\|c^\perp\|}{\|c\|} \leq \frac{\epsilon}{s_2} = \frac{\|\sqrt{\frac{1}{p}} M' c'\|}{s_2 \|c\|} = \frac{\|\sqrt{\frac{1}{p}} M'\|}{s_2} \|c'\|. $$

(80)

This is the desired upper bound for the reconstruction error. 

\square
In the above expression, \( \left\| \sqrt{\frac{1}{p} M'} \right\| \) is bounded by an \( O(1) \) number. Indeed, according to Eq. (63) and Eq. (64), we have:

\[
x^T \left( \frac{1}{p} M_{\text{tot}}^T M_{\text{tot}} \right) x = x^T (I - C^T) B (I - C)x + x^T \hat{C}^T B \hat{C} x
\leq \frac{1}{3} x^T \left[ (I - C^T)(1 - C) + \hat{C}^T \hat{C} \right] x
= \frac{1}{3} x^T \left( 2 - C^T - C \right) x.
\]

According to definition Eq. (56), we have:

\[
|x^T Cx| = \frac{1}{D} \left| \text{Tr} \left[ Q(0)^\dagger Q(t) \right] \right| \leq \frac{1}{D} \sqrt{\text{Tr}(Q(0)^\dagger Q(0)) \text{Tr}(Q(t)^\dagger Q(t))} = \|x\|^2.
\]

where \( Q(t) = \sum x_\alpha O_\alpha(t) \) (recall that \( \text{Tr}(O_\alpha(0)O_\beta(0)) = D \delta_{\alpha \beta} \)). Similarly \( |x^T C^T x| \leq \|x\|^2 \). Therefore, \( x^T \left( \frac{1}{p} M_{\text{tot}}^T M_{\text{tot}} \right) x \leq \frac{1}{4} \|x\|^2 \), which implies

\[
\left\| \frac{1}{p} M' \right\| \leq \left\| \frac{1}{p} M_{\text{tot}} \right\| \leq \frac{2}{\sqrt{3}}.
\]

About the second smallest singular value \( s_2 \), we can prove \( s_2 \neq 0 \) just from the existence of a gap of \( M_{\text{tot}} \).

**Claim 5.** If \( M_{\text{tot}} \) has a gap \( \lambda \) between its minimal singular value 0 and 2nd minimal singular value, then \( M \) has at most one singular value in \([0, \sqrt{2}]\). Therefore \( s_2 \geq \sqrt{2} \).

**Proof.** If not, assuming \( x_1 \) and \( y_1 \) are two (orthogonal) singular vectors of \( M \) with singular values \( s_x \) and \( s_y \), \( s_x, s_y \in [0, \sqrt{2}] \). Denote \( x = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \), \( y = \begin{pmatrix} y_1 \\ 0 \end{pmatrix} \). Then

\[
\| M_{\text{tot}} x \| = \| M x_1 \| = s_x \|x\|.
\]

We decompose \( x \) as \( x^\parallel + x^\perp \) (similarly for \( y \)), where \( \| \) means parallel to \( \begin{pmatrix} c \\ \epsilon \end{pmatrix} \), the singular vector of \( M_{\text{tot}} \) with singular value 0. We have

\[
\frac{\lambda}{\sqrt{2}} \|x\| \geq s_x \|x\| = \| M_{\text{tot}} x \| \geq \lambda \|x^\perp\|,
\]

hence \( \|x^\perp\| < \frac{1}{\sqrt{2}} \|x\| \) and \( \|x^\parallel\| > \frac{1}{\sqrt{2}} \|x\| \). Similar inequalities hold for \( y \). Therefore

\[
|(x, y)| = |(x^\parallel, y^\parallel) + (x^\perp, y^\perp)| \geq |(x^\parallel, y^\parallel)| - |(x^\perp, y^\perp)| > 0.
\]

However \( (x, y) = (x_1, y_1) = 0 \), which leads to a contradiction. \( \square \)

The gap \( \lambda \) of \( M_{\text{tot}} \) can be very small due to the presence of longer-range interactions in \( \mathcal{O}' \). However, we can use the operator expansion as before to get a more reasonable estimation on \( s_2 \). In this case, Eq. (61) is still true for the \( M \) here \( s \in \mathcal{O} \) only includes those in \( \mathcal{O} \), not \( \mathcal{O}' \). We also have:

\[
s_2^2 \geq \sigma_2 (A) = \sigma_2 ((I - C^T) B (I - C)) \geq \frac{1}{3} l_{\text{max}}^2 \sigma_2 ((I - C^T)(I - C)).
\]

Note that here \( Cc \neq c \) so we may no longer have 0 singular value. Here \( l_{\text{max}} \) is the max length of an operator in \( \mathcal{O} \) (instead of \( \mathcal{O} \cup \mathcal{O}' \)), so it is still small.

From ETH (Eq. (54)), we have \( C_{\alpha \beta} = \frac{c_{\alpha} c_{\beta}}{\|c\|^2 + \|c'\|^2} + O(\frac{1}{2^2}) \equiv (C_0)_{\alpha \beta} + O(\frac{1}{2^2}) \). Since the spectrum of \( (I - C_0^T)(I - C_0) \) is

\[
\left( 1 - \frac{\|c\|^2}{\|c\|^2 + \|c'\|^2} \right)^2, 1, 1, \ldots, 1,
\]

\( n \) is
we get:

\[ s_2^2 \geq \left( \frac{1}{3} \right)^{l_{\max}} - O\left( \frac{1}{L} \right). \]  

(89)

[1] I.M. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*, Modern Birkhäuser Classics (Birkhäuser Boston, 2009).
[2] David A. Cox, John B. Little, and Donal O’Shea, *Ideals, varieties, and algorithms an introduction to computational algebraic geometry and commutative algebra* (Springer Int. Publ., Switzerland, 2015).
[3] Boris Mityagin, “The zero set of a real analytic function,” (2015), arXiv:1512.07276 [math.CA].
[4] T. Tony Cai and Anru Zhang, “Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics,” *Ann. Statist.* 46, 60–89 (2018).
[5] Yichen Huang, Fernando G. S. L. Brandão, and Yong-Liang Zhang, “Finite-size scaling of out-of-time-ordered correlators at late times,” *Phys. Rev. Lett.* 123, 010601 (2019).
[6] Álvaro M. Alhambra, Jonathon Riddell, and Luis Pedro García-Pintos, “Time evolution of correlation functions in quantum many-body systems,” *Phys. Rev. Lett.* 124, 110605 (2020).