PHYSICAL REVIEW A **81**, 022113 (2010)

Unitary equilibrations: Probability distribution of the Loschmidt echo

Lorenzo Campos Venuti[1] and Paolo Zanardi[1,2]

[1]*Institute for Scientific Interchange, Villa Gualino, Viale Settimio Severo 65, I-10133 Torino, Italy*

[2]*Department of Physics and Astronomy and Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089-0484, USA*

(Received 3 July 2009; published 16 February 2010)

Closed quantum systems evolve unitarily and therefore cannot converge in a strong sense to an equilibrium state starting out from a generic pure state. Nevertheless for large system size one observes *temporal typicality*. Namely, for the overwhelming majority of the time instants, the statistics of observables is practically indistinguishable from an effective equilibrium one. In this paper we consider the Loschmidt echo (LE) to study this sort of unitary equilibration after a quench. We draw several conclusions on general grounds and on the basis of an exactly solvable example of a quasifree system. In particular we focus on the whole probability distribution of observing a given value of the LE after waiting a long time. Depending on the interplay between the initial state and the quench Hamiltonian, we find different regimes reflecting different equilibration dynamics. When the perturbation is small and the system is away from criticality the probability distribution is Gaussian. However close to criticality the distribution function approaches a double-peaked, universal form.

DOI: 10.1103/PhysRevA.81.022113 PACS number(s): 03.65.Yz, 05.30.−d

I. INTRODUCTION

A sudden change of the parameters governing the evolution of a closed quantum many-body system gives typically rise to a complex and fascinating dynamics. This so-called Hamiltonian *quench* is attracting an increasing amount of attention [1–4]. The reason for such an interest is, at least, twofold; in the first place this out of equilibrium phenomenon has been recently observed in cold atom systems [5,6]. Second, at a more conceptual level, the equilibration dynamics of a quenched quantum system plays a role in the very foundations of statistical mechanics [7–12]. New insights can be gained on the fundamental question about the emergence of a thermal behavior in closed quantum systems.

In this paper we study a prototypical dynamical quantity for a quantum quench: the Loschmidt echo (LE). This quantity is defined by the square modulus of the scalar product, of the time-evolved, (out of equilibrium) quantum state with the initial (equilibrium) one, e.g., Hamiltonian ground state.

In spite of the simplicity of its definition the Loschmidt echo $\mathcal{L}$, or closely related quantities, convey a great deal of information in a variety of physical problems; for example $\mathcal{L}$ has been intensively studied in the context of Fermi edge singularities in x-ray spectra of metals [13], quantum chaos [14,15], decoherence [16–18], and more recently quantum criticality [19] and out-of-equilibrium fluctuations [20,21].

Typically (cf. Fig. 1) the Loschmidt echo rapidly decays from its maximum value $\mathcal{L} = 1$ at $t = 0$ and, after an initial transient starts oscillating erratically around the same well defined value. For finite size systems after a sufficiently long time a pattern of collapses and revivals is observed due to the almost-periodic nature of the underlying quantum dynamics. On the other hand for infinite volume systems the Hamiltonian spectrum generically becomes continuous and an asymptotic value $\mathcal{L}_\infty$ (coinciding with the average one) is eventually reached.

The main goal here is to investigate the statistical properties of the Loschmidt echo seen as a random variable over the observation time interval $[0, T]$. One of the key properties is that a small variance, by standard probability theory arguments, guarantees that for the overwhelming majority of times, $\mathcal{L}(t)$ sticks very close to its average value [7,10,12]. This is the sense in which one can speak about *equilibration* dynamics and corresponding "equilibrium properties" of a finite system that is evolving unitarily and therefore cannot have attractive fixed points.

We shall show how the features of the probability distribution of the Loschmidt echo depend on a rich interplay between the initial state and quench Hamiltonian on the one hand and the system's size and observation time on the other. In particular we will focus on the potential role that the vicinity of quantum critical points may have on the features of the Loschmidt echo probability distribution function [4,17,19]. This latter analysis will be mostly carried over by exploiting exact results for quasifree spin chain i.e., the quantum Ising model [19].

The paper is organized as follows: in Sec. II we give the general setting. Later we introduce a relaxation time T_R and discuss the universality content of the $\mathcal{L}(t)$ before this time scale. In Sec. II C we define and study other relevant time scales, the time T_1 for necessary for observing the correct average, and revival times where large portion of $\mathcal{L}(t)$ are back in phase. In Sec. II D we give explicit formulas for the moments of the LE assuming the nonresonant hypothesis. In Sec. III we concentrate on a particular example and prove all the general results advocated so far for an exactly solvable case. Moreover we discover three universal behaviors for the whole LE probability distribution function. We draw some parallels with another natural quenched observable: the magnetization. Finally Sec. V is devoted to conclusions and outlook.

II. GENERAL BEHAVIOR

Let us start by recalling a few elementary yet crucial facts. If $H = \sum_n E_n \Pi_n$ is the system's Hamiltonian (Π_n's = spectral projections) the closed-system dynamics is described by the time-evolution superoperator $\mathcal{U}_t = e^{-it\mathcal{H}}, \mathcal{H}(X) = [H, X]$. This superoperator is thought here of as a map of the space of trace-class operators X into itself ($\|X\|_1 := \mathrm{tr}\sqrt{X^\dagger X} < \infty$). Closed quantum systems cannot equilibrate in the strong sense,

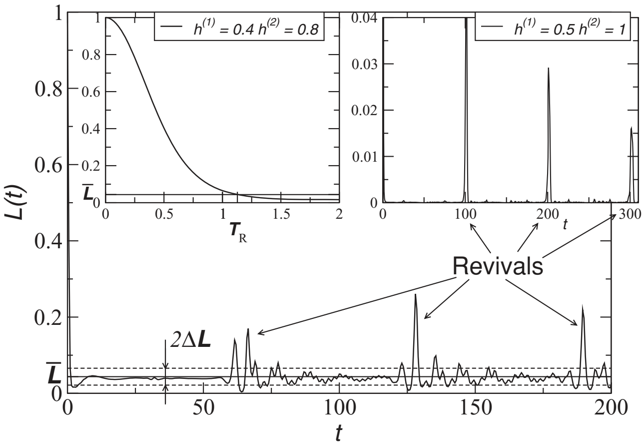


FIG. 1. Typical behavior of the Loschmidt echo for the Ising model in transverse field. All curves refer to a size of $L = 100$. In the upper left panel the relaxation time T_R is indicated. Using arguments as in Sec. III D (see also notes [23,24]) one is able to show that, in the quantum Ising model at criticality, the first revival time is exactly $T_{\rm rev} = L$ (upper left panel).

as unitary evolutions $\mathcal{U}_t$ do not have nontrivial i.e., nonfixed, limit points in the norm topology for $t \to \infty$ [32]. One may then wonder whether a weaker form of convergence can be achieved for $t \to \infty$.

Let us then consider the expectation value of an observable $A(t) := \mathrm{tr}[\mathcal{U}_t(\rho_0)\mathrm{A}]$ and write the spectral resolution of the superoperator $\mathcal{U}_t$ as a formal sum $\mathcal{U}_t = \sum_{\mathcal{E}} e^{-it\mathcal{E}} |\mathcal{E}\rangle\rangle\langle\langle\mathcal{E}|$, here $\mathcal{E}(|\mathcal{E}\rangle\rangle)$ denote the eigenvalues (eigenvector) of $\mathcal{H}$. In finite dimensions the kernel of $\mathcal{H}$ is spanned by the Π_n's and gives rise to a time-independent contribution to $A(t)$ i.e., $A_\infty \sum_n := \mathrm{tr}(\Pi_n \rho_0 \Pi_n A)$; the point is now to understand whether the remaining components involving the nontrivial time-dependent factors $\exp(-i\mathcal{E}t)(\mathcal{E} \neq 0)$ admits a limit for $t \to \infty$. In finite dimensions the $\mathcal{E}$ are a (finite) discrete set of differences of Hamiltonian eigenvalues, e.g., $\mathcal{E} = E_n - E_m$, and correspondingly $A(t) - A_\infty$ is a quasiperiodic function: the *long-time limit* of $A(t)$ does not exist. On the other hand in the infinite-dimensional case the spectrum of $\mathcal{H}$ can be continuous and, if the function $\hat{A}(\mathcal{E})\ \langle\rho_0, \mathcal{E}\rangle\langle\mathcal{E}, A\rangle$ is sufficiently well behaved, using the Riemann-Lebesgue lemma, $\lim_{t\to\infty} \int d\mathcal{E} \hat{A}(\mathcal{E}) \exp(-i\mathcal{E}t) = 0$. Therefore in this case

$$\lim_{t\to\infty} A(t) = \overline{A(t)} = A_\infty \tag{1}$$

where $\overline{A(t)} := \lim_{T\to\infty} \frac{1}{T}\int_0^T A(t)dt$ denotes the time average over an infinite time interval. While this convergence cannot be achieved uniformly for the *whole* set of system's observables (in that it would imply strong convergence) it can be proven for specific families of A's, e.g., local ones [22].

There is a third form of convergence that one can consider here: the convergence in probability. In the following we shall consider the above defined $A(t)$ as a random variable over the real line of t endowed with the uniform measure dt/T with $T \to \infty$. Suppose we have a sequence of $\{B_L\}_L$ (think of L as the system size), we say that the B_L's converge to zero in probability if

$$\lim_{L\to\infty} \Pr\{t \in \mathbf{R}/\,|B_L(t)| \geqslant \epsilon\} = 0, \quad \forall \epsilon > 0. \tag{2}$$

The meaning of this type of convergence should be clear: for large L the probability of observing a value of $B_L(t)$ different from zero is vanishingly small. In other words the fractions of t's for which $B_L(t) \neq 0$ is going to zero for $L \to \infty$. As a matter of fact this is the type of convergence, with $B_L(t) = A_L(t) - A_{L,\infty}$, that has been considered in [10,12]. The stochastic convergence [Eq. (2)] implies that the probability distributions of the random variables A_L i.e., $P_L(\alpha) := \overline{\delta[\alpha - A_L(t)]}$ are converging to the one of $A_\infty := \lim_{L\to\infty} A_{L,\infty}$ i.e., $\lim_{L\to\infty} P_L(\alpha) = \delta(\alpha - A_\infty)$. *In this context we say that the initial state* ρ_0 relaxes or equilibrates to $\rho_{\rm eq}$ if it happens that $A_\infty = \mathrm{tr}(A\rho_{\rm eq})$ [33].

A typical strategy to demonstrate this kind of unitary equilibration is to prove Eq. (2) by showing that (i) $\overline{B_L(t)} = 0$ (ii) $\mathrm{var}(B_L)$ goes to zero sufficiently fast for $L \to \infty$. If this is the case one can use a basic probability theory result $\Pr\{t \in \mathbf{R}/|B_L(t) - \overline{B_L(t)}| \geqslant \epsilon\} \leqslant \mathrm{var}(B_L)/\epsilon^2$. Since, under the assumption (ii) the RHS of this latter relation can be made arbitrarily small and given (i), Eq. (2) holds true.

Notice that yet another way to formulate the kind of convergence [Eq. (2)] is by means of the concept of typicality [8,9]; the probability of observing a "nontypical" value i.e., one that deviates significantly from the mean one becomes negligible in the large L limit.

A. Loschmidt echo

The time dependent quantity we are going to focus on in the rest of this paper is the Loschmidt echo (LE),

$$\mathcal{L}(t) = |\langle\psi|e^{-itH}|\psi\rangle|^2, \tag{3}$$

where the state $|\psi\rangle$ is possibly, but not necessarily, the ground state of the Hamiltonian H at a different coupling. In the sequel, statistical averages are always taken with respect to this state.

In the following we will consider $\mathcal{L}(t)$ as a random variable with uniformly distributed $t \geqslant 0$. Ideally we are interested not only in the first moment but in the whole probability distribution function. The probability of $\mathcal{L}$ to have value in Ω is given by $P(\mathcal{L} \in \Omega) = \lim_{T\to\infty} T^{-1}\mu(\mathcal{L}^{-1}(\Omega) \cap [0, T])$. For those x for which the probability density is well defined, it is given by

$$P(x) = \overline{\delta[x - \mathcal{L}(t)]} = \lim_{T\to\infty} \frac{1}{T} \sum_{\substack{0<t_n<T \\ \mathcal{L}(t_n)=x}} \frac{1}{\left|\frac{d\mathcal{L}}{dt} t_n\right|}.$$

In a realistic setting it can be convenient for the experimenter to prepare many copies of the state $|\psi\rangle$, let each of them evolve with H and measure the LE's after given times $t_{{\rm mes},i}$. In this situation the experimenter will naturally form the histogram of the measured data which gives an approximation to $P(x)$. Mathematically the distribution $P(x)$ can as well be defined by its moments. The kth moment of this probability distribution is given by $\mu_k := \int x^k P(x)dx = \overline{\mathcal{L}^k(t)}$. Notice that the Loschmidt echo can be written as $\mathcal{L}(t) = \langle\rho_\psi, e^{-i\mathcal{H}t}(\rho_\psi)\rangle$, where: $\rho_\psi := |\psi\rangle\langle\psi|$, $\mathcal{H}(X) := [H, X]$ and $\langle X, Y\rangle := \mathrm{tr}(X^\dagger Y)$ denotes the Hilbert-Schmidt scalar product. From this it follows $\mathcal{L}^n(t) = \langle\rho_\psi^{\otimes n}, e^{-i\mathcal{H}^{(n)}t}(\rho_\psi^{\otimes n})\rangle$, where $\mathcal{H}^{(n)} := \sum_{i=1}^n {}^{\otimes(i-1)} \otimes \mathcal{H} \otimes^{\otimes(n-i)}$. Performing the time average [34] one finds $\mu_n = \langle\rho_\psi^{\otimes n}, \mathcal{P}^{(n)}(\rho_\psi^{\otimes n})\rangle$, where $\mathcal{P}^{(n)}$ projects onto the kernel of $\mathcal{H}^{(n)}$. In particular the time average

$\bar{\mathcal{L}} = \mu_1$ is given by

$$\bar{\mathcal{L}} = \langle \rho_\psi, \mathcal{P}^{(1)}(\rho_\psi)\rangle = \langle \mathcal{P}^{(1)}(\rho_\psi), \mathcal{P}^{(1)}(\rho_\psi)\rangle = \mathrm{tr}\big(\rho_{\mathrm{eq}}^2\big), \quad (4)$$

where $\rho_{\mathrm{eq}} := \mathcal{P}^{(1)}(\rho_\psi)$. From the general discussion in Sec. II we know that $\mathcal{P}^{(1)}(X) = \sum_n \Pi_n X \Pi_n$. The effective equilibrium state $\rho_{\mathrm{eq}} = \mathcal{P}^{(1)}(\rho_0)$ is just the ρ_0 totally dephased in the H eigenbasis.

Experimentally to obtain this state one has to perform a projective measurement of H on ρ_0 (with no postselection). This remark suggests a way to give to the average [Eq. (4)] a *direct operational meaning*. Let us first rewrite $\bar{\mathcal{L}}$ as $\mathrm{Tr}(S[\mathcal{P}^{(1)}(\rho_\psi)\otimes\mathcal{P}^{(1)}(\rho_\psi)])$ where S denotes the *swap* operator over $\mathcal{H}^{\otimes 2}$. Now the protocol goes as follows: (i) prepare two copies of ρ_ψ (ii) measure H on both copies, this prepares the state $\Omega\mathcal{P}^{(1)}(\rho_\psi)^{\otimes 2}$ (iii) resort to the interferometric techniques described e.g., in [25] to extract $\mathrm{Tr}(S\Omega)$. Alternatively, by observing that S is an Hermitian operator, i.e., an observable, instead of (iii) one could just (iii)' measure S and obtain its expectation value in Ω. Of course this procedure can be generalized straightforwardly to the nth moment μ_n [35] Therefore we see that the probability distribution P can be in principle reconstructed with higher and higher accuracy by a series of direct measurements on multiple copies of the system without the necessity of measuring the time-series of $\mathcal{L}(t)$.

B. Short time regime and criticality

As already pointed out, typically the LE decays from its maximum value 1 at $t = 0$ and, after an initial transient, starts oscillating erratically around its mean value. In this section we will analyze the universality content of this initial transient and its dependence on the interplay between the initial state $|\psi\rangle$ and the evolving Hamiltonian H.

We start by noticing that the LE Eq. (3) is the square modulus of a characteristic function $\chi(t) = \langle e^{-itH}\rangle$ which is the Fourier transform of the energy probability distribution: $\hat{\chi}(\omega) \equiv \langle\delta(H-\omega)\rangle$. Both χ and the LE can be expressed in terms of the cumulants of H,

$$\chi(t) = \exp\sum_{n=1}^{\infty}\frac{(-it)^n}{n!}\langle H^n\rangle_c, \quad (5)$$

$$\mathcal{L}(t) = \exp 2\sum_{n=1}^{\infty}\frac{(-t^2)^n}{(2n)!}\langle H^{2n}\rangle_c, \quad (6)$$

where $\langle\cdot\rangle_c$ stands for the connected average with respect to $|\psi\rangle$. The sums above starts from $n = 1$ because the zero order cumulant is zero: $\langle H^0\rangle_c = 0$. Since H is a local operator, i.e., a sum of local "variables," we can expect in some circumstances, the central limit theorem (CLT) to apply. More specifically the version of the CLT we are going to consider here, is the following. *In the thermodynamic limit, the probability distribution of the rescaled variable* $Y \equiv (H - \langle H\rangle)/\sqrt{\langle H^2\rangle_c}$ tends to a Gaussian (with variance 1 and mean zero). In other words, all but the second connected moments of Y tend to zero when the volume goes to infinity.

When the CLT applies, for sufficiently large system sizes, the distribution of H will be of the form

$$\tilde{\chi}(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(\omega-\langle H\rangle)^2}{2\sigma^2}\right], \quad \sigma^2 \equiv \langle H^2\rangle_c.$$

One can systematically compute corrections to this formula and order them as inverse powers of the system size L. Fourier transforming back we obtain the characteristic function and the LE

$$\chi(t) = e^{it\langle H\rangle}e^{-1/2t^2\sigma^2} \Rightarrow \mathcal{L}(t) = e^{-t^2\sigma^2}. \quad (7)$$

It may seem that this expression for the LE could have be obtained right away by keeping only the first term in the expansion of the exponential in Eq. (6),

$$\mathcal{L}(t) \simeq 1 - t^2\sigma^2 \simeq e^{-t^2\sigma^2}.$$

This is simply a quadratic approximation, that does not rely on the CLT. Its validity requires $t^2\sigma^2 \ll 1$ that in turn implies $t \ll 1/\sqrt{\langle H_2\rangle}$ (as will be explained later this means roughly $t \ll L^{-d/2}$ where d is the space dimension). From Fig. 1 we see that typically $\mathcal{L}(t)$ decays from 1 and after an initial transient, starts oscillating around an average value $\bar{\mathcal{L}}$ which will be computed below. Now, when the CLT applies, Eq. (7) can help us to define this transient or *relaxation* time T_R given by $e^{-T_R^2\sigma^2} = \bar{\mathcal{L}}$. Roughly after this time one starts seeing oscillations in $\mathcal{L}(t)$. Since in general (see below) one has $\bar{\mathcal{L}} \simeq e^{-fL^d}$, and when the CLT applies σ^2 scales like the system volume, the relaxation time scales as $T_R = \sqrt{-\ln\bar{\mathcal{L}}/\sigma^2} = O(L^0)$. The situation is different when one considers small variations of the parameters δh. In this limit the average LE is related to the well studied ground-state fidelity $F = |\langle\psi^{(1)} \mid \psi^{(2)}\rangle|$. More precisely one has $\bar{\mathcal{L}} \simeq F^4$ [18]. Close to critical points the behavior of the fidelity is dictated by the scaling dimension Δ of the most relevant operator in H with respect to the critical state $|\psi\rangle$ [26]. The precise scaling is the following: $F \sim 1 - \mathrm{const}\times\delta h^2 L^{2(d+\zeta-\Delta)}$, where ζ is the dynamical critical exponent. Similarly one can show (see below) that $\sigma^2 \sim L^{2(d-\Delta)}$. All in all this amounts to saying that the relaxation time for small variation δh around a critical point (roughly $\delta h \ll L^{-(d+\zeta-\Delta)}$) increases from $O(1)$ to $T_R \sim L^\zeta$. In the thermodynamic limit instead, i.e., taking first the limit $L\to\infty$, using standard scaling arguments, one can show that for δh small and h close to the critical point h_c, the relaxation times diverges as $T_R \sim |h-h_c|^{-\zeta\nu}$, ν being the correlation length exponent. In Fig. 2 we plot the relaxation time for a concrete example that will be studied thoroughly in Sec. III, the Ising model in transverse field. There one has $\Delta = \zeta = \nu = 1$, and so the singularities observed in Fig. 2 are of simple algebraic type: $T_R \sim |h-h_c|^{-1}$ around the critical points $h_c = \pm 1$.

Let us now turn to discuss when we expect the CLT to work. Consider first the case when $|\psi\rangle$ is the ground state of a gapped Hamiltonian and the connected energy correlators go to zero exponentially fast: $\langle H(x)H(y)\rangle\langle H(x)\rangle \overset{|x-y|\to\infty}{\to} \langle H(y)\rangle$ (exponential clustering). In this case the connected averages of H scale as the volume: $\langle H^n\rangle_c \sim L^d$. As a consequence the cumulants of the rescaled variable satisfy $\langle Y^n\rangle_c \sim L^{-(nd-2d)/2}$ for $n \geqslant 2$, which immediately implies the CLT in the sense given above.

Therefore we can have violation of the CLT only in the gapless case when the state $|\psi\rangle$ is critical or when clustering fails. Let us then consider a critical state $|\psi\rangle$. Connected averages have a regular extensive part and a singular part which scales according to the most relevant component of

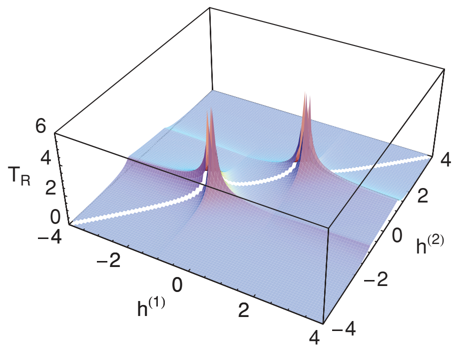


FIG. 2. (Color online) Relaxation time for the Loschmidt echo in the Ising model in transverse field. The critical points are at $h_c = \pm 1$. Clearly we observe divergences at critical points when δh is small. On the line $h^{(1)} = h^{(2)}$ $\mathcal{L}(t) = 1$ and so there is no relaxation or even dynamics.

H with scaling dimension Δ. When $|\psi\rangle$ is the ground state of H at a different coupling, Δ is the scaling dimension of the perturbation δH to the critical Hamiltonian. At leading order, we can write $\langle H^n \rangle_c \sim A_n L^d + B_n L^{n(d-\Delta)}$, and so the rescaled variable satisfies

$$\langle Y^n \rangle_c \sim \frac{A_n L^d + B_n L^{n(d-\Delta)}}{(A_2 L^d + B_2 L^{2(d-\Delta)})^{n/2}}.$$

If $\Delta < d/2$ the cumulants of the rescaled variable Y do not go to zero but to universal constants [27,28] given by

$$\langle Y^n \rangle_c \to \frac{B_n}{B_2^{n/2}}.$$

In this case the probability distribution of the energy is a, non-Gaussian, universal distribution. This kind of universal behavior has been observed for instance in [29] on an example where the scaling dimension is $\Delta = 1/8$. In the opposite situation where $\Delta > d/2$, all the cumulants of Y go to zero except for the first two, and the distribution function approaches a Gaussian in the large size limit. In the intermediate case $\Delta = d/2$ the cumulants of Y do not go zero but to a constant which is however not universal due to extensive contributions coming from the denominator. We recall that in d dimensional, zero temperature quantum mechanics, operators are classified into relevant, irrelevant, and marginal if their scaling dimension is, respectively, smaller, larger, or equal to $d + \zeta$ where ζ is the dynamical exponent. Hence we see that, even in the critical case, we observe deviation from the Gaussian behavior only if the perturbation δH is sufficiently relevant, specifically $\Delta < d/2$.

To finish let us remind the reader that the CLT also breaks down when clustering fails. To summarize, when the CLT applies, the LE tends to a Gaussian and plotting the function $\mathcal{L}_L(t/\sqrt{\langle H^2 \rangle_2})$ for different sizes L one should observe data collapse (see Fig. 3).

C. Equilibration and long-time behavior

After having discussed the short-time behavior of the LE related to the initial transient, let us now turn to its long-time behavior. We first rewrite Eq. (3) in the eigenbasis of

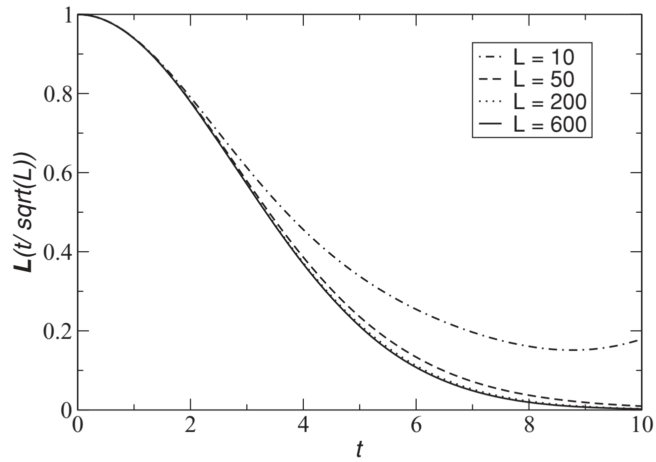


FIG. 3. Rescaled Loschmidt echo $\mathcal{L}_L(t/\sqrt{L})$ for the Ising model in transverse field. The state which defines the average is critical: $h^{(1)} = 1$ while H is at $h^{(2)} = 2.5$. The same data collapse feature is observed when choosing different $h^{(i)}$s, although the variance of this Gaussian is sensitive to that.

$$H = \textstyle\sum_n E_n |n\rangle\langle n|$$

$$\mathcal{L}(t) = \sum_{n,m} p_n p_m e^{-it(E_n - E_m)}, \qquad (8)$$

where $p_n = |\langle \psi \mid n \rangle|^2$.

If the spectrum of H is non degenerate the superoperator $\mathcal{P}^{(1)}$ acts as a dephasing in the Hamiltonian eigenbasis, i.e., $\mathcal{P}^{(1)}(X) = \sum_n \langle n|X|n\rangle |n\rangle\langle n|$. In other words the time average of the exponentials in Eq. (8) gives simply $\delta_{n,m}$ and Eq. (4) reduces to $\bar{\mathcal{L}} = \sum_n p_n^2$. As is well known this quantity is the purity of an equilibrium, dephased, state: $\rho_{\text{eq}} = \sum_n p_n |n\rangle\langle n|$.

Time scales. In the preceding section we already defined a relevant time scale, the relaxation time T_R which is $O(1)$ off-criticality while $T_R = O(L^\zeta)$ in the critical case and for sufficiently small variations $\delta h \ll L^{-(d+\zeta-\Delta)}$.

In some situations it is useful to consider a finite observation time T. We will write $\bar{\mathcal{L}}$ to indicate the corresponding average. It is natural to ask about the interplay between the observation time T and the linear size of the system L. In other words in general $\lim_{L\to\infty}\lim_{T\to\infty}\overline{\mathcal{L}_L} \neq \lim_{T\to\infty}\lim_{L\to\infty}\overline{\mathcal{L}_L}$. Since taking larger system sizes has the effect of sending the revival times to infinity and the LE attunes its maximum value 1 at $t = 0$, typically the function $\lim_{L\to\infty}\mathcal{L}_L(t)$ has only one large peak at $t = 0$ whereas $\mathcal{L}_L(t)$ has peaks at all the revival times. Correspondingly we expect $\lim_{L\to\infty}\lim_{T\to\infty}\overline{\mathcal{L}_L} > \lim_{T\to\infty}\lim_{L\to\infty}\overline{\mathcal{L}_L}$. This expectation has been confirmed for the case of the one-dimensional quantum Ising model, see Sec. III.

Another question which is relevant in the measurement process is how large must the observation time be to effectively measure $\bar{\mathcal{L}}$? That is, what is the condition to have $\overline{\mathcal{L}^{T_1}} = \bar{\mathcal{L}}$ or more generally what is the smallest time T_n such that one observes $\overline{[(\mathcal{L})^n]^{T_n}} = \overline{(\mathcal{L})^n}$? Let us focus on T_1. Looking at Eq. (8) one realizes that it suffices to have $T \gg \Delta_{\text{min}}^{-1}$, where Δ_{min} is the smallest gap in the whole spectrum i.e., $\Delta_{\text{min}} = \min_{n,m}(E_n - E_m)$. We can address this question for the class of quasifree Fermi systems in d spatial dimensions.

In this case the energy has the form $E_n = \sum_k n_k \Lambda_k$ where k is a d-dimensional quasimomentumlike label and we can assume the one-particle energy Λ_k to be positive. Then the gap is given by $\Delta_{\min} = \min_{\beta_k} |\sum_k \beta_k \Lambda_k|$ with $\beta_k = 0, \pm 1$. By choosing $\beta_k = (-1)^k$, one obtains a $\Delta_{\min}$ which is exponentially small in L in those (frequent) cases where Λ_k is an analytic function of $\boldsymbol{k}$ [30]. However in quasifree systems the weights p_n decrease exponentially with the number of excitations in n. In practice the highest weight is given for energy differences between the one and zero particle spectra: $\Delta^{(1,0)} = \min_{\boldsymbol{k}} \Lambda_{\boldsymbol{k}}$ which is a constant of order 1 in the gapful case, while typically scales as L^{-1} for the critical case. The next largest amount of spectral weight is attained at a gap which is a difference between one-particle energies $\Delta^{(1,1)} = \min_{\boldsymbol{k},\boldsymbol{q}} |\Lambda_{\boldsymbol{k}} - \Lambda_{\boldsymbol{q}}|$. It will be favorable to have $\boldsymbol{k}$ and $\boldsymbol{q}$ nearby in the region where $\Lambda_{\boldsymbol{k}}$ is flat or almost flat. So we get $\Delta^{(1,1)} = \min_{\boldsymbol{k}} |\Lambda_{\boldsymbol{k}} - \Lambda_{\boldsymbol{k}+\delta \boldsymbol{k}}| \simeq \min_{\boldsymbol{k}} |\nabla_{\boldsymbol{k}} \Lambda_{\boldsymbol{k}} \cdot \delta \boldsymbol{k}|$. This gap is at least of order of L^{-2} (or at least $O(L^{-3})$ if there exists a $\boldsymbol{k}$ vector such that $\nabla_{\boldsymbol{k}} \Lambda_{\boldsymbol{k}} = 0$). From this discussion we estimate that, at least in quasifree systems, to have $\overline{\mathcal{L}^{T_1}} \simeq \bar{\mathcal{L}}$ one must take $T_1 = O(1)$ in the gapful case, while one has $T_1 = O(L)$ at criticality. If one needs $\overline{\mathcal{L}^{T_1}} \simeq \bar{\mathcal{L}}$ with a larger degree of precision than one must choose considerably larger time: $T_1 = O(L^2)$ [or $T_1 = O(L^3)$ if $\nabla_{\boldsymbol{k}} \Lambda_{\boldsymbol{k}} = 0$ has a solution within the allowed set of $\boldsymbol{k}$ vectors]. Related time scales are *revival times*. We define a revival time to be that particular time for which a large portion of spectral weight $p_n p_m$ has revived. More precisely $T_{\mathrm{rev}} \omega_{\mathrm{peak}} = 2\pi$, where ω_{peak} is a particular frequency $E_n - E_m$ such that the weight $p_n p_m$ is large. From the discussion above we expect $T_{\mathrm{rev}} = O(1)$ when H has a gap above the ground state, while $T_{\mathrm{rev}} = O(L)$ when H is critical. These expectations have been confirmed (see Fig. 1) on the hand of a solvable model that will be discussed in the next sections.

D. Moments of the Loschmidt echo

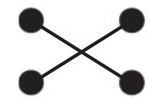
Having computed the time averaged LE we can now turn to higher moments. In doing this one has to distinguish cases where $n = m$ from those where $E_n = E_m$ in Eq. (8). So we write

$$\mathcal{L}(t) = \bar{\mathcal{L}} + X(t),$$
$$X(t) = \sum_{n \neq m} p_n p_m e^{-it(E_n - E_m)},$$

and the nth moment is given by

$$\overline{[\mathcal{L}(t)]^n} = \sum_{k=0}^{n} \binom{n}{k} \bar{\mathcal{L}}^{n-k} \overline{[X(t)]^k}.$$

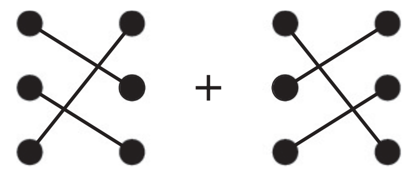
The computation of the average $\overline{[X(t)]^k}$ can be done assuming a *strong nonresonance* condition. With this we mean the following. We say H satisfies a k-nonresonance condition if the only way to fulfill $\sum_{l=1}^{k} E_{i_l} - E_{j_l} = 0$ is to match the E_i's to the E_j's. Strong nonresonance is k-nonresonance for any k. Note that this condition cannot be fulfilled when k becomes of the order of the Hilbert's space dimension. Now to compute $\overline{[X(t)]^k}$ draw $2k$ points in two rows of length k. Imagine $E_i (E_j)$ are the points at the left (right). Now draw all possible contraction between i's and j's (no contraction among i's or j's since they have the same sing), which are $k!$, but *keep only those sets of contractions where there is no horizontal line*. This requirement corresponds to the constraint $i_l \neq j_l, l = 1, \ldots, k$. For example for $[X(t)]^2$ we have only one contribution



so

$$\overline{[X(t)]^2} = \sum_{i_1 \neq i_2} p_{i_1}^2 p_{i_2}^2$$

For $\overline{[X(t)]^3}$ we have two diagrams,



Both these diagrams give the same contribution (simply swap i's with j's) and the result is

$$\overline{[X(t)]^3} = 2 \sum_{\substack{i_1 \neq i_2 \\ i_2 \neq i_3, i_3 \neq i_1}} p_{i_1}^2 p_{i_2}^2 p_{i_3}^2.$$

The number of terms in $\overline{[X(t)]^k}$, $N(k)$ is the number of all permutations without fixed points and is given by

$$N(k) = \sum_{j=2}^{k} (-1)^{k-j} \binom{k}{j} (j! - 1).$$

However, among these $N(k)$ terms, many of them give different contributions. Look for instance at $\overline{[X(t)]^4}$,

$$\overline{[X(t)]^4} = 3 \left(\sum_{i_1 \neq i_2} p_{i_1}^2 p_{i_2}^2 \right)^2 + 6 \sum_{\substack{i_1 \neq i_2, i_2 \neq i_3 \\ i_3 \neq i_4, i_4 \neq i_1}} p_{i_1}^2 p_{i_2}^2 p_{i_3}^2 p_{i_4}^2$$

Correctly one has $3 + 6 = N(4) = 9$.

We collect here the first three moments

$$\mu_1 = \bar{\mathcal{L}},$$
$$\mu_2 = \bar{\mathcal{L}}^2 + \sum_{i_1 \neq i_2} p_{i_1}^2 p_{i_2}^2,$$
$$\mu_3 = \bar{\mathcal{L}}^3 + 3\, \bar{\mathcal{L}} \sum_{i_1 \neq i_2} p_{i_1}^2 p_{i_2}^2 + 2 \sum_{\substack{i_1 \neq i_2 \\ i_2 \neq i_3, i_3 \neq i_1}} p_{i_1}^2 p_{i_2}^2 p_{i_3}^2,$$

while the cumulants are

$$\kappa_1 = \bar{\mathcal{L}}$$
$$\kappa_2 = \sum_{i_1 \neq i_2} p_{i_1}^2 p_{i_2}^2$$
$$\kappa_3 = 2 \sum_{\substack{i_1 \neq i_2 \\ i_2 \neq i_3, i_3 \neq i_1}} p_{i_1}^2 p_{i_2}^2 p_{i_3}^2.$$

We can notice that each term in $\overline{[X(t)]^k}$ has the same form of $\bar{\mathcal{L}}^k$ except for a number of nonresonance constraints of the form $i_l \neq i_m$. Correspondingly $\overline{[X(t)]^k} < N(k)\bar{\mathcal{L}}^k$. Using now $\sum_{k=0}^{n} \binom{n}{k} N(k) = n!$ we obtain the simple bound $\overline{\mathcal{L}^n} < n!\bar{\mathcal{L}}^n$ for $n \geqslant 2$. This means that

$$e^{\lambda\bar{\mathcal{L}}} \leqslant \tilde{\chi}(\lambda) = \overline{e^{\lambda\mathcal{L}}} < \frac{1}{1-\bar{\mathcal{L}}\lambda},$$

and so we obtained a bound on the characteristic function $\tilde{\chi}$. We see that, when $\bar{\mathcal{L}} \to 0$ the probability distribution of $\mathcal{L}$ becomes a delta function at zero (the characteristic function becomes identically one). The distribution function of the upper bound is

$$\vartheta(x)\frac{e^{-x/\bar{\mathcal{L}}}}{\bar{\mathcal{L}}}.$$

Later we will encounter situations where this function gives a good approximation to the Loschmidt echo probability distribution.

III. ISING MODEL IN TRANSVERSE FIELD

From now on we will give a detailed description of the Loschmidt echo for the case of an exactly solvable model. The model we consider is the Ising model in transverse field with Hamiltonian

$$H = -\sum_i \left(\sigma_i^x \sigma_{i+1}^x + h\sigma_i^z\right).$$

This model can be mapped to quasifree fermions and so diagonalized exactly. At zero temperature we distinguish two phases: (i) an ordered one in the longitudinal direction in which $\langle\sigma_i^x\sigma_j^x\rangle m^2$, for $|h| < 1$, and (ii) a paramagnetic phase for $|h| > 1$. The points $h = \pm 1$ are critical points where the system is described by a conformal invariant field theory with central charge $c = 1/2$.

The LE is given in this case by [16] (superscripts, inserted here for clarity, refer to different values of the coupling constant h)

$$\mathcal{L}(t) = |\langle\psi^{(1)}|e^{-itH^{(2)}}|\psi^{(1)}\rangle|^2 = \prod_{k>0}\left[1 - \sin^2\left(\vartheta_k^{(1)} - \vartheta_k^{(2)}\right)\sin^2\left(\Lambda_k^{(2)}t/2\right)\right], \quad (9)$$

where $\tan(\vartheta_k^{(i)}) = -\sin(k)/[h^{(i)} + \cos(k)]$ and the single particle fermionic dispersion is $\Lambda_k^{(i)} = 2\sqrt{[h^{(i)} + \cos(k)]^2 + \sin(k)^2}$. The band minimum (maximum) is at $E_m = 2\,\min\{|1 - h^{(2)}|, |1 + h^{(2)}|\}(E_M = 2\max\{|1 - h^{(2)}|, |1 + h^{(2)}|\})$. Finally, for periodic boundary conditions that will be used throughout, the quasimomenta satisfy $k_n = \pi(2n+1)/L, n = 0, 1, \ldots, L/2 - 1$. Exploiting the fact that H decomposes into a direct sum of $L/2$ blocks 4×4, we are able to compute the complete dephased equilibrium state $\rho_{\rm eq} = \sum_n p_n |n\rangle\langle n|$. The result is

$$\rho_{\rm eq} = \sum_{\alpha\in\mathbb{Z}_2^L} p(\alpha)\,|\alpha\rangle\,\langle\alpha|, \quad (10)$$

where the multi-index α is $\alpha = (\alpha_1, \ldots, \alpha_L)$, $\alpha_i = 0, 1$, the state is $|\alpha\rangle = \otimes_{k>0}|\alpha_k\rangle$ with $|0_k\rangle = \cos(\vartheta_k^{(2)}/2)|0,0\rangle_{k,-k} - i\sin(\vartheta_k^{(2)}/2)|1,1\rangle_{k,-k}$ and $|1_k\rangle = i\sin(\vartheta_k^{(2)}/2)|0,0\rangle_{k,-k} - \cos(\vartheta_k^{(2)}/2)|1,1\rangle_{k,-k}$. Finally the weights are given by

$$p(\alpha) = \prod_{k>0}\frac{\tan^{2\alpha_k}(\delta\vartheta_k/2)}{1 + \tan^2(\delta\vartheta_k/2)}. \quad (11)$$

Using Eq. (10) together with Eq. (11) one can show (cf. Ref. [31]) that the dephased state has the following totally factorized form:

$$\rho_{\rm eq} = \bigotimes_{k>0}\left(a_k\,|0_k\rangle\,\langle 0_k| + b_k\,|1_k\rangle\,\langle 1_k|\right),$$

where $a_k = [1 + \tan^2(\delta\vartheta_k/2)]^{-1}$, and $b_k = 1 - a_k$.

A. Short time regime

Let us first discuss the short-time, transient regime. Looking at the function $\ln\,\mathcal{L}(t)$ one can readily see that all its n derivatives at $t = 0$ are the Riemann sums of a summable function irrespective of the $h^{(i)}$s being critical. This means that all the derivatives of $\ln\mathcal{L}(t)$ grow linearly with L, and together with Eq. (6) implies that the cumulants are linear even at criticality, i.e., $\langle H^{2n}\rangle_c \propto L$. The same result could have been derived by noting that for $|h| \neq 1$ the system is gapful and clustering. When $|\psi^{(1)}\rangle$ is critical, i.e., $|h^{(1)}| = 1$, the scaling dimension of $\delta H = H^{(2)} - H^{(1)} = -\delta h\sum_i \sigma_i^z$ is one and so according to the reasoning in Sec. II B the cumulants of H grows linearly with L.

Accordingly the CLT applies. Since all the cumulants, including the variance, grow as L, plotting the function $\mathcal{L}_L(t/\sqrt{L})$ for different sizes L, one observes data collapse. This behavior is illustrated in Fig. 3. Clearly the plot reproduces a Gaussian with variance $\lim_{L\to\infty}\langle H^2\rangle_c/L$.

B. Long time, large sizes, and the order of limits

Consider now a physical situation where an experimenter computes $\overline{\mathcal{L}_L}$. We inserted the labels T and L to stress that both size and expectation time are finite, as is required in a true experiment. Here we want to study the interplay between T and L. Consider first the case where we send L to infinity, or more physically L is the largest scale of our system. In this situation the spectrum of H is practically continuous, and we can write the LE as

$$\mathcal{L}(t) = \exp\left\{\frac{L}{2\pi}\int_0^\pi \ln\left[1 - \sin^2\left(\vartheta_k^{(1)} - \vartheta_k^{(2)}\right)\,\sin^2\left(\Lambda_k^{(2)}t/2\right)\right]dk\right\} \equiv e^{-Ls(t)}.$$

In this approximation we sent all the revival times to infinity and so the function $\mathcal{L}(t)$ is no longer almost periodic, but rather tends to a precise limit as $t \to \infty$. We can calculate the limit $s(\infty)$ and also the first corrections, as $t \to \infty$. The procedure is outlined in the Appendix. The result is

$$s(t) = s(\infty) - \frac{A_m}{|t|^{3/2}}\cos\left(tE_m + \frac{3}{4}\pi\right) + (m \leftrightarrow M), \quad (12)$$

where $s(\infty)$ is the limiting value and $A_{m/M}$ are constants which depend on $h^{(i)}$ and are given in the appendix. The result (A1) has already been found in [4], here we provide the explicit form of the asymptotic value, $s(\infty)$.

Since the function $\mathcal{L}(t) = e^{-Ls(t)}$ has a limit at infinity, its time average is precisely this limit $\bar{\bar{\mathcal{L}}} e^{-Ls(\infty)}$. More precisely the distribution function becomes a delta function $P(x) = \delta(x - e^{-Ls(\infty)})$.

Consider now performing first the time average of Eq. (9). According to the discussion in Sec. II C, this requires at least observation times as large as $T \gg L$ (if we are at criticality). The result in this case is (for the explicit computation see the following section)

$$\bar{\mathcal{L}} = \exp \sum_{k>0} \log\left[1 - \sin^2\left(\vartheta_k^{(2)} - \vartheta_k^{(1)}\right)/2\right]. \tag{13}$$

If now L is large, we can approximate the sum with the integral: $\bar{\mathcal{L}} = e^{-Lg(h^{(1)},h^{(2)})}$. Calling $\delta\vartheta_k = \vartheta_k^{(2)} - \vartheta_k^{(1)}$ the two functions, g and $s(\infty)$ are given by

$$g = -\frac{1}{2\pi}\int_0^\pi \ln[1 - \sin^2(\delta\vartheta_k)/2]dk, \tag{14}$$

$$s(\infty) = -\frac{1}{\pi}\int_0^\pi \ln\{[1 + |\cos(\delta\vartheta_k)|]/2\}dk. \tag{15}$$

We observed that the two averages e^{-Lg}—obtained by first doing the time average and then taking large L—or $e^{-Ls(\infty)}$—obtained by first considering L large and then doing the time average—are qualitatively very similar for most values of the parameters $h^{(i)}$. The only region where there is an appreciable difference is when $h^{(1)}$ and $h^{(2)}$ correspond to different phases (either $|h^{(1)}| < 1$ and $|h^{(2)}| > 1$ or vice-versa).

C. Moments of the Loschmidt echo in presence of degeneracy

Having the explicit form of the LE Eq. (9) we can compute its time average and also other moments. Since the Ising model is mapped to a free Fermi system on a finite lattice, and given the form of the quasiparticle dispersion Λ_k, its spectrum is nondegenerate. In other words the Ising Hamiltonian is 1-nondegenerate. However, for the same reason, it is not k-nondegenerate for $k \geqslant 2$. This means that to compute moments higher than the first, we really need to use the explicit form Eq. (9) and cannot rely on the results of Sec. II C. For the first moment this problem does not arise, and we can either use $\bar{\mathcal{L}} = \sum_n p_n^2$ or do the time average of Eq. (9). Correctly the results coincide, and they rely on the fact that $\sum_k (n_k - m_k)\Lambda_k = 0$, implies $n_k = m_k$, i.e., that the spectrum is nondegenerate. The result is

$$\bar{\mathcal{L}} = \prod_{k>0}\left[1 - \sin^2\left(\vartheta_k^{(1)} - \vartheta_k^{(2)}\right)/2\right]. \tag{16}$$

For later convenience we define $\alpha_k = \sin^2(\vartheta_k^{(1)} - \vartheta_k^{(2)})$. To compute higher moments we first rewrite Eq. (9) as

$$\mathcal{L}(t) = \prod_{k>0}[1 + X_k(t)],$$

$$X_k(t) = -\alpha_k \sin^2(\Lambda_k t/2) = \sum_{\beta=0,\pm1} c_\beta^k e^{i\beta\Lambda_k t},$$

$$c_0^k = -\frac{\alpha_k}{2}, \quad c_{\pm1}^k = \frac{\alpha_k}{4}.$$

We write the nth power of the LE as

$$[\mathcal{L}(t)]^n = \prod_{k>0}\left[1 + Y_k^{(n)}(t)\right],$$

$$Y_k^{(n)}(t) = \sum_{m=1}^{n}\binom{n}{m}[X_k(t)]^m \equiv \sum_{\gamma=0,\pm1,\ldots,\pm m} g_{\gamma,k}^{(n)} e^{i\gamma\Lambda_k t}.$$

Now, when computing M products of $Y_k^{(n)}$ terms, only the $\gamma = 0$ term will survive after taking the time average. In other words

$$\overline{Y_{k_1}^{(n)}\cdots Y_{k_M}^{(n)}} = g_{0,k_1}^{(n)}\cdots g_{0,k_M}^{(n)},$$

and so we have

$$\overline{[\mathcal{L}(t)]^n} = \prod_{k>0}\left(1 + g_{0,k}^{(n)}\right).$$

An explicit formula for $g_{0,k}^{(n)}$ is

$$g_{0,k}^{(n)} = \sum_{m=1}^{n}\binom{n}{m}\sum_{\substack{\beta_1,\ldots,\beta_m\\ \sum\beta_i=0}} c_{\beta_1}^k\cdots c_{\beta_m}^k.$$

Noting that

$$\sum_{\substack{\beta_1,\ldots,\beta_m\\ \sum\beta_i=0}} c_{\beta_1}^k\cdots c_{\beta_m}^k = \sum_{\substack{n_1\\ n_0+2n_1=m}}\frac{m!}{(n_1!)^2\, n_0!}\left(c_0^k\right)^{n_0}\left(c_1^k\right)^{2n_1},$$

we obtain

$$g_{0,k}^{(n)} = \sum_{m=1}^{n}\left(\frac{-\alpha_k}{4}\right)^m\binom{n}{m}\frac{\partial_t^m(2t - t^2 - 1)_{t=0}^m}{m!}$$
$$= \sum_{m=1}^{n}\left(\frac{-\alpha_k}{4}\right)^m\binom{n}{m}\binom{2m}{m}.$$

For example we have

$$g_{0,k}^{(1)} = -\frac{\alpha_k}{2},$$

$$g_{0,k}^{(2)} = -\alpha_k + \frac{3}{8}\alpha_k^2,$$

$$g_{0,k}^{(3)} = -\frac{3}{2}\alpha_k + \frac{9}{8}\alpha_k^2 - \frac{5}{16}\alpha_k^3,$$

$$g_{0,k}^{(4)} = -2\alpha_k + \frac{9}{4}\alpha_k^2 - \frac{5}{4}\alpha_k^3 + \frac{35}{128}\alpha_k^4.$$

The variance of the LE is then given by

$$\overline{\Delta\mathcal{L}^2} = \prod_{k>0}\left(1 - \alpha_k + \frac{3}{8}\alpha_k^2\right) - \prod_{k>0}\left(1 - \alpha_k + \frac{1}{4}\alpha_k^2\right). \tag{17}$$

The first moment and the variance are plotted in Fig. 4. One should note that close to the critical points $h_c = \pm 1$ there appears a small region δh where the variance is large.

Equation (17) gives explicitly the variance in a case where the nonresonant hypothesis is violated. Since $|\alpha_k| \leqslant 1$ generally the variance is given by the difference between two exponentially small quantities and so, a fortiori, is exponentially small in the system size L. However, looking at Fig. 4 one notes a small region of parameter close to the

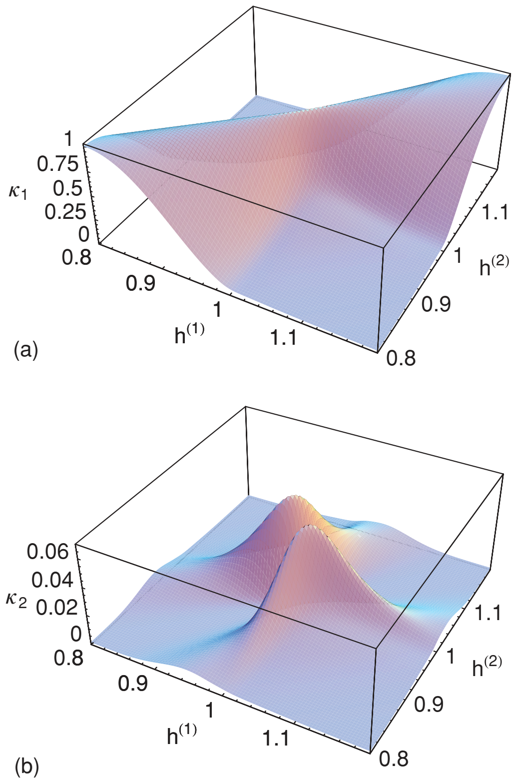


FIG. 4. (Color online) From top to bottom, mean, and variance of Loschmidt echo at size $L = 100$. The region where the variance is large shrinks when increasing the system size L. The height of the peak however remains constant.

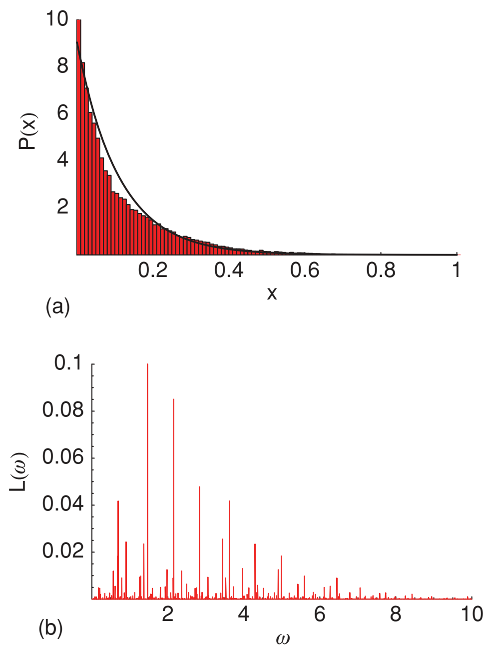


FIG. 5. (Color online) Approximate exponential behavior. Parameters are $L = 18$, $h^{(1)} = 0.3$, $h^{(2)} = 1.4$. When δh is large this behavior is observed even for moderate sizes (here $L = 18$), (upper panel). The thick line reproduces $\vartheta(x)e^{-x/\bar{\mathcal{L}}}/\bar{\mathcal{L}}$ with $\bar{\mathcal{L}}$ given by Eq. (16). In the lower panel we plot the Fourier series $\bar{\mathcal{L}}_{\text{disc}}(\omega)$ given by Eq. (20).

critical points, where the variance is large. As we will see this fact has important consequences.

It is now interesting to compare the result in Eq. (17) with what would have been obtained assuming a nonresonant condition. Clearly the second moment computed assuming nonresonance has less terms than the correct one. Since for the LE all the contributions are positive, the nonresonant result ought to be smaller. In other words, for the variance we must have $\overline{\Delta\mathcal{L}^2} \geqslant \overline{\Delta\mathcal{L}^2_{\text{nr}}}$.

Comparison with the nonresonant result. To compute the variance assuming nonresonance we use

$$\overline{\mathcal{L}^2_{\text{nr}}} = \bar{\mathcal{L}}^2 + 2\sum_{i<j} p_i^2 p_j^2 = 2\bar{\mathcal{L}}^2 - \sum_i p_i^4.$$

Using this formula together with Eq. (11) we obtain

$$\overline{\Delta\mathcal{L}^2_{\text{nr}}} = \bar{\mathcal{L}}^2 - \prod_{k>0}\left(1 - \alpha_k + \frac{1}{8}\alpha_k^2\right)$$
$$= \prod_{k>0}\left(1 - \alpha_k + \frac{1}{4}\alpha_k^2\right) - \prod_{k>0}\left(1 - \alpha_k + \frac{1}{8}\alpha_k^2\right).$$

We have verified that the inequality $\overline{\Delta\mathcal{L}^2} \geqslant \overline{\Delta\mathcal{L}^2_{\text{nr}}}$ holds for all values of the coupling constants $h^{(1)}$ and $h^{(2)}$. However the qualitative behavior of $\overline{\Delta\mathcal{L}^2_{\text{nr}}}$ is very similar to the true variance $\overline{\Delta\mathcal{L}^2}$.

D. Loschmidt echo distribution function

We now turn to consider the whole probability distribution of the LE. As we have noted earlier, for any L finite being the spectrum discrete, $\mathcal{L}(t)$ is an almost-periodic function. Actually $\mathcal{L}(t)$ belongs to a smaller class, since it is a trigonometric polynomial. In any case, most results we will present are valid for the larger class of almost-periodic functions.

We now give the results for the whole LE probability distribution function. We have observed three kinds of universal behavior emerging in different, well defined regimes. (i) Exponential behavior where the probability distribution is well approximated by $\vartheta(x)e^{-x/\bar{\mathcal{L}}}/\bar{\mathcal{L}}$ (Fig. 5), (ii) Gaussian behavior, (Fig. 6), and (iii) a universal double peaked function (Fig. 7). More precisely we have the following scenario:

(I) δh large. In this case, for L moderately large, the distribution is approximately exponential. The feature is more pronounced when $h^{(1)}, h^{(2)}$ are in different phases, the limiting case being $h^{(1)} \approx \pm h^{(2)}$.

(II) δh small. In this case we have to distinguish two situations:

i. $h^{(i)}$ close to the critical point:
(a) $L \ll |h^{(i)} - 1|^{-1} \propto \xi$, universal double-peaked distribution. Note that $L \ll \xi$ is the so-called quasicritical regime.
(b)$L \gg |\delta h|^{-1}$, exponential distribution.

ii. Off critical:
(a) $L \gg |\delta h|^{-2}$, exponential distribution.
(b) Otherwise Gaussian.

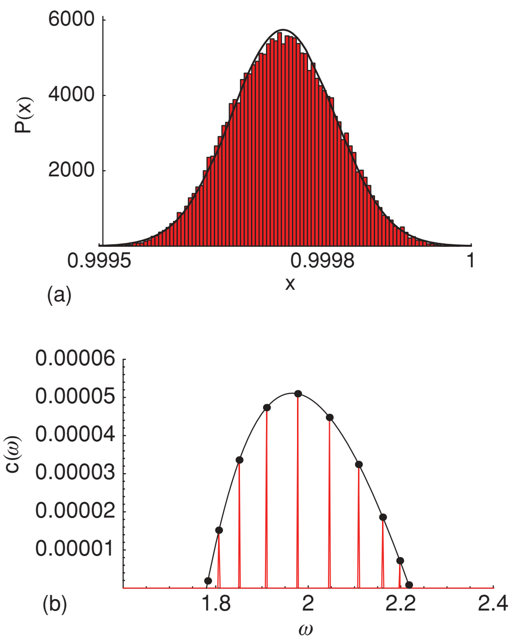


FIG. 6. (Color online) Gaussian behavior for δh small but away from criticality. Parameters are $L = 20$, $h^{(1)} = 0.1$, $h^{(2)} = 0.11$. Note that the distribution function is an extremely peaked Gaussian (upper panel). The thick line is a Gaussian with mean and variance given by Eqs. (16) and (17). In the bottom panel one can notice that many frequencies contribute to the LE. The one-particle contribution $c(\omega)$, Eq. (22), is given by the black dots while the red curve gives the true spectral decomposition $\bar{\mathcal{L}}_{\rm disc}(\omega)$ given by Eq. (20).

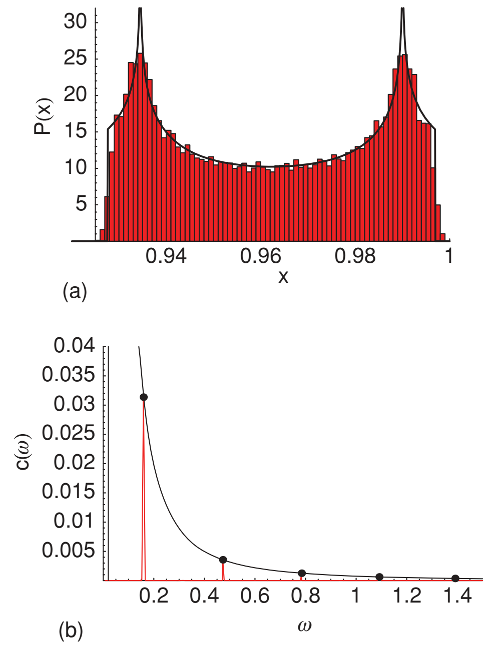


FIG. 7. (Color online) Typical double-peak structure behavior. Parameters are $L = 40$, $h^{(1)} = 0.99$, $h^{(2)} = 1.01$. Upper panel: probability distribution (histogram) together with the result of the approximation given in Eq. (18) (thick line). The mean $\bar{\mathcal{L}}$ is taken from Eq. (9) while the parameters A and B are obtained by computing the two largest spectral weights [see Eq. (21)]. Lower panel: the light curve (red online) is the Fourier series $\bar{\mathcal{L}}_{\rm disc}(\omega)$ (projected spectral density), black curve shows the highest coefficient $c(\omega)$ given by Eq. (22), together with allowed frequencies (black dots).

In other words, say that we fixed $h^{(i)}$ in order to have either a Gaussian or a double-peaked distribution. We can always find an L large enough such that the distribution becomes exponential in both cases. However, for the Gaussian case we must reach considerably larger sizes: $L \gg |\delta h|^{-2}$ (compare Figs. 8 and 9).

We have observed an exponential distribution in the region of parameters where the average LE is much smaller than one: $\bar{\mathcal{L}} \ll 1$. Due to the bound $\overline{\mathcal{L}^2} < 2\bar{\mathcal{L}}^2$, one has $\Delta\mathcal{L} < \bar{\mathcal{L}}^2$ so that when $\bar{\mathcal{L}} \ll 1$ even the variance is small. Since the LE is supported in [0, 1] and in particular $\mathcal{L}(t)$ must be positive we expect in the region $\bar{\mathcal{L}} \ll 1$ a distribution with positive support, with a large peak very close to zero, and rapidly decaying tail. We have verified that an exponential distribution of the form $\vartheta(x)e^{-x/\bar{\mathcal{L}}}/\bar{\mathcal{L}}$ gives a pretty good approximation in the region $\bar{\mathcal{L}} \ll 1$. Note in any case, that the exponential form is always an approximation. In particular, for $x \to 0$ the true distribution $P(x)$ tends to zero for any value of the parameters. This happens since we have always $\mathcal{L}(t) > 0$ strictly. And so generally $0 < \mathcal{L}_{\min} \leqslant \mathcal{L} \leqslant 1$. This feature can be accounted for by adding a (small) ϵ term to the exponential: $P(x) \propto e^{-x/\bar{\mathcal{L}} - \varepsilon/x}$.

Let us now investigate the conditions under which the first moment is small and so we expect an approximately exponential behavior. Looking at Eq. (13) we see that $\bar{\mathcal{L}} \ll 1$ holds when $Lg(h^{(1)}, h^{(2)}) \gg 1$ where g is given by Eq. (14). Clearly the function g is zero (its minimum) when $h^{(1)} = h^{(2)}$, and is quadratic in the difference on the diagonal. Quite interestingly the function g is appreciably different from zero only when $h^{(1)}$ and $h^{(2)}$ correspond to different phases (i.e., either $|h^{(1)}| < 1$ and $|h^{(2)}| > 1$ or vice versa) so that in these cases we observe exponential behavior even for moderately small lattices. When $h^{(1)}$ is close to $h^{(2)}$, but away from critical points, the function g is quadratic in the difference δh so that $\bar{\mathcal{L}} \approx \exp(-\text{const} \times L\delta h^2)$. Hence to have $\bar{\mathcal{L}} \ll 1$ and so to observe approximate exponential behavior we obtain the relation $L \gg \delta h^{-2}$. At criticality and for δh small instead,

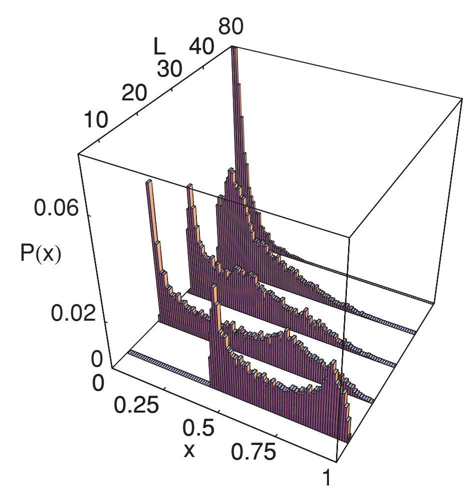


FIG. 8. (Color online) Double-peak distribution approaching an exponential one when increasing system size L at fixed $h^{(i)}$. Parameters are $h^{(1)} = 0.9$, $h^{(2)} = 1.2$ and chain length are $L = 10, 20, 30, 40$.

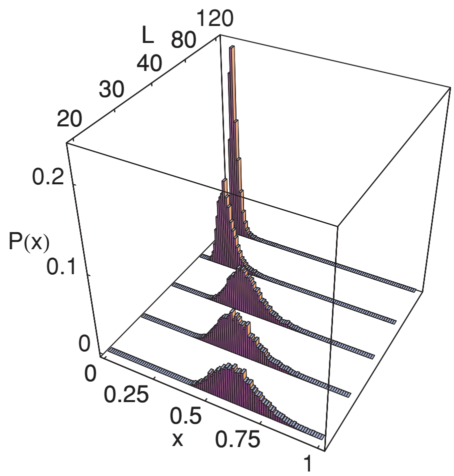


FIG. 9. (Color online) Gaussian distribution approaching an exponential one when increasing system size L at fixed $h^{(i)}$. Parameters are $h^{(1)} = 0.2, h^{(2)} = 0.6$ and chain length are $L = 20, 30, 40, 80, 120$.

we can use $\bar{\mathcal{L}} \approx F^4$ where F is the fidelity which scales as $F \sim 1 - \text{const} \times \delta h^2 L^{2(d+\zeta-\Delta)}$ (see Sec. II B and Ref. [26]). In the quantum Ising model we are considering we have $d = \zeta = \nu = \Delta = 1$ and so the average behaves as $\bar{\mathcal{L}} \approx \exp(-\text{const} \times L^2 \delta h^2)$. This means that for δh small around a critical point $h_c = \pm 1$, the condition to have an exponential distribution becomes $L \gg \delta h^{-1}$.

We now turn to consider the origin of the double-peak shaped distribution function. As we have already noticed, the LE is a (finite) sum of cosines with given frequencies and amplitude. We can imagine a situation where only few frequencies contribute to the LE. In the limiting case, only two nonzero terms. That means that the LE can be approximated by

$$\mathcal{L}(t) = \overline{\mathcal{L}} + A\cos(\omega_A t) + B\cos(\omega_B t). \tag{18}$$

where we can assume A, B positive.

We have devoted some attention to the probability distribution generated by such a function. If ω_A and ω_B are rationally dependent, the function is periodic and the distribution function has square root singularities at all values of $\mathcal{L}$ where $\partial_t \mathcal{L}(t) = 0$. However in our case the frequencies $\omega_{A/B}$ are always rationally independent. In this case the vector $\boldsymbol{x}(t) = (\omega_A t, \omega_B t)$ wraps around the torus in a uniform way. We can then invoke ergodicity and transform the time average into a "phase space" average [in this case the phase space is $\boldsymbol{x} = (x_1, x_2)$]. Hence the probability distribution function is given by

$$P(\mathcal{L}(t) = \ell) = \frac{1}{(2\pi)^2}\int_0^{2\pi} dx_1 \int_0^{2\pi} dx_2 \delta\left[\mathcal{L}(x_1, x_2) - \ell\right].$$

By using Eq. (18) this probability density can be written as

$$P(\ell + \bar{\mathcal{L}}) = \frac{1}{\pi^2 A}\int_{\max\{-1,(\ell-B)/A\}}^{\min\{1,(\ell+B)/A\}} \times \frac{dz}{\sqrt{\left(\frac{\ell+B}{A} - z\right)\left(z - \frac{\ell-B}{A}\right)(1-z^2)}}. \tag{19}$$

The integral above can be expressed in terms of elliptic functions, but we would not need its explicit expression. Typically the function (19) is double-peak shaped function (see Fig. 7), with support in $[\bar{\mathcal{L}} - |A+B|, \bar{\mathcal{L}} + |A+B|]$ and two peaks at $\ell = \bar{\mathcal{L}} \pm |A-B|$. The divergence at the peaks position is of logarithmic type: close to the peaks, setting $\ell = \bar{\mathcal{L}} \pm |A-B| + \epsilon$, one has

$$P(\ell) = -\frac{\ln(\epsilon)}{2\pi^2\sqrt{|AB|}} + O(1).$$

Note that we never observe the limiting case where $B = 0$ and $\mathcal{L}(t)$ becomes a periodic function. This means that even a very small spectral weight on B cannot be discarded. On the other hand the distribution function (19) seems to be quite stable against the presence of other oscillating terms with small spectral weight. This stability can be seen in Fig. 7 where one clearly has at least three frequencies with reasonable spectral weight, but the probability density is still well approximated by a double-peaked structure.

Spectral analysis. To understand the behavior of $P(\mathcal{L} = x)$ we do a spectral analysis of $\mathcal{L}(t)$ to see which frequencies contribute most. In fact, for almost-periodic functions there is a similar Fourier decomposition as for periodic functions. The Fourier expansion is given in this case by $\hat{\mathcal{L}}_{\text{disc}}(\omega) = \overline{\mathcal{L}(t)e^{i\omega t}}$. Taking into account Eq. (8) $\hat{\mathcal{L}}_{\text{disc}}(\omega)$ can be written as

$$\hat{\mathcal{L}}_{\text{disc}}(\omega) = \sum_{n,m} p_n p_m \delta_{\omega, E_n - E_m}. \tag{20}$$

We would like to know which frequencies have the largest weight. This is achieved by expanding the product in Eq. (9) [36]

$$\begin{aligned}\mathcal{L}(t) &= 1 + \sum_{k>0} X_k(t) + \sum_{k_1>k_2>0} X_{k_1}(t)X_{k_2}(t) + \cdots \\ &= \bar{\mathcal{L}}\sum_{k>0}\tilde{X}_k(t) + \sum_{k_1>k_2>0}\tilde{X}_{k_1}(t)\tilde{X}_{k_2}(t) + \cdots \\ \text{with } \tilde{X}_k(t) &= \sum_{\beta=\pm1} c_\beta^k e^{i\beta\Lambda_k t} = \frac{\alpha_k}{2}\cos(\Lambda_k\, t).\end{aligned}$$

Now, since each $\tilde{X}_k(t)$ is smaller than 1/2 in modulus, it is reasonable to approximate the LE with the first two terms of this expansion and we obtain

$$\mathcal{L}(t) \simeq \bar{\mathcal{L}} + \sum_{k>0}\frac{\alpha_k}{2}\cos(\Lambda_k\, t). \tag{21}$$

In this approximation we only wrote the zero-frequency contribution, which corresponds to the mean, and the contribution coming from the one-particle spectrum. The next term has also contributions coming from the two particle spectrum. To be more precise, call $E^{(n)}$ the energy of the n-particle spectrum then $E_a^{(1)} - E_b^{(0)} \propto \Lambda_k$, (first order contribution), while $E_a^{(1)} - E_b^{(1)} \propto \Lambda_{k_1} - \Lambda_{k_2}$, and $E_a^{(2)} - E_b^{(0)} \propto \Lambda_{k_1} + \Lambda_{k_2}$ (second order contribution with less spectral weight).

Note that we expect Eq. (21) to be approximately valid (with a different form for the amplitudes and the frequencies) also for nonintegrable models in which a one-particle approximation works well.

Now, if there is a regime where the amplitudes $\alpha_k/2$ are highly peaked around few quasimomenta, in the limiting case only two, then the LE can be approximated as in Eq. (18) and we expect a double peaked distribution function. So we are led to study the (one-particle) amplitude function

$$c(\omega) \equiv \left.\frac{\alpha_k}{2}\right|_{\omega=\Lambda_k}, \quad \omega \in [E_m, E_M]. \tag{22}$$

Generally $c(\omega)$ is a bell-shaped function, starting linearly from the band minimum E_m, reaching a maximum value and then decreasing to zero at the band maximum E_M. It is not difficult to show that [23], when δh is small and for roughly $|1-h^{(2)}| \lesssim 10^{-1}$, $c(\omega)$ starts developing a peak, the width of which being proportional to $|1-h^{(2)}|$. In the limiting case $h^{(2)}=1$, $c(\omega)$ has its maximum at $E_m=0$ and then decreases monotonically to $E_M=2$. So, for δh small and for $|1-h^{(2)}| \lesssim 10^{-1}$, $c(\omega)$ is a peaked function. In order to have few frequencies fall within the peak, and so to have large spectral weight on few frequencies, we must additionally have $L|1-h^{(2)}| \ll 1$. This is easily seen analyzing the dispersion Λ_k for $h^{(2)}$ close to the critical point [24]. All in all, the conditions to have a double-peak distribution, are δh small and $L|1-h^{(2)}| \ll 1$. The feature is more pronounced when the h_i are not precisely critical. In fact even though $c(\omega)$ is most peaked when $h^{(2)}=1$ (and the peak is at $\omega=0$), we have to remember that the allowed values of ω are $\omega_n=\Lambda_{k_n}$ where $k_n=\pi(2n+1)/L$, and the smallest frequency is $\omega_1=\Lambda_{\pi/L}$. If we perturb $h^{(2)}$ from 1 the peak of $c(\omega)$ shifts to the right, approaching ω_1, so it is favorable to have $h^{(2)} \neq 1$. Not surprisingly, the conditions to have a double-peak probability density, coincide with having a large variance (see Fig. 4 bottom panel).

Generally fixing $h^{(i)}$ and increasing the size L, one eventually violates the quasicritical condition $L \ll \xi$. At this stage the double-peak feature tends to disappear and the distribution approaches an exponential one. This can be clearly seen in Fig. 8. From this figure one can have the impression that the "double-peak feature" is a prerequisite of short sizes, since in this case one has few frequencies anyway. As we have tried to explain instead, this feature survives for larger sizes, provided we shrink δh sufficiently (Fig. 7).

Instead, when δh is small but $h^{(i)}$ are far from the critical point, then $c(\omega)$ is not peaked, and many frequencies have a large spectral weight (see Fig. 6). In this case the distribution becomes Gaussian. The emergence of a Gaussian distribution can be qualitatively understood in the following way. First write the LE according to its spectral decomposition $\mathcal{L}(t)=\sum_n A_n e^{it\omega_n}$ where the amplitudes are precisely given by $A_n=\hat{\mathcal{L}}_{\rm disc}(\omega_n)$ and are positive. When the frequencies are rationally independent the variables $x_n=t\omega_n$ wrap uniformly around a large dimensional torus. Then one can consider each $A_n \times e^{itx_n}$ as an independent random variable. The assumption δh is small but $h^{(i)}$ away from criticality corresponds to say that $\mathcal{L}(t)$ can be considered as a sum of *many* independent random variables, giving rise to a Gaussian distribution as a consequence of the central limit theorem. When $L \gg |\delta h|^{-2}$ the conditions of independence breaks down and we recover an approximate exponential behavior.

At this point it is worth to comment on the relation between the different regimes observed (exponential, Gaussian, double peak) and the equilibration dynamics. As we have recalled, to have equilibration in probability, it is sufficient to have a small variance. In the exponential regime (given by $\bar{\mathcal{L}} \ll 1$) the variance is given approximately by $k_2 \approx \bar{\mathcal{L}}^2$ and is exponentially small in the system size. In the Gaussian regime one can have a large mean but the variance will still be small (cf. Fig. 6). Hence in these regimes (exponential and Gaussian) the conditions to have unitary equilibration in the sense specified in Sec. II, are satisfied: for the large majority of times one will observe the LE very close to its mean. On the contrary in the double-peak regime the variance is large. In fact, studying numerically the variance given by Eq. (17) we have verified that the condition to have a large variance is precisely the same the defines the double-peak regime (cf. Fig. 4 bottom panel). The lack of equilibration in the double-peak regime can also be understood very intuitively. In this regime the system oscillates among very few eigenstates of the evolution Hamiltonian and equilibration cannot take place in any sense. Conversely in the exponential and double-peak cases, the system has access to many different states, and equilibration can take place in probabilistic sense.

IV. PROBABILITY DISTRIBUTION FUNCTION FOR THE MAGNETIZATION

In the same spirit we can compute the probability distribution of a local operator. The first candidate that comes to mind is the transverse magnetization. We computed the following time-dependent observable $m(t)=\langle\psi^{(1)}|e^{itH^{(2)}}\sigma_i^z e^{-itH^{(2)}}|\psi^{(1)}\rangle$. Using again Eq. (10) one obtains [37]

$$m(t) = \frac{1}{L}\sum_k \cos\left(\vartheta_k^{(2)}\right)\cos(\delta\vartheta_k) + \sin\left(\vartheta_k^{(2)}\right)\sin(\delta\vartheta_k)\cos\left(t\Lambda_k^{(2)}\right), \tag{23}$$

where the quasimomenta range now in the whole Brillouin zone: $k=\pi(2n+1)/L$, $n=0,1,\ldots,L-1$. Correctly, when $h^{(1)}=h^{(2)}$ we recover the zero temperature equilibrium result $\langle\sigma_i^z\rangle = L^{-1}\sum_k \cos(\vartheta_k)$.

From Eq. (23) we see that $m(t)$ can be written—exactly—as a constant term plus an oscillating part with frequencies given by the single particle spectrum Λ_k. The discussion on characteristic times becomes simplified as all time scales are uniquely determined by $\Lambda_k^{(2)}$. For example the time T_1 necessary to observe the correct mean: $\overline{m^{T_1}}=\bar{m}$ must simply satisfy $T_1 \gg \text{gap}^{-1}$ which means $T_1 \gg L$ in the quasicritical regime $|h^{(2)}-1|^{-1} \gg L$, while it suffices to have $T_1 \gg O(1)$ away from criticality. Given Eq. (23) it is not difficult to compute the mean and the variance, which are given by

$$\bar{m} = \frac{1}{L}\sum_k \cos\left(\vartheta_k^{(2)}\right)\cos(\delta\vartheta_k), \tag{24}$$

$$\overline{\Delta m^2} = \frac{1}{L^2}\sum_k \sin^2\left(\vartheta_k^{(2)}\right)\sin^2(\delta\vartheta_k). \tag{25}$$

Some comments are in order here. First fixing $h^{(1)}, h^{(2)}$ the variance (Fig. 10) goes to zero as L^{-1} and not exponentially

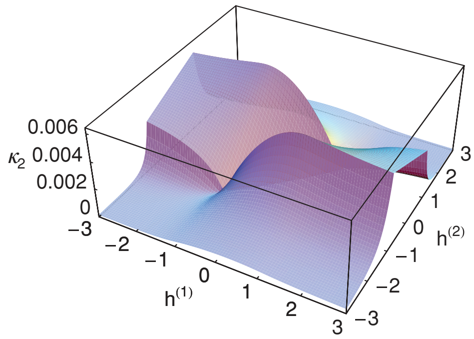


FIG. 10. (Color online) Variance of the magnetization as given by Eq. (10) for $L = 80$. A signature of criticality are the cusps at $h^{(2)} = \pm 1$.

fast as was the case for the LE. Second, the spectral weight associated to the frequency $\Lambda_k^{(2)}$ is $\sin(\vartheta_k^{(2)})\sin(\delta\vartheta_k)$. We observed that there are always many frequencies with large spectral weight. In other words the spectral weight function is never peaked.

These comments suggest us to expect a Gaussian behavior for the probability distribution function of the magnetization irrespective of the parameters approaching critical values. This has indeed been observed (Fig. 11).

V. CONCLUSIONS

The unitary character of the dynamics of a closed quantum system implies that whatever relevant notion of equilibration one might have has to be a subtle one. In this paper we investigated the unitary equilibration of a quantum system after a sudden change of its Hamiltonian parameter. To this aim we used a prototypical time-dependent quantity: the Loschmidt echo. We established how the global features of $\mathcal{L}$ depend on the physical properties of the initial state preparation and on those of the quench Hamiltonian. The central object of our analysis is given by the long-time probability distribution for $\mathcal{L}$: $P(x) = \overline{\delta(\mathcal{L}(t) - x)} := \lim_{T\to\infty} T^{-1}\int_0^T \delta(\mathcal{L}(t) - x)dt$. Broadly speaking concentration phenomena for P correspond to quantum equilibration.

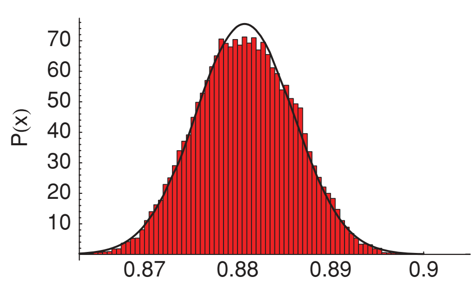


FIG. 11. (Color online) The probability distribution for the magnetization has only Gaussian behavior. Here parameters are $L = 40$, $h^{(1)} = 0.9$, $h^{(2)} = 1.01$. The continuous line is a Gaussian with mean and variance given by Eqs. (24) and (25).

Here below for the reader's sake we summarize the main findings of the paper.

(i) Resorting to a cumulant expansion we characterized the short-time behavior of $\mathcal{L}(t)$. Different regimes can be identified depending on the most relevant scaling dimension of the quench Hamiltonian. When the central limit theorem (CLT) applies one has a Gaussian decay over a time scale $O(1)$ for gapped systems. For the critical case the time-scale becomes $O(L^\zeta)$ in a small region $|\delta h| \ll L^{-(d+\zeta-\Delta)}$. At critical points the CLT can be violated and $\mathcal{L}(t)$ takes a universal non-Gaussian form for sufficiently relevant perturbations.

(ii) We discussed the general structure of the higher momenta of P i.e., $\mu_k := \overline{\mathcal{L}^k(t)} = \int x^k P(x)dx$ using the so-called nonresonant hypothesis. We showed that all the μ_k are bounded by those corresponding to a Poissonian distribution, i.e., $P(x) = \vartheta(x)\exp(-x/\bar{\mathcal{L}})/\bar{\mathcal{L}}$.

(iii) Using exact results for the quantum Ising chain we rigorously analyzed the interplay between the chain length L and the averaging time T. In particular we showed how the limit $\lim_{T\to\infty}$ and the thermodynamical one, i.e., $\lim_{L\to\infty}$ do not commute. While in finite systems the $\mathcal{L}(t)$ is an almost-periodic function, in the thermodynamical limit $\mathcal{L}_\infty := \lim_{t\to\infty}\mathcal{L}(t)$ exists and $P(x) \to \delta(x - \mathcal{L}_\infty)$. We explicitly computed $\mathcal{L}_\infty$ and the way it is asymptotically approached for large t. We gave a general closed form for the exact μ_k's and compared with that obtained with the nonresonant hypothesis.

(iv) For the quantum Ising chain we numerically investigated $P(x)$. We identified three universal regimes (a) an exponential one (P is Poissonian) when L is the largest scale of the system; (b) a Gaussian one for intermediate L and initial state and quench parameters close and off-critical; (c) a double-peak shape for P when the parameters are close to each other and close to criticality. This result holds in the *quasicritical* region $L|h^{(i)} - 1| \ll 1$, $(i = 1, 2)$.

(v) Finally, for the sake of the comparison of the Loschmidt echo with a prototypical observable, we computed the time-dependent magnetization after the quench and studied its long-time statistics. In this case only a Gaussian regime appears to be reachable.

We have shown that the Loschmidt echo encodes sophisticated information about the quantum equilibration dynamics. For finite system vastly different time scales arise: short-time relaxation is intertwined with a complex a pattern of collapses and revivals and eventually Poincare recurrences. Unveiling how these phenomena depend on spectral properties of the underlying Hamiltonians, is one of the key challenges in the way to understand emergent thermal behavior in closed quantum systems.

ACKNOWLEDGMENTS

The authors gratefully acknowledge discussions with H. Saleur and A. Winter. Supported by NSF Grants No. PHY-803304 and DMR-0804914 and European project COQUIT FP7-ICT-2007-C, n. 233747. LCV wishes to thank S. Fortunato, F. Radicchi, and A. Lancichinetti for providing a result on the permutation group and D. Burgarth for useful discussions.

APPENDIX

Asymptotic of $s(t)$

Here we want to compute the asymptotic of $s(t)$ for $t \to \infty$, that is the integral,

$$s(t) = -\frac{1}{2\pi}\int_0^\pi \ln\left[1-\sin^2\left(\vartheta_k^{(1)}-\vartheta_k^{(2)}\right)\sin^2\left(\Lambda_k^{(2)}t/2\right)\right]dk$$

We can go to energy integration setting $\Lambda_k^{(2)}=\omega$

$$s(t) = -\frac{1}{2\pi}\int_{E_m}^{E_M}\ln[1-\alpha(\omega)\sin^2(\omega t/2)]\rho(\omega)d\omega.$$

where $E_m = 2\min\{|1+h^{(2)}|,|1-h^{(2)}|\}$ and $E_M = 2\max\{|1+h^{(2)}|,|1-h^{(2)}|\}$. To be explicit,

$$\rho(\omega) = \frac{2\omega}{\sqrt{\left(\omega^2-E_m^2\right)\left(E_M^2-\omega^2\right)}},$$

$$\alpha(\omega) = \frac{\left(\omega^2-E_m^2\right)\left(E_M^2-\omega^2\right)(h^{(2)}-h^{(1)})^2}{4(h^{(2)})^2[4(h^{(2)}-h^{(1)})(1-h^{(1)}h^{(2)})+h^{(1)}\omega^2]\omega^2}.$$

Note that $\alpha(\omega)$ is zero at the band's edge, positive otherwise (and smaller than 1 in modulus). Instead $\rho(\omega)$ has square root (van Hove) singularities at the band edges as a result of the quadratic dispersion at those points (when $h^{(2)}\neq 1$). When $h^{(2)}=1$ the dispersion is linear at the bottom of the band but still quadratic at the upper band edge, hence in this case only the square root singularity at the upper band edge survives.

Then expand the logarithm into an infinite series. Using the Riemann-Lebesgue lemma we can show that,

$$\lim_{t\to\infty}\int f(w)\,[\sin(wt)]^{2k}\,dw = 2^{-2k}\binom{2k}{k}\int f(w)\,d\omega,$$

provided that f is summable. The resulting series can be summed

$$-\sum_{k=1}^{\infty}\frac{x^k}{k}2^{-2k}\binom{2k}{k} = 2\ln\left(\frac{1+\sqrt{1-x}}{2}\right),\quad \text{if}\quad |x|<1.$$

So finally

$$\lim_{t\to\infty}s(t) = -\frac{1}{\pi}\int_{E_m}^{E_M}\ln\left[\frac{1+\sqrt{1-\alpha(\omega)}}{2}\right]\rho(\omega)d\omega$$
$$= -\frac{1}{\pi}\int_0^\pi\ln\left[\frac{1+\sqrt{1-\alpha(k)}}{2}\right]dk.$$

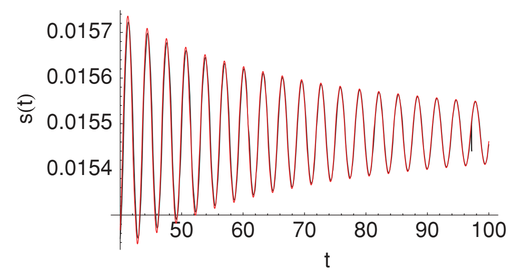


FIG. 12. (Color online) Typical behavior of the function $s(t)$ (black line) in the thermodynamic limit. The red line is the approximation given by Eq. (A1). Parameters are $h^{(1)}=1.3$, $h^{(2)}=2$.

To compute the first correction to the limit note that $\alpha(\omega)^k$ smoothen the singularity at the band edge, so that for the leading correction we need only $k=1$ in the expansion of the logarithm. The evaluation of the oscillating integral is done with a saddle point technique. The result is

$$s(t) \simeq s(\infty) - \frac{1}{4\pi}\int_{E_m}^{E_M}\alpha(\omega)\rho(\omega)\cos(\omega t)d\omega$$
$$\simeq s(\infty) - \frac{A_m}{|t|^{3/2}}\cos\left(tE_m+\frac{3}{4}\pi\right)+(m\leftrightarrow M),\quad \text{(A1)}$$

with constants given by (we assumed here $h^{(2)}>0$ so that the band minimum is $E_m = 2|1-h^{(2)}|$)

$$A_m = \frac{1}{16\sqrt{\pi}}\frac{(h^{(1)}-h^{(2)})^2}{(1-h^{(1)})^2(h^{(2)})^{3/2}\sqrt{|1-h^{(2)}|}},$$

$$A_M = -\frac{1}{16\sqrt{\pi}}\frac{(h^{(1)}-h^{(2)})^2}{(1+h^{(1)})^2(h^{(2)})^{3/2}\sqrt{|1+h^{(2)}|}}.$$

The result (A1) should be the same as the square modulus of Eq. (12) in Ref. [4]. However in Ref. [4] there appears only one frequency, corresponding to the lowest band edge. The discrepancy probably arises from a continuum approximation which discards the effect of the van Hove singularity present at the upper band edge. As we have seen both terms give similar contributions. In particular, even at criticality, the van Hove singularity at the upper band edge survives. In Fig. 12 one can appreciate the validity of the approximation (A1).

[1] P. Calabrese and J. Cardy, Phys. Rev. Lett. **96**, 136801 (2006).
[2] M. A. Cazalilla, Phys. Rev. Lett. **97**, 156403 (2006).
[3] S. R. Manmana, S. Wessel, R. M. Noack, and A. Muramatsu, Phys. Rev. Lett. **98**, 210405 (2007).
[4] A. Silva, Phys. Rev. Lett. **101**, 120603 (2008).
[5] T. Kinoshita, T. Wegner, and D. S. Weiss, Nature (London) **440**, 900 (2006).
[6] S. Hoffenberth, I. Lesanovsky, B. Fisher, T. Schumm, and J. Schmiedmayer, Nature (London) **449**, 324 (2007).
[7] H. Tasaki, Phys. Rev. Lett. **80**, 1373 (1998).
[8] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Phys. Rev. Lett. **96**, 050403 (2006).
[9] P. Reimann, Phys. Rev. Lett. **99**, 160404 (2007).
[10] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).
[11] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).
[12] N. Linden, S. Popescu, A. J. Short, and A. Winter, Phys. Rev. E **79**, 061103 (2009).
[13] K. D. Schotte and U. Schotte, Phys. Rev. **182**, 479 (1969).
[14] T. Prosen, Phys. Rev. Lett. **80**, 1808 (1998).

[15] R. A. Jalabert and H. M. Pastawski, Phys. Rev. Lett. **86**, 2490 (2001).

[16] H. T. Quan, Z. Song, X. F. Liu, P. Zanardi, and C. P. Sun, Phys. Rev. Lett. **96**, 140604 (2006).

[17] D. Rossini, T. Calarco, V. Giovannetti, S. Montangero, and R. Fazio, J. Phys. A: Math. Theor. **40**, 8033 (2007).

[18] D. Rossini, T. Calarco, V. Giovannetti, S. Montangero, and R. Fazio, Phys. Rev. A **75**, 032333 (2007).

[19] P. Zanardi, H. T. Quan, X. Wang, and C. P. Sun, Phys. Rev. A **75**, 032109 (2007).

[20] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).

[21] J. Kurchan, e-print arXiv:cond-mat/0007360.

[22] M. Cramer, A. Flesch, I. P. McCulloch, U. Schollwöck, and J. Eisert, Phys. Rev. Lett. **101**, 063001 (2008).

[23] Expanding up to second order in δh around $h^{(2)}$, $c(\omega)$ can be well approximated by $c(\omega) = \frac{1}{8h_2^2\omega^4}(\omega^2 - E_m^2)(E_M^2 - \omega^2)\delta h^2 + O(\delta h^4)$. This function has a flex at $\omega_{\text{flex}} \simeq 3.61 - h^{(2)}$. Therefore the width of the peak is small compared to the bandwith roughly when $1 - h^{(2)} \lesssim 10^{-1}$.

[24] When $h^{(2)}$ is close to criticality, say $h^{(2)} = 1 + \Delta h$, the band is approximately $\Lambda_k^{(2)} \simeq (1 + \Delta h/2)\sin(k/2)$. The number of frequencies which fall in the peak is given by the n satisfying $\Lambda_{k_n}^{(2)} = \omega_{\text{flex}}$. At first order we obtain $n = O(L\Delta h)$, and so to have few frequencies in the peak we must have $Lh^{(2)} - 1 \ll 1$.

[25] D. K. L. Oi, Phys. Rev. Lett. **91**, 067902 (2003).

[26] L. Campos Venuti and P. Zanardi, Phys. Rev. Lett. **99**, 095701 (2007).

[27] V. Privman, P. Hohenberg, and A. Aharony, *Phase Transition and Critical Phenomena* (Academic Press, London, 1989), Vol. 14.

[28] V. Aji and N. Goldenfeld, Phys. Rev. Lett. **86**, 1007 (2001).

[29] A. Lamacraft and P. Fendley, Phys. Rev. Lett. **100**, 165706 (2008).

[30] More precisely, assume the lattice is bipartite and the quasi-momenta satisfy some quantization in order to be roughly equally spaced: $k = 2\pi n/L$. Then $\sum_{\boldsymbol{k}} (-1)^{\boldsymbol{k}} \Lambda_{\boldsymbol{k}}$ is the difference between two multidimensional Riemann sums. If $\Lambda_{\boldsymbol{k}}$ is analytic, both these sums converge exponentially fast to their integral, and so their difference is exponentially small in L.

[31] E. Barouch, B. M. McCoy, and M. Dresden, Phys. Rev. A **2**, 1075 (1970).

[32] Here below we sketch why this is so. Let us suppose that $\rho_\infty = \lim_{t\to\infty}\mathcal{U}_t(\rho_0)$. Obviously ρ_∞ is a fixed point for $\mathcal{U}_t$, i.e., $\mathcal{U}_t(\rho_\infty) = \mathcal{U}_t(\lim_{u\to\infty}\mathcal{U}_u(\rho_0)) = \lim_{u\to\infty}\mathcal{U}_{t+u}(\rho_0) = \lim_{u\to\infty}\mathcal{U}_u(\rho_0) = \rho_\infty$. Using unitary invariance of the trace norm it follows that $\|\mathcal{U}_t(\rho_0) - \rho_\infty\|_1 = \|\mathcal{U}_t(\rho_0 - \rho_\infty)\|_1 = \|\rho_0 - \rho_\infty\|_1$ and therefore $\lim_{t\to\infty}\|\mathcal{U}_t(\rho_0) - \rho_\infty\|_1 = 0$ implies $\rho_0 = \rho_\infty$.

[33] Notice that in the infinite-dimensional case discussed above Eq. (1) implies $P(\alpha) = \delta(\alpha - A_\infty)$. Indeed from Eq. (1) it follows for any continuous f that $\overline{f[A(t)]} = f(A_\infty)$. Whence $P(\alpha) = \overline{\delta[\alpha - A(t)]} = \overline{1/2\pi \int e^{i\lambda\alpha}e^{-i\lambda A(t)}} = 1/2\pi \int e^{i\lambda\alpha}\overline{e^{-i\lambda A(t)}} = 1/2\pi \int e^{i\lambda\alpha}e^{-i\lambda A_\infty} = \delta(\alpha - A_\infty)$.

[34] $\mathcal{L}^n(t) = \langle\rho_\psi^{\otimes n}, \mathcal{P}^{(n)}(\rho_\psi^{\otimes n})\rangle + \langle\rho_\psi^{\otimes n}, e^{-it\tilde{\mathcal{H}}^{(n)}}(\rho_\psi^{\otimes n})\rangle$ where $\tilde{\mathcal{H}}^{(n)} = (-\mathcal{P}^{(n)})\mathcal{H}^{(n)}(-\mathcal{P}^{(n)})$. The time average of the second term is $\langle\rho_\psi^{\otimes n}, F_T^n(\rho_\psi^{\otimes n})\rangle/T$ where $F_T^n = (-i\tilde{\mathcal{H}}^{(n)})^{-1}[e^{-iT\tilde{\mathcal{H}}^{(n)}} -]$. Since F_T^n is a bounded operator its expectation value divided by T goes to zero when $T \to \infty$.

[35] In (ii) one has to measure $H^{(n)} := \sum_{i=1}^n {}^{\otimes(i-1)} \otimes H \otimes^{\otimes(n-i)}$ in $\rho_\psi^{\otimes n}$.

[36] A more precise approximation valid also in region where δh is large is given by $\mathcal{L}(t) = \prod_{k>0}(1 + c_0^k + \tilde{X}_k(t)) \simeq \bar{\mathcal{L}}(1 + \sum_{k>0}(1 + c_0^k)^{-1}\tilde{X}_k(t))$. The two expressions coincide when δh is small.

[37] This is precisely Eq. (5.6) of (31) in the zero temperature limit.