Slowest local operators in quantum spin chains

Hyungwon Kim, Mari Carmen Bañuls, J. Ignacio Cirac, Matthew B. Hastings, and David A. Huse

Slowest local operators in quantum spin chains

Hyungwon Kim,[1, 2] Mari Carmen Bañuls,[3] J. Ignacio Cirac,[3] Matthew B. Hastings,[4, 5] and David A. Huse[1]

[1]*Physics Department, Princeton University, Princeton, NJ 08544, USA*
[2]*Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA*
[3]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany*
[4]*Station Q, Microsoft Research, Santa Barbara, CA 93106-6105, USA*
[5]*Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, USA*

We numerically construct slowly relaxing local operators in a nonintegrable spin-1/2 chain. Restricting the support of the operator to M consecutive spins along the chain, we exhaustively search for the operator that minimizes the Frobenius norm of the commutator with the Hamiltonian. We first show that the Frobenius norm bounds the time scale of relaxation of the operator at high temperatures. We find operators with significantly slower relaxation than the slowest simple "hydrodynamic" mode due to energy diffusion. Then, we examine some properties of the nontrivial slow operators. Using both exhaustive search and tensor network techniques, we find similar slowly relaxing operators for a Floquet spin chain; this system is hydrodynamically "trivial", with no conservation laws restricting their dynamics. We argue that such slow relaxation may be a generic feature following from locality and unitarity.

I. INTRODUCTION

It has been proposed that an isolated quantum many-body system with a few local conservation laws can still thermalize, in the sense that local observables approach their thermal equilibrium values [1–3] set by conserved quantities. Quantum many-body systems with extensive number of local conserved quantities (integrable systems) still relax to stationary states, which are described by the diagonal ensemble [4]. These stationary states of integrable systems can sometimes be equivalent to a generalized Gibbs ensemble [5–8] but not in general [9–11]. See Ref. 12 for a review. Experimental studies have increased the interest in these questions [13, 14]. One proposed theoretical explanation of thermalization of a generic quantum many-body system is the eigenstate thermalization hypothesis (ETH) [1–3, 15–21], which argues that many-body eigenstates of thermalizing nonintegrable Hamiltonians have local reduced density operators that are thermal; then, at long times, dephasing between different energy eigenstates brings small subsystems to thermal equilibrium.

However, not all systems show this local thermalization in an accessible time scale [22]. One possibility is that the ETH is false for the system of Ref. 22 but it seems unlikely [20]; another possibility is that the time scales required to thermalize locally are too long to be numerically accessible. A question is: how can such slow thermalization arise, when a nonintegrable system like the one studied has no conserved local quantities other than energy?

In this work, we illustrate how such slow relaxation can emerge by showing that indeed slow, almost-conserved local operators *are* present in many nonintegrable systems (note that Ref. 22 studied slowly thermalizing *initial states*). We construct these operators numerically, by explicitly searching for operators with a small commutator with the Hamiltonian. For numerical reasons discussed below, much of our work focuses on the Frobenius norm rather than the operator norm to measure the commutator, but we also discuss the operator norm and make a connection to Ref. 22.

One such operator with small commutator is the one that results from the thermal diffusion due to a spatially-smooth inhomogeneity in the energy density. This operator and its commutator with the Hamiltonian match the expectations from diffusive hydrodynamics, with the square of the Frobenius norm of the commutator decreasing as $\sim M^{-2}$ for the slowest such operator on a sub-region containing M spins. However, we show numerically that these are *not* the operators with the smallest commutator. We construct operators whose commutator with the Hamiltonian, for the accessible system sizes of $M \sim 100$, is quantitatively and substantially smaller (slower) than that of the simple diffusive mode, and appears to decrease with a *larger* power of length than the diffusive $\sim M^{-2}$. Thus we find "unexpected almost-conserved quantities" for these systems.

To further understand the presence of such approximately conserved quantities we turn to a Floquet spin chain where energy is not conserved. We also find slowly relaxing operators in this Floquet system. In fact, we argue that some slowly relaxing operators must be present in any Floquet system or more generally in any quantum circuit. However, these slowly relaxing operators are in a sense morally similar to the slowly relaxing operator in a Hamiltonian system describing energy fluctuations: these slow operators are present in any such system, so long as the unitary dynamics is local. They themselves do not inhibit relaxation of the local density matrices "as fast as possible" i.e., on a time scale proportional to the length of the interval; (see Ref. 23 for a proof that this happens for random local circuits). Thus, the real surprise is our numerical observation that there are other operators in

some nonintegrable Hamiltonian systems (and possibly in some quantum circuits) with even slower relaxation.

The rest of this work is organized as follows. In section II, we study a model Hamiltonian system, where energy conservation is the only local conservation law. We explain how to find the local operator of length M that gives the smallest Frobenius norm of the commutator with the Hamiltonian and we connect such quantity to thermalization time of local operators. Then, we show how the Frobenius norm decreases as we increase the length M of the operator. By comparing with the simple diffusive mode, we establish that there exist some local operators that do thermalize slower than diffusion. Next, we study the operator norm of the commutator with the Hamiltonian to make a more natural connection to the previous work. In section III, we study systems without any local conservation law. We study a Floquet system, which is a natural counterpart to the Hamiltonian system without energy conservation, and do the same analysis to find slow operators. We find slowly relaxing operators, whose thermalization time is bounded below by $1/M$. Then, we show that this phenomenon is generally true by finding the same scaling in random quantum circuits. In section IV, we summarize what we have found.

II. HAMILTONIAN SYSTEM: PRESENCE OF LOCAL CONSERVATION LAW

A. Model

As a nonintegrable model Hamiltonian, we choose a spin-1/2 Ising chain with both longitudinal and transverse fields:

$$H = \sum_{i=-\infty}^{\infty} g\sigma_i^x + h\sigma_i^z + \sigma_i^z \sigma_{i+1}^z \ , \qquad (1)$$

where σ_i^x and σ_i^z are Pauli matrices of the spin at site i. Ref. 22 has found a nonthermalizing state for this model within the accessible time scale. We choose $(g, h) = (0.905, 0.809)$, at which this model is known to be robustly nonintegrable even for a relatively small system size [24]. See appendix for another set of parameters and different model Hamiltonians.

B. Method

We consider local operators supported on a finite interval of M consecutive sites. (We have also studied translationally invariant operators [25].) Since every traceless Hermitian operator $\hat{A}_M$ can be expressed by a linear combination of $4^M - 1$ traceless Hermitian basis operators, we write

$$\hat{A}_M = \sum_{\ell=1}^{4^M-1} c_\ell \hat{O}_{(M,\ell)} \ , \qquad (2)$$

where c_ℓ is a real number and $\hat{O}_{(M,\ell)}$ is the corresponding basis operator. We choose $\hat{O}_{(M,\ell)}$ to be mutually orthogonal using the Hilbert-Schmidt inner product so that $\text{tr}(\hat{O}_{(M,\ell)}\hat{O}_{(M,k)}) = 0$ for $\ell \neq k$.

The dynamics of an operator comes from the commutator with the Hamiltonian. Therefore, we want to *minimize* the magnitude of $[\hat{A}_M, H]$ to construct a slowly relaxing local operator of length M. We use the square of the Frobenius norm, $\text{tr}(\hat{O}\hat{O}^\dagger)$, to quantify the commutator since it gives a quadratic form (Eq. (3)) of which we can readily find the minimum. Although the operator norm generally controls the dynamics of arbitrary states at arbitrary time, its numerical minimization is very challenging. As we show below, the (square of) Frobenius norm can actually bound the thermalization time scale at infinite temperature, where we expect the dynamics to be fastest. Furthermore, we can get the upper bound of the operator norm by using the operators that minimize the Frobenius norm. The behavior of upper bounds is consistent with the results we obtained using the Frobenius norm.

We minimize the following:

$$\begin{aligned}
f(\hat{A}_M) &= \frac{\text{tr}([\hat{A}_M, H][\hat{A}_M, H]^\dagger)}{\text{tr}(\hat{A}_M \hat{A}_M^\dagger)} \\
&= \sum_{\ell, k} \frac{c_\ell c_k \text{tr}([\hat{O}_{(M,\ell)}, H][\hat{O}_{(M,k)}, H]^\dagger)}{\sum_j c_j^2 \text{tr}((\hat{O}_{(M,j)})^2)} \ .
\end{aligned} \qquad (3)$$

We define $\lambda(M)$ to be the minimum of $f(\hat{A}_M)$: $\lambda(M) = \min\{f(\hat{A}_M)\}$, and we call the corresponding $\hat{A}_M$ the slowest operator acting on M sites. Since the Hamiltonian has time-reversal symmetry, we can consider even and odd operators under time-reversal separately. It turns out that for $M \geq 4$, the minimizer of $f(\hat{A}_M)$ always comes from the even sector. Up to $M = 11$, we obtain exact results and for larger M we minimize Eq. (3) using a matrix product operator (MPO) ansatz [26] for $\hat{A}_M$. For $M \leq 28$, we find the values of $\lambda(M)$ have converged within 1% error.

First, let's understand the physical meaning of $f(\hat{A}_M)$. We consider an initial mixed state $\rho = I/Z + \epsilon \hat{A}_M$, where Z is the normalization factor, I is the identity and ϵ is chosen to make ρ nonnegative. $\hat{A}_M$ serves as a small inhomogeneity in the infinite temperature ensemble and is assumed to have unit Frobenius norm. Let's define $a_M(t)$ as the expectation value of $\hat{A}_M/\epsilon$ at time t: $a_M(t) = (1/\epsilon)\text{tr}(\rho \hat{A}_M(t)) = \text{tr}(\hat{A}_M \hat{A}_M(t))$, where $\hat{A}_M(t)$ is in the Heisenberg picture. [Note that $a_M(0) = 1$.] Using Cauchy-Schwarz inequality, we have the following:

$$\left| \frac{d^2 a_M(t)}{dt^2} \right| = |\text{tr}([\hat{A}_M(t), H][\hat{A}_M, H])| \leq f(\hat{A}_M) \ , \quad (4)$$

Then, we can bound the distance between $a_M(t)$ and

$a_M(0)$.

$$|a_M(t) - a_M(0)| = \left| \int_0^t d\tau \int_0^\tau d\tau' \frac{d^2 a_M(\tau')}{d\tau'^2} \right| \leq \frac{f(\hat{A}_M)t^2}{2} . \tag{5}$$

In thermodynamic limit and finite M, the thermal expectation value is $a_M^{th} = \text{tr}(\hat{A}_M) = 0$. Therefore, the following inequality holds:

$$|a_M(t) - a_M^{th}| \geq |a_M(0) - a_M^{th}| - |a_M(0) - a_M(t)|$$

$$\geq 1 - \frac{f(\hat{A}_M)t^2}{2} . \tag{6}$$

Consequently, $f(\hat{A}_M)$ bounds the thermalization time scale τ of $\hat{A}_M$ from below by $\tau \geq f(\hat{A}_M)^{-1/2}$, and small $\lambda(M)$ implies a long thermalization time of $\hat{A}_M$. In addition, since $\lambda(M)$ is the minimum of Eq. (4) at $t = 0$, the optimal $\hat{A}_M$ is *the* slowest operator at early time.

C. Results and Comparison with energy diffusion

Figure 1 (a) plots $\lambda(M)$. It is clear that $\lambda(M)$ decreases with M as it should. The data can be well-fitted by two functional forms; power-law decay with exponent 2.62 and a logarithmic correction to $1/M^2$, which is, more precisely, a fit to $a/(M \ln(bM))^2$, with a and b as fitting parameters. In either case, the rate of decrease with M is *faster* than $1/M^2$, which is the scaling of the slowest diffusive energy mode as we show now.

A slowly relaxing energy mode can be constructed by considering an energy modulation of wavelength $2M$:

$$\hat{E}_M = \sum_i c_M(i) h_i = \sum_{i=-M/2}^{M/2} \cos\left(\frac{i\pi}{M}\right)(g\sigma_i^x + h\sigma_i^z)$$

$$+ \sum_{i=-M/2}^{M/2-1} \cos\left(\frac{(i+1/2)\pi}{M}\right) \sigma_i^z \sigma_{i+1}^z , \tag{7}$$

where h_i is the energy density operator $(g\sigma_i^x + h\sigma_i^z + 1/2(\sigma_{i-1}^z \sigma_i^z + \sigma_i^z \sigma_{i+1}^z))$ and $c_M(i)$ is the cosine modulation function restricted to lie in $[-M/2, M/2]$. Since the energy is conserved, we can use the continuity equation: $dh_i/dt = -\nabla \cdot \mathbf{j}_i$, where $\mathbf{j}_i$ is the energy current density at site i: $\mathbf{j}_i = g(\sigma_i^y \sigma_{i+1}^z - \sigma_{i+1}^y \sigma_i^z)$. Combining this with the Heisenberg equation of motion, we have the following.

$$i[H, \hat{E}_M] = \frac{d}{dt}\hat{E}_M = -\sum_i c_M(i)\partial_i \cdot \mathbf{j}_i \tag{8}$$

$$= \sum_i \mathbf{j}_i \partial_i c_M(i) \simeq -\frac{\pi}{M}\sum_i s_M(i) h_i, \tag{9}$$

where ∂_i is the discrete spatial derivative and $s_M(i)$ is the sine modulation. Therefore,

$$\frac{\text{tr}([H, \hat{E}_M][H, \hat{E}_M]^\dagger)}{\text{tr}(\hat{E}_M \hat{E}_M^\dagger)} \sim 1/M^2 . \tag{10}$$
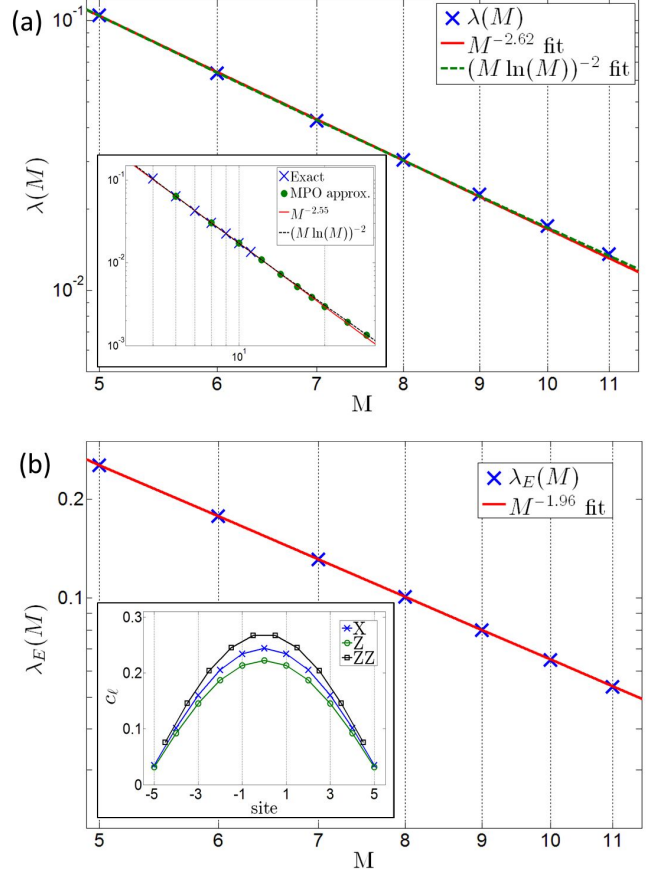

FIG. 1. (color online) (a) Decay of $\lambda(M)$, minimum of Eq. (3), as a function of M with power-law and logarithmic correction fits. Inset: $\lambda(M)$ vs. M for best MPO results (green dots). It can also be well fitted by both power-law with almost the same exponent (-2.62 for exact and -2.55 for MPO) and logarithmic correction. (b) $\lambda_E(M)$, the minimum of Eq. (3) with only terms in the Hamiltonian, vs. M. M^{-2} decay is consistent with thermal diffusion. Inset: Structure of the optimal operator for $M = 11$. $X = \sigma_\ell^x$, $Z = \sigma_\ell^z$, and $ZZ = \sigma_\ell^z \sigma_{\ell+1}^z$. We locate the $\sigma_\ell^z \sigma_{\ell+1}^z$ term at $\ell + 1/2$. Coefficients are normalized as $\sum_\ell c_\ell^2 = 1$. It shows a clear sinusoidal energy modulation with the expected wavelength $2M$, and with the relative ratios equal to those in the Hamiltonian.

Here, we adapt a more conservative approach. We do the same numerical search as before but restrict the operator space within the terms in the Hamiltonian so we only consider $3M - 1$ basis operators instead of $4^M - 1$. Figure 1 (b) plots $\lambda_E(M)$, the minimum of $f(\hat{A}_M)$ in this restricted space. As expected, we have almost perfect $1/M^2$ scaling. Since there are only a few basis operators, we can easily look at the details of the structure of the optimal operator. The inset of Figure 1 (b) is the structure of the optimal operator in terms of the local terms in Hamiltonian. It is indeed of the form of Eq. (7) and thus the energy modulation of the longest wavelength is the slowest energy mode. Note also the fact that apart from displaying a different scaling, $\lambda(M)$ is much smaller than $\lambda_E(M)$.

This hydrodynamic description, however, only considers operators that are *linear* in energy density operators. Combinations of higher power of energy density terms, such as $h_i h_j$, $h_i h_j h_k$, ..., may yield the slowest operator given by exhaustive search. Although we are unable to obtain the exact slowest operators, we find a deviation of the scaling of $\lambda(M)$ from M^{-2} by adding nonlinear energy density operators. Furthermore, most of the weight (measured by Hilbert-Schmidt norm) of the slowest operator turns out to be energy density modulation. In actuality, we can obtain slow operators (not slowest) by "dressing" the modulated form. See Appendix for details.

D. Operator Norm

The operator norm is mathematically more convenient for studies of time scale since we can directly interpret the relaxation of the operator in terms of Lieb-Robinson bounds [27, 28]. Optimizing the operator norm numerically is, however, very challenging. Therefore, we have minimized the (square of) Frobenius norm. Nevertheless, once we have an operator, it is easy to compute the operator norm of the commutator with the Hamiltonian.

Our first task is to relate the operator norm with the thermalization time scale. Let us assume that $\hat{A}_M$ satisfies the following:

$$||[\hat{A}_M, H]|| \leq \chi(M), \tag{11}$$

where H is the Hamiltonian and $||\ldots||$ means the operator norm of the argument and $\chi(M)$ is some nonnegative valued function. Then, using the Heisenberg equation of motion, we have

$$\left|\left|\frac{d}{dt}\hat{A}_M(t)\right|\right| = \left|\left|e^{-iHt}\left(\frac{d}{dt}\hat{A}_M(t)\right)e^{iHt}\right|\right|$$
$$= ||[\hat{A}_M(t=0), H]|| \leq \chi(M) \tag{12}$$

$$||\hat{A}_M(t) - \hat{A}_M(0)|| = \left|\left|\int_0^t \left(\frac{d}{d\tau}\hat{A}_M(\tau)\right)d\tau\right|\right|$$
$$\leq \int_0^t ||\frac{d}{d\tau}\hat{A}_M(\tau)||d\tau \leq \chi(M)t , \tag{13}$$

where we have used the fact that e^{-iHt} is a norm-preserving unitary operator. This inequality bounds the distance of an operator evolving under Hamiltonian dynamics from its initial configuration. Note that we have previously used the *square* of the Frobenius norm to bound the distance from the thermal value, where we find the bound of the thermalization time is given by $\sqrt{\lambda(M)}$.

Next, we consider an initial state ρ such that

$$|\langle \hat{A}_M \rangle_0 - \langle \hat{A}_M \rangle_\beta| = \gamma(M) , \tag{14}$$

where $\langle \ldots \rangle_0$ is the expectation value of the initial condition, $\langle \ldots \rangle_\beta$ is the thermal expectation value and $\gamma(M)$ is
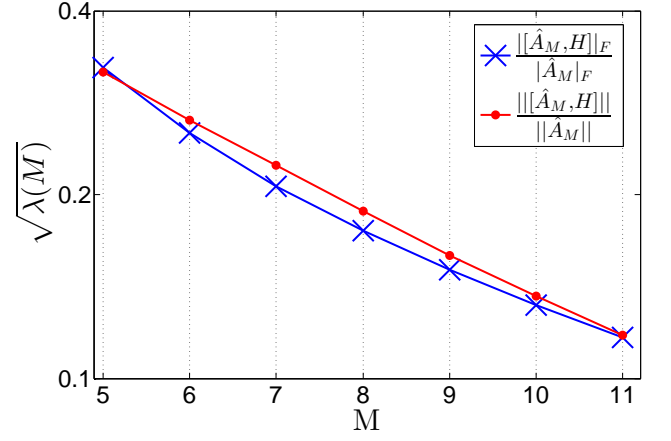

FIG. 2. (color online) Normalized magnitude of $[\hat{A}_M, H]$ measured by the Frobenius norm ($|\ldots|_F$) and the operator norm ($||\ldots||$) in log-linear plot. $\hat{A}_M$ is obtained by minimizing the Frobenius norm of the commutator with Hamiltonian as is in the main text. Two norms give quantitatively similar values.

some nonnegative valued function. Now we can estimate the distance between the thermal expectation value and the expectation value at time t (for $t \leq \gamma(M)/\chi(M)$):

$$|\langle \hat{A}_M \rangle_t - \langle \hat{A}_M \rangle_\beta|$$
$$\geq \left| |\langle \hat{A}_M \rangle_0 - \langle \hat{A}_M \rangle_\beta| - |\langle \hat{A}_M \rangle_t - \langle \hat{A}_M \rangle_0| \right|$$
$$\geq \gamma(M) - t\chi(M) \tag{15}$$

where $\langle \ldots \rangle_t$ is the expectation value at time t. Therefore, if we have a sequence of M-body operators $\{\hat{A}_M\}$ for which the operator norm of the commutator with the Hamiltonian decays fast with M and an initial state which does not allow a fast decrease of $\gamma(M)$, the time scale of thermalization of $\hat{A}_M$ is

$$\tau_M \geq \frac{\gamma(M)}{\chi(M)} . \tag{16}$$

In particular, if $\chi(M)$ decreases faster than a power law with M, thermalization may take longer than polynomial time in contrast to the case of diffusion.

Figure 2 plots the decay of the Frobenius norm and the operator norm of $[\hat{A}_M, H]$, where $\hat{A}_M$ is obtained in the main text by minimizing the Frobenius norm. It shows that the value of λ measured by the operator norm is numerically similar to that of the Frobenius norm. It appears that the operator norms computed for these operators exhibit exponential-like decrease with M but the range of M is not enough to draw a conclusion.

1. Initial State and Thermalization Time Scale

Ref. 22 reports nonthermalization of spin polarized initial state along x direction and weak thermalization of spin polarized initial state along z direction. Having slowest operators from the Frobenius norm, we study
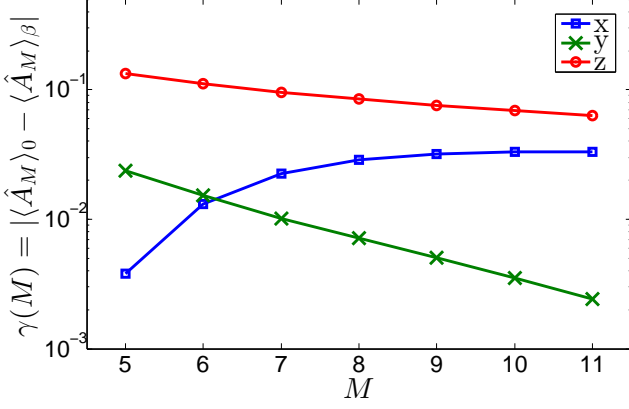
FIG. 3. (color online) Distance between expectation values of initial spin-polarized states and corresponding thermal states. For the x-polarized state (blue square), the distance does not decrease as M increases. $\hat{A}_M$ is obtained by minimizing the Frobenius norm of the commutator with Hamiltonian. Label means that direction along which the initial state is polarized.

their expectation values with respect to spin-polarized states. Figure 3 plots the distance between the expectation values of spin-polarized initial states and corresponding thermal states. The operator $\hat{A}_M$ is obtained by minimizing the Frobenius norm as we explained before. It is noteworthy that the distance does not decrease for x-polarized state. From Eq.(16) and Figs. 2 and 3, we see that thermalization time scale increases fast for x-polarized state, which is consistent with Ref.22. However, it needs some care to make direct connection to Ref. 22 since our operator is *not* translation invariant and we do not minimize the operator norm. We leave this problem for later studies.

III. SYSTEMS WITHOUT LOCAL CONSERVATION LAW

A. Floquet system

Refs. 29–31 test whether the energy conservation is important in the slow relaxation of local operators. We adopt the same Floquet operator used in Ref. 20:

$$U_F = \exp(-iH_x\tau)\exp(-iH_z\tau) , \quad (17)$$

where H_x is the σ^x part ($\sum_i g\sigma_i^x$) and H_z is the σ^z part ($\sum_i h\sigma_i^z + \sigma_i^z\sigma_{i+1}^z$) of the Hamiltonian. We choose $\tau = 0.8$. Although U_F does not conserve energy, it acts *locally* on the system. This type of Floquet operator is shown to thermalize a local operator to the infinite temperature ensemble [20, 32].

B. Method

We minimize the square of the Frobenius norm of the commutator with the Floquet operator. Up to $M = 11$,

we obtain exact results and for larger $M \le 16$ we use the MPO ansatz to find the minimum of the following

$$g(\hat{A}_M) = \frac{\text{tr}([\hat{A}_M, U_F][\hat{A}_M, U_F]^\dagger)}{\text{tr}(\hat{A}_M\hat{A}_M^\dagger)} . \quad (18)$$

Again, we define $\lambda(M) = \min\{g(\hat{A}_M)\}$ and call the corresponding $\hat{A}_M$ the slowest operator.

Let's relate $g(\hat{A}_M)$ with the thermalization time scale. As in the case of Hamiltonian system, we consider the same initial state; $\rho = I/Z + \epsilon\hat{A}_M$. Since the time step in the Floquet system is discrete, we define $a_M^{(N)} = \text{tr}(\hat{A}_M(N)\hat{A}_M)$, with $a_M^{(0)} = 1$ and $\hat{A}_M(N) = U_F^{\dagger N}\hat{A}_M U_F^N$. Using Cauchy-Schwarz inequality, we can show that

$$|a_M^{(n+1)} - a_M^{(n)}| = |\text{tr}([\hat{A}_M(n), U_F]\hat{A}_M U_F^\dagger)| \le \sqrt{g(\hat{A}_M)} . \quad (19)$$

Then, the following inequality follows:

$$|a_M^{(N)} - a_M^0| \le \sum_{n=0}^{N-1}|a_M^{(n+1)} - a_M^{(n)}| \le N\sqrt{g(\hat{A}_M)} . \quad (20)$$

Since the thermal expectation value $\hat{a}_M$ is indeed zero in the Floquet system [20], we have,

$$|a_M^{(N)} - a_M^{th}| \ge |a_M^{(0)}| - |a_M^{(N)} - a_M^{(0)}| \ge 1 - N\sqrt{g(\hat{A}_M)} . \quad (21)$$

Therefore, $g(\hat{A}_M)$ bounds the thermalization time of $\hat{A}_M$ from below by $N \ge g(\hat{A}_M)^{-1/2}$. Again, small $\lambda(M)$ implies a long thermalization time.

To extend to larger systems, we optimize the commutator over operators of a specific form. We consider operators in the space spanned by $U^n\hat{O}U^{-n}$ for $\hat{O}$ a traceless Hermitian operator acting on a single site and taking a finite number of powers of unitary operator U (this case, the Floquet U_F). Considering the operator of the following filtered form,

$$\tilde{A}_N = \sum_{n=-N}^{N} c_n U^n\hat{O}U^{-n} , \quad (22)$$

where $\tilde{A}_N$ is supported on $M = 2N + 1$ sites for the Floquet system, we find $\tilde{A}_N$ that gives $\tilde{\lambda}(M) = \min\{\text{tr}([\tilde{A}_N, U][\tilde{A}_N, U]^\dagger)/\text{tr}(\tilde{A}_N\tilde{A}_N^\dagger)\}$ by varying the coefficients $\{c_n\}$ and $\hat{O}$. This method can be considered as a version of the Lanczos method as it also works in a Krylov subspace, but here we use a tensor network method to approximate $\text{tr}(\hat{O}^\dagger U^n\hat{O}U^{-n})$. Thus $\tilde{\lambda}$ is a variational upper bound of the true minimum λ. Interestingly, this simple model with only $2N + 4$ real variables ($2N + 1$ for c_n's and 3 for $\hat{O}$) agrees very well with the available exact results ($M = 11$ or $N = 5$). Moreover, we find that the filtered $\hat{A}_N$ that gives $\tilde{\lambda}(M)$ is obtained from the same
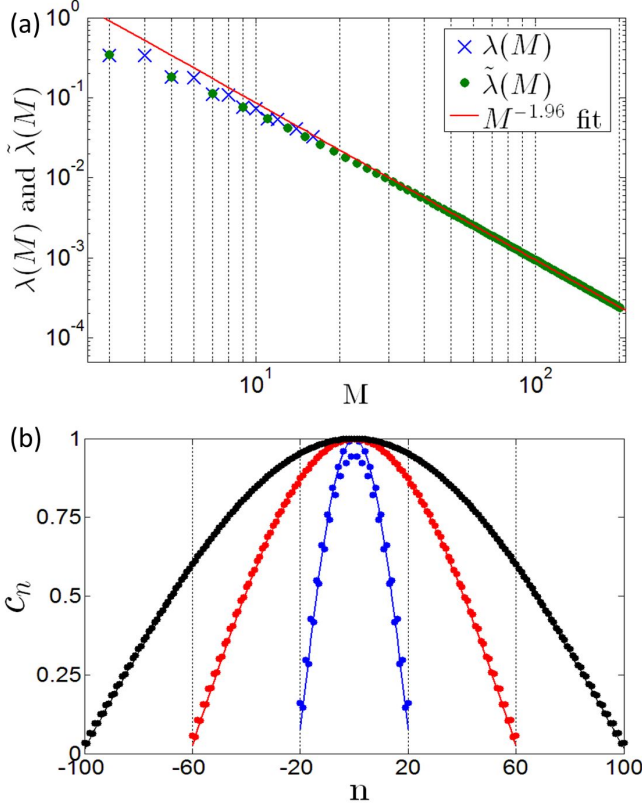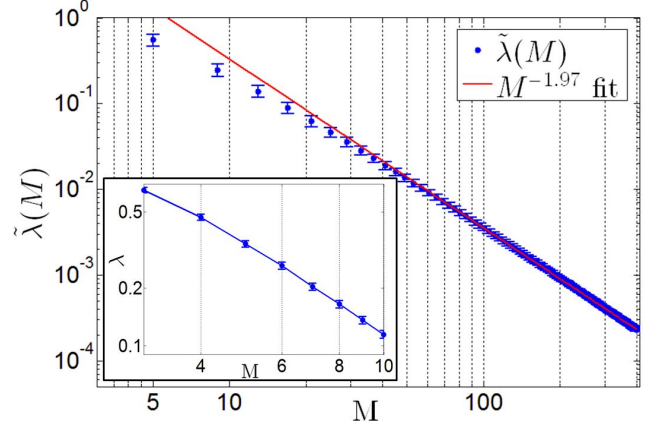
FIG. 5. (color online) $\tilde{\lambda}(M)$ for a random circuit model computed by matrix product method with the filtered of operators (Eq. (23)). $1/2\sigma^z$ is chosen for the single site Hermitian operator. We averaged 50 realizations of random circuits. Asymptotically, the data follows very closely M^{-2}. Inset: $\lambda(M)$ is the result of exhaustive search without using the filtered form for the same model. These values cannot be matched by the simple filtered operator we consider.

Comparing Figure 4 (a) with Figure 1 (a), we see that $\lambda(M)$ decreases *slower* than that in Hamiltonian cases, which indicates that the Floquet system thermalizes the local operator *faster*. Since the only apparent difference between the Hamiltonian system and the Floquet system is the existence of the energy conservation, we attribute this faster relaxation to the absence of conservation law. Nevertheless, Figure 4 clearly exhibits that the rate of relaxation in the Floquet system becomes slower as the support M increases, thus $\hat{A}_M$ is again an approximately conserved quantity.

D. Quantum Circuit

To determine whether this phenomenon is more generally true, we also study a family of quantum circuits, each composed of two rounds, where in the first round, gates act on pairs of sites $..., (1, 2), (3, 4), ...$ and on the second round gates act on pairs $..., (0, 1), (2, 3), (4, 5), ...$ We choose all gates in a given round to be the same, but choose them randomly. The results are shown in Fig. 5. We find again that even in this random case, slow operators are present.

To study larger systems, we again apply the matrix product method to the same form of the operators:

$$\tilde{A}_N = \sum_{n=-N}^{N} c_n U^n \hat{O} U^{-n} \ . \qquad (23)$$

Since our random circuit consists of two alternating non-commuting unitaries, the support of $\tilde{A}_N$ is now $4N + 1$ instead of $2N + 1$. In this random circuits, generally there is no single site Hermitian $\hat{O}$ that matches the exact

FIG. 4. (color online) (a) Decay of $\lambda(M)$ (blue cross), the exact minimum of Eq. (18) and $\tilde{\lambda}(M)$ (green dot), a variational upper bound found in the form of Eq. (23). The fractional difference between $\lambda(M)$ and $\tilde{\lambda}(M)$ is less than $1/200$ for $M \leq 11$. $\tilde{\lambda}(M)$ decreases close to M^{-2} asymptotically. Note that there is a strong parity effect in the exact result: $\lambda(2N - 1) \simeq \lambda(2N)$. (b) Structure of coefficients c_n (dots) for $N = 20$ (blue, most narrow), 60 (red, intermediate width), and 100 (black, widest). A cosine function with wavelength $4N + 2$ fits well except near $n = 0$ of $N = 20$.

single site operator $\hat{O}$ for all values of N. For the given model and the parameters, $\hat{O} = 0.04\sigma^x - 0.66\sigma^y + 0.75\sigma^z$. Note that $\tilde{A}_N$ can only be supported on odd number of sites. It turns out that the exact operator of even support can be approximated by a simple symmetric extension: $\tilde{A}_N \otimes \sigma^0 + \sigma^0 \otimes \tilde{A}_N$, where σ^0 is an identity. Therefore, $\lambda(2N - 1) \simeq \lambda(2N)$. It is noteworthy that the filtering method is unable to capture the slowest dynamics in the Hamiltonian system.

C. Results

Figure 4 (a) is the plot of $\lambda(M)$, the minimum value of Eq. 18, and $\tilde{\lambda}(M)$, which was obtained by the form of Eq. (23). $\tilde{\lambda}(M)$ asymptotically decreases very close to M^{-2} and its optimal distribution of $\{c_n\}$ is close to a cosine modulation (Figure 4 (b)), which we expect to be a generic feature of quantum circuits with local dynamics (see the Appendix).

calculation for a small system size, unlike in the Floquet operator we studied in the main text. Therefore, we just choose $\hat{O} = 1/2\sigma^z$ as a single site Hermitian operator. The weights c_n again agree very well a cosine shape and λ decays close to $1/M^2$. This hints that $1/M^2$ would be a generic feature.

Such a decay $1/M^2$ would be exact with a cosine if $\text{tr}(\hat{O}U^n\hat{O}U^{-n}) = 0$ for all $n \neq 0$. We are unable to prove this decay in general but show weaker results in appendix. Finally, we do not preclude possibilities that there might be some quantum systems without local conservation law where $\lambda(M)$ decreases faster than $1/M^2$, which would be very interesting.

IV. SUMMARY AND OUTLOOK

We have numerically constructed a series of local operators that relax slower than local energy fluctuations do. Although we can approximate these slow operators by adding nonlinear energy density operators or assuming special filtered forms, their exact origin requires further exploration. We have also performed the same analysis in similar systems without energy conservation, e.g. Floquet systems, again finding slowly relaxing operators. These operators present a new class of observables; if they can be studied experimentally, they may reveal unsuspected slow relaxation.

Our method is by no means restricted to the particular Hamiltonians or Floquet operators we studied or to non-translationally invariant operators. As explained in the Appendix, it is very easy to apply this method to find instead the slowest translationally-invariant operator. For any given any spin Hamiltonian, Floquet operator, or quantum circuit, our method should find the slowest local operator of given length M. If unknown local conserved operators in terms of spin operators exist, this procedure should be able to detect them. Therefore, this may be used to find out unknown local conserved operators or unsuspected slow dynamics (e.g. quasi many-body localization in translation invariant systems [33, 34]) if the system of interest has such things.

V. ACKNOWLEDGEMENT

Appendix A: Structure of Slowest Operators and Construction of slow operators in Hamiltonian System

The main conclusion of this work is that the minimum value of the commutator with Hamiltonian measured by Frobenius norm decreases faster than expected from hydrodynamics type arguments as we increase the range of operators M. First, let's understand why this is nontrivial.

There are exponentially many $(4^M - 1)$ linearly orthogonal operators for a given M. Thus, it may not seem surprising that there exists a sequence of operators whose commutator with the Hamiltonian show Frobenius norm with fast decreasing scaling. However, what is important is that $\lambda(M)$ is the *minimum* value of commutator with the Hamiltonian, which is not a random sequence. For instance, it is possible to construct a series of operators of range up to M (range 1, range 2, ..., range M operators), such that their commutators with the Hamiltonian scale with a larger exponent than the one we found. However, they are not the operators that minimize the commutator with Hamiltonian and thus not the slowest operator at a given range. We do not call them slow operators. What we found, instead, is that the scaling of the *slowest* operators is different from hydrodynamics, which is usually considered to be the slowest mode of a given range (wavelength).

1. Structure of slowest operator

Although our understanding of the nature of the slowest operators in a Hamiltonian system is incomplete, we can extract some useful information by analyzing the operators found. First, we look at how different the slowest operator is from the energy density modulation, which is expected to be the slowest mode from hydrodynamics. Once we find the slowest operator, we can decompose the operator in an operator basis (Eq. (2) in the main text) and study the magnitude of the different components, $c_\ell = \text{tr}(\hat{A}_M \hat{O}^\dagger_{M,\ell})/\text{tr}(\hat{O}_{M,\ell}\hat{O}^\dagger_{M,\ell})$, where normalization is $\sum_{\ell=1}^{4^M-1}(c_\ell)^2 = 1$. It turns out that the slowest operator consists mainly of linear energy density operators; $(\sigma_i^x, \sigma_i^z, \sigma_i^z\sigma_{i+1}^z)$. For example, when $M = 11$, the square sum of c_ℓ's of energy density operators is 0.87, which is remarkably large given the fact that there are only 32 such operators out of $4^{11} - 1$ basis operators. However, the relative magnitude of c_ℓ's does not exactly follow a cosine modulation, although it shows some similarity, so there exists some kind of "dressing" to energy modulation.

Next, we look at the overlap of the slowest operator with the energy density operators of higher powers. It is easier to analyze the operator in terms of two site energy

density operator, h_n:

$$h_n = \frac{1}{2}\left(g(\sigma_n^x + \sigma_{n+1}^x) + h(\sigma_n^z + \sigma_{n+1}^z)\right) + \sigma_n^z\sigma_{n+1}^z .$$
(A1)

Normalizing h_n to have a unit Frobenius norm, we compute the overlap o_n (we omit the range index M for brevity) resolved by the position n.

$$o_n^1 = |\mathrm{tr}(\hat{A}_M h_n)| ,$$
(A2)

where the superscript 1 means the overlap of the linear order of h_n. Note that $\mathrm{tr}(h_n h_{n+1}) \neq 0$ thus o_n contains some contribution from h_{n-1} and h_{n+1}. The overlap of quadratic order of energy density operator $o_n^2(x)$ (superscript 2 means the quadratic order) is obtained by,

$$o_n^2(x) = |\mathrm{tr}(\hat{A}_M h_n h_{n+x})| .$$
(A3)

We can continue this to arbitrary powers of h_n's. Since $(h_n)^2$ is not identity, the $o_n^2(x)$ terms may contain some contribution already counted in o_n^1. Despite this certain degree of double-counting, this decomposition reveals another structure in the slowest operator. Figure 6 is the plot of o_n^1 and $o_n^2(x)$ ($x = 0, 1, 2, 3$) of the slowest operator obtained by the MPO ansatz for $M = 28$. It is clear that the slowest operator takes more contribution from local operators (for quadratic order, smaller x) than non-local operators (large x). For the same order and x, their relative magnitudes are similar to cosine modulation but not exactly the same. These features suggest that the slowest operator can be built by adding nontrivial dressing to energy density modulation and its higher powers.

In the following subsection, we build a sequence of slow operators that shows a fast decaying scaling based on the above observation. In both cases, we are able to construct slow operators that are close to the slowest ones but unable to make them as close to have the same scaling.

2. Construction of Slow Operators

a. Nonlinear energy density operators

Figure 6 implies that the slowest operator contains nontrivial contributions from higher powers of the energy density operators. Therefore, we build operators consisting of nonlinear powers of energy density operators. For a given power $\alpha(= 1, 2, \ldots)$ and range M, we construct operators of the following form:

$$\hat{B}_{M,1} = \sum_n c_n h_n$$
(A4)

$$\hat{B}_{M,2} = \sum_{n \geq m} c_{n,m} P(n,m) h_n h_m$$
(A5)

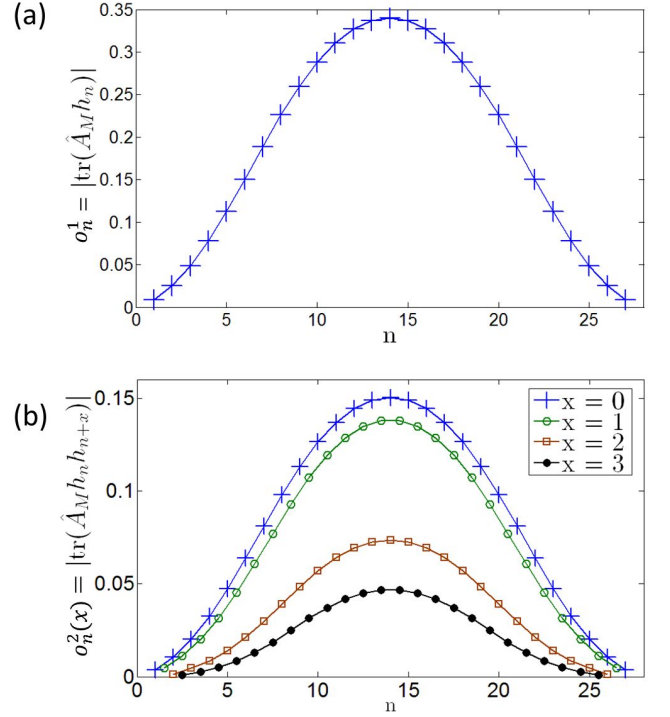$$\hat{B}_{M,3} = \sum_{n \geq m \geq l} c_{n,m,l} P(n,m,l) h_n h_m h_l ,$$
(A6)

FIG. 6. (color online) $\hat{A}_M$ is computed using MPO with bond dimension 140 and $M = 28$. (a) o_n^1 is the overlap between the slowest operator $\hat{A}_M$ and the energy density operator h_n. The shape resembles a cosine modulation (highest in the middle and the lowest at the edges) but does not exactly match it. (b) $o_n^2(x)$ is the overlap between the slowest operator and $h_n h_{n+x}$. The weight decreases as the separation x increases. For each x, the shape is similar to the other cases.

and so on. Here $P(\ldots)$ is the symmetric permutation operator that makes each term Hermitian. Here we allow h_n to be identity so that $\hat{B}_{M,\alpha}$ includes all $\hat{B}_{M,\beta}$ where $\beta \leq \alpha$. The range of each term in the summation is restricted within M. We optimize the coefficients $c_{\ldots}$ to have the minimum commutator with the Hamiltonian in the Frobenius measure. Note that $\hat{B}_{M,1}$ should just be an energy modulation as we saw in the main text.

Figure 7 is the plot of $\lambda(M)$, the minimum value of the square of the Frobenius norm of the commutator with the Hamiltonian for powers 1, 2, and 3. At linear power, $\lambda(M)$ decreases with M^{-2} as expected from conventional hydrodynamics. For the quadratic order, however, we do not see a faster decay. It has instead a smaller exponent for the range we have constructed, although it should decay at least as fast as M^{-2}, since the quadratic order operator contains the linear order operator. Starting from cubic power, we see a signature of slow operators, where $\lambda(M)$ decreases faster than M^{-2}. How these nonlinear order of energy density operators contribute to slow relaxation remains unclear but our results suggest that they have some nontrivial consequences.
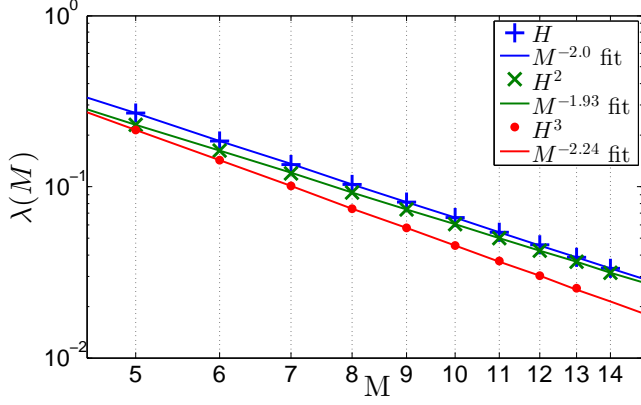
FIG. 7. (color online) $\lambda(M)$ is the minimum value of the commutator with the Hamiltonian measured by the square of Frobenius norm. Operators are constructed as Eq. (A6). For power 1, $\lambda(M) \sim M^{-2}$ as expected. For power 2, we do not make slow operators. From power 3, we start seeing a sequence of slow operators whose scaling of $\lambda(M)$ decreases faster than M^{-2}.

b. Filtered operator

Another way of constructing slow operators is the filtering of energy modulation. A conventional filtering method (see Eq.(23)) is discussed in the main text where we construct the slowest operator of the Floquet system. In the Hamiltonian case, however, the method should be modified since in principle for any $t \neq 0$, $\exp(-iHt)$ is a global operator that makes a local operator of range M act on every site. In addition, it turns out that it is not easy to build slow operators for a Hamiltonian system if we start from a single site operator. Instead, we use the cosine modulation of the energy density operator as starting operator and trace out the external sites to restrict the range of operators to be M at every step. We construct an operator of the following form:

$$\hat{A}_{M,N} = \sum_{n=-N}^{N} c_n \hat{a}_{M,n} \qquad (A7)$$

$$\hat{a}_{M,n} = \mathrm{tr}_{\backslash M}(e^{iHn\delta t} \hat{a}_{M,n-1} e^{-iHn\delta t}) , \qquad (A8)$$

where $\hat{a}_{M,0}$ is the range M cosine modulation of energy density operator, $\mathrm{tr}_{\backslash M}$ traces out the region outside M consecutive spins and δt is the small time step. First we obtain the $2N+1$ $\hat{a}_{M,n}$'s and then we compute the coefficients c_n that minimize the Frobenius norm of the commutator with the Hamiltonian. For a sufficiently small δt, we can approximate e^{iHt} with H acting on $M+2$ sites (another site at each edge) with negligible error at each step. We took $\delta t = 0.667$ and checked that results do not change by increasing the approximate range of e^{iHt}. The underlying idea of this construction is that we start from the expected slowest mode (energy modulation) and then "filter out" the fast component, if exits, at each small time evolution. We make the total evolution step number N be proportional to the range M.
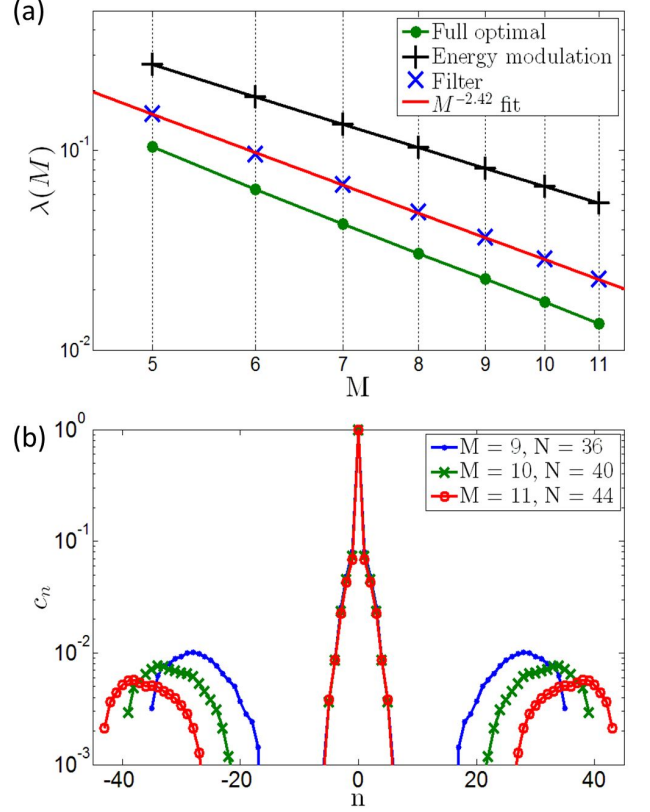

FIG. 8. (color online) Filtering operator of the Hamiltonian system. (a) $\lambda(M)$ is the minimum value of the square of the Frobenius norm of the commutator of the filtered operator (Eq. (A7)) with Hamiltonian. It decreases faster than the diffusion scaling, M^{-2}. Therefore, they are slow operators. (b) The coefficients of slowest filtered operator. It is highly peaked near $n = 0$ as expected. In addition, they get small but nonnegligible contributions from the later step operators.

Figure 8 (a) is the plot of $\lambda(M)$, which is the square of the Frobenius norm of the commutator between the slowest operator found by the filtering scheme and the Hamiltonian. It decreases faster than energy diffusion and thus we can say that these are indeed slow operators. The coefficients c_n are highly peaked around $n = 0$, which is consistent with the previous observation that the largest contribution to the slowest mode comes from the energy density modulation. Nevertheless, we are unable to get the exact slowest operator by varying δn, N, and $\hat{a}_{M,0}$.

Appendix B: Results of another set of parameters and different model Hamiltonians

1. Another set of parameters

In this section, we show that the results in the main text do not depend on the parameter choice. We choose another set of parameters $(g, h) = (-1.05, 0.5)$, which is
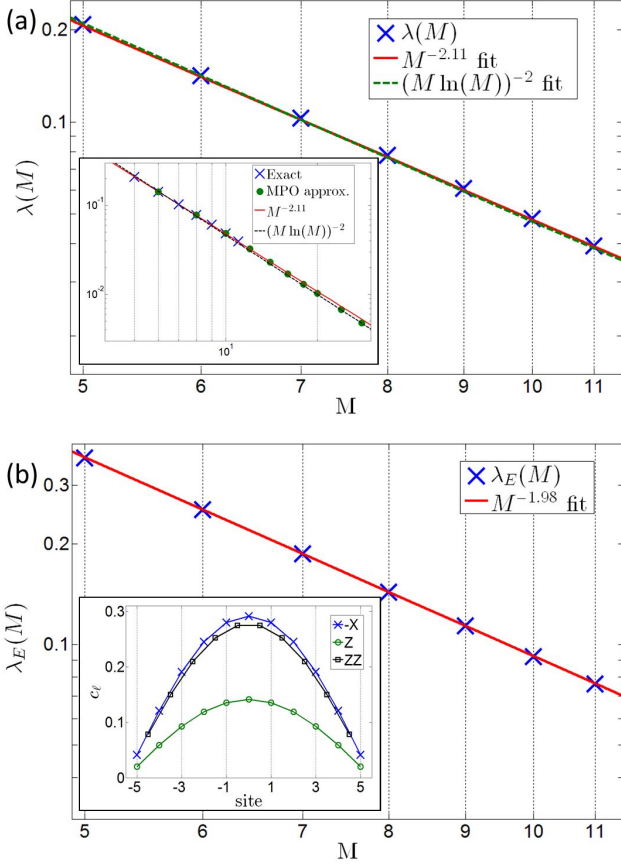
FIG. 9. (color online) (a) $\lambda(M)$ is the minimum value of the Frobenius norm of the commutator with Hamiltonian. Here we choose the parameters to be $(g,h) = (-1.05, 0.5)$. λ still decreases faster than $1/M^2$. Inset: $\lambda(M)$ vs. M using tensor network method (MPO approximation). The results are indistinguishable from available exact results. (b) $\lambda_E(M)$ is the same as when we restrict the search space within the terms in the Hamiltonian. As expected, it has M^{-2} scaling and the structure is cosine modulation (inset).

the parameter choice of Ref. 22.

Figure 9 (a) is the minimum value of Eq.(3) in the main text with the other parameter choice. We can see that it decays faster than $1/M^2$. As is the case of the parameter choice in the main text, the data can be well-fitted by two methods; a power-law and a logarithmic correction to $1/M^2$. Since the power-law exponent could depend on the parameter choice, we do not attempt to draw a strong conclusion from this data except that $\lambda(M)$ decreases faster than $1/M^2$, the scaling of the diffusive energy mode.

Figure 9 (b) is the plot of the minimum value of Eq.(3) in the main text when only terms in Hamiltonian are allowed for two sets of parameters. Unlike the case where all operators are used, the decay scaling remains the same as $1/M^2$ as expected from the hydrodynamics. Therefore, we again explicitly demonstrate that the longest wavelength energy modulation is the slowest mode of a conserved quantity.
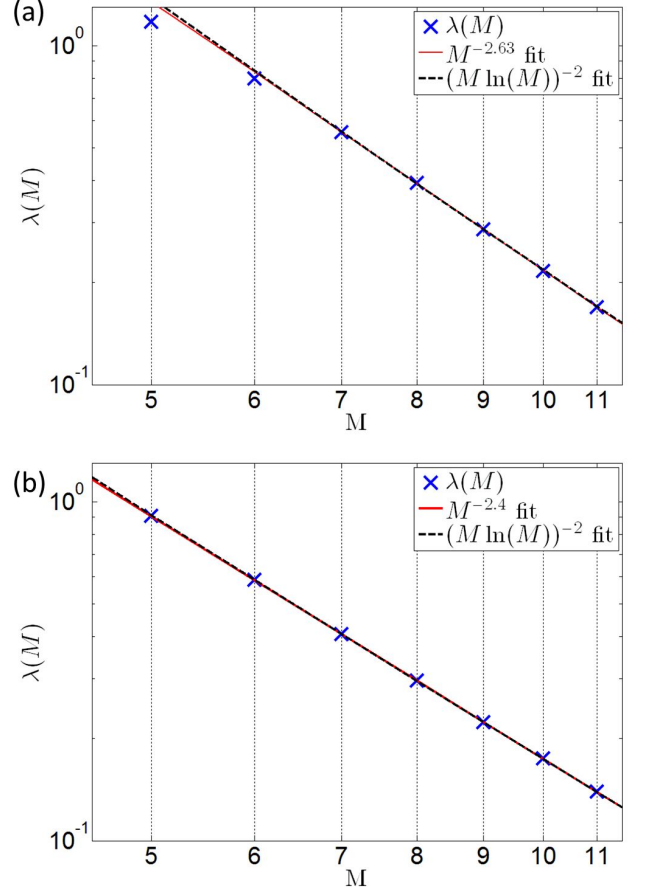

FIG. 10. (color online) $\lambda(M)$ is the minimum value of the Frobenius norm of the commutator with Hamiltonian, Eq. (B1). (a) $(J_z, g, h) = (0.5, 0.8, 0.4)$. (b) $(J_z, g, h) = (0, 0.8, 0.4)$. Both cases, $\lambda(M)$ decreases faster than M^{-2} and thus there exist nontrivial slow operators.

2. Different model Hamiltonians

In this section, we consider another nonintegrable Hamiltonian and show that there exist slow operators with nontrivial scaling of $\lambda(M)$.

We consider the XXZ model with fields along x and z directions.

$$H = \sum_i \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + J_z \sigma_i^z \sigma_{i+1}^z + g\sigma_i^x + h\sigma_i^z .$$
(B1)

This model is nonintegrable when at least two of three parameters (J_z, g, h) are nonzero. We choose two cases: $(J_z, g, h) = (0.5, 0.8, 0.4)$ and $(J_z, g, h) = (0, 0.8, 0.4)$. These choices have fewer discrete symmetries than the other two possibilities; when only $g = 0$, this model conserves total spin along z direction and when only $h = 0$, this model is symmetric under all spin flip (Ising symmetry).

Figure 10 is the plot of $\lambda(M)$ vs. M, where $\lambda(M)$ is the smallest value of the Frobenius norm of the commutator with the Hamiltonian for an operator of range M. We

can clearly see that in both cases $\lambda(M)$ decreases faster than diffusive scaling, M^{-2}. Therefore, there exists some slow operator that relaxes slower than diffusion in this nonintegrable model. We have obtained similar results for the other two possibilities of parameter choices and there we found larger exponent, which may result from existence of other symmetries.

Appendix C: Results of Matrix Product Operators with various bond dimensions

In this section, we present more results of the matrix product ansatz with various bond dimensions for the Hamiltonian system. For a fixed bond dimension D, we compute the matrix product operator (MPO) that minimizes $\lambda(M)$. This gives another upper bound to the true minimum value of given support M. Fig. 11 shows the results of $D = 4, 20, 80$, and 140. Since the energy modulation can be expressed by a matrix product operator with bond dimension 3, operators of bond dimension 4 contain the energy modulation. This explains what we see in the Fig. 11 where $\lambda(M) \sim M^{-2}$ (same behavior as the energy diffusion) at $D = 4$. For growing bond dimensions, we see increasingly larger deviations from the diffusion scaling (M^{-2}). This supports our conclusion in the main text that there exists nontrivial slowly relaxing operators in the nonintegralbe Hamiltonian system.

Appendix D: Translation invariant operators

In the main text, we have studied non-translationally invariant operators in order to compare with the visualized sinusoidal energy modulation. However, the general procedure of minimizing the square of the Frobenius norm of the commutator with Hamiltonian is not restricted to non-translationally invariant system. In addition, translation invariant system is more directly connected to Ref. 22.

First, we define the length M translation invariant operator $\hat{B}_M$ as following:

$$\hat{B}_M = \sum_i \hat{A}_{M,i} , \tag{D1}$$

where $\hat{A}_{M,i}$ is the length M traceless Hermitian operator supported on sites from i to $i + M - 1$. Now, we just search for an optimal operator $\hat{B}_M$ that minimizes the Frobenius norm of the commutator with Hamiltonian as Eq. (3). What changes here is that once we expand $\hat{A}_{M,i}$ in terms of basis operators as Eq.(2), the denominator in Eq. (3) becomes different since basis operators from different sites may no longer be orthogonal. Therefore, finding the minimum becomes a generalized eigenvalue problem.

Figure 12 is the results of minimum $\lambda(M)$ for the translation invariant operators. Again, we see that the decay
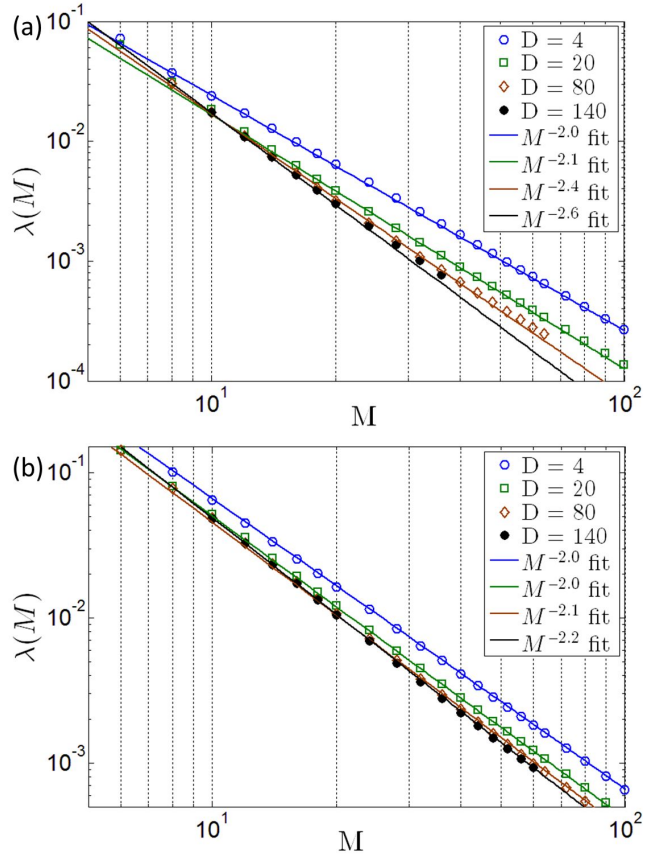

FIG. 11. (color online) (a) $\lambda(M)$ computed by the matrix product ansatz at fixed bond dimension D. Straight lines are power-law fittings to the $\lambda(M)$'s computed. At $D = 80$, we already see significant deviation from the diffusion scaling. This is a strong evidence that we have operators relaxing slower than the hydrodynamic mode. (b) Same calculation with a different parameter set; $(g, h) = (-1.05, 0.5)$. Qualitative features are the same.

rate of $\lambda(M)$ is faster than M^{-2} although it is not simply connected to M^{-2} decay of energy modulation.

Appendix E: Existence of Slow Operators for Arbitrary Quantum Circuits

We now show that a slowly relaxing operator must always exist, on an interval of length M with the relaxation rate going to zero as M gets large. Let U be the unitary of the quantum circuit restricted to an interval of length slightly larger than M (in this way, we can consider only finite dimensional spaces).

Let $\mathcal{E}(\hat{O})$ be a super-operator defined by $\mathcal{E}(\hat{O}) = U\hat{O}U^{\dagger}$. The space of operators can be regarded as a vector space, with inner product $(A, B) = \text{tr}(A^{\dagger}B)$, and with $\mathcal{E}(\hat{O})$ being a linear operator on this space. The operator $\mathcal{E}$ is non-Hermitian but it is a normal operator, since its Hermitian conjugate is equal to $\mathcal{E}^{\dagger}(\hat{O}) = U^{\dagger}\hat{O}U$ and $\mathcal{E}(\mathcal{E}^{\dagger}(\hat{O})) = \mathcal{E}^{\dagger}(\mathcal{E}(\hat{O})) = \hat{O}$ and so $[\mathcal{E}, \mathcal{E}^{\dagger}] = 0$. Let
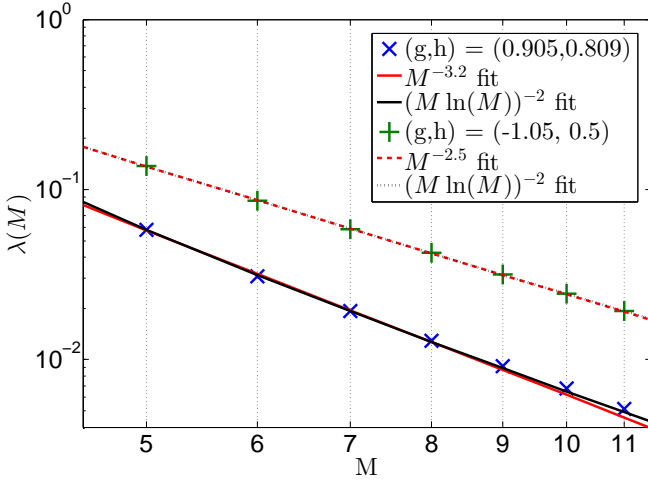
FIG. 12. (color online) ($\lambda(M)$ of translation invariant operators for two sets of parameter choices. Both cases, $\lambda(M)$ decreases faster than M^{-2}.

$\mathcal{E}_h = \frac{1}{2}(\mathcal{E} + \mathcal{E}^\dagger)$ and $\mathcal{E}_a = \frac{1}{2i}(\mathcal{E} - \mathcal{E}^\dagger)$.

We begin by constructing an operator $\hat{O}$ supported on an interval of length M such that $\mathcal{E}_h(\hat{O}) = x\hat{O} + \hat{\epsilon}$ for some x and for $\hat{\epsilon} = \mathcal{O}(1/M)$. Let A be any traceless operator supported on a single site in the center of the interval. We consider the Krylov space generated by the vectors $A, \mathcal{E}_h(A), \mathcal{E}_h^2(A), \dots$. Let K be the number of vectors we take; K will be proportional to M and will be chosen such that all these operators are supported on the interval of length M. Since $\mathcal{E}_h$ is Hermitian, using the Lanczos procedure we can write it as a tridiagonal matrix T in this Krylov space. We claim that there exists a vector v supported on the first $K-1$ vectors $w_1, \dots, w_{K-1}$ such that $|Tv - xv|^2/|v|^2 \leq \mathcal{O}(1/K^2)$. To verify this claim, let ψ be any eigenvector of T with at least half its weight on the first $K/2$ vectors with $\psi = \sum_k a_k v_k$; let x be the corresponding eigenvalue; let $v = \sum_k a_k(1 - k/K)v_k$. The basis in which the matrix is tridiagonal has basis vectors $w_1, w_2, \dots$, with w_k being in the span of the first k vectors $A, \mathcal{E}_h(A), \dots$. Hence, the vector v gives us the desired operator $\hat{O}$.

Now consider the operators

$$hatO_\pm = (\sqrt{1 - x^2} \pm \mathcal{E}_a)\hat{O}, \qquad (E1)$$

with $\hat{O}$ normalized such that $||\hat{O}|| = 1$. Note that $\mathcal{E}(\hat{O}_\pm) = \mathcal{E}_h(\hat{O}_\pm) + i\mathcal{E}_a(\hat{O}_\pm) = x\hat{O}_\pm + i(\sqrt{1 - x^2}\mathcal{E}_a \pm \mathcal{E}_a^2)\hat{O} + \mathcal{O}(1/M)$. Using the fact that $\mathcal{E}_h^2 + \mathcal{E}_a^2$ is equal to the identity super-operator, $(\sqrt{1 - x^2}\mathcal{E}_a \pm \mathcal{E}_a^2)\hat{O} = \sqrt{1 - x^2}\mathcal{E}_a \pm (1 - x^2))\hat{O} + \mathcal{O}(1/M) = \pm\sqrt{1 - x^2}\hat{O}_\pm + \mathcal{O}(1/M)$. Hence,

$$\mathcal{E}(\hat{O}_\pm) = z\hat{O}_\pm + \mathcal{O}(1/M) \qquad (E2)$$

with

$$z = x \pm i\sqrt{1 - x^2}, \qquad (E3)$$

so that $|z| = 1$. At least one of the two operators $\hat{O}_\pm$ must have non-negligible norm. Let X be the corresponding operator, normalized to have norm 1. Hence, we have constructed an operator X supported on the interval of length M such that $\mathcal{E}(X) = zX + \mathcal{O}(1/M)$.

This already implies that there is some operator X which is slowly relaxing but perhaps oscillating; i.e., since X is an approximate eigenoperator of $\mathcal{E}$, if $z = 1$ then X changes slowly over time, while if $z \neq 1$, then the expectation value of X oscillates.

In fact, we can always construct an operator Y which is an approximate eigenoperator of $\mathcal{E}$. Here is one way. Consider many disjoint intervals of length M. Let $X_1, X_2, \dots$ be the operators on these intervals with eigenvalues $z_1, z_2, \dots$ Choose some subset S of these intervals such that the product of the z_i on that subset is close to 1: $\prod_{i \in S} z_i \approx 1$. Then, let $Y = \prod_{i \in S} X_i$. This requires some analytic estimates to determine the support of Y required: since Y is a product of many operators, the error (in that each X_i is only an approximate eigenoperator) may add, so the support of Y may scale as a fairly large polynomial in the error. We leave this estimate for later.

Appendix F: $1/M^2$ scaling of filtered operators

We show $1/M^2$ scaling of filtered operators of the form Eq. (23) if $\text{tr}(\hat{O}U^n\hat{O}U^{-n}) = 0$ for all $n \neq 0$, where $\hat{O}$ is a traceless Hermitian acting on a single site. First, observe that

$$\frac{\text{tr}([\tilde{A}_N, U][\tilde{A}_N, U])}{\text{tr}(\tilde{A}_N\tilde{A}_N^\dagger)}$$
$$= 2 - \Bigg(\sum_{n,m=-N}^{N} c_n c_m \text{tr}(U^{n-m-1}\hat{O}U^{-n+m+1}\hat{O}$$
$$+ U^{n-m+1}\hat{O}U^{-n+m-1}\hat{O})\Bigg)/\text{tr}(\tilde{A}_N\tilde{A}_N^\dagger) .$$
$$(F1)$$

Therefore, if $\text{tr}(\hat{O}U^n\hat{O}U^{-n}) = \delta_{n,0}\text{tr}(\hat{O}\hat{O}^\dagger)$, the above expression simplifies to

$$\frac{\text{tr}([\tilde{A}_N, U][\tilde{A}_N, U])}{\text{tr}(\tilde{A}_N\tilde{A}_N^\dagger)} = 2\left(1 - \frac{\sum_{n=-N}^{N-1} c_n c_{n+1}}{\sum_{n=-N}^{N} c_n^2}\right) . \quad (F2)$$

This is a trivial quadratic optimization problem and the solution is $c_n = \cos(n\pi/(2N + 2))$. Therefore, the minimum value $\tilde{\lambda}$ for sufficiently large N is

$$\tilde{\lambda} = \min\left[\frac{\text{tr}([\tilde{A}_N, U][\tilde{A}_N, U])}{\text{tr}(\tilde{A}_N\tilde{A}_N^\dagger)}\right]$$
$$= 2 - 2\cos\left(\frac{\pi}{2N + 2}\right) \simeq \frac{\pi^2}{4(N + 1)^2} \sim \frac{1}{M^2} , \quad (F3)$$

where we used the fact that the support $M = 4N + 1$.

Ref. 20 has shown that the U_F (the Floquet operator) thermalizes a local operator at infinite temperature, and thus $\text{tr}(\hat{O}U^n\hat{O}U^{-n}) = 0$ for a sufficiently large n. For the Floquet system, therefore, the above condition is approximately satisfied for a large enough N. Figure 4 in the main text shows that for $N = 60$ and 100, the form is very close to the cosine and the scaling for large $M = 2N + 1$ closely follows M^{-2} scaling. For a general random circuit, Figure 5 again shows M^{-2} scaling. The structure of $\{c_n\}$ is found to follow cosine modulation similar to the Floquet case.

One example of quantum circuits that satisfies the condition that $\text{tr}(\hat{O}U^n\hat{O}U^{-n}) = 0$ for all $n \neq 0$ is the swap operator U_{sw}.

$$U_{sw} = \prod_n V_{2n,2n+1} \prod_m V_{2m-1,2m} , \qquad \text{(F4)}$$

where $V_{x,y}$ swaps the spins $\mathbf{S}_x$ and $\mathbf{S}_y$; $V_{x,y}|\mathbf{S}_x, \mathbf{S}_y\rangle =$

$|\mathbf{S}_y, \mathbf{S}_x\rangle$. Then, for an $\hat{O}$ acting on site 0, U_{sw} moves $\hat{O}$ to the site $-2n$ and U_{sw}^{-n} moves $\hat{O}$ to the site $2n$ and thus the condition is satisfied. In this case, we can write every step analytically and prove $1/M^2$ scaling and the cosine modulation. In a generic quantum circuit, however, these features are seen only for sufficiently large M. Furthermore, in a translation invariant system, U_{sw} has an exact local conservation law. For instance, any translation invariant single site operator commutes with U_{sw} and is thus conserved.

In the main text, we have connected $\lambda(M)$ to the thermalization time scale by $1/\tau \sim \sqrt{\lambda(M)}$ (Eq.(21)). At first glance, $1/\tau \sim \sqrt{\lambda(M)} \sim 1/M$ may seem trivial for an operator of support M by Lieb-Robinson bound type argument, but an important point here is that we computed the relaxation of an operator by the Frobenius norm, not the operator norm by which the Lieb-Robinson bound has been computed [27].

[1] J. M. Deutsch, "Quantum statistical mechanics in a closed system," Phys. Rev. A **43**, 2046–2049 (1991).

[2] M Srednicki, "Chaos and quantum thermalization," Phys. Rev. E **50**, 888–901 (1994).

[3] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii, "Thermalization and its mechanism for generic isolated quantum systems," Nature **452**, 854–858 (2008).

[4] Noah Linden, Sandu Popescu, Anthony J. Short, and Andreas Winter, "Quantum mechanical evolution towards thermal equilibrium," Phys. Rev. E **79**, 061103 (2009).

[5] Marcos Rigol, Vanja Dunjko, Vladimir Yurovsky, and Maxim Olshanii, "Relaxation in a completely integrable many-body quantum system: An ab initio study of the dynamics of the highly excited states of 1d lattice hardcore bosons," Phys. Rev. Lett. **98**, 050405 (2007).

[6] Pasquale Calabrese, Fabian H. L. Essler, and Maurizio Fagotti, "Quantum quench in the transverse-field ising chain," Phys. Rev. Lett. **106**, 227203 (2011).

[7] Christian Gogolin, Markus P. Müller, and Jens Eisert, "Absence of thermalization in nonintegrable systems," Phys. Rev. Lett. **106**, 040401 (2011).

[8] Maurizio Fagotti, Mario Collura, Fabian H. L. Essler, and Pasquale Calabrese, "Relaxation after quantum quenches in the spin-$\frac{1}{2}$ heisenberg xxz chain," Phys. Rev. B **89**, 125101 (2014).

[9] B. Wouters, J. De Nardis, M. Brockmann, D. Fioretto, M. Rigol, and J.-S. Caux, "Quenching the anisotropic heisenberg chain: Exact solution and generalized gibbs ensemble predictions," Phys. Rev. Lett. **113**, 117202 (2014).

[10] B. Pozsgay, M. Mestyán, M. A. Werner, M. Kormos, G. Zaránd, and G. Takács, "Correlations after quantum quenches in the xxz spin chain: Failure of the generalized gibbs ensemble," Phys. Rev. Lett. **113**, 117203 (2014).

[11] G. Goldstein and N Andrei, "Failure of the GGE hypothesis for integrable models with bound states," ArXiv e-prints (2014), arXiv:1405.4224v2 [cond-mat.quantumgas].

[12] M. Mestyan, B. Pozsgay, G. Takacs, and M.A. Werner, "Quenching the XXZ spin chain: quench action approach versus generalized Gibbs ensemble," ArXiv e-prints (2014), arXiv:1412.4787v1 [cond-mat.stat-mech].

[13] Anatoli Polkovnikov, Krishnendu Sengupta, Alessandro Silva, and Mukund Vengalattore, "Colloquium," Rev. Mod. Phys. **83**, 863–883 (2011).

[14] V.I. Yukalov, "Equilibration and thermalization in finite quantum systems," Laser Phys. Lett. **8**, 485 (2011).

[15] Lea F. Santos and Marcos Rigol, "Localization and the effects of symmetries in the thermalization properties of one-dimensional quantum systems," Phys. Rev. E **82**, 031130 (2010).

[16] M. Rigol and M. Srednicki, "Alternatives to eigenstate thermalization," Phys. Rev. Lett. **108**, 110601 (2012).

[17] M Kruczenski and S. Khlebnikov, "Thermalization of isolated quantum systems," ArXiv e-prints (2013), arXiv:1312.4612v2.

[18] W. Beugeling, R. Moessner, and Masudul Haque, "Finite-size scaling of eigenstate thermalization," Phys. Rev. E **89**, 042112 (2014).

[19] S. Sorg, L. Vidmar, L. Pollet, and F. Heidrich-Meisner, "Relaxation and thermalization in the one-dimensional bose-hubbard model: A case study for the interaction quantum quench from the atomic limit," Phys. Rev. A **90**, 033606 (2014).

[20] Hyungwon Kim, Tatsuhiko N. Ikeda, and David A. Huse, "Testing whether all eigenstates obey the eigenstate thermalization hypothesis," Phys. Rev. E **90**, 052105 (2014).

[21] G. Goldstein and N Andrei, "Stron eigenstate thermalization hypothesis," ArXiv e-prints (2014), arXiv:1408.3589v1 [cond-mat.stat-mech].

[22] M. C. Bañuls, J. I. Cirac, and M. B. Hastings, "Strong and weak thermalization of infinite nonintegrable quantum systems," Phys. Rev. Lett. **106**, 050405 (2011).

[23] F. G. S. L. Brandao, A. Harrow, and M. Horodecki, "Local random quantum circuits are approximate polynomial-designs," ArXiv e-prints (2012), arXiv:1208.0692v1 [quant-ph].

[24] Hyungwon Kim and David A. Huse, "Ballistic spreading of entanglement in a diffusive nonintegrable system," Phys. Rev. Lett. **111**, 127205 (2013).

[25] We have also performed similar analysis in the infinite chain for translationally invariant operators where the action of the local operator is the same on all sites. Most of our main results (Figures 1 (a), 4 (a)) remain true. Since it is easier to visualize local operators and to directly compare with diffusive energy mode for translationally non-invariant operators, we present the results on an infinite chain where the local operator is placed on M specific consecutive sites. See the Appendix.

[26] B Pirvu, V Murg, J I Cirac, and F Verstraete, "Matrix product operator representations," New Journal of Physics **12**, 025012 (2010).

[27] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).

[28] S. Bravyi, M. B. Hastings, and F. Verstraete, "Lieb-robinson bounds and the generation of correlations and topological quantum order," Phys. Rev. Lett. **97**, 050401 (2006).

[29] Luca D'Alessio and Marcos Rigol, "Long-time behavior of isolated periodically driven interacting lattice systems," Phys. Rev. X **4**, 041048 (2014).

[30] Achilleas Lazarides, Arnab Das, and Roderich Moessner, "Equilibrium states of generic quantum systems subject to periodic driving," Phys. Rev. E **90**, 012110 (2014).

[31] Pedro Ponte, Anushya Chandran, Z. Papić, and Dmitry A. Abanin, "Periodically driven ergodic and many-body localized quantum systems," Annals of Physics **353**, 196 (2015).

[32] T. Prosen, "Ruelle resonances in quantum many-body dynamics," J. Phys. A:Math. Gen **35**, L737 (2002).

[33] Wojciech De Roeck and F. Huveneers, "Scenario for delocalization in translation-invariant systems," Phys. Rev. B **90**, 165137 (2014).

[34] N.Y. Yao, C. R. Laumann, J. I. Cirac, M.D. Lukin, and J.E. Moore, "Quasi Many-body Localization in Translation Invariant Systems," ArXiv e-prints (2014), arXiv:1410.7407v1 [cond-mat.dis-nn].