Prethermalization and Thermalization in Generic Isolated Quantum Systems

Krishnanand Mallayya,1,2 Marcos Rigol,1,2 and Wojciech De Roeck1,2

1Department of Physics, Pennsylvania State University, University Park, Pennsylvania, USA 16802
2Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA
(Received: October 31, 2018)

Prethermalization has been extensively studied in systems close to integrability. We propose a more general, yet conceptually simpler, setup for this phenomenon. We consider a possibly nonintegrable—reference dynamics, weakly perturbed so that the perturbation breaks at least one conservation law of the reference dynamics. We argue then that the evolution of the system proceeds via intermediate (generalized) equilibrium states of the reference dynamics. The motion on the manifold of equilibrium states is governed by an autonomous equation, flowing towards global equilibrium in a time of order \( g^{-2} \), where \( g \) is the perturbation strength. We also describe the leading correction to the time-dependent reference equilibrium state, which is, in general, of order \( g \). The theory is well-confirmed in numerical calculations of model Hamiltonians, for which we use a numerical linked cluster expansion and full exact diagonalization.

I. INTRODUCTION

Prethermalization [1] has emerged over the last decade as an interesting and ubiquitous phenomenon in the dynamics of ultracold quantum gases in one-dimensional geometries [2–6]. In general, it refers to a separation of timescales: Some systems far from equilibrium quickly relax to long-lived (non-)thermal states (not true thermal equilibrium states) on short timescales, before eventually relaxing to the expected true thermal equilibrium states on much longer timescales.

There are some general instances of this phenomenon that are well understood, analytically and numerically. A first example are quenches in isolated noninteracting (integrable) systems, in which interactions (integrability-breaking perturbations) of strength \( g \) are turned on [7–13]. By dephasing, observables quickly settle to quasi-steady states that are well described by generalized Gibbs ensembles (GGEs) [14–18] of the non-interacting systems. The observables then relax to the thermal equilibrium values (thermalize) in a much longer time scale \( \propto g^{-2} \), as predicted by kinetic Boltzmann-like equations. The latter can be derived applying time-dependent perturbation theory to the GGEs [19, 20], and, physically, describe the effects of collisions between (quasi-)particles. It was recently shown numerically that, weakly breaking integrability in strongly interacting integrable systems also results in thermalization rates \( \propto g^{-2} \) [21]. A second example are isolated weakly interacting driven systems, for which GGEs and kinetic Boltzmann-like equations are also relevant [22, 23]. A third example are periodically driven systems at high frequency. In general, they quickly reach a time-periodic state that can be identified as a Gibbs state corresponding to an effective Hamiltonian, before relaxing to the thermal (infinite-temperature [24, 25]) state in a time scale that is exponentially long in the frequency of the drive [26–29].

In this work, we argue that the first two examples above are instances of a more universal phenomenon, a phenomenon that occurs whenever an equilibrating dynamics (to a thermal- or GGE-like state) is weakly perturbed so that, at least, one of the conserved quantities in the original dynamics is no longer conserved in the weakly perturbed one. A similar point of view was put forward in Refs. [30, 31] for open quantum systems. We first present this conclusion in a loose manner in Sec. II, and then, in Secs. III and IV, we develop a systematic treatment of weakly perturbed systems in which a conservation law is broken. The remainder of the manuscript is dedicated to demonstrating, in the context of numerical experiments, the validity and accuracy of this systematic treatment. Section V, in which we introduce the models, quenches, observables, etc, is the preamble to the numerical experiments, which are reported in Sec. VI. A summary of our results is presented in Sec. VII, while the Appendices report details of our analytical and numerical calculations.

II. PREAMBLE

Setup – We consider isolated extended quantum systems that are translationally invariant. For the sake of concreteness, one can think of chains with \( L \gg 1 \) sites. We have a pair of Hamiltonians \( \hat{H}_0 \) and \( \hat{H} \), in which the latter is a small perturbation of the former

\[
\hat{H} = \hat{H}_0 + g\hat{V},
\]

where \( g \) small. We assume that the (reference) Hamiltonian \( \hat{H}_0 \) has a conservation law, say \( \hat{Q} \), which is not shared by \( \hat{H} \), namely, \( [\hat{H}_0, \hat{Q}] = 0 \) but \( [\hat{V}, \hat{Q}] \neq 0 \). We are interested in cases in which \( \hat{H}_0, \hat{Q}, \hat{H}, \) and \( \hat{V} \) are extensive operators. In order to take the thermodynamic limit, it is helpful to deal with the intensive counterparts of those operators: \( \hat{h}_0 = \hat{H}_0/L, \; \hat{q} = \hat{Q}/L, \; \hat{h} = \hat{H}, \) etc. For simplicity, we restrict our analysis to cases in which \( \hat{Q} \) has a discrete spectrum; \( \hat{Q} \) can be the total particle number, the total magnetization, etc.
Assumption of fast equilibration – We consider cases in which the dynamics generated by $H_0$, even when the system is far from equilibrium, results in fast equilibration. Namely, we assume that, for experimentally relevant translational-invariant initial states $\hat{\rho}_i$, observables converge within a time $\tau^*$ to the predictions of an ensemble (e.g., microcanonical) of statistical mechanics $\hat{\rho}_{e_o,q}$, with $(e_0, q) = \left( (\hat{H}_0)_{\hat{\rho}_i}, (\hat{q})_{\hat{\rho}_i} \right)$. The restriction “experimentally relevant” is put to avoid cases in which the initial state is a macroscopic superposition of states at different densities $(e_0, q)$ (see Ref. [32] for a discussion of this issue in the context of quantum quenches). The restriction to translation-invariant states is used to avoid having leading effects in the equilibration dynamics that are $L$-dependent, e.g., particle or energy transport, which would complicate the picture.

Initial prethermalization – Even though there is no standard perturbation theory around a genuinely interacting $H_0$, it seems reasonable to assume that the effect of the perturbation is not felt on times $\tau \ll 1/g$, and that, for such short times, one can meaningfully approximate the dynamics under $\hat{H}$ by the reference dynamics generated by $\hat{H}_0$. Therefore, by the assumption of fast equilibration, one should observe a fast initial relaxation of observables toward the predictions of $\hat{\rho}_{e_o,q}$. The latter state can be very different to the thermal equilibrium ensemble $\hat{\rho}_e$ associated to $\hat{H}$. $\hat{\rho}_{e_o,q}$ differs from $\hat{\rho}_e$ by the fact that $q$ is an additional constraint [of course, it also differs because $e_0$ is the energy density of $H_0$ and not that of $H$, but that change is only $\mathcal{O}(g)$].

The thermalization rate – To determine how fast the true equilibrium is approached, let us start from $\hat{\rho}_{e_o,q}$ and look at the change in $\langle \hat{Q}(\tau) \rangle_{e_o,q}$. We do this perturbatively:

$$\langle \hat{Q}(\tau) \rangle_{e_o,q} = \langle \hat{Q} \rangle_{e_o,q} + \mathcal{O}(g^2 \tau^* L)$$

where we used that $[\hat{H}_0, \hat{Q}] = [\hat{H}_0, \hat{\rho}_{e_o,q}] = 0$. The precise error estimate $\mathcal{O}(-)$ will only be argued for later. The important point to be highlighted from Eq. (2) is that the leading order correction to $\langle \hat{Q} \rangle_{e_o,q}$ vanishes because $\langle [\hat{V}, \hat{Q}] \rangle_{e_o,q} = 0$, by the cyclic property of the trace and the fact that $[\hat{Q}, \hat{\rho}_{e_o,q}] = 0$. This suggests that the thermalization rate is $\propto g^2$. Indeed, carrying out the expansion one order further, we recognize Fermi’s golden rule. Finally, if we had replaced $\hat{Q}$ by a general observable $\hat{O}$, the first order term does not vanish. It results in a universal deviation of $\langle \hat{O} \rangle$ in the instantaneous state from that of $\hat{\rho}_{e_o,q}$.

III. SLOW DYNAMICS OF APPROXIMATELY CONSERVED QUANTITIES

In this section, we present a derivation of an approximate autonomous equation governing the dynamics of the densities $(e_0, q)$. This is Eq. (21), or, at a more abstract level, Eq. (21). Our derivation is not mathematically rigorous, it uses physical assumptions, and goes substantially beyond the heuristics presented above. Its validity and accuracy are confirmed by numerical calculations in Sec. VI.

A. Slow variables

We identify the densities $(e_0, q)$ as slow variables. To set up a controlled derivation, we need a projection map from states $\rho$ to $(e_0, q)$. It turns out, however, that it is more natural to start with a projection $\mathcal{P}$ from $\hat{\rho}$ to a probability distribution $p$ on $(e_0, q)$. Indeed, each such a $p$ can be lifted to a $\hat{\rho}$ by

$$\hat{\rho}_p = \int d e_0 dq p(e_0, q) \hat{\rho}_{e_o,q},$$

and the physically most natural case is, of course, when $p(e_0, q)$ is a Dirac delta distribution $\delta(e_0 - e_0^0) \delta(q - q^*)$, cf., the discussion on ‘experimentally relevant states’ in Section II. To construct the map $\mathcal{P}$, a natural approach is to measure $\hat{H}_0$ and $\hat{Q}$:

$$p_p(e_0, q) \propto \text{Tr} \left[ \hat{\rho} (\hat{H}_0 \approx e_0 L) \mathcal{P}(\hat{Q} \approx q L) \right],$$

where $P(\hat{H}_0 \approx e_0 L)$ is a spectral projection of $\hat{H}_0$ on the interval $[e_0 L - \delta_E, e_0 L + \delta_E]$ with a resolution $\delta_E$ that is much larger than the level spacing, but smaller than any relevant energy scale (like, e.g., the energy per site). For $P(\hat{Q} \approx q L)$, we can simply take the projection of $\hat{Q}$ on the eigenvalue closest to $q L$. Then we can set

$$\mathcal{P} \hat{\rho} = \int d e_0 dq p_p(e_0, q) \hat{\rho}_{e_o,q},$$

where “$\int d e_0 dq$” represents a sum with the aforementioned resolution.

With translationally invariant quantum many-body systems in mind, we use a more drastic definition of $\mathcal{P} \hat{\rho}$, which relies on insights from eigenstate thermalization for nonintegrable (quantum chaotic) systems [32–34], and generalized eigenstate thermalization for integrable systems [15, 18, 35–37]. We do not distinguish between nonintegrable systems, for which the number of conserved quantities in the thermodynamic limit is $\mathcal{O}(1)$, from integrable systems, for which the number of conserved quantities in the thermodynamic limit is infinite. In translationally invariant (nonintegrable or integrable) quantum many-body systems, even a single eigenstate of $\hat{H}_0$ and $\hat{Q}$ is a bona fide equilibrium ensemble, so we take the resolution $\delta_E$ above to be zero. In that case, $\mathcal{P}$ is the projection onto the diagonal ensemble [32], its action on $\hat{\rho}$ results in

$$\mathcal{P} \hat{\rho} = \sum_\alpha \left( \langle E^0_\alpha | \hat{\rho} | E^0_\alpha \rangle | E^0_\alpha \rangle \langle E^0_\alpha \rangle \right),$$

where $|E^0_\alpha\rangle$ are the eigenkets of $\hat{H}_0$ and $\hat{Q}$, and we have assumed that either there are no degeneracies in the
many-body energy spectrum or that, if present, they are few and unimportant. This is generically the case in many-body interacting quantum systems [20].

We define $\mathcal{P} = 1 - \mathcal{P}$, and note that both $\mathcal{P}, \mathcal{P}$ are projectors, namely, $\mathcal{P}^2 = \mathcal{P}$ and $\mathcal{P}^2 = \mathcal{P}$. A warning is in order, despite the fact that $\mathcal{P}\hat{\rho}$ and $\mathcal{P}\hat{\rho}$ encode all the information about the microscopic distributions of $E_0$ and $Q$ (whose widths are expected to be subextensive in $E_0$ and $Q$ [32]), all that information is not needed and our analysis does not allow one to keep track of it in time. Our results only depend on, and keep track of, the distribution of densities $\rho(e_{0}, q)$.

### B. ‘Mori-Zwanzig’ approach

Let us introduce the Liouville superoperator $\mathcal{L} = -i[\hat{H}, \cdot]$. Then, following Mori-Zwanzig [38, 39], see Appendix A, the theory of linear ordinary differential equations (ODEs) gives us the following rewriting of the $\mathcal{P}$-projected Liouville equation $\partial_{\tau}\mathcal{P}\hat{\rho}(\tau) = \mathcal{L}\mathcal{P}\hat{\rho}(\tau)$

\begin{equation}
\partial_{\tau}\mathcal{P}\hat{\rho}(\tau) = \mathcal{P}\mathcal{L}\mathcal{P}\hat{\rho}(\tau) + \int_{0}^{\tau} d\tau^{'\prime}\mathcal{P}\mathcal{L}_{\tau^{'\prime}}\mathcal{P}\hat{\rho}(\tau - s) + \mathcal{P}\mathcal{L}e^{\tau^{'\prime}\mathcal{L}_{\tau^{'\prime}}}\mathcal{P}\hat{\rho}_{I},
\end{equation}

where $\hat{\rho}_{I} \equiv \hat{\rho}(\tau = 0)$. To bring some structure to this equation, we now split

\begin{equation}
\mathcal{L} = \mathcal{L}_{0} + \mathcal{L}_{1}, \quad \text{with} \quad \mathcal{L}_{0} = -i[\hat{H}_{0}, \cdot] , \quad \mathcal{L}_{1} = -ig[\hat{V}, \cdot],
\end{equation}

and note the properties

\begin{equation}
\mathcal{L}_{0}\mathcal{P} = \mathcal{P}\mathcal{L}_{0} = 0, \quad \mathcal{P}\mathcal{L}_{1}\mathcal{P} = 0.
\end{equation}

They follow from elementary considerations. This allows to recast the above equation of motion (EOM) as

\begin{equation}
\partial_{\tau}\mathcal{P}\hat{\rho}(\tau) = \int_{0}^{\tau} d\tau^{'\prime}\mathcal{A}_{\tau}\mathcal{P}\hat{\rho}(\tau - s) + \mathcal{Y}_{\tau}\mathcal{P}\hat{\rho}_{I},
\end{equation}

where we introduced

\begin{equation}
\mathcal{A}_{\tau} = \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} + \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} + \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} ; \quad \mathcal{Y}_{\tau} = \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} + \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} \mathcal{P}.
\end{equation}

The object $\mathcal{A}_{\tau}$ represents a memory kernel. For completeness, we also report the expression for $\mathcal{P}\hat{\rho}(\tau)$, see Appendix A,

\begin{equation}
\mathcal{P}\hat{\rho}(\tau) = e^{T_{\mathcal{P}}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau}}\mathcal{P}\hat{\rho}_{I} + \int_{0}^{\tau} d\tau^{'\prime} e^{T_{\mathcal{P}}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau}}\mathcal{P}\hat{\rho}(\tau - s) + \mathcal{P}\mathcal{L}e^{T_{\mathcal{P}}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau} + \mathcal{P}\mathcal{L}_{1}\mathcal{L}_{0}^{\tau}}\mathcal{P}\hat{\rho}_{I}.
\end{equation}

It should be stressed that all the previous equations are exact, and, hence, they are not particularly useful. In the next section, we make some motivated approximations.

### C. Equilibration

Our only assumption is that the dynamics generated by $\hat{H}_{0}$ results in fast equilibration of observables to the equilibrium state that is characterized by the expectation value of $H_{0}$, $E_{0} = \text{Tr}(\bar{\rho}_{0}\hat{H}_{0})$, and of $Q$, $Q = \text{Tr}(\bar{\rho}_{0}\hat{Q})$. Once again, we stress that we do not distinguish between nonintegrable reference Hamiltonians $\hat{H}_{0}$, for which observables equilibrate to the thermal predictions (thermalize [20]), from integrable reference Hamiltonians, for which observables equilibrate to the GGE predictions (exhibit generalized thermalization [15]). For our purposes, it does not matter whether equilibration is towards a traditional ensemble of statistical mechanics or towards a GGE: in both cases it means that, for generic (few-body) observables $\hat{O}$ and initial states $\bar{\rho}_{I}$,

\begin{equation}
\left| \frac{\text{Tr}[\hat{\rho}_{I}\hat{O}_{0}(\tau)] - \text{Tr}[\bar{\rho}_{0}\hat{O}_{\tau}]}{\text{Tr}[\bar{\rho}_{0}\hat{O}_{\tau}]} \right| = f(\tau) \rightarrow 0,
\end{equation}

where $\hat{O}_{0}(\tau) = e^{i\tau\hat{H}_{0}}\bar{\rho}_{0}e^{-i\tau\hat{H}_{0}}$. We define an equilibration timescale $\tau^{*}(\epsilon)$, which is understood as the time $\tau$ at which $f(\tau) \approx \epsilon$ for some small dimensionless $\epsilon$. By fast equilibration, we mean that $\tau^{*}(\epsilon) \sim O(1)$ in the relevant time units of the problem. In Sec. VI, numerical simulations of a translationally invariant one-dimensional (1D) Hamiltonian indeed show that, in times that are $O(1)$ in the relevant time-scale, observables reach values that are very close to the predictions of $\bar{\rho}_{0,q}$.

We do not expect $\epsilon$ to be arbitrarily small because, even if the Hamiltonian and the initial state are translationally invariant, the unavoidable coupling to hydrodynamic modes in interacting systems renders the approach of local observables to equilibrium polynomial, $f(\tau) \propto \tau^{-d/2}$, as argued in Ref. [40] on the basis of fluctuating hydrodynamics. In such a case, one expects

\begin{equation}
\tau^{*}(\epsilon) \approx e^{-2/d}. \quad \text{(14)}
\end{equation}

Our numerical results suggests that, for the models and observables studied, hydrodynamics tails may set in when $\epsilon$ is very small, at times that are beyond the reach of our numerical calculations.

The way $\tau^{*}$ enters our analysis is that we approximate, for any $\hat{\rho}$,

\begin{equation}
e^{\tau\mathcal{L}_{0}}\hat{\rho} \approx \mathcal{P}e^{\tau\mathcal{L}_{0}}\hat{\rho}_{I}, \quad \text{whenever} \quad \tau > \tau^{*}, \quad \text{(15)}
\end{equation}

accepting an error $O(\epsilon)$. This approximation amounts to assuming that the system has equilibrated after a time $\tau^{*}$ with respect to the dynamics generated by $\hat{H}_{0}$. Hence, at time $\tau^{*}$, we replace the density matrix of the system by that of the diagonal ensemble. Of course, the usual caveats typical of irreversibility apply: this replacement can only be correct when dealing with local or few-body observables $\hat{O}$, or sums thereof, not for (special) many-body operators such as the spectral projectors of $\hat{H}_{0}$ [20].

### D. Born approximation

We can now state our main assumption as a weak coupling condition, namely,

\begin{equation}
g\tau^{*} \ll 1. \quad \text{(16)}
\end{equation}
To see this assumption at work, let us expand the exponential \( e^{(L_\alpha + \mathcal{P}L_\beta)} \) in the definition of \( \mathcal{A}_s \) in Eq. (11)

\[
\mathcal{A}_s = \mathcal{P}L_1 e^{L_\alpha} L_1 \mathcal{P} + \int_0^\tau ds \mathcal{P}L_1 e^{(s-s_1) L_\alpha} \mathcal{P} e^{s_1 L_\alpha} L_1 \mathcal{P} + \ldots
\]

In order for this expansion in powers of \( g \) to be meaningful, one needs to make sure that the series above can be resummed such that it is linear in \( L \). For example, since every \( L_1 \) carries a factor \( L \), it appears that the first and second term in the equation above are of order \( L^2 \) and \( L^3 \), respectively. However, one can write these terms as sums of [and integrals over \((e_0, q)] truncated correlation functions \( \langle V_0(s)V_\alpha^{(e_0, q)} \rangle \) and \( \langle V_0(s)V_\alpha^{(e_0, q)} \rangle \), see, e.g., the remark following Eq. (B4). These truncated correlation functions are of order \( L \) due to clustering properties of the \((e_0, q)\)-equilibrium ensembles. In higher orders, such considerations become more complicated and we refer the interested reader to mathematical work establishing a meaningful expansion [41, 42]. Assuming that one recovers the linearity in \( L \) to all orders, for any \( \tau \geq \tau^* \),

\[
\int_0^\tau ds \mathcal{A}_s = \int_0^\tau ds \mathcal{P}L_1 e^{sL_\alpha} L_1 \mathcal{P} + \mathcal{O}(g^3 \tau^* L),
\]

where the first term is \( \mathcal{O}(g^2 \tau^* L) \). One then sees that the expansion is meaningful if

\[
g^3 \tau^* L \ll g^2 \tau^* L,
\]

which is, of course, equivalent to Eq. (16).

### E. Markov approximation

The discussion above has shown that the lowest order approximation to \( \int_0^\infty ds \mathcal{A}_s \), namely,

\[
\mathcal{K} = \int_0^\infty ds \mathcal{P}L_1 e^{sL_\alpha} L_1 \mathcal{P},
\]

is \( \mathcal{O}(g^2 \tau^* L) \), with the factor \( L \) originating from the spatial sum in \( L_1 \). The physical meaning of this factor is that the quantities that change smoothly in time are the densities \((e_0, q)\) rather than \((E_0, Q)\) themselves, or, alternatively, local observables. We then see that the superoperator \( \mathcal{K} \) is \( \mathcal{O}(g^2 \tau^*) \), and this is the rate at which \( \hat{\rho} \) changes. Recalling that the time-integral defining \( \mathcal{K} \) reaches its \( \tau = \infty \) value at \( \tau \approx \tau^* \), leads to the conclusion that, in Eq. (10), we can approximate \( \mathcal{P} \hat{\rho}(\tau - s) \) by \( \mathcal{P} \hat{\rho}(\tau) \), making an error \( \mathcal{O}(g^2 \tau^* L) \) (which is small by our weak-coupling assumption \( g \tau^* \ll 1 \)). A further simplification occurs by observing that \( \mathcal{Y}_\tau \) is proportional to \( g \), and vanishes fast as \( \tau \geq \tau^* \), so one might set \( \mathcal{Y}_\tau = 0 \) for times \( \tau \geq \tau^* \). If we make these two approximations, then the EOM [Eq. (10)] reads

\[
\partial_\tau \mathcal{P} \hat{\rho}(\tau) = \mathcal{K} \mathcal{P} \hat{\rho}(\tau), \quad \tau \gg \tau^*.
\]

A conservative look at the validity of Eq. (21), in particular to the approximation \( \mathcal{Y}_\tau = 0 \), shows that one should trust Eq. (21) only for times \( \tau \gg \tau^* \) (where the subindex ‘\( \tau^* \)’ stands for ‘transient’) and \( \tau^* \) is determined by \( g^2 \tau^* \gg g f(\tau^*) \), as follows by comparing the magnitude of the two terms in Eq. (10).

### F. The autonomous equation

The meaning of Eq. (21), in which \( \mathcal{K} \) is a superoperator acting on density matrices, is simplified by the presence of \( \mathcal{P} \). As mentioned before, \( \mathcal{P} \) projects onto distributions of densities \( p(e_0, q) \), and so Eq. (21) is actually an evolution equation on such \( p \). It is easier, and more natural, to guess the form of this evolution equation than to derive it from the formula above, so we will do the former here and relegate the latter to Appendix B.

Given \( \hat{\rho}_{e_0, \tau} \), we need to find the rate of change of \((e_0, q)\). The natural path is to use Fermi’s golden rule, within which \( e_0 \) does not change in time. The rate of change of \( q \), called the ‘drift’ \( d \), is given by

\[
d = 2\pi g^2 \int dE_0^{(0)} \delta(E_0^{(0)} - E_0^{(\tau)}) \frac{Q_0^{(0)} - Q_0^{(\tau)}}{L} |\langle E_0^{(0)}|\hat{V}|E_0^{(\tau)}\rangle|^2,
\]

where the integral is understood as a sum and the delta function \( \delta(E_0^{(0)} - E_0^{(\tau)}) \) selects an interval of energies with a width smaller than any relevant energy scale but much larger than the level spacing. The state \( |E_0^{(0)}\rangle \) is taken to have \((e_0, q)\), so that, by means of eigenstate thermalization (or its generalized version for integrable systems), \( d = d(e_0, q) \) does not depend on \(|E_0^{(0)}\rangle \) but only on \((e_0, q)\). Even though \( \hat{V} = \mathcal{O}(L) \), the drift \( d(e_0, q) \) is \( \mathcal{O}(1) \) because of decay of spatial correlations, see Appendix B. We hence obtain the autonomous equation

\[
\partial_\tau q(q) = d[e_0(q), q(q)], \quad \partial_{e_0}(e_0(q)) = 0.
\]

The corresponding evolution equation for the distributions \( p_r(e_0, q) \) is then

\[
\partial_\tau p_r(e_0, q) = -d[e_0(q), q] \partial_{q_r}(e_0, q),
\]

which lifts naturally to an evolution on \( \hat{\rho} \). The stationary solution of Eq. (23) as \( \tau \to \infty \) is denoted by \((e_0, q^*)\), where \( q^* = q^*(e_0) \) is determined by

\[
d[e_0(q^*), q^*(e_0)] = 0.
\]

Within Fermi’s golden rule [Eq. (22)], this is recognized as a detailed balance condition, indicating that \( q^*(e_0) \) is the equilibrium value of \( q \) given \( e_0 \). In other words, \( q^* \) is determined by maximizing the entropy at fixed \( e_0 \) and, hence, the resulting ensemble \( \hat{\rho}_{e_0, q^*} \) is equivalent to one in which no constraint on \( \hat{q} \) is imposed:

\[
\hat{\rho}_{e_0, q^*} \approx \hat{\rho}_{e_0}.
\]

This also shows that the asymptotic state \( \hat{\rho}_{e_0, q^*} \) is close to the global equilibrium state \( \hat{\rho}_e \) for which \( \langle \hbar \rangle = e = e_0 + \mathcal{O}(g) \). In Sec. IV, we quantify the \( \mathcal{O}(g) \) difference between \( \hat{\rho}_{e_0, q^*} \) and \( \hat{\rho}_e \).
G. Corrections

In principle, our scheme allows to compute higher-order corrections in $g$ to the autonomous Eqs. (21) and (23). In particular, one can expand $A_n$ in a power series in terms of order $(gt^\ast)^n L$, $n \geq 1$, with the terms for $n = 1$ and 2 given in Sec. III D. Integrating over $s$, one then gets a series for $K_n$ and hence for the drift coefficient $d[e_0, q]$. It is, however, far from obvious that such corrections are useful, as: (i) we have made approximations at several points, not only in truncating $A_n$, but also, e.g., replacing $P\hat{\rho}(\tau - s)$ for $s \leq \tau^\ast$ by $P\hat{\rho}(\tau)$, and (ii) a not-so-fast decay of $f(\tau)$ would imply that the $\mathcal{Y}_s$ term in Eq. (10) remains potentially important at long times, obscuring any precise corrections to the Markovian part.

Rigorous work in the context of open quantum systems [41–43] implies that one can meaningfully compute corrections, but only if the condition $\int_0^\infty dt |f(\tau)| < \infty$ is satisfied.

IV. DYNAMICS OF LOCAL OBSERVABLES

So far, we have focused on the evolution of the quasi-conserved quantity $\hat{Q}$, which we found to be described by the autonomous Eq. (23). We now turn our attention to more general observables. Our main finding is that their time-dependent expectation values are described by a ‘deformed equilibrium ensemble’, with an $O(g)$ deviation from $\rho_{c_0,q}(\tau)$. This deviation has a universal form.

A. Correction to $P\hat{\rho}(\tau)$

At first sight, the evolution of generic observables is slaved by the evolution of the slow variables $(e_0, q)$. Indeed, the general picture is that the evolution of the density matrix $\hat{\rho}(\tau)$ takes place in the space of equilibrium density matrices $\hat{\rho}_{c_0,q}(\tau)$, leading to the prediction

$$\text{Tr}[\hat{\rho}(\tau)\hat{O}] \approx \text{Tr}[\hat{O}P\hat{\rho}(\tau)] = \langle \hat{O} \rangle_{c_0,q}(\tau).$$

(27)

In the last equality, we assumed that $p(e_0, q)$ is concentrated at a single value. Yet, strictly speaking, the previous section dealt with $P\hat{\rho}(\tau)$, rather than with $\hat{\rho}(\tau)$, so Eq. (27) is in need of a justification. To provide it, we return to the formalism in Sec. III B, and write [see Eq. (12)]

$$\hat{\rho}(\tau) = P\hat{\rho}(\tau) + \tilde{P}\hat{\rho}(\tau) = P\hat{\rho}(\tau) + e^{sP\mathcal{L}}P\hat{\rho}_1 + \int_0^\tau ds e^{sP\mathcal{L}}P\mathcal{L}P\hat{\rho}(\tau - s).$$

(28)

The first term in Eq. (28) is dominant for large $\tau \geq \tau_0$. It amounts to the approximation resulting in Eq. (27). The second term in the last line of Eq. (28) is transient, decaying as $gf(\tau)$, as discussed also in Sec. III E. The third term is the correction we are interested in. Let us write it explicitly, when paired with an observable $\hat{O}$, namely $\int_0^\tau ds \text{Tr}[\hat{O}P e^{sP\mathcal{L}}\mathcal{L}P\hat{\rho}(\tau - s)]$. By expanding $e^{sP\mathcal{L}}P$ in powers of $g$, we get the leading contribution

$$\int_0^\tau ds \text{Tr}[\hat{O}P e^{sP\mathcal{L}}\mathcal{L}P\hat{\rho}(\tau - s)] = ig\int_0^\tau ds \text{Tr}[[\hat{V}_0(-s), \hat{O}]P\hat{\rho}(\tau - s)].$$

(29)

One can further simplify this expression by approximating $P\hat{\rho}(\tau - s)$ by $P\hat{\rho}(\tau)$, justified by the reasoning in Sec. III E, and by again assuming that the distribution $p(e_0, q)$, corresponding to $P\hat{\rho}(\tau)$, is concentrated at a single density $[e_0, q(\tau)]$. Then the last line in Eq. (29) reads

$$ig\int_0^\infty ds [[\hat{V}_0(-s), \hat{O}]_{e_0,q}(\tau)].$$

(30)

By the assumption of fast equilibration at $s \approx \tau^\ast$, this expression is $O(g\tau^\ast)$, i.e., it is a small correction to $\langle \hat{O} \rangle_{c_0,q}(\tau)$. Even though it is subleading, this correction has a universal form, as we explain in the next section. Finally, note that Eq. (30), obtained following a naive first order perturbation theory, was already written in Sec. II for $Q$ [see Eq. (2)]. In that case, the $s$-independent integrand vanishes identically.

B. Susceptibility

To understand the correction in Eq. (30), consider the equilibrium ensemble $\hat{\rho}_{e_0,q^*}$ with $(e_0, q^*)$ the $\tau \to \infty$ solution of Eq. (23), or, alternatively, with $q^*$ determined by maximizing the entropy at fixed $e_0$. The ensemble $\hat{\rho}_{e_0,q^*}$ is a small perturbation of the real equilibrium ensemble $\hat{\rho}_c$ corresponding to $\hat{H}$, with $|\hat{h}|_{\hat{\rho}_c} = |\hat{h}|_{\hat{\rho}_{e_0,q^*}} + O(g)$. We can relate these two ensembles by linear response theory. Indeed, we can imagine starting with $\hat{\rho}_{e_0,q^*}$ at $\tau = 0$, and switching on the perturbation $g\hat{V}$. The system will then evolve precisely to $\hat{\rho}_c$, and the change in the expectation value of observables $\hat{O}$ should be described by linear response theory. In particular, the stationary change is described by the zero-frequency response coefficient (also known as the susceptibility), which is exactly Eq. (30).

This discussion should also clarify how thermalization with respect to $\hat{H}$ is reconciled with our treatment, which is based on equilibration with respect to $\hat{H}_0$. The global equilibrium state $\hat{\rho}_c$ is obtained as a universal correction to the state $\hat{\rho}_{e_0,q^*}$.

C. Deformed equilibrium ensembles from Fermi’s golden rule

In Sec. IV A, we used the last term in Eq. (28) to derive the $O(g)$ correction to observables in $\hat{\rho}(\tau)$ from their expectation values in $\hat{\rho}_{c_0,q}(\tau)$. Here, we point out that this is also what the autonomous Eq. (23) dictates.

Let us compute the time derivative of $q(\tau)$:

$$\partial_\tau q(\tau) = \text{Tr}[\hat{q}\mathcal{L}\hat{\rho}(\tau)] = \text{Tr}[\hat{q}\mathcal{L}_1\hat{\rho}(\tau)],$$

(31)
where, in the last equality, we used that $[\hat{H}_0, \hat{Q}] = 0$. Previously, we evaluated this time derivative in a different way: we approximated $\dot{\rho}(\tau)$ by $\dot{\rho}_{eq, q}(\tau)$ (assuming, again, that $\rho(0,q)$ is a Dirac delta function) and we arrived at

$$\partial_{\tau} q(\tau) = \text{Tr}[\hat{q} K \dot{\rho}_{eq, q}(\tau)].$$

(32)

This is a more abstract rendering of the autonomous Eq. (23). At this point, one can ask which form of $\dot{\rho}(\tau)$ would reconcile Eqs. (31) and (32). By the cyclic property of the trace, and using that $[\hat{q}, \dot{\rho}_{eq, q}(\tau)] = 0$, we note that Eq. (31) does not depend on the leading contribution $\dot{\rho}_{eq, q}(\tau)$ to $\dot{\rho}(\tau)$. It depends only on the correction term. We then see that Eq. (31) reduces to Eq. (32) if we choose the correction term to be

$$\int_0^\infty ds \bar{E}_0 e^s \mathcal{L}_1 \dot{\rho}_{eq, q}(\tau),$$

which is the correction we used in Sec. IV A to derive Eq. (30). A spectacular corollary of this reasoning is the observation that, if $\dot{\rho}(\tau)$ were exactly commuting with $\hat{q}$, then the rate of change vanishes, i.e., $\partial_{\tau} q(\tau) = 0$.

V. MODELS, QUENCHES, AND OBSERVABLES

In the rest of the manuscript, we test the previous ideas and analytical results by means of numerical calculations. We focus on the dynamics of nonintegrable models of strongly interacting hard-core bosons in one-dimensional (1D) lattices (which can be mapped onto spin-1/2 Hamiltonians). We study the effects of breaking particle-number conservation [the U(1) symmetry in the corresponding spin models].

A. Models

Quantum dynamics are studied under time-independent Hamiltonians $\hat{H}_\alpha$ of the form

$$\hat{H}_\alpha = \hat{H}_0 + g_\alpha \hat{V}_\alpha,$$

(34)

where the reference Hamiltonian $\hat{H}_0$, which we take to be nonintegrable, commutes with the total particle-number operator $\hat{N}$, $[\hat{H}_0, \hat{N}] = 0$. The perturbations, $g_\alpha \hat{V}_\alpha$, do not commute with $\hat{N}$, $[\hat{V}_\alpha, \hat{N}] \neq 0$.

We take $\hat{H}_0$ to be the $tV$-$t'V'$ Hamiltonian for hard-core bosons in 1D lattices [44, 45], with nearest (next-nearest) neighbor hopping $t$ ($t'$) and interaction $V$ ($V'$):

$$\hat{H}_0 = \sum_i \left[ -t \left( \hat{b}^\dagger_i \hat{b}_{i+1} + \text{H.c.} \right) - t' \left( \hat{b}^\dagger_i \hat{b}_{i+2} + \text{H.c.} \right) \right] + V \left( \hat{n}_i - \frac{1}{2} \right) \left( \hat{n}_{i+1} - \frac{1}{2} \right) + V' \left( \hat{n}_i - \frac{1}{2} \right) \left( \hat{n}_{i+2} - \frac{1}{2} \right),$$

(35)

where $\hat{b}^\dagger_i$ ($\hat{b}_i$) is the hard-core boson creation (annihilation) operator and $\hat{n}_i = \hat{b}^\dagger_i \hat{b}_i$ is the number operator at site $i$. When $t' = V' = 0$, $\hat{H}_0$ is integrable (in the spin language, it is the Hamiltonian of the spin-1/2 XXZ chain) [46]. Here, we focus on cases in which $t' = V' \neq 0$ so that $\hat{H}_0$ is nonintegrable [44, 45].

We consider two perturbations $g_\alpha \hat{V}_\alpha$, with $\alpha = 1, 2$. The first one is

$$g_1 \hat{V}_1 = g_1 \sum_i \left( \hat{b}_i + \frac{1}{2} \left( \hat{b}_{i+1} + \hat{b}^\dagger_{i+1} \right) + \text{H.c.} \right).$$

(36)

It will be important later that the presence of nearest neighbor hopping terms in $\hat{V}_1$ make $\langle E_0^\alpha | \hat{V}_1 | E_0^\alpha \rangle \neq 0$ for typical eigenkets of $\hat{H}_0$.

The second perturbation we consider is

$$g_2 \hat{V}_2 = g_2 \sum_i \left( \hat{b}_i + \frac{1}{2} \hat{b}_{i+1} + \text{H.c.} \right).$$

(37)

This perturbation only contains terms that change the particle number. Hence, $\langle E_0^\alpha | \hat{V}_2 | E_0^\alpha \rangle = 0$ for all eigenkets $| \alpha \rangle$ of $\hat{H}_0$ and $\hat{N}$.

B. Initial states and description after equilibration

We study the quantum dynamics of initial states $\dot{\rho}_I$ that are far from equilibrium with respect to both $\hat{H}_0$ and $\hat{H}_\alpha$. This is achieved by choosing $\dot{\rho}_I$ to be thermal equilibrium states of initial Hamiltonians $\hat{H}_\alpha$, such that $[\hat{H}_I, \hat{H}_\alpha] \neq 0$ and $[\hat{H}_I, \hat{H}_0] \neq 0$. Dynamics are generated by quantum quenches in which, at time $\tau = 0$, one suddenly changes $\hat{H}_I \rightarrow \hat{H}_\alpha$ and lets the systems evolve unitarily. We consider systems that are translationally invariant before and after the quench.

The time-evolving density matrix after the quench can be written as $\dot{\rho}(\tau) = e^{-i \hat{H}_0 \tau} \dot{\rho}_I e^{i \hat{H}_0 \tau}$. We are interested in the dynamics of observables $\hat{O}$, whose expectation values are given by $O(\tau) = \text{Tr} [\hat{O} \dot{\rho}(\tau)]$. At long times, one expects observables to equilibrate at the values predicted by the diagonal ensemble (DE), $O_{\text{DE}} = \text{Tr} [\rho_{\text{DE}} \hat{O}]$, where $\rho_{\text{DE}} = \lim_{\tau \rightarrow \infty} (1/\tau) \int_0^\tau d\tau \dot{\rho}(\tau)$ is the density matrix of the diagonal ensemble [32]. When written in the basis of eigenkets $| E_0^\alpha \rangle$ of $\hat{H}_\alpha$, $\rho_{\text{DE}}$ takes the form

$$\dot{\rho}_{\text{DE}} = \sum_i \langle E_0^\alpha | \rho_I | E_i^\alpha \rangle | E_i^\alpha \rangle \langle E_i^\alpha |.$$

(38)

For nonintegrable (quantum chaotic) systems, because of eigenstate thermalization [32–34], one expects the predictions of the diagonal ensemble to match those of traditional statistical mechanics ensembles, namely, we expect observables to thermalize [20]. For the $tV$-$t'V'$ Hamiltonian for hard-core bosons in 1D lattices, eigenstate thermalization and thermalization were studied in Ref. [44], while quantum chaos was studied in Ref. [45]. This means that, in our systems, we can also describe observables after equilibration by means of the grand canonical ensemble (GE) characterized by a temperature $T$, and, when
particle number is a conserved quantity, by a chemical potential $\mu$. The grand canonical ensemble density matrices have the form

$$
\hat{\rho}_{\text{GE}} = \begin{cases} 
\frac{e^{-\hat{H}_\alpha/T}}{\text{Tr}[e^{-\hat{H}_\alpha/T}]}, & \text{when } g_\alpha \neq 0, \\
\frac{e^{-(\hat{H}_\alpha+\mu\hat{N})/T}}{\text{Tr}[e^{-(\hat{H}_\alpha+\mu\hat{N})/T}]}, & \text{when } g_\alpha = 0.
\end{cases}
$$

When $g_\alpha \neq 0$, $T$ is fixed by the (conserved) energy of the time-evolving state

$$
\text{Tr}[\hat{\rho}_{\text{GE}}\hat{H}_\alpha] = \text{Tr}[\hat{\rho}_I\hat{H}_\alpha].
$$

When $g_\alpha = 0$, $T$ and $\mu$ are determined by the (conserved) energy [Eq. (40)] and by the (conserved) particle number in the time-evolving state

$$
\text{Tr}[\hat{\rho}_{\text{GE}}\hat{N}] = \text{Tr}[\hat{\rho}_I\hat{N}].
$$

Due to particle-hole symmetry in the Hamiltonians $\hat{H}_\alpha$, when $g_\alpha \neq 0$ (namely, in the absence of particle-number conservation), our systems after equilibration are always at half filling irrespective of the initial filling $n_I$. We consider initial fillings $n_I \neq 1/2$, which means that the filling must change during the dynamics when $g_\alpha \neq 0$.

From our analysis in the previous sections, we expect that, for small values of $g_\alpha$, the dynamics of generic observables follow a two-step process towards thermalization: (i) fast relaxation driven by $\hat{H}_0$ (prethermal dynamics), and (ii) slower, nearly exponential, relaxation to the thermal equilibrium predictions (thermalization dynamics). At long times, close to thermal equilibrium, hydrodynamics may become dominant and algebraic relaxation is expected to take place [40]. That regime is not resolved in this work.

As discussed in Sec. IV, the near-exponential dynamics following prethermalization can be described by projecting $\hat{\rho}(\tau)$ in the basis of the eigenkets of $\hat{H}_0$, up to an $O(\mu)$ correction. This projected state is a diagonal ensemble of $\hat{H}_0$, whose density matrix [see Eq. (6)] we denote as

$$
\hat{\rho}_0(\tau) \equiv \mathcal{P}\hat{\rho}(\tau) = \sum_i \langle \langle E_i^0|\hat{\rho}(\tau)|E_i^0\rangle \rangle |E_i^0\rangle\langle E_i^0|,
$$

where $|E_i^0\rangle$ are the eigenkets of $\hat{H}_0$.

C. Numerical Linked Cluster Expansion

In what follows, we use the numerical linked cluster expansion (NLCE) approach introduced in Ref. [21] (see Refs. [47, 48] for NLCE studies in two dimensions) to study the quantum dynamics of various observables in our translationally invariant 1D systems in the thermodynamic limit.

NLCEs were originally introduced to study systems in thermal equilibrium [49], and allow one to obtain the expectation value of extensive observables per site ($\mathcal{O} = O/L$), in the thermodynamic limit ($L \to \infty$), as sums over the contributions of the connected clusters that can be embedded in the lattice. Given the connected clusters $c$, which can be embedded in the lattice in $M(c)$ ways per site and have weights $W_O(c)$, one obtains $\mathcal{O}$ in the following way

$$
\mathcal{O} = \sum_c M(c) \times W_O(c).
$$

$W_O(c)$ is computed, for each cluster $c$, from the expectation value of the observable $\hat{O}$ in the cluster ($\mathcal{O'}$) using the inclusion-exclusion principle:

$$
W_O(c) = \mathcal{O'} - \sum_{s \subseteq c} W_O(s),
$$

where the sum is over all connected sub-clusters of $c$. For the smallest cluster, $W_O(c) = \mathcal{O'}$. For each cluster, $\mathcal{O'} = \text{Tr}[\hat{\rho}' \hat{O}]$, where $\hat{\rho}'$ is the density matrix of the relevant ensemble in the cluster. In NLCEs, $\mathcal{O'}$ is calculated exactly numerically using full exact diagonalization.

In our calculations, the density matrix of the initial state in each cluster, $\hat{\rho}_I$, is taken to be the grand canonical density matrix set by the initial Hamiltonian in each cluster $\hat{H}_I^c$ (our initial states are in thermal equilibrium with respect to $\hat{H}_I$). In all quenches, we take $\hat{H}_I$ to be the $t$-$t'$-$V$-$V'$ model in Eq. (35). Since $[\hat{H}_I^c, \hat{N}^c] = 0$, where $\hat{N}^c$ is the total particle number operator in the cluster,

$$
\hat{\rho}_I^c = \frac{e^{-(\hat{H}_I^c+\mu_I\hat{N}^c)/T_I}}{\text{Tr}[e^{-(\hat{H}_I^c+\mu_I\hat{N}^c)/T_I}]},
$$

Also, since $\hat{H}_I^c$ is particle-hole symmetric, we need $\mu_I \neq 0$ in order to have initial states with filling $n_I \neq 1/2$.

To calculate the time evolution of the expectation values of the observables, $\mathcal{O}(\tau)$, where $\tau$ denotes the time after the quench, the density matrix of each cluster is evolved with the Hamiltonian after the quench $\hat{H}_I^c$

$$
\hat{\rho}^c(\tau) = \left(e^{-i\hat{H}_I^c\tau}\right) \hat{\rho}_I^c \left(e^{i\hat{H}_I^c\tau}\right),
$$

and the NLCE calculation is carried out as usual [21]. Our Hamiltonians after the quench have the form in Eq. (34). Similarly, in order to obtain NLCE results after the quench for the diagonal ensemble, the grand canonical ensemble, and in the projected basis of $\hat{H}_0$, we use $\hat{\rho}_{\text{DE}}, \hat{\rho}_{\text{GE}},$ and $\hat{\rho}_0(\tau)$ from Eqs. (38), (39), and (42), respectively, for each cluster $c$.

NLCEs have been used to study quenches in the $t$-$t'$-$V$-$V'$ model [Eq. (35)] to understand the dynamics [21], and the description of observables after equilibration [50], at the integrable point $t' = V' = 0$ and away from it $t' = V' \neq 0$. The presence of next-nearest neighbor hoppings and interactions makes it possible to have different building blocks to construct the clusters in the NLCE. In Ref. [51], it was shown that maximally connected clusters
– built adding contiguous sites and all possible bonds, starting from one site – are optimal for studying quenches in this model \cite{51}. Here, as in Refs. \cite{21, 50}, we use that \(N\) is the order of NLCE, and we denote the value of an observable \(O(\tau)\) evaluated with NLCE to order \(l\) as \(O_l(\tau)\).

D. Observables

We study three observables which have properties that make them qualitatively distinct in the context of dynamics and description after equilibration.

The first observable is the total particle number
\[N(\tau) = \text{Tr}[\hat{\rho}(\tau)\hat{N}],\]
whose value per site, the particle filling, is denoted as \(n(\tau)\). This is a conserved quantity with respect to the reference Hamiltonian \(\hat{H}_0\), i.e., it only changes during dynamics under \(\hat{H}_\alpha\) after the quench if \(g_\alpha \neq 0\).

The second observable is the nearest neighbor hopping correlation function
\[K(\tau) = \text{Tr}[\hat{\rho}(\tau)\hat{K}],\]
whose value per site is denoted as \(k(\tau)\), where
\[\hat{K} = \sum_i \left( \hat{b}_i^\dagger \hat{b}_{i+1} + \hat{b}_{i+1}^\dagger \hat{b}_i \right).\]

\(k(\tau)\) is a local observable, closely related to the kinetic energy per site. It changes during the dynamics independently of whether \(\hat{N}\) is conserved or not.

The third observable is the distribution function
\[M_k(\tau) = \text{Tr}[\hat{\rho}(\tau)\hat{M}_k],\]
whose value per site is the momentum distribution function, denoted as \(m_k(\tau)\). \(\hat{M}_k\) is the unnormalized (to make each \(k\) component extensive) Fourier transform of the one-body density matrix, given by
\[\hat{M}_k = \sum_{j,j'} e^{ik(j-j')} \hat{b}_j^\dagger \hat{b}_{j'}.\]

\(m_k(\tau)\) is a non-local observable that changes during the dynamics independently of whether \(\hat{N}\) is conserved or not. This observable is of particular interest because it is regularly measured (using time-of-flight expansion) in experiments with ultracold quantum gases \cite{52}. \(m_k(\tau)\) was the observable used in Ref. \cite{2} to show lack of thermalization in 1D Bose gases with contact interactions, and in Ref. \cite{6} to study prethermalization and thermalization 1D Bose gases with dipolar interactions.

E. Parameters used in the calculations

The initial state is taken to be a thermal equilibrium state at temperature \(T_I = 10\) and chemical potential \(\mu_1 = 2\) (similar results are obtained for other values of \(T_I\), not too low, and \(\mu_1\)). For an initial Hamiltonian \(\hat{H}_I\) with nearest and next-nearest neighbor coupling parameters \(t_I = 0.5, V_I = 1.5\), and \(t'_I = V'_I = 0.7\). After the quench, \(\hat{H}_\alpha\) has coupling parameters \(t = V = 1\) (these set the energy scale in our calculations), \(t' = V' = 0.7\), and \(g_\alpha \in (0, 0.12)\). For these parameters after the quench, \(\hat{H}_0\) \cite{44, 45} and \(\hat{H}_\alpha\) are quantum chaotic and the system thermalizes for all values of \(g_\alpha\) (see Ref. \cite{21} for an NLCE study of the quench dynamics when \(g_\alpha = 0\)).

For \(n(\tau)\) and \(k(\tau)\), we carry out the NLCE up to the 17th order for quenches with \(g \neq 0\) (the largest Hamiltonian sector that needs to be diagonalized has 65792 states after exploiting reflection symmetry), and, thanks to particle number conservation, up to the 19th order for quenches with \(g = 0\) (the dimension of the largest Hamiltonian sector that needs to be diagonalized is 46252). For \(m_k(\tau)\), the NLCE is carried out to one order lower than for \(n(\tau)\) and \(k(\tau)\), namely, up to the 16th order for \(g \neq 0\) and up to the 18th order for \(g = 0\). This is because of the overhead generated by the calculation of the dynamics of all the matrix elements of the one-body density matrix [see Eqs. (50) and (51)].

VI. NUMERICAL RESULTS

A. Dynamics of the particle filling

In Fig. 1(a), we show the evolution of the particle filling under \(\hat{H}_I\) for three values of \(g_1\). Results are shown for the last two orders of the NLCE up to \(\tau = 100\). In those quenches, we expect the convergence errors for \(n(\tau)\) to be below 0.01\% for times \(\tau \lesssim 4\), and to remain low (below 1\%) up to times \(\tau \approx 16\) (see Appendix C). For \(\tau \gtrsim 16\), the results for the last two orders of the NLCE can be seen to (slightly) deviate from each other in Fig. 1(a). In all the plots in Fig. 1(a), \(n(\tau)\) can be seen to approach \(n_{DE} = 1/2\). For \(g_1 = 0.12\), \(n(\tau = 100) \approx n_{DE}\).

The approach of \(n(\tau)\) to \(n_{DE} = 0.5\) is exponential. This is apparent in Fig. 1(b), where we plot the normalized “distance” to equilibrium
\[\delta_{DE}^I[n(\tau)] = \left| \frac{n(\tau) - n_{DE}}{n_{DE}} \right|.\]

We fit \(\delta_{DE}^I[n(\tau)]\) to an exponential function \(\propto \exp[-\Gamma_{NLCE}^I(g_1)\tau]\) to obtain the thermalization rate \(\Gamma_{NLCE}^I(g_1)\). In order to gain an understanding of the accuracy of the obtained rates, we carry out fits in the time interval \(\tau \in [1, 16]\) for \(l = 17\) [the corresponding fits are shown in Fig. 1(b) as thin continuous lines], and in the interval \(\tau \in [1, 6]\) for \(l = 16\). The rates obtained in those calculations are shown in Fig. 1(c), as NLCE-17...
Next, we study the dynamics of the nearest neighbor hopping

\[ \dot{\rho}(\tau) = 0 \]

whenever the state $\hat{\rho}(\tau)$ commutes exactly with $\hat{Q}$. The rate also vanishes if $\hat{\rho}(\tau)$ is time-reversal invariant about $\tau$. Both conditions apply to our initial states $\hat{\rho}_1$. As a result, there is a narrow plateau in $n(\tau)$ for $\tau \leq 1$. This plateau is best seen in Fig. 7. This is why, to obtain the rates reported in Fig. 1(c), we fit $n(\tau)$ at times $\tau \geq 1$.

Next, we show that the values obtained for $\Gamma_{\text{NLCE}}(g_1)$ are in agreement with the ones predicted by Fermi’s golden rule. From Eq. (22), changing $\hat{Q} \rightarrow \hat{N}$ and $\hat{V} \rightarrow \hat{V}_\alpha$, one can write

\[
\dot{n}(\tau) = \frac{2\pi g^2}{L} \sum_{\langle ij \rangle} \delta(E^0_i - E^0_j) (N_j - N_i) P_i(\tau) \times \left| \langle E^0_j | \hat{V}_\alpha | E^0_i \rangle \right|^2,
\]

where $\dot{n}(\tau) = \frac{dn}{d\tau} |_{\tau = \tau_0}$, $E^0_i$ are the eigenkets of $H_0$ with energies $E^0_i$, $N_i = \langle E^0_i | \hat{N} | E^0_i \rangle$, and $P_i(\tau) = \langle E^0_i | \rho(\tau) | E^0_i \rangle$. To evaluate Eq. (53) numerically, we replace the delta function by a sum over energies $E^0_i$ that lie within a small energy window $[E^0_i - \Delta E/2, E^0_i + \Delta E/2]$ (see Appendix D).

The thermalization rates are estimated by computing

\[
\Gamma_{\text{Fermi}}(g_a) = -\frac{\dot{n}(\tau)}{n(\tau) - n_{\text{DE}}},
\]

When $|n(\tau) - n_{\text{DE}}| \ll n_{\text{DE}}$, $\Gamma_{\text{Fermi}}(g_a)$ becomes independent of $\tau$ and $n(\tau)$ relaxes exponentially. This condition is, to a good degree, satisfied for our choice of initial state, for which $\delta_{\text{DE}}[n(\tau)] < 0.09$ [see Fig. 1(b)]. Our calculations of $\Gamma_{\text{Fermi}}(g_a)$ are done using full exact diagonalization in chains with $L = 17$ and 18 sites, and periodic boundary conditions. We identify a range of values for $\Delta E$ and $\tau$ for which the results for $\Gamma_{\text{Fermi}}(g_a)$ are robust against the choice of $\Delta E$, $\tau$, and $L$ (see Appendix D).

In Fig. 1(c), we report our results for $\Gamma_{\text{Fermi}}(g_a)$. They are in excellent agreement with $\Gamma_{\text{NLCE}}(g_1)$. We should add that, for quenches $\hat{H}_I \rightarrow \hat{H}_2$, Eq. (53) predicts the same leading $O(g_a^2)$ dynamics as under $\hat{H}_I$. This is the case because the terms that change the total particles number are same in $\hat{V}_1$ and $\hat{V}_2$. Hence, $n(\tau)$ is the same for both Hamiltonians up to corrections $O(g_a^3)$.

\[ \text{rate} \]

\[ \text{rate} = \text{rate} \]

\[ \text{rate} = \text{rate} \]

**B. Dynamics of the nearest neighbor hopping**

Next, we study the dynamics of the nearest neighbor hopping: $k(\tau)$ [see Eq. (49)]. In contrast to the particle filling studied in Sec. VIA, the nearest neighbor hopping exhibits dynamics even if $g_\alpha = 0$ after the quench.
FIG. 2. Dynamics of the nearest neighbor hopping $k(\tau)$ when $g_\alpha = 0$ after the quench. Results are shown for the 18th (NLCE-18) and 19th (NLCE-19) orders of the NLCE. We also report, as horizontal lines, the results for the 19th order of the diagonal ensemble (DE19) and for the 19th order of the grand canonical ensemble (GE19).

Figure 2 shows the 18th (NLCE-18) and 19th (NLCE-19) orders of the NLCE for $k(\tau)$ after a quench in which $g_\alpha = 0$. The results of both orders are in excellent agreement with each other, and equilibrate rapidly to the prediction of the diagonal ensemble (horizontal dashed line). The predictions of the diagonal and grand canonical (horizontal contiguous line) ensembles are very close to each other, indicating thermalization. This is expected because $H_0$ after the quench is nonintegrable. The small difference between the diagonal and the grand canonical ensemble results are due to the lack of convergence of the NLCE for the former (the latter is fully converged) [50]. Those differences decrease with increasing the order of the NLCE. In what follows, we use the grand canonical ensemble predictions to probe thermalization.

In Fig. 3, we show $k(\tau)$ after quenches $\hat{H}_I \rightarrow \hat{H}_1$ (left panels) and $\hat{H}_I \rightarrow \hat{H}_2$ (right panels). The dynamics are qualitatively similar in both cases. They can be split in two regimes: (i) fast (prethermal) dynamics driven by $\hat{H}_0$ (note that, at times $\tau \lesssim 10$, dynamics in Fig. 3 are nearly identical to those in Fig. 2), and (ii) a slower (thermalization) dynamics controlled by the strength of the perturbation. During the latter regime, the system approaches [and reaches for $g_1 = g_2 = 0.12$ and $\tau = 100$, see Figs. 3(b) and 3(d)] the prediction of the grand canonical ensemble $\hat{\rho}_{GE}$ corresponding to $\hat{H}_\alpha$ after the quench (with the temperature set by the initial state).

The slow approach to thermal equilibrium can be well described using the projected state $\hat{\rho}_0(\tau)$ from Eq. (42). In Fig. 3, we show the results for the projected dynamics along with those for the actual dynamics. As follows from the discussion in Sec. IV, the results for the projected dynamics approach those of a thermal equilibrium state of $\hat{H}_0$, which we denote as $\hat{\rho}_{GE\alpha}$. We compute the temperature $T_\alpha$ in $\hat{\rho}_{GE\alpha}$ using the expectation value of $\hat{H}_0$ in the thermal equilibrium state $\hat{\rho}_{GE}$ of $\hat{H}_\alpha$

$$\text{Tr}[\hat{\rho}_{GE\alpha} \hat{H}_0] = \text{Tr}[\hat{\rho}_{GE} \hat{H}_0].$$

The chemical potential in $\hat{\rho}_{GE\alpha}$ is $\mu_0 = 0$ because, for $g_\alpha \neq 0$, the systems after equilibration are at half filling.

In the left panels in Fig. 3, for quenches $\hat{H}_I \rightarrow \hat{H}_1$, one can see the advanced offsets (see Sec. IV) between the actual dynamics and the projected dynamics. The offsets are much smaller for quenches $\hat{H}_I \rightarrow \hat{H}_2$, whose results are shown in the right panels in Fig. 3. The offsets between the actual and projected dynamics remain constant at long times and are, essentially, the difference between the predictions of $\hat{\rho}_{GE}$ and $\hat{\rho}_{GE\alpha}$.

The difference in the offsets generated by quenches $\hat{H}_I \rightarrow \hat{H}_1$ and $\hat{H}_I \rightarrow \hat{H}_2$ can be understood using Eq. (30), replacing $\hat{O} \rightarrow \hat{K}$ ($\hat{K}$ is defined in Eq. (49)) and $\hat{V} \rightarrow \hat{V}_\alpha$. Since $\hat{K}$ and $\hat{\rho}_0(\tau)$ are block diagonal in the particle number basis, only the presence of terms in $V_\alpha$ that do not change the particle number can produce an $O(\alpha)$ correction. Such terms are present in $V_1$ (the hopping terms), see Eq. (36), but are absent in $V_2$, see Eq. (37). This means that, to leading order, $\Delta k_1(\tau) \propto \alpha$ while $\Delta k_2(\tau) \propto \alpha^2$ with $\beta \geq 2$.

In Fig. 4, we show the long-time offsets between the actual and projected dynamics as functions of $g_\alpha$ (for positive and negative values of $g_\alpha$) in the quenches $\hat{H}_I \rightarrow \hat{H}_1$ [Fig. 4(a)] and $\hat{H}_I \rightarrow \hat{H}_2$ [Fig. 4(b)]. The offsets are computed as the difference between the predictions of $\hat{\rho}_{GE}$ and $\hat{\rho}_{GE\alpha}$ for the equilibrated results (see the horizontal lines in Fig. 3). Those predictions are converged to ma-
chine precision in the 17th order of the NLCE shown in Fig. 4. Figure 4(a), for quenches $\hat{H}_1 \rightarrow \hat{H}_2$, makes apparent the presence of a leading linear correction and of a subleading quadratic one. Figure 4(b) shows the absence of the linear correction for quenches $\hat{H}_1 \rightarrow \hat{H}_2$. There, the leading correction is quadratic, and our numerical results allow us to identify a subleading cubic correction, which leads to a weak asymmetry about $g_2 = 0$.

Comparing the results reported in Fig. 1(a), and in Figs. 3(a) and 3(b), for $g_1 \equiv 0.06$ and $0.12$, it becomes apparent that $k(\tau)$ equilibrates (reaches the long-time horizontal line prediction) faster than $n(\tau)$. This can be understood to be the result of the nearest neighbor hopping being particle-hole symmetric (as the Hamiltonians $\hat{H}_\alpha$ are). This means that, close to equilibrium, the difference between $k(\tau)$ and $k_{\text{GE}}$ can only be a function of even powers of the difference between $n(\tau)$ and $n = 1/2$. Say the leading even power in the difference between $n(\tau)$ and $n = 1/2$ entering in $k(\tau)$ is 2, then $k(\tau) \simeq 2g_{\text{Fermi}}^2(\nu_n)k(\tau)$. Our numerical results for the thermalization rates of $k(\tau)$, not shown, support the correctness of this simple analysis.

C. Dynamics of the momentum distribution

Here, we study the dynamics of momentum distribution function $m_k(\tau)$ [Eq. (50)] for quenches $\hat{H}_1 \rightarrow \hat{H}_2$. Similar to the dynamics of $k(\tau)$, the evolution of $m_k(\tau)$ can be split into a fast prethermalization dynamics, and a slower relaxation to the thermal equilibrium prediction at a rate controlled by the strength of the perturbation.

In Fig. 5, we show results for the 15th (NLCE-15) and 16th (NLCE-16) orders of the NLCE for $m_k(\tau)$ at four times ($\tau = 0, 2, 10$, and $100$) after quenches in which $g_1 = 0.03$ (left panels in Fig. 5) and $g_1 = 0.12$ (right panels in Fig. 5). In all panels in Fig. 5, we also show results for the 15th (NLCE-15) and 16th (NLCE-16) orders of the NLCE for the projected dynamics, and of the grand canonical ensemble for the projected dynamics ($\text{GE}_0$), all evaluated at the 16th order of the NLCE, in the basis of $\hat{H}_0$ [see the legend in (a)], and for the projected dynamics in the basis of $\hat{H}_0$. The predictions of the diagonal ensemble ($\text{DE}$) and the grand canonical ensemble ($\text{GE}$) for the dynamics, and of the grand canonical ensemble for the projected dynamics ($\text{GE}_0$), are shown in (d) and (h) [see the legends in (d)]. The inset in (h) shows the predictions of the latter three ensembles about $k = \pi$.
and remain so at long times [see Figs. 5(c), 5(d), 5(g), and 5(h)].

In Figs. 5(d) and 5(h), we also show the 16th order NLCE prediction of the diagonal and the grand canonical ensembles for the long-time dynamics, and of the grand canonical ensemble for the long-time projected dynamics [see Eq. (55) and the discussion surrounding this equation]. For \( g_1 = 0.03 \) [Fig. 5(d)], the results of the three equilibrium ensembles are nearly indistinguishable from each other. They differ from those of the dynamics at \( \tau = 100 \) (they all become nearly indistinguishable from each other at later times). For \( g_1 = 0.12 \) [Fig. 5(h)], the results for the three equilibrium ensembles and for the dynamics at \( \tau = 100 \) agree with each other. This, in contrast to the results for \( g_1 = 0.03 \), makes apparent that \( m_k(\tau) \) thermalizes faster with increasing the magnitude of \( g_1 \) (as expected). Also, as expected from our discussion in Sec. IV and for \( k(\tau) \) (see Fig. 3) in Sec. VI B, there is an \( O(g_1) \) offset between the results for \( m_k(\tau) \) in the dynamics and in the projected dynamics. The magnitude of this offset is momentum dependent, as made apparent by the inset in Fig. 5(d), and is too small to be resolved at the scales used in the main panels of Fig. 5.

The operator corresponding to the momentum distribution function, unlike the one for the nearest neighbor hopping, is not particle-hole symmetric. This implies that, as \( m_k(\tau) \) approaches its equilibrium value in the diagonal ensemble \( m_k^{\text{DE}} \), to leading order

\[
m_k(\tau) - m_k^{\text{DE}} \propto \left[ n(\tau) - \frac{1}{2} \right].
\]  

Thus, we expect \( m_k \) to thermalize with the same rate \( \Gamma_{\text{Fermi}}(g_1) \) given by Eq. (54). To quantify the “distance” to equilibrium for \( m_k(\tau) \), we compute

\[
\delta_k^{\text{DE}}[m(\tau)] = \frac{\sum_k \left| m_k(\tau) - m_k^{\text{DE}} \right|}{\sum_k m_k^{\text{DE}}}. 
\]  

In Fig. 6(a), we show \( \delta_k^{\text{DE}}[m(\tau)] \) evaluated at the 15th (NLCE-15) and 16th (NLCE-16) orders of the NLCE. \( \delta_k^{\text{DE}}[m(\tau)] \) and \( \delta_k^{\text{DE}}[m(\tau)] \) can be seen to decay close to exponentially, although the convergence of the results for \( \delta_k^{\text{DE}}[m(\tau)] \) in Fig. 6(a) is not as good as for \( \delta_k^{\text{DE}}[n(\tau)] \) in Fig. 1(b). This is understandable because: (i) we are able to calculate one order lower for \( m_k(\tau) \) than for \( n(\tau) \), and (ii) \( m_k(\tau) \) probes correlations at all distances while \( n(\tau) \) is local, and is a thermodynamic quantity.

We fit \( \delta_k^{\text{DE}}[m(\tau)] \) to an exponential function \( \propto \exp[-\Gamma_{\text{NLCE}}^{\text{DE}}(g_1)\tau] \) to obtain the thermalization rates \( \Gamma_{\text{NLCE}}^{\text{DE}}(g_1) \) for the momentum distribution function. The fits are carried out in the interval \( \tau \in [2, 6] \) for NLCE-16 [shown as thin continuous lines in Fig. 6(a)] and in the interval \( \tau \in [2, 5] \) for NLCE-15. The rates \( \Gamma_{\text{NLCE}}^{\text{DE}}(g_1) \) are reported in Fig. 6(b) for \( g_1 \in [0.03, 0.12] \). In Fig. 6(b), we also plot the rates \( \Gamma_{\text{Fermi}}(g_1) \) obtained by evaluating Fermi’s golden rule [see Eqs. (53) and (54)] using full exact diagonalization in chains with \( L = 18 \) (Fermi-18) and periodic boundary conditions. The error bars show the standard deviation from averages over different choices of \( \Delta E \) and \( \tau \) (see Appendix D).
full exact diagonalization in chains with $L = 18$ sites and periodic boundary conditions (see Sec. VIA and Appendix D). [The rates $\Gamma^{\text{Fermi}}(g_1)$ were also reported in Fig. 1(c).] Figure 6(b) shows that, as advanced, the thermalization rates for the momentum distribution function are the same (within our computational errors) as the ones for the particle filling. A power law fit to the rates $\Gamma_{16}^{\text{NLCE}}(g_1)$ is also shown in Fig. 6(b). We find that $\Gamma_{16}^{\text{NLCE}}(g_1) \propto g_1^\beta$ with $\beta = 2.00$, in agreement with the numerical results in Sec. VI A, and with the analytical ones in Secs. III and IV.

VII. SUMMARY

We have put forward a conceptually simple scenario for prethermalization and thermalization in generic isolated quantum systems with a weakly broken conservation law. This scenario applies equally to noninteracting and strongly interacting integrable systems in which integrability is weakly broken, an example of the latter was recently shown numerically to exhibit a thermalization rate that is $O(g^2)$. [21], as well as to nonintegrable ones in which a conservation law is weakly broken. The weak perturbation allows the system to equilibrate to a state that is a (generalized) thermal equilibrium state of the unperturbed system. The properties of such a state are determined by the slowly changing value of the quasi-conserved quantity. The separation of time scales leads to a universal description, two aspects of which stand out:

(i) The dynamics of the quasi-conserved quantity is described by an autonomous equation that can be constructed from Fermi’s golden rule in unperturbed equilibrium ensembles. This equation is the generalization of the nonlinear Boltzmann equation appearing in weakly interacting quantum systems.

(ii) The deviation of observables in the instantaneous state from the prediction of the unperturbed equilibrium ensemble is described by first order perturbation theory. This generalizes the concept of ‘deformed GGE’ that was described in Ref. [11] for integrable systems in the presence of a weak integrability breaking perturbation.

Our theoretical results, as well as several special behaviors related to the initial state selected, properties of the perturbations that break the conservation law, and properties of the observables studied, were validated by our numerical calculations.

ACKNOWLEDGMENTS

M.R. and W.D.R. are thankful to Fabian Essler for many insightful discussions over the years about this topic. They motivated this work. We acknowledge support from the National Science Foundation, under Grant No. PHY-1707482 (K.M and M.R.), and in part under Grant No. PHY-1748958. The computations were carried out at the Institute for CyberScience at Penn State.

Appendix A: The Mori-Zwanzig formalism

To stress the algebra of the problem, we use a notation different from that in the main text. We consider a linear space $V$ with a projector $P : V \to V$, $P^2 = P$, and a linear evolution equation

$$\dot{a}(\tau) = Ma(\tau), \quad a(\tau) \in V, \quad M : V \to V. \quad (A1)$$

Then, we write $P_1 = P$, $P_2 = (1 - P)$, $a_i = P_i a$, and $M_{ij} = P_i M P_j$. As the projector is time-independent, the evolution equation can be cast as a system of two coupled equations:

$$\dot{a}_1(\tau) = M_{11} a_1(\tau) + M_{12} a_2(\tau), \quad (A2)$$

$$\dot{a}_2(\tau) = M_{21} a_1(\tau) + M_{22} a_2(\tau). \quad (A3)$$

We first formally solve the second of these equations

$$a_2(\tau) = e^{\tau M_{22}} a_2(0) + \int_0^\tau d\tau_1 e^{\tau_1 M_{22}} M_{21} a_1(\tau - \tau_1), \quad (A4)$$

and insert Eq. (A4) into Eq. (A2), yielding

$$\dot{a}_1(\tau) = M_{11} a_1(\tau) + M_{12} e^{\tau M_{22}} a_2(0)$$

$$+ \int_0^\tau d\tau_1 e^{\tau_1 M_{22}} M_{21} a_1(\tau - \tau_1). \quad (A5)$$

See Eqs. (7) and (12) in the main text. The relevance of this framework to irreversible phenomena was realized in Refs. [38, 39].

Appendix B: From $\mathcal{K}$ to the autonomous equation

Here we look into the superoperator $\mathcal{K}$ [see Eq. (20)],

$$\mathcal{K} = \int_0^{\pm \infty} ds \mathcal{L}_1 e^{s \mathcal{L}_0} \mathcal{L}_1 P, \quad (B1)$$

in more detail. Each $\mathcal{L}_1$ contains a commutator $-i q [\hat{V}, \cdot]$, so in total there are four terms, depending on whether $q \hat{V}$ in each $\mathcal{L}_1$-term acts on the left or on the right. We decompose

$$\mathcal{K} = \mathcal{K}_{\text{gain}} + \mathcal{K}_{\text{loss}}, \quad (B2)$$

where the ‘gain’ operator contains the two cases in which the $\hat{V}$’s act on different sides, and the ‘loss’ operator contains the two cases in which they act on the same side.

For both operators, we can recast the two cases into a single formula by extending the integration range of $s$ from $[0, \infty)$ to $(-\infty, \infty)$. One then has that

$$\mathcal{K}_{\text{gain}} \rho = g^2 \int_0^\infty ds \int dq d\rho \int_{-\infty}^\infty \rho \int_{-\infty}^\infty ds \text{Tr} \left[ \hat{P}_{\epsilon, q} \hat{V}_0(s) \mathcal{P} \rho \hat{V} \right], \quad (B3)$$
while for $K_{\text{loss}}$ acting on $\hat{\rho}$ one has
\[
K_{\text{loss}}\hat{\rho} = -g^2 \int dq d\epsilon q \rho_{q\epsilon} \int_{-\infty}^{\infty} ds \text{Tr} \left[ \hat{P}_{q\epsilon} \hat{V}_0(s) \hat{V} \hat{P} \right],
\]
where we abbreviated $\hat{P}_{q\epsilon} = P(\hat{H}_0 \approx \epsilon L) P(\hat{Q} \approx q L)$ (see Section III A), and we used the invariance of $\mathcal{P} \hat{\rho}$ and $\hat{P}_{q\epsilon}$ under the evolution generated by $\hat{H}_0$. Note that we can assume, without loss of generality, that $\hat{V}$ satisfies $\langle V \rangle_{q\epsilon, q} = 0$ for all $(\epsilon, q)$, because a term with nonzero mean would have canceled out in the commutators [alternatively, it would cancel between Eqs. (B3) and (B4)]. For this reason, we can replace all correlation functions below by truncated correlation functions.

Recall that we identified the space of $\mathcal{P} \hat{\rho}$ with distributions $p$ on $(\epsilon, q)$. Indeed, any $\mathcal{P} \hat{\rho}$ is of the form $\hat{\rho}_p$ with distribution $p$, see Eq. (3). Therefore, we can now abuse notation and interpret $K$ as a kernel on the space of distributions $p$. We then have
\[
K_{\text{loss}}(\epsilon' q') = \int d\epsilon dq \mathcal{K}(\epsilon, q; \epsilon', q') p(\epsilon, q),
\]
and, from the above formulas, we identify
\[
\mathcal{K}(\epsilon, q; \epsilon', q') = g^2 \int_{-\infty}^{\infty} ds \text{Tr} \left[ \hat{P}_{q\epsilon} \hat{V}_0(s) \hat{P}_{q', \epsilon'} \right],
\]
where $\hat{P}_{q}$ are spectral projections of $\hat{Q}$ and the sums runs over the eigenvalues of $\hat{Q}$. Note that the admissible values of $\delta Q$ are $O(1)$ because $\hat{V}$ is a sum of $O(1)$ local terms. For natural examples of $\hat{Q}$, such as the total particle number operator, we see that also $\hat{V}^{\delta Q}$ is a sum of local terms. In that case Eq. (B7) is manifestly of order $O(L)$, as a truncated correlation function in equilibrium. The first term in Eq. (B6) can be written as
\[
r_{q\epsilon, q'}(\delta Q) = g^2 \int_{-\infty}^{\infty} ds \text{Tr} \left[ \hat{V}^{\delta Q}(s) \rho_{q\epsilon} \hat{V}^{\delta Q} \right],
\]
with $\epsilon' = \epsilon$ and $q' = q = \delta Q / L$. Indeed, the condition $\epsilon' = \epsilon$ is enforced by the integral over $s$. The second term in Eq. (B6) has a contribution at $\epsilon = \epsilon'$, $q = q'$ only (by the cyclic property of the trace, the projector $\hat{P}_{q, q'}$ is put next to the density matrix $\rho_{q\epsilon, q'}$) and its value is
\[
\sum_{\delta Q} r_{q\epsilon, q'}(\delta Q).
\]
One could also have guessed this value because the process generated by $K$ conserves probability, which translates to
\[
\int dq d\epsilon q \mathcal{K}(\epsilon, q; \epsilon', q') = 0.
\]
We have now explicitly interpreted the process generated by $K$ as a jump process, with jump rates $r_{\epsilon}\rho_{q, q'}(\delta Q) \geq 0$ for jumps in the density $q$ of order $1 / L$. The link to the ‘drift’ computed in Sec. III F is by
\[
d(\epsilon, q) = \sum_{\delta Q} \frac{1}{L} r_{\epsilon, q}(\delta Q) \]
The expressions in Sec. III F are recovered by writing Eq. (B8) in terms of eigenkets of $\hat{H}_0$, and this eventually yields Eq. (24).

Appendix C: Convergence of NLCE and exact diagonalization

All the numerical results reported in the main text, but the relaxation rates $\Gamma_{\text{Fermi}}$ computed using Fermi’s golden rule and full exact diagonalization, were obtained using NLCE calculations. The basics of NLCEs was summarized in Sec. V C, and relevant parameters for the NLCE calculations (orders, largest Hilbert spaces involved, etc) were mentioned in Sec. V E. Here we discuss the convergence of the NLCE calculations and finite-size effects in the full exact diagonalization calculations.

All our full exact diagonalization calculations are carried out in chains with periodic boundary conditions. We use translational symmetry to block-diagonalize the Hamiltonian, which allows us to study larger chains than within the NLCE calculations. In the exact diagonalization calculations when $g_{\alpha} \neq 0$, in the absence of particle number conservation, the largest Hamiltonian sector diagonalized has 14602 states (for $L = 18$). When $g_{\alpha} = 0$, in the presence of particle number conservation, the largest Hamiltonian sector diagonalized has 9252 states (for $L = 20$). We only report exact diagonalization results for $g_{\alpha} = 0$ in Fig. 9(a). The matrices involved in our full exact diagonalization calculations are complex for sectors with total quasi-momentum $k \neq 0$ and $\pi$, and the results reported contain the contribution from all $L$ quasi-momentum sectors.

In Fig. 7, we show the evolution of the particle filling $n(\tau)$ for $g_{1} = 0.12$ (the fastest changing case studied) as obtained in the last four orders of the NLCE, and in the three largest chains solved using full exact diagonalization. In the scale of the figure, all the results are nearly indistinguishable from each other up to $\tau \approx 10$. Beyond that time, but not too far from it, the NLCE results can be seen to oscillate in the order of the expansion, even orders of the NLCE are above the odd orders (see the inset in Fig. 7). With increasing the order of the NLCE, one finds that the amplitude of the oscillation decreases and the results converge at longer times. This is similar to what happens in thermal equilibrium calculations, in which the NLCEs converge at lower temperatures as one increases the order of the expansion [49]. The exact diagonalization results, on the other hand, can be seen to approach the NLCE ones monotonically with increasing
the chain size. In Fig. 7, the $n(\tau)$ results in the last order of the NLCE and in the largest periodic chain diagonalized are nearly indistinguishable up to $\tau \approx 20$.

To gain a more quantitative understanding of the convergence of the particle filling $n_L(\tau)$ calculations within NLCE, where $l$ is the order of the expansion, and of finite-size effects in the exact diagonalization calculations of $n_L(\tau)$, where $L$ is the chain size, we compute the relative differences

$$\Delta^\text{NLCE}_l[n(\tau)] = \frac{|n_{l-1}(\tau) - n_l(\tau)|}{|n_l - 1/2|}$$  \hspace{1cm} (C1)

between the NLCE results at order $l$ and the last order calculated $l = 17$, and

$$\Delta^\text{ED}_L[n(\tau)] = \frac{|n_{L-18}(\tau) - n_L(\tau)|}{|n_L - 1/2|}$$  \hspace{1cm} (C2)

between the results for chains with $L$ sites and the largest chain $L = 18$ diagonalized. We also compute the relative differences between the last order of the NLCE and the largest periodic chain diagonalized

$$\Delta^\text{NLCE-ED}_l[n(\tau)] = \frac{|n_{l-17}(\tau) - n_{L-18}(\tau)|}{|n_L - 1/2|}.$$  \hspace{1cm} (C3)

In Eqs. (C1)-(C3), $n_I$ is the initial particle filling, which is obtained within machine precision at the 17th order of the NLCE, and 1/2 is the filling in the diagonal ensemble.

In Fig. 8, we plot the relative differences defined in Eqs. (C1) and (C2) for $l = 15$ and 16 of the NLCE, and for $L = 16$ and 17 in the exact diagonalization calculations, also for $g_1 = 0.12$ as in Fig. 7. Some points to be highlighted from the plots in Fig. 8 are: (i) The NLCE results for $l = 16$ are likely converged within machine precision up to $\tau \approx 2$, while the exact diagonalization ones for $L = 17$ are likely converged within machine precision only up to about one half of that time ($\tau \approx 1$). (ii) The relative differences between the various orders of the NLCE and between the various exact diagonalization calculations are below 0.01% for $\tau \lesssim 4$. (iii) At times $\tau \approx 16$, the relative differences are in all cases below 2% (they are smaller between the exact diagonalization results than between the NLCE ones).

The results for the relative difference $\Delta^\text{NLCE-ED}_l[n(\tau)]$ are qualitatively similar to those for $\Delta^\text{NLCE}_l[n(\tau)]$ and $\Delta^\text{ED}_L[n(\tau)]$. We find that $\Delta^\text{NLCE-ED}_l[n(\tau)] \leq 0.01\%$ for times $\tau \lesssim 5$. We use times $\tau \leq 5$ for the exact diagonalization calculation of the rates from Fermi’s golden rule in Appendix D. We also find that $\Delta^\text{NLCE-ED}_l[n(\tau)] \lesssim 0.5\%$ for times $\tau \lesssim 16$. We use times $\tau \leq 16$ in the fits to obtain the rates from the 17th order of the NLCE dynamics in Sec. VI B.

Next, for the nearest neighbor correlations $k(\tau)$, we discuss the convergence of the NLCE calculations $k_l(\tau)$, and finite-size effects in the full exact diagonalization calculations $k_L(\tau)$. $k(\tau)$, being a correlation function, is more challenging to obtain accurately than $n(\tau)$, which is a thermodynamic quantity.

In Fig. 9(a), we show dynamics after the quench $\hat{H}_f \rightarrow H_0$ for $k_l(\tau)$ with $l = 17, 18$ and 19 (the latter two are also reported in Fig. 2) and for $k_L(\tau)$ with $L = 18, 19$ and 20. $k_l(\tau)$ and $k_L(\tau)$ are almost indistinguishable from each other up to times $\tau \approx 5$ (the earliest times are not shown to gain dynamical range in the $y$ axis). For $\tau \gtrsim 5$, the NLCE and exact diagonalization results split and each equilibrate to the prediction of the corresponding diagonal ensemble, $k^{\text{DE}}_{l=19}$ (DENLCE-19) and $k^{\text{DE}}_{L=20}$ (DEED-20), respectively. As in Ref. [21], the NLCE results for the diagonal ensemble (DENLCE-19) are closer to the grand canonical ensemble ones $s^{\text{DE}}_{l=19}$ (GENLCE-19).
FIG. 9. Time evolution of the nearest neighbor hopping $k(τ)$ obtained in the last three orders $l$ of the NLCE (NLCE-$l$ in the legends), and in the three largest chains with $L$ sites and periodic boundary conditions diagonalized (ED-$L$ in the legends). Results for the diagonal ensemble are shown for the last order of the NLCE (DE$_{\text{NLCE-}17}$) and the largest chain diagonalized (DE$_{\text{ED-}20}$), while results for the grand canonical ensemble are shown for the last order of the NLCE (GE$_{\text{NLCE-}17}$), which are converged to machine precision. (a) Quenches $\hat{H}_1 \rightarrow \hat{H}_0$, for $l = 17$, 18, and 19 (NLCE-$17$, NLCE-$18$, and NLCE-$19$), and for $L = 18$, 19, and 20 (ED-$18$, ED-$19$, and ED-$20$), along with the diagonal ensembles for $l = 19$ (DE$_{\text{NLCE-}19}$) and $L = 20$ (DE$_{\text{ED-}20}$), and the grand canonical ensemble for $l = 19$ (GE$_{\text{NLCE-}19}$). (b) Quenches $\hat{H}_1 \rightarrow \hat{H}_1$, in which $g_1 = 0.12$, for $l = 15$, 16, 17 (NLCE-$15$, NLCE-$16$, NLCE-$17$), and for $L = 17$, 18, 19 (ED-$17$, ED-$18$, and ED-$19$), along with the diagonal ensembles for $l = 17$ (DE$_{\text{NLCE-}17}$) and $L = 18$ (DE$_{\text{ED-}18}$), and the grand canonical ensemble for $l = 17$ (GE$_{\text{NLCE-}17}$). Higher orders in the NLCE, and larger chains in exact diagonalization, are calculated in (a) than in (b) thanks to particle number conservation in the former.

than the exact diagonalization results for the diagonal ensemble (DE$_{\text{ED-}20}$). The differences between them are due to lack of convergence of the diagonal ensemble within NLCEs and finite-size effects for the diagonal ensemble within exact diagonalization. The NLCE results for the grand canonical ensemble are converged within machine precision. Consequently, at long times, the NLCE is expected to be more accurate than exact diagonalization. At intermediate times $5 ≤ τ ≤ 15$, $k_{l=16}(τ)$ and $k_{l=17}(τ)$ are almost indistinguishable from each other, while $k_{l=17}(τ)$ shifts upward with increasing $L$, toward the NLCE predictions. This suggests that NLCE is also more accurate than exact diagonalization at intermediate times.

In Fig. 9(b), we show dynamics after the quench $\hat{H}_1 \rightarrow \hat{H}_1$ ($g_1 = 0.12$) for $k_l(τ)$ with $l = 15$, 16, and 17 (the latter two are also reported in Fig. 3) and for $k_{L}(τ)$ with $L = 16$, 17, and 18. $k_l(τ)$ and $k_{L}(τ)$ are almost indistinguishable from each other up to times $τ ≈ 4$ (again, the earliest times are not shown to gain dynamical range in the $y$ axis). For $τ ≥ 4$, the NLCE and exact diagonalization results again split, as for the quench $\hat{H}_1 \rightarrow \hat{H}_0$, despite the fact that both equilibrate to diagonal ensemble results that are very close to each other. $k_{l=17}^{\text{DE}}$ (DE$_{\text{NLCE-}17}$) and $k_{l=18}^{\text{DE}}$ (DE$_{\text{ED-}18}$) in Fig. 9(b) are almost indistinguishable from each other and from the grand canonical ensemble result $k_{l=17}^{\text{GE}}$ (GE$_{\text{NLCE-}17}$). At times $5 ≤ τ ≤ 15$, $k_{l=16}(τ)$ and $k_{l=17}(τ)$ are almost indistinguishable, while $k_{l}(τ)$ shifts upward toward the NLCE predictions with increasing $L$. As for the quenches $\hat{H}_1 \rightarrow \hat{H}_0$, these results suggest that NLCE is more accurate than exact diagonalization at intermediate times in quenches $\hat{H}_1 \rightarrow \hat{H}_1$.

The discrepancy between the NLCE and exact diagonalization results at intermediate times after quenches $\hat{H}_1 \rightarrow \hat{H}_1$ is likely a manifestation of finite-size effects in $k_{l}(τ)$ that result from the “prethermal” dynamics seen in Fig. 9(a), which equilibrates to a diagonal ensemble value ($k_{l=20}^{\text{DE}}$) that is lower than the expected grand canonical ensemble one. The results in Fig. 9(b) also make apparent that, because of finite-size effects, the thermalization rates for $k(τ)$ are smaller in exact diagonalization, and increasing with increasing $L$, than in NLCE.

Appendix D: Thermalization rates from Fermi’s golden rule

We use full exact diagonalization in chains with $L$ sites and periodic boundary conditions to evaluate Eq. (53). For the numerical calculation, the delta function is replaced by a coarse-graining procedure leading to the following modified version of Eq. (53)

$$f_{\Delta E}(τ) = \frac{2πg_1^2}{LΔE} \sum_i P_i(τ) \times \sum_{|E_i − E_j| ≤ ΔE/2} \left| \langle E_j^0 | \hat{V}_0 | E_i^0 \rangle \right|^2 (N_j − N_i),$$

where $|E_i^0\rangle$ ($|E_j^0\rangle$) are the eigenkets of $\hat{H}_0$ with energy $E_i^0$ ($E_j^0$), $N_i = \langle E_i^0 | \hat{N} | E_i^0 \rangle$, and $P_i(τ) = \langle E_0^0 | \hat{ρ}(τ) | E_i^0 \rangle$.

In Fig. 10, we plot $f_{\Delta E}(τ)$, evaluated in chains with $L = 17$ and $L = 18$ for quenches $\hat{H}_1 \rightarrow \hat{H}_1$ [see Eq. (36)], as a function of $ΔE$. We show results at two times, $τ = 1$
FIG. 10. $f_{\Delta E}(\tau)$, see Eq. (D1), evaluated for a wide range of energy windows $\Delta E$ for quenches $\hat{H}_i \rightarrow \hat{H}_1$ [see Eq. (36)] with $g_1 = 0.03, 0.06,$ and $0.12$. $f_{\Delta E}(\tau)$ is computed using full exact diagonalization of chains with $L = 18$ (see legends) and $L = 17$ (dotted lines) and periodic boundary conditions. The vertical lines delimit the range $\Delta E / L \in [0.03, 0.09]$ in which $f_{\Delta E}(\tau)$ is approximately constant.

in Fig. 10(a) and $\tau = 5$ in Fig. 10(b), and for three values of $g_1$ at each time. Our main finding in Fig. 10 is that, in the interval $\Delta E / L \in [0.03, 0.09]$, $f_{\Delta E}(\tau)$ is nearly independent of $\Delta E$, and is approximately the same in chains with $L = 17$ and $L = 18$. Similar results were obtained at times $\tau \in [0, 5]$, for which we showed in Appendix C that our best NLCE and exact diagonalization results for $n(\tau)$ differ by less than a 0.01%.

Having identified an appropriate range of values of $\Delta E$, we compute the rate

$$\Gamma_{\Delta E, \tau}(g_1) = \frac{f_{\Delta E}(\tau)}{n(\tau) - n_{DE}},$$

which is defined following Eq. (54).

Figure 11 shows $\Gamma_{\Delta E, \tau}(g_1)$ vs $\tau$ for three values of $g_1$, and for three values of $\Delta E / L$ for each value of $g_1$. The rates for each value of $g_1$ decrease slowly with increasing $\tau$, and are very close to each other for the three values of $\Delta E / L$ shown. Given the results for $n(\tau)$ at times $\tau \leq 1$ in Fig. 9, which exhibit a plateau-like behavior discussed in Sec. VIA, the rates $\Gamma_{\text{Fermi}}(g_1)$ for $\tau \leq 1$ are not meaningful. Hence, the rates $\Gamma_{\text{Fermi}}(g_1)$ reported in Sec. VIA were obtained averaging $\Gamma_{\Delta E, \tau}(g_1)$ over the results for $\tau = 1, 1.5, \ldots, 5$ (9 values), and for $\Delta E / L = 0.03, 0.032, \ldots, 0.09$ (31 values), for a total of 279 values entering each average. In Fig. 11, we report the averages as horizontal lines. For each average, we also compute the standard deviation. In Fig. 1(c), we report the averages, and the standard deviations (as errorbars), for different values of $g_1$ and for $L = 17$ and $L = 18$. 

![Diagram](image_url)

FIG. 11. $\Gamma_{\Delta E, \tau}(g_1)$ [see Eq. (D2)] evaluated for quenches $\hat{H}_i \rightarrow \hat{H}_1$ [see Eq. (36)] with $g_1 = 0.03, 0.06,$ and $0.12$. The results reported are from full exact diagonalization of chains with $L = 18$ and periodic boundary conditions. The average thermalization rate (horizontal line) is obtained by averaging $\Gamma_{\Delta E, \tau}(g_1)$ over $\tau = 1, 1.5, \ldots, 5$ (9 values) and over $\Delta E / L = 0.03, 0.032, \ldots, 0.09$ (31 values), for a total of 279 values of $\Gamma_{\Delta E, \tau}(g_1)$ entering each average reported.

[1] J. Berges, S. Borsányi, and C. Wetterich, Phys. Rev. Lett. 93, 142002 (2004).
[2] T. Kinoshita, T. Wenger, and D. S. Weiss, Nature 440, 900 (2006).
[3] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. A. Smith, E. Demler, and J. Schmiedmayer, Science 337, 1318 (2012).
[4] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, Science 348, 207 (2015).
[5] T. Langen, T. Gasenzer, and J. Schmiedmayer, J. Stat. Mech., 064009 (2016).
[6] Y. Tang, W. Kao, K.-Y. Li, S. Seo, K. Mallayya, M. Rigol, S. Gopalakrishnan, and B. L. Lev, Phys. Rev. X 8, 021030 (2018).
[7] M. Moeckel and S. Kehrein, Phys. Rev. Lett. 100, 175702 (2008); Ann. Phys. 324, 2146 (2009).
[8] M. Eckstein, M. Kollar, and P. Werner, Phys. Rev. Lett. 103, 056403 (2009).
[9] M. Kollar, F. A. Wolf, and M. Eckstein, Phys. Rev. B 84, 054304 (2011).
[10] N. Nessi, A. Iucci, and M. A. Cazalilla, Phys. Rev. Lett. 113, 210402 (2014).
[11] F. H. L. Essler, S. Kehrein, S. R. Manmana, and N. J.
[12] B. Bertini, F. H. L. Essler, S. Groha, and N. J. Robinson, Phys. Rev. Lett. 115, 180601 (2015); Phys. Rev. B 94, 245117 (2016).
[13] M. Fagotti and M. Collura, arXiv:1507.02678.
[14] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, Phys. Rev. Lett. 98, 050405 (2007).
[15] L. Vidmar and M. Rigol, J. Stat. Mech., 064007 (2016).
[16] F. H. L. Essler and M. Fagotti, J. Stat. Mech., 064002 (2016).
[17] M. A. Cazalilla and M.-C. Chung, J. Stat. Mech., 064004 (2016).
[18] J.-S. Caux, J. Stat. Mech., 064006 (2016).
[19] M. Stark and M. Kollar, arXiv:1308.1610.
[20] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Adv. Phys. 65, 239 (2016).
[21] K. Mallayya and M. Rigol, Phys. Rev. Lett. 120, 070603 (2018).
[22] A. Lazarides, A. Das, and R. Moessner, Phys. Rev. Lett. 112, 150401 (2014).
[23] E. Canovi, M. Kollar, and M. Eckstein, Phys. Rev. E 93, 012130 (2016).
[24] L. D’Alessio and M. Rigol, Phys. Rev. X 4, 041048 (2014).
[25] A. Lazarides, A. Das, and R. Moessner, Phys. Rev. E 90, 012110 (2014).
[26] M. Bukov, L. D’Alessio, and A. Polkovnikov, Adv. Phys. 64, 139 (2015).
[27] D. A. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, Phys. Rev. B 95, 014112 (2017).
[28] D. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, Commun. Math. Phys. 354, 809 (2017).
[29] T. Mori, T. Kuwahara, and K. Saito, Phys. Rev. Lett. 116, 120401 (2016).
[30] Z. Lenarčič, F. Lange, and A. Rosch, Phys. Rev. B 97, 024302 (2018).
[31] F. Lange, Z. Lenarčič, and A. Rosch, Phys. Rev. B 97, 165138 (2018).
[32] M. Rigol, V. Dunjko, and M. Olshanii, Nature 452, 854 (2008).
[33] J. M. Deutsch, Phys. Rev. A 43, 2046 (1991).
[34] M. Srednicki, J. Stat. Phys. 33, 783 (1985).
[35] A. C. Cassidy, C. W. Clark, and M. Rigol, Phys. Rev. Lett. 106, 140405 (2011).
[36] J.-S. Caux and F. H. L. Essler, Phys. Rev. Lett. 110, 257203 (2013).
[37] T. Ishii and T. Mori, arXiv:1810.00625.
[38] R. Zwanzig, J. Chem. Phys. 33, 1338 (1960).
[39] H. Mori, Prog. Theor. Phys. 33, 423 (1965).
[40] J. Lux, J. Müller, A. Mitra, and A. Rosch, Phys. Rev. A 89, 053608 (2014).
[41] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, Phys. Rev. X 8, 021069 (2018).
[42] M. Rigol, Phys. Rev. Lett. 103, 100403 (2009).
[43] L. F. Santos and M. Rigol, Phys. Rev. E 81, 036206 (2010).
[44] M. A. Cazalilla, R. Citro, T. Giamarchi, E. Orignac, and M. Rigol, Rev. Mod. Phys. 83, 1405 (2011).
[45] I. G. White, B. Sundar, and K. R. A. Hazzard, arXiv:1710.07696.
[46] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, Phys. Rev. X 8, 021069 (2018).
[47] M. Rigol, Phys. Rev. Lett. 97, 187202 (2006); Phys. Rev. E 75, 061118 (2007).
[48] M. Rigol, Phys. Rev. Lett. 112, 170601 (2014); Phys. Rev. Lett. 116, 100601 (2016).
[49] K. Mallayya and M. Rigol, Phys. Rev. E 95, 033302 (2017).
[50] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. 80, 885 (2008).