

Prethermalization, Shadowing Breakdown, and the Absence of Trotterization Transition in Quantum Circuits

Marko Žnidarič

*Physics Department, Faculty of Mathematics and Physics, University of Ljubljana,
1000 Ljubljana, Slovenia*



(Received 27 May 2025; revised 12 November 2025; accepted 17 March 2026; published 22 April 2026)

One of the premier utilities of present-day noisy quantum computers is simulation of many-body quantum systems. We study how long in time is such a discrete-time simulation representative of a continuous time Hamiltonian evolution; namely, a finite time step introduces so-called Trotterization errors. We demonstrate that the truncated operator propagator (Ruelle-Pollicott resonances) is a powerful tool to that end, as well as to study prethermalization and discrete-time crystals, including finding those phenomena at large gate duration. We show that the effective energy is more stable than suggested by Trotter errors—a manifestation of prethermalization—while all other observables are not. Even the most stable observable though deteriorates in the thermodynamic limit. Different than in classical systems with the strongest chaos, where the faithfulness time (the shadowing time) can be infinite, in quantum many-body chaotic systems it is finite. A corollary of our results is also that, opposite of previous claims, there is no Trotterization transition in nonintegrable many-body quantum systems. We demonstrate our results on a one-dimensional (1D) kicked Ising model, as well as on a 1D kicked XX model and a 2D kicked Ising model. The truncated propagator is also used to calculate the energy diffusion constant in the tilted-field Ising model with high accuracy.

DOI: [10.1103/69d7-83xh](https://doi.org/10.1103/69d7-83xh)

Subject Areas: Nonlinear Dynamics,
Quantum Physics,
Statistical Physics

I. INTRODUCTION

One of the defining properties of chaotic systems [1] is exponential complexity of their time evolution as quantified by positive dynamical entropies. A manifestation of that is, for instance, the colloquial “butterfly effect”—two close initial conditions will separate exponentially fast in time. Considering that typical chaotic systems are not analytically solvable, and one therefore has to resort to the indignity of numerical simulations, a pertinent question is how representative are such noisy numerical trajectories of true trajectories? At first sight it would seem that the exponential separation of trajectories renders any noisy numerical simulation futile. However, it is not as grim.

First, we should specify what is it that we really want from a noisy simulation. Comparing a noisy and a true trajectory both starting from the same specific initial condition is not what we should in most cases care about.

This is especially true in a many-body system, where we do not have access to all degrees of freedom, and in quantum mechanics where the stress is on (ensemble) averages rather than on individual pure states. For instance, calculating the overlap of two states, often called the fidelity (also known as the Loschmidt echo), is not the most informative quantity in a many-body system (unless one is, e.g., interested in the ground-state physics of a gapped system): Small fidelity does not necessarily mean that few-body observables that we care about have gone away. Furthermore, the properties of a specific exact initial state are really irrelevant; measurable quantities away from phase transitions should be robust to small variations. What one should care about is that the noisy evolution of a particular initial state is still *representative* of a true non-noisy evolution, i.e., that its results are similar to an exact average. It would be enough for the noisy evolution to be similar to a true evolution from some nearby initial condition. Precisely that property is captured by the famous shadowing property of chaotic trajectories [1] that comes to our rescue when doing noisy simulations of chaotic systems.

Shadowing means that, while it is true that a trajectory $f_\epsilon^t(\mathbf{x}_0)$ starting from $\mathbf{x}_0$ will in the presence of noise ϵ diverge exponentially fast from a true trajectory $f_0^t(\mathbf{x}_0)$

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

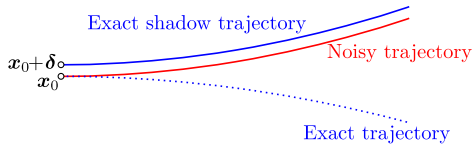


FIG. 1. Shadowing in classical chaotic systems. A true shadow trajectory (blue) from a slightly perturbed initial condition closely follows (shadows) the noisy trajectory (red), that otherwise exponentially separates from the original exact trajectory (dotted blue).

starting from $\mathbf{x}_0$, there exists a nearby initial condition $\mathbf{x}'_0 = \mathbf{x}_0 + \delta$ for which a true trajectory $f^t_0(\mathbf{x}'_0)$ closely follows the original noisy trajectory, $f^t_0(\mathbf{x}'_0) \approx f^t_\epsilon(\mathbf{x}_0)$. In other words, while a noisy trajectory is not a good representative of an exact trajectory, it closely follows a “shadow trajectory” that starts from a slightly perturbed initial condition (Fig. 1). The shadowing time is the time up to which one can find the shadow trajectory that follows the noisy trajectory. For uniformly hyperbolic [2] systems the above intuitive picture can be made rigorous in the form of the shadowing lemma [3,4], saying that the shadowing time in such systems is infinite. Uniformly hyperbolic systems are in a way the nicest chaotic systems (strongest chaos), an example being the Arnold’s cat map, in which noisy trajectories are representative of exact ensemble averages. For classical systems that are not uniformly hyperbolic, the shadowing time is finite [5].

We address the question of the shadowing time in quantum many-body systems. This is an important theoretical question, that is moreover very relevant for noisy quantum computers and simulators, whose best-case present use [6] is to simulate [7] other quantum systems [8–12], including the prethermalization regime [13] that we shall discuss. If one wants to demonstrate the supremacy of quantum computation, the number of qubits has to be large, and the simulation has to be faithful to true dynamics. Simulating the many-body quantum Hamiltonian $H_0 = \sum_j h_j$ with local interactions h_j can be implemented by splitting the time evolution propagator,

$$U(T) = \exp(-iH_0T), \quad (1)$$

into small time steps of duration τ , thereby approximating the Hamiltonian evolution by a quantum circuit,

$$U(T) \approx U^t_\tau, \quad U_\tau = \prod_j \exp(-ih_j\tau), \quad t = \frac{T}{\tau}, \quad (2)$$

where each term $\exp(-ih_j\tau)$ is now, for instance, a two-qubit gate that can be implemented on a quantum computer. A natural question is how is with the shadowing property in such many-body quantum systems; how long can a given simulation [rhs of Eq. (2)] be trusted to represent the true dynamics [lhs of Eq. (2)]?

Characterizing shadowing properties of a many-body quantum system is a formidable task, first because of all the intricacies that quantum mechanics brings, and second because of the system size limit $L \rightarrow \infty$, where even for classical systems things are not clear. We will therefore tackle a simpler question. We will not address the full shadowing, which is a statement on closeness of states, i.e., of all observables, but rather a question of direct experimental relevance that will nevertheless shed light also on the shadowing time. We will (i) focus on a specific type of noise, namely, errors [14–16] due to Trotterizing Hamiltonian evolution Eq. (2), and (ii) measure closeness by simple expectation values or autocorrelation functions of local observables (rather than some state norm).

A. Summary of results

We will find that even in the optimal situation—looking at the best-behaving observable—a noisy simulation follows a true simulation only up to a finite but long time, meaning that the shadowing time is necessarily finite. More specifically, under Trotterization errors there is a special observable, namely, the energy H_0 , that is very stable and resilient to errors, more than the simple Trotter error $\mathcal{O}(\tau^2)$ would suggest. Despite its stability though, with increasing system size L the stability will deteriorate. While in a finite system it might look that H_0 is stable, in the thermodynamic limit (TDL) of $L \rightarrow \infty$ and time $t \rightarrow \infty$, no matter how small τ , the energy H_0 will not be conserved, implying that the noisy simulation is fundamentally different than the true one. As a consequence, perfect infinite-time shadowing, like the one holding for classical systems with the strongest chaos, is impossible under Trotterization errors in locally interacting quantum many-body systems. The relevant timescale of the above instability is nothing but the prethermalization [17,18] time T (for a review of prethermalization, see Ref. [19]) and is given by the largest eigenvalue of the momentum-resolved truncated operator propagator. As predicted theoretically, it is exponentially large in the gate frequency $1/\tau$, but the benefit of our approach will be to give an exact quantitative prediction for T , not just a bound. The corresponding eigenvector is, on the other hand, an almost-conserved effective energy which is for small τ close to H_0 . Note that the prethermalization time and the shadowing time are different concepts. The shadowing time is related to the worst-case stability—thermalization of observable that deviates first, whereas we are focusing on the best-case scenario, i.e., thermalization of H_0 . An alternative viewpoint is that we are interested in the breaking of a single conserved operator H_0 due to Trotterization errors.

The truncated propagator therefore offers a way to obtain the effective Hamiltonian useful in studies of Floquet systems [20], and is an explicit and concrete alternative to various divergent operator expansions like the Baker-Campbell-Hausdorff (BCH) or the Floquet-Magnus series

[21]. We also point out and numerically demonstrate that in order to infer the correct behavior in the thermodynamic limit the Heisenberg time t_H (the inverse level spacing) has to be much larger than any relevant dynamical timescale in the system, specifically, the long prethermalization time T . It is another example when relying on small-system results, for instance, obtained by exact diagonalization, can lead to incorrect conclusions. Previous claims [16,22,23] of a small Trotter step transition (or a crossover) in system's properties, like the scaling of errors, were a finite-size effect caused by a finite averaging time (i.e., being in the incorrect limit of $T > t_H$). In the TDL there is no transition.

Another utility of the truncated propagator is to study discrete-time crystals (DTCs) [24–26]; see Ref. [27] for a review. Because we have access to exact T we will be able to address the question of the stability and the thermodynamic limit behavior of the DTC in a nearest-neighbor kicked Ising model away from its integrable $\tau = \pi$ point. We find that the prethermalization, underlying the robustness of DTC, is rather prevalent in the kicked Ising model and not limited only to small values of τ . Finally, looking at the quasimomentum dependence of the spectral gap, we will be able to calculate the diffusion constant, either during a finite but long prethermalization time or in the Hamiltonian limit. Because we directly work in the thermodynamic limit $L \rightarrow \infty$, the value of the energy diffusion constant for the chaotic Ising model is calculated with higher precision than with most other methods. We also briefly show results for the kicked XX model and a two-dimensional (2D) kicked Ising model, demonstrating generality of our results.

The truncated operator propagator method, even in our present brute-force implementation whose complexity grows exponentially with the truncated operator size r , offers a procedure that can treat difficult problems with long timescales, like prethermalization or diffusion, better than existing methods, obtaining quantitatively precise results in the thermodynamic limit with relatively modest numerical effort.

II. OBSERVABLES AND TRUNCATED OPERATOR PROPAGATOR

As our criterion of closeness of U to its “noisy” approximation U_τ^t , we shall look at how close are either expectation values of observables, or more generally, correlation functions of observables. Expectation values will be evaluated at infinite temperature,

$$\langle A \rangle \equiv \text{tr}(A)/2^L, \quad (3)$$

where 2^L is the Hilbert space size of L qubits. Compared to using averaging over a specific state, such an infinite-temperature average over density operator $\rho = \mathbb{1}/2^L$ is unbiased, easier to treat theoretically, and is also experimentally relevant. Because of self-averaging it can be

replaced at large L by an average over a single random state $|\psi\rangle$, $\langle A \rangle \approx \langle \psi | A | \psi \rangle$. More informative than a simple expectation value is the autocorrelation function $C_A(t)$ of observable A ,

$$C_A(t) \equiv \frac{1}{L} \langle A(0)A(t) \rangle, \quad (4)$$

where we will be interested in A 's that are extensive sums of local densities,

$$A = \sum_j a_j, \quad (5)$$

such that the prefactor $1/L$ in Eq. (4) ensures that $C_A(t)$ becomes system size independent for large L . Concretely, we shall take for A either the total magnetization $Z = \sum_j \sigma_j^z$, the energy H_0 , or the staggered magnetization $S = \sum_j (-1)^j \sigma_j^z$.

To evaluate $C_A(t)$ we need the time evolved observable $A(t) = U_\tau^{-t} A U_\tau^t$, which can be written in the Hilbert space of operators as $|A(t)\rangle\rangle = \mathcal{M}^t |A\rangle\rangle$, where $\mathcal{M}$ is a linear (super)operator propagator that propagates observables (operators) by one step (one U_τ). With this notation the correlation function is simply

$$C_A(t) = \langle\langle A | \mathcal{M}^t | A \rangle\rangle, \quad (6)$$

where the Hilbert-Schmidt inner product is $\langle\langle A | B \rangle\rangle \equiv \text{tr}(A^\dagger B)/\text{tr}(\mathbb{1})$. The expectation value is, on the other hand, $\langle A(t) \rangle = \langle\langle \mathbb{1} | \mathcal{M}^t | A \rangle\rangle$. In systems that have a mixing property (connected) correlation functions must decay to zero, while in chaotic ones without conserved operators one moreover expects this decay to be exponential for local A . One might wonder how can one obtain decaying correlation functions from the properties of $\mathcal{M}$? This is not obvious because $\mathcal{M}$ is unitary and therefore its spectrum is on a unit circle—one would expect that, e.g., exponential decay of $C_A(t)$ will be according to Eq. (6) connected to eigenvalues of $\mathcal{M}$ that satisfy $|\lambda| < 1$. Unitarity of $\mathcal{M}$, however, is not in contradiction with the apparent relaxation (equilibration): Namely, we expect relaxation of local observables and related decay of correlation functions; under unitary evolution no information is lost, it just gets spread out into the full Hilbert space, i.e., “disappears” in nonlocal operators.

In order to recover the decay of correlations from $\mathcal{M}$, one must look at analytical properties of its resolvent $(\mathcal{M} - z\mathbb{1})^{-1}$ inside a unit circle of complex z . This idea is known under the name of Ruelle-Pollicott (RP) resonances [28–30]. One way to get access to possible singularities inside a unit circle is to limit oneself to an appropriate functional space—in our case it will be the space of local operators.

A. Truncated propagator

The truncated operator propagator is the main technical tool of this work and will enable us to quantitatively study the prethermalization, DTC, and more, thereby resolving several outstanding questions.

One way to break the unitarity of $\mathcal{M}$ is to truncate it to operators with local support. Doing that one obtains a finite dimensional truncated $\mathcal{M}$ whose eigenvalues λ_j move inside a unit circle [31]. The largest one in modulus λ_1 will, provided it is isolated from the rest, determine the asymptotic decay of correlations as $C_A(t) \sim \lambda_1^t$, and is called a Ruelle-Pollicott resonance [28,29], a resonance inside a unit circle in the analytical continuation of the propagator resolvent. In Ref. [31] $\mathcal{M}$ was truncated to translationally invariant operators ($k = 0$ in our notation) and studied in the kicked Ising model. Besides chaotic systems, the method can also be used to find constants of motion [32], for instance, in Refs. [34,35] in cellular

automata and in Ref. [36] to discover that all magnetization conserving brickwall qubit circuits are integrable.

We shall use an extension of the above method to an arbitrary (quasi)momentum k [37], a refinement that has been already used to find new unconventional symmetries in integrable circuits [38]. Momentum dependence will allow us to treat any observable, not just translationally invariant ones, as well as deduce transport properties, for instance, the value of the diffusion constant. Namely, we are interested in translationally invariant H_0 without any spatial inhomogeneities, such that also its Trotterization U_τ will have homogeneous gates. In other words, U_τ commutes with a shift by one site operator S (the method can be generalized to systems invariant only under a shift by more than one site [39]), and therefore $\mathcal{M}$ is block diagonal in the eigenbasis of S . Blocks can be labeled by eigenvalues e^{ik} of S , where $k \in [-\pi, \pi]$ is a lattice quasimomentum (we shall often call it just momentum). The operator basis in a given momentum block consists of operators $A^{(\alpha)}$,

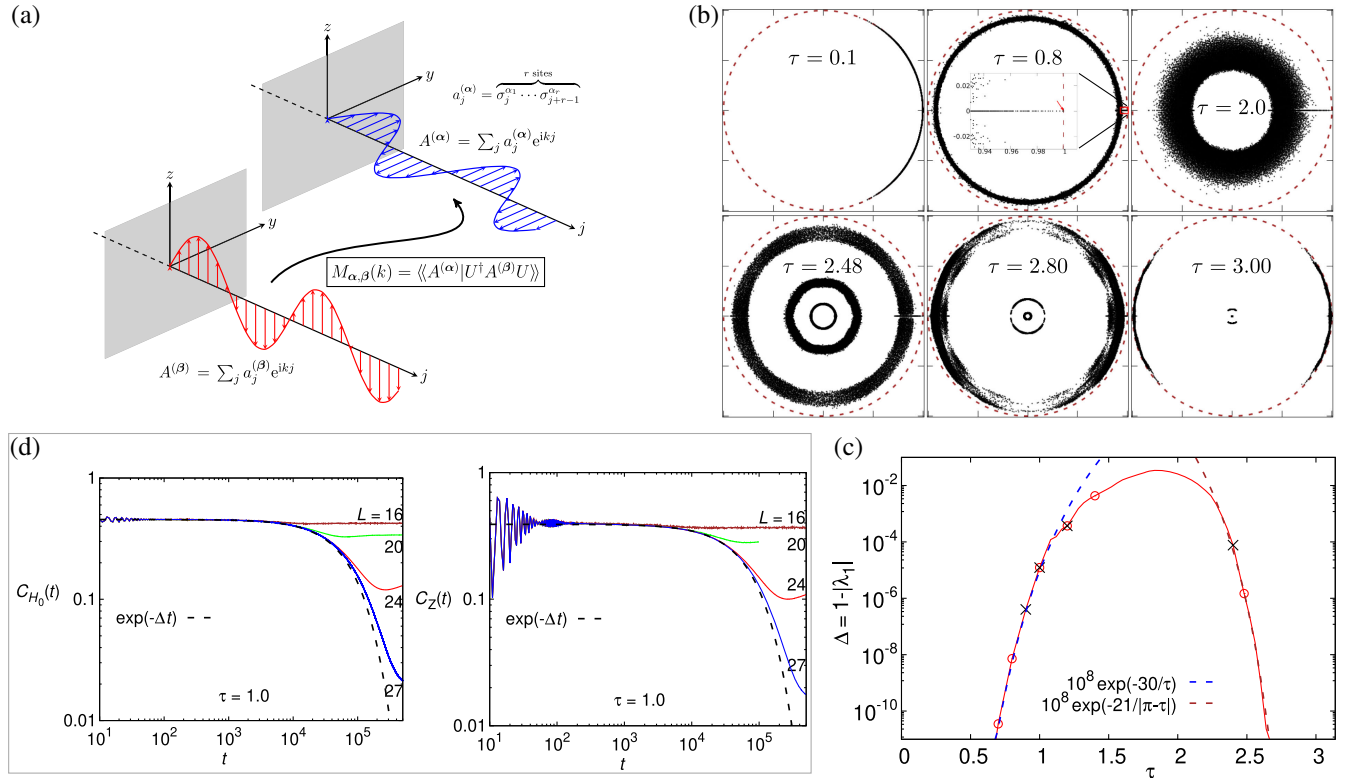


FIG. 2. Truncated propagator and Ruelle-Pollicott resonances. (a) Momentum-resolved operator propagator $M(k)$ is truncated to infinite system size operators with their density supported on at most r sites (sine-shaped curves suggest quasimomentum modulation of local densities, e.g., of $a_j^{(\alpha)} = \sigma_j^z$ having $r = 1$). (b) Spectrum of nonunitary $M(k = 0)$ lies inside a dashed unit circle. Shown are spectra at truncation $r = 8$ and different gate duration τ in the kicked Ising model Eq. (11). Spectrum can be used to identify prethermalization (e.g., $\tau = 0.8$) as well as a discrete-time crystal ($\tau = 3.0$). The largest eigenvalue (e.g., enlarged inset with red arrow in the top middle frame) is shown in (c). (c) Dependence of the spectral gap Δ on τ (full curve is for $r = 9$). Circles and crosses mark τ for which we show in other figures convergence with r , and correlation functions, respectively. (d) Autocorrelation function of energy H_0 Eq. (10) and magnetization $Z = \sum_j \sigma_j^z$ (right). Dotted curve is exponential decay with the rate Δ from (c), showing that the inverse gap $T = 1/\Delta$ gives the duration of the prethermalization plateau. In the thermodynamic limit $L \rightarrow \infty$ the plateau at large time disappears for any τ —there is no Trotterization transition with τ . Data are for the kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

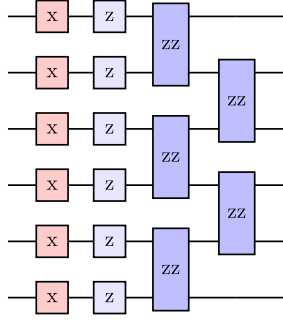


FIG. 3. Kicked Ising quantum circuit. Sequence of gates for one Floquet step of the kicked Ising model Eq. (11).

$$A^{(\alpha)} = \sum_j e^{ikj} a_j^{(\alpha)}, \quad a_j^{(\alpha)} = \sigma_j^{\alpha_1} \cdots \sigma_{j+r-1}^{\alpha_r}, \quad (7)$$

where the sum runs over all site indices j (in an infinite system), and $\sigma^{\alpha_j} \in \{\mathbb{1}, \sigma^x, \sigma^y, \sigma^z\}$ are Pauli matrices (the local operator basis) specified by a vector of indices $\alpha = (\alpha_1, \dots, \alpha_r)$. Matrix elements of the operator propagator are [see Fig. 2(a) for illustration]

$$M_{\alpha,\beta}(k) = \langle\langle A^{(\alpha)} | U_\tau^\dagger A^{(\beta)} U_\tau \rangle\rangle. \quad (8)$$

While it looks that in order to calculate $M(k)$ one has to deal with infinite-size operators $A^{(\alpha)}$, because of a sharp light cone in quantum circuits the initial local density $a_j^{(\alpha)}$ can in one step spread over at most a fixed number s of additional sites to the left or to the right of its initial support. For our specific kicked Ising circuit (Fig. 3) where U_τ contains only a single layer of commuting nearest-neighbor two-qubit gates, we have $s = 1$. It is therefore enough to work with propagation on just $r + 2$ sites. One can in fact implement evolution of local operators *in situ* using only r sites instead of $r + 2$, see Appendix A 2 for details, which enables us to study larger truncation sizes r , or simply speeding up calculations compared to tracking evolution on full $r + 2$ sites. Denoting $b(\beta) \equiv U_\tau^\dagger a_1^{(\beta)} U_\tau$, Eq. (8) explicitly evaluates to

$$M_{\alpha,\beta}(k) = \frac{1}{2^{r+2}} \text{tr}[(a_1^{(\alpha)} + a_0^{(\alpha)} e^{ik} + a_2^{(\alpha)} e^{-ik}) b(\beta)]. \quad (9)$$

Such an operator propagator has a symmetry $M^\dagger(U) = M(U^\dagger)$ and is without truncation a unitary operator (for $k = 0$ and Hermitian basis, like the Pauli basis Eq. (7), it is real orthogonal). If we truncate the basis to all local densities supported on at most r consecutive sites (1D systems), that is vector indices α and β run over $N = 3 \times 4^{r-1}$ possible local densities (3 instead of 4 is because the first index α_1 should not describe $\mathbb{1}$, otherwise the density would be supported on at most $r - 1$ sites, and could be obtained by a translation of other basis

elements), we get a truncated operator propagator that is the main workhorse of this work. Because of the truncation we are dropping some matrix elements of an infinite-size unitary operator, resulting in a finite non-unitary matrix of size $N \times N$. For more details on implementation, see Refs. [37,39], and for *in situ* propagation, see Appendix A 2.

B. Kicked Ising model

Properties of $M(k)$ [Eq. (9)] will allow us to answer all the questions that we highlighted in the Introduction. Let us demonstrate its spectral properties on the tilted field Ising model, which is a canonical many-body system that is chaotic for an appropriate parameter choice, e.g., $h_x = h_z = 1$ [40].

The Hamiltonian is

$$H_0 = J \sum_j \frac{\sigma_j^z \sigma_{j+1}^z}{2} + h_z \sigma_j^z + h_x \sigma_j^x. \quad (10)$$

Evolution $U(\mathcal{T})$ can be implemented on a quantum computer by splitting H_0 into two terms, one containing σ^z operators and the other σ^x . The one-step propagator shown in Fig. 3 is

$$\begin{aligned} U_\tau &\equiv U_z U_x, \\ U_x &= \prod_j e^{-i\tau h_x \sigma_j^x / 2}, \\ U_z &= e^{-i\tau \sum_j \sigma_j^z \sigma_{j+1}^z / 4} \prod_j e^{-i\tau h_z \sigma_j^z / 2}. \end{aligned} \quad (11)$$

Alternatively, the same U_τ can be thought of as being a propagator of a time-periodic Hamiltonian,

$$H = J \sum_j \left[\frac{\sigma_j^z \sigma_{j+1}^z}{2} + h_z \sigma_j^z + h_x \sigma_j^x \sum_{n=-\infty}^{\infty} \tau \delta(t - n\tau) \right], \quad (12)$$

called the kicked Ising model. RP resonances in the truncated propagator of the kicked Ising model have been studied in Refs. [31,41], and more recently in Ref. [37].

There are three parameters: τ , h_x , and h_z . For generic parameters the model is nonintegrable without any local conserved quantities. We measure time in units of τ , i.e., our t takes integer values, and we use parameters $J = \frac{1}{2}$, $h_x = h_z = 1$, the same as, e.g., Refs. [16,22]. Only in Sec. III D 2 on transport will we use different ones in order to compare with values of the diffusion constant at parameters found in the literature.

C. Spectrum of the truncated propagator

Let us briefly describe what the spectrum of $M(k)$ [Eq. (9)] looks like—in later sections we will describe in detail how it can serve as a useful tool to study Floquet

systems (and also Hamiltonian systems in an appropriate limit).

In Fig. 2(b) we show a set of exemplary spectra of $M(k)$ for $k = 0$ and the truncation size $r = 8$. Because we have $h_x = h_z = 1$ the spectrum $\sigma(M, \tau)$ of $M(k)$ has a symmetry $\sigma(M, \tau) = \sigma(M, 2\pi - \tau)$ [i.e., is symmetric around $\tau = \pi$, $\sigma(M, \pi + \tau) = \sigma(M, \pi - \tau)$], meaning that it is enough to study parameter range $\tau \in [0, \pi)$. In all cases the spectrum is within a unit circle, its form though heavily depends on the gate duration τ . At very small $\tau = 0.1$ the propagator U_τ is close to the identity and therefore all eigenvalues of $M(k)$ are close to the unit circle (within a band of width $1 - |\lambda_j| \sim \tau^2$). At small $\tau = 0.8$, which will be a regime of most interest to us, one has a ring of eigenvalues, with those closest to the unit circle lying on the real axis. Of particular interest is the largest one $|\lambda_1|$ because it will determine the asymptotic exponential decay rate of correlation functions [see the inset in Fig. 2(b)].

At $\tau > 2$ the spectrum qualitatively changes and is made out of three rings. Point $\tau = \pi$ is special because it is integrable (two otherwise different integrable points merge into one because of $h_x = h_z = 1$) [37,41]: Singular values of $M(k)$ are all 1 [$U_x = \prod_j (-i\sigma_j^x)$, $U_z = \prod_j (-i\sigma_j^z) \prod_p (1 - i\sigma_p^z \sigma_{p+1}^z)/\sqrt{2}$], while eigenvalues of $M(k)$ are $10 \times 4^{r-3}$ equal to $+1$, the same number is -1 , while the rest is 0 (i.e., a fraction $5/24$ of all is $+1$). The fact that nonzero eigenvalues for $\tau \rightarrow \pi$ bunch around ± 1 is a signature of a discrete-time crystal [24–26], specifically of the model discussed in, e.g., Ref. [43]. As we shall see, expectation values of some observables will exhibit oscillations with period 2 (Sec. III E). Spectrum of $M(k)$ will therefore tell us a lot of interesting things.

III. RESULTS

A. Leading Ruelle-Pollicott resonance and prethermalization

The most important quantity is the eigenvalue of $M(k)$ with the largest modulus. It determines the long-time decay rate of correlation functions of generic observables (those not orthogonal to the corresponding eigenvector) under evolution by circuit U_τ . Furthermore, as we shall see, it also determines the duration of the prethermalization phase, as well as a range of times and length scales on which one can observe transient diffusion during the prethermalization phase, being a Floquet remnant of energy diffusion in the Hamiltonian H_0 . Other methods have been used to find “slowest” observables, like minimizing the commutator with U_τ [44] (fastest decay at short times), see also Refs. [45,46], or using time averaging [47].

In Fig. 2(c) we show numerical results for the largest eigenvalue of $M(k)$, which turns out to be from $k = 0$ (due to reflection symmetry there are two parity subsectors at $k = 0$, and the shown λ_1 is from the even subsector; see also Ref. [37] for more details on parity sectors). We can see that

the gap (distance to the unit circle) is rather small, and λ_1 is always isolated from the second largest eigenvalue (will be shown later). We have checked that within the shown precision λ_1 is converged with respect to increasing the truncation support r (see Appendix A). We shall especially focus on small τ behavior where the gap decreases rapidly with increasing gate frequency $1/\tau$, being a manifestation of prethermalization [17,18] in which one has a parametrically long regime in which the system looks like being thermal with respect to an almost-conserved Hamiltonian; see Ref. [19] for a review. While one can prove generic bounds on the prethermalization time T [17,18,48,49], they do not give an explicit prediction that could be used in a concrete model. The truncated operator propagator provides that.

For small τ we can try to write the propagator U_τ in terms of the Floquet Hamiltonian H_F ,

$$U_\tau = e^{-iH_F\tau} \approx e^{-iH'_F\tau}. \quad (13)$$

The exact Floquet Hamiltonian H_F is in general nonlocal and cannot be obtained by, e.g., BCH expansion because the series is divergent; see Appendix B. However, up to prethermalization time T we can write the propagator approximately in terms of an almost-conserved operator H'_F . While one can in principle obtain H'_F via a BCH expansion, such a procedure is fraught with a certain freedom of what is the “optimal” truncation of an otherwise divergent series. The truncated operator propagator provides a more direct way— H'_F is nothing but the eigenvector corresponding to λ_1 .

One expects that the prethermalization duration T is exponentially large in $1/\tau$ [17,18,48,49]. A hand-waving argument is that if the frequency $1/\tau$ is larger than the local energy density h , it will take of order $\sim 1/(h\tau)$ simultaneous spin “flips” in order to absorb the energy $1/\tau$. Therefore, such a process will only occur at a high order $n \sim 1/\tau$ in the BCH expansion of H_F , leading to an exponentially small probability, and as a consequence to an exponentially long time T when the heating begins:

$$T \sim e^{a/\tau}. \quad (14)$$

The truncated operator propagator gives, importantly, exact quantitative prediction; namely,

$$T = 1/\Delta, \quad \Delta \equiv 1 - |\lambda_1|, \quad (15)$$

where λ_1 is the largest eigenvalue of $M(k)$. For the chosen parameters the obtained numerical dependence can be described well by $T \approx 10^{-8} \exp(30/\tau)$, see Fig. 2(c) (in the range $T \in [10^4, 10^{12}]$; for smaller τ the gap gets too small for double precision arithmetic, while we could not reach convergence at available r using higher precision arithmetic). The corresponding right eigenvector is equal to an almost-conserved operator H'_F (that could be

approximately calculated in principle using the BCH series; see Ref. [37] for demonstration).

That T is exactly the correct length of the prethermalization phase is shown by calculating autocorrelation of various observables numerically exactly in finite systems (see Appendix A for numerical details). We specifically focus on the energy H_0 [Eq. (10)] and the total magnetization Z . Both have a large overlap with the eigenvector corresponding to λ_1 , and should therefore decay very slowly with time t . This is visible in Fig. 2(d). The theoretical dashed line shown in plots is exponential decay,

$$C_A(t) \asymp B_A e^{-\Delta t}, \quad (16)$$

without any fitting parameters. Its decay rate is Δ read off from Fig. 2(c), while its amplitude B_A is calculated exactly to few lowest orders in τ using the BCH expansion, obtaining (for $J = 1/2$, $h_x = h_z = 1$)

$$B_{H_0} = \frac{9}{16} - \frac{3}{32}\tau^2 + \dots, \quad B_Z = \frac{4}{9} - \frac{4}{81}\tau^2 + \dots; \quad (17)$$

see Appendix B for details. We can see that this theory perfectly agrees with numerics, including at other values of τ (see Appendix A for data).

Important for our discussion of shadowing is behavior after prethermalization ends, $t \gg T$. We can see in Fig. 2(d) that at finite L a second lower plateau is reached, which, however, decreases to 0 as one increases L . In the thermodynamic limit, where one has to first send $L \rightarrow \infty$ and only then $t \rightarrow \infty$, there is therefore no plateau in correlation functions, nor in the magnetization itself. This is different than what has been observed in Refs. [16,22,23]—different conclusions there were a consequence of using a finite averaging time over which one calculated, e.g., magnetization [i.e., the averaging time was smaller than the prethermalization time T [Eq. (15)], or, using too small L for the chosen averaging time]. In the TDL of chaotic systems there is therefore no Trotterization transition due to mysterious “quantum localization.” What was observed in Ref. [16] is a simple manifestation of a single almost-conserved operator due to prethermalization. Let us remark that the Trotterization transition in integrable systems [50], for instance in XXZ-type integrable circuits, is of a completely different origin—it is a consequence of different phases between which the whole analytical integrability structure changes [36] and thereby also physical properties [38].

If the shadowing time for quantum many-body systems would be infinite, then the energy H_0 should be stable; i.e., there should exist a shadowing trajectory under noisy U_τ for which the expectation value of H_0 would stay close to the exactly constant value obtained under Hamiltonian evolution by U . Because this is not the case [Fig. 2(d)] we can conclude that the shadowing in quantum many-body systems under Trotter errors cannot be infinite. This is, in

fact, in line with expectations, classical and quantum. In quantum Floquet systems one can argue that eventually the system should “heat up” (see, e.g., Refs. [51,52]). In classical systems one, for instance, also finds that the shadowing time is finite, and in particular gets smaller if there is a large variation in the values of Lyapunov exponents [5,53]. Many-body systems, where in the classical case one has a whole spectrum of Lyapunov exponents, have in a way too much “freedom” to be uniformly hyperbolic, and therefore an infinite shadowing time.

B. Finite-size effects and the Heisenberg time

From data like those in Fig. 2(d) we can see that at smaller L , like $L = 16$ or 20 , even at large time autocorrelation functions do not decay but rather reach a second nonzero plateau value $C_A(\infty)$. Why is that? How large L do we need in order to observe the correct behavior in the TDL where correlation functions do decay? In a finite system the longest possible timescale is called the Heisenberg time and is given by the inverse level spacing. In a many-body system it has a leading exponential dependence on L , $t_H \sim 2^L$. At larger times $t \gg t_H$ quantum system resolves its discrete spectrum and essentially starts to behave like it is “integrable,” even if its true dynamics is chaotic (e.g., one has recurrences, correlations do not decay, etc., due to the spectral resolution being a sum of a finite number of frequencies). One should therefore never look at properties of quantum systems at $t \gg t_H$ —at best one will sample finite-size effects.

Being careful about t_H is especially important in systems with long timescales: For instance, in our case the prethermalization time T is exponentially large in $1/\tau$ [see, e.g., Fig. 2(c)]; another case is systems with disorder where dynamics gets exponentially slow with increasing disorder strength [54]. For instance, in Fig. 2(d) we can see that at $L = 16$ it indeed seems that even at $t \sim 10^6$ correlations do not decay, and one would be tempted to conclude that one is in a nonmixing regime. However, that is not correct. For larger L correlations do start to decay. The reason for nondecay at $L = 16$ is simply that one is not in the regime $T \ll t_H$. This is another example where relying on an exact diagonalization, either full spectrum or just an often-used diagonal approximation, would completely incorrectly diagnose the long-time dynamics in the thermodynamic limit [e.g., plateau in $L = 16$ data in Fig. 2(d)].

In order to see the correct physics in the thermodynamic limit, one should have

$$T = \frac{1}{\Delta} \ll t_H \sim 2^L. \quad (18)$$

We numerically demonstrate correctness of Eq. (18) in Fig. 4. From numerical simulation data we read out the value of $C_Z(t)$ at large time, where a second plateau is reached for finite L . Plotting this value $C_Z(t \rightarrow \infty)$ as a

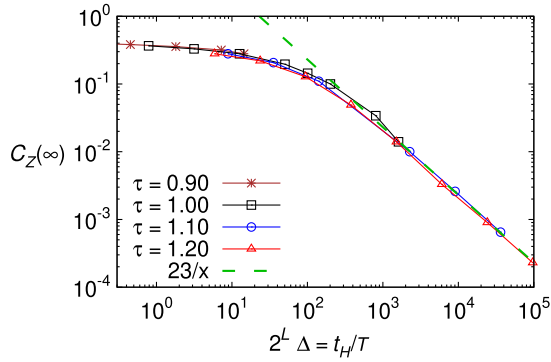


FIG. 4. Prethermalization vs Heisenberg time. Finite-size asymptotic value of magnetization correlation function [e.g., from Fig. 2(c) we can read $C_Z(\infty) \approx 0.1$ for $L = 24$ and $\tau = 1.0$, reached at $t > 10^5$] is shown for various sizes $L = 14$ – 28 and τ . To observe the correct thermodynamic limit $\lim_{t \rightarrow \infty} \lim_{L \rightarrow \infty}$, in which $C_Z(\infty)$ decays as $\propto 1/2^L$ (green line), one must have $T \ll t_H$ —the prethermalization time $T = 1/\Delta$ has to be smaller than the Heisenberg time $t_H \sim 2^L$. Data are for the kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

function of scaled time t_H/T , where T is read off from Fig. 2(c), we get a good collapse of data for all L and τ (i.e., T). We indeed see that in order to really see decaying correlations, where $C_Z(\infty) \sim 1/2^L$, one must have $t_H \gg T$. Different conclusions in Refs. [16,22,23] are due to violation of this condition. Regarding the stability of expectation of H_0 , the good news is that the prethermalization time is exponentially large in $1/\tau$ and so one does not need excessively small time step τ to observe long stability; however, if one wants to study the TDL, one should still respect $T \ll t_H$ and therefore also not take a too small τ at a given L ; the minimal meaningful time step τ scales in a finite system as $1/L$ [Eq. (18)].

While computational complexity of finding the largest eigenvalue of $M(k)$ in our implementation (see Appendix A) grows exponentially with the support r , i.e., the necessary CPU time increases by ≈ 4 when r is increased by 1, it is still much faster than any other known method used for prethermalization. For instance, taking as an example the kicked Ising at $\tau = 1.0$ in Fig. 2, exact pure-state evolution used to get the magnetization correlation function for $L = 27$ until $t = 5 \times 10^5$ took 660 h (≈ 27 days on a single node of a standard computational cluster), while our calculation of the gap Δ at $r = 10$ took ≈ 6 min with the *in situ* propagation (while it takes ≈ 9 h using slower propagation on $r + 2$ sites; see Appendix A for details). At $r = 10$ the result $\Delta = 1.20 \times 10^{-5}$ is essentially converged with respect to r (see Appendix A), and even at much smaller truncation $r = 6$, where the calculation takes only a few seconds, we already get an almost correct $\Delta = 1.33 \times 10^{-5}$. On the other hand, for pure-state evolution decreasing L is “dangerous”: $L = 24$ is barely enough [Fig. 2(d)] to give

us a hint of the correct prethermalization behavior (as discussed, this is the reason for previous incorrect claims) but still takes ≈ 4 days. Using exact diagonalization, which is only possible at even smaller systems like $L = 16$, would, on the other hand, give a completely incorrect physics at long times. In short, the truncated propagator method gives correct, high-precision results with much less numerical effort than other methods, while even at lower precision r , as opposed to other methods, it still identifies the correct physics (small Δ and prethermalization).

C. Fast decay and no prethermalization of generic observables

Here we want to answer the question, which observables do exhibit prethermalization? With a momentum-resolved propagator, which covers propagation of any observable, that is easy. We just have to check how many eigenvalues of $M(k)$ are exponentially close to a unit circle. As we shall see, there is only one.

Let us check if there are any other very small gaps of $M(k)$ which would imply that there is another almost-conserved local operator. Numerically calculating the second largest eigenvalue in the $k = 0$ sector, or more precisely, the second one in the even $k = 0$ and the largest in the odd $k = 0$ sector, we see in Fig. 5(a) that those eigenvalues are always separated from the prethermalization one $\lambda_1(k = 0)$. A similar situation occurs also for nonzero k . The spectral gap for all those eigenvalues scales as $\sim \tau^2$ for small time steps, which is the same as the scaling of the Trotterization error $\mathcal{O}(\tau^2)$. We can therefore conclude that H'_F is the only almost-conserved local operator for which the prethermalization occurs.

All other observables that are orthogonal to H'_F will deteriorate on timescales that are in line with Trotter errors, that is, much faster than the prethermalization time T . Note that while a typical observable from the even zero momentum sector will have a nonzero overlap with H'_F (which though, depending on what we mean by typical, i.e., the measure, can be small), any observable with a nonzero k will have zero overlap. Therefore, allowing observables with nonzero momentum a typical observable will not exhibit prethermalization. This is numerically demonstrated in Fig. 5(b). We pick a staggered magnetization $S \equiv \sum_j (-1)^j \sigma_j^z$ and observe that the correlation function for finite τ starts to differ from the one obtained for Hamiltonian evolution generated by H_0 much more rapidly than, e.g., H_0 or Z in Fig. 2(d) which do exhibit prethermalization. For the largest $\tau = 0.8$ shown, for which the staggered magnetization visibly deviates already at $t \approx 22$, the total magnetization Z would follow the Hamiltonian result up to very large prethermalization $T \approx 10^8$. A typical observable is therefore much more sensitive to Trotterization errors because it does not exhibit prethermalization.

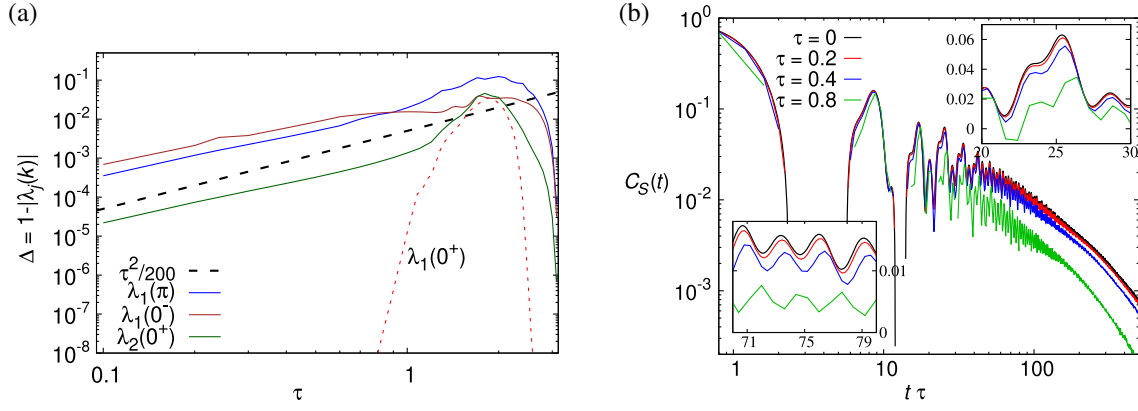


FIG. 5. Single almost-conserved operator. (a) Largest eigenvalues of $M(k)$ for different quasimomenta k : even and odd sector in $k = 0$ and $k = \pi$. Only one gap is very small [dashed red; also Fig. 2(c)], all other scale as $\sim \tau^2$ for small τ . (b) Correlation function of staggered magnetization $S = \sum_j (-1)^j \sigma_j^z$ deteriorates fast already for small τ ($L = 32$, finite-size effects are negligible on shown times; black curve is Hamiltonian evolution by H_0). The two insets are enlargements on specific time windows.

D. Diffusion constant from the truncated propagator

1. Finite gate times

Up to large time T the operator H'_F behaves as if it would be conserved. Similarly as one can study energy transport under Hamiltonian evolution by H_0 , one can also ask about transport properties of H'_F in the circuit U_τ for $t \ll T$ when H'_F is quasiconserved.

Assuming H_0 is chaotic, and thereby also H'_F , we would expect this transport to be diffusive. This is indeed the case and can be seen in the momentum dependence of the spectrum of $M(k)$. Details about the momentum-dependent spectrum of the truncated propagator in the presence of diffusion due to U(1) symmetry will be published elsewhere [39], for discussion of correlation functions see also recent Ref. [55], here we just look at $\lambda_1(k)$. Namely, the leading eigenvalue of $M(k)$ gives the asymptotic decay rate of correlation functions of operators with finite k . In particular, diffusion should be reflected in the decay rate of appropriate density correlations being Dk^2 . Such a

scaling can be obtained by Fourier transforming diffusion equation; or more explicitly, the correlation function of the diffusive conserved density q in momentum space decays as $\langle q(t, k)q(0, k) \rangle \sim e^{-Dk^2}$ [56].

We therefore have that the largest eigenvalue of the truncated operator propagator M at $k = 0$ determines the prethermalization duration $T = 1/\Delta = 1/(1 - |\lambda_1(k = 0)|)$, while its momentum dependence $1 - \lambda_1(k) \approx Dk^2$ determines diffusion constant D . This can be seen in Fig. 6, from which we could, for instance, read out that a finite-time D (until prethermalization ends) at $\tau = 0.6$ is $D \approx 0.13 \times 0.6 = 0.078$ (this is diffusion constant in units of time where one unit of time equals one propagation step of duration τ ; per unit of physical time, i.e., where one makes $1/\tau$ propagation steps in a unit of time, its value would instead be $D \approx 0.13$). We can in fact see that (i) one has quadratic dependence only at small k , i.e., diffusion is expectedly exhibited only on large hydrodynamic scales, and that (ii) k must also not be too small in order to see

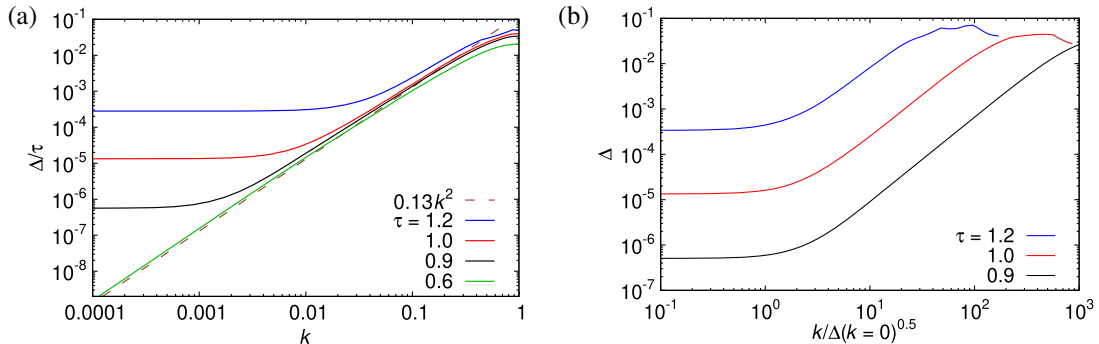


FIG. 6. Diffusion during prethermalization. (a) Dependence of the gap $\Delta = 1 - |\lambda_1(k)|$ on quasimomentum k [at $k = 0$ one gets Fig. 1(c), $r = 6$]. For small $k > k_0$ one has diffusive quadratic dependence $\Delta \approx D(\tau)\tau k^2$. (b) Scaled k : diffusion can be observed at quasimomenta larger than $k_0 \sim \Delta^{1/2}$. The prefactor $D(\tau \rightarrow 0)$ is the diffusion constant of Hamiltonian H_0 . Kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

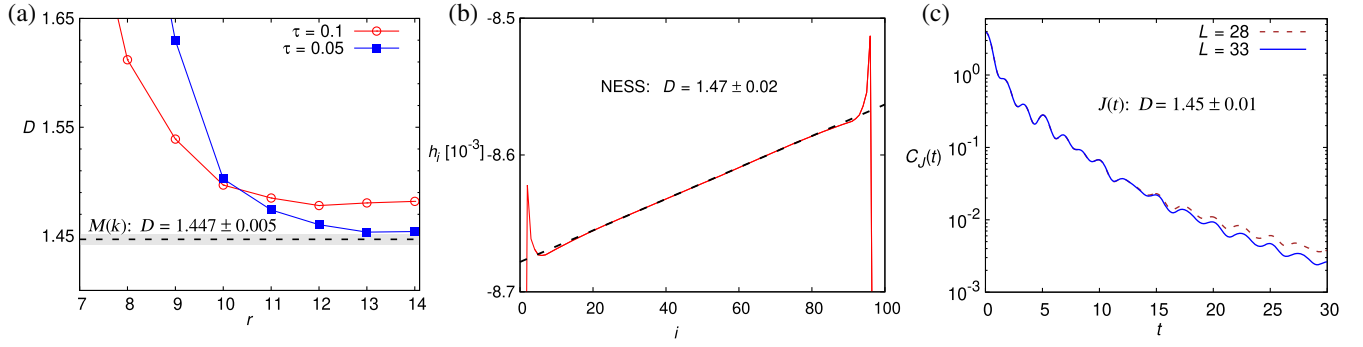


FIG. 7. Diffusion constant for the Hamiltonian tilted Ising model. (a) Using the truncated propagator in the limit of small τ and k , getting $D = [1 - \lambda_1(k)]/(k^2\tau)$ at $k = 0.02$. (b) Lindblad NESS with $L = 100$ and high-temperature boundary parameters. Shown is the energy density profile (the NESS current j_p is independent of site p). (c) Green-Kubo and the energy current autocorrelation function. All three methods give consistently the same value of D ($\mathcal{J} = 1.0$, $h_x = 1.4$, $h_z = 0.9045$).

diffusive $\Delta \sim k^2$. Looking at data for different τ we see that the k_0 above which one sees diffusive scaling goes as $\sqrt{\Delta}$ [Fig. 6(b)]. We remark that we do not find any subdiffusively scaling eigenvalues, which is different than what is found by minimizing the commutator [44], i.e., in the short-time behavior.

2. Hamiltonian limit

For small τ the prefactor in front of $\sim k^2$ is trivially proportional to τ because in the Hamiltonian limit of $\tau \rightarrow 0$ one should recover the diffusion constant of the Hamiltonian H_0 which is measured in units of physical time $\tau\tau$, and not in time steps t as obtained from the gap of U_τ . To check that D obtained in such a way from $M(k)$ is indeed equal to the diffusion constant of H_0 , we have calculated D directly for the Hamiltonian by two other methods, one is the Lindblad equation with boundary driving where one studies a true nonequilibrium steady state (NESS), while the other is the Green-Kubo formula using the equilibrium current autocorrelation function.

In order to facilitate comparison with existing data we here take slightly different parameters; namely, we use $H'_0 = \sum_j \mathcal{J} \sigma_j^x \sigma_{j+1}^x + h_x \sigma_j^x + h_z \sigma_j^z$, with $\mathcal{J} = 1$ and $h_x = 1.4$, $h_z = 0.9045$. A number of recent papers have values of D at infinite temperature for those parameters, with values in the range $D = 1.38$ – 1.55 ; specifically, Ref. [57] gets $D \approx 1.40$, Ref. [58] $D \approx 1.55$, Krylov method of Ref. [59] $D \approx 1.38$, with the most accurate value seemingly in Ref. [60] at $D \approx 1.445$.

Our results for D obtained from a small k behavior of $\lambda_1(k)$ of the truncated propagator are in Fig. 7(a). Taking small $\tau = 0.05$ we get $D = [1 - \lambda_1(k = 0.02)]/(k^2\tau) \approx 1.454$ for truncation size $r = 13$, while doing extrapolation [61] to $r \rightarrow \infty$ gets us $D \approx 1.447 \pm 0.005$ (similar value is obtained for $k = 0.01$). We see that our estimate agrees well with Ref. [60], while other methods are apparently less accurate. For larger $\tau = 0.1$, when the Trotterization is not negligible, the converged value is a bit larger and reflects a

finite-time diffusion constant in a circuit with long pre-thermalization time.

We shall now compare the above value of D with the one obtained by two additional completely different methods that study how the energy current density j_p behaves:

$$j_p = h_x \mathcal{J} (\sigma_p^y \sigma_{p+1}^z - \sigma_{p-1}^z \sigma_p^y). \quad (19)$$

The first method relies on the scaling of energy current density j_p in a NESS, where according to Fourier's law one should have for NESS expectations $j_p = -D \nabla h$, with h being the energy density and all quantities mean expectation value in the NESS. To induce a NESS we couple two spins at each chain end to a bath described effectively by two-site Lindblad operators. We essentially repeat calculation from Ref. [62] (see Appendix A) for the parameters used here. Results are shown in Fig. 7(b). We also observed (data not shown) that as one lowers the temperature, D increases (e.g., decreasing the energy density to $h \approx -0.02$ we get about 1% larger D). We can see [Fig. 7(b)] that the energy density profile is indeed linear, as one would expect for a diffusive system with constant D . Likewise, the NESS current density is independent of site p within about 1% precision. Using system size $L = 100$ and dividing the NESS energy current density value $j_p \approx 1.69 \times 10^{-6}$ by the fitted gradient $\nabla h \approx 1.15 \times 10^{-6}$, we get $D \approx 1.47$. Note that close to chain edges there are boundary effects on close to 20 sites, which is perhaps related to the truncation support r on which λ_1 converges [Fig. 7(a)].

The second method studies the equilibrium energy current, specifically the autocorrelation function of the extensive current $J = \sum_p j_p$. According to the Green-Kubo formula [63] the diffusion constant D is

$$D = \frac{1}{\chi} \int_0^\infty \lim_{L \rightarrow \infty} \frac{1}{L} \langle J(t) J(0) \rangle dt, \quad (20)$$

where $\chi = \text{tr}(H'_0 H'_0)/L = \mathcal{J}^2 + h_x^2 + h_z^2$. Calculating correlation function $C_J(t)$ numerically exactly (Appendix A)

in a finite system of size $L = 33$, we get $D \approx 1.45 \pm 0.01$ [the error here stems from the fact that at finite L such $C_J(t)$ agrees with the one in the TDL only up to a finite time ≈ 25 ; see Fig. 7(c) and Appendix A].

We therefore see that all three methods give the same value of D within their $\approx 1\%$ – 2% accuracy, which furthermore agrees with the value reported in Ref. [60]. We note that the chosen parameters studied in this part seem relatively “easy” in a sense that not too large t or L are necessary to get good convergence and accuracy. For other parameters, where for instance $C_J(t)$ oscillates and also takes negative values, situation might be more difficult [64].

While our truncated propagator method scales exponentially with r , it still gives more precise values of D with less numerical effort than other methods. Let us compare concrete numbers. Despite the gap being very small due to small $\tau = 0.05$ and $k = 0.02$, computation to get Δ takes ≈ 1.7 h at $r = 10$, while at $r = 13$ it takes ≈ 3 days and requires ≈ 30 Gb of memory. To get the NESS shown in Fig. 7(b) with $L = 100$ it, on the other hand, takes ≈ 19 days, while to get the $L = 33$ data for the current autocorrelation function in Fig. 7(c) it takes ≈ 6 days. Therefore, the truncated propagator method is significantly faster than other methods even using the largest $r = 13$, where it results in much more precise D . At smaller $r = 10$ – 11 , where its precision would be comparable to the NESS method, it is ~ 100 times faster. Moreover, it seems that at least in the studied case, obtaining higher precision results is easier with the truncated propagator. To increase precision in Fig. 7(c), i.e., increase L substantially beyond 33, is nearly impossible due to prohibitive memory requirements (Appendix A). To increase the precision of the NESS results, one would likewise have to increase L . Because of finite-size corrections scaling as $1/L$ [65] one would need $L \approx 400$ to decrease the error down to 0.005 estimated for our truncated propagator method at $r = 13$. This would be very time-consuming because in order to preserve precision of the matrix product operator (MPO) description we would also need a much larger bond size than $\chi = 200$ (Appendix A) rendering each relaxation step to reach NESS much slower. The reason that for a similar precision one, on one hand, needs $L \approx 400$ with NESS while already $r = 13$ suffices for the truncated propagator method is because in the truncated operator method there are no finite-size effects at any r —it always works exactly in the thermodynamic limit—and, furthermore, the convergence with r seems to be rather fast [Fig. 7(a)] [61].

E. Prethermalization at large τ and DTC

In Fig. 2(c) we can see that the gap is very small also for non-small τ (brown dashed curve) when one approaches the integrable point $\tau = \pi$. Let us explore if this is also an instance of prethermalization in the sense that there is an almost-conserved operator and that appropriate correlation functions exhibit long prethermal plateau.

At the integrable point itself we can see that Eq. (11) simplifies to $U_x = \prod_j (-i\sigma_j^x)$, $U_z = \prod_j (-i\sigma_j^z) \prod_j (1 - i\sigma_j^z \sigma_{j+1}^z) / \sqrt{2}$, or all together

$$U_\pi = (-1)^L e^{-i(\pi/4) \sum_j \sigma_j^x \sigma_{j+1}^x} e^{-i(\pi/2) \sum_j \sigma_j^y}. \quad (21)$$

This is nothing but the integrable kicked transversal field Ising model; see, e.g., Ref. [66]. In the presence of a disorder it would be an example of a much studied discrete-time crystal (also the so-called π spin glass) [24–26]. Note that our generator Eq. (11) with τ close to π is not exactly equal to the one obtained by perturbing the angle π in the two terms in Eq. (21), an often studied model of DTC [43,67–70]; however, as we shall see, its properties are rather similar. We also in particular have a specific value of the field in the σ^y direction that simplifies it even more (as well as a particular prefactor in front of the zz term [43]). Namely, because the zz and the y terms commute, one can rewrite it as

$$U_\pi^{2p} = e^{-i\pi(p/2) \sum_j \sigma_j^x \sigma_{j+1}^x}. \quad (22)$$

For odd p each two-site term gives just a relative sign between the even and odd parity sectors, while for even p it is identity, for instance, $U_\pi^4 = \mathbb{1}$. If one stroboscopically looks at every fourth step the situation at $\tau = \pi$ is actually similar to the one at $\tau = 0$: at $\tau = 0$ one has trivial $U_0 = \mathbb{1}$, whereas at $\tau = \pi$ the propagator U_π is a fourth root of identity.

Numerically calculating the correlation function, for instance, of H_0 [Eq. (10)], we indeed see (Fig. 8) long prethermalization plateau, exactly as was the case for small τ . The gap again perfectly predicts the thermalization time $T = 1/\Delta$, with the scaling being [Fig. 2(c)] $T \approx 10^{-8} \exp(21/|\pi - \tau|)$. This case of large τ is also interesting for another reason: Namely, because the propagator is close to a root of identity, we can expect an almost periodic behavior in time, that is, breaking of the time-translation symmetry, i.e., a discrete time crystal. It has been known [71] that prethermalization is a way to obtain DTC in homogeneous systems; however, not in (short-ranged) one-dimensional systems. In our case we can observe that the total magnetization oscillates (Fig. 8) with a period 2. Similar behavior has been observed before in the kicked Ising model, see Refs. [43,67–69], seemingly at odds with a one-dimensional nature of the system (for classical 1D DTC see Ref. [72]). We in fact have three different regimes in the behavior of $Z(t)$ [see Fig. 8(b)]: (i) initial fast decay of magnetization until a minimum at $t \approx L$ (see Ref. [70] for a detailed discussion of this regime), (ii) intermediate prethermal times $L \ll t \ll T$ where one has a DTC-like behavior, though the size of the plateau in the average spin magnetization $Z(t)/L$ here decays with L as $\sim 1/L$, and (iii) asymptotic times $t \gg T$ where prethermalization breaks down and $Z(t)$ decays to a

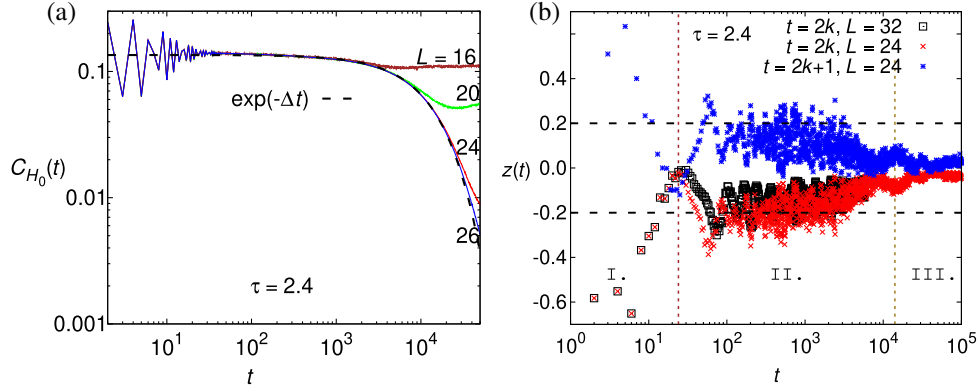


FIG. 8. Small frequency prethermalization and an almost DTC. Prethermalization and discrete-time crystal at large $\tau = 2.40$. (a) Correlation function at infinite temperature shows prethermalization. (b) Magnetization $z(t) = \langle Z(t) \rangle / L$, starting from a state $|11\dots 1\rangle$ with $L = 24$, plotted with different colors at odd and even steps, and $L = 32$ (for clarity only even times are shown). First vertical brown line is at $t = 24$ and separates regimes of fast decay and that of a prethermal DTC-like behavior, while the second vertical line is at $1/\Delta$, indicating the end of prethermalization (for more details about regimes (i)–(iii) see text). Data are for the kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

value that is exponentially small in L . Although one does not have a true DTC in the thermodynamic limit in regime (ii), it could still be of (experimental) interest as the plateau for specific initial states, like $|1\dots 1\rangle$, is still relatively large, say >0.1 at $L = 32$, despite being rather far away from the exact DTC at $\tau = \pi$. Previous works [43,67–69] mostly focused on smaller deviations from π , where one has to be extra careful about finite-size effects; see also discussion in Ref. [70]. Because the truncated propagator gives an easy way to obtain the prethermalization duration $T = 1/\Delta$, it also gives an exact prediction on the robustness of potential DTCs, avoiding confusing a finite- L DTC with an exact one in the TDL. As discussed (Sec. III B), if one is interested in the TDL, one should be aware of T and of the Heisenberg time, and look at appropriate L and t .

Worth noting is also that the scaling of the second largest eigenvalues in different sectors [data like those in Fig. 5(a)] this time goes as $\sim |\pi - \tau|^3$, i.e., the power is 3 rather than 2. Finally, let us also note in passing that for the integrable kicked Ising model at $\tau = \pi$ [Eq. (21)] a local Floquet Hamiltonian Eq. (13) does not exist despite the system not being ergodic, in fact, it is integrable with the two factors in Eq. (21) commuting; see Appendix B for details. While a standard argument for the absence of a local H_F in chaotic Floquet systems is that they heat up [51,52], the reverse is obviously not true—the absence of heating to infinite temperature does not imply the existence of a local H_F (for a single-particle example, see Ref. [73]).

F. 2D kicked Ising model

So far we have demonstrated the ideas on the 1D kicked Ising model. The methods and conclusions though are model independent and here we briefly comment and show data for some other models.

One could specifically wonder if the kicked Ising model is generic enough, for instance, it has a single layer of

commuting two-qubit gates (Fig. 3), whereas one might be interested in, say, a brickwall circuit where one has 2 layers of two-qubit gates (with gates commuting within each layer). To that end we have implemented a model where on top of kicked Ising gates we add another layer of commuting xx gates, thereby making the model more generic. In the limit of small τ such a model corresponds to a 1D XX model in a tilted magnetic field. More detailed results are presented in Appendix C; here we just summarize main findings. At small τ one again finds prethermalization with its timescale given by the leading Ruelle-Pollicott resonance of the truncated propagator. Close to $\tau = \pi$ behavior though is different than in the kicked Ising model. While the timescale again diverges, it does so only quadratically in $1/(\tau - \pi)$ [see Fig. 16(d)], and not exponentially like in the kicked Ising model. This is so because the kicked XX model is not a DTC at $\tau = \pi$. In fact, behavior is similar to the kicked Ising model at small τ in sectors with nonzero quasimomentum k where there are no almost-conserved operators (Fig. 5). From the momentum dependence of λ_1 one can again get the energy diffusion constant which agrees with the one obtained from the integral of the current autocorrelation function via the Green-Kubo formula (Fig. 17).

The truncated operator propagator can be used also in 2D models and we focus here on an example of a 2D kicked Ising model. Spins are now located on a 2D square lattice, with a one-step propagator being

$$\begin{aligned}
 U_\tau &= U'_{zz} U'_z U'_x, \\
 U'_x &= \prod_{i,j} e^{-i\tau h_x \sigma_{i,j}^x}, \\
 U'_z &= \prod_{i,j} e^{-i\tau h_z \sigma_{i,j}^z}, \\
 U'_{zz} &= e^{-i\tau \sum_{\text{n.n.}} \sigma_{i,j}^z \sigma_{i',j'}^z},
 \end{aligned} \tag{23}$$

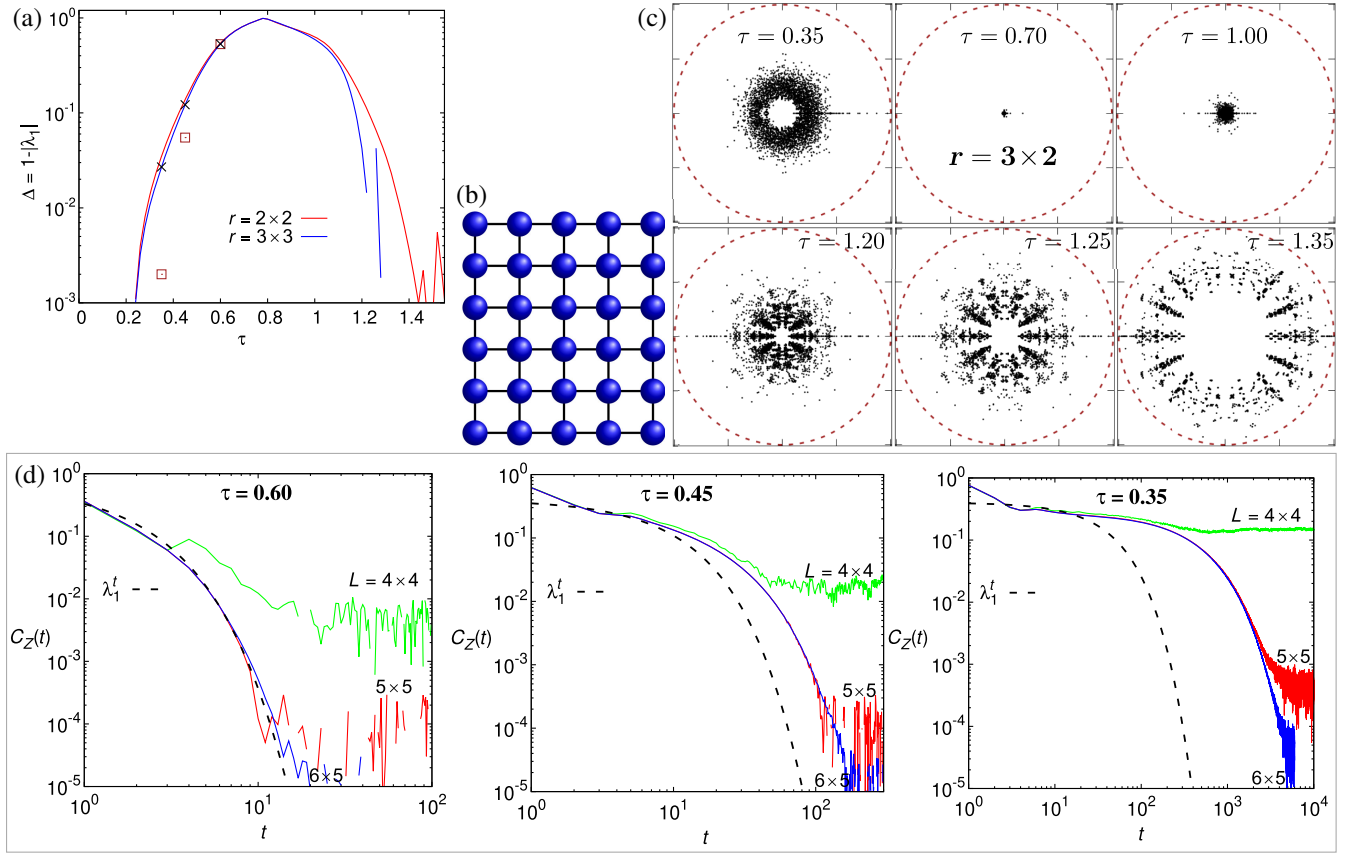


FIG. 9. 2D kicked Ising model on a square lattice [illustrated in (b)]. (a) Dependence of the gap on τ . Three crosses mark the empirical decay used as dashed black exponential curves for three τ in (d), while brown squares are best fitted exponentials to numerical data for largest size $L = 6 \times 5$ in (d). Panel (c) shows spectra of $M(\mathbf{k} = 0)$ for truncation $r = 3 \times 2$.

where the sum in the U'_{zz} is over nearest neighbors on a 2D square lattice [Fig. 9(b)]. Similar to the 1D model we calculate the operator propagator in an infinite system truncated to a basis of local operators supported on r sites. Because we are in 2D, the local basis operators are now from r sites on an $n \times m = r$ grid. We show results for the largest eigenvalue for the $r = 2 \times 2$ and 3×3 truncations in Fig. 9(a). Because one has translational invariance in x as well as in y direction, resulting in two conserved quasi-momentum components k_x and k_y , construction of independent basis elements is a bit more involved. One has to ensure that all basis elements are independent, i.e., that they cannot be obtained from each other by translations. In 1D it was enough to fix, e.g., the first-site operator to be a nonidentity; in 2D this is not enough. Details can be found in Appendix D. Let us just mention that the size of the basis approaches $4^{r=mn}$ for large m and n . We also use *in situ* numerical propagation of local basis operators so that we can easily study $r = 3 \times 3$ truncation.

Dynamics of the 2D model probably warrants a separate study; here we focus on the key features in the zero quasimomentum sector ($k_x = k_y = 0$), setting also $h_z = h_z = 1$ and varying only τ (note that compared to the 1D kicked Ising model we do not use factors $1/2$ or $1/4$ in the

exponents). At small τ one finds prethermalization with the prethermalization time being exponentially large in $1/\tau$, finding no Trotterization transition [Fig. 9(a)]. The decay timescale at larger $\tau = 0.60$, obtained from the exact numerical simulation of finite systems of size up to $L = 6 \times 5$ spins with periodic boundary conditions, agrees with the one obtained from the truncated propagator at truncation size $r = 3 \times 3$ [Figs. 9(a) and 9(d)]. At smaller τ one can see a rapid increase of the prethermalization time, e.g., between $\tau = 0.45$ and $\tau = 0.35$ it increases by almost 2 decades, though small linear truncation size (2 or 3) is not yet large enough to get a precise value of the prethermalization time (this is actually similar as in 1D). Nevertheless, one can see a clear rapid decrease of the gap with decreasing τ [Fig. 9(a)]. Similarly as in 1D relying on small systems, e.g., with $L = 4 \times 4 = 16$ spins, would incorrectly indicate a plateau at large times [green curves in Fig. 9(d)].

There are two other interesting points worth highlighting. The first is that at $\tau = \pi/4$ all eigenvalues of $M(\mathbf{k} = 0)$ are zero; i.e., the gap shown in Fig. 9(a) is exactly one. While this is not specific to 2D, and has already been noted in the 1D kicked Ising model [37], it means that all autocorrelation functions of extensive local operators decay

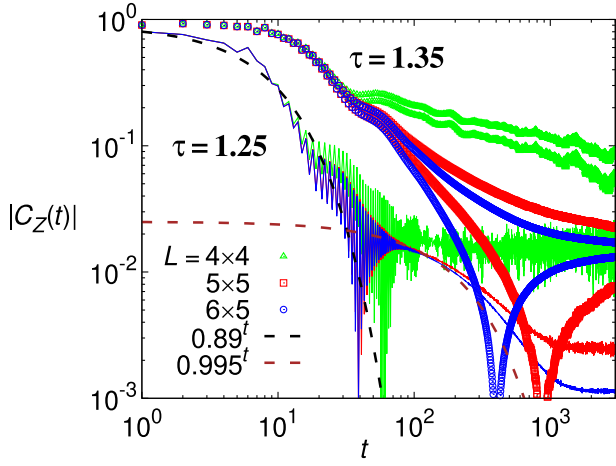


FIG. 10. Transition in the behavior of magnetization correlation function in the 2D kicked Ising model around $\tau \approx 1.3$. Top set of data (points) is for $\tau = 1.35$, while the lower ones (curves) are for $\tau = 1.25$.

immediately. We state some further properties of M at $\tau = \pi/4$ and $\tau = \pi/2$ in Appendix D.

The second point is an apparent transition in the dynamics around $\tau \approx 1.3$. Namely, for $\tau > \pi/4$ one notices that the largest absolute eigenvalue becomes negative, as opposed to being positive for $\tau < \pi/4$ [see, e.g., the spectrum for $\tau = 1.0$ in Fig. 9(c)]. As a consequence, correlation functions will oscillate between negative and positive values between even and odd times. What is even more interesting is that around $\tau \approx 1.3$ there seems to be a qualitative change in the spectrum [Fig. 9(c)] which no longer has a noisy ring or disk core (like in both 1D models). This change is reflected also in the decay of correlation functions. In Fig. 10 we can see that at $\tau = 1.25$ the autocorrelation function of magnetization has two distinct decay regimes: Up to $t \approx 50$ it decays fast in an oscillatory manner (C_Z at even and odd times have opposite signs), whereas for $t > 50$ it is positive and decays slower—perhaps still exponentially, though our largest system size is not large enough to verify that with confidence. On the other hand, for $\tau = 1.35$ the decay is much slower, has a single regime (C_Z is always positive), with the asymptotic decay perhaps being more likely a power law rather than exponential. This could be due to a continuum of eigenvalues approaching 1 in modulus [see Ref. [37] for an example where a continuum of eigenvalues (a branch cut) causes a power law decay].

IV. CONCLUSION

We have shown that the truncated propagator of extensive local operators is a very versatile method to study Floquet systems. It determines evolution of correlation functions; specifically, through the largest eigenvalue and the associated eigenvector, it gives an explicit prediction for

the prethermalization time and the effective (local) Floquet Hamiltonian governing prethermalization. This allowed us to resolve the question about reliability of approximating Hamiltonian evolution by a small-step Trotterization. Except for a single almost conserved observable all other orthogonal observables do not exhibit prethermalization, and even the almost conserved observable decays in the thermodynamic limit. This means that due to the many-body dynamics the so-called shadowing time is always finite, and that there is no Trotterization transition in chaotic quantum systems. Previous observations of Trotterization transition were an artifact of finite-size effects. One has to ensure that the Heisenberg time is larger than the (long) prethermalization time. Because the limits $L \rightarrow \infty$ and $t \rightarrow \infty$ do not commute, in order to get the correct thermodynamic results it is essential that our method works directly in the thermodynamic limit. If one would instead use small systems, for instance, exact diagonalization, one would get incorrect results.

The truncated propagator can also be used to identify prethermal discrete-time crystals. Just looking at the spectrum will already give an indication if there is a prethermal DTC [eigenvalues very close to roots of 1, e.g., those around -1 in Fig. 2(c) at $\tau = 3.0$]. Studying chaotic kicked Ising model we find, perhaps surprisingly, that the prethermalized regime takes up the larger part of the gate duration interval $\tau \in [0, \pi]$. For instance, prethermalization is absent (say, gap $\Delta > 10^{-3}$, i.e., $T < 10^3$) only for roughly $\tau \in [1.5, 2.2]$. Glimpses of this rather slow thermalization in the 1D kicked Ising model can be seen already in data presented in Ref. [41], and are present also in the 2D kicked Ising model. Prethermalization seems to be more common than previously thought and is not limited only to large frequencies (small τ).

Momentum dependence of the largest eigenvalue also contains information about transport. We have used it to calculate the energy diffusion constant during the long prethermalization phase, as well as for the Hamiltonian model obtained in the limit $\tau \rightarrow 0$. Despite simply using small τ to approximate Hamiltonian dynamics, which causes the gap to be very small and difficult to calculate, the obtained diffusion constant is very precise and the method is faster than previously used methods.

While we always used numerics to obtain the spectrum of the truncated propagator, it would be interesting to see if exact solutions are possible for specific models. Coming up with a better numerical method that would allow one to get the relevant eigenvalues for truncation support cutoff larger than about $r = 14$ sites, feasible with the *in situ* method that we use, would also be a valuable extension. Note though that whether one has a prethermalization or DTC can be already inferred at small r , and so the method can be used as a simple tool to look for those regimes. In the kicked Ising model we in particular find a prethermalized DTC behavior for a wide range of parameters around $\tau = \pi$

(roughly $|\pi - \tau| < 0.9$), with the effect though slowly decaying with increasing system size.

We have also studied the kicked XX model, where prethermalization is found only for small τ , as well as 2D kicked Ising model finding an interesting transition in the decay of correlation functions.

ACKNOWLEDGMENTS

The author would like to thank Urban Duh for work on related project [39], and acknowledge Grants No. J1-4385, No. J1-70049, and No. P1-0402 from Slovenian Research Agency (ARIS).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: NUMERICS

1. Spectrum of $M(k)$

The size of matrix $M(k)$ is $N = \frac{3}{4}4^r$. For smaller r the size is small enough to numerically diagonalize the whole matrix. For larger r one can use methods targeting only the largest eigenvalue (in modulus), or few largest, like the simple power method for λ_1 , or Lanczos-based methods (we have used the ARPACK package [74]) that construct a tridiagonal matrix representing M on a Krylov subspace. The method does not require the whole matrix, which would be memorywise prohibitive, but rather needs only application of M on a vector. This can be simply done by simulating the action of our circuit on $r + 2$ sites; see Eq. (9), that is, consecutively applying individual gates (Fig. 3). While the implementation on $r + 2$ sites is straightforward, one can improve on speed and memory by directly performing evolution *in situ* on just r sites as described in the next subsection.

In Fig. 11 we check convergence of $\lambda_1(k = 0^+)$ that determines the prethermalization time with the support size r , confirming that T reported in Fig. 2(c) is the asymptotic one.

2. *In situ* propagation of operators

To get the action of the truncated propagator we need the overlaps of the propagated local operator $U^\dagger a_1^{(\beta)} U$, where $a_1^{(\beta)}$ lives on r sites, with local operators that are either on the same r sites or are shifted by one site to the left or to the right (for evolutions that can shift by at most 1 site, like the kicked Ising model); see Eq. (9). A simple way to do that is simply to evolve the operator $a_1^{(\beta)}$ on $r + 2$ sites, starting by padding one identity at the left and the right edge,

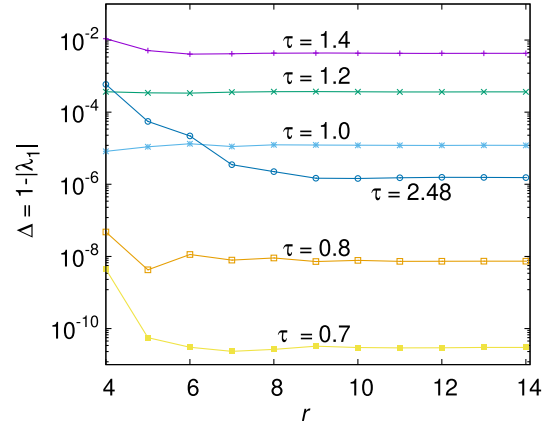


FIG. 11. Convergence of the gap with r for selected τ [symbols in Fig. 2(c)].

calculating the full $U^\dagger a_1^{(\beta)} U$ that now lives on $r + 2$ sites, and then projecting it to the relevant operators with support on r sites. However, that is expensive: We have to do evolution on $r + 2$ sites and each additional site increases memory and computational requirements by a factor of ≈ 4 . Much faster is to do the *in situ* evolution on just the relevant r sites. In the following we describe how to do that, having in mind the 1D kicked Ising model.

One uses the fact that an initial operator with support on r sites can increase its support only at its edges. Specifically, to get the contribution without any shift, $a_1^{(\beta)} \rightarrow a_1^{(\alpha)}$, the two-qubit gate acting on the boundary [dashed lines in Fig. 12(a)] has to map at the left edge $\mathbb{1}_0 x_1$, where “ x ” indicates any nonidentity operator, $x \in \{\sigma^x, \sigma^y, \sigma^z\}$, to again the identity at the leftmost site [see Fig. 12(a)]. For the kicked Ising model where the

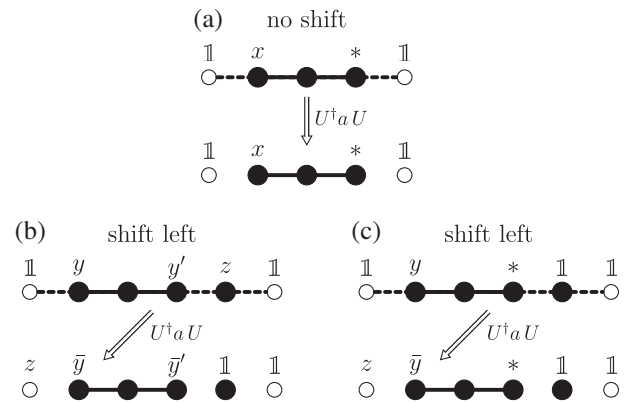


FIG. 12. Truncated propagator construction: *in situ* propagation of local operators on just r sites. The initial operator is supported on r sites (black circles) padded with an identity outside its support (white circles). Executing a two-qubit zz gate on dashed lines has to execute a mapping indicated by the arrow. Afterward one has to perform the remaining two-qubit gates (full lines). See text for details.

two-qubit gate has only the zz interaction, the only possibility is $\mathbb{1}_0 x_1 \rightarrow \mathbb{1}_0 x_1$. At the right edge one has a similar situation, with the only difference that the last operator can also be the identity, resulting in $(*)_r \mathbb{1}_{r+1} \rightarrow (*)_r \mathbb{1}_{r+1}$, where $(*) \in \{\sigma^x, \sigma^y, \sigma^z, \mathbb{1}\}$. If we want to achieve a mapping of the initial operator to a shifted one supported on r consecutive sites, we have to get an identity at its tail end, and a nonidentity operator at its front end. For the zz gate in question such a mapping can be achieved only by $\mathbb{1}_0 y_1 \rightarrow \sigma_0^z \bar{y}_1$, where “ y ” stands for $y \in \{\sigma^x, \sigma^y\}$, and its negation $\bar{y}$ is σ^x if $y = \sigma^y$ and σ^y if $y = \sigma^x$. This is sketched in Fig. 12(b). For a shift to the left there is another possibility, namely, the rightmost site can already have an identity operator, resulting in a scheme in Fig. 12(c) [for a shift to the right Fig. 12(b) is the only option]. We can see that in both Figs. 12(b) and 12(c) it is enough to keep track of only r sites—at the beginning on r black circles, and after performing the first step, transformation on all dashed links, only on the leftmost r sites, afterward executing the remaining gates (full-line links in Fig. 12).

3. Correlation functions

We calculate correlation functions in finite systems numerically exactly by evolving a pure random initial state $|\psi\rangle$ in up to $L = 33$ spins by repeatedly applying one- and two-qubit gates (see Fig. 3). For large L such a single random state expectation value is a good approximation for an infinite temperature average. Namely, finite-size fluctuations in a correlation function scale as $\sim 1/2^{L/2}$, and are, e.g., $\sim 10^{-5}$ for $L = 32$. We always look at values of $C_A(t)$ that are larger than that [e.g., Fig. 5(b)]. Note that in some cases, like studies of transport, there can be other boundary finite-size effects that are larger; see Fig. 7(c) for an example. Therefore, we always check convergence with L , making sure to have finite-size effect under control. Memory requirements of the method are exponential in L (keeping track of two pure states, each of size $\mathcal{N} = 2^L$),

and reach ≈ 400 Gb for $L = 33$, while each gate application takes $\mathcal{O}(\mathcal{N})$ operations.

In Figs. 13 and 14 we check that correlation functions decay to zero in the TDL also for other values of τ . At $\tau = 0.9$ the prethermalization time is so large that the simulation takes a lot of time and we cannot afford to simulate as large systems. Nevertheless, $L = 25$ already reveals that the same prethermalization theory still applies, i.e., there is no Trotterization transition, despite being well below the claimed transition (crossover) point in Ref. [16]. At $\tau = 1.2$ in Fig. 14 the decay is rather clear, including for a specific fully polarized initial state.

In calculating the energy current autocorrelation function Eq. (19), in Fig. 7(c) we have simply used the kicked Ising model with a small time step $\tau = 0.05$. Comparing results at $L = 32$ between $\tau = 0.05$ and $\tau = 0.02$, we can see that, e.g., up to $t = 30$ the difference in D is about 0.002 due to a finite Trotterization step approximation of Hamiltonian evolution. On the other hand, comparing different L at $\tau = 0.05$ we can estimate that the $L = 33$ results do correctly reproduce the correlation function in the thermodynamic limit up to $t \approx 25$ with precision in D of about 0.002 (this time scales diffusively $\propto L^2$ with system size). Because $C_J(t)$ at $t = 25$ is already quite small, taking into account that the remaining integral from $t = 25$ to ∞ would bring only about ≈ 0.01 , we can estimate the diffusion constant by the integral of $L = 33$ data up to $t = 30$, which gives 1.452, and estimate error at 0.01.

One can in fact use the Green-Kubo integral up to a finite time t up to which we trust to have the correct $C_J(t)$ in the TDL as a lower bound on D because $C_J(t)$ for the chosen parameters is always positive, thereby estimating $D > 1.45$. On the other hand, convergence of the truncated operator propagator method [Fig. 7(a)] seems to be monotonously from above, so using the $r = 12$ value of $D = 1.462$ we can estimate $D < 1.46$. Both methods together therefore suggest tighter estimate $D = 1.45 - 1.46$.

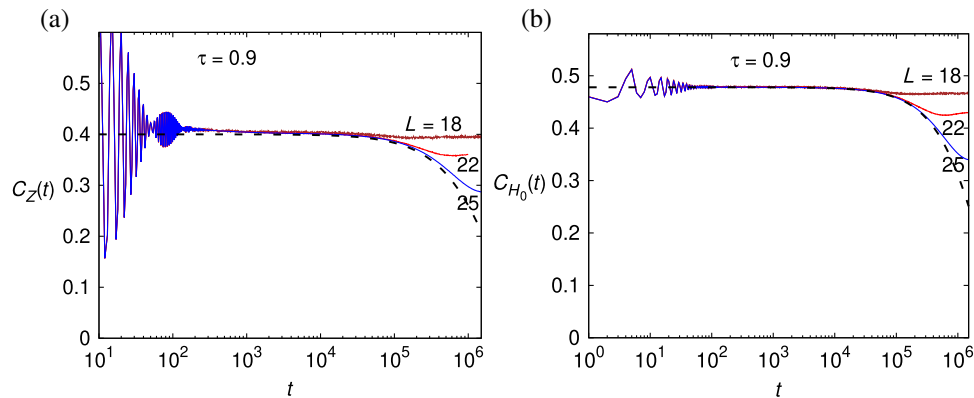


FIG. 13. Autocorrelation functions at $\tau = 0.9$ again decay to zero in the TDL. Theoretical dashed lines (no fitting parameters) are exponential decay agreeing with numerics for large L . All for kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

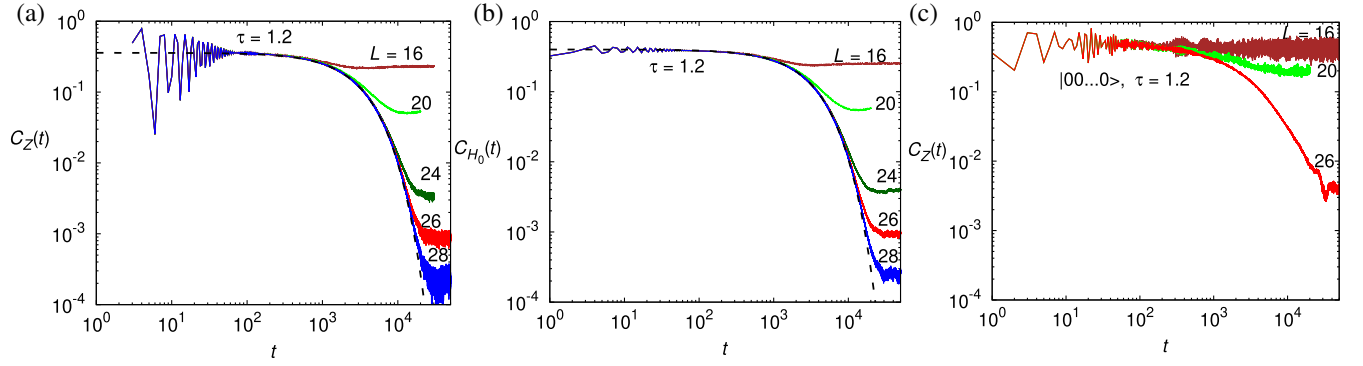


FIG. 14. Autocorrelation functions at $\tau = 1.2$ and random initial states [(a),(b)], as well as for a polarized initial state $|00 \dots 0\rangle$ in (c). All for kicked Ising model Eq. (11) with $J = \frac{1}{2}$, $h_x = h_z = 1$.

4. Boundary driven systems and NESS

To obtain a NESS and study transport properties of the chaotic Ising model H_0 , specifically its diffusion constant, we use a boundary driven Lindblad equation. The method has been used many times since its first use on the very same Ising model [62] (for different parameter values). Let us briefly outline its working.

Evolution of the density operator $\rho(t)$ is described by the Lindblad master equation,

$$\frac{d\rho}{dt} = i[\rho, H] + \sum_j ([L_j \rho, L_j^\dagger] + [L_j, \rho L_j^\dagger]), \quad (\text{A1})$$

where L_j are Lindblad operators. For our simulations we want to induce the flow of energy in the NESS, and the Lindblad operators act on two boundary spins at each end. They are chosen in such a way that the dissipator $\mathcal{L}$ given solely by Lindblad operators on those two spins would induce a canonical state of two boundary spins. More precisely, L_j at, say, the left edge are such that $\mathcal{L}(\rho_L) = 0$, where ρ_L is the reduced density operator obtained by tracing over 4 out of 6 spins, $\rho_L \sim \text{tr}(e^{-H/T_L})$, where H is the Hamiltonian on 6 spins. See Ref. [62] for more details. We use Lindblad (bathlike) “temperature” parameters $T_L = 295$ and $T_R = 300$, inducing an almost infinite-temperature NESS in which the average energy density is small.

The NESS, which is unique, is obtained by evolving $\rho(t)$ in time (Trotter step in the fourth order scheme is 0.05) until convergence to the NESS is obtained (for $L = 100$ this relaxation time is 5×10^3). Evolution is performed by writing ρ in a matrix product operator form with matrices of maximal bond size χ [$\chi = 200$ for data in Fig. 7(b)], with numerical complexity scaling as $\mathcal{O}(\chi^3)$. Application of a single two-site gate is performed using standard time-evolved block decimation (TEBD) techniques [75]. Convergence with the MPO bond dimension seems to be rather slow (definitely slower than in many other studies of spin transport with the same method). The value of

diffusion constant that we get, $D \approx 1.47$, seems to be a bit on the larger side compared to the other two methods.

APPENDIX B: BCH SERIES

Writing all propagators in $U = U_z U_x$ in terms of generators as

$$U_z = e^{-iA\tau}, \quad U_x = e^{-iB\tau}, \quad U = e^{-iH_F\tau}, \quad (\text{B1})$$

the Baker-Campbell-Hausdorff series gives

$$H_F = A + B + \frac{(-i\tau)}{2} E_3 + \frac{(-i\tau)^2}{12} ([A, E_3] - [B, E_3]) + \dots, \quad (\text{B2})$$

where $E_3 \equiv [A, B]$, and we explicitly wrote out terms up to third order in products of A and B . More systematically, the series can be written in such a way that all next order terms are a single commutator of two terms occurring in previous orders, i.e.,

$$H_F = \sum_{j \geq 1} (-i\tau)^{|j|-1} z_j E_j \equiv \sum_{m=0} (-i\tau)^m H_m, \quad (\text{B3})$$

where $E_j \equiv [E_{j'}, E_{j''}]$ with appropriate $j', j'' < j$, starting with $E_1 \equiv A$, $E_2 \equiv B$ and $|1| = |2| = 1$, while z_j are rational numbers. Coefficients up to order 20 are provided in Ref. [76] [$|j| \leq 20$, i.e., $m \leq 19$ in Eq. (B3)]. For instance, the highest order $|j| = 20$ contains 52377 terms E_j , while all 20 orders together have $\max(j) = 111013$ terms.

We have symbolically calculated all $H_0, \dots, H_{19}$ in an infinite system, each H_m being a translational sum of densities, $H_m = \sum_{\alpha} c^{(\alpha)} \sum_j \sigma_j^{(\alpha)}$, with the Hilbert-Schmidt (i.e., Frobenius) norm being $|H_m|^2 = \sum_{\alpha} (c^{(\alpha)})^2$. Norms of individual orders are shown in Fig. 15 for different τ . Note that the τ dependence of H_F is trivial [Eq. (B3)]; the nontrivial parts are H_m . A similar plot has been shown for

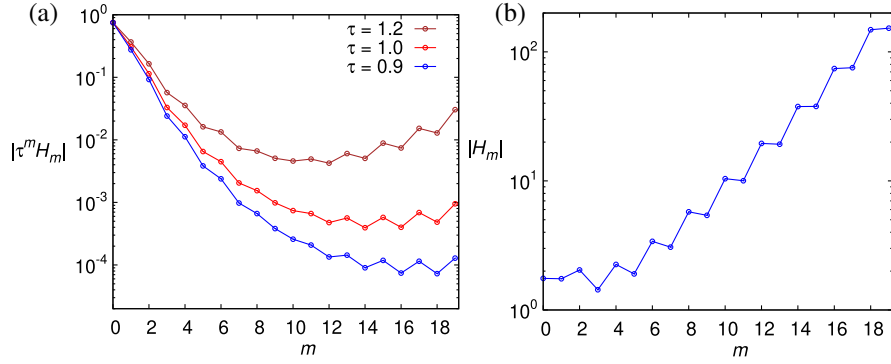


FIG. 15. Divergence of the BCH series for the kicked Ising model. (a) Chaotic case in the prethermalization regime. (b) Integrable case at the DTC point $\tau = \pi$. We use $J = 1/2$, $h_x = h_z = 1$ in Eq. (12).

different parameters in Ref. [37], and for a finite system in Ref. [18].

We can also calculate the prethermalization finite-time plateau. If H'_F would be conserved (which it is not), the plateau in energy correlation function of H_0 would be equal to the square of the projection of H_0 to H'_F , which can be written as

$$\langle C_{H_0} \rangle = \frac{[\text{tr}(H_0 H'_F)]^2}{\text{tr}(H'^2_F)}. \quad (\text{B4})$$

Evaluating traces to few lowest orders in τ , we get ($J = \frac{1}{2}$, $h_x = h_z = 1$)

$$\langle C_{H_0} \rangle \approx \frac{9}{16} - \frac{3}{32} \tau^2 - \frac{13}{1152} \tau^4 + \dots, \quad (\text{B5})$$

for instance, evaluating to $\langle C_{H_0} \rangle \approx 0.457$ for $\tau = 1.0$ used in Fig. 1(d). The leading term above is simply $\text{tr}(H_0^2) = J^2(h_x^2 + h_z^2 + \frac{1}{4})$, which would be the plateau in the autonomous limit.

Doing similar projection for the magnetization correlation function gets

$$\langle C_Z \rangle = \frac{[\text{tr}(Z H'_F)]^2}{\text{tr}(H'^2_F)} \approx \frac{4}{9} - \frac{4}{81} \tau^2 - \frac{73}{14580} \tau^4 + \dots, \quad (\text{B6})$$

evaluating to $\langle C_Z \rangle \approx 0.39$ for $\tau = 1.0$ shown in Fig. 1(d). Again, the leading order term, obtained for the Hamiltonian limit $\tau = 0$, is simply $h_z^2/(h_z^2 + h_x^2 + 1/4)$.

1. Integrable transversal Ising

The integrable kicked Ising propagator at $\tau = \pi$ is very special; it can be written [Eq. (21)] as $U_\pi = (-1)^L e^{iA} e^{iB}$ (τ is here included in A and B), with the two unitary factors commuting. Despite that, it cannot be written as $U_\tau = e^{-iH_F \tau}$ with a local Floquet Hamiltonian H_F . Namely, we have calculated successive orders in the BCH series in an infinite system for this special transversal Ising model, seeing a clear divergence with m [Fig. 15(b)]. The structure

of different orders is here simple. At even order $m = 2k$ there are exactly $2(k+1)$ nonzero terms in H_m : namely, one σ^y , $k+1$ terms of form $\sigma^z(\sigma^y \dots \sigma^y)\sigma^z$ with a varying number of consecutive σ^y (from 0 to k), and k terms of form $\sigma^x(\sigma^y \dots \sigma^y)\sigma^x$ with the number of consecutive σ^y varying from 0 to $k-1$ (a structure inherited from Jordan-Wigner transformation). For odd $m = 2k+1$ one has $2(k+1)$ nonzero Pauli strings, namely, $\sigma^x(\sigma^y \dots)\sigma^z + \sigma^z(\sigma^y \dots)\sigma^x$, with the number of σ^y going from 0, 1, ..., k . H_F is therefore nonlocal in Pauli operators. We also remind (Sec. II B) that at $\tau = \pi$, regardless of k , the truncated propagator has all nonzero eigenvalues equal to $|\lambda_j| = 1$ (a fraction 5/24 of eigenvalues is $+1$, and the same number -1); i.e., all eigenvectors are either exactly conserved by U_π or are from the kernel (a fraction 7/12 of all operator space size).

APPENDIX C: 1D KICKED XX MODEL

The one-step propagator we take is [Fig. 16(a)]

$$\begin{aligned} U_\tau &\equiv U_{zz} U_z U_x U_{xx}, \\ U_x &= \prod_j e^{-i\tau h_x \sigma_j^x / 2}, \\ U_z &= \prod_j e^{-i\tau h_z \sigma_j^z / 2}, \\ U_{zz} &= e^{-i\tau \sum_j \sigma_j^z \sigma_{j+1}^z / 4}, \\ U_{xx} &= e^{-iJ_x \tau \sum_j \sigma_j^x \sigma_{j+1}^x / 4}. \end{aligned} \quad (\text{C1})$$

Compared to the kicked Ising model we have an additional layer of commuting xx gates. We use $h_x = h_z = 1$, $J_x = 1$. Results for the spectrum of the truncated propagator, correlation function in prethermalization regime of small τ , and dependence of the gap on τ are shown in Fig. 16. As discussed in the main text, one notable difference compared to the 1D kicked Ising model is behavior close to $\tau = \pi$. Here the timescale is only quadratically large in $1/|\tau - \pi|$, rather than exponentially as in the kicked Ising model. This

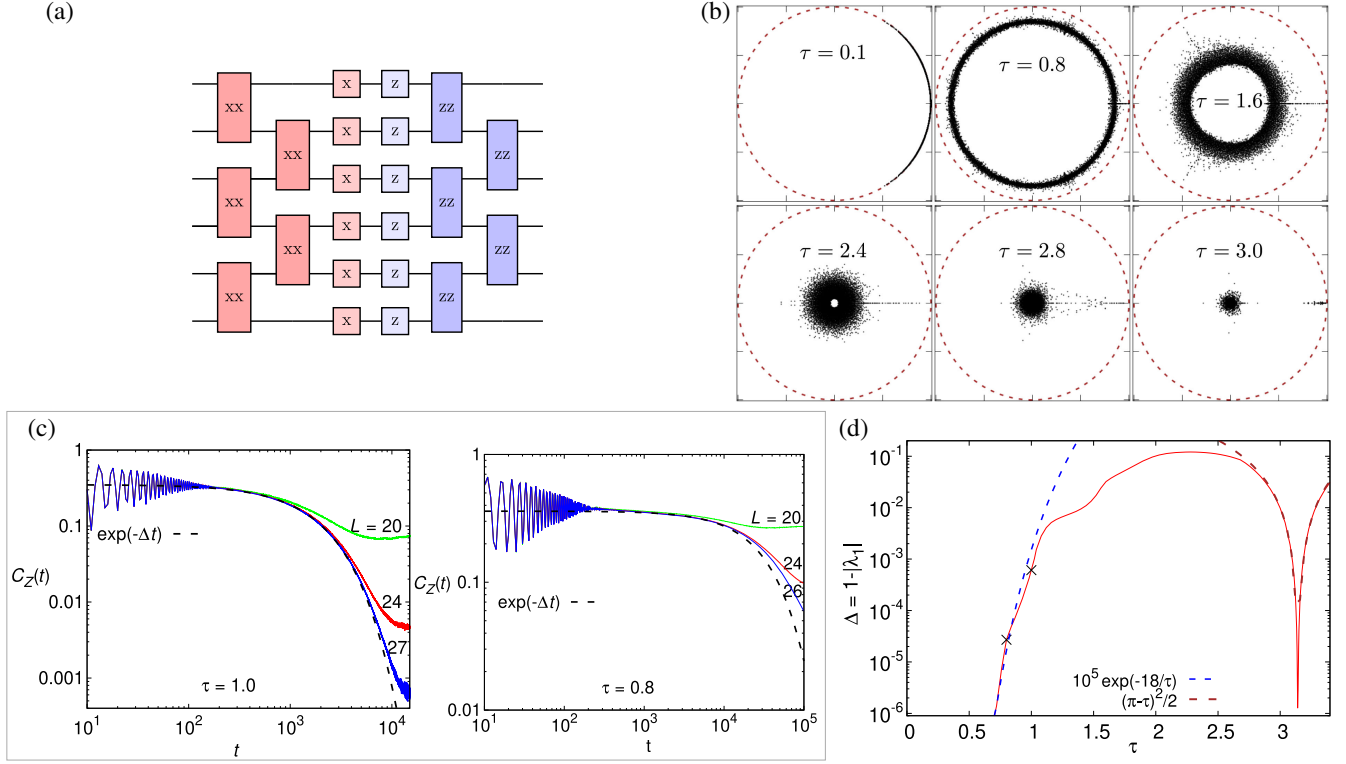


FIG. 16. 1D kicked XX model Eq. (C1). Panel (a) shows the circuit, (b) the spectra of truncated propagators for $r = 7$, (c) autocorrelation function of magnetization [black dashed curves use the gap from (d)], and (d) the gap for $r = 8$. Two crosses indicate the decay rate for dashed curves for τ in (c).

is because at $\tau = \pi$ the kicked XX model is not a DTC; i.e., all nontrivial eigenvalues of M go toward $\lambda = 1$, as opposed to the DTC behavior in the 1D and 2D kicked Ising models where they go toward ± 1 .

1. Energy diffusion

We can again study energy transport in the Hamiltonian limit by taking small τ , where dynamics approximates Hamiltonian

$$H'_0 = \sum_j \mathcal{J}_z \sigma_j^z \sigma_{j+1}^z + \mathcal{J}_x \sigma_j^x \sigma_{j+1}^x + h_z \sigma_j^z + h_x \sigma_j^x. \quad (\text{C2})$$

In numerics we will use $\mathcal{J}_z = \mathcal{J}_x = h_x = h_z = 1$. The energy current is

$$j_p = h_x \mathcal{J}_z (\sigma_{p-1}^y \sigma_p^z - \sigma_p^z \sigma_{p-1}^y) \quad (\text{C3})$$

$$+ h_z \mathcal{J}_x (\sigma_{p-1}^x \sigma_p^y - \sigma_p^y \sigma_{p-1}^x) + 2\mathcal{J}_z \mathcal{J}_x (\sigma_{p-1}^x \sigma_p^y \sigma_{p+1}^z - \sigma_p^z \sigma_{p-1}^y \sigma_{p+1}^x). \quad (\text{C4})$$

Results for the diffusion constant D obtained from the truncated propagator using $\tau = 0.05$ and $k = 0.01$ and $k = 0.02$ are in Fig. 17(a), and agree with D obtained by evaluating the integral of the energy current autocorrelation function shown in Fig. 17(b). The energy current

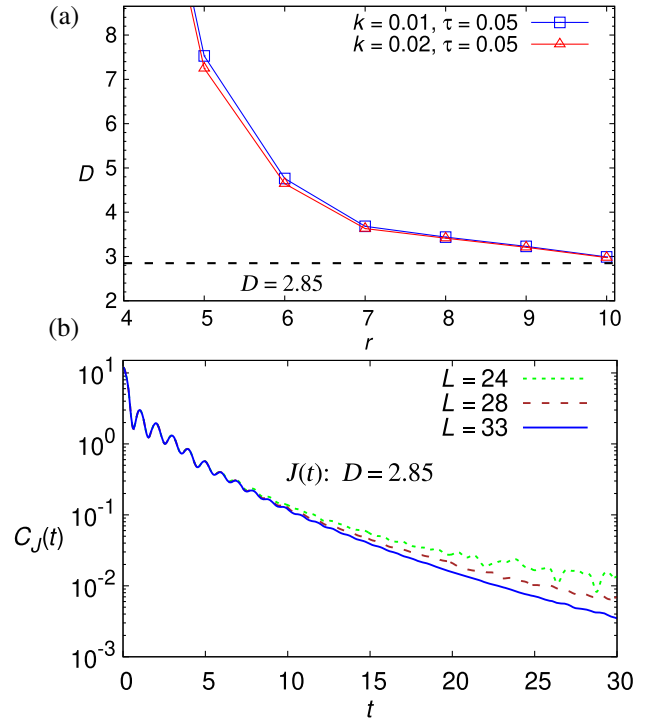


FIG. 17. Energy diffusion constant in the XX model Eq. (C2). (a) From the truncated propagator, (b) from current autocorrelation function.

autocorrelation function is at first sight similar to the one for the 1D kicked Ising model in Fig. 7(c), the difference here being more regular oscillations visible at smaller t .

APPENDIX D: 2D KICKED ISING MODEL

1. Local operator basis

Let us illustrate the local basis construction on a simple case of a $r = 2 \times 2 = 4$ sites. All basis elements can be grouped into individual classes according to the position of all identity operators. For $r = 2 \times 2$ there are $N_{\text{cl}} = 10$ classes that we list as symbols arranged in a 2×2 pattern indicating their position in a 2D grid, 1 denoting the identity operator and “x” an arbitrary nonidentity Pauli matrix:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & x & x & 1 & 1 & x & x & x & 1 & x & x & 1 & x & x & x & x \\ 1 & x' & x & x' & 1 & x' & 1 & x' & x & 1' & 1 & x' & x & x' & x & x' & x & x' & x & x \end{pmatrix}. \quad (\text{D1})$$

The number of basis elements in a given class is equal to 3^p , where p is the number of “x” symbols, resulting in the total basis size for 2×2 truncation being $N = 228 = \sum_j^{N_{\text{cl}}} 3^{p(j)}$. All other arrangements of operators on 2×2 sites, for instance, $\begin{pmatrix} x & 1 \\ x & 1 \end{pmatrix}$ that is not a member of the above basis, can be obtained from the basis in Eq. (D1) by translations in x and y directions. Compared to the 1D case basis we can see that we have two additional classes [fifth and ninth in Eq. (D1)] in which the lower rightmost operator is not necessarily a nonidentity.

For larger truncation sizes r procedure is conceptually the same but combinatorially more complicated; however, basis can be easily constructed using translations to eliminate elements. We list the basis sizes and the number of classes in Table I. For an $r = n \times m$ truncation support the basis size is a bit less than 4^{nm} , while the number of classes is upper bounded by 2^{nm} [for large n and m it approaches these two bounds, while the exact formula for the number of classes for, e.g., $r = n \times 2$ basis is $[5 + 6(2^{n-2} - 1)]2^{n-1}$, and for $r = n \times 3$ it is $[11 + 14(2^{n-2} - 1)]2^{2n-2}$].

To obtain M one needs to propagate the local basis. An easy way is to, similarly as in 1D, include in propagation all sites directly connected to the r basis sites (a layer of commuting two-site gates, like in kicked Ising models, can increase support only by one site to its neighbors). For instance, for a $r = 2 \times 2$ basis this requires 8 additional sites (Fig. 18), demanding in total propagation on $r + 8 = 12$ sites. While this is numerically feasible, larger bases would not be. Therefore we for 2D kicked Ising model use an *in situ* propagation on just r sites. This allows us to treat also $r = 3 \times 3$ basis with ease. The *in situ* propagation for a 2D support is a bit more nuanced than in 1D, but follows the same idea.

2. Special point $\tau = \pi/4$

One has $e^{-i(\pi/4)\sigma_j^z \sigma_{j+1}^z} = (\mathbb{1} - i\sigma_j^z \sigma_{j+1}^z)/\sqrt{2}$, and using $h_x = h_z = 1$ also $e^{-i(\pi/4)\sigma_j^x} = (\mathbb{1} - i\sigma_j^x)/\sqrt{2}$, and $e^{-i(\pi/4)\sigma_j^y} = (\mathbb{1} - i\sigma_j^y)/\sqrt{2}$. The two one-qubit gates together perform a permutation of Pauli operators, namely, map $\sigma^x \rightarrow \sigma^z$, $\sigma^y \rightarrow \sigma^x$, $\sigma^z \rightarrow \sigma^y$. The two-qubit zz gate, on the other hand, has four 2×2 blocks; in one it maps

$$\mathbb{1}\sigma^y \rightarrow \sigma^z\sigma^x, \quad \sigma^z\sigma^x \rightarrow -\mathbb{1}\sigma^y, \quad (\text{D2})$$

and in the second it maps

$$\mathbb{1}\sigma^x \rightarrow -\sigma^z\sigma^y, \quad \sigma^z\sigma^y \rightarrow \mathbb{1}\sigma^x. \quad (\text{D3})$$

The remaining two blocks are obtained by swapping the first and second qubits (the gate is symmetric with respect to the exchange of qubits). All other Pauli products ($\mathbb{1}\sigma^z, \sigma^z\mathbb{1}, \sigma^x\sigma^y, \sigma^y\sigma^x, \sigma^x\sigma^x, \sigma^y\sigma^y, \sigma^z\sigma^z, \mathbb{1}\mathbb{1}$) are mapped back to themselves, $ab \rightarrow ab$. All put together this causes that at $\tau = \pi/4$ the truncated propagator at $k = 0$ has all eigenvalues equal to zero, some of which form Jordan blocks. One has an interesting situation where taking a finite block corresponding to operators with support on $\leq r$ sites of an otherwise unitary operator (spectrum on a unit circle) results in a truncated propagator with all eigenvalues being 0. For instance, for $r = 2 \times 2$ truncation there is one 2×2 Jordan block for each basis class, i.e., there are

TABLE I. The local operator basis sizes for different supports r in 2D square lattice with translational invariance by one site, e.g., in 2D kicked Ising model.

	Truncation r				
	2×2	3×2	3×3	4×3	4×4
Basis size N	228	3792	254 208	16 453 632	4 261 675 008
Classes N_{cl}	10	44	400	3392	57 856

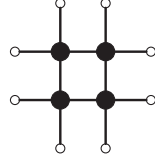


FIG. 18. For $r = 2 \times 2$ truncation and 2D kicked Ising model the initial operator supported on black circles can spread at most by 1 site in each direction (white circles).

10 [Eq. (D1)], in which a product of σ^z is mapped to a product of σ^y at same positions, for instance,

$$\sigma^z \sigma^z \rightarrow \sigma^y \sigma^y, \quad (\text{D4})$$

whereas a product of σ^y is mapped to an operator with larger support, concretely, σ^y are mapped to σ^x at the same position, times a product of σ^y at all sites neighboring the initial σ^y (the above $\sigma^y \sigma^y$ is mapped to a product of two σ^x and six new σ^y at sites connected to the original σ^y).

If one has $h_x \neq 1$, structure is a bit similar. Many eigenvalues are still zero, but Jordan blocks are destroyed. For $r = 2 \times 2$ one gets exactly one nonzero eigenvalue for each Jordan block. Namely, the eigenvalue corresponding to the class with p Pauli σ^z is $\cos^p(h_x \pi/2)$, with the corresponding eigenvector having single-site operator a ,

$$a = -\sin(h_x \pi/2) \sigma^y + \cos(h_x \pi/2) \sigma^z, \quad (\text{D5})$$

at the location of each “x” in our class description Eq. (D1). All of the above properties are insensitive to the value of h_z . Properties of M at $\tau = \pi/4$ are therefore rather interesting and warrant further study.

3. Special point $\tau = \pi/2$

Here one has $e^{-i(\pi/2)\sigma_j^z \sigma_{j+1}^z} = -i\sigma_j^z \sigma_{j+1}^z$, $e^{-i(\pi/2)\sigma_j^x} = -i\sigma_j^x$, and $e^{-i(\pi/2)\sigma_j^z} = -i\sigma_j^z$. Because of coordination number 4 on a 2D square lattice, the two-qubit gates square to one, so that we get $U'_{\pi/2} = \prod_j (i\sigma_j^y)$, i.e., a simple product propagator. The two-step propagator is, on the other hand, trivial $(U'_{\pi/2})^2 = (-1)^L \mathbb{1}$; i.e., one has a DTC. This means that all eigenvalues of $U'_{\pi/2}$ are ± 1 . This carries over to the spectrum of M which also has all its eigenvalues equal to ± 1 . Close to $\tau = \pi/2$, truncations we use, e.g., 3×3 , are not large enough to get an accurate estimate for $|\lambda_1|$, for instance, jagged behavior in Fig. 9(a) (for some τ the largest eigenvalue can even have the modulus larger than 1).

(folding) manifold. A uniformly hyperbolic system is one with stretching exponents uniformly larger than 1; i.e., the stretching never becomes zero, or, stable and unstable directions never become degenerate and coincide.

- [3] D. V. Anosov, *Geodesic flows on closed Riemannian manifolds of negative curvature*, Proc. Steklov Inst. Math. **90**, 1 (1967).
- [4] R. Bowen, *ω -limit sets for axiom A diffeomorphism*, J. Diff. Eqns. **18**, 333 (1975).
- [5] T. Sauer, C. Grebogi, and J. A. Yorke, *How long do numerical chaotic solutions remain valid?*, Phys. Rev. Lett. **79**, 59 (1997).
- [6] T. Hoefler, T. Häner, and M. Troyer, *Disentangling hype from practicality: On realistically achieving quantum advantage*, Commun. ACM **66**, 82 (2023).
- [7] A. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, *Practical quantum advantage in quantum simulation*, Nature (London) **607**, 667 (2022).
- [8] Y. Kim et al., *Evidence for the utility of quantum computing before fault tolerance*, Nature (London) **618**, 500 (2023).
- [9] E. Rosenberg et al., *Dynamics of magnetization at infinite temperature in a Heisenberg spin chain*, Science **384**, 48 (2024).
- [10] P. Zhang et al., *Emergence of steady quantum transport in a superconducting processor*, Nat. Commun. **15**, 10115 (2024).
- [11] O. Shtanko, D. S. Wang, H. Zhang, N. Harle, A. Seif, R. Movassagh, and Z. Mineev, *Uncovering local integrability in quantum many-body dynamics*, Nat. Commun. **16**, 2552 (2025).
- [12] R. Hagshenas et al., *Digital quantum magnetism at the frontier of classical simulations*, arXiv:2503.20870.
- [13] Y. Yang, A. Christianen, S. Collinvent, V. Smelyanskiy, M. C. Banuls, T. E. O'Brien, D. S. Wild, and J. I. Cirac, *Simulating prethermalization using near-term quantum computers*, PRX Quantum **4**, 030320 (2023).
- [14] A. M. Childs and Y. Su, *Nearly optimal lattice simulation by product formulas*, Phys. Rev. Lett. **123**, 050503 (2019).
- [15] L. Clinton, J. Bausch, and T. Cubitt, *Hamiltonian simulation algorithms for near-term quantum hardware*, Nat. Commun. **12**, 4989 (2021).
- [16] M. Heyl, P. Hauke, and P. Zoller, *Quantum localization bounds Trotter errors in digital quantum simulation*, Sci. Adv. **5**, eaau8342 (2019).
- [17] 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).
- [18] 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).
- [19] W. W. Ho, T. Mori, D. A. Abanin, and E. G. Dalla, *Torricelli Quantum and classical Floquet prethermalization*, Ann. Phys. (Amsterdam) **454**, 169297 (2023).
- [20] A. Eckardt, *Colloquium: Atomic quantum gases in periodically driven optical lattices*, Rev. Mod. Phys. **89**, 011004 (2017).

[1] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993).

[2] Hyperbolic systems are, roughly speaking, systems where at each point there is an unstable (stretching) and a stable

- [21] S. Blanes, F. Casas, J. A. Oteo, and J. Ros, *The Magnus expansion and some of its applications*, *Phys. Rep.* **470**, 151 (2009).
- [22] T. N. Ikeda, S. Sugiura, and A. Polkovnikov, *Robust effect of ground state in a nonintegrable Floquet quantum circuit*, *Phys. Rev. Lett.* **133**, 030401 (2024).
- [23] P. M. Schindler and M. Bukov, *Geometric Floquet theory*, *Phys. Rev. X* **15**, 031037 (2025).
- [24] V. Khemani, A. Lazarides, R. Moessner, and S. L. Sondhi, *Phase structure of driven quantum systems*, *Phys. Rev. Lett.* **116**, 250401 (2016).
- [25] D. V. Else, B. Bauer, and C. Nayak, *Floquet time crystals*, *Phys. Rev. Lett.* **117**, 090402 (2016).
- [26] C. W. von Keyserlingk and S. L. Sondhi, *Phase structure of one-dimensional interacting Floquet systems. II. Symmetry-broken phases*, *Phys. Rev. B* **93**, 245146 (2016).
- [27] M. P. Zaletel, M. Lukin, C. Monroe, C. Nayak, F. Wilczek, and N. Y. Yao, *Colloquium: Quantum and classical discrete time crystals*, *Rev. Mod. Phys.* **95**, 031001 (2023).
- [28] M. Pollicott, *On the rate of mixing in axiom A flows*, *Invent. Math.* **81**, 413 (1985).
- [29] D. Ruelle, *Resonances in chaotic dynamical systems*, *Phys. Rev. Lett.* **56**, 405 (1986).
- [30] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, Cambridge, England, 1998).
- [31] T. Prosen, *Ruelle resonances in quantum many-body systems*, *J. Phys. A* **35**, L737 (2002).
- [32] A somewhat similar method has been very recently suggested in Ref. [33]. For circuits their matrix K is equal to $K = 2 \cdot \mathbb{1} - M - M^T$, where M is our operator propagator at $k = 0$.
- [33] J. Pawłowski, J. Herbrych, and M. Mierzejewski, *Algorithm for finding local integrals of motion in quantum lattice models in the thermodynamic limit*, *Phys. Rev. B* **112**, 155130 (2025).
- [34] K. Klobas, *Exact time-dependent solutions of interacting systems*, Ph.D. thesis, University of Ljubljana, 2020.
- [35] R. Sharipov, M. Koterle, S. Grozdanov, and T. Prosen, *Ergodic behaviors in reversible 3-state cellular automata*, *arXiv:2503.16593*.
- [36] M. Žnidarič, U. Duh, and L. Zadnik, *Integrability is generic in homogeneous $U(1)$ -invariant nearest-neighbor qubit circuits*, *Phys. Rev. B* **112**, L020302 (2025).
- [37] M. Žnidarič, *Momentum-dependent quantum Ruelle-Pollicott resonances in translationally invariant many-body systems*, *Phys. Rev. E* **110**, 054204 (2024).
- [38] M. Žnidarič, *Inhomogeneous $SU(2)$ symmetries in homogeneous integrable $U(1)$ circuits and transport*, *Nat. Commun.* **16**, 4336 (2025).
- [39] U. Duh and M. Žnidarič, *Ruelle-Pollicott resonances of diffusive $U(1)$ -invariant qubit circuits*, *SciPost Phys.* **20**, 061 (2026).
- [40] T. Prosen and M. Žnidarič, *Is efficiency of classical simulations of quantum dynamics related to integrability?*, *Phys. Rev. E* **75**, 015202(R) (2007).
- [41] T. Prosen, *Chaos and complexity of quantum motion*, *J. Phys. A* **40**, 7881 (2007).
- [42] Note that we use $J = 1/2$ in order to be able to directly compare timescales to, e.g., Ref. [16], which means that the prefactor in front of the $z - z$ term is $\tau/4$; i.e., $\tau = 0.8$ can indeed be considered “small”.
- [43] S. Choudhury, *Route to extend the lifetime of a discrete time crystal in a finite spin chain without disorder*, *Atoms* **9**, 25 (2021).
- [44] H. Kim, M. C. Bañuls, J. I. Cirac, M. B. Hastings, and D. A. Huse, *Slowest local operators in spin chains*, *Phys. Rev. E* **92**, 012128 (2015).
- [45] C.-J. Lin and O. I. Motrunich, *Explicit construction of quasiconserved local operator of translationally invariant nonintegrable quantum spin chain in prethermalization*, *Phys. Rev. B* **96**, 214301 (2017).
- [46] E. Izotova, *Local versus translationally invariant slowest operators in quantum Ising spin chains*, *Phys. Rev. E* **108**, 024138 (2023).
- [47] M. Mierzejewski, T. Prosen, and P. Prelovšek, *Approximate conservation laws in perturbed integrable lattice models*, *Phys. Rev. B* **92**, 195121 (2015).
- [48] 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).
- [49] 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).
- [50] E. Vernier, B. Bertini, G. Giudici, and L. Piroli, *Integrable digital quantum simulation: Generalized Gibbs ensembles and Trotter transitions*, *Phys. Rev. Lett.* **130**, 260401 (2023).
- [51] L. D’Alessio and M. Rigol, *Long-time behavior of isolated periodically driven interacting lattice systems*, *Phys. Rev. X* **4**, 041048 (2014).
- [52] A. Lazarides, A. Das, and R. Moessner, *Equilibrium states of generic quantum systems subject to periodic driving*, *Phys. Rev. E* **90**, 012110 (2014).
- [53] C. Grebogi, S. M. Hammel, J. A. Yorke, and T. Sauer, *Shadowing of physical trajectories in chaotic dynamics: Containment and refinement*, *Phys. Rev. Lett.* **65**, 1527 (1990).
- [54] R. K. Panda, A. Scardicchio, M. Schulz, S. R. Taylor, and M. Žnidarič, *Can we study the many-body localization transition?*, *Europhys. Lett.* **128**, 67003 (2019).
- [55] E. McCulloch, J. A. Jacoby, C. von Keyserlingk, and S. Gopalakrishnan, *Subexponential decay of local correlations from diffusion-limited dephasing*, *arXiv:2504.05380*.
- [56] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Adison-Wesley, Reading, MA, 1990).
- [57] T. Rakovszky, C. W. von Keyserlingk, and F. Pollmann, *Dissipation-assisted operator evolution method for capturing hydrodynamic transport*, *Phys. Rev. B* **105**, 075131 (2022).
- [58] C. Artiano, C. Fleckenstein, D. Aceituno Chavez, T. K. Kvornig, and J. H. Bardarson, *Efficient large-scale many-body quantum dynamics via local-information time evolution*, *PRX Quantum* **5**, 020352 (2024).
- [59] J. Wang, M. H. Lamann, R. Steinigeweg, and J. Gemmer, *Diffusion constants from the recursion method*, *Phys. Rev. B* **110**, 104413 (2024).

- [60] S. Yi-Thomas, B. Ware, J.D. Sau, and C.D. White, *Comparing numerical methods for hydrodynamics in a one-dimensional lattice spin model*, *Phys. Rev. B* **110**, 134308 (2024).
- [61] Convergence is quite fast with r , perhaps going as $\sim 1/r^6$ for $r \geq 10$ (or even exponential in r). The estimated error bar is conservative. Note that despite rather small gaps, $\approx 3 \times 10^{-5}$, we can with our in situ propagation method and ARPACK treat large support sizes of $r = 13$ where the size of M is $N \approx 5 \times 10^7$.
- [62] T. Prosen and M. Žnidarič, *Matrix product simulations of non-equilibrium steady states of quantum spin chains*, *J. Stat. Mech.* (2009) P02035.
- [63] N. Pottier, *Nonequilibrium Statistical Physics: Linear Irreversible Processes* (Oxford University Press, New York, 2010).
- [64] It would be beneficial to have a set of reference cases (“easy” and “hard”), for instance, in the kicked Ising model, for which one could compare precise values of diffusion constant in order to assess quality of different numerical methods (Ref. [60] is a step in this direction).
- [65] M. Žnidarič, *Nonequilibrium steady-state Kubo formula: Equality of transport coefficients*, *Phys. Rev. B* **99**, 035143 (2019).
- [66] T. Prosen, *Exact time-correlation functions of quantum Ising chain in a kicking transversal magnetic field*, *Prog. Theor. Phys. Suppl.* **139**, 191 (2000).
- [67] A. Pizzi, D. Malz, G. De Tomasi, J. Knolle, and A. Nunnenkamp, *To heat or not to heat: Time crystallinity and finite-size effects in clean Floquet systems*, *Phys. Rev. B* **102**, 214207 (2020).
- [68] W. C. Yu, J. Tangpanitanon, A. W. Glaetzle, D. Jaksch, and D. G. Angelakis, *Discrete time crystal in globally driven interacting quantum systems without disorder*, *Phys. Rev. A* **99**, 033618 (2019).
- [69] C. Fleckenstein and M. Bukov, *Thermalization and prethermalization in periodically kicked quantum spin chains*, *Phys. Rev. B* **103**, 144307 (2021).
- [70] M. Collura, A. De Luca, D. Rossini, and A. Lerose, *Discrete time-crystalline response stabilized by domain-wall confinement*, *Phys. Rev. X* **12**, 031037 (2022).
- [71] D. V. Else, B. Bauer, and C. Nayak, *Prethermal phases of matter protected by time-translation symmetry*, *Phys. Rev. X* **7**, 011026 (2017).
- [72] Z.-Y. Li and Y.-R. Zhang, *Prethermal discrete time crystals in one-dimensional classical Floquet systems with nearest-neighbor interactions*, *Phys. Rev. B* **112**, 134313 (2025).
- [73] T. Haga, *Divergence of the Floquet-Magnus expansion in a periodically driven one-body system with energy localization*, *Phys. Rev. E* **100**, 062138 (2019).
- [74] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users’ Guide* (Society for Industrial and Applied Mathematics, 1998).
- [75] G. Vidal, *Efficient simulation of one-dimensional quantum many-body systems*, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [76] F. Casas and A. Murua, *An efficient algorithm for computing the Baker-Campbell-Hausdorff series and some of its applications*, *J. Math. Phys. (N.Y.)* **50**, 033513 (2009).