Fragile many-body ergodicity from action diffusion

Thudiyangal Mithun [1,2] Carlo Danieli [3,2] M. V. Fistul,[2,4,5] B. L. Altshuler,[6,2] and Sergej Flach[2,7]

[1]*Department of Mathematics and Statistics, University of Massachusetts, Amherst, Massachusetts 01003-4515, USA*
[2]*Center for Theoretical Physics of Complex Systems, Institute for Basic Science, Daejeon 34051, Korea*
[3]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, 01187 Dresden, Germany*
[4]*Theoretische Physik III, Ruhr-Universität Bochum, Bochum 44801, Germany*
[5]*Russian Quantum Center, National University of Science and Technology "MISIS", 119049 Moscow, Russia*
[6]*Physics Department, Columbia University, New York, New York 10027, USA*
[7]*Basic Science Program (IBS School), Korea University of Science and Technology (UST), Daejeon 34113, Korea*

Weakly nonintegrable many-body systems can restore ergodicity in distinctive ways depending on the range of the interaction network in action space. Action resonances seed chaotic dynamics into the networks. Long-range networks provide well connected resonances with ergodization controlled by the individual resonance chaos time scales. Short-range networks instead yield a dramatic slowing down of ergodization in action space, and lead to rare resonance diffusion. We use Josephson junction chains as a paradigmatic study case. We exploit finite time average distributions to characterize the thermalizing dynamics of actions. We identify an action resonance diffusion regime responsible for the slowing down. We extract the diffusion coefficient of that slow process and measure its dependence on the proximity to the integrable limit. Independent measures of correlation functions confirm our findings. The observed fragile diffusion is relying on weakly chaotic dynamics in spatially isolated action resonances. It can be suppressed, and ergodization delayed, by adding weak action noise, as a proof of concept.

I. INTRODUCTION

We address the question of the way a macroscopic nonintegrable system will slow down its thermalization dynamics upon approaching an integrable limit. The conventional perception of evolving dynamical systems with a macroscopic number of degrees of freedom (DoF) is them being in a state of thermal equilibrium, i.e., ergodic. This assumes all allowed microstates having the same probability. It goes along with trajectories visiting the vicinity of all points of the available phase space (i.e., the phase space subject to constraints due to integrals of motion such as, e.g., the energy), and infinite time averages equaling available phase space averages [1]. Statistical physics approaches were paved by Gibbs and Boltzmann and provide a straight connection between microcanonical dynamics and the emergence of canonical distributions [1]. The more interesting it is to study cases when this connection is not evident, eroding, or even missing. This can happen (i) for dynamics in the proximity of an integrable limit, (ii) for dynamics in the proximity of nonergodic sets of measure zero (such as periodic orbits), and (iii) for dynamics driven out of ergodicity due to additional constraints (e.g., condensation). Fermi-Pasta-Ulam-Tsingou problems [2–6] can be associated with (ii), and non-Gibbs states for interacting Bose lattice gases [7–11] with (iii). As for (i), the celebrated Kolmogorov-Arnold-Moser (KAM) theorem is available [12], but applies to systems with finite numbers of DoF and dictates weakly nonintegrable dynamics to be nonergodic on a finite measure set of invariant tori, while being ergodic on the complementary one (Arnold diffusion [13]). The KAM borders are assumed to quickly diminish with increasing DoF numbers [12]. What lies beyond those borders for macroscopic systems? The expectation that Gibbs and Boltzmann take over, was shattered by recent results on many-body localization [14,15], which show that certain quantum many-body systems can resist thermalization at finite distance from integrable limits. With most analytical results being non-rigorous, and computations notoriously heavy due to exploding Hilbert space dimensions, the weakly touched field of ergodization and thermalization of corresponding classical many-body systems is in the focus of this work.

Networks of weakly coupled superconducting grains are one of the few paradigmatic examples of systems where the above scenarios have been considered [16–19]. Also, related networks of interacting anharmonic oscillators were used to argue for and show the existence of two different classes of nonintegrable perturbations of an integrable Hamiltonian $H_0(\{J_k\})$, with a countable set of actions J_k (k being an integer) [20]. Nonintegrable perturbations $H_1(\{J_k, \Theta_k\})$ typically span long-range or short-range networks (LRN or SRN) in the

action-angle space (here Θ_k are the canonically conjugated angles). A reference action J_k in that network is coupled to $R_k \times L_k$-tuples of other action-angle pairs. L_k is typically a single digit integer. R_k, however, can be intensive (SRN) or extensive (LRN) [20].

Let us consider a typical LRN of weakly interacting actions, and translationally invariant two-body interactions $L_k = 3$, $R_k \sim N^2$ with N being the volume (system size) [20]. Such cases turn up when considering translationally invariant systems (extended eigenmodes in the absence of interactions) and weak short-range contact nonlinear interactions, as opposed to long-range couplings leading to critical behavior [21]. Chaotic dynamics can develop in a given L-tuple on a time scale $T_\Lambda = 1/\Lambda$ where Λ is the typical (largest) Lyapunov exponent in the system. Chaotic dynamics develops due to nonlinear resonances which take place when ratios of network matrix elements to certain frequency differences are large [22]. In the proximity to an integrable limit the network matrix elements scale with $1/N$ [20]. Therefore the probability for an L-tuple to be resonant will be $\pi_r/N \ll 1$ where $\pi_r \ll 1$ is an intensive measure of the distance to the integrable limit. However, there are $R_k \sim N^2$ L-tuples in which one given action is involved. Therefore the probability Π_k that the reference action is in resonance with at least one of its L-tuples (i.e., satisfying the resonance condition) turns exponentially close to unity: $\Pi_k \approx 1 - (1 - \pi_r/N)^{R_k} \approx 1 - e^{-R_k \pi_r/N}$ for such a macroscopic LRN. It follows that LRNs thermalize homogeneously in action space, see, e.g., Ref. [20].

On the contrary, SRNs show anomalously slow ergodization and thermalization dynamics in proximity to an integrable limit [18,20]. Since R_k is now intensive and N independent, the resonance probability $\Pi_k \sim \pi_r$ is small, resonances are rare, and thermalization is delayed until resonances were able to migrate through the whole system. Thermalization is expected to be a highly inhomogeneous process in action space.

In this work we quantitatively describe the dynamics of thermalization in the SRN regime by making use of finite time average (FTA) distributions. We (a) observe an alternate regime of action diffusion, and extract the diffusion coefficient as a function of the proximity to the integrable limit, (b) show the connection between the dynamics of FTA distributions and autocorrelation functions and predict and observe algebraic decay of correlations in time, and (c) finally predict the diffusion delay through action noise destroying resonances and provide computational evidence of the delay.

II. MODEL, DISTRIBUTIONS, MEASURES

We consider the Hamiltonian

$$H(q, p) = \sum_{n=1}^{N} \left[\frac{p_n^2}{2} + E_J(1 - \cos(q_{n+1} - q_n)) \right], \quad (1)$$

describing the dynamics of a chain of N superconducting islands with nearest neighbor Josephson coupling in its classical limit. This model is equivalent to a one-dimensional *XY* chain or simply a coupled rotor chain with rotor momenta p_n and angles q_n. E_J controls the strength of Josephson coupling and will be compared to the energy density $h = H/N$. The

equations of motion of Eq. (1) read

$$\dot{q}_n = p_n, \quad \dot{p}_n = E_J[\sin(q_{n+1} - q_n) + \sin(q_{n-1} - q_n)]. \quad (2)$$

We apply periodic boundary conditions $p_1 = p_{N+1}$ and $q_1 = q_{N+1}$. The system has two conserved quantities: the total energy H and the total angular momentum $P = \sum_{n=1}^{N} p_n$. We will choose $P = 0$ without loss of generality. A SRN limit is obtained for $E_J/h \to 0$, with $H_0 = \sum_{n=1}^{N} \frac{p_n^2}{2}$ and $H_1 = \sum_{n=1}^{N} E_J(1 - \cos(q_{n+1} - q_n))$. The actions $J_n \equiv p_n$, and the angles $\Theta_l \equiv q_n$. Note that the opposite limit $E_J/h \to \infty$ (not further studied in this work) yields a LRN network as discussed in the introduction.

The microcanonical dynamics of Eq. (1) explores the available phase space Γ. For an observable $f(\vec{X})$, its phase space average is $\langle f \rangle \equiv \frac{1}{Z} \int f(\vec{X})d\Gamma$, $Z = \int d\Gamma$. Here $\vec{X}$ is a point in Γ. The ergodicity property is tested quantitatively by showing that the infinite time average of any observable $f(\vec{X})$ will be equal to its phase space average $\langle f \rangle$. Lacking infinite times, we rather compute finite time averages, which depend on both the averaging time T, and the initial condition $\vec{X}_0$:

$$f_T(\vec{X}_0) = \frac{1}{T} \int_0^T f(\vec{X}(t))dt, \quad \vec{X}(t = 0) = \vec{X}_0. \quad (3)$$

For an ergodic system it follows

$$\lim_{T \to \infty} f_T(\vec{X}_0) = \langle f \rangle, \quad (4)$$

for any choice of $\vec{X}_0$ except for a subset of measure zero. Dense scanning of all initial points $\vec{X}_0$ over Γ yields the *finite time average distribution* $\rho(f; T)$ of the finite time averages $f_T(\vec{X}_0)$. It is a function of f, parametrically depends on T, and is characterized by its central moments

$$\mu_m(T) = \int (f - \mu)^m \rho(f; T)df, \quad \mu_0 = 1 \quad (5)$$

with $\mu = \int f \rho(f; T)df$. It follows that the first moment $\mu \equiv \langle f \rangle$ is invariant under variation of the averaging time T. All higher moments will in general depend on T. For an ergodic system it follows that

$$\lim_{T \to \infty} \mu_m(T) \to 0, \quad m \geqslant 2. \quad (6)$$

In our studies ρ is close to a Gaussian distribution, which allows us to focus on μ_2.

We use the actions (momenta) p_n as the relevant slow observables in the SRN proximity to the integrable limit $E_J \to 0$: $f \equiv p_n$. With $P = 0$ it follows $\mu_1 = 0$. The second moment $\mu_2(T)$ is then simply the variance of ρ, and further related to the momentum-momentum auto-correlation function $\mathcal{R}(t) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau p_l(\tau)p_l(t + \tau)d\tau$ as $\mu_2(T) = \frac{1}{T} \int_0^T \mathcal{R}(t)dt$. Under the usual assumption that the correlation function will have an exponential decay at large enough times (with ways to even weaken the requirement) the conclusion is that $\mu_2(T \to \infty) \sim 1/T$. This also follows from assuming that at large enough times the finite time averages are obtained from averaging over uncorrelated numbers with some distribution function. With the central limit theorem we arrive at a standard deviation of the order of $1/\sqrt{T}$ and a corresponding variance of the order of $1/T$.
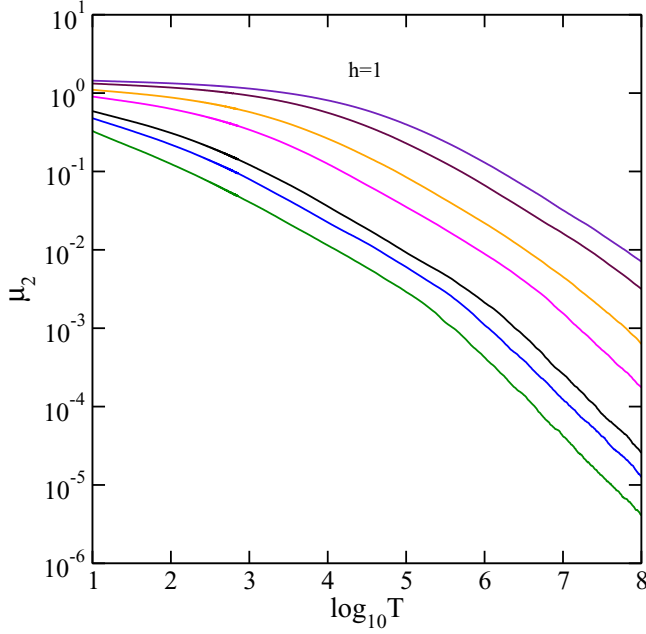
FIG. 1. The time dependence of the second moment $\mu_2(T)$ for various values of E_J: $E_J = 0.25, 0.3, 0.4, 0.5, 0.7, 0.8, 1.0$ from top to bottom. Here $h = 1$, $N = 1024$ and $R = 192$.

FIG. 2. The time dependence of the second moment $\mu_2(T)$ for different system sizes N: $N = 2^6, 2^7, 2^8, 2^9, 2^{10}, 2^{11}, 2^{12}, 2^{13}$ from bottom to top. Here, $h = 1$, $E_J = 0.7$, and $R = 192$.

III. RESULTS

The details of the integration methods are outlined in Appendix A. We numerically integrate R trajectories using symplectic integrators [23], where each initial point is different and was chosen by setting $q_n = 0$, drawing uncorrelated random values of p_n from a Maxwell distribution, constraining $P = 0$, rescaling all momenta such that the desired energy density h is obtained, and giving each trajectory a prethermalization run of $t_{prethermal} = 10^6$. Since all actions p_n are statistically equivalent, we measure them all and add all data into one pool which is used to compute ρ. Figure 1 shows $\mu_2(T)$ for $h = 1$, $N = 1024$ and $0.25 \leqslant E_J \leqslant 1$. The variance $\mu_2(T)$ resists decay up to some characteristic ergodization time scale T_E, after which it turns decaying as expected, signaling restoration of ergodicity. Note that this time scale T_E was assessed in Ref. [18] and is an intensive time scale.

A close inspection of the size dependence of $\mu_2(T)$ is shown in Fig. 2 for $h = 1$, $E_J = 0.7$ and a variety of system sizes. We find that $\mu_2(T_E \leqslant T \leqslant T_D)$ loosely follows a $1/\sqrt{T}$ diffusive decay, which is followed by the anticipated $1/T$ decay for $T_D \leqslant T$. The new time scale $T_D(N)$ is evidently system-size dependent. To support that finding, we plot $\delta = d(log_{10}\mu_2)/d(log_{10}T)$ versus T in Fig. 3. The curves clearly show intermediate saturation on a plateau with $\delta \approx -0.5$, and a subsequent decay at $T_D(N)$ down to $\delta = -1$. Since T_D is increasing with system size, we conjecture that $T_D(N \rightarrow \infty) \rightarrow \infty$, extending the $1/\sqrt{T}$ decay in $\mu_2(T)$ to infinite times for infinite size. In turn this implies that the correlation function $\mathcal{R}(t) \sim 1/\sqrt{t}$ without any exponential cutoff in the same limit.

Let us discuss possible mechanisms leading to the observed behavior of $\mu_2(T)$ for different system sizes by studying the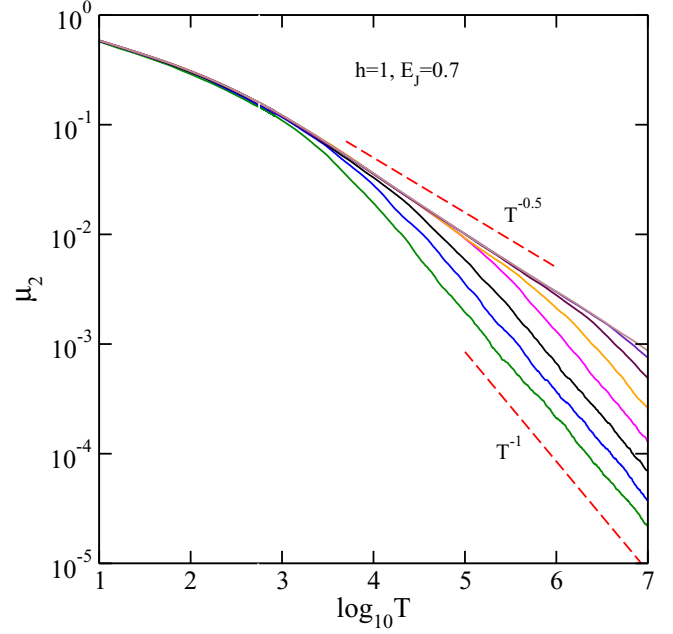 presence and the propagation of chaotic resonances along the chain; similarly to former studies, e.g., [24–26] in both classical and quantum many-body systems. We consider the occurrence of a chaotic resonance to take place when first-order perturbation theory for the evolution of a given rotor at site n breaks down. A simple calculation reported in Appendix B yields $\Delta_n^+ < E_J$ and $\Delta_n^- < E_J$ with $\Delta_n^\pm = |p_n(p_n - p_{n\pm1})|$. The presence of such resonantly coupled triplets of grains along the network generate chaotic dynamics and results in a Lyapunov exponent whose inverse yields a time scale $T_\Lambda \leqslant 10$ on the studied interval $0.25 \leqslant E_J \leqslant 1$ [18]. It follows that $T_\Lambda \ll T_E, T_D$. The resonance probability can be easily computed as (see Appendix B for details).

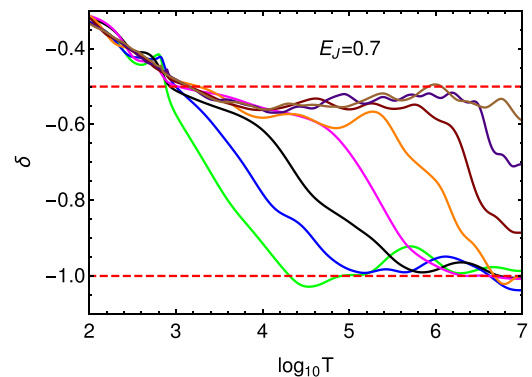$$\pi_r \approx \frac{1}{4}\left(\frac{E_J}{h}\right)^2. \qquad (7)$$

FIG. 3. The time dependence of the parameter $\delta = d(log_{10}\mu_2)/d(log_{10}T)$ for the curves in Fig. 2. System size increases at $\delta = -0.8$ from left to right. The red dashed lines represent $\delta = -0.5$ and $\delta = -1$.
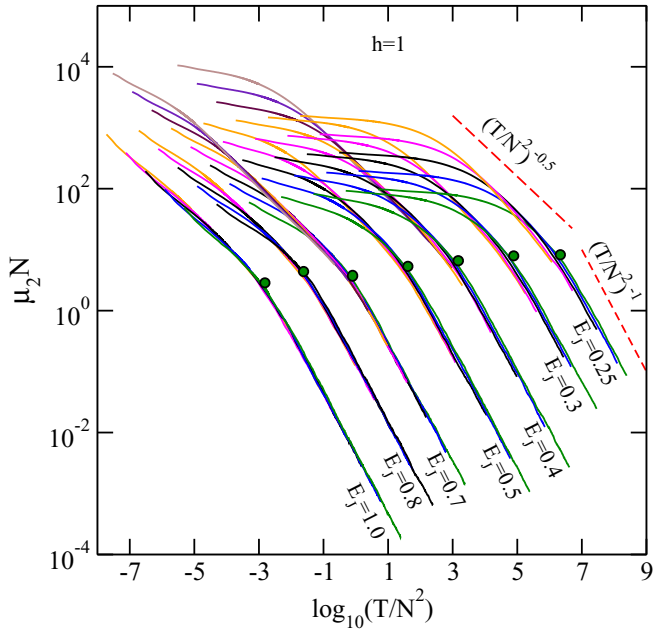
FIG. 4. The rescaled time dependence of $\mu_2 N$ versus bT/N^2 for various E_J: $E_J = 1.0, 0.8, 0.7, 0.5, 0.4, 0.3, 0.25$ from left to right. The corresponding shift factor $b = 0.01, 0.1, 1, 10, 100, 1000, 10000$ is introduced for better visibility of the curves. Here $h = 1$.

At variance to the LRN cases, the SRN resonances are rare and inhomogeneously distributed over the system at any time. The typical distance between consecutive chaotic triplets $l_r \sim (h/E_J)^2$ grows with reducing E_J turning the resonances more sparse and rare [27]. The assumption of partial thermalization of actions involved in the rare resonances will still not lead to any substantial observation of the onset of ergodization, simply because resonances are rare and separated by non-chaotic (regular) regions. To onset ergodicity instead, chaotic resonances have to diffusively migrate throughout the entire system [18]. This presumably happens due to an incoherent detuning of the momenta of rotors in a neighborhood of a given resonance. Once such a neighboring rotor is sufficiently detuned, it could become resonant with its own neighborhood forming a new resonance.

To confirm that we observe action diffusion, we rescale $\mu_2 \rightarrow \mu_2 N$ and $T \rightarrow T/N^2$, as shown in Fig. 4. For a given value of E_J we observe very good collapse of all curves onto one master curve for $T \gtrsim T_E$. The master curves in Fig. 4 show the turnover from diffusive $1/\sqrt{T}$ to asymptotic $1/T$ decay at the time T_D. The diffusion process assumes that a diffusion coefficient $D \sim N^2/T_D$ can be read off a fit of T_D. We found the inverse D^{-1} from the intersection of the fit of the diffusive $1/\sqrt{T}$ and the asymptotic $1/T$ trends (marked with green dots in Fig. 4). The measured values of the diffusion coefficient D are then reported as a function of E_J in Fig. 5, which appears to be reasonably close to a power law over the analyzed interval. To check whether the asymptotic behavior of D may turn into an exponential behavior rather than power law [28–34] we plot the data in linear-log scale in the inset of Fig. 5. From the available data we conclude that a power law is more close to the obtained data.
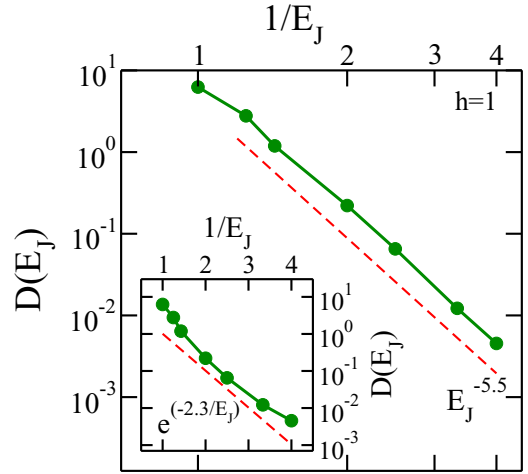

FIG. 5. Diffusion coefficient D vs E_J for fixed $h = 1$ in log-log scale (main) and linear-log scale (inset). The red dashed lines guide the eye and represent a power-law trend $D \sim E_J^{-5.5}$ (main) and an exponential trend $D \sim \exp(-2.3/E_J)$ (inset).

Our analysis shows that the dynamics of resonances starts with a diffusion process between chaotic triplets, i.e., on a length scale $l_r \sim \sqrt{DT_E}$. After that the diffusion continues until all fluctuations stored in N/l_r nonresonant patches each of the size l_r, were exchanged and reached a given location in the system. This happens for $T_D \sim N^2/D$ [35].

If the above scenario is correct, we can expect to delay the diffusion, relaxation, and ergodization process, if we manage to efficiently destroy resonances, before they had time to diffuse. To check this prediction, we take a system with $N = 512$ and $E_j = 0.7$ at $h = 1$. Every time interval $T_\Lambda \approx 10$ we randomly pick a site n, and increment or decrement its momentum p_n by a given value Δ_p with equal probability [36]. On average a given site n_0 is reached on a time $T_\Lambda N \approx 5000$. If $\Delta_p \approx E_J/h$ we expect to efficiently
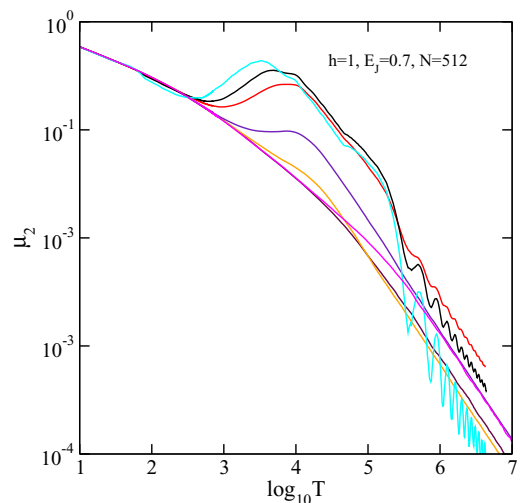

FIG. 6. The time dependence of $\mu_2(T)$ in the presence of a random kick process with $\Delta_p = 0$ (magenta), 0.3 (maroon), 0.6 (orange), 1 (indigo), 2 (red), 3 (black), 5 (cyan). Here $N = 512$, $h = 1$ and $E_J = 0.7$.

detune and destroy a nonlinear resonance. The effect should become visible for $T \approx 5000$. The results in Fig. 6 are excellently reproducing the prediction. The kicks will also generate new resonances at another location, such that the average number of resonances will not change. Our results confirm that resonance diffusion is at the origin of the ergodization process. It is exactly this diffusive process which is efficiently harmed, destroyed and delayed by the above kicking procedure.

IV. CONCLUSIONS

Combining our results with previous studies shows that weakly nonintegrable many-body systems can restore ergodicity in distinctive ways depending on the range of the interaction network in action space. It all starts with action resonances seeding chaotic dynamics into the networks. While long-range networks provide well connected resonances with ergodization controlled by the characteristic individual resonance chaos time scales, short-range networks instead yield a dramatic slowing down of ergodization in action space, and lead to rare resonance diffusion. We used Josephson junction chains as a paradigmatic study case and exploited finite time average distributions to characterize the thermalizing dynamics of actions. The slowing down of the thermalization dynamics upon approaching the integrable limit results in a decreasing of an effective diffusion constant which is related to heat conductivity. This slowing down appears to follow a power law in the distance from the integrable limit, rather than an exponential one. We identify an action resonance diffusion regime responsible for the slowing down. The observed fragile diffusion is relying on weakly chaotic dynamics in spatially isolated action resonances. We were able to successfully delay and suppress it by adding weak action noise, as a proof of concept. Among a number of intriguing open questions, we mention the search for further distinct classes of nonintegrable action networks (neither short nor long ranged), and the impact of quantization on the fragile short-range network dynamics in the vicinity of an integrable limit.

We also note that we studied thermalization or restoring of ergodicity for generic initial data, which excite all available actions to typically nonzero Gibbs distributed values at some finite temperature. This is conceptually different from exciting only one or a few actions to nonzero values, as done in a plethora of studies of the Fermi-Pasta-Ulam-Tsingou problem [5,37] where only one or a few modes (actions) are excited. For the short- range network considered here, a few momenta would be excited to nonzero values only. That will lead straight to the excitation of perturbed discrete breathers [38,39] with solution specific slow relaxation. For the long-range network case the situation is essentially similar, as single mode excitations in a weakly nonlinear system will get closer to exact q-breather states [4].

ACKNOWLEDGMENTS

APPENDIX A: NUMERICAL INTEGRATION DETAILS

We split Eq.(1) into

$$A = \sum_{n=1}^{N} \frac{p_n^2}{2}, \qquad B = E_J \sum_{n=1}^{N} (1 - \cos(q_{n+1} - q_n)). \quad \text{(A1)}$$

As discussed in Ref. [23], this separation leads to a symplectic integration scheme called $ABA864$, where

$$ABA864(\Delta t) = e^{a_1 \Delta t L_A} e^{b_1 \Delta t L_B} e^{a_2 \Delta t L_A} e^{b_2 \Delta t L_B} e^{a_3 \Delta t L_A}$$
$$\times e^{b_3 \Delta t L_B} e^{a_4 \Delta t L_A} e^{b_4 \Delta t L_B} e^{a_4 \Delta t L_A} e^{b_3 \Delta t L_B}$$
$$\times e^{a_3 \Delta t L_A} e^{b_2 \Delta t L_B} e^{a_2 \Delta t L_{AZ}} e^{b_1 \Delta t L_B} e^{a_1 \Delta t L_A},$$

and $a_1 \approx 0.071$, $a_2 \approx 0.241$, $a_3 \approx 0.521$, $a_4 \approx -0.334$, $b_1 \approx 0.183$, $b_2 \approx 0.311$, $b_3 \approx -0.027$, and $b_4 \approx 0.065$; see Ref. [40] for the accurate values of these coefficients. We chose a time-step Δt that ensure that relative energy error $|\frac{H(t)-H(t=0)}{H(t=0)}| \sim 10^{-5}$ during all simulations.

The operators $e^{\Delta t L_A}$ and $e^{\Delta t L_B}$ propagate the set of initial conditions (q_n, p_n) from Eq. (A1) at the time t to the final values (q'_n, p'_n) at the time $t + \Delta t$:

$$e^{\Delta t L_A} : \begin{cases} q'_n = q_n + p_n \Delta t \\ p'_n = p_n \end{cases}$$

$$e^{\Delta t L_B} : \begin{cases} q'_n = q_n \\ p'_n = p_n + E_J[\sin(q_{n+1} - q_n) + \sin(q_{n-1} - q_n)]\Delta t. \end{cases} \quad \text{(A2)}$$

APPENDIX B: RESONANCES

We expand the solutions of Eq. (2) in a Taylor series in powers of the small parameter E_J and initial conditions $p_n(0) = p_n$, $q_n(0) = 0$:

$$\{q_n(t), p_n(t)\} = \sum_{\nu} E_J^{\nu} \{q_n^{(\nu)}(t), p_n^{(\nu)}(t)\}. \quad \text{(B1)}$$

The zero order part of the integrable limit reads

$$p_n^{(0)}(t) = p_n, \ q_n^{(0)} = p_n t. \quad \text{(B2)}$$

and the first-order correction term for the momenta is

$$p_n^{(1)}(t) = \frac{\cos(p_n - p_{n-1})t}{p_n - p_{n-1}} + \frac{\cos(p_n - p_{n+1})t}{p_n - p_{n+1}}. \quad \text{(B3)}$$

The perturbation expansion is assumed to fail when the first-order term is comparable to the zero-order one leading to the following resonance condition:

$$|p_n(p_n - p_{n\pm 1})| \lesssim E_J. \quad \text{(B4)}$$

To ensure chaotic dynamics sets in, we need both resonance conditions to be satisfied, since a pair (dimer) of two coupled momenta will still conserve the total momentum and therefore stay integrable. Consider one of the conditions and rewrite them using the energy $E_n = p_n^2/2$ which is assumed to be distributed according to the Gibbs distribution $P_E(E) = \beta e^{-\beta E}$.

Here $\beta = 1/h$. The resonance condition then reads

$$|E_n \pm \sqrt{E_n E_{n+1}}| \leqslant E_J/2. \qquad (B5)$$

Denoting $E_{n+1} = E_n + \Delta$, assuming $|\Delta| \ll E_n$, choosing the $-$ sign and noting that the distribution of $|\Delta|$ is still the Gibbs one $P_{|\Delta|}(|\Delta|) = \beta e^{-\beta|\Delta|}$ we arrive at the probability of that resonance $|\Delta| \leqslant E_J$ to be satisfied:

$$P_{r1} = \tfrac{1}{2}(1 - e^{-E_J/h}). \qquad (B6)$$

The probability of satisfying both and therefore encountering a chaotic resonance amounts to

$$\pi_r = \frac{1}{4}(1 - e^{-E_J/h})^2 \approx \frac{1}{4}\left(\frac{E_J}{h}\right)^2. \qquad (B7)$$

[1] K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).

[2] J. Ford, The Fermi-Pasta-Ulam problem: Paradox turns discovery, Phys. Rep. **213**, 271 (1992).

[3] T. P. Weissert, *The Genesis of Simulation in Dynamics: Pursuing the Fermi-Pasta-Ulam Problem* (Springer-Verlag, New York, NY, 1997).

[4] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, q-Breathers and the Fermi-Pasta-Ulam Problem, Phys. Rev. Lett. **95**, 064102 (2005).

[5] G. Gallavotti, *The Fermi-Pasta-Ulam Problem: A Status Report*, Vol. 728 (Springer, New York, 2007).

[6] C. Danieli, D. K. Campbell, and S. Flach, Intermittent many-body dynamics at equilibrium, Phys. Rev. E **95**, 060202(R) (2017).

[7] K. Ø. Rasmussen, T. Cretegny, P. G. Kevrekidis, and N. Grønbech-Jensen, Statistical Mechanics of a Discrete Nonlinear System, Phys. Rev. Lett. **84**, 3740 (2000).

[8] S. Iubini, R. Franzosi, R. Livi, G.-L. Oppo, and A. Politi, Discrete breathers and negative-temperature states, New J. Phys. **15**, 023032 (2013).

[9] T. Mithun, Y. Kati, C. Danieli, and S. Flach, Weakly Nonergodic Dynamics in the Gross-Pitaevskii Lattice, Phys. Rev. Lett. **120**, 184101 (2018).

[10] A. Yu. Cherny, T. Engl, and S. Flach, Non-Gibbs states on a Bose-Hubbard lattice, Phys. Rev. A **99**, 023603 (2019).

[11] S. Iubini, L. Chirondojan, G.-L. Oppo, A. Politi, and P. Politi, Dynamical Freezing of Relaxation to Equilibrium, Phys. Rev. Lett. **122**, 084102 (2019).

[12] J. Moser, *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics*, Vol. 1 (Princeton University Press, Princeton, NJ, 2001).

[13] F. Diacu and P. Holmes, *Celestial Encounters: The Origins of Chaos and Stability*, Vol. 22 (Princeton University Press, Princeton, NJ, 1999).

[14] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Metal insulator transition in a weakly interacting many-electron system with localized single-particle states, Ann. Phys. **321**, 1126 (2006).

[15] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[16] D. Escande, H. Kantz, R. Livi, and S. Ruffo, Self-consistent check of the validity of Gibbs calculus using dynamical variables, J. Stat. Phys. **76**, 605 (1994).

[17] M. Pino, L. B. Ioffe, and B. L. Altshuler, Nonergodic metallic and insulating phases of Josephson junction chains, Proc. Natl. Acad. Sci. USA **113**, 536 (2016).

[18] T. Mithun, C. Danieli, Y. Kati, and S. Flach, Dynamical Glass and Ergodization Times in Classical Josephson Junction Chains, Phys. Rev. Lett. **122**, 054102 (2019).

[19] W. De Roeck and F. Huveneers, Glassy dynamics in strongly anharmonic chains of oscillators, C. R. Phys. **20**, 419 (2019).

[20] C. Danieli, T. Mithun, Y. Kati, D. K. Campbell, and S. Flach, Dynamical glass in weakly nonintegrable Klein-Gordon chains, Phys. Rev. E **100**, 032217 (2019).

[21] M. Belger, S. De Nigris, and X. Leoncini, Slowing down of so-called chaotic states: 'Freezing' the initial state, Discontinuity Nonlinearity Complex. **5**, 447 (2016).

[22] B. V. Chirikov, A universal instability of many-dimensional oscillator systems, Phys. Rep. **52**, 263 (1979).

[23] C. Danieli, B. M. Manda, T. Mithun, and Ch. Skokos, Computational efficiency of numerical integration methods for the tangent dynamics of many-body Hamiltonian systems in one and two spatial dimensions, Math. Eng. **1**, 447 (2019).

[24] D. M. Basko, Weak chaos in the disordered nonlinear schrodinger chain: Destruction of Anderson localization by Arnold diffusion, Ann. Phys. **326**, 1577 (2011).

[25] W. De Roeck, F. Huveneers, M. Müller, and M. Schiulaz, Absence of many-body mobility edges, Phys. Rev. B **93**, 014203 (2016).

[26] F. Huveneers, Classical and quantum systems: Transport due to rare events, Ann. Phys. **529**, 1600384 (2017).

[27] D. M. Basko, Local nature and scaling of chaos in weakly nonlinear disordered chains, Phys. Rev. E **86**, 036202 (2012).

[28] C. Giardinà, R. Livi, A. Politi, and M. Vassalli, Finite Thermal Conductivity in 1D Lattices, Phys. Rev. Lett. **84**, 2144 (2000).

[29] V. Oganesyan, A. Pal, and D. A. Huse, Energy transport in disordered classical spin chains, Phys. Rev. B **80**, 115104 (2009).

[30] S. Flach, M. V. Ivanchenko, and N. Li, Thermal conductivity of nonlinear waves in disordered chains, PRAMANA J. Phys. **77**, 1007 (2011).

[31] Y. Li, N. Li, and B. Li, Temperature dependence of thermal conductivities of coupled rotator lattice and the momentum diffusion in standard map, Eur. Phys. J. B **88**, 182 (2015).

[32] W. De Roeck and F. Huveneers, Asymptotic localization of energy in nondisordered oscillator chains, Commun. Pure Appl. Math. **68**, 1532 (2015).

[33] S. Iubini, S. Lepri, R. Livi, and A. Politi, Coupled transport in rotor models, New J. Phys. **18**, 083023 (2016).

[34] M. Kumar, A. Kundu, M. Kulkarni, D. A. Huse, and A. Dhar, Transport, correlations, and chaos in a classical disordered anharmonic chain, Phys. Rev. E **102**, 022130 (2020).

[35] J. T. Edwards and D. J. Thouless, Numerical studies of localization in disordered systems, J. Phys. C **5**, 807 (1972).

[36] Each time one momentum is changed by a given value Δ_p, the other momenta are corrected such that the total momentum P is again zero. After that all momenta are slightly rescaled to reach the original energy density h. Hence, the distance from the integrable limit is not changed by the kick procedure.

[37] E. Fermi, J. Pasta, and S. Ulam, Los Alamos Report LA-1940, 1955 (unpublished); in *Collected Papers of Enrico Fermi*, edited by E. Segré (University of Chicago Press, Chicago, 1965), Vol. 2, p. 978.

[38] S. Flach and C. R. Willis, Discrete breathers, Phys. Rep. **295**, 181 (1998).

[39] S. Flach and A. V. Gorbach, Discrete breathers - advances in theory and applications, Phys. Rep. **467**, 1 (2008).

[40] S. Blanes, F. Casas, A. Farres, J. Laskar, J. Makazaga, and A. Murua, New families of symplectic splitting methods for numerical integration in dynamical astronomy, Appl. Num. Math. **68**, 58 (2013).