Stability of quantum statistical ensembles with respect to local measurements

Walter Hahn[*] and Boris V. Fine[*]

Skolkovo Institute of Science and Technology, Skolkovo Innovation Centre, Nobel Street 3, Moscow 143026, Russia
and Institute for Theoretical Physics, Philosophenweg 19, 69120 Heidelberg, Germany

We introduce a stability criterion for quantum statistical ensembles describing macroscopic systems. An ensemble is called "stable" when a small number of local measurements cannot significantly modify the probability distribution of the total energy of the system. We apply this criterion to lattices of spins-1/2, thereby showing that the canonical ensemble is nearly stable, whereas statistical ensembles with much broader energy distributions are not stable. In the context of the foundations of quantum statistical physics, this result justifies the use of statistical ensembles with narrow energy distributions such as canonical or microcanonical ensembles.

I. INTRODUCTION

An isolated classical system always has a fixed value of the total energy. In contrast, an isolated quantum system can be in a superposition of states with different total energies [1]. This entails the following difficulty for the foundations of statistical physics: A typical isolated macroscopic quantum system is generally expected to thermalize under the action of its internal dynamics for the overwhelming majority of initial nonequilibrium states appearing in nature or created in a laboratory [2–12]. Thermalization implies that the density matrix of any small subsystem within the large system approaches a form consistent with the canonical Gibbs density matrix (canonical ensemble) for the large system. A necessary condition for an isolated many-particle quantum system to thermalize is a sufficiently narrow initial probability distribution of its total energy, as is, for example, the case for the canonical and the microcanonical statistical ensembles (see below). At the same time, for an isolated quantum system, the probability of occupying an energy eigenstate remains unchanged with time. Thus, a large isolated quantum system with a broad initial distribution of the total energy thermalizes not to a conventional equilibrium state with a well-defined temperature but rather to a mixture or superposition of states with different temperatures. The problem now is that the initial states characterized by the narrow distribution of total energy are *not* necessarily the most probable ones. For example, non-Gibbs equilibrium for small subsystems emerges when initial states are selected in the Hilbert space of a typical many-particle system among quantum superpositions with a given energy expectation value and without any constraint on the width of the energy window for participating eigenstates [13–16]. The latter condition defines the "quantum microcanonical" (QMC) ensemble [13,17–21].

Given the above considerations, why do the initial nonequilibrium quantum states of macroscopic systems *not* normally exhibit the broad participation of energy eigenstates and hence non-Gibbs statistics for small subsystems? In this paper, we address the above question by introducing the criterion of stability of quantum statistical ensembles with respect to local measurements and then apply this criterion to lattices of spins

1/2. Similar ideas in other contexts have been considered in Refs. [22,23].

A quantum statistical ensemble for a macroscopic system is defined by the probability $p(E)$ of occupying an eigenstate of total energy E. Given the density of energy states $\nu(E)$, the probability distribution of the total energy is

$$g(E) = p(E)\nu(E). \qquad (1)$$

We call $g(E)$ broad when $w_g/(E_{\text{av}} - E_{\text{min}}) \sim 1$, where w_g^2 is the variance of $g(E)$, E_{av} is the average energy, and E_{min} is the ground-state energy of the system. As shown in Appendix A, a canonical ensemble with a positive temperature T is narrow from the above perspective, because, in this case, $w_g/(E_{\text{av}} - E_{\text{min}}) \sim 1/\sqrt{N_s} \ll 1$, where $N_s \sim 10^{23}$ is the number of particles or microscopic subsystems in the system, cf. Ref. [16].

The article is organized as follows. Section II provides the definition of the stability criterion for quantum statistical ensembles. In Sec. III, the effects of local measurements are discussed qualitatively. Section IV includes an analytical investigation of the stability of quantum ensembles for lattices of spins-1/2, including noninteracting spins in a magnetic field and systems of interacting spins. Section V presents the results of numerical investigations for systems of interacting spins. Final remarks and conclusions are given in Sec. VI. Longer derivations are included in the Appendices.

II. DEFINITION OF THE STABILITY CRITERION

Let us now observe that accidental measurements of microscopic particles in a macroscopic system cannot be excluded under any foreseeable natural or experimental conditions. We, therefore, introduce the following stability criterion: *A physically realizable quantum statistical ensemble describing a stationary state of a macroscopic system must be stable with respect to a small number of any arbitrarily chosen local measurements within the system. The measurement is called "local" if the measured quantity is localized in the three-dimensional physical space [24]. The number of measurements n is called small if $n \ll \sqrt{N_s}$. The ensemble is called stable if*

$$\Delta G(n) \equiv \int_{-\infty}^{+\infty} |g_n(E) - g_0(E)|dE \ll 1, \qquad (2)$$

---

[*]w.hahn@thphys.uni-heidelberg.de, b.fine@skoltech.ru

where $g_0(E)$ and $g_n(E)$ are the probability distributions before and after the measurements, respectively.

The measurements in question are implied to occur naturally, for example, when a passing photon becomes entangled with the system and accidentally measured later. Without the measurement, the above process would describe a decoherence event [25–30].

In the following, we simplify the analysis by assuming random instantaneous projective measurements of individual particles in the system [31–35]. Furthermore, we mainly focus on "strong instability," which we define as the case when fewer than 10 measurements lead to $\Delta G \gtrsim 0.1$ independent of N_s.

III. EFFECTS OF LOCAL MEASUREMENTS

A. Narrowing vs. broadening

Qualitatively, measurements can lead to both narrowing and broadening of $g(E)$. The broadening effect of a single measurement is due to the off-diagonal elements of the projection operator describing the measurement in the basis of the total-energy eigenstates. The narrowing effect originates from correlations between the total energy of the system and the measurement outcomes. Indeed, broad probability distributions $g(E)$ can be considered as a mixture of microcanonical (or canonical) ensembles corresponding to different temperatures $T(E)$, which, in turn, imply different probability distributions of local variables. When a given measurement outcome is more likely for one represented temperature than for another, the postmeasurement distribution will be narrower than the initial one [32]. In terms of the energy scales, the increase of the variance w_g^2 due to the off-diagonal elements of a local projection operator should normally be of the order of ϵ_1^2, where ϵ_1 is an appropriately chosen single-particle energy. The reason is that the quantum projections corresponding to the local measurements are one-particle or a few particle operators. At the same time, the decrease of w_g^2 for broad $g(E)$ can easily be of the order $\epsilon_1^2 N_s^2$, i.e., much larger. Since we focus primarily on broad $g(E)$, we neglect the broadening effect of the measurements unless explicitly specified otherwise. We further limit our derivations to $g(E)$ satisfying inequality

$$\left| \frac{dg(E)}{dE} \right| \lesssim \frac{g(E)}{w_{\text{can}}}, \qquad (3)$$

where w_{can} is the width of the energy distribution corresponding to the canonical ensemble with the same average energy as that of $g(E)$.

B. Heating effect of measurements

For an ensemble with narrow energy distribution $g(E)$, such as the canonical ensemble, the above-mentioned broadening effect also leads to heating, defined as the drift of the average energy E_{av} in the direction of larger entropy $S(E_{\text{av}})$ (larger value of $\ln[\nu(E_{\text{av}})]$). This means that E_{av} increases for positive temperatures and decreases for negative ones. Such a behavior is consistent with the second law of thermodynamics, because local measurements of individual particles can be viewed as small nonadiabatic perturbations of the system. For positive temperatures T, the increase of E_{av} due to one measurement is of the order of the one-particle energy ϵ_1.

A broad ensemble can be considered as a mixture of canonical ensembles with different temperatures. The heating for each of the contributing canonical ensembles occurs as explained above. Therefore, the overall heating effect for a broad ensemble corresponds to the combination of the heating effects for the individual canonical ensembles. In a system with a finite Hilbert space, the asymptotic shape of $g(E)$ corresponds to the canonical ensemble at infinite temperature, which is, in turn, proportional to the density of states $\nu(E)$ of the system.

C. Coarse graining of the energy axis

Our definition of $g(E)$ implies averaging over energy bins whose width Δ_e satisfies the following inequalities:

$$\epsilon_1 \ll \Delta_e \ll T(E_{\text{av}})\sqrt{C_V(E_{\text{av}})}. \qquad (4)$$

In the case of negative temperatures, $|T(E_{\text{av}})|$ should be used instead.

The left inequality in Eq. (4) together with the restriction $n \ll \sqrt{N_s}$ and the inequality (3) allows us to neglect heating described in the preceding part. The right inequality in Eq. (4) assures that the energy eigenstates within each bin approximately correspond to the same density matrices of small subsystems within the entire system considered.

The probability distribution $g_n(E)$ is defined in terms of the above energy bins as follows:

$$g_n(E) = \frac{1}{\Delta_e} \sum_k^{\text{bin}(E)} (\rho_n)_{kk}, \qquad (5)$$

where the sum is taken over all energy eigenstates within the given bin and $(\rho_n)_{kk}$ are the diagonal elements of the density matrix ρ_n represented in the basis of the total-energy eigenstates.

IV. LATTICES OF SPINS-1/2

Now we consider a lattice of N_s spins-1/2 and examine how multiple random local measurements affect $g(E)$. We implement an individual local measurement by selecting a random spin at a random time and then measuring its projection on a random axis. The measurements are assumed to occur very rarely with constant average rate per spin τ_m^{-1} (τ_m is much longer than the characteristic time of microscopic dynamics). We label measurements by index n. Each measurement is characterized by the parameters $\{m_n, \vartheta_n, \varphi_n\}$, where m_n labels the lattice site of the measured spin and (ϑ_n, φ_n) are the polar and azimuthal spherical angles indicating the orientation of the spin after the measurement.

A. Projection operator

The projection operator $\mathcal{P}_n$, which represents the nth measurement, is defined as

$$\mathcal{P}_n \equiv \cdots \mathbf{1}_{m_n-1} \otimes (|\vartheta_n\varphi_n\rangle\langle\vartheta_n\varphi_n|)_{m_n} \otimes \mathbf{1}_{m_n+1} \cdots, \qquad (6)$$

where $\mathbf{1}_i$ is the unit matrix acting on the Hilbert space associated with the spin at lattice site i and $|\vartheta_n\varphi_n\rangle = \cos\left(\frac{\vartheta_n}{2}\right)|\uparrow\rangle_z + \sin\left(\frac{\vartheta_n}{2}\right)e^{i\varphi_n}|\downarrow\rangle_z$ is the quantum state of a spin polarized into the direction given by the spherical angles (ϑ_n, φ_n). The operator

$(|\vartheta_n \varphi_n\rangle\langle\vartheta_n\varphi_n|)_{m_n}$ acts on the Hilbert space associated with the spin at lattice site m_n. The operator $\mathcal{P}_n$ satisfies the relations $\mathcal{P}_n^\dagger = \mathcal{P}_n$ and $\mathcal{P}_n\mathcal{P}_n = \mathcal{P}_n$.

The operator $\mathcal{P}_n$ is related to the operator S_n of the m_nth spin projection in the direction (ϑ_n,φ_n) [36] as

$$\mathcal{P}_n = \tfrac{1}{2}\mathbf{1} + S_n, \tag{7}$$

where $\hbar = 1$.

B. Evolution of the energy distribution $g(E)$: General formalism

We denote the density matrix of the total system after n measurements as ρ_n and assume that the initial density matrix ρ_0 is diagonal in the basis of the energy eigenstates (see Appendix B). The transformation from ρ_{n-1} to ρ_n reads

$$\rho_n = \frac{\mathcal{P}_n e^{-i\mathcal{H}(t_n - t_{n-1})} \rho_{n-1} e^{i\mathcal{H}(t_n-t_{n-1})} \mathcal{P}_n^\dagger}{\mathrm{Tr}\left[\mathcal{P}_n e^{-i\mathcal{H}(t_n-t_{n-1})} \rho_{n-1} e^{i\mathcal{H}(t_n-t_{n-1})} \mathcal{P}_n^\dagger\right]}, \tag{8}$$

where t_n is the time of the nth measurement. The corresponding probability distribution of the total energy after n measurements $g_n(E)$ can be expressed as

$$g_n(E) = \frac{1}{B}\left[\mathcal{P}_1^\dagger e^{i\mathcal{H}(t_2-t_1)} \mathcal{P}_2^\dagger \cdots \mathcal{P}_n^\dagger\right.$$
$$\left.\times \mathcal{P}_n \cdots \mathcal{P}_2 e^{-i\mathcal{H}(t_2-t_1)} \mathcal{P}_1\right]_{\mathrm{diag}}(E)\, g_0(E), \tag{9}$$

where B is a normalization factor, and $[\cdots]_{\mathrm{diag}}(E)$ denotes the diagonal elements of the operator in the energy basis averaged over suitably chosen energy bins introduced in Sec. III C. The derivation of Eq. (9) is given in Appendix B.

Since the measured spins are typically far away from each other, the effect of individual measurements in Eq. (9) normally factorizes, which leads to

$$g_n(E) = \frac{1}{B_n}[\mathcal{P}_n]_{\mathrm{diag}}(E)\, g_{n-1}(E), \tag{10}$$

where B_n is a normalization factor. The derivation of Eq. (10) is given in Appendix C. Below we consider the case when two measured spins are accidentally close to each other such that the corresponding effect does not factorize.

The stability measure (2) averaged over all possible outcomes of n measurements reads (see Appendix F)

$$\overline{\Delta G}(n) = \int \left|\prod_{i=1}^{n} \frac{1}{B_i}[\mathcal{P}_i]_{\mathrm{diag}}(E) - 1\right| g_0(E)\,dE, \tag{11}$$

where the bar denotes the result of averaging.

C. Noninteracting spins in a magnetic field

Let us now turn to the example of noninteracting spins in magnetic field H_z with Hamiltonian

$$\mathcal{H} = -H_z \sum_i S_{iz}, \tag{12}$$

where S_{ix}, S_{iy}, and S_{iz} are the spin operators on the ith lattice site. In this case, the outcome of a single spin measurement normally correlates with the total energy of the system and,
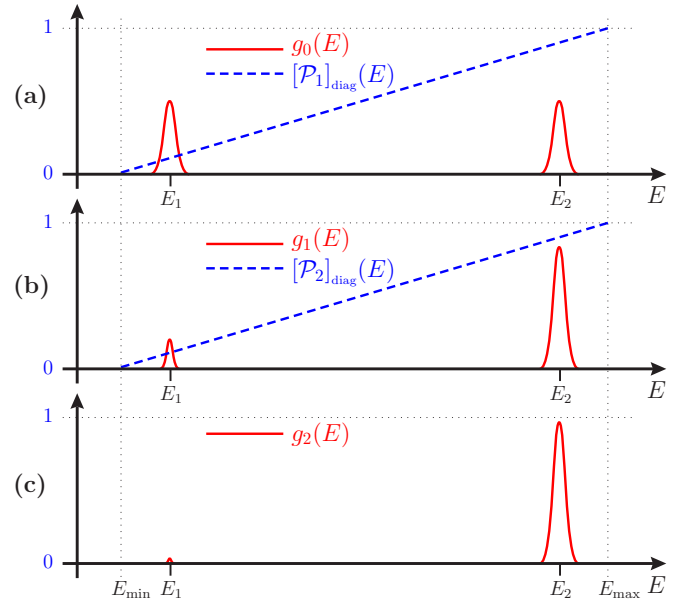

FIG. 1. Schematic representation of the evolution of a two-peak energy distribution $g_n(E)$ (solid red lines) governed by Eq. (10): $g_0(E) = 1/2[\delta(E - E_1) + \delta(E - E_2)]$; $g_1(E) \cong [\mathcal{P}_1]_{\mathrm{diag}}(E)\, g_0(E)$; and $g_2(E) \cong [\mathcal{P}_2]_{\mathrm{diag}}(E)\, g_1(E)$. Here $[\mathcal{P}_1]_{\mathrm{diag}}(E)$ and $[\mathcal{P}_2]_{\mathrm{diag}}(E)$ (dashed blue lines) correspond to two single-spin measurements with respective outcomes $\vartheta_1 = \pi$ and $\vartheta_2 = \pi$ substituted in Eq. (13).

therefore, leads to a significant narrowing of $g(E)$ governed by Eq. (10). A calculation based on Eq. (7) gives (see Appendix D)

$$[\mathcal{P}_n]_{\mathrm{diag}}(E) = \frac{1}{2} - \cos(\vartheta_n)\frac{E}{E_{\max} - E_{\min}}, \tag{13}$$

where $E_{\max} = H_z N_s/2$ and $E_{\min} = -H_z N_s/2$.

The action of transformation (10) consists of "cutting" $g_{n-1}(E)$ by function $[\mathcal{P}_n]_{\mathrm{diag}}(E)$ and then renormalizing the result. This "cutting" normally makes $g_n(E)$ narrower than $g_{n-1}(E)$. The outcome of the next measurement can, in principle, lead to the opposite effect, but it is more probable that it will lead to further narrowing, because the probability of subsequent measurement outcomes is determined by the narrower $g_n(E)$. After many iterations, the drastic narrowing of $g(E)$ becomes overwhelmingly probable.

Figure 1 schematically illustrates the effect of a sequence of transformations (10) applied to the initial two-peak distribution

$$g_0(E) \approx \tfrac{1}{2}[\delta(E - E_1) + \delta(E - E_2)], \tag{14}$$

where $\delta(\dots)$ is a Dirac δ function [37]. In this case, one of the two peaks dominates $g_n(E)$ for $n \to \infty$. Figure 2 presents computed $\overline{\Delta G}(n)$ for the above $g_0(E)$. In each case, we obtain that, after nine measurements, $\overline{\Delta G}(n) > 0.1$ independent of the number of spins in the system, which implies "strong instability." In Appendix F, we obtained the analytical approximation

$$\overline{\Delta G}(n) \approx \sqrt{1 - e^{-\lambda n}}, \tag{15}$$

where $\lambda \cong u^2(E_2 - E_1)^2$ with $u \equiv \overline{|d[\mathcal{P}_n]_{\mathrm{diag}}(E)/dE|} \sim 1/(E_{\max} - E_{\min})$. This approximation is illustrated in Fig. 2.
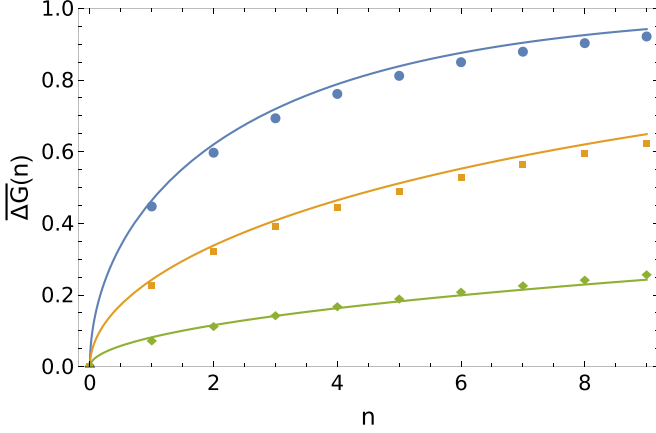
FIG. 2. Averaged ensemble stability measure $\overline{\Delta G}(n)$ as a function of the number of measurements n for a two-peak initial distribution $g_n(E)$. Points represent exact numerically computed results as explained in Appendix E. Lines correspond to the approximated expression (15) with $\lambda = 0.3 \frac{(E_1 - E_2)^2}{(E_{\max} - E_{\min})^2}$. Different colors represent different pairs of values for (E_1, E_2): blue (circles) $(-0.9, 0.9)$, yellow (squares) $(-0.9, 0.0)$, and green (rhombi) $(-0.9, -0.6)$ in units where $E_{\min} = -1$ and $E_{\max} = 1$.

Let us now consider the initial Gaussian distribution

$$g_0(E) \cong \exp\left[-\frac{(E - E_0)^2}{2w_{g,0}^2}\right], \qquad (16)$$

defined by parameters E_0 and $w_{g,0}$, where $w_{g,0} \ll E_{\max} - E_{\min}$. After n measurements, $g_n(E)$ remains approximately Gaussian with the width $w_{g,n}$ following the relation

$$\overline{\frac{1}{w_{g,n}^2}} = \frac{1}{w_{g,0}^2} + u^2 n, \qquad (17)$$

which is derived in Appendix G. This relation leads to $\overline{\Delta G}(n) \sim 1$ when $\overline{1/w_{g,n}^2} \sim 2/w_{g,0}^2$, which corresponds to the number of measurements $n_{cr} \sim 1/(w_{g,0}^2 u^2)$. According to our criterion, the border case for the ensemble stability corresponds to $n_{cr} \sim \sqrt{N_s}$, which implies that the ensemble is unstable for $w_{g,0} \gg (E_{\max} - E_{\min})/4\sqrt{N_s}$. An ensemble with $w_{g,0} \lesssim (E_{\max} - E_{\min})/\sqrt[4]{N_s}$ may still become narrower due to measurements, but the criterion calls it "stable," because the narrowing is relatively slow.

1. Absolute stability of the canonical ensemble

We finally note that, for $w_{g,0} \sim \epsilon_1 \sqrt{N_s}$, where $\epsilon_1 \equiv (E_{\max} - E_{\min})/N_s$ is the characteristic single-spin energy, the decrease of the variance as a result of one measurement is $w_{g,1}^2 - w_{g,0}^2 \approx w_{g,0}^4 u^2 \sim \epsilon_1^2$. This is of the same order of magnitude as the earlier mentioned increase of w_g^2 associated with the broadening effect of $g(E)$ caused by the off-diagonal elements of the projection operator. Therefore, it is reasonable to expect that, for some $w_{g,0} \sim \epsilon_1 \sqrt{N_s}$, the narrowing effect of measurements would compensate the broadening effect, and hence such an ensemble is absolutely stable with respect to measurements. Remarkably, this $w_{g,0}$ is of the order of the width of the canonical ensemble for $T \gtrsim \epsilon_1$ [38,39].

2. Characteristic ensemble-narrowing time

Given that the system is measured on average once per time τ_m/N_s, the above estimates for λ and for n_{cr} imply that the characteristic time to gain $\Delta G \sim 1$ for a broad ensemble with the initial variance $w_{g,0}^2$ is

$$\tau_c \sim \tau_m \frac{\epsilon_1^2 N_s}{w_{g,0}^2}. \qquad (18)$$

For a macroscopic system, τ_c is, therefore, extremely short unless $w_{g,0} \leqslant \epsilon_1 \sqrt{N_s}$.

D. Interacting spins

Now we turn to the Hamiltonian of nearest-neighbor interaction:

$$\mathcal{H} = -\sum_{i<j} J_x S_{ix} S_{jx} + J_y S_{iy} S_{jy} + J_z S_{iz} S_{jz}, \qquad (19)$$

where J_x, J_y, and J_z are the coupling constants. In contrast to the previous case (12), the outcome of a single-spin measurement here is not correlated with the total energy of the system, i.e., $[\mathcal{P}_n]_{\text{diag}}(E) = \text{const}$, and, hence, does not induce narrowing of $g(E)$. At least two accidental measurements sufficiently close in space and time are required for this. Let us consider two such measurements n and $n-1$ at times $t_n > t_{n-1}$. The same treatment that led to Eq. (10) now gives (see Appendix H)

$$g_n(E) = \frac{1}{B_n}[\mathcal{A}_{n,n-1}^\dagger \mathcal{A}_{n,n-1}]_{\text{diag}}(E)\, g_{n-2}(E), \qquad (20)$$

where $\mathcal{A}_{n,n-1} = \mathcal{P}_n e^{-i\mathcal{H}(t_n - t_{n-1})} \mathcal{P}_{n-1}$. Substituting Eq. (7), we obtain in Appendix H

$$\begin{aligned} &[\mathcal{A}_{n,n-1}^\dagger \mathcal{A}_{n,n-1}]_{\text{diag}}(E) \\ &= \tfrac{1}{4} + \tfrac{1}{2}[\{S_{n-1}(t_{n-1}), S_n(t_n)\}]_{\text{diag}}(E) \\ &\quad + [S_{n-1}(t_{n-1}) S_n(t_n) S_{n-1}(t_{n-1})]_{\text{diag}}(E), \end{aligned} \qquad (21)$$

where $\{\cdots, \cdots\}$ is the anticommutator.

1. Relation to equilibrium spin correlation functions

Once the spin orientations obtained in the $(n-1)$-st and the nth measurements are specified and the energy E is fixed, the last two terms on the right-hand side of Eq. (21) can be considered equilibrium spin correlation functions for the microcanonical ensemble corresponding to energy E [40]. For macroscopic systems, these correlation functions equal, in turn, the correlation functions for the canonical ensemble with temperature $T(E)$.

To give an example, let us assume that the outcome of the first measurement is $\vartheta_1 = 0$, $\varphi_1 = 0$ (spin 1 points into the positive z direction) and the outcome of the second measurement is $\vartheta_2 = \frac{\pi}{2}$, $\varphi_2 = 0$ (spin 2 points into the positive x direction). For the two-spin term, we then obtain

$$[\{S_2(t_2), S_1(t_1)\}]_{\text{diag}}(E) = \langle\{S_x(\vec{r}_2, t_2), S_z(\vec{r}_1, t_1)\}\rangle_{T(E)}, \qquad (22)$$

where $\vec{r}_n$ is the position of the nth measured spin.

2. Characteristic ensemble-narrowing time

The above spin correlation functions depend on the time delay and the distance between the two measurements. Therefore, the cutting function in Eq. (21) is strongly influenced by the presence of long-range magnetic order and integrals of motion. We characterize the overall behavior of the spin correlation functions by correlation time $\tau_{\text{corr}}(E)$ and correlation length $\xi(E)$. If a system has constants of motion associated with spin polarizations, this corresponds to an infinite $\tau_{\text{corr}}(E)$. Likewise, if a system exhibits long-range magnetic ordering in a range of energies E, this means that for this range of energies $\xi(E)$ is infinite. [It is worth noting that the correlation functions entering Eq. (21) differ from the conventional correlation functions of spin fluctuations: For the latter the product of single-spin expectation values is subtracted.]

Ordered phase. If a significant part of a broad $g_0(E)$ corresponds to temperatures within the magnetically ordered phase, then the characteristic time τ_c to gain $\Delta G \sim 1$ can be estimated by expression (18). The reason is that, in the magnetically ordered phase, $\xi(E)$ is infinite and, therefore, each measurement correlates with all previous ones in the sense that, for all subsequent measurements k, $[S_n(t_n)S_k(t_k)]_{\text{diag}}(E) \neq 0$. The overall situation resembles the case of Hamiltonian (12), with the external magnetic field replaced by the local field created by the ordered neighbors of each spin.

Nonordered phase. In the nonordered (paramagnetic) phase, $\xi(E)$ is of the order of the nearest-neighbor distance (except for the energy range in the proximity to the magnetic phase transitions). We make the assumption justified by the final result that $\tau_c \ll \tau_{\text{corr}}(E)$ for all energies E, where $\tau_{\text{corr}}(E)$ is the correlation time introduced above. Therefore, we set the time delay entering the cutting function in Eq. (21) effectively to zero. Since $\tau_c \ll \tau_{\text{corr}}$, the outcome of the nth measurement addressing a nearest-neighbor of a previously measured spin is correlated with the total energy of the system. The probability that the nth measurement does not address a nearest-neighbor of a previously measured spin is $P_n = 1 - (n-1)\frac{N_{\text{NN}}}{N_s}$ for $n \ll N_s$, where N_{NN} is the number of nearest neighbors. The probability that, among n measurements, there is no nearest-neighbor pair measured, is

$$P(n) = \prod_{k=1}^{n} P_n = \prod_{k=1}^{n}\left[1 - (k-1)\frac{N_{\text{NN}}}{N_s}\right]$$
$$\approx \exp\left(-\frac{N_{\text{NN}}}{N_s}\sum_{k=1}^{n}k\right) \approx e^{-n^2\frac{N_{\text{NN}}}{N_s}} \quad (23)$$

for $1 \ll n \ll N_s$. Using the relation $n = \frac{N_s t}{\tau_m}$, we finally obtain the probability that, after time t, no pair of nearest neighbors is measured,

$$P(t) \approx e^{-N_{\text{NN}}N_s\frac{t^2}{\tau_m^2}}. \quad (24)$$

Accordingly, $1 - P(t)$ is the probability to measure at least one pair of nearest neighbors. Therefore, for the paramagnetic phase and $w_{g,0} \sim E_{\max} - E_{\min}$, we obtain $\tau_c \sim \tau_m/\sqrt{N_s}$ [41]. If only single-spin measurements were allowed, then the latter ensemble would be called stable according to our criterion. However, the criterion admits any local measurements, including, for example, the measurements of the total spin of

two neighboring spins. For such measurements, the estimate (18) remains valid, thus rendering the ensemble unstable.

We additionally note here that applying an external magnetic field to the paramagnetic phase would drastically shorten τ_c, because, in this case, single-spin measurements described by Eq. (13) would cause a much faster ensemble narrowing.

E. General case

Beyond the Hamiltonians (12) and (19), we expect that, for a broad class of systems with short-range interactions, local measurements whose possible outcomes correlate with the total energy of the system have a narrowing effect comparable to that of single-spin measurements for Hamiltonian (12). In the general case, the analog of the cutting function $[\mathcal{P}_n]_{\text{diag}}(E)$ is much more difficult to calculate directly but, otherwise, can be reasonably expected to be nonlinear with the energy-dependent slope of the order of $1/(\epsilon_1 N_s)$. The latter implies that the estimate (18) for the ensemble-narrowing time and the argument for the proximity of the canonical ensemble to the absolute stability threshold remain valid.

V. NUMERICAL INVESTIGATION OF SPIN SYSTEMS

In order to substantiate the conjectures made in the previous section, we numerically investigate a periodic chain of 24 spins-1/2 governed by the Hamiltonian (19) with $J_x = 0.47$, $J_y = -0.37$, and $J_z = -0.79$. The spins are measured in pairs: An odd-numbered measurement is done on a randomly selected spin, while one of its nearest neighbors is chosen for the next even-numbered measurement. The time delays $t_n - t_{n+1}$ are chosen randomly from the interval [0,2]. The numerical techniques used for the calculations are described in Appendix I.

In Fig. 3, we show a typical evolution of $g(E)$ starting with a two-peak $g_0(E)$ corresponding to the superposition of two quantum states representing canonical ensembles with temperatures $T_1 = 0.1$ and $T_2 = -0.1$. In this example, one peak of $g(E)$ becomes significantly suppressed, which represents the narrowing sketched in Fig. 1.
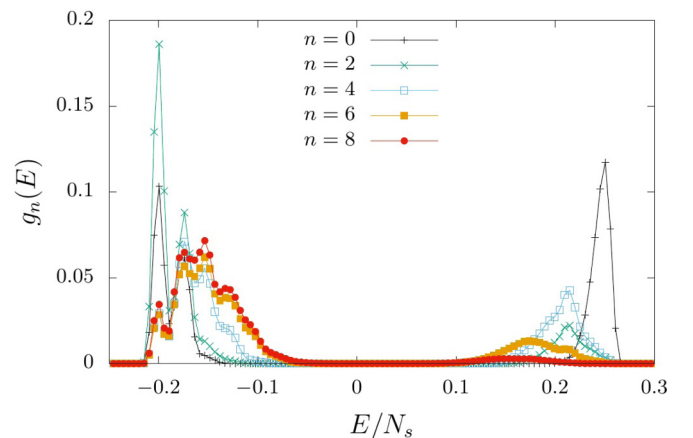
FIG. 3. The energy distribution $g_n(E)$ is shown after each even-numbered measurement as indicated in the legend. The system size is $N_s = 24$.
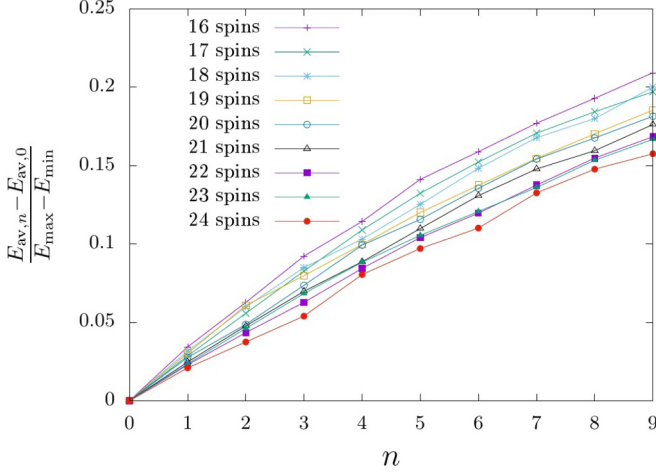
FIG. 4. The quantity $\frac{E_{av,n} - E_{av,0}}{E_{max} - E_{min}}$ is shown for the interacting spin system introduced in Sec. V. The energy distribution corresponds to the Gibbs distribution with the temperature $T = 0.1$. Different system sizes were chosen as indicated in the legend. The plotted values have been averaged over many iterations. The points are connected by lines in order to guide the eye.

Adapting estimate (18) to the present case of spin-pair measurements, we substitute $N_s = 24$, pair measurement time $\tau_m \approx 2N_s$, pair energy $\epsilon_1^2 \approx 0.25(J_x^2 + J_y^2 + J_z^2) = 0.25$, and $w_g \approx 0.2N_s$ (see Fig. 3) to obtain from Eq. (18) $\tau_c \sim 12$ corresponding to $n \sim 12$ and within a factor of two consistent with the characteristic value $n \sim 6$ extracted from Fig. 3.

We further tested the finite-size scaling of heating in small spin systems. As expected, the individual peaks in Fig. 3 also exhibit significant finite-size effects, namely peak broadening and the drift of the peak maxima associated with heating, which prevent us from computing $\Delta G(n)$ representative of the macroscopic limit. However, as we show below, the finite-size scaling of the heating is consistent with our expectation that they can be neglected in the macroscopic limit.

In order to quantify the heating for the small system sizes available in numerical investigations, we calculate the deviation of the average energy E_{av} from its initial value $E_{av,0}$ for a one-peak $g(E)$ corresponding to the canonical ensemble with the temperature $T = 0.1$. In Fig. 4, we show the evolution of $\frac{E_{av,n} - E_{av,0}}{E_{max} - E_{min}}$ as a function of n. In this case, we randomly measure spins individually (i.e., not in pairs) after time delays chosen randomly from the interval [0,2]. The results indicate that the heating becomes relatively weaker with increasing system sizes, which is in agreement with the general expectations from our analytical considerations.

VI. CONCLUSIONS

Let us now make two final remarks.

(1) According to our stability criterion, ensembles with broad $g(E)$ are not stable. In particular, the QMC ensemble [13] is not stable with respect to local measurements, because it implies $g(E)$ having the form of two widely separated peaks corresponding to $T = 0$ and $T = \infty$, respectively [13].

(2) Our stability analysis can also be applied to unconventional quantum ensembles in experiments with non-

macroscopic isolated quantum clusters [22,42–48]. Adapting the obtained results to such systems implies that the above experiments should avoid (i) external magnetic fields, (ii) long-range order, and (iii) local constants of motion.

In conclusion, we have shown that even relatively rare local measurements impose strict constraints on quantum statistical ensembles. We introduced a stability criterion, according to which quantum statistical ensembles characterized by a total energy distribution significantly broader than that of a canonical ensemble are unstable. This result justifies the use of statistical ensembles with narrow $g(E)$ for equilibrium description of macroscopic systems.

APPENDIX A: WIDTH OF $g(E)$ FOR THE CANONICAL ENSEMBLE

For the canonical ensemble,

$$g(E) = \frac{1}{Z} e^{-\beta E} \nu(E) = \frac{1}{Z} e^{-\beta E + S(E)}, \quad (A1)$$

where $Z = \int_{-\infty}^{\infty} e^{-\beta E} \nu(E) dE$ is the partition function, $\beta \equiv \frac{1}{T}$ ($k_B = 1$) is the inverse temperature, and $S(E) \equiv \ln\left[\frac{\nu(E)}{\nu_0}\right]$ is the entropy (ν_0 is an unimportant constant).

The distribution $g(E)$ can be approximated by a Gaussian function $\exp\left[-\frac{(E - E_{av})^2}{2w_g^2}\right]$, where E_{av} is defined by the condition $\beta = S'(E_{av})$ with $S'(E) \equiv \frac{dS(E)}{dE}$. Expanding the exponent in Eq. (A1) at E_{av} up to the second order, we obtain

$$w_g = \sqrt{-\frac{1}{\frac{d^2 S(E)}{dE^2}\big|_{E_{av}}}} = T(E_{av}) \sqrt{\frac{dE(T)}{dT}\bigg|_{T(E_{av})}}$$

$$= T(E_{av}) \sqrt{C_V[T(E_{av})]}, \quad (A2)$$

where $C_V(T) \equiv \frac{dE(T)}{dT}$ is the total specific heat of the system, $T(E_{av}) = \frac{1}{\beta(E_{av})}$, and $\beta(E)$ is the inverse function of $E(\beta)$ defined above. Typically, $C_V \sim N_s$. Therefore, $\frac{w_g}{E_{av} - E_{min}} \sim \frac{1}{\sqrt{N_s}}$, which means that $g(E)$ is sharply peaked.

APPENDIX B: DERIVATION OF EQ. (9)

We denote the initial density matrix of the macroscopic system as ρ_0. We choose ρ_0 to be diagonal in the basis of the energy eigenstates $|E_i\rangle$,

$$\rho_0 = \sum_i (\rho_0)_{ii} |E_i\rangle\langle E_i|. \quad (B1)$$

By doing this, we assume that the off-diagonal elements of ρ_0, even if initially present, would have a negligible effect on the subsequent evolution of $g(E)$ due to the rapid dephasing between states having macroscopically different values of energy.

The transformation from ρ_{n-1} to ρ_n given in Eq. (8) can be iterated to obtain the transformation from ρ_0 to ρ_n,

$$(\rho_n)_{kl} = \frac{1}{B} \langle E_k | \mathcal{O} \rho_0 \mathcal{O}^\dagger | E_l \rangle, \tag{B2}$$

where $(\rho_n)_{kl} = \langle E_k | \rho_n | E_l \rangle$, B is a normalization factor, and

$$\mathcal{O} = \mathcal{P}_n e^{-i\mathcal{H}(t_n - t_{n-1})} \mathcal{P}_{n-1} \cdots \mathcal{P}_2 e^{-i\mathcal{H}(t_2 - t_1)} \mathcal{P}_1 e^{-i\mathcal{H}t_1}. \tag{B3}$$

We are interested in the total-energy distribution and, therefore, focus on $(\rho_n)_{kk}$, i.e., the diagonal elements of the density matrix in the total-energy basis. Substituting Eq. (B3) and Eq. (B1) into Eq. (B2), we obtain after a transformation

$$(\rho_n)_{kk} = \frac{1}{B} \langle E_k | \mathcal{O} \left(\sum_i (\rho_0)_{ii} | E_i \rangle \langle E_i | \right) \mathcal{O}^\dagger | E_k \rangle, \tag{B4}$$

$$= \frac{1}{B} (\rho_0)_{kk} \langle E_k | \mathcal{O} | E_k \rangle \langle E_k | \mathcal{O}^\dagger | E_k \rangle$$

$$+ \frac{1}{B} \langle E_k | \mathcal{O} \left(\sum_{i, i \neq k} (\rho_0)_{ii} | E_i \rangle \langle E_i | \right) \mathcal{O}^\dagger | E_k \rangle. \tag{B5}$$

Using the identity $|E_k\rangle\langle E_k| = \mathbf{1} - \sum_{i, i \neq k} |E_i\rangle\langle E_i|$, we further rewrite Eq. (B5) as

$$(\rho_n)_{kk} = \frac{1}{B} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$$

$$- \frac{1}{B} \sum_{i, i \neq k} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger | E_i \rangle \langle E_i | \mathcal{O} | E_k \rangle$$

$$+ \frac{1}{B} \sum_{i, i \neq k} (\rho_0)_{ii} \langle E_i | \mathcal{O}^\dagger | E_k \rangle \langle E_k | \mathcal{O} | E_i \rangle. \tag{B6}$$

Now we introduce the coarse graining of the energy axis, which means that we divide the energy axis into bins of width Δ_e introduced above. Substituting Eq. (B6) into Eq. (5), we obtain

$$g_n(E) = \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$$

$$- \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_{i, i \neq k} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger | E_i \rangle \langle E_i | \mathcal{O} | E_k \rangle$$

$$+ \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_{i, i \neq k} (\rho_0)_{ii} \langle E_i | \mathcal{O}^\dagger | E_k \rangle \langle E_k | \mathcal{O} | E_i \rangle, \tag{B7}$$

where $\sum$ [without bin(E)] denotes a sum which is not restricted to the bin.

Now we show that the first term in Eq. (B7) makes the main contribution to $g_n(E)$, while the last two terms in Eq. (B7) mainly compensate each other. To show this, we split each unrestricted sum into two sums, where one sum extends over the energy eigenstates within the bin, while the other one extends over the energy eigenstates outside the bin. The latter sum is to be denoted as $\sum^{\overline{\mathrm{bin}(E)}}$. We also use the relation $\langle E_k | \mathcal{O}^\dagger | E_i \rangle \langle E_i | \mathcal{O} | E_k \rangle = |\langle E_i | \mathcal{O} | E_k \rangle|^2$. Hence, we obtain for

the last two terms in Eq. (B7)

$$- \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_{i, i \neq k}^{\mathrm{bin}(E)} (\rho_0)_{kk} |\langle E_i | \mathcal{O} | E_k \rangle|^2$$

$$- \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_i^{\overline{\mathrm{bin}(E)}} (\rho_0)_{kk} |\langle E_i | \mathcal{O} | E_k \rangle|^2$$

$$+ \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_{i, i \neq k}^{\mathrm{bin}(E)} (\rho_0)_{ii} |\langle E_k | \mathcal{O} | E_i \rangle|^2$$

$$+ \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} \sum_i^{\overline{\mathrm{bin}(E)}} (\rho_0)_{ii} |\langle E_k | \mathcal{O} | E_i \rangle|^2. \tag{B8}$$

Now the two terms in Eq. (B8) where both sums are restricted to bin(E) cancel each other. This can be readily seen after exchanging the summation indices i and k in one of these two terms. The remaining terms in Eq. (B7) can be further rewritten as

$$g_n(E) = \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$$

$$- \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} (\rho_0)_{kk} \left(\sum_i^{\overline{\mathrm{bin}(E)}} |\langle E_i | \mathcal{O} | E_k \rangle|^2 \right)$$

$$+ \frac{1}{\Delta_e B} \sum_i^{\overline{\mathrm{bin}(E)}} (\rho_0)_{ii} \left(\sum_k^{\mathrm{bin}(E)} |\langle E_k | \mathcal{O} | E_i \rangle|^2 \right). \tag{B9}$$

The last two terms of Eq. (B9) contain off-diagonal elements $|\langle E_i | \mathcal{O} | E_k \rangle|^2$ corresponding to transitions between energy bins, because E_k and E_i lie in different bins. Let us denote the characteristic energy range $|E_k - E_i|$ of the off-diagonal elements $\langle E_k | \mathcal{O} | E_i \rangle$ as $\Delta_\mathcal{O}$. This range is limited by the condition

$$\Delta_\mathcal{O} \sim \epsilon_1 n \ll \Delta_e. \tag{B10}$$

Therefore, only small energy intervals of length $\Delta_\mathcal{O}$ near the boundaries between the bins contribute to the sums. As shown in Fig. 5, we label these energy intervals as X_-, $\overline{X}_-$, X_+ and $\overline{X}_+$. With such notations, Eq. (B9) can be rewritten as

$$g_n(E) = \frac{1}{\Delta_e B} \sum_k^{\mathrm{bin}(E)} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$$

$$- \frac{1}{\Delta_e B} \sum_k^{X_-} (\rho_0)_{kk} \left(\sum_i^{\overline{X_-}} |\langle E_i | \mathcal{O} | E_k \rangle|^2 \right)$$

$$- \frac{1}{\Delta_e B} \sum_k^{X_+} (\rho_0)_{kk} \left(\sum_i^{\overline{X_+}} |\langle E_i | \mathcal{O} | E_k \rangle|^2 \right)$$

$$+ \frac{1}{\Delta_e B} \sum_i^{\overline{X_-}} (\rho_0)_{ii} \left(\sum_k^{X_-} |\langle E_k | \mathcal{O} | E_i \rangle|^2 \right)$$

$$+ \frac{1}{\Delta_e B} \sum_i^{\overline{X_+}} (\rho_0)_{ii} \left(\sum_k^{X_+} |\langle E_k | \mathcal{O} | E_i \rangle|^2 \right). \tag{B11}$$
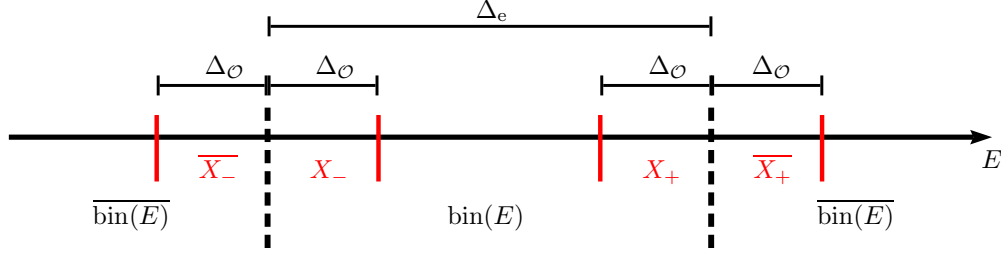
FIG. 5. Schematic representation of the energy intervals $\mathrm{bin}(E)$, $\overline{\mathrm{bin}(E)}$, X_-, $\overline{X}_-$, X_+, and $\overline{X}_+$, introduced in the text. The characteristic size of the intervals Δ_e and $\Delta_{\mathcal{O}}$ are indicated above.

Now we show that the last four terms in Eq. (B11) can be neglected in comparison with the first one, provided $g_0(E)$ does not change too fast. Specifically, we impose the condition (3), which together with the inequalities (4) yields

$$\left| \frac{dg_0(E)}{dE} \right| \lesssim \frac{g_0(E)}{w_{\mathrm{can}}} \ll \frac{g_0(E)}{\Delta_e}, \qquad (\mathrm{B12})$$

where $w_{\mathrm{can}} = T(E_{\mathrm{av}})\sqrt{C_V(E_{\mathrm{av}})}$ is the width of the energy distribution corresponding to the canonical ensemble with the same initial average energy as that of $g_0(E)$. The condition in Eq. (B12) must be satisfied within the energy interval, where $g_0(E)$ is large enough to make a non-negligible contribution to the normalization integral $\int_{-\infty}^{\infty} g_0(E) dE = 1$.

In the first term of Eq. (B11), we expect that, even if $(\rho_0)_{kk}$ and $\langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$ fluctuate with respect to their bin-averaged values, they do it in an uncorrelated way. According to Eq. (5), the bin-averaged value of $(\rho_0)_{kk}$ is $\Delta_e g_0(E)/N_{\mathrm{bin}}(E)$, where $N_{\mathrm{bin}}(E)$ is the number of states within the bin. We define the bin average of $\langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$ as

$$[\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E) \equiv \frac{1}{N_{\mathrm{bin}}(E)} \sum_{k}^{\mathrm{bin}(E)} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle. \qquad (\mathrm{B13})$$

Given the right inequalities in Eqs. (5) and (B12), both bin averages change very weakly over the bin size Δ_e. Therefore, we can approximate the entire first term in Eq. (B11) as

$$\frac{1}{\Delta_e B} \sum_{k}^{\mathrm{bin}(E)} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle \approx \frac{1}{B} [\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E) \, g_0(E). \qquad (\mathrm{B14})$$

Each of the remaining four terms in Eq. (B11) has comparable values. Let us estimate the first of them. We use the following inequality

$$\sum_{i}^{\overline{X}_-} |\langle E_i | \mathcal{O} | E_k \rangle|^2 \leqslant \sum_{i} |\langle E_i | \mathcal{O} | E_k \rangle|^2 = \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle, \qquad (\mathrm{B15})$$

where, as before, the second sum extends over all energy eigenstates of the system. Employing this inequality together

with the assumptions used for deriving Eq. (B14), we obtain

$$\frac{1}{\Delta_e B} \sum_{k}^{X_-} (\rho_0)_{kk} \left(\sum_{i}^{\overline{X}_-} |\langle E_i | \mathcal{O} | E_k \rangle|^2 \right)$$

$$\leqslant \frac{1}{\Delta_e B} \sum_{k}^{X_-} (\rho_0)_{kk} \langle E_k | \mathcal{O}^\dagger \mathcal{O} | E_k \rangle$$

$$\approx \frac{1}{\Delta_e B} \sum_{k}^{X_-} (\rho_0)_{kk} [\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E)$$

$$\approx \frac{1}{B} [\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E) \frac{\Delta_{\mathcal{O}}}{\Delta_e} g_0(E). \qquad (\mathrm{B16})$$

Since $\frac{\Delta_{\mathcal{O}}}{\Delta_e} \ll 1$, Eqs. (B11), (B14), and (B16) imply that

$$g_n(E) \approx \frac{1}{B} [\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E) \, g_0(E), \qquad (\mathrm{B17})$$

which is the same as Eq. (9).

APPENDIX C: DERIVATION OF EQ. (10)

Let us now consider the cutting function in Eq. (B17) or, equivalently, in Eq. (9)

$$[\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E)$$

$$= \left[\mathcal{P}_1^\dagger e^{i\mathcal{H}(t_2-t_1)} \mathcal{P}_2^\dagger \cdots \mathcal{P}_n^\dagger \mathcal{P}_n \cdots \mathcal{P}_2 e^{-i\mathcal{H}(t_2-t_1)} \mathcal{P}_1 \right]_{\mathrm{diag}}(E)$$

$$= \left[\mathcal{P}_1^\dagger(t_1) \mathcal{P}_2^\dagger(t_2) \cdots \mathcal{P}_n^\dagger(t_n) \mathcal{P}_n(t_n) \cdots \mathcal{P}_2(t_2) \mathcal{P}_1(t_1) \right]_{\mathrm{diag}}(E), \qquad (\mathrm{C1})$$

where, in the last expression, we used $\mathcal{P}_k^\dagger(t_k) \equiv e^{i\mathcal{H}t_k} \mathcal{P}_k^\dagger e^{-i\mathcal{H}t_k}$, cf. Eq. (B13). Given that $n \ll \sqrt{N_s}$, it is rather unlikely that among the measured spins $m_1, m_2, \ldots, m_n$ there are two spins which are close to each other. Therefore, we assume the corresponding measurement operators $\mathcal{P}_k^\dagger(t_k)$ to commute. Rearranging the operators in the above expression and using $\mathcal{P}_k^\dagger(t_k) \mathcal{P}_k^\dagger(t_k) = \mathcal{P}_k^\dagger(t_k)$, we obtain

$$[\mathcal{O}^\dagger \mathcal{O}]_{\mathrm{diag}}(E) = [\mathcal{P}_n(t_n) \cdots \mathcal{P}_2(t_2) \mathcal{P}_1(t_1)]_{\mathrm{diag}}(E). \qquad (\mathrm{C2})$$

On the right-hand side of the above expression, there is the n-particle correlation function for the microcanonical ensemble. For systems with short-range interaction, we expect that correlation functions of mutually independent quantities

factorize, which leads to

$$[\mathcal{P}_n(t_n) \cdots \mathcal{P}_2(t_2)\mathcal{P}_1(t_1)]_{\text{diag}}(E)$$
$$= [\mathcal{P}_n(t_n)]_{\text{diag}}(E) \cdots [\mathcal{P}_2(t_2)]_{\text{diag}}(E)[\mathcal{P}_1(t_1)]_{\text{diag}}(E). \quad (C3)$$

This allow us to treat measurements iteratively. Since $g(E)$ does not change with time between the measurements, the result does not depend on the particular values of $t_1, \ldots, t_n$ and, according to Eq. (B13), $\mathcal{P}_k(t_k) = \mathcal{P}_k(0) = \mathcal{P}_k$. Therefore, we obtain

$$g_n(E) = \frac{1}{B_n}[\mathcal{P}_n]_{\text{diag}}(E) \, g_{n-1}(E), \quad (C4)$$

which is the same as Eq. (10).

For macroscopic systems, the above analysis does not depend on the particular value of τ_m. However, for small clusters, Eq. (C4) is valid as long as the delay between measurements τ_m is much longer than the characteristic time scales of the system's microscopic dynamics. These characteristic time scales are, for example, of the order of $\frac{1}{H_z}$ for Hamiltonian (12) or of the order of $\frac{1}{J} = \frac{1}{\sqrt{J_x^2 + J_y^2 + J_z^2}}$ for Hamiltonian (19).

There may be anomalous situations, when the expression (9) for the measurement of a pair of neighboring spins cannot be approximated by the result of two successive applications of Eq. (C4) no matter how long the delay between the two measurements is. An example is the Ising model defined by $J_x = J_y = 0$ in Hamiltonian (19) which has anomalously many local integrals of motion (z components of spins).

APPENDIX D: DERIVATION OF EQ. (13)

The Hamiltonian $\mathcal{H} = -H_z \sum_i S_{iz}$ with $H_z > 0$ for spins in a magnetic field is diagonal in the product basis of individual spin states $|\uparrow\rangle_z$ and $|\downarrow\rangle_z$. Therefore, we obtain for the expectation value of the z polarization of spin at lattice site m_n,

$$\langle S_{nz} \rangle = \frac{\langle (S_z)_{\text{tot}} \rangle}{N_s} = -\frac{E}{N_s H_z} = -\frac{E}{E_{\max} - E_{\min}}. \quad (D1)$$

With $\langle S_n \rangle = \langle S_{nz} \rangle \cos(\vartheta_n)$, this yields

$$[S_n]_{\text{diag}}(E) = -\cos(\vartheta_n)\frac{E}{E_{\max} - E_{\min}}, \quad (D2)$$

which, taking into account the relation $\mathcal{P}_n = \frac{1}{2}\mathbf{1} + S_n$, results in Eq. (13).

APPENDIX E: CALCULATION OF THE RESULTS SHOWN IN FIG. 2

Substituting $g_0(E) = \frac{1}{2}[\delta(E - E_1) + \delta(E - E_2)]$ into Eq. (11), we obtain

$$\overline{\Delta G}(n) = \frac{1}{2}\overline{\left|\prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E_1) - 1\right|}$$
$$+ \frac{1}{2}\overline{\left|\prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E_2) - 1\right|}, \quad (E1)$$

where the bar denotes the result of averaging. The probability for obtaining the given measurement outcomes equals the normalization factor $\prod_{i=1}^{n} B_i$. Given Eq. (13), the above

expression only depends on the polar angles $\vartheta_1, \ldots, \vartheta_n$ describing the measurement outcome. This leads to

$$\overline{\Delta G}(n) = \frac{1}{2}\left[\int_0^{\pi}\prod_{i=1}^{n}B_i\left|\prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E_1) - 1\right|\sin(\vartheta_1)\right.$$
$$\cdots \sin(\vartheta_n)\,d\vartheta_1\cdots d\vartheta_n$$
$$+ \int_0^{\pi}\prod_{i=1}^{n}B_i\left|\prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E_2) - 1\right|\sin(\vartheta_1)$$
$$\left.\cdots \sin(\vartheta_n)\,d\vartheta_1\cdots d\vartheta_n\right] \quad (E2)$$

$$= \frac{1}{2}\left[\int_0^{\pi}\left|\prod_{i=1}^{n}[\mathcal{P}_i]_{\text{diag}}(E_1) - \prod_{i=1}^{n}B_i\right|\sin(\vartheta_1)\right.$$
$$\cdots \sin(\vartheta_n)\,d\vartheta_1\cdots d\vartheta_n$$
$$+ \int_0^{\pi}\left|\prod_{i=1}^{n}[\mathcal{P}_i]_{\text{diag}}(E_2) - \prod_{i=1}^{n}B_i\right|\sin(\vartheta_1)$$
$$\left.\cdots \sin(\vartheta_n)\,d\vartheta_1\cdots d\vartheta_n\right]. \quad (E3)$$

Using Eq. (13), we further obtain

$$\prod_{i=1}^{n}B_i = \int\prod_{i=1}^{n}[\mathcal{P}_i]_{\text{diag}}(E)g_0(E)dE$$
$$= \frac{1}{2}\prod_{i=1}^{n}\left(\frac{1}{2} - \cos(\vartheta_i)\frac{E_1}{E_{\max} - E_{\min}}\right)$$
$$+ \frac{1}{2}\prod_{i=1}^{n}\left(\frac{1}{2} - \cos(\vartheta_i)\frac{E_2}{E_{\max} - E_{\min}}\right). \quad (E4)$$

Substituting the last expression into Eq. (E3), the two summands in Eq. (E3) are identical and this leads to

$$\overline{\Delta G}(n) = \frac{1}{2}\int_0^{\pi}\left[\left|\prod_{i=1}^{n}\left(\frac{1}{2} - \cos(\vartheta_i)\frac{E_1}{E_{\max} - E_{\min}}\right)\right.\right.$$
$$\left.\left. - \prod_{i=1}^{n}\left(\frac{1}{2} - \cos(\vartheta_i)\frac{E_2}{E_{\max} - E_{\min}}\right)\right|\right]$$
$$\sin(\vartheta_1)\cdots\sin(\vartheta_n)\,d\vartheta_1\cdots d\vartheta_n. \quad (E5)$$

Integrating the above expression numerically, we obtain the results shown in Fig. 2.

APPENDIX F: DERIVATION OF THE APPROXIMATION FOR $\overline{\Delta G}(n)$ PLOTTED IN FIG. 2

In this section, we derive the analytical approximation for $\overline{\Delta G}(n)$ used in Fig. 2. From relation (C4), we obtain

$$g_n(E) = \prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E) \, g_0(E). \quad (F1)$$

The substitution of this expression into Eq. (2) leads to

$$\Delta G(n) = \int_{-\infty}^{+\infty}\left|\prod_{i=1}^{n}\frac{1}{B_i}[\mathcal{P}_i]_{\text{diag}}(E) - 1\right|g_0(E)dE, \quad (F2)$$

which, after averaging, gives Eq. (11).

For the initial distribution $g_0(E) = \frac{1}{2}[\delta(E - E_1) + \delta(E - E_2)]$, Eq. (F1) leads to $g_n(E) = p_1\delta(E - E_1) + p_2\delta(E - E_2)$ with some probabilities p_1 and p_2 ($p_1 + p_2 = 1$). We further note that for $n \to \infty$, either $p_1 \to 0$ or $p_2 \to 0$, such that $\Delta G(n) = |p_1 - p_2| \to 1$.

In order to find an approximate expression for $\overline{\Delta G}(n)$, let us first observe that

$$\overline{\Delta G^2}(n) \approx \overline{[\Delta G(n) - \Delta G(n-1)]^2} + \overline{\Delta G^2}(n-1), \quad \text{(F3)}$$

because $\overline{\Delta G(n) - \Delta G(n-1)} \approx 0$. Therefore, we make the ansatz $\overline{\Delta G^2}(n) = \sum_{i=1}^{n} \overline{\gamma_i^2}$, where $\gamma_i = \Delta G(i) - \Delta G(i-1)$. Since the individual summands $\overline{\gamma_i^2}$ become smaller as $\overline{\Delta G^2}(n)$ approaches 1, we make the rough approximation $\overline{\gamma_i^2} \cong 1 - \overline{\Delta G^2}(n)$, which, in the continuum limit for n, leads to the differential equation $\frac{d\overline{\Delta G^2}(n)}{dn} = \lambda[1 - \overline{\Delta G^2}(n)]$, where λ is some constant. Assuming that $\overline{\Delta G^2}(n) \approx \overline{\Delta G}^2(n)$, we obtain

$$\overline{\Delta G}(n) \approx \sqrt{1 - e^{-\lambda n}}. \quad \text{(F4)}$$

We further adopt an approximation

$$\lambda = \kappa\, u^2 (E_1 - E_2)^2, \quad \text{(F5)}$$

where $u \equiv \overline{|d[\mathcal{P}_n]_{\mathrm{diag}}(E)/dE|}$, and κ is a fitting parameter. The expression in Eq. (F5) can be justified by the following considerations. The parameter λ must be equal to 0 when $E_1 = E_2$. Also, λ must remain invariant under a sign change of $E_1 - E_2$. Therefore, the lowest-order term allowed in an analytical expansion of λ around 0 is proportional to $(E_1 - E_2)^2$. Assuming that the value of λ is only controlled by u and $E_1 - E_2$, we conclude that u must also enter quadratically in order for λ to be dimensionless.

In Fig. 2, all three curves have been plotted with the same value for the parameter κ such that $\lambda = 0.3\,(\frac{E_1-E_2}{E_{\max}-E_{\min}})^2$.

The above approximation can be further supported by a more detailed calculation of $\overline{\Delta G}(n)$ in which case the expression for $\Delta G(n)$ in Eq. (F2) is to be averaged over all possible measurement outcomes. For this average, $\Delta G(n)$ must be weighed by the probability for a given set of n measurement outcomes, which equals the normalization coefficient $\prod_{i=1}^{n} N_i = \int \prod_{i=1}^{n} [\mathcal{P}_i]_{\mathrm{diag}}(E)\, g_0(E) dE$. Let us now make a simplifying assumption that the spin measurements are only done along the z direction. Consequently, there are two possible measurement outcomes: positive ($\vartheta_n = 0$) and negative ($\vartheta_n = \pi$). According to Eq. (D2), the projection operator is thus $[\mathcal{P}_i]_{\mathrm{diag}}(E) = \frac{1}{2} \pm \frac{E}{E_{\max}-E_{\min}}$. After averaging over these two possibilities for each measurement, we obtain

$$\overline{\Delta G}(n) = \frac{1}{2} \sum_{k=0}^{n} \left| D_{np_1}(k) - D_{np_2}(k) \right|, \quad \text{(F6)}$$

where $D_{np}(k) = \binom{n}{k} p^k (1-p)^{n-k}$ is the binomial distribution and $p_i = \frac{1}{2} - \frac{E_i}{E_{\max}-E_{\min}}$. The value of $\overline{\Delta G}(n)$ is governed by the overlap as function of k between the two binomial distributions $D_{np_1}(k)$ and $D_{np_2}(k)$. With an increasing n, this overlap decreases approximately exponentially, which is consistent with the asymptotic behavior of $\overline{\Delta G}(n)$ that follows from Eq. (F4) for large n.

APPENDIX G: NARROWING OF A GAUSSIAN PROBABILITY DISTRIBUTION

For a Gaussian distribution

$$g_{n-1}(E) \cong \exp\left[-\frac{(E - E_{\mathrm{av}})^2}{2w_{g,n-1}^2} \right] \quad \text{(G1)}$$

with the variance $w_{g,n-1}^2 \ll (E_{\max} - E_{\min})^2$, the "cutting" by the linear function $[\mathcal{P}_n]_{\mathrm{diag}}(E)$ can be expressed as

$$g_n(E) \sim [\mathcal{P}_n]_{\mathrm{diag}}(E)\, e^{-\frac{(E-E_{\mathrm{av}})^2}{2w_{g,n-1}^2}} = e^{\ln([\mathcal{P}_n]_{\mathrm{diag}}(E)) - \frac{(E-E_{\mathrm{av}})^2}{2w_{g,n-1}^2}}. \quad \text{(G2)}$$

It changes the width according to the relation

$$\frac{1}{w_{g,n}^2} - \frac{1}{w_{g,n-1}^2} = \left[\frac{[\mathcal{P}_n]'_{\mathrm{diag}}(E_{\mathrm{av}})}{[\mathcal{P}_n]_{\mathrm{diag}}(E_{\mathrm{av}})} \right]^2, \quad \text{(G3)}$$

where $[\mathcal{P}_n]'_{\mathrm{diag}}(E_{\mathrm{av}}) = \frac{d[\mathcal{P}_n]_{\mathrm{diag}}(E_{\mathrm{av}})}{dE_{\mathrm{av}}}$. From Eq. (G3), we obtain

$$\overline{\left[\frac{1}{w_{g,n}^2} \right]} \approx \frac{1}{w_{g,0}^2} + u^2\, n, \quad \text{(G4)}$$

where $u = \overline{|[\mathcal{P}_n]'_{\mathrm{diag}}(E_{\mathrm{av}})|}$, and we use $[\mathcal{P}_n]_{\mathrm{diag}}(E_{\mathrm{av}}) \sim 1$.

APPENDIX H: DERIVATION OF Eq. (21)

If two measurements $n-1$ and n accidentally occur close in space and time, the corresponding projection operators $\mathcal{P}_{n-1}(t_{n-1})$ and $\mathcal{P}_n(t_n)$ normally do not commute, and the effect of the two measurements does not factorize. Assuming that all other measured spins are far away from m_{n-1} and m_n, we obtain for the term describing the effect of the two measurements entering the expression in Eq. (C3) $[\mathcal{A}^\dagger_{n,n-1}\mathcal{A}_{n,n-1}]_{\mathrm{diag}}(E)$, where $\mathcal{A}_{n,n-1} = \mathcal{P}_n e^{-i\mathcal{H}(t_n-t_{n-1})}\mathcal{P}_{n-1}$. The expression for the modification of $g(E)$ analogous to Eq. (C4) reads

$$g_n(E) = \frac{1}{B_n}[\mathcal{A}^\dagger_{n,n-1}\mathcal{A}_{n,n-1}]_{\mathrm{diag}}(E)\, g_{n-2}(E). \quad \text{(H1)}$$

For the function $[\mathcal{A}^\dagger_{n,n-1}\mathcal{A}_{n,n-1}]_{\mathrm{diag}}(E)$, we obtain

$$
\begin{aligned}
&[\mathcal{A}^\dagger_{n,n-1}\mathcal{A}_{n,n-1}]_{\mathrm{diag}}(E) \\
&= [\mathcal{P}^\dagger_{n-1}(t_{n-1})\mathcal{P}^\dagger_n(t_n)\mathcal{P}_n(t_n)\mathcal{P}_{n-1}(t_{n-1})]_{\mathrm{diag}}(E) \\
&= [\mathcal{P}_{n-1}(t_{n-1})\mathcal{P}_n(t_n)\mathcal{P}_{n-1}(t_{n-1})]_{\mathrm{diag}}(E). \quad \text{(H2)}
\end{aligned}
$$

Substituting Eq. (7) leads to Eq. (21), which we rewrite here [49],

$$
\begin{aligned}
&[\mathcal{A}^\dagger_{n,n-1}\mathcal{A}_{n,n-1}]_{\mathrm{diag}}(E) \\
&= \frac{1}{4} + \frac{1}{2}[\{S_{n-1}(t_{n-1}), S_n(t_n)\}]_{\mathrm{diag}}(E) \\
&\quad + [S_{n-1}(t_{n-1})S_n(t_n)S_{n-1}(t_{n-1})]_{\mathrm{diag}}(E), \quad \text{(H3)}
\end{aligned}
$$

where $\{S_{n-1}(t_{n-1}), S_n(t_n)\} \equiv S_{n-1}(t_{n-1})S_n(t_n) + S_n(t_n)S_{n-1}(t_{n-1})$. When deriving Eq. (H3), we assumed that the outcomes of single-spin measurements are not correlated with the total energy as explained in the main text. This is formally expressed as $[S_n]_{\mathrm{diag}}(E) = 0$. We further note that the three-spin term in Eq. (H3) should typically be significantly smaller than the two-spin term.

APPENDIX I: NUMERICAL INVESTIGATION OF SPIN SYSTEMS

For the numerical investigation of interacting spin systems, we use the approach established in Refs. [50–52]. This approach is based on the direct integration of the Schrödinger equation and, hence, does not involve the diagonalization of the Hamiltonian matrix.

For the numerical study, we employ the property of quantum typicality, which means that, for quantum systems with very large Hilbert spaces, a single wave function representing a certain ensemble, for example, the ensemble given by $g(E)$, leads to the same average values of interest as the overall ensemble, cf. Eq. (H3). Therefore, we study the stability of energy distributions by considering randomly chosen wave functions.

We prepare the two-peak $g_0(E)$ shown in Fig. 3 by creating a quantum superposition of two pure states, each corresponding to a canonical ensemble, one with temperature $T_1 = 0.1$ and the other with temperature $T_2 = -0.1$. In order to prepare the state at a temperature T_1 (or T_2), we use the method of imaginary-time evolution. First, we randomly select a wave function corresponding to the infinite-temperature limit [51]. Second, we use the relation $e^{-\frac{\mathcal{H}}{T}}|\psi\rangle = e^{-i\mathcal{H}t}|\psi\rangle$, where $t = -i\frac{1}{T}$. For the imaginary time evolution, we use the same technique as for the one for the real-time evolution introduced in Refs. [51,53,54] and also explained below.

For the real-time evolution, we use the fourth-order Runge-Kutta method [51,53,54]

$$|\psi(t + \Delta t)\rangle \approx |\psi(t)\rangle + |v_1\rangle + |v_2\rangle + |v_3\rangle + |v_4\rangle, \quad (I1)$$

where $|v_1\rangle = -i\mathcal{H}|\psi(t)\rangle\Delta t$, $|v_2\rangle = -\frac{1}{2}i\mathcal{H}|v_1\rangle\Delta t$, $|v_3\rangle = -\frac{1}{3}i\mathcal{H}|v_2\rangle\Delta t$, and $|v_4\rangle = -\frac{1}{4}i\mathcal{H}|v_3\rangle\Delta t$. For the Schrödinger equation, this method corresponds to the fourth-order Taylor expansion of the time-evolution operator. It was empirically shown in Ref. [51] that this method is indeed accurate for the time intervals of interest in the present calculations.

In order to measure individual spins, we first choose a random direction (ϑ_n, φ_n) and then, using Eq. (6), calculate $\mathcal{P}_n|\psi\rangle$.

For the calculation of the energy distribution $g_n(E)$ after the nth measurement at time t_n, we perform the following Fourier transform [52]

$$g_n(E) \cong \int \langle \psi(t + t_n)|\psi(t_n)\rangle \, e^{iEt} \, dt. \quad (I2)$$

In order to do this, we first generate a discrete time series by time evolving the wave function immediately after each measurement. In our calculations, the length of the time steps is 0.025 and the number of these time steps is 2048. After generating the time series, we multiply it by the Kaiser-Bessel window function in order to improve the energy resolution. The Kaiser-Bessel window function is defined as

$$\mathcal{K}(k) \equiv \frac{I_0\left[\pi\alpha\sqrt{1 - \left(\frac{2k}{N-1} - 1\right)^2}\right]}{I_0(\pi\alpha)}, \quad (I3)$$

where I_0 is the zeroth-order modified Bessel function of the first kind, α is a non-negative real number which determines the shape of the window, and $N = 2048$. In our calculations, we use $\alpha = 3$. Finally, we Fourier transform the resulting time series and normalize the distribution in order to obtain $g_n(E)$.

---

[1] E. Schrödinger, *Statistical Thermodynamics*, 2nd ed. (Cambridge University Press, Cambridge, 1952).

[2] M. A. Cazalilla and M. Rigol, New J. Phys. **12**, 055006 (2010).

[3] C. Kollath, A. M. Läuchli, and E. Altman, Phys. Rev. Lett. **98**, 180601 (2007).

[4] S. Genway, A. F. Ho, and D. K. K. Lee, Phys. Rev. Lett. **105**, 260402 (2010).

[5] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics* (Springer-Verlag, Berlin, 2009).

[6] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Phys. Rev. Lett. **96**, 050403 (2006).

[7] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).

[8] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).

[9] M. C. Bañuls, J. I. Cirac, and M. B. Hastings, Phys. Rev. Lett. **106**, 050405 (2011).

[10] A. V. Ponomarev, S. Denisov, and P. Hänggi, Phys. Rev. Lett. **106**, 010405 (2011).

[11] M. Rigol, V. Dunjko, and M. Olshanii, Nature **452**, 854 (2008).

[12] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Adv. Phys. **65**, 239 (2016).

[13] B. V. Fine, Phys. Rev. E **80**, 051130 (2009).

[14] B. Fresch and G. J. Moro, J. Chem. Phys. **133**, 034509 (2010).

[15] B. Fresch and G. J. Moro, J. Chem. Phys. **133**, 034510 (2010).

[16] B. V. Fine and F. Hantschel, Phys. Scr. **2012**, 014078 (2012).

[17] W. K. Wootters, Found. Phys. **20**, 1365 (1990).

[18] D. C. Brody and L. P. Hughston, J. Math. Phys. **39**, 6502 (1998).

[19] G. Aarts, G. F. Bonini, and C. Wetterich, Nucl. Phys. B **587**, 403 (2000).

[20] B. Fresch and G. J. Moro, J. Phys. Chem. A **113**, 14502 (2009).

[21] M. P. Müller, D. Gross, and J. Eisert, Commun. Math. Phys. **303**, 785 (2011).

[22] A. Shimizu and T. Miyadera, Phys. Rev. Lett. **89**, 270403 (2002).

[23] H. C. Donker, H. De Raedt, and M. I. Katsnelson, Phys. Rev. B **93**, 184426 (2016).

[24] In principle, the criterion allows for a broader definition of local measurements in the sense that linear combinations of local measurements can also be admitted.

[25] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).

[26] E. Joos and H. D. Zeh, Z. Phys. B **59**, 223 (1985).

[27] E. Joos, H. D. Zeh, C. Kiefer, D. J. W. Giulini, J. Kupsch, and I. O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin Heidelberg, 2003).

[28] A. J. Leggett, J. Phys.: Condens. Matter **14**, R415 (2002).

[29] S. Yuan, M. I. Katsnelson, and H. De Raedt, Phys. Rev. B **77**, 184301 (2008).

[30] M. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition* (Springer, Berlin Heidelberg, 2007).

[31] C. Gardiner and P. Zoller, *Quantum Noise*, 2nd ed. (Springer, Berlin Heidelberg, 1999).

[32] M. Ballesteros, M. Fraas, J. Fröhlich, and B. Schubnel, J. Stat. Phys. **162**, 924 (2016).

[33] C. Guerlin, J. Bernu, S. Deléglise, C. Sayrin, S. Gleyzes, S. Kuhr, M. Brune, J.-M. Raimond, and S. Haroche, Nature (London) **448**, 889 (2007).

[34] M. Bauer and D. Bernard, Phys. Rev. A **84**, 044103 (2011).

[35] H. Maassen and B. Kümmerer, Lecture Notes-Monograph Series **48**, 252 (2006).

[36] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley Publishing Company, Reading, MA, 1994).

[37] The delta-function is an approximation for a peak with a width which is much smaller than the distance between the two peaks but still large enough to satisfy the assumptions behind the derivations of Eqs. (9) and (10).

[38] U. Marzolino and D. Braun, Phys. Rev. A **88**, 063609 (2013).

[39] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, Phys. Rev. X **4**, 031019 (2014).

[40] Spin correlation functions entering Eq. (21) are defined as $\langle S_{i\mu} S_{j\nu} \rangle$, where $\langle \cdots \rangle$ represents the microcanonical average. This definition is to be contrasted with the correlation functions for spin fluctuations $\langle S_{i\mu} S_{j\nu} \rangle - \langle S_{i\mu} \rangle \langle S_{j\nu} \rangle$. The difference is essential in the presence of magnetic order.

[41] Since $\sqrt{N_s}$ is very large, τ_c is certainly much smaller than any reasonable correlation time τ_{corr}. This is a consistency check of our initial assumption $\tau_c \ll \tau_{\text{corr}}$.

[42] P. Schindler, D. Nigg, T. Monz, J. T. Barreiro, E. Martinez, S. X. Wang, S. Quint, M. F. Brandl, V. Nebendahl, C. F. Roos, M. Chwalla, M. Hennrich, and R. Blatt, New J. Phys. **15**, 123012 (2013).

[43] L. Hackermuller, K. Hornberger, B. Brezger, A. Zeilinger, and M. Arndt, Nature **427**, 711 (2004).

[44] F. Jelezko, T. Gaebel, I. Popa, A. Gruber, and J. Wrachtrup, Phys. Rev. Lett. **92**, 076401 (2004).

[45] A. N. Wenz, G. Zürn, S. Murmann, I. Brouzos, T. Lompe, and S. Jochim, Science **342**, 457 (2013).

[46] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwock, J. Eisert, and I. Bloch, Nat. Phys. **8**, 325 (2012).

[47] T. Kinoshita, T. Wenger, and D. S. Weiss, Nature **440**, 900 (2006).

[48] K. Ji and B. V. Fine, Phys. Rev. Lett. **107**, 050401 (2011).

[49] The bracket does not denote a commutator in the expression (H3). For the definition of $[\cdots]_{\text{diag}}(E)$, we refer the reader to the discussion following Eq. (9).

[50] W. Zhang, N. Konstantinidis, K. A. Al-Hassanieh, and V. V. Dobrovitski, J. Phys.: Condens. Matter **19**, 083202 (2007).

[51] T. A. Elsayed and B. V. Fine, Phys. Rev. Lett. **110**, 070404 (2013).

[52] A. Hams and H. De Raedt, Phys. Rev. E **62**, 4365 (2000).

[53] R. Steinigeweg, A. Khodja, H. Niemeyer, C. Gogolin, and J. Gemmer, Phys. Rev. Lett. **112**, 130403 (2014).

[54] R. Steinigeweg, J. Gemmer, and W. Brenig, Phys. Rev. Lett. **112**, 120601 (2014).