Prethermalization and Thermalization in Isolated Quantum Systems

Krishnanand Mallayya,[1,2] Marcos Rigol,[1,2] and Wojciech De Roeck[3,2]

[1]*Department of Physics, Pennsylvania State University, University Park, Pennsylvania 16802, USA*
[2]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA*
[3]*Instituut voor Theoretische Fysica, KULeuven, 3001 Leuven, Belgium*

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.

Subject Areas: Condensed Matter Physics,
Quantum Physics,
Statistical Physics

I. INTRODUCTION

Prethermalization [1] has emerged over the past 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 is quenches in isolated noninteracting (integrable) systems, in which interactions (integrability-breaking perturbations) of strength g are turned on [7–15]. By dephasing, observables quickly settle to quasisteady states that are well described by generalized Gibbs ensembles (GGEs) [16–20] of the noninteracting systems. The observables then relax to the thermal equilibrium values (thermalize) in a much longer timescale $\propto g^{-2}$, as predicted by kinetic Boltzmann-like equations. The latter can be derived applying time-dependent perturbation theory to the GGEs [21,22] 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 $\propto g^2$ [23]. A second example is isolated weakly interacting driven systems, for which GGEs and kinetic Boltzmann-like equations are also relevant [24,25]. A third example is 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 [26,27]) state in a timescale that is exponentially long in the frequency of the drive [28–33].

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 dynamics. A similar point of view was put forward in Refs. [34,35] 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 paper 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 and discussion of our results is presented in Sec. VII, while the Appendixes 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 includes the former plus a weak perturbation,

$$\hat{H} = \hat{H}_0 + g\hat{V}, \tag{1}$$

where g is 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: $\hat{h}_0 = \hat{H}_0/L$, $\hat{q} = \hat{Q}/L$, $\hat{h} = \hat{H}/L$, etc. In our numerical calculations, $\hat{Q}$ ($\hat{q}$) is the total number of particles (the filling), or, in the spin language, the total magnetization (the site magnetization).

Assumption of fast equilibration.—We consider cases in which the dynamics generated by $\hat{H}_0$, 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 τ^* to the predictions of an ensemble (e.g., microcanonical) of statistical mechanics $\hat{\rho}_{e_0,q}$, with $(e_0, q) = (\langle \hat{h}_0 \rangle_{\hat{\rho}_I}, \langle \hat{q} \rangle_{\hat{\rho}_I})$. 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. [36] 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 $\hat{H}_0$, it is reasonable to assume that the effect of the perturbation is small at times $\tau \ll 1/g$, in the sense that, for such 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 $\hat{\rho}_{e_0,q}$. The latter state can be very different from the thermal equilibrium ensemble $\hat{\rho}_e$ associated to $\hat{H}$. $\hat{\rho}_{e_0,q}$ 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 $\hat{H}_0$ and not that of $\hat{H}$, but that difference is only $\mathcal{O}(g)$].

Thermalization rate.—To determine how fast the true equilibrium is approached, let us start from $\hat{\rho}_{e_0,q}$ and look at the change in $\langle \hat{Q}(\tau) \rangle_{e_0,q}$. We do this perturbatively:

$$\langle \hat{Q}(\tau) \rangle_{e_0,q} = \langle \hat{Q} \rangle_{e_0,q} + ig\tau \langle [\hat{V}, \hat{Q}] \rangle_{e_0,q} + \mathcal{O}(g^2 \tau \tau^* L), \tag{2}$$

where we use that $[\hat{H}_0, \hat{Q}] = [\hat{H}_0, \hat{\rho}_{e_0,q}] = 0$. The precise error estimate $\mathcal{O}(\cdot)$ will only be argued for later. The important point to be highlighted from Eq. (2) is that the leading order correction to $\langle \hat{Q} \rangle_{e_0,q}$ vanishes because $\langle [\hat{V}, \hat{Q}] \rangle_{e_0,q} = 0$, by the cyclic property of the trace and the fact that $[\hat{Q}, \hat{\rho}_{e_0,q}] = 0$. 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 $\hat{\rho}_{e_0,q}$.

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. (23), 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 $\hat{\rho}$ 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 p can be lifted to a $\hat{\rho}$ by

$$\hat{\rho}_p = \int de_0 dq\, p(e_0, q)\hat{\rho}_{e_0,q}, \tag{3}$$

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

$$p_{\hat{\rho}}(e_0, q) \propto \mathrm{Tr}[\hat{\rho} P(\hat{H}_0 \approx e_0 L)P(\hat{Q} \approx qL)], \tag{4}$$

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 δ_E that is much larger than the level spacing, but smaller than any relevant energy scale (like, e.g., the energy per site). For $P(\hat{Q} \approx qL)$, we can simply take the projection of $\hat{Q}$ on the eigenvalue closest to qL. Then we can set

$$\mathcal{P}\hat{\rho} = \int de_0 dq\, p_{\hat{\rho}}(e_0, q)\hat{\rho}_{e_0,q}, \tag{5}$$

where $\int de_0 dq$ represents a sum with the aforementioned resolution. We define $\bar{\mathcal{P}} = 1 - \mathcal{P}$, and note that both $\mathcal{P}$, $\bar{\mathcal{P}}$ are projectors, namely, $\mathcal{P}^2 = \mathcal{P}$ and $\bar{\mathcal{P}}^2 = \bar{\mathcal{P}}$. A warning is in order, despite the fact that $\mathcal{P}\hat{\rho}$ and $\bar{\mathcal{P}}\hat{\rho}$ 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 [36]), 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 $p(e_0, q)$.

The above definition of the projection $\mathcal{P}$ is motivated by ensembles of statistical mechanics, which are expected to describe many-body quantum systems after equilibration. An equivalent definition can be motivated purely from quantum mechanics. If one takes $\hat{\rho}$ and evolves it under the unitary evolution dictated by $\hat{H}_0$, observables equilibrate to the predictions of the so-called diagonal ensemble $\hat{\rho}_{\mathrm{DE}}$ [22]:

$$\hat{\rho}_{\mathrm{DE}} = \sum_{E^0} (\langle E^0 | \hat{\rho} | E^0 \rangle) | E^0 \rangle \langle E^0 |, \tag{6}$$

where $|E^0\rangle$ are the eigenkets $\hat{H}_0$ 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 unimportant. This is generically the case in interacting many-body quantum systems [22].

The map $\hat{\rho} \to \hat{\rho}_{\mathrm{DE}}$ is a projection as well; let us call it $\mathcal{P}_{\mathrm{DE}}$. It provides an alternative path to identifying the slow variables. This is the case because, in recent years, we have come to realize that the ensembles defined by $\mathcal{P}\hat{\rho}$ and $\mathcal{P}_{\mathrm{DE}}\hat{\rho}$ are equivalent when it comes to their predictions for observables (few-body or local operators) in large systems. This is a consequence of eigenstate thermalization for nonintegrable (quantum chaotic) systems [36–38], and generalized eigenstate thermalization for integrable systems [17,20,39–41]. Hence, it does not really matter whether one uses the projection $\mathcal{P}$ or $\mathcal{P}_{\mathrm{DE}}$, so we do not actively distinguish between them in most of our analytical derivations. In our numerical calculations we use $\mathcal{P}_{\mathrm{DE}}$, as $\hat{\rho}_{\mathrm{DE}}$ can be calculated exactly in the same way for nonintegrable and integrable systems. $\mathcal{P}$, on the other hand, demands that one identifies beforehand the relevant conserved quantities (not necessary for $\mathcal{P}_{\mathrm{DE}}$). Keeping all this is mind, in what follows we do not actively 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.

B. Mori-Zwanzig approach

Let us introduce the Liouville superoperator $\mathcal{L} = -i[\hat{H}, \cdot]$. Then, following Mori-Zwanzig [42,43], see Appendix A, the theory of linear ordinary differential equations gives us the following rewriting of the $\mathcal{P}$-projected Liouville equation

$\partial_\tau \hat{\rho}(\tau) = \mathcal{L}\hat{\rho}(\tau)$:

$$\partial_\tau \mathcal{P}\hat{\rho}(\tau) = \mathcal{P}\mathcal{L}\mathcal{P}\hat{\rho}(\tau)$$
$$+ \int_0^\tau ds\mathcal{P}\mathcal{L}e^{s\bar{\mathcal{P}}\mathcal{L}}\mathcal{L}\mathcal{P}\hat{\rho}(\tau - s) + \mathcal{P}\mathcal{L}e^{\tau\bar{\mathcal{P}}\mathcal{L}}\bar{\mathcal{P}}\hat{\rho}_I, \tag{7}$$

where $\hat{\rho}_I \equiv \hat{\rho}(\tau = 0)$, and we use that the projector $\mathcal{P}$ is time independent. To bring some structure to this equation, we now split

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1, \quad \text{with} \quad \mathcal{L}_0 = -i[\hat{H}_0, \cdot], \mathcal{L}_1 = -ig[\hat{V}, \cdot], \tag{8}$$

and note the properties

$$\mathcal{L}_0\mathcal{P} = \mathcal{P}\mathcal{L}_0 = 0, \qquad \mathcal{P}\mathcal{L}_1\mathcal{P} = 0. \tag{9}$$

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

$$\partial_\tau \mathcal{P}\hat{\rho}(\tau) = \int_0^\tau ds\mathcal{A}_s\mathcal{P}\hat{\rho}(\tau - s) + \mathcal{Y}_\tau\bar{\mathcal{P}}\hat{\rho}_I, \tag{10}$$

where we introduce

$$\mathcal{A}_s = \mathcal{P}\mathcal{L}_1 e^{s(\mathcal{L}_0 + \bar{\mathcal{P}}\mathcal{L}_1)}\mathcal{L}_1\mathcal{P} \quad \text{and} \quad \mathcal{Y}_\tau = \mathcal{P}\mathcal{L}_1 e^{\tau(\mathcal{L}_0 + \bar{\mathcal{P}}\mathcal{L}_1)}\bar{\mathcal{P}}. \tag{11}$$

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

$$\bar{\mathcal{P}}\hat{\rho}(\tau) = e^{\tau\bar{\mathcal{P}}\mathcal{L}}\bar{\mathcal{P}}\hat{\rho}_I + \int_0^\tau ds\bar{\mathcal{P}}e^{s\bar{\mathcal{P}}\mathcal{L}}\mathcal{L}\mathcal{P}\hat{\rho}(\tau - s). \tag{12}$$

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 $\hat{H}_0$, $E_0 = \mathrm{Tr}(\hat{\rho}_I\hat{H}_0)$, and of $\hat{Q}$, $Q = \mathrm{Tr}(\hat{\rho}_I\hat{Q})$. 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 [22]), from integrable reference Hamiltonians, for which observables equilibrate to the GGE predictions (exhibit generalized

thermalization [17]). 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 $\hat{\rho}_I$,

$$\frac{|\text{Tr}[\hat{\rho}_I \hat{O}_0(\tau)] - \text{Tr}[\hat{\rho}_{e_0,q}\hat{O}]|}{|\text{Tr}[\hat{\rho}_{e_0,q}\hat{O}]|} =: f(\tau) \to 0, \qquad (13)$$

where $\hat{O}_0(\tau) = e^{i\tau\hat{H}_0}\hat{O}e^{-i\tau\hat{H}_0}$. We define an equilibration timescale $\tau^*(\epsilon)$, which is understood as the time τ at which $f(\tau) \approx \epsilon$ for some small dimensionless ϵ. By fast equilibration, we mean that $\tau^*(\epsilon) \sim \mathcal{O}(1)$ in the relevant time units of the problem. In Sec. VI, numerical simulations of a translationally invariant one-dimensional Hamiltonian indeed show that, in times that are $\mathcal{O}(1)$ in the relevant timescale, observables reach values that are very close to the predictions of $\hat{\rho}_{e_0,q}$.

We do not expect ϵ to be arbitrarily small because, even if the Hamiltonian and the initial state are translational 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. [44] on the basis of fluctuating hydrodynamics. In such a case, one expects

$$\tau^*(\epsilon) \approx \epsilon^{-2/d}. \qquad (14)$$

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

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

$$e^{\tau\mathcal{L}_0}\hat{\rho} \approx \mathcal{P}e^{\tau\mathcal{L}_0}\hat{\rho}, \quad \text{whenever } \tau > \tau^*, \qquad (15)$$

accepting an error $\mathcal{O}(\epsilon)$. This approximation amounts to assuming that the system has equilibrated after a time τ^* with respect to the dynamics generated by $\hat{H}_0$. Hence, at time τ^*, we replace the density matrix of the system by $\hat{\rho}_{e_0,q}$. 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 $\hat{H}_0$ [22].

D. Born approximation

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

$$g\tau^* \ll 1. \qquad (16)$$

To see this assumption at work, let us expand the exponential $e^{s(\mathcal{L}_0 + \bar{\mathcal{P}}\mathcal{L}_1)}$ in the definition of $\mathcal{A}_s$ in Eq. (11):

$$\mathcal{A}_s = \mathcal{P}\mathcal{L}_1 e^{s\mathcal{L}_0}\mathcal{L}_1\mathcal{P}$$
$$+ \int_0^s ds_1 \mathcal{P}\mathcal{L}_1 e^{(s-s_1)\mathcal{L}_0}\bar{\mathcal{P}}\mathcal{L}_1 e^{s_1\mathcal{L}_0}\mathcal{L}_1\mathcal{P} + \cdots. \qquad (17)$$

Note that, for example, the first term on the rhs could also be written as $\mathcal{P}\mathcal{L}_1\bar{\mathcal{P}}e^{s\mathcal{L}_0}\bar{\mathcal{P}}\mathcal{L}_1\mathcal{P}$ or even $\mathcal{P}\mathcal{L}_1\bar{\mathcal{P}}e^{s\bar{\mathcal{P}}\mathcal{L}_0\bar{\mathcal{P}}}\bar{\mathcal{P}}\mathcal{L}_1\mathcal{P}$, which makes apparent that the intermediate evolution acts on the fast degrees of freedom. 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 $\mathcal{L}_1$ carries a factor L, it appears that the first and second terms 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 $\langle \hat{V}_0(s)\hat{V}\rangle^c_{e_0,q}$ and $\langle \hat{V}_0(s)\hat{V}_0(s_1)\hat{V}\rangle^c_{e_0,q}$; 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 [45,46]. Assuming that one recovers the linearity in L to all orders, for any $\tau \geq \tau^*$,

$$\int_0^\tau ds\mathcal{A}_s = \int_0^\tau ds\mathcal{P}\mathcal{L}_1 e^{s\mathcal{L}_0}\mathcal{L}_1\mathcal{P} + \mathcal{O}(g^3\tau^{*2}L), \qquad (18)$$

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

$$g^3\tau^{*2}L \ll g^2\tau^* L, \qquad (19)$$

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,

$$\mathcal{K} = \int_0^\infty ds\mathcal{P}\mathcal{L}_1 e^{s\mathcal{L}_0}\mathcal{L}_1\mathcal{P}, \qquad (20)$$

is $\mathcal{O}(g^2\tau^*L)$, with the factor L originating from the spatial sum in $\mathcal{L}_1$. 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 $\mathcal{O}(g^2\tau^{*2})$

(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 equation of motion [Eq. (10)] reads

$$\partial_\tau \mathcal{P}\hat{\rho}(\tau) = \mathcal{K}\mathcal{P}\hat{\rho}(\tau), \qquad \tau \gg \tau^*. \qquad (21)$$

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 $\tau \gg \tau_{\text{tr}}$ (where the subindex tr stands for transient) and τ_{tr} is determined by $g^2\tau^* \gg gf(\tau_{\text{tr}})$, as follows by comparing the magnitude of the two terms in Eq. (10).

F. 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 $\hat{\rho}_{e_0,q}$, 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. Having a single eigenket $|E^0\rangle$ in mind, the rate of change of q, called the "drift" d, is given by

$$d(E^0) = 2\pi g^2 \int d\bar{E}^0 \delta(E^0 - \bar{E}^0) \frac{Q_{\bar{E}^0} - Q_{E^0}}{L} |\langle E^0|\hat{V}|\bar{E}^0\rangle|^2, \qquad (22)$$

where $Q_{E^0} = \langle E^0|\hat{Q}|E^0\rangle$, the integral is understood as a sum, and the delta function $\delta(E^0 - \bar{E}^0)$ selects an interval of energies with a width smaller than any relevant energy scale but much larger than the level spacing. In principle, one should average $d(E^0)$ over E^0, with the distribution provided by statistical mechanics $\hat{\rho}_{e_0,q}$, or by quantum mechanics, i.e., the diagonal ensemble $\hat{\rho}_{\text{DE}}$. However, once again, because of eigenstate thermalization (or its generalized version for integrable systems), we expect the average to be unnecessary for large L. Hence, the drift $d = d(e_0, q)$ does not depend on $|E^0\rangle$ 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 then obtain the autonomous equation

$$\partial_\tau q(\tau) = d[e_0(\tau), q(\tau)], \qquad \partial_\tau e_0(\tau) = 0. \qquad (23)$$

The corresponding evolution equation for the distributions $p_\tau(e_0, q)$ is then

$$\partial_\tau p_\tau(e_0, q) = -d[e_0, q]\partial_q p_\tau(e_0, q), \qquad (24)$$

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

$$d[e_0, q^*(e_0)] = 0. \qquad (25)$$

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:

$$\hat{\rho}_{e_0,q^*} \approx \hat{\rho}_{e_0}. \qquad (26)$$

This also shows that the asymptotic state $\hat{\rho}_{e_0,q^*}$ is close to the global equilibrium state $\hat{\rho}_e$ for which $\langle\hat{h}\rangle = e = e_0 + \mathcal{O}(g)$. 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 one to compute higher-order corrections in g to the autonomous Eqs. (21) and (23). In particular, one can expand $\mathcal{A}_s$ in a power series in terms of order $(g\tau^*)^n L$, $n \geq 1$, with the terms for $n = 1$ and 2 given in Sec. III D. Integrating over s, one then gets a series for $\mathcal{K}$, 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 $\mathcal{A}_s$, but also, e.g., replacing $\mathcal{P}\hat{\rho}(\tau - s)$ for $s \leq \tau^*$ by $\mathcal{P}\hat{\rho}(\tau)$, and (ii) a not-so-fast decay of $f(\tau)$ would imply that the $\mathcal{Y}_\tau$ 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 [45–47] implies that one can meaningfully compute corrections, but only if the condition $\int_0^\infty d\tau|f(\tau)| < \infty$ is satisfied.

IV. DYNAMICS OF LOCAL OBSERVABLES

So far, we have focused on the evolution of the quasiconserved 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 $\mathcal{O}(g)$ difference from $\hat{\rho}_{e_0,q(\tau)}$. This deviation has a universal form.

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

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 $\hat{\rho}_{e_0,q(\tau)}$, leading to the prediction

$$\mathrm{Tr}[\hat{\rho}(\tau)\hat{O}] \approx \mathrm{Tr}[\hat{O}\mathcal{P}\hat{\rho}(\tau)] = \langle\hat{O}\rangle_{e_0,q(\tau)}. \tag{27}$$

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 $\mathcal{P}\hat{\rho}(\tau)$, 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)]

$$\hat{\rho}(\tau) = \mathcal{P}\hat{\rho}(\tau) + \bar{\mathcal{P}}\hat{\rho}(\tau)$$

$$= \mathcal{P}\hat{\rho}(\tau) + e^{\tau\bar{\mathcal{P}}\mathcal{L}}\bar{\mathcal{P}}\hat{\rho}_I + \int_0^\tau ds\,\bar{\mathcal{P}}e^{s\bar{\mathcal{P}}\mathcal{L}}\mathcal{L}\mathcal{P}\hat{\rho}(\tau-s). \tag{28}$$

The first term in Eq. (28) is dominant for large $\tau \geq \tau_{\mathrm{tr}}$. 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, $\int_0^\tau ds\,\mathrm{Tr}[\hat{O}\bar{\mathcal{P}}e^{s\bar{\mathcal{P}}\mathcal{L}}\mathcal{L}\mathcal{P}\hat{\rho}(\tau-s)]$. By expanding $e^{s\bar{\mathcal{P}}\mathcal{L}}$ in powers of g, we get the leading contribution:

$$\int_0^\tau ds\,\mathrm{Tr}[\hat{O}\,\bar{\mathcal{P}}\,e^{s\mathcal{L}_0}\mathcal{L}\mathcal{P}\hat{\rho}(\tau-s)]$$

$$= ig\int_0^\tau ds\,\mathrm{Tr}\{[\hat{V}_0(-s),\hat{O}]\mathcal{P}\hat{\rho}(\tau-s)\}. \tag{29}$$

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

$$ig\int_0^\infty ds\,\langle[\hat{V}_0(-s),\hat{O}]\rangle_{e_0,q(\tau)}. \tag{30}$$

By the assumption of fast equilibration at $s \approx \tau^*$, this expression is $\mathcal{O}(g\tau^*)$; i.e., it is a small correction to $\langle\hat{O}\rangle_{e_0,q(\tau)}$. 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 $\hat{Q}$ [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 $\hat{\rho}_e$ corresponding to $\hat{H}$, with $\langle\hat{h}\rangle_{\hat{\rho}_e} = \langle\hat{h}\rangle_{\hat{\rho}_{e_0,q^*}} + \mathcal{O}(g)$. 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 $\hat{\rho}_e$, 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 $\hat{\rho}_e$ 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 use the last term in Eq. (28) to derive the $\mathcal{O}(g)$ correction to observables in $\hat{\rho}(\tau)$ from their expectation values in $\hat{\rho}_{e_0,q(\tau)}$. Here, we point out that this is also what the autonomous Eq. (23) dictates.

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

$$\partial_\tau q(\tau) = \mathrm{Tr}[\hat{q}\mathcal{L}\hat{\rho}(\tau)] = \mathrm{Tr}[\hat{q}\mathcal{L}_1\hat{\rho}(\tau)], \tag{31}$$

where, in the last equality, we use that $[\hat{H}_0, \hat{Q}] = 0$. Previously, we evaluated this time derivative in a different way: we approximated $\hat{\rho}(\tau)$ by $\hat{\rho}_{e_0,q(\tau)}$ [assuming, again, that $p(e_0, q)$ is a Dirac delta function] and we arrived at

$$\partial_\tau q(\tau) = \mathrm{Tr}[\hat{q}\mathcal{K}\hat{\rho}_{e_0,q(\tau)}]. \tag{32}$$

This is a more abstract rendering of the autonomous Eq. (23). At this point, one can ask which form of $\hat{\rho}(\tau)$ would reconcile Eqs. (31) and (32). By the cyclic property of the trace, and using that $[\hat{q}, \hat{\rho}_{e_0,q(\tau)}] = 0$, we note that Eq. (31) does not depend on the leading contribution $\hat{\rho}_{e_0,q(\tau)}$ to $\hat{\rho}(\tau)$. 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

$$\int_0^\infty ds\,\bar{\mathcal{P}}e^{s\mathcal{L}_0}\mathcal{L}_1\hat{\rho}_{e_0,q(\tau)}, \tag{33}$$

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 $\hat{\rho}(\tau)$ 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 paper, we test the previous ideas and analytical results in numerical experiments. We focus on

the dynamics of nonintegrable models of strongly interacting hard-core bosons in one-dimensional 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

$$\hat{H}_\alpha = \hat{H}_0 + g_\alpha \hat{V}_\alpha, \tag{34}$$

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 t-V-t'-V' Hamiltonian for hard-core bosons in 1D lattices [48,49], with nearest (next-nearest) neighbor hopping t (t') and interaction V (V'):

$$\hat{H}_0 = \sum_i \left[-t(\hat{b}_i^\dagger \hat{b}_{i+1} + \text{H.c.}) - t'(\hat{b}_i^\dagger \hat{b}_{i+2} + \text{H.c.}) \right.$$
$$\left. + V\left(\hat{n}_i - \frac{1}{2}\right)\left(\hat{n}_{i+1} - \frac{1}{2}\right) + V'\left(\hat{n}_i - \frac{1}{2}\right)\left(\hat{n}_{i+2} - \frac{1}{2}\right) \right], \tag{35}$$

where $\hat{b}_i^\dagger$ ($\hat{b}_i$) is the hard-core boson creation (annihilation) operator and $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ 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) [50]. Here, we focus on cases in which $t' = V' \neq 0$, so that $\hat{H}_0$ is nonintegrable [48,49].

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

$$g_1 \hat{V}_1 = g_1 \sum_i \left[\hat{b}_i + \frac{1}{2}(\hat{b}_i \hat{b}_{i+1} - \hat{b}_i^\dagger \hat{b}_{i+1}) + \text{H.c.} \right]. \tag{36}$$

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

The second perturbation we consider is

$$g_2 \hat{V}_2 = g_2 \sum_i \left(\hat{b}_i + \frac{1}{2} \hat{b}_i \hat{b}_{i+1} + \text{H.c.} \right). \tag{37}$$

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

B. Initial states and description after equilibration

We study the quantum dynamics of initial states $\hat{\rho}_I$ that are far from equilibrium with respect to both $\hat{H}_0$ and $\hat{H}_\alpha$. This is achieved by choosing $\hat{\rho}_I$ to be thermal equilibrium

states of initial Hamiltonians $\hat{H}_I$, 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 system 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 $\hat{\rho}(\tau) = e^{-i\hat{H}_\alpha \tau} \hat{\rho}_I e^{i\hat{H}_\alpha \tau}$. We are interested in the dynamics of observables $\hat{O}$, whose expectation values are given by $O(\tau) = \text{Tr}[\hat{\rho}(\tau)\hat{O}]$. At long times, one expects observables to equilibrate at the values predicted by the diagonal ensemble (DE), $O_{\text{DE}} = \text{Tr}[\hat{\rho}_{\text{DE}}\hat{O}]$, where $\hat{\rho}_{\text{DE}} = \lim_{\tau' \to \infty}(1/\tau') \int_0^{\tau'} d\tau \hat{\rho}(\tau)$ is the density matrix of the diagonal ensemble [36]. When written in the basis of eigenkets $|E_i^\alpha\rangle$ of $\hat{H}_\alpha$, $\hat{\rho}_{\text{DE}}$ takes the form

$$\hat{\rho}_{\text{DE}} = \sum_i (\langle E_i^\alpha | \hat{\rho}_I | E_i^\alpha \rangle) |E_i^\alpha\rangle \langle E_i^\alpha|. \tag{38}$$

For nonintegrable (quantum chaotic) systems, because of eigenstate thermalization [36–38], one expects the predictions of the diagonal ensemble to match those of traditional statistical mechanics ensembles; namely, we expect observables to *thermalize* [22]. For the t-V-t'-V' Hamiltonian for hard-core bosons in 1D lattices, eigenstate thermalization and thermalization were studied in Ref. [48], while quantum chaos was studied in Ref. [49]. 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 μ. The grand canonical ensemble density matrices have the form

$$\hat{\rho}_{\text{GE}} = \begin{cases} \dfrac{e^{-\hat{H}_\alpha/T}}{\text{Tr}[e^{-\hat{H}_\alpha/T}]} & \text{when } g_\alpha \neq 0 \\[2mm] \dfrac{e^{-(\hat{H}_\alpha + \mu \hat{N})/T}}{\text{Tr}[e^{-(\hat{H}_\alpha + \mu \hat{N})/T}]} & \text{when } g_\alpha = 0. \end{cases} \tag{39}$$

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

$$\text{Tr}[\hat{\rho}_{\text{GE}}\hat{H}_\alpha] = \text{Tr}[\hat{\rho}_I \hat{H}_\alpha]. \tag{40}$$

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

$$\text{Tr}[\hat{\rho}_{\text{GE}}\hat{N}] = \text{Tr}[\hat{\rho}_I \hat{N}]. \tag{41}$$

Because of 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_α, 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 [44]. 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 $\mathcal{O}(g)$ correction. This projected state is a diagonal ensemble of $\hat{H}_0$, whose density matrix [see Eq. (6)] we denote as

$$\hat{\rho}_0(\tau) \equiv \mathcal{P}_{\mathrm{DE}}\hat{\rho}(\tau) = \sum_i (\langle E_i^0|\hat{\rho}(\tau)|E_i^0\rangle)|E_i^0\rangle\langle E_i^0|, \quad (42)$$

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

C. Numerical linked cluster expansion (NLCE)

In what follows, we use the numerical linked cluster expansion (NLCE) approach introduced in Ref. [23] (see Refs. [51,52] 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 [53], 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

$$\mathcal{O} = \sum_c M(c) \times W_O(c). \quad (43)$$

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

$$W_O(c) = O^c - \sum_{s \subset c} W_O(s), \quad (44)$$

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

In our calculations, the density matrix of the initial state in each cluster $\hat{\rho}_I^c$ 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,

$$\hat{\rho}_I^c = \frac{e^{-(\hat{H}_I^c + \mu_I \hat{N}^c)/T_I}}{\mathrm{Tr}[e^{-(\hat{H}_I^c + \mu_I \hat{N}^c)/T_I}]}. \quad (45)$$

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 τ denotes the time after the quench, the density matrix of each cluster is evolved with the Hamiltonian after the quench $\hat{H}_\alpha^c$,

$$\hat{\rho}^c(\tau) = (e^{-i\hat{H}_\alpha^c \tau})\hat{\rho}_I^c(e^{i\hat{H}_\alpha^c \tau}), \quad (46)$$

and the NLCE calculation is carried out as usual [23]. 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 $\hat{\rho}_{\mathrm{DE}}^c$, $\hat{\rho}_{\mathrm{GE}}^c$, and $\hat{\rho}_0^c(\tau)$ 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 [23], and the description of observables after equilibration [54], 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. [55], 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. Here, as in Refs. [23,54], we use that NLCE in our calculations (there is only one maximally connected cluster with a given number of sites). The number of sites in the largest cluster considered defines 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,

$$N(\tau) = \mathrm{Tr}[\hat{\rho}(\tau)\hat{N}], \quad (47)$$

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 one-body nearest neighbor correlation,

$$K(\tau) = \mathrm{Tr}[\hat{\rho}(\tau)\hat{K}], \qquad (48)$$

whose value per site is denoted as $k(\tau)$, where

$$\hat{K} = \sum_i (\hat{b}_i^\dagger \hat{b}_{i+1} + \hat{b}_{i+1}^\dagger \hat{b}_i). \qquad (49)$$

$k(\tau)$ 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,

$$M_k(\tau) = \mathrm{Tr}[\hat{\rho}(\tau)\hat{M}_k], \qquad (50)$$

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

$$\hat{M}_k = \sum_{j,j'} e^{ik(j-j')} \hat{b}_j^\dagger \hat{b}_{j'}. \qquad (51)$$

$m_k(\tau)$ is a nonlocal one-body 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 [56]. $m_k(\tau)$ was the observable used in Ref. [2] to show lack of thermalization in 1D Bose gases with contact interactions, and in Ref. [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 $\mu_I = 2$ (similar results are obtained for other values of T_I, not too low, and μ_I), 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, $\hat{H}_0$ [48,49] and $\hat{H}_\alpha$ are quantum chaotic and the system thermalizes for all values of g_α (see Ref. [23] for a 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 $g_\alpha \neq 0$ (the largest Hamiltonian sector that needs to be diagonalized has 65 792 states after exploiting reflection symmetry), and, thanks to particle number conservation, up to the 19th order for quenches with $g_\alpha = 0$ (the dimension of the largest Hamiltonian

sector that needs to be diagonalized is 46 252). 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 $g_\alpha \neq 0$ and up to the 18th order for $g_\alpha = 0$. 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 $\hat{H}_1$ 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 $n_{\mathrm{DE}} = 1/2$. For $g_1 = 0.12$, $n(\tau = 100) \approx n_{\mathrm{DE}}$.

The approach of $n(\tau)$ to $n_{\mathrm{DE}} = 0.5$ is exponential. This is apparent in Fig. 1(b), where we plot the normalized "distance" to equilibrium:

$$\delta_l^{\mathrm{DE}}[n(\tau)] = \left| \frac{n_l(\tau) - n_{\mathrm{DE}}}{n_{\mathrm{DE}}} \right|. \qquad (52)$$

We fit $\delta_l^{\mathrm{DE}}[n(\tau)]$ to an exponential function $\propto \exp[-\Gamma_l^{\mathrm{NLCE}}(g_1)\tau]$ to obtain the thermalization rate $\Gamma_l^{\mathrm{NLCE}}(g_1)$. 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 and NLCE-16, respectively. They agree with each other within the errors of the fits. This suggests that our calculation of $\Gamma^{\mathrm{NLCE}}(g_1)$ is robust. A power-law fit to the rates obtained for $l = 17$ is also shown in Fig. 1(c). We find that $\Gamma_{17}^{\mathrm{NLCE}}(g_1) \propto g_1^\beta$ with $\beta = 1.99$, in agreement with the analytical results in Sec. III.

In closing Sec. IV C, we argued that the rate $\dot{q}(\tau) = 0$ whenever the state $\hat{\rho}(\tau)$ commutes exactly with $\hat{Q}$. The rate also vanishes if $\hat{\rho}(\tau)$ is time-reversal invariant about τ. Both conditions apply to our initial states $\hat{\rho}_I$. As a result, there is a narrow plateau in $n(\tau)$ for $\tau \leq 1$. This plateau is best seen in Fig. 9. 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 $\Gamma^{\mathrm{NLCE}}(g_1)$ 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 for $\dot{n}(\tau) = dn/d\tau'|_{\tau'=\tau}$:
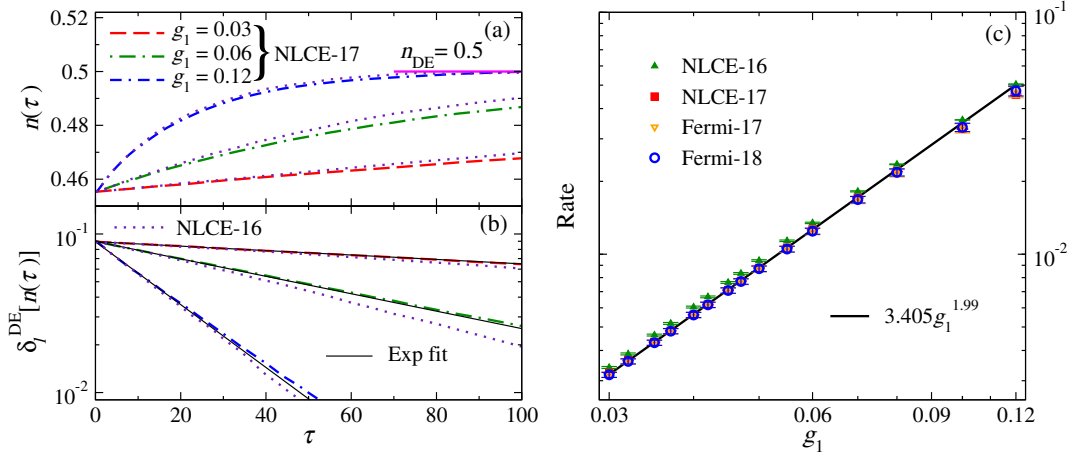
FIG. 1. Dynamics of the particle filling after quantum quenches $\hat{H}_I \rightarrow \hat{H}_1$. (a) Dynamics of the particle filling $n(\tau)$ and (b) dynamics of the "distance" to equilibrium $\delta_l^{DE}[n(\tau)]$; see Eq. (52). NLCE results are shown for $l = 17$ [NLCE-17, see legend in (a)] and $l = 16$ (NLCE-16, dotted lines). Straight lines in (b) depict fits to the results for $l = 17$ in the interval $\tau \in [1, 16]$ and to exponential functions $\propto \exp[-\Gamma_{17}^{NLCE}(g_1)\tau]$. (c) Thermalization rates $\Gamma^{NLCE}(g_1)$ obtained from fits as the ones in (b), for $l = 17$ in the interval $\tau \in [1, 16]$ (NLCE-17), and for $l = 16$ in the interval $\tau \in [1, 6]$ (NLCE-16), reported for $g_1 \in [0.03, 0.12]$. Error bars show 95% confidence bounds for the fits. The straight line is the result of a fit to $\Gamma_{17}^{NLCE}(g_1) \propto g_1^{\beta}$ for $g_1 \in [0.03, 0.06]$. The open symbols show the rates $\Gamma^{Fermi}(g_1)$ obtained evaluating Fermi's golden rule [see Eqs. (53) and (54)] using full exact diagonalization in chains with $L = 18$ (Fermi-18) and $L = 17$ (Fermi-17), and periodic boundary conditions. The error bars show the standard deviation from averages over different choices of ΔE and τ (see Appendix D).

$$\dot{n}(\tau) = \frac{2\pi g_\alpha^2}{L} \sum_{i,j} \delta(E_j^0 - E_i^0)(N_j - N_i)P_i^0(\tau)$$
$$\times |\langle E_j^0|\hat{V}_\alpha|E_i^0\rangle|^2, \tag{53}$$

where $|E_i^0\rangle$ are the eigenkets of $\hat{H}_0$ with energies E_i^0, $N_i = \langle E_i^0|\hat{N}|E_i^0\rangle$, and we average over the diagonal ensemble distribution, $P_i^0(\tau) = \langle E_i^0|\hat{\rho}(\tau)|E_i^0\rangle$. To evaluate Eq. (53) numerically, we replace $\sum_j \delta(E_j^0 - E_i^0)$ by a sum over energies E_j^0 that lie within a small energy window $[E_i^0 - \Delta E/2, E_i^0 + \Delta E/2]$ (see Appendix D).

The thermalization rates are estimated by computing

$$\Gamma^{Fermi}(g_\alpha) = -\frac{\dot{n}(\tau)}{n(\tau) - n_{DE}}. \tag{54}$$

When $|n(\tau) - n_{DE}| \ll n_{DE}$, $\Gamma^{Fermi}(g_\alpha)$ becomes independent of τ and $n(\tau)$ relaxes exponentially. This condition is, to a good degree, satisfied for our choice of initial state, for which $\delta_l^{DE}[n(\tau)] < 0.09$ [see Fig. 1(b)]. Our calculations of $\Gamma^{Fermi}(g_\alpha)$ 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 ΔE and τ for which the results for $\Gamma^{Fermi}(g_\alpha)$ are robust against the choice of ΔE, τ, and L (see Appendix D).

In Fig. 1(c), we report our results for $\Gamma^{Fermi}(g_1)$. They are in excellent agreement with $\Gamma^{NLCE}(g_1)$. We should add that, for quenches $\hat{H}_I \rightarrow \hat{H}_2$, Eq. (53) predicts the same leading $\mathcal{O}(g_\alpha^2)$ dynamics as under $\hat{H}_1$. This is the case because the

terms that change the total particles number are the same in $\hat{V}_1$ and $\hat{V}_2$. Hence, $n(\tau)$ is the same for both Hamiltonians up to corrections $\mathcal{O}(g_\alpha^3)$.

B. Dynamics of the one-body nearest neighbor correlation

Next, we study the dynamics of the one-body nearest neighbor correlation, $k(\tau)$ [see Eq. (49)]. In contrast to the particle filling studied in Sec. VI A, the nearest neighbor correlation $k(\tau)$ exhibits dynamics even if $g_\alpha = 0$ after the quench.

Figure 2 shows the 18th (NLCE-18) and 19th (NLCE-19) orders of the NLCE for $k(\tau)$ after a quench in which $g_\alpha = 0$.
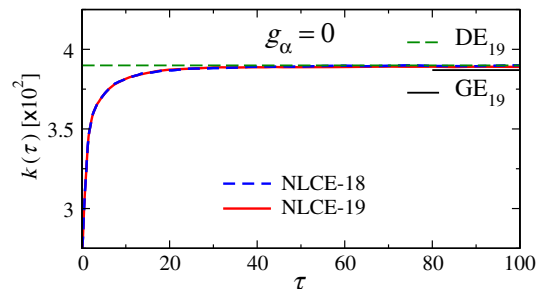

FIG. 2. Dynamics of the one-body nearest neighbor correlation $k(\tau)$ when $g_\alpha = 0$ after the quench. Results are shown for the 18th (NLCE-18) and 19th (NLCE-19) orders of the NLCE. We also report, as horizontal lines, the results for the 19th order of the diagonal ensemble (DE$_{19}$) and for the 19th order of the grand canonical ensemble (GE$_{19}$).

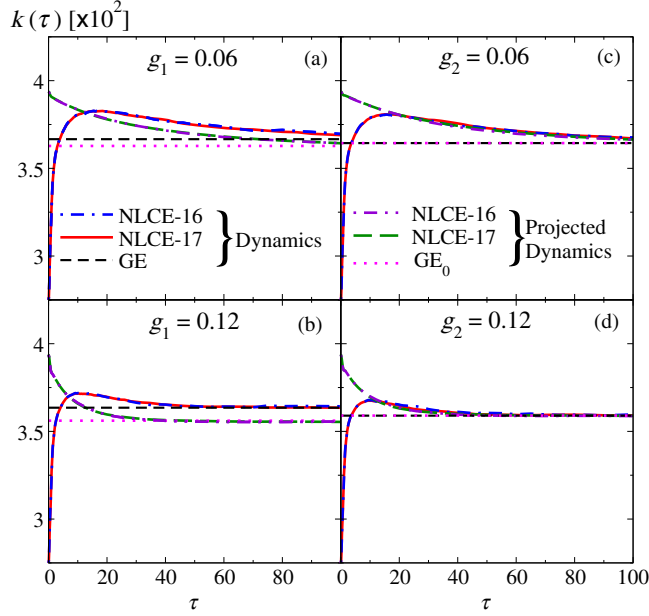FIG. 3. Dynamics of the one-body nearest neighbor correlation $k(\tau)$ for (a) $g_1 = 0.06$, (b) $g_1 = 0.12$, (c) $g_2 = 0.06$, and (d) $g_2 = 0.12$. Results are shown for the 16th (NLCE-16) and 17th (NLCE-17) orders of the NLCE, both for the dynamics [see legends in (a)] and for the dynamics in the projected basis of $\hat{H}_0$ [see legends in (c)]. The horizontal lines show the results for the grand canonical ensemble corresponding to the original dynamics (GE) and to the projected dynamics (GE$_0$), both evaluated at the 17th order of the NLCE.

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 $\hat{H}_0$ after the quench is nonintegrable. The small difference between the diagonal and the grand canonical ensemble results is due to the lack of convergence of the NLCE for the former (the latter is fully converged) [54]. 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-hand panels) and $\hat{H}_I \rightarrow \hat{H}_2$ (right-hand 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 $\hat{\rho}_{\mathrm{GE}}$ 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 $\hat{\rho}_{\mathrm{GE}_0}$. We compute the temperature T_0 in $\hat{\rho}_{\mathrm{GE}_0}$ using the expectation value of $\hat{H}_0$ in the thermal equilibrium state $\hat{\rho}_{\mathrm{GE}}$ of $\hat{H}_\alpha$:

$$\mathrm{Tr}[\hat{\rho}_{\mathrm{GE}_0}\hat{H}_0] = \mathrm{Tr}[\hat{\rho}_{\mathrm{GE}}\hat{H}_0]. \qquad (55)$$

The chemical potential in $\hat{\rho}_{\mathrm{GE}_0}$ is $\mu_0 = 0$ because, for $g_\alpha \neq 0$, the systems after equilibration are at half filling.

In the left-hand panels in Fig. 3, for quenches $\hat{H}_I \rightarrow \hat{H}_1$, one can see the advanced offsets (see Sec. IV) between the 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-hand 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 $\hat{\rho}_{\mathrm{GE}}$ and $\hat{\rho}_{\mathrm{GE}_0}$.

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 $\hat{\rho}_0(\tau)$ [$\hat{\rho}_0(\tau)$ is defined in Eq. (42)] are block diagonal in the particle number basis, only the presence of terms in $\hat{V}_\alpha$ that do not change the particle number can produce an $\mathcal{O}(g)$ correction. Such terms are present in $\hat{V}_1$ (the hopping terms), see Eq. (36), but are absent in $\hat{V}_2$, see Eq. (37). This means that, to leading order, $\Delta k_1(\tau) \propto g_1$ while $\Delta k_2(\tau) \propto g_2^\beta$ with $\beta \geq 2$.

In Fig. 4, we show the long-time offsets between the actual and projected dynamics as functions of g_α (for positive and negative values of g_α) 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 $\hat{\rho}_{\mathrm{GE}}$ and $\hat{\rho}_{\mathrm{GE}_0}$ for the equilibrated results (see the horizontal lines in Fig. 3). Those predictions are converged to machine precision in the 17th order of the NLCE shown in Fig. 4. Figure 4(a), for quenches $\hat{H}_I \rightarrow \hat{H}_1$, 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 $\hat{H}_I \rightarrow \hat{H}_2$. 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 $g_1 = 0.06$ 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 one-body nearest neighbor correlation being particle-hole symmetric
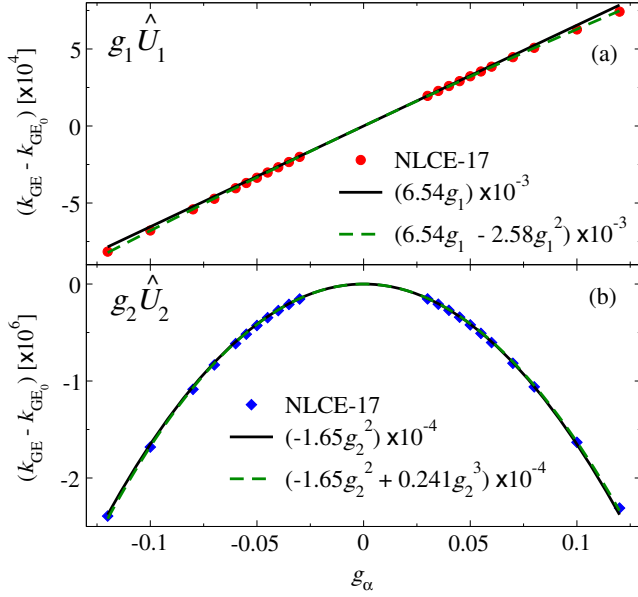
FIG. 4. Difference in the equilibrium (grand canonical ensemble) results between the actual dynamics (GE) and the projected dynamics (GE$_0$) for the one-body nearest neighbor correlation ($k_{GE} - k_{GE_0}$) evaluated at the 17th order of the NLCE. (a) $g_1 \hat{V}_1$ perturbation (symbols), with linear and linear plus quadratic fits (lines). (b) $g_2 \hat{V}_2$ perturbation (symbols), with quadratic and quadratic plus cubic fits (lines).

FIG. 5. Momentum distribution function $m_k(\tau)$ [see Eq. (50)] at four times after quenches in which $\hat{H}_I \to \hat{H}_1$ with $g_1 = 0.03$ [(a)–(d)] and $g_1 = 0.12$ [(e)–(h)]. $m_k(\tau)$ is evaluated at $\tau = 0$ [(a),(e)], $\tau = 2$ [(b),(f)], $\tau = 10$ [(c),(g)], and $\tau = 100$ [(d),(h)]. Results are reported for the 15th (NLCE-15) and 16th (NLCE-16) orders of the NLCE. We show $m_k(\tau)$ for the dynamics [see the legends in (a)], and for the projected dynamics in the basis of $\hat{H}_0$ [see the legends in (b)]. The predictions of the diagonal ensemble (DE) and the grand canonical ensemble (GE) for the dynamics, and of the grand canonical ensemble for the projected dynamics (GE$_0$), all evaluated at the 16th order of the NLCE, are shown in (d) and (h) [see the legends in (d)]. The inset in (h) shows the predictions of the latter three ensembles about $k = \pi$.

(as the Hamiltonians $\hat{H}_\alpha$ are). This means that, close to equilibrium, the difference between $k(\tau)$ and k_{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 $k(\tau) - k_{GE}$ is 2, then $\dot{k}(\tau) \simeq 2\Gamma^{\text{Fermi}}(g_\alpha)[k(\tau) - k_{GE}]$. 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 the momentum distribution function $m_k(\tau)$ [Eq. (50)] for quenches $\hat{H}_I \to \hat{H}_1$. 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-hand panels in Fig. 5) and $g_1 = 0.12$ (right-hand 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 of $m_k(\tau)$ in the basis of $\hat{H}_0$ [the dynamics dictated by $\hat{\rho}_0(\tau)$, see Eq. (42)]. Figures 5(b) and 5(f) show that the momentum distribution in the original and projected dynamics become nearly indistinguishable from each other after short times ($\tau \gtrsim 2$),
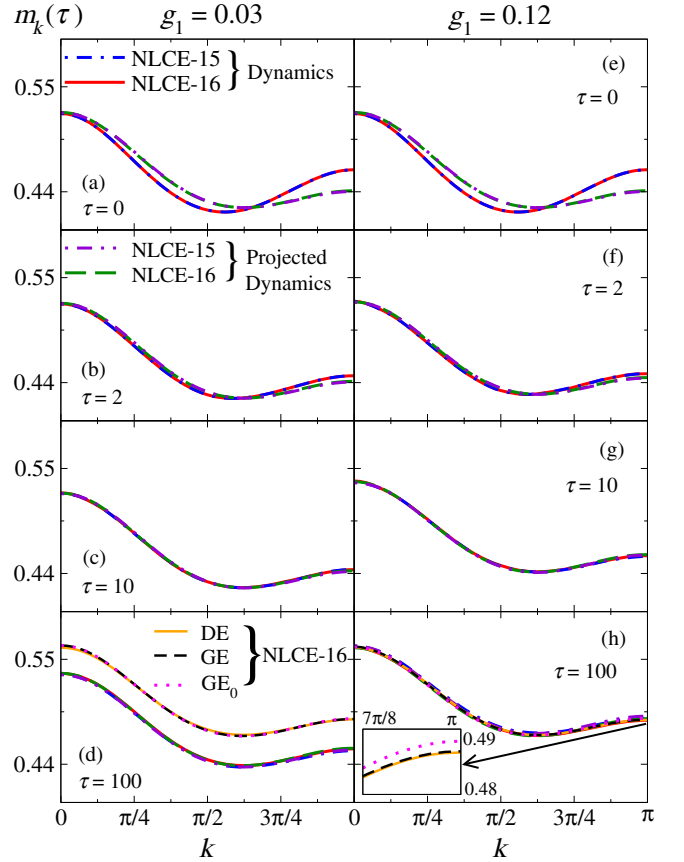
and remain so at long times [see Figs. 5(c), 5(d), 5(g), and 5(h)].

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
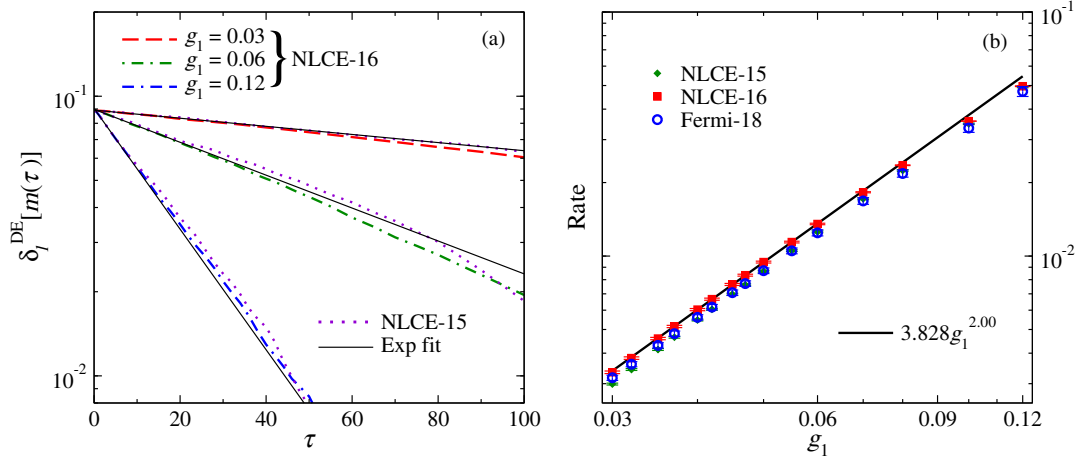
FIG. 6. Dynamics of the momentum distribution function $m_k(\tau)$ [Eq. (50)] after quenches $\hat{H}_I \rightarrow \hat{H}_1$. (a) Dynamics of the "distance" to equilibrium $\delta^{DE}[m(\tau)]$; see Eq. (57). NLCE results are shown for $l = 16$ [NLCE-16] and $l = 15$ (NLCE-15, dotted lines). Straight lines in (a) depict fits to the results for $l = 16$ in the interval $\tau \in [2, 6]$ and to exponential functions $\propto \exp[-\Gamma_{16}^{NLCE}(g_1)\tau]$. (b) Thermalization rates $\Gamma^{NLCE}(g_1)$ of $m_k(\tau)$ (filled symbols) obtained from fits as the ones in (a), for $l = 16$ in the interval $\tau \in [2, 6]$ (NLCE-16), and for $l = 15$ in the interval $\tau \in [2, 5]$ (NLCE-16), reported for $g_1 \in [0.03, 0.12]$. Error bars show 95% confidence bounds for the fits. The straight line is the result of a fit to $\Gamma_{16}^{NLCE}(g_1) \propto g_1^\beta$ for $g_1 \in [0.03, 0.06]$. The open circles show the rates $\Gamma^{Fermi}(g_1)$ [also reported in Fig. 1(c)] 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 ΔE and τ (see Appendix D).

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 $k(\tau)$ in Sec. VI B (see Fig. 3), there is an $\mathcal{O}(g_1)$ 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 one-body nearest neighbor correlation, is not particle-hole symmetric. This implies that, as $m_k(\tau)$ approaches its equilibrium value in the diagonal ensemble m_k^{DE}, to leading order

$$m_k(\tau) - m_k^{DE} \propto \left[n(\tau) - \frac{1}{2} \right]. \qquad (56)$$

Thus, we expect m_k to thermalize with the same rate $\Gamma^{Fermi}(g_1)$ given by Eq. (54). To quantify the "distance" to equilibrium for $m_k(\tau)$, we compute

$$\delta_l^{DE}[m(\tau)] = \frac{\sum_k |m_k^l(\tau) - m_k^{l,DE}|}{\sum_k m_k^{l,DE}}. \qquad (57)$$

In Fig. 6(a), we show $\delta_l^{DE}[m(\tau)]$ evaluated at the 15th (NLCE-15) and 16th (NLCE-16) orders of the NLCE. $\delta_{15}^{DE}[m(\tau)]$ and $\delta_{16}^{DE}[m(\tau)]$ can be seen to decay close to exponentially, although the convergence of the results for $\delta_l^{DE}[m(\tau)]$ in Fig. 6(a) is not as good as for $\delta_l^{DE}[n(\tau)]$ 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 $\delta_l^{DE}[m(\tau)]$ to an exponential function $\propto \exp[-\Gamma_l^{NLCE}(g_1)\tau]$ to obtain the thermalization rates $\Gamma_l^{NLCE}(g_1)$ 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 $\Gamma_l^{NLCE}(g_1)$ are reported in Fig. 6(b) for $g_1 \in [0.03, 0.12]$. In Fig. 6(b), we also plot the rates $\Gamma^{Fermi}(g_1)$ obtained by evaluating Fermi's golden rule [see Eqs. (53) and (54)] using full exact diagonalization in chains with $L = 18$ sites and periodic boundary conditions (see Sec. VI A and Appendix D). [The rates $\Gamma^{Fermi}(g_1)$ 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 $\Gamma_{16}^{NLCE}(g_1)$ is also shown in Fig. 6(b). We find that $\Gamma_{16}^{NLCE}(g_1) \propto g_1^\beta$ 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 AND DISCUSSION

We put forward a conceptually simple scenario for prethermalization and thermalization in 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, as well as to nonintegrable systems 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 quasiconserved quantity (or quantities). The separation of timescales leads to a universal description, two aspects of which stand out.

(i) The dynamics of the (or each) quasiconserved 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. [12] 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 numerical experiments in systems for which both the reference and the perturbed dynamics are nonintegrable.

A systematic study of integrable systems in which integrability is weakly broken is beyond the scope of this work. As mentioned in the Introduction, noninteracting systems in the presence of weak integrability-breaking interactions were studied in Refs. [7–15]. There have also been studies of integrable systems mappable onto noninteracting ones in the presence of an integrability-breaking perturbation [57]. For a strongly interacting integrable system, $\hat{H}_0$ in Eq. (35) with $t = V = 1$ and $t' = V' = 0$ (which is not mappable onto a noninteracting model), in Ref. [23] it was shown numerically that weakly breaking integrability by making $g \equiv t' = V' \neq 0$ results in thermalization rates $\propto g^2$.

To check the applicability of our theory to strongly interacting integrable reference dynamics, in Fig. 7 we show results for dynamics under the same reference Hamiltonian as in Ref. [23] when one breaks integrability with the perturbation $g_2 \hat{V}_2$ in Eq. (37). Those results are the equivalent of results reported in Figs. 1 and 3, for quenches $\hat{H}_I \to \hat{H}_2$ in which $t'_I = V'_I = 0$ in $\hat{H}_I$ and $t' = V' = 0$ in $\hat{H}_2$. The main panel in Fig. 7 shows that the thermalization rate for the particle filling is $\propto g_2^2$, and that it agrees with the Fermi golden rule prediction [see Eqs. (53) and (54)] for small values of g_2. The upper inset in Fig. 7 shows that the particle filling approaches exponentially its thermal equilibrium value, as seen in Fig. 1(b) for the nonintegrable reference dynamics. The lower inset in Fig. 7 shows the
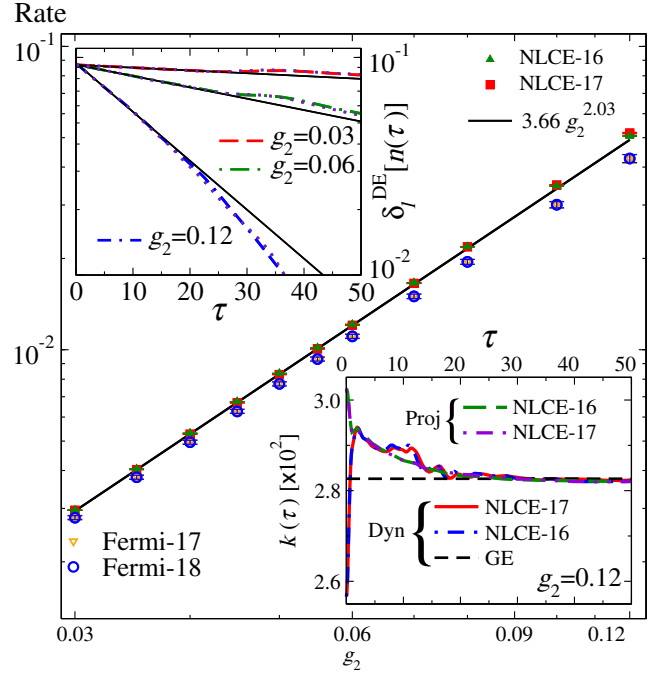

FIG. 7. Dynamics after quantum quenches $\hat{H}_I \to \hat{H}_2$, for $t'_I = V'_I = 0$ in $\hat{H}_I$ and $t' = V' = 0$ in $\hat{H}_2$ (integrable reference dynamics). Main panel: Thermalization rates $\Gamma^{\text{NLCE}}(g_2)$ for the particle filling dynamics (filled symbols) obtained from fits as the ones shown in the upper inset. The straight line is the result of a fit to $\Gamma_{17}^{\text{NLCE}}(g_2) \propto g_2^\beta$. The open symbols show the rates $\Gamma^{\text{Fermi}}(g_2)$, see Eqs. (53) and (54), obtained using full exact diagonalization (see Appendix D). Upper inset: $\delta_l^{\text{DE}}[n(\tau)]$, see Eq. (52), versus τ. The straight lines depict fits to exponential functions $\propto \exp[-\Gamma_{17}^{\text{NLCE}}(g_2)\tau]$. The legends in the main panel and in the upper inset follow Fig. 1, and the fits are done in the same intervals as in Fig. 1. Lower inset: Dynamics, and projected dynamics in the basis of $\hat{H}_0$, of $k(\tau)$ for $g_2 = 0.12$. The horizontal line shows the result for the grand canonical ensemble corresponding to the original dynamics. The legends are identical to those in Fig. 3.

dynamics, and the projected dynamics in the basis of $\hat{H}_0$, of $k(\tau)$ for $g_2 = 0.12$. They are in excellent agreement with each other, like in Fig. 3(d) for the nonintegrable reference dynamics. In short, there are no qualitative differences between the results reported in Figs. 1 and 3 for the nonintegrable reference dynamics and in Fig. 7 for the integrable reference dynamics.

We should stress that all our numerical results for the projected dynamics were obtained within the diagonal ensemble, which allowed us to study indistinctively quantum chaotic reference dynamics in Sec. VI, and strongly interacting integrable reference dynamics in Fig. 7. However, because of eigenstate thermalization in quantum chaotic systems, and its generalized version in integrable systems, one can equally well use traditional ensembles of statistical mechanics for the projected dynamics when the reference dynamics is quantum chaotic [22] and generalized Gibbs
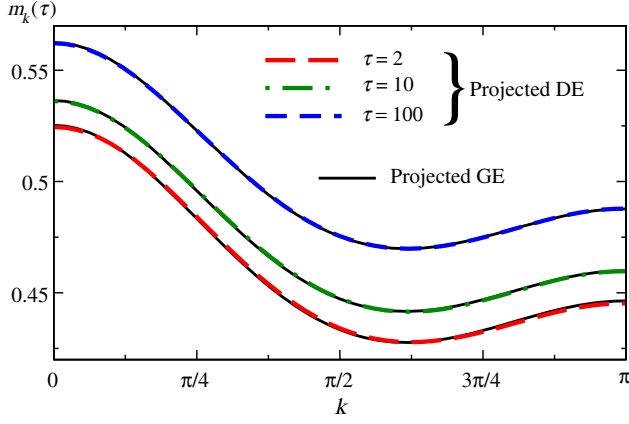
FIG. 8.   Projected dynamics of $m_k(\tau)$ in the basis of $\hat{H}_0$ at three times after the quench $\hat{H}_I \to \hat{H}_1$, with $g_1 = 0.12$. We compare results for the projected dynamics within the diagonal ensemble (Projected DE), also reported in Fig. 5, with those obtained within the grand canonical ensemble (Projected GE). The temperature and the chemical potential of the grand canonical ensemble are set so that the energy density e_0 and the particle filling n in this ensemble match those in the diagonal ensemble. The results reported are for the 16th order of the NLCE.

ensembles when the reference dynamics is integrable [17]. As an example, in Fig. 8 we compare the results for the projected dynamics in the diagonal ensemble reported in Fig. 5 for $g_1 = 0.12$ at times $\tau = 2$, 10, and 100 with those obtained replacing the diagonal ensemble by the grand canonical ensemble with the same energy density e_0 and particle filling n. As expected, the results obtained within both ensembles are indistinguishable from each other.

APPENDIX A: 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 $\mathcal{P}: V \to V$, $\mathcal{P}^2 = \mathcal{P}$, and a linear evolution equation,

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

Then, we write $\mathcal{P}_1 = \mathcal{P}$, $\mathcal{P}_2 = (1 - \mathcal{P})$, $a_i = \mathcal{P}_i a$, and $M_{ij} = \mathcal{P}_i M \mathcal{P}_j$. As the projector is time independent, the evolution equation can be cast as a system of two coupled equations:

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

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

We first formally solve the second of these equations,

$$a_2(\tau) = e^{\tau M_{22}}a_2(0) + \int_0^\tau ds\, e^{sM_{22}} M_{21}a_1(\tau - s), \quad \text{(A4)}$$

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

$$\dot{a}_1(\tau) = M_{11}a_1(\tau) + M_{12}e^{\tau M_{22}}a_2(0)$$
$$+ \int_0^\tau ds\, M_{12}e^{sM_{22}} M_{21}a_1(\tau - s). \quad \text{(A5)}$$

See Eqs. (7) and (12) in the main text. The relevance of this framework to irreversible phenomena was noticed in Refs. [42,43].

APPENDIX B: FROM $\mathcal{K}$ TO THE AUTONOMOUS EQUATION

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

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

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

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

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

$$\mathcal{K}_{\text{gain}}\hat{\rho} = g^2 \int_0^\infty de_0 dq \hat{\rho}_{e_0,q} \int_{-\infty}^\infty ds\, \text{Tr}[\hat{P}_{e_0,q}\hat{V}_0(s)\mathcal{P}\hat{\rho}\,\hat{V}], \quad \text{(B3)}$$

while for $\mathcal{K}_{\text{loss}}$ acting on $\hat{\rho}$ one has

$$\mathcal{K}_{\text{loss}}\hat{\rho} = -g^2 \int de dq \hat{\rho}_{e_0,q} \int_{-\infty}^\infty ds\, \text{Tr}[\hat{P}_{e_0,q}\hat{V}_0(s)\hat{V}\mathcal{P}\hat{\rho}], \quad \text{(B4)}$$

where we abbreviate $\hat{P}_{e_0,q} = P(\hat{H}_0 \approx e_0 L)P(\hat{Q} \approx qL)$ (see Sec. III A), and we use the invariance of $\mathcal{P}\hat{\rho}$ and $\hat{P}_{e_0,q}$ under the evolution generated by $\hat{H}_0$. Note that we can assume,

without loss of generality, that $\hat{V}$ satisfies $\langle \hat{V} \rangle_{e_0,q} = 0$ for all (e_0, 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 (e_0, 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 $\mathcal{K}$ as a kernel on the space of distributions p. We then have

$$\mathcal{K}p(e_0', q') = \int de_0 dq \mathcal{K}(e_0, q; e_0', q') p(e_0, q), \quad (\text{B5})$$

and, from the above formulas, we identify

$$\mathcal{K}(e_0, q; e_0', q') = g^2 \int\limits_{-\infty}^{\infty} ds \text{Tr}[\hat{P}_{e_0',q'}\hat{V}_0(s)\hat{\rho}_{e_0,q}\hat{V}]$$

$$- g^2 \int\limits_{-\infty}^{\infty} ds \text{Tr}[\hat{P}_{e_0',q'}\hat{V}_0(s)\hat{V}\hat{\rho}_{e_0,q}]. \quad (\text{B6})$$

To unravel this further, we introduce

$$\hat{V}^{\delta Q} = \sum_{Q,Q'} \delta_{Q',Q+\delta Q}\hat{P}_{Q'}\hat{V}\hat{P}_Q, \quad (\text{B7})$$

where $\hat{P}_Q$ are spectral projections of $\hat{Q}$ and the sum runs over the eigenvalues of $\hat{Q}$. Note that the admissible values of δQ are $\mathcal{O}(1)$ because $\hat{V}$ is a sum of $\mathcal{O}(1)$ 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 $\mathcal{O}(L)$, as a truncated correlation function in equilibrium. The first term in Eq. (B6) can be written as

$$r_{e_0,q}(\delta Q) = g^2 \int\limits_{-\infty}^{\infty} ds \text{Tr}[\hat{V}_0^{\delta Q}(s)\hat{\rho}_{e_0,q}\hat{V}^{-\delta Q}], \quad (\text{B8})$$

with $e_0' = e_0$ and $q' - q = \delta Q/L$. Indeed, the condition $e_0' = e_0$ is enforced by the integral over s. The second term in Eq. (B6) has a contribution at $e_0 = e_0'$, $q = q'$ only (by the cyclic property of the trace, the projector $\hat{P}_{e_0',q'}$ is put next to the density matrix $\hat{\rho}_{e_0,q}$), and its value is

$$\sum_{\delta Q} r_{e_0,q}(\delta Q). \quad (\text{B9})$$

One could also have guessed this value because the process generated by $\mathcal{K}$ conserves probability, which translates to

$$\int dq de_0' \mathcal{K}(e_0, q; e_0', q') = 0. \quad (\text{B10})$$

We have now explicitly interpreted the process generated by $\mathcal{K}$ as a jump process, with jump rates $r_{e_0,q}(\delta Q) \geq 0$ for jumps in the density q of order $1/L$. The link to the "drift" computed in Sec. III F is by

$$d(e_0, q) = \sum_{\delta Q} \delta Q \frac{1}{L} r_{e_0,q}(\delta Q). \quad (\text{B11})$$

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 Γ^{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 $14\,602$ 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. 11(a). The matrices involved in our full exact diagonalization calculations are complex for sectors with total quasimomentum $k \neq 0$ and π, and the results reported contain the contribution from all L quasimomentum sectors.

In Fig. 9, 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. 9). 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
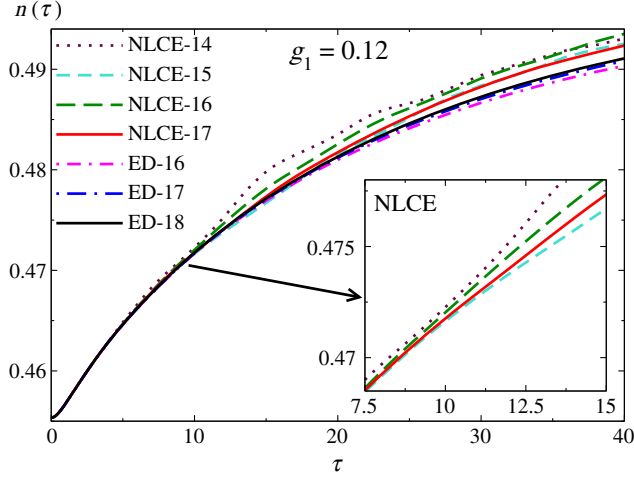
FIG. 9. Time evolution of the filling $n(\tau)$, for $g_1 = 0.12$, obtained in the last four orders l of the NLCE (NLCE-l in the legends), and in the three largest chains with L sites and periodic boundary conditions solved using full exact diagonalization (ED-L in the legends). The inset shows an enlargement of the NLCE results between $\tau = 7.5$ and 15.

happens in thermal equilibrium calculations, in which the NLCEs converge at lower temperatures as one increases the order of the expansion [53]. The exact diagonalization results, on the other hand, can be seen to approach the NLCE ones monotonically with increasing the chain size. In Fig. 9, 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 (ED) calculations of $n_L(\tau)$, where L is the chain size, we compute the relative differences,

$$\Delta_l^{\text{NLCE}}[n(\tau)] = \frac{|n_{l=17}(\tau) - n_l(\tau)|}{|n_I - 1/2|}, \quad \text{(C1)}$$

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

$$\Delta_L^{\text{ED}}[n(\tau)] = \frac{|n_{L=18}(\tau) - n_L(\tau)|}{|n_I - 1/2|} \quad \text{(C2)}$$

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:

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

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.
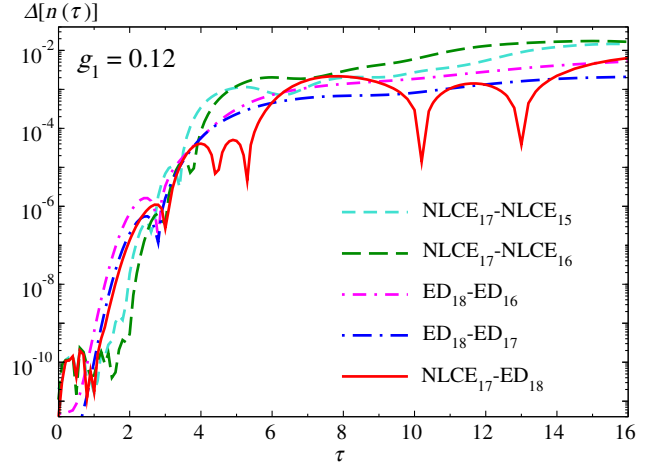

FIG. 10. Relative differences, defined in Eqs. (C1)–(C3), for $g_1 = 0.12$. The exact diagonalization results are obtained in chains with periodic boundary conditions.

In Fig. 10, 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. 9. Some points to be highlighted from the plots in Fig. 10 are the following: (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 $\Delta^{\text{NLCE-ED}}[n(\tau)]$ are qualitatively similar to those for $\Delta_l^{\text{NLCE}}[n(\tau)]$ and $\Delta_L^{\text{ED}}[n(\tau)]$. We find that $\Delta^{\text{NLCE-ED}}[n(\tau)] \leq 0.01\%$ for times $\tau \leq 5$. 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 $\Delta^{\text{NLCE-ED}}[n(\tau)] \lesssim 0.5\%$ for times $\tau \leq 16$. 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 one-body 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. 11(a), we show dynamics after the quench $\hat{H}_I \rightarrow \hat{H}_0$ 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 $\tau \gtrsim 5$,
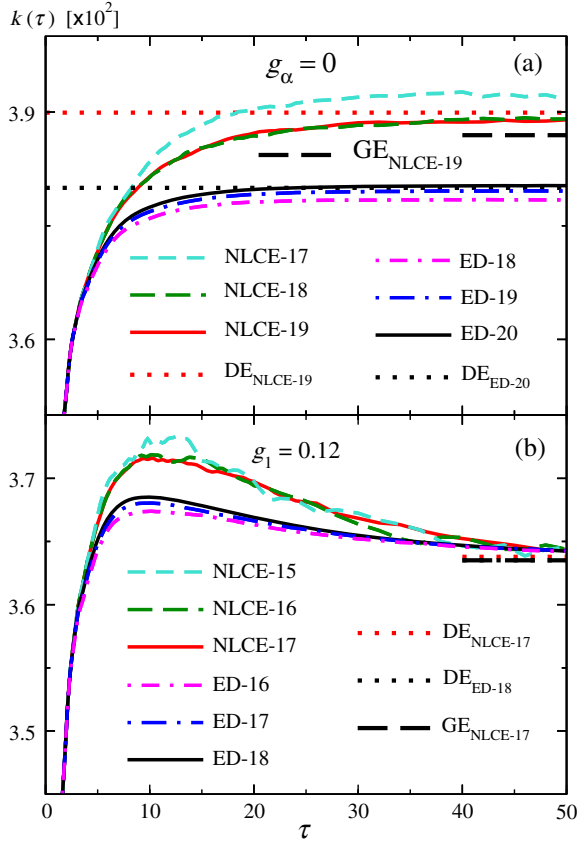
FIG. 11. Time evolution of the one-body nearest neighbor correlation $k(\tau)$ obtained in the last three orders l of the NLCE (NLCE-l in the legends), and in the three largest chains with L sites and periodic boundary conditions diagonalized (ED-L in the legends). Results for the diagonal ensemble are shown for the last order of the NLCE (DE$_{NLCE-l}$) and the largest chain diagonalized (DE$_{ED-L}$), while results for the grand canonical ensemble are shown for the last order of the NLCE (GE$_{NLCE-l}$), which are converged to machine precision. (a) Quenches $\hat{H}_I \to \hat{H}_0$, 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$_{NLCE-19}$) and $L = 20$ (DE$_{ED-20}$), and the grand canonical ensemble for $l = 19$ (GE$_{NLCE-19}$). (b) Quenches $\hat{H}_I \to \hat{H}_1$, 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, ED-19), along with the diagonal ensembles for $l = 17$ (DE$_{NLCE-17}$) and $L = 18$ (DE$_{ED-18}$), and the grand canonical ensemble for $l = 17$ (GE$_{NLCE-17}$). 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.

the NLCE and exact diagonalization results split and each equilibrate to the prediction of the corresponding diagonal ensemble, $k_{l=19}^{DE}$ (DE$_{NLCE-19}$) and $k_{L=20}^{DE}$ (DE$_{ED-20}$), respectively. As in Ref. [23], the NLCE results for the diagonal ensemble (DE$_{NLCE-19}$) are closer to the grand canonical ensemble ones, $k_{l=19}^{GE}$ (GE$_{NLCE-19}$), than the exact diagonalization results for the diagonal ensemble

(DE$_{ED-20}$). 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 $5 \lesssim \tau \lesssim 15$, $k_{l=16}(\tau)$ and $k_{l=17}(\tau)$ are almost indistinguishable from each other, while $k_L(\tau)$ 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. 11(b), we show dynamics after the quench $\hat{H}_I \to \hat{H}_1$ ($g_1 = 0.12$) for $k_l(\tau)$ with $l = 15$, 16, and 17 (the latter two are also reported in Fig. 3) and for $k_L(\tau)$ with $L = 16$, 17, and 18. $k_l(\tau)$ and $k_L(\tau)$ are almost indistinguishable from each other up to times $\tau \approx 4$ (again, the earliest times are not shown to gain dynamical range in the y axis). For $\tau \gtrsim 4$, the NLCE and exact diagonalization results again split, as for the quench $\hat{H}_I \to \hat{H}_0$, despite the fact that both equilibrate to diagonal ensemble results that are very close to each other. $k_{l=17}^{DE}$ (DE$_{NLCE-17}$) and $k_{L=18}^{DE}$ (DE$_{ED-18}$) in Fig. 11(b) are almost indistinguishable from each other and from the grand canonical ensemble result $k_{l=17}^{GE}$ (GE$_{NLCE-17}$). At times $5 \lesssim \tau \lesssim 15$, $k_{l=16}(\tau)$ and $k_{l=17}(\tau)$ are almost indistinguishable, while $k_L(\tau)$ shifts upward toward the NLCE predictions with increasing L. As for the quenches $\hat{H}_I \to \hat{H}_0$, these results suggest that NLCE is more accurate than exact diagonalization at intermediate times in quenches $\hat{H}_I \to \hat{H}_1$.

The discrepancy between the NLCE and exact diagonalization results at intermediate times after quenches $\hat{H}_I \to \hat{H}_1$ is likely a manifestation of finite-size effects in $k_L(\tau)$ that result from the "prethermal" dynamics seen in Fig. 11(a), which equilibrates to a diagonal ensemble value ($k_{L=20}^{DE}$) that is lower than the expected grand canonical ensemble one. The results in Fig. 11(b) also make apparent that, because of finite-size effects, the thermalization rates for $k(\tau)$ 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):

$$f_{\Delta E}(\tau) = \frac{2\pi g_\alpha^2}{L \Delta E} \sum_i P_i(\tau)$$
$$\times \sum_{|E_j - E_i| \leq \Delta E/2} |\langle E_j^0|\hat{V}_\alpha|E_i^0\rangle|^2 (N_j - N_i), \quad (D1)$$
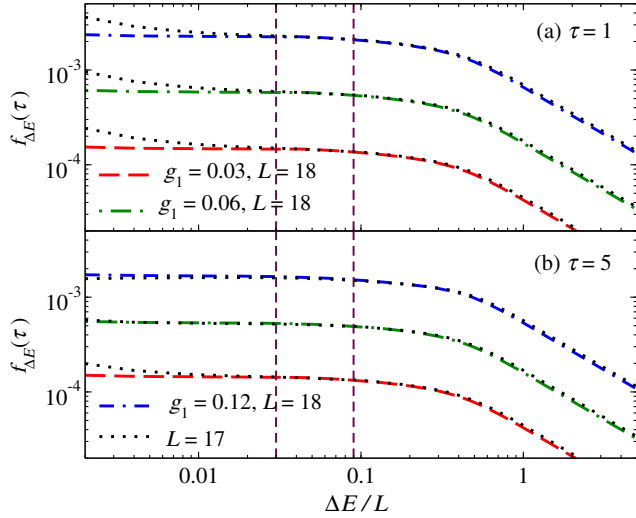
FIG. 12. $f_{\Delta E}(\tau)$, see Eq. (D1), evaluated for a wide range of energy windows ΔE for quenches $\hat{H}_I \to \hat{H}_1$ [see Eq. (36)] with $g_1 = 0.03$, 0.06, and 0.12. $f_{\Delta E}(\tau)$ is computed using full exact diagonalization of chains with $L = 18$ (see legends) and $L = 17$ (dotted 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.

FIG. 13. $\Gamma_{\Delta E,\tau}(g_1)$ [see Eq. (D2)] evaluated for quenches $\hat{H}_I \to \hat{H}_1$ [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 $\tau = 1, 1.5, \ldots, 5$ (nine 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.

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 $P_i(\tau) = \langle E_i^0|\hat{\rho}(\tau)|E_i^0\rangle$.

In Fig. 12, we plot $f_{\Delta E}(\tau)$, evaluated in chains with $L = 17$ and $L = 18$ for quenches $\hat{H}_I \to \hat{H}_1$ [see Eq. (36)], as a function of ΔE. We show results at two times, $\tau = 1$ in Fig. 12(a) and $\tau = 5$ in Fig. 12(b), and for three values of g_1 at each time. Our main finding in Fig. 12 is that, in the interval $\Delta E/L \in [0.03, 0.09]$, $f_{\Delta E}(\tau)$ is nearly independent of ΔE, 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 0.01%.

Having identified an appropriate range of values of ΔE, we compute the rate

$$\Gamma_{\Delta E,\tau}(g_\alpha) = -\frac{f_{\Delta E}(\tau)}{n(\tau) - n_{DE}}, \qquad (D2)$$

which is defined following Eq. (54).

Figure 13 shows $\Gamma_{\Delta E,\tau}(g_1)$ versus τ 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 τ, 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. 11, which exhibit a plateaulike behavior discussed in Sec. VI A, the rates $\Gamma^{Fermi}(g_\alpha)$ for $\tau \leq 1$ are not meaningful. Hence, the rates $\Gamma^{Fermi}(g_\alpha)$ reported in Sec. VI A were obtained averaging $\Gamma_{\Delta E,\tau}(g_\alpha)$ over the results for $\tau = 1, 1.5, \ldots, 5$ (nine values), and for $\Delta E/L = 0.03, 0.032, \ldots, 0.09$ (31 values),
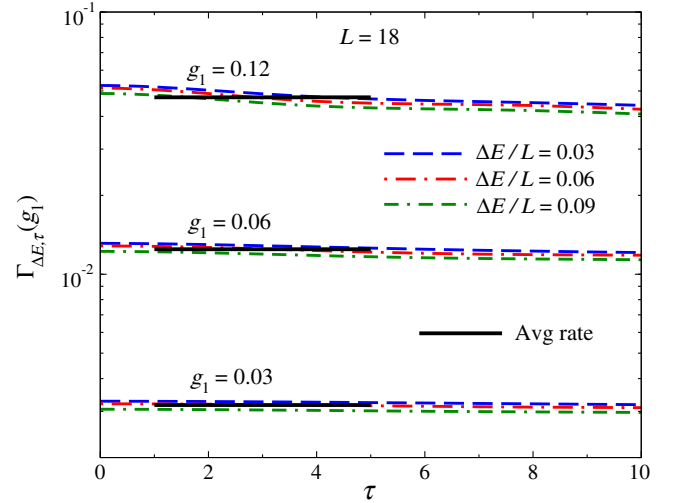
for a total of 279 values entering each average. In Fig. 13, we report the averages as horizontal lines. For each average, we also compute the standard deviation. In Fig. 1(c), we report the averages, and the standard deviations (as error bars), for different values of g_1 and for $L = 17$ and $L = 18$.

The rates $\Gamma^{Fermi}(g_2)$ reported in Fig. 7 were obtained averaging $\Gamma_{\Delta E,\tau}(g_2)$ over the results for $\tau = 1, 1.5, \ldots, 5$ (nine values), and for $\Delta E/L = 0.100, 0.102, \ldots, 0.15$, $0.16, \ldots, 0.20$ (31 values), for a total of 279 values entering each average.

[1] J. Berges, S. Borsányi, and C. Wetterich, *Prethermalization*, Phys. Rev. Lett. **93**, 142002 (2004).

[2] T. Kinoshita, T. Wenger, and D. S. Weiss, *A Quantum Newton's Cradle*, Nature (London) **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, *Relaxation and Prethermalization in an Isolated Quantum System*, 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, *Experimental Observation of a Generalized Gibbs Ensemble*, Science **348**, 207 (2015).

[5] T. Langen, T. Gasenzer, and J. Schmiedmayer, *Prethermalization and Universal Dynamics in Near-Integrable Quantum Systems*, J. Stat. Mech. (2016) 064009.

[6] Y. Tang, W. Kao, K.-Y. Li, S. Seo, K. Mallayya, M. Rigol, S. Gopalakrishnan, and B. L. Lev, *Thermalization Near Integrability in a Dipolar Quantum Newton's Cradle*, Phys. Rev. X **8**, 021030 (2018).

[7] M. Moeckel and S. Kehrein, *Interaction Quench in the Hubbard Model*, Phys. Rev. Lett. **100**, 175702 (2008); *Real-Time Evolution for Weak Interaction Quenches in Quantum Systems*, Ann. Phys. (Amsterdam) **324**, 2146 (2009).

[8] M. Eckstein, M. Kollar, and P. Werner, *Thermalization after an Interaction Quench in the Hubbard Model*, Phys. Rev. Lett. **103**, 056403 (2009).

[9] M. Kollar, F. A. Wolf, and M. Eckstein, *Generalized Gibbs Ensemble Prediction of Prethermalization Plateaus and Their Relation to Nonthermal Steady States in Integrable Systems*, Phys. Rev. B **84**, 054304 (2011).

[10] M. Tavora and A. Mitra, *Quench Dynamics of One-Dimensional Bosons in a Commensurate Periodic Potential: A Quantum Kinetic Equation Approach*, Phys. Rev. B **88**, 115144 (2013).

[11] N. Nessi, A. Iucci, and M. A. Cazalilla, *Quantum Quench and Prethermalization Dynamics in a Two-Dimensional Fermi Gas with Long-Range Interactions*, Phys. Rev. Lett. **113**, 210402 (2014).

[12] F. H. L. Essler, S. Kehrein, S. R. Manmana, and N. J. Robinson, *Quench Dynamics in a Model with Tuneable Integrability Breaking*, Phys. Rev. B **89**, 165104 (2014).

[13] B. Bertini, F. H. L. Essler, S. Groha, and N. J. Robinson, *Prethermalization and Thermalization in Models with Weak Integrability Breaking*, Phys. Rev. Lett. **115**, 180601 (2015); *Thermalization and Light Cones in a Model with Weak Integrability Breaking*, Phys. Rev. B **94**, 245117 (2016).

[14] M. Fagotti and M. Collura, *Universal Prethermalization Dynamics of Entanglement Entropies after a Global Quench*, arXiv:1507.02678.

[15] P. Reimann and L. Dabelow, *Typicality of Prethermalization*, Phys. Rev. Lett. **122**, 080603 (2019).

[16] M. Rigol, V. Dunjko, V. Yurovsky, and M. 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 Hard-Core Bosons*, Phys. Rev. Lett. **98**, 050405 (2007).

[17] L. Vidmar and M. Rigol, *Generalized Gibbs Ensemble in Integrable Lattice Models*, J. Stat. Mech. (2016) 064007.

[18] F. H. L. Essler and M. Fagotti, *Quench Dynamics and Relaxation in Isolated Integrable Quantum Spin Chains*, J. Stat. Mech. (2016) 064002.

[19] M. A. Cazalilla and M.-C. Chung, *Quantum Quenches in the Luttinger Model and Its Close Relatives*, J. Stat. Mech. (2016) 064004.

[20] J.-S. Caux, *The Quench Action*, J. Stat. Mech. (2016) 064006.

[21] M. Stark and M. Kollar, *Kinetic Description of Thermalization Dynamics in Weakly Interacting Quantum Systems*, arXiv:1308.1610.

[22] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *From Quantum Chaos and Eigenstate Thermalization to Statistical Mechanics and Thermodynamics*, Adv. Phys. **65**, 239 (2016).

[23] K. Mallayya and M. Rigol, *Quantum Quenches and Relaxation Dynamics in the Thermodynamic Limit*, Phys. Rev. Lett. **120**, 070603 (2018).

[24] A. Lazarides, A. Das, and R. Moessner, *Periodic Thermodynamics of Isolated Quantum Systems*, Phys. Rev. Lett. **112**, 150401 (2014).

[25] E. Canovi, M. Kollar, and M. Eckstein, *Stroboscopic Prethermalization in Weakly Interacting Periodically Driven Systems*, Phys. Rev. E **93**, 012130 (2016).

[26] L. D'Alessio and M. Rigol, *Long-Time Behavior of Isolated Periodically Driven Interacting Lattice Systems*, Phys. Rev. X **4**, 041048 (2014).

[27] A. Lazarides, A. Das, and R. Moessner, *Equilibrium States of Generic Quantum Systems Subject to Periodic Driving*, Phys. Rev. E **90**, 012110 (2014).

[28] M. Bukov, L. D'Alessio, and A. Polkovnikov, *Universal High-Frequency Behavior of Periodically Driven Systems: From Dynamical Stabilization to Floquet Engineering*, Adv. Phys. **64**, 139 (2015).

[29] D. A. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, *Effective Hamiltonians, Prethermalization, and Slow Energy Absorption in Periodically Driven Many-Body Systems*, Phys. Rev. B **95**, 014112 (2017).

[30] D. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, *A Rigorous Theory of Many-Body Prethermalization for Periodically Driven and Closed Quantum Systems*, Commun. Math. Phys. **354**, 809 (2017).

[31] T. Mori, T. Kuwahara, and K. Saito, *Rigorous Bound on Energy Absorption and Generic Relaxation in Periodically Driven Quantum Systems*, Phys. Rev. Lett. **116**, 120401 (2016).

[32] T. Kuwahara, T. Mori, and K. Saito, *Floquet-Magnus Theory and Generic Transient Dynamics in Periodically Driven Many-Body Quantum Systems*, Ann. Phys. (Amsterdam) **367**, 96 (2016).

[33] F. Machado, G. D. Meyer, D. V. Else, C. Nayak, and N. Y. Yao, *Exponentially Slow Heating in Short and Long-Range Interacting Floquet Systems*, arXiv:1708.01620.

[34] Z. Lenarčič, F. Lange, and A. Rosch, *Perturbative Approach to Weakly Driven Many-Particle Systems in the Presence of Approximate Conservation Laws*, Phys. Rev. B **97**, 024302 (2018).

[35] F. Lange, Z. Lenarčič, and A. Rosch, *Time-Dependent Generalized Gibbs Ensembles in Open Quantum Systems*, Phys. Rev. B **97**, 165138 (2018).

[36] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and Its Mechanism for Generic Isolated Quantum Systems*, Nature (London) **452**, 854 (2008).

[37] J. M. Deutsch, *Quantum Statistical Mechanics in a Closed System*, Phys. Rev. A **43**, 2046 (1991).

[38] M. Srednicki, *Chaos and Quantum Thermalization*, Phys. Rev. E **50**, 888 (1994).

[39] A. C. Cassidy, C. W. Clark, and M. Rigol, *Generalized Thermalization in an Integrable Lattice System*, Phys. Rev. Lett. **106**, 140405 (2011).

[40] J.-S. Caux and F. H. L. Essler, *Time Evolution of Local Observables After Quenching to an Integrable Model*, Phys. Rev. Lett. **110**, 257203 (2013).

[41] T. Ishii and T. Mori, *Strong Eigenstate Thermalization within a Generalized Shell in Noninteracting Integrable Systems*, arXiv:1810.00625.

[42] R. Zwanzig, *Ensemble Method in the Theory of Irreversibility*, J. Chem. Phys. **33**, 1338 (1960).

[43] H. Mori, *Transport, Collective Motion, and Brownian Motion*, Prog. Theor. Phys. **33**, 423 (1965).

[44] J. Lux, J. Müller, A. Mitra, and A. Rosch, *Hydrodynamic Long-Time Tails after a Quantum Quench*, Phys. Rev. A **89,** 053608 (2014).

[45] E. B. Davies, *Markovian Master Equations*, Commun. Math. Phys. **39,** 91 (1974).

[46] W. De Roeck and A Kupiainen, *Diffusion for a Quantum Particle Coupled to Phonons in $d \geq 3$*, Commun. Math. Phys. **323,** 889 (2013).

[47] J. Lukkarinen and H. Spohn, *Kinetic Limit for Wave Propagation in a Random Medium*, Arch. Ration. Mech. Anal. **183,** 93 (2007).

[48] M. Rigol, *Breakdown of Thermalization in Finite One-Dimensional Systems*, Phys. Rev. Lett. **103,** 100403 (2009).

[49] L. F. Santos and M. Rigol, *Onset of Quantum Chaos in One-Dimensional Bosonic and Fermionic Systems and Its Relation to Thermalization*, Phys. Rev. E **81,** 036206 (2010).

[50] M. A. Cazalilla, R. Citro, T. Giamarchi, E. Orignac, and M. Rigol, *One Dimensional Bosons: From Condensed Matter Systems to Ultracold Gases*, Rev. Mod. Phys. **83,** 1405 (2011).

[51] I. G. White, B. Sundar, and K. R. A. Hazzard, *Quantum Dynamics from a Numerical Linked Cluster Expansion*, arXiv:1710.07696.

[52] E. Guardado-Sanchez, P. T. Brown, D. Mitra, T. Devakul, D. A. Huse, P. Schauß, and W. S. Bakr, *Probing the Quench Dynamics of Antiferromagnetic Correlations in a 2D Quantum Ising Spin System*, Phys. Rev. X **8,** 021069 (2018).

[53] M. Rigol, T. Bryant, and R. R. P. Singh, *Numerical Linked-Cluster Approach to Quantum Lattice Models*, Phys. Rev. Lett. **97,** 187202 (2006); *Numerical Linked-Cluster Algorithms. I. Spin Systems on Square, Triangular, and Kagomé Lattices*, Phys. Rev. E **75,** 061118 (2007).

[54] M. Rigol, *Quantum Quenches in the Thermodynamic Limit*, Phys. Rev. Lett. **112,** 170601 (2014); *Fundamental Asymmetry in Quenches Between Integrable and Nonintegrable Systems*, Phys. Rev. Lett. **116,** 100601 (2016).

[55] K. Mallayya and M. Rigol, *Numerical Linked Cluster Expansions for Quantum Quenches in One-Dimensional Lattices*, Phys. Rev. E **95,** 033302 (2017).

[56] I. Bloch, J. Dalibard, and W. Zwerger, *Many-Body Physics with Ultracold Gases*, Rev. Mod. Phys. **80,** 885 (2008).

[57] M. Marcuzzi, J. Marino, A. Gambassi, and A. Silva, *Prethermalization in a Nonintegrable Quantum Spin Chain after a Quench*, Phys. Rev. Lett. **111,** 197203 (2013); *Prethermalization from a Low-Density Holstein-Primakoff Expansion*, Phys. Rev. B **94,** 214304 (2016).