Nonlinear perturbation of Random Matrix Theory

Klaus M. Frahm[1] and Dima L. Shepelyansky[1]

[1] *Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, 31062 Toulouse, France*

(Dated: December 22, 2022)

We consider a system of linear oscillators, or quantum states, described by Random Matrix Theory and analyze how its time evolution is affected by a nonlinear perturbation. Our numerical results show that above a certain chaos border a weak or moderate nonlinearity leads to a dynamical thermalization of a finite number of degrees of freedom with energy equipartition over linear eigenmodes as expected from the laws of classical statistical mechanics. The system temperature is shown to change in a broad range from positive to negative values and the dependence of system characteristics on the initial injected energy is determined. Below the chaos border the dynamics is described by the Kolmogorov-Arnold-Moser integrability. Due to universal features of Random Matrix Theory we argue that the obtained results describe the generic properties of its nonlinear perturbation.

PACS numbers:

In far 1872, 150 years ago, Boltzmann developed the theory of statistical mechanics and thermalization originated from the dynamical laws of classical motion of many-body systems [1]. This result led to the famous Boltzmann-Loschmidt dispute on a possibility of thermalization and time irreversibility emerging from the reversible dynamical equations of particle motion [2, 3] (see also [4]). The modern resolution of this dispute is based on the theory of dynamical chaos for generic nonlinear systems characterized by a positive maximal Lyapunov exponent and Kolmogorov-Sinai entropy leading to an exponential instability of motion (see e.g. [5–8]). This instability leads to an exponential growth of errors which breaks time reversibility (see e.g. an example in [9]).

The first numerical studies of how ergodicity, dynamical thermalization and energy equipartition appear in an oscillator system with moderate nonlinearity were reported by Fermi, Pasta, Ulam in 1955 [10]. The conclusion was that "The results show very little, if any, tendency toward equipartition of energy between the degrees of freedom." [10]. It was argued in [11] that in the continuum limit the Fermi-Pasta-Ulam (FPU) problem is close to the Korteweg-de Vries equation with stable soliton solutions shown to be completely integrable [12], as well as the nonlinear Schrödinger equation [13]. In addition, at weak nonlinearity the FPU α-model is close to the completely integrable Toda lattice [14, 15]. Another explanation of equipartition absence in the FPU problem was given in [16–18] showing that below a certain strength of nonlinear interactions between oscillator modes the system is located in the regime of Kolmogorov-Arnold-Moser (KAM) integrability and only above this border an overlap of nonlinear resonances takes place with emergence of chaos and thermalization. Numerical simulations demonstrated a dynamical thermalization with energy equipartition reported in [17, 18]. Thus, even 50 years after [10], various regimes of nonlinear dynamics of the FPU problem are actively discussed by the community of dynamical systems [19] (see e.g. recent [20]). The variety of studies clearly demonstrates that this model played an important role in the investigations of nonlinear dynamics but also that it has multiple specific features indicating that it does not belong to a class of generic oscillator systems with nonlinear interactions.

To construct a generic model of many-body oscillator systems with nonlinear interactions between oscillators we take insight from quantum mechanics of many-body systems whose spectral properties are described by Random Matrix Theory (RMT) invented by Wigner for a description of spectra of complex nuclei, atoms and molecules [21]. At present RMT finds applications in multiple areas of physics [22, 23] including systems of quantum chaos whose dynamics is chaotic in the classical limit [24, 25]. The properties of RMT eigenvalues and eigenstates were established in various studies and are well known. The RMT eigenstates are ergodic, i.e. uniformly distributed on the N-dimensional unit sphere, and the level spacing statistics is described by the universal RMT distribution [21–25]. Due to the linearity of the Schrödinger equation the time evolution of a wave function ψ described by a RMT Hamiltonian also describes a time evolution of a system of N linear oscillators with random linear couplings. By its own, due to the universal properties of RMT, it is interesting to understand how a nonlinear perturbation affects RMT evolution.

With the aim to understand the effects of nonlinear perturbation of RMT we consider a simple model described by the Schrödinger equation with a Hamiltonian given by a random matrix with an additional nonlinear interaction between linear modes:

$$i\hbar\frac{\partial\psi_n(t)}{\partial t} = \hat{H}_0\psi_n(t) + \beta|\psi_n(t)|^2\psi_n(t) \quad . \qquad (1)$$

Here $\hat{H}_0$ is an RMT matrix of size N taken from the Gaussian Orthogonal Ensemble (GOE) [22] whose matrix elements $H_{n,n'}$ have zero mean and variance $\langle H^2_{n,n'}\rangle = (1+\delta_{n,n'})/(4(N+1))$ such that the averaged density of states is given by the the semi-circle law $dm/dE = \frac{2N}{\pi}\sqrt{1-E^2}$ with typical eigenvalues in the interval $E_m \in [-1,1]$ (we use dimensionless units with $\hbar = 1$), β is a

dimensionless constant characterizing the nonlinear interaction strength in the original basis n.

The eigenmodes of $\hat{H}_0$ at energies E_m are $\phi_m(n)$ which are ergodic with a uniform distribution on the N-dimensional unit sphere. The time evolution of the wave function can be expressed in the basis of eigenmodes as $\psi_n(t) = \sum_{m=1}^{N} C_m(t)\,\phi_m(n)$ with coefficients $C_m(t)$ giving the occupation probability $\rho_m = \langle |C_m(t)|^2 \rangle$ (with some long time or ensemble average; see below). The time evolution (1) has two integrals of motion being the probability norm $\sum_n |\psi_n(t)|^2 = 1$ and total energy $E = \sum_n [< \psi_n(t)|\hat{H}_0|\psi_n(t) > + (\beta/2)|\psi_n(t)|^4]$. At $\beta = 0$ the model (1) can be viewed as a quantum system or as a classical system of coupled linear oscillators whose Hamiltonian in the basis of oscillator eigenmodes is $H = \sum E_m C_m^*(t)\, C_m(t)$ where C_m, C_m^* is a pair of conjugated variables and E_m plays the role of oscillator frequencies. Since RMT captures the universal features of quantum and linear oscillator systems we expect that the model (1) describes the universal properties of oscillator systems with chaotic dynamics induced by weak or moderate nonlinear couplings between oscillators. We call the model (1) Nonlinear Random Matrix model (NLIRM).

Above a certain chaos border with $\beta > \beta_c$ a moderate nonlinearity destroys KAM integrability leading to chaotic dynamics with a positive maximal Lyapunov exponent λ. The nonlinear frequency shift is $\delta\omega \sim \beta|\psi_n|^2 \sim \beta/N$ and, as it was argued in [26–29], a developed chaos takes place when this shift $\delta\omega$ becomes comparable to a typical energy spacing between energies (or frequencies) of the linear system $\Delta\omega \sim 1/N$. Thus $\delta\omega > \Delta\omega$ implies chaos with the chaos border $\beta_c = \text{const.} \sim 1$ being independent of system size N.

The issue of dynamical thermalization in finite size nonlinear lattices with disorder was studied in [29, 30]. The time evolution in these systems is described by the Discrete Anderson Nonlinear Schrödinger Equation (DANSE) with hopping between nearby sites. In the linear case the disorder leads to Anderson localization of modes [31] which is well visible when the localization length ℓ is smaller than the system size N. In this respect our RMT model (1) is rather different since the linear modes are delocalized and ergodic in a vector space of dimension N. We expect that our model (1) is generic and captures also certain features of the models of Bose-Einstein condensate (BEC) evolution in the chaotic Bunimovich stadium [32] or the Sinai oscillator [33] described by the nonlinear Gross-Pitaevskii equation (GPE) [34]. Indeed, the linear eigenmodes of these systems have properties of quantum chaos similar to RMT [24, 25]. There are however also certain differences discussed below.

For the GPE models [32, 33] it is natural to assume that the dynamical thermalization induced by moderate nonlinearity leads to the Bose-Einstein (BE) distribution of probabilities ρ_m over quantum levels of the linear system. In the limit of high temperature T this distribution is reduced to a classical energy equipartition (EQ) distribution [4, 35]. For the DANSE type models [29, 30]

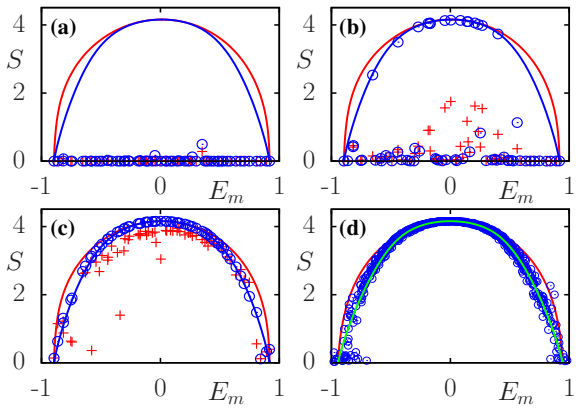

FIG. 1: Entropy S versus energy E_m of the initial state m at $t = 0$ for one RMT realisation at $N = 64$ and $\beta = 0.02$ (a), $\beta = 0.1$ (b) and $\beta = 1$ (c) or 10 RMT realisations at $\beta = 1$ (d). The entropy S is computed from ρ_m averaged over the time range $2^{23} \leq t \leq 2^{24}$ (blue/black $\circ$) and $2^{16} \leq t \leq 2^{17}$ (red/grey $+$ in (a), (b)) or $2^{11} \leq t \leq 2^{12}$ (red/grey $+$ in (c)). The theory curves $S(E)$ for BE (red/grey) and EQ (blue/black in (a),(b),(c) or green in (d)) are from ρ_m values of (2) with E_m values of the used RMT realisation (a), (b), (c) or a fictitious spectrum according to the semi-circle law in (d) (where E_m is the solution of $m - 1/2 = M(E_m)$, $m = 1, \ldots, N$ with $M(E)$ being the integrated density of states).

the quantum Gibbs (QG) distribution was proposed to explain numerically obtained results. In fact QG and BE distributions give very close thermalization properties and we mainly discuss the BE case here. Thus there are two options for the thermalized distributions of probabilities ρ_m:

$$\rho_m = \frac{1}{\exp[(E_m - \mu)/T] - 1}\ (\text{BE}), \rho_m = \frac{T}{E_m - \mu}\ (\text{EQ}). \quad (2)$$

Here T is the system temperature and $\mu(T)$ is the chemical potential dependent on temperature. The parameters T and μ are determined by the norm and energy conservation $\sum_m \rho_m = 1$ and $\sum_m E_m \rho_m = E$ (for E we assume the case of weak or moderate nonlinearity which gives only a weak contribution to the total energy). The entropy S of the system is determined by the usual relation [4, 35]: $S = -\sum_m \rho_m \ln \rho_m$ with the implicit theoretical dependencies on temperature $E(T)$, $S(T)$, $\mu(T)$. The derivation of (2) is given in Supplementary Material (SupMat).

Based on classical statistical mechanics [4, 35] the dynamical thermalization should lead to the EQ distribution (2) since DANSE, GPE [29, 30, 32, 33] and NLIRM (1) models describe classical nonlinear fields without second quantization. In contrast, in [29, 30, 32, 33] it was argued that a moderate nonlinearity plays a role of an effective nonlinear thermostate that leads to quantum BE or QG distributions (2).

Of course, both BE and EQ approaches (2) give different thermal characteristics leading to a contradiction

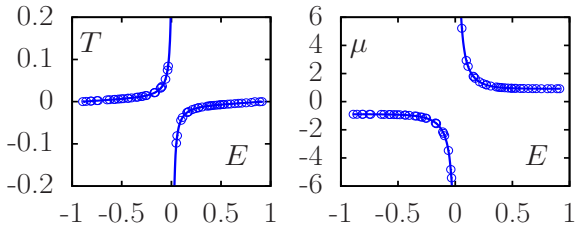

FIG. 2: Dependence of T and μ on energy E for EQ ansatz (2) (curves); data points are for $\beta = 1$, $N = 64$ and time range $2^{23} \leq t \leq 2^{24}$, with T and μ determined from norm and numerical entropy values (same RMT realisation as in Fig. 1 (c)).

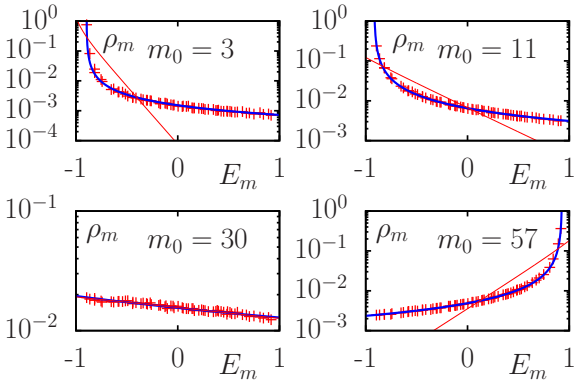

FIG. 3: Dependence of $\rho_m(E_m)$ on E_m for 4 initial states at $m_0 = 3, 11, 30, 56$ for $\beta = 1$, $N = 64$ and time average range $2^{23} \leq t \leq 2^{24}$. The blue/black curve shows theory of EQ ansatz with with $\rho_{\rm EQ}(E) = T/(E-\mu)$. The red/grey line shows BE ansatz theory $\rho_{\rm BE}(E) = 1/(\exp[(E-\mu)/T] - 1)$. The related T, μ theory (2) values are given in SupMat Fig. S4 for BE and EQ cases.

discussed in detail in [30, 32, 33]. The main argument in favor of the BE (or the QG) ansatz was based on a reasonably good agreement of numerical data for entropy vs energy with the theoretical thermal dependence $S(E)$ given by the BE (or QG) ansatz. The quantities S and E are extensive (self averaging) and it was argued that their analysis is more preferable as compared to the direct study of the strongly fluctuating probabilities ρ_m [29, 30, 32, 33]. Here we show that the ergodicity of RMT eigenstates of $\hat{H}_0$ allows to reduce significantly the fluctuations and to obtain stable results for ρ_m that are clearly described by the EQ ansatz (2).

The numerical integration of (1) is done with the symplectic scheme of order 4 [36] using a step size $\Delta t = 0.1$ up to maximal times $t = 4 \times 10^6 - 1.3 \times 10^8$ with exact norm conservation and energy conservation with accuracy $\sim 10^{-8}$ (see SupMat for more details). As initial condition, we choose an eigenmode $\phi_m(n)$ of $\hat{H}_0$ at some index m (some times also noted m_0) such that the energy remains close to the initial energy $E \approx E_m$. Examples of the time dependence $S(t)$ are shown in SupMat Fig. S1 demonstrating a steady-state regime reached at times $t > 10^4$ for $\beta = 1$. The obtained dependence $S(E)$ is shown in Fig. 1 at different β values for a specific RMT realisation and two time scales and also for 10 RMT realisations at $\beta = 1$. At small values $\beta = 0.02, 0.1$ the system is close to an integrable KAM regime [7, 8] while at $\beta = 1$ essentially all modes are thermalized (see Fig. 1, SupMat Fig. S2 and additional material in [37]). These results show that the critical border for thermalization is located at $\beta_c \sim 0.1$ independent of N. However, the exact determination of β_c is a rather complicated task due to the presence of many-body nonlinear effects like e.g. the Arnold diffusion [7, 8]. Also at the spectral borders $E \approx \pm 1$ the spacing between energies E_m increases according to the semicircle law [22] and therfore it is more difficult to reach thermalization there.

An important feature of Fig. 1 is that the theory curves $S(E)$ obtained with BE and EQ ansatz (2) are rather close to each other. Thus due to fluctuations of numerical data for $S(E)$ it is difficult to determine which theory BE or EQ describes better the numerical data. However, the data points are significantly closer to the BE-curve, especially for moderate energies $|E| \approx 0.5 - 0.8$ where both curves are somewhat different (the difference between the QG and BE $S(E)$ curves, not visible on graphical precision, is ~ 0.003 at the spectral borders and much smaller at other E values, so that we discuss mainly the BE case).

For the QE ansatz the dependencies $T(E)$, $\mu(E)$, obtained by the solution of the equations for energy and norm for a given RMT spectrum, are shown in Fig. 2 (SupMat Fig. S3 for the BE ansatz) for the thermalized regime at $\beta = 1$. The numerical points obtained from E and norm values are by definition exactly located on the theory curves. If instead of E we use the numerical data of S then the points slightly deviate from the theory (Fig. 2 and SupMat Fig. S3) but T and μ values themselves are drastically different between BE and EQ cases.

The most direct way to distinguish between BE and EQ cases is to compare the probability dependence $\rho_m(E)$ with the theory (2). Such a comparison is shown in Fig. 3 for 4 initial states at m=m_0, $\beta = 1$ and $N = 64$ (more data are in SupMat Fig. S4 and [37]). The dynamical thermalization clearly follows the EQ ansatz and not at all the BE one, except for an initial state at $E_{m_0} \approx 0$ where both approaches are equivalent. This observation is in agreement with the classical statistical mechanics [4, 35]. The probabilities ρ_m for all initial energies E_{m_0} are shown in Fig. 4 with a good agreement between numerical data and the EQ ansatz (see [37] for figures as Fig. 3 for all m_0 values). The statistical distribution $p(x)$ of fluctuations of the rescaled quantity $x = (E_{m_0} - \mu)|C_m(t)|^2/T$ (with μ, T from the EQ ansatz for the energy E_{m_0}) also follows the Boltzmann law $p(x) = \exp(-x)$ (see SupMat Fig. S5).

In Fig. 5 we show the energy dependence of the maximal positive Lyapunov exponent λ_m on energy E_m of initial state m for different β values (more data are in

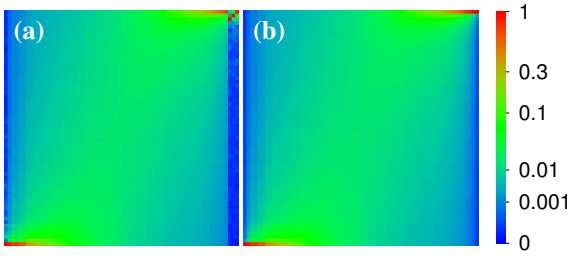

FIG. 4: Density plot of ρ_m for parameters of Fig. 3 with initial state index $1 \leq m_0 \leq 64$ in the x-axis and $1 \leq m \leq 64$ in the y-axis. The colorbar shows ρ_m values in a nonlinear scale to increase the visibility of small ρ_m values. Panel (a) shows numerical data for $\beta = 1$, $N = 64$; panel (b) shows the EQ ansatz $\rho_{\rm EQ}(E_m)$ (see also Fig. 3).

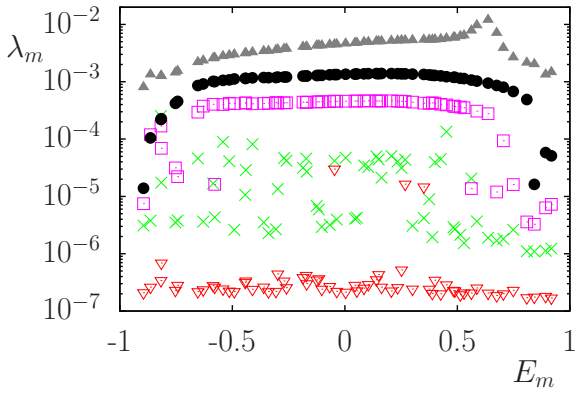

FIG. 5: Lyapunov exponent λ_m dependence on E_m with m being the index of the initial state for $N = 64$; λ_m is determined from the fit $\ln \|\Delta\psi(t)\| = a + b\ln(t) + \lambda_m t$ for $\beta = 2$ (grey ▲; top), 1 (black •), 0.5 (pink □), 0.1 (green ×) at $t \leq 2^{22}$; $\beta = 0.02$ for $t \leq 2^{27}$ (red ▽; bottom).

SupMat Figs. S6-S10 and [37]). In the thermalized phase $\beta = 1$ we have a smooth variation of λ with E_m while below or close to the thermalization border at $\beta = 0.1$ high λ_m values appear only at specific E_m values. We attribute this to the existence of triplets of energies with very close E_m values. Indeed, in a hypothetic case of 3 equal E_m values the KAM theory is not valid and developed chaos exists at arbitrary small β values as it shown in [26, 38]. Nonetheless, in RMT there is level repulsion and double or triple degeneracies are forbidden leaving place only to quasi-degeneracy of levels so that KAM becomes valid at $\beta \to 0$. Thus for $\beta = 0.02$ we have typically λ_m approaching to zero with increasing time. Our preliminary results show that in the thermal phase at larger $|T|$ (if $E_m \approx 0$) we have an approximate dependence $\lambda \sim \beta^\eta / N^\nu$ with $\eta \approx 1.52, \nu \approx 1.72$ (see SupMat Figs. S6-S10). However, the Lyapunov exponent dependence on β and N requires further more detailed studies.

Finally, we discuss the reasons why the nature of thermal equipartition, BE or EQ, was so difficult to establish in previous studies [29, 30, 32, 33]. One of them is the proximity of $S(E)$ curves for both approaches. At the same time the direct determination of the $\rho_m(E)$ dependence is rather difficult due to significant fluctuations, as it was pointed out previously. These fluctuations are especially large for the DANSE case at a large disorder ($W = 4$ in [29]) when the localization length ℓ is significantly smaller than system size N ($\ell/N \approx 0.1$ at $N = 64$). We illustrate this in SupMat Figs. S11-S12 showing that at smaller disorder $W = 2$ with larger localization length ℓ the fluctuations of ρ_m are reduced and at long times we have an agreement of $\rho_m(E)$ with the EQ ansatz and strong deviations from the BE ansatz. For NLIRM model (1) the linear eigenmodes are ergodic, i.e. no localization, and the fluctuations of $\rho_m(E)$ are significantly reduced that allows to distinguish clearly between EQ and BE cases.

The cases of GPE in the Bunimivich stadium [32] and the Sinai-oscillator trap [33] are somewhat different. Indeed, in these models the spectrum of the linear system is unbounded so that, even if linear eigenstates are in the quantum chaos regime, the probability spreading to high energies is rather slow due to small coupling transitions induced by nonlinearity between states with significantly different energies. Thus in these systems there is a formation of a relatively compact probability packet at low energies which spreads to high energies very slowly in time. Such an energy packet of ρ_m gives $S(E)$ values compatible with the curve of the BE ansatz however the fluctuations of $\rho_m(E)$ are very strong with a significant difference from the BE distribution at high energies (see e.g. Fig.5 in [32] and Figs. 8,11 in [33]). To analyze these features in more detail, we add to the diagonal RMT matrix element $H_{n,n}$ an additional diagonal energy fn with a constant $f > 0$. Then the variation of linear energies fN becomes rather large and exceeds significantly those of the RMT case. The results for this model at $\beta = 1$, $f = 0.25$ show that at times $t = 2^{15}$ for $N = 32$ (or $t = 2^{20}$ for $N = 64$) the probabilities $\rho_m(E)$ form a compact packet of approximate BE shape and the EQ thermal distribution is reached (with fluctuations) only at very large times $t = 2^{27}$ (see SupMat Figs. S13, S14). Such large time scales were out of reach in [32, 33] due to the complexity of the numerical integration of GPE.

In conclusion, we showed that a nonlinear perturbation of RMT leads to dynamical thermalization with energy equipartition corresponding to the laws of classical statistical mechanics [4, 35]. Such a thermalization appears due to dynamical chaos in finite systems with moderate or large number of degrees of freedom at weak or moderate perturbation of a linear RMT system. At very weak perturbations the system dynamics is characterized by a quasi-integrable KAM regime. We argue that the proposed NLIRM model captures the generic features of dynamical thermalization in systems weakly perturbed by classical nonlinear fields. Of course, for finite many-body quantum systems with second quantization the interactions lead to quantum dynamical thermalization and distributions of Bose-Einstein for bosons or Fermi-Dirac for fermions, as it has been demonstrated in numerical stud-

ies [39] and [40–42] respectively.

Acknowledgments: This work has been partially supported through the grant NANOX N^o ANR-17-EURE-0009 in the framework of the Programme Investissements d'Avenir (project MTDINA). This work was granted access to the HPC resources of CALMIP (Toulouse) under the allocation 2022-P0110.

[1] L. Boltzmann, *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, Wiener Berichte **66**, 275 (1872).

[2] J. Loschmidt, *Über den Zustand des Wärmegleichgewichts eines Systems von Körpern mit Rücksicht auf die Schwerkraft*, Sitzungsberichte der Akademie der Wissenschaften, Wien **II 73**, 128 (1876).

[3] L. Boltzmann, *Über die Beziehung eines allgemeine mechanischen Satzes zum zweiten Haupsatze der Wärmetheorie*, Sitzungsberichte der Akademie der Wissenschaften, Wien **II 75**, 67 (1877).

[4] J.E. Mayer, M. Goeppert-Mayer, *Statistical mechanics*, John Wiley & Sons, N.Y. (1977).

[5] V. Arnold, A. Avez, *Ergodic problems of classical mechanics*, Benjamin, N.Y. (1968).

[6] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic theory*, Springer-Verlag, N.Y. (1982).

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

[8] A. Lichtenberg and M. Lieberman, *Regular and Chaotic Dynamics*, Springer, N.Y. (1992).

[9] D. L. Shepelyansky, *Some statistical properties of simple classically stochastic quantum systems*, Physica D **8**, 208 (1983).

[10] E.Fermi, J.Pasta and S.Ulam, *Studies of non linear problems*, Los Alamos Report LA-1940 (1955); published later in E.Fermi *Collected papers*, E.Serge (Ed.) **2**, 491, Univ. Chicago Press, Chicago IL (1965); see also historical overview in T.Dauxois *Fermi, Pasta, Ulam and a mysterious lady*, Phys. Today **61(1)**, 55 (2008).

[11] N.J. Zabusky and M.D. Kruskal, *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett. **15**, 240 (1965).

[12] C. S. Gardner, J. M. Greene, M. D. Kruskal and R. M. Miura, *Method for solving the Korteweg - de Vries equation*, Phys. Rev. Lett. **19**, 1095 (1967).

[13] V.E. Zakharov and A.B. Shabat, *Interaction between solitons in a stable medium*, Sov. Phys. JETP **37(5)**, 823 (1973).

[14] M. Toda, *Studies of a non-linear lattice*, Phys. Reports **18(1)**, 1 (1975).

[15] G. Benettin, H. Christodoulidi and A. Ponno, *The Fermi-Pasta-Ulam problem and its underlying integrable dynamics*, J. Stat. Phys. **152**, 195 (2013).

[16] B.V. Chirikov and F.M. Izrailev, *Statistical properties of a non-linear string*, Sov. Phys. Doklady **11(1)**, 30 (1966).

[17] B.V.Chirikov, F.M.Izrailev and V.A.Tayursky, *Numerical experiments on statistical behavior of dynamical systems with a few degrees of freedoms*, Comp. Comm. Phys. **5**, 11 (1973).

[18] R. Livi, M. Pettini, S. Ruffo and A. Vulpiani, *Chaotic behavior in nonlinear Hamiltonian systems and equilibrium statistical mechanics*, J. Stat. Phys. **48**, 539 (1987).

[19] G.Gallavotti (Ed.), *The Fermi-Pasta-Ulam problem: a status report*, Lect. Notes Phys. **728**, Springer, Berlin (2008).

[20] M. Gallone, M. Marian, A. Ponno and S. Ruffo, *Burgers turbulence in the Fermi-Pasta-Ulam-Tsingou chain*, Phys. Rev. Lett. **129**, 114101 (2022).

[21] E.P. Wigner, *Random matrices in physics*, SIAM Review **9(1)**, 1 (1967).

[22] M.L. Mehta, *Random matrices*, Elsvier, Amsterdam (2004).

[23] T. Guhr, A. Müller-Groeling and H.A. Weidenmüller, *Random Matrix Theories in quantum physics: common concepts*, Phys.Rep. **299**, 189 (1998).

[24] O. Bohigas, M.-J. Giannoni and C. Schmit, *Characterization of chaotic quantum spectra and universality of level fluctuation laws*, Phys. Rev. Lett. **52**, 1 (1984).

[25] F. Haake, *Quantum signatures of chaos*, Springer, Berlin (2010).

[26] B.V. Chirikov and D.L. Shepelyanskii, *Dynamics of some homogeneous models of classical Yang-Mills fields*, Sov. J. Nucl. Phys. **36(6)**, 908 (1982).

[27] D.L. Shepelyansky, *Delocalization of quantum chaos by weak nonlinearity*, Phys. Rev. Lett. **70**, 1787 (1993).

[28] I.Garcia-Mata and D.L.Shepelyansky, *Delocalization induced by nonlinearity in systems with disorder*, Phys. Rev. E **79**, 026205 (2009).

[29] M. Mulansky, K. Ahnert, A. Pikovsky and D.L. Shepelyansky, *Dynamical thermalization of disordered nonlinear lattices*, Phys. Rev. E **80**, 056212 (2009).

[30] L. Ermann and D.L. Shepelyansky, *Quantum Gibbs distribution from dynamical thermalization in classical nonlinear lattices*, New J. Phys. **15**, 123004 (2013).

[31] P.W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109**, 1492 (1958).

[32] L. Ermann, E. Vergini and D.L. Shepelyansky, *Dynamical thermalization of Bose-Einstein condensate in Bunimovich stadium*, Europhys. Lett. **111**, 50009 (2015).

[33] L. Ermann, E. Vergini and D.L. Shepelyansky, *Dynamics and thermalization of a Bose-Einstein condensate in a Sinai-oscillator trap*, Phys. Rev. A **94**, 013618 (2016).

[34] L. Pitaevskii and S. Stringari, *Bose-Einstein condensation*, Oxford Univ. Press, Oxford (2003).

[35] L.D. Landau and E.M. Lifshitz, *Statistical physics*, Wiley, New York (1976).

[36] E. Forest and Ronald D. Ruth, *Fourth-order symplectic integration.* Physica D. 43: 105 (1990); `https://cloudfront.escholarship.org/dist/prd/content/qt35h9v2k9/qt35h9v2k9.pdf` (Accessed Dec 2022).

[37] `https://www.quantware.ups-tlse.fr/QWLIB/nonlinrmt/index.html` (Accessed Dec 22, 2022).

[38] M. Mulansky, K. Ahnert, A. Pikovsky and D.L. Shepelyansky, *Strong and weak chaos in weakly nonintegrable many-body Hamiltonian systems*, J. Stat. Phys. **145**, 1256 (2011).

[39] P. Schlageck and D.L. Shepelyansky, *Dynamical ther-*

malization in Bose-Hubbard systems, Phys. Rev. E **93**, 012126 (2016).
[40] A.R. Kolovsky and D.L. Shepelyansky, *Dynamical thermalization in isolated quantum dots and black holes*, Europhys. Lett. **117**, 10003 (2017).
[41] K.M. Frahm and D.L. Shepelyansky, *Dynamical decoherence of a qubit coupled to a quantum dot or the SYK black hole*, Eur. Phys. J. B **91**, 257 (2018).
[42] K.M. Frahm, L. Ermann and D.L. Shepelyansky, *Dynamical thermalization of interacting fermionic atoms in a Sinai-oscillator trap*, MDPI Condens. Matter **4**, 76 (2019).

Supplementary Material for

Nonlinear perturbation of Random Matrix Theory

by K. M. Frahm and D. L. Shepelyansky
Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, 31062 Toulouse, France

I. STATISTICAL CLASSICAL THEORY

Since the "nonlinear Schrödinger equation" (1) has two integrals of motion, neglecting the energy $\sim \beta/N$ due to the nonlinear term and assuming global chaos for sufficiently large β, we expect that the system becomes ergodic or thermalizes on the manifold fixed by the two constraints :

$$\sum_m E_m |C_m|^2 = E \quad , \quad \sum_m |C_m|^2 = 1$$

where C_m are the coefficients of the state in the expansion of the eigenbasis of the matrix $\hat{H}_0$. This situation corresponds in principle to a micro canonical ensemble with an additional constraint which is technically quite complicated. One can use $|C_1|^2 = 1 - \sum_{m=2}^N |C_m|^2$ to remove the first coordinate C_1 from the phase space to obtain a pure micro canonical ensemble for $C_2, \ldots, C_N$ with:

$$E - E_1 = \sum_{m=2}^N (E_m - E_1)|C_m|^2$$

but there is still the condition $\sum_{m=2}^N |C_m|^2 = 1 - |C_1|^2 \leq 1$ which creates technical complications. For small temperature or energy (with E being close to E_1, assuming an ordered eigenvalue spectrum $E_1 < E_2 < \ldots < E_N$) one can neglect this condition and in this case it is not difficult to show by standard text book techniques of statistical physics that in the limit $N \gg 1$ the marginal distribution of a field C_m (integrating out the other fields of the micro-canonical ensemble) is a (complex) Gaussian

$$p(C_m) \sim \exp\left(-\frac{(E_m - E_1)|C_m|^2}{T_{mc}} \right)$$

with the micro-canonical temperature $T_{mc} = (E - E_1)/N$ and providing the equipartition average : $\rho_{m,mc} = \langle |C_m|^2 \rangle = T_{mc}/(E_m - E_1)$.

However, for larger energies the additional inequality for the coefficients C_m cannot be neglected. Therefore, we treat the system as a grand-canonical ensemble, which is equivalent for $N \gg 1$. In this approach the fields C_m can freely fluctuate and the constraints are only verified

in average. The classical grand canonical partition function is given by

$$Z = \int \prod_m d^2C_m \, \exp\left(-\frac{1}{T}\sum_m (E_m-\mu)|C_m|^2\right)$$
$$\sim T^N \prod_m \frac{1}{E_m-\mu} \quad \Rightarrow$$
$$\ln(Z) = N\ln(T) - \sum_m \ln(E_m-\mu) + \text{const.}$$

with two parameters being the (grand canonical) temperature T and the chemical potential μ which are determined by the implicit equations

$$1 = \sum_m \rho_m \quad , \quad E = \sum_m E_m \rho_m \tag{S.1}$$

with ρ_m being the statistical average :

$$\rho_m = \langle |C_m|^2 \rangle = \frac{T}{E_m-\mu} \equiv \rho_{EQ}(E_m) \; .$$

Here we have either $T > 0$ and $\mu < E_1$ or $T < 0$ and $\mu > E_N$ in order to have well defined Gaussian integrals in the partition function and only solutions for T and μ satisfying this condition are valid. From

$$E-\mu = \left\langle \sum_m (E_m-\mu)|C_m|^2 \right\rangle = T^2 \frac{\partial \ln(Z)}{\partial T}$$
$$= T^2 \frac{N}{T} \quad \Rightarrow \quad T = \frac{E-\mu}{N} \tag{S.2}$$

we find that μ is a solution of the implicit equation:

$$1 = T\sum_m \frac{1}{E_m-\mu} = \frac{1}{N}\sum_m \frac{E-\mu}{E_m-\mu} \; . \tag{S.3}$$

For a given value of E and a given spectrum E_m this equation can be solved numerically by standard techniques and using (S.2) we also obtain T once μ is known. Depending on the sign of $E - \sum_m E_m < 0$ (or > 0) we have either $\mu < E_1$ and $T > 0$ (or $\mu > E_N$ and $T < 0$) as unique and physically valid solution (mathematically there are typically many other but invalid solutions of (S.3) in the interval $E_1 < \mu < E_N$). Once $\mu(E)$ and $T(E)$ are known one can use ρ_m to compute the entropy

$$S_{EQ}(E) = -\sum_m \rho_{EQ}(E_m)\ln(\rho_{EQ}(E_m)) \; .$$

This expression was used to compute the theoretical $S(E)$ curves in the equi-partition approach based on the grand-canonical classical theory shown in Figs. 1, S2, S11, S13 for various examples.

We mention that the grand canonical temperature (S.2) is similar to the micro-canonical temperature if we replace $E_1 \to \mu$ and it is not difficult to verify that in the limit $E \searrow E_1$ we have $\mu \nearrow E_1$ with $T \searrow 0$ (or if $E \nearrow E_N \quad \Rightarrow \quad \mu \searrow E_N$ with $T \nearrow 0$; see also Figs. 2 and S3). Also the micro-canonical expression for ρ_m provides numerically correct $S(E)$ curves (identical to the grand canonical curve) for the lower 20%-30% of the energy spectrum where $\mu \approx E_1$ with a rather good accuracy.

The Bose-Einstein ansatz with

$$\rho_m = \rho_{BE}(E) \equiv \frac{1}{e^{(E_m-\mu)/T}-1} \tag{S.4}$$

cannot be directly justified by the classical field approach. From a purely formal point of view it can be obtained by replacing in the partition function $|C_m|^2 \to c_m$ with integer c_m and replacing the Gaussian integrations by sums over $c_m = 0, 1, 2, \ldots$ thus resulting in (S.4). In the framework of this approach T and μ are computed by solving numerically the implicit equations (S.1) with $\rho_m = \rho_{BE}(E_m)$ which is technically a bit more complicated as for the EQ case. In the limit of large $|T|$ we can expand in (S.4) the exponential and both approaches become equivalent.

The difference between both approaches in the $S(E)$ curves is not very strong but the numerical data of long time averages of $\rho_m = \langle |C_m(t)|^2 \rangle$ clearly show the validity of the EQ model provided the state is sufficiently thermalized as can be seen in Figs. 3, and S4.

Furthermore, according to both the micro-canonical and grand-canonical approaches the statistical distribution of C_m is a complex Gaussian which corresponds to an exponential distribution of $|C_m|^2$, i.e. the distribution of the rescaled variable $x = (E_m-\mu)|C_m|^2/T$ is theoretically $p(x) = \exp(-x)$ which is clearly confirmed by the numerical data for quite large values of x as can be seen in Fig. S5 providing an additional confirmation of the classical model.

Both approaches require the use of a given fixed energy spectrum E_m which is typically obtained by diagonalizing a certain realisation of an RMT matrix (or another matrix for the variants as DANSE or the model with additional diagonal elements). However, in Fig. 1 (d), we show the data for 10 different RMT realisations which would provide individually slightly different $S(E)$ curves. For this figure we used, for both theoretical $S(E)$ curves, a fictitious spectrum with E_m being the solution

of $m-1/2 = M(E_m)$ for $m = 1, \ldots N$ where

$$M(E) = \frac{2N}{\pi}\int_0^E \sqrt{1-E'^2}\,dE' = \frac{N}{2} + \frac{N}{\pi}\left(\arcsin(E) + E\sqrt{1-E^2}\right)$$

is the integrated density of states of the RMT semi-circle law such that $M(-1) = 0$ and $M(1) = N$. This fictitious spectrum corresponds to a constant uniform level spacing in the unfolded spectrum.

The link between the radius (here being unity) of the semi-circle law of a GOE matrix and the variance of its matrix elements $\langle H^2_{n,n'}\rangle = (1+\delta_{n,n'})/(4(N+1))$ is rather standard [22]. However, it can be easily verified by computing the average

$$\left\langle \mathrm{Tr}(\hat{H}_0^2)\right\rangle = \frac{1}{4(N+1)}\big(2N + N(N-1)\big) = \frac{N}{4}$$

which should coincide with

$$\sum_m \left\langle E_m^2\right\rangle = \frac{2N}{\pi}\int_{-1}^{1} dE\, E^2\sqrt{1-E^2} = \frac{N}{4}\ .$$

II. SYMPLECTIC INTEGRATOR

Here we remind some basic facts about symplectic integrators and the particular implementation for our case. For further details, its derivation, we refer for example to [36], especially for the 4th order variant.

A. General method

Let A and B two non-commuting operators of a general Lie algebra for which it is possible to compute exactly and efficiently (by some exact numerical/analytical method) $\exp(tA)$ and $\exp(tB)$ individually and for arbitrary values of t (or more precisely these operators applied to some given vector of function) while the numerical problem to compute $\exp[t(A+B)]$ is very difficult (very inefficient) or even impossible (as far as an exact method is concerned).

To solve this problem it is sufficient to compute $\exp[\Delta t(A+B)]$ for small Δt (with some given precision) and then to apply: $\exp[t(A+B)] = \exp[\Delta t(A+B)]^n$ with $n = t/\Delta t$ (assuming that t is an integer multiple of Δt). To compute $\exp[\Delta t(A+B)]$ approximately one can write:

$$\exp[\Delta t(A+B)] \approx \prod_{j=1}^{p}\Big[\exp(d_j\Delta t A)\,\exp(c_j\Delta t B)\Big]$$

where the product is ordered with increasing j-values from right to left. The coefficients c_j, d_j, $j = 1, \ldots, p$ are determined such that the error (for one step) is $\sim (\Delta t)^{p+1}$ for a given order p and implying a global error $\sim (\Delta t)^p$ (for many steps and fixed t). The simplest case is $p = 1$ with $c_1 = d_1 = 1$ corresponding to the usual Trotter formula. For $p = 2$, we have the symmetrized Trotter formula with $c_1 = 0$, $c_2 = 1$, $d_1 = d_2 = \frac{1}{2}$. For $p = 3$ there is a non-symmetric solution which can also be found in [36] (see references therein for the proper credit) but which is not really simpler (with all 6 coefficients being different from zero) than the fourth order solution. For $p = 4$ there is a symmetric solution which according to [36] is:

$$c_1 = 0,\ c_2 = c_4 = 2x+1,\ c_3 = -4x-1,$$
$$d_1 = d_4 = x + 0.5,\ d_2 = d_3 = -x$$

where $x = (2^{1/3} + 2^{-1/3} - 1)/6$ is the real solution of $48x^3 + 24x^2 - 1 = 0$. Note that these coefficients verify the sum rule $\sum_j c_j = \sum_j d_j = 1$ due to the first order terms in both exponential expressions. The fourth order formula requires as the third order formula the multiplication of 6 exponential factors for one step if one uses an optimization to merge the d_4-factor with the d_1-factor of the next step (a similar optimization is possible for the symmetrized Trotter formula).

In typical applications one applies this method to solve numerically the time evolution a classical Hamiltonian or quantum system where the Hamiltonian is a sum of two terms $H_1 + H_2$ for which the individual exponentials (of either the Liouville operator associated to H_j or $-iH_j$, $j = 1, 2$) can be computed analytically or by an efficient exact numerical method.

The advantage of the method is that it respects the symplectic/unitary symmetry of the problem. Furthermore, even if one chooses a low order variant with a not so small time step Δt, one can argue that the approximate time evolution (with respect to "$A+B$") represents in reality the *exact* time evolution of a slightly different operator $S \approx A + B$ such that $\exp(\Delta t S)$ coincides exactly with the above product of exponential terms and that many physical features of the modified time evolution are still very relevant since they apply to the same "class" of systems.

B. Numerical implementation

In our case, we chose $A = -i\hat{H}_0$ (in the quantum point of view or the Liouville operator associated to $\hat{H}_0$ in the classical point of view) and $B = -i\hat{V}(\psi)$ where $\hat{V}(\psi)$ is

an effective potential depending on ψ and with matrix elements $V_{n,n'}(\psi) = \beta|\psi_n|^2\,\delta_{n,n'}$. In this case $e^{-it\hat{V}(\psi)}$ provides the *exact* time evolution of the pure nonlinear equation (assuming $\hat{H}_0 = 0$):

$$\frac{\partial\psi_n(t)}{\partial t} = -i\beta|\psi_n(t)|^2\,\psi_n(t) \quad \Rightarrow$$
$$\psi(t) = e^{-it\beta|\psi(0)|^2}\,\psi_n(0)$$

which can be easily verified by writing $\psi_n = r_n\,e^{-i\theta_n}$ such that $\dot{r}_n = 0 \quad \Rightarrow \quad r_n(t) = r_n(0) = \text{const.}$ and $\dot{\theta} = \beta r_n^2 \quad \Rightarrow \quad \theta(t) = \theta(0) + t\beta r_n^2(0)$. The conservation of $|\psi_n(t)| = \text{const.}$ (for the pure nonlinear equation) is a feature of the particular form of the nonlinear term and due to this $V(\psi)$ does not depend on ψ nor on t (during the purely nonlinear time evolution) and the time evolution due to the quantum exponential of $-it\hat{V}(\psi)$ coincides exactly with the time evolution of the exponential of the classical Liouville operator associated to the nonlinear term.

In the numerical implementation, we choose a certain initial condition of the state in representation of the eigenbasis of $\hat{H}_0$, e.g. $C_m(0) = \delta_{m,m_0}$ with m_0 being the index of the initial state. Then, we apply the first exponential factor with coefficient d_1 (and given value of Δt) which corresponds to $e^{-iE_m d_1 \Delta t}\,C_m \to C_m$. Then, using the unitary matrix that diagonalizes $\hat{H}_0$, we transform $C_m \to \psi_n$ and we apply the exponential factor with c_2 (since $c_1 = 0$ if $p = 2$ or $p = 4$) which corresponds to $e^{-ic_2\Delta t\beta|\psi_n|^2}\,\psi_n \to \psi_n$ which represents *exactly* the purely nonlinear time evolution. Then we transform $\psi_n \to C_m$ and apply the next exponential factor with coefficient d_2 etc. (If one uses a non-symmetric variant, with $c_1 \neq 0$, for $p = 1$ or $p = 3$ one has first to transform the initial condition to ψ_n, apply the first c_1-factor and transform back to C_m.)

We have implemented and tested all four variants of the method. In particular, we have verified that the classical energy is conserved, i.e. its residual numerical fluctuations ($\sim 10^{-8}$ for the fourth order variant at $\Delta t = 0.1$) scale with $(\Delta t)^p$ and also that the errors of other quantities scale with $(\Delta t)^p$. For the case of a RMT with an extra diagonal where the values E_m become larger, we have also tested the precision by comparing some data with $\Delta t = 0.0125$ (for reduced iteration times) which does not change the values of S etc. (apart from statistical fluctuations).

III. ADDITIONAL FIGURES

In this section, we present additional Figures for the main part of this article. Some of the following captions also contain physical discussions or additional information for figures in the main part; in particular the values of T and μ for both approaches and the four states shown in Fig. 3 are given in the caption of Fig. S4 below.

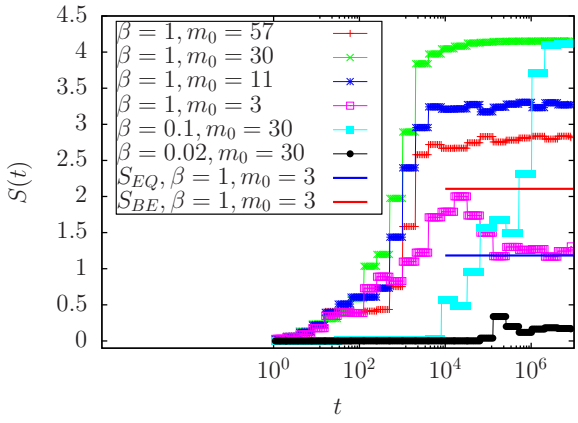

FIG. S1: Time dependence of the entropy $S(t)$ for the four states shown in Fig. 3 and two other states at $\beta = 0.02, 0.1$ with initial index $m_0 = 30$ (for $\beta = 0.02, 0.1$ the other three index values $m_0 = 3, 11, 57$ correspond to very small entropy $S(t)$ values clearly below the case $m_0 = 30$). The entropy is computed from $S(t) = -\sum_m \rho_m \ln(\rho_m)$ where ρ_m is obtained as the time average $\rho_m = \langle|C_m(t)|^2\rangle$ for successive time intervals with increasing lengths by a factor of two corresponding to the plateau intervals of constant $S(t)$ visible in the figure. The thick horizontal lines represent the theoretical entropy S_{EQ} for EQ (blue) and S_{BE} for BE (red) for the energy of the state at $\beta = 1$ and $m_0 = 3$ (pink open squares). At intermediate times $t \approx 2 \times 10^4$ the entropy of this state is close to S_{BE} while at longer times $t \geq 10^6$ it decreases to S_{EQ} showing that the EQ ansatz describes the correct long time thermalization but also that at intermediate times the entropy is larger and closer to the BE ansatz. The other states $m_0 = 11, 30, 57$ at $\beta = 1$ thermalize rather quickly at $t \geq 10^4$-10^5 to their final value S_{EQ} (with $S_{BE} \approx S_{EQ}$ for $m_0 = 30$). For $\beta = 0.1$ the state $m_0 = 30$ (cyan full squares) thermalizes to the same entropy value as with $\beta = 1$ (green crosses) but only for very long time scales $t \geq 10^6$.

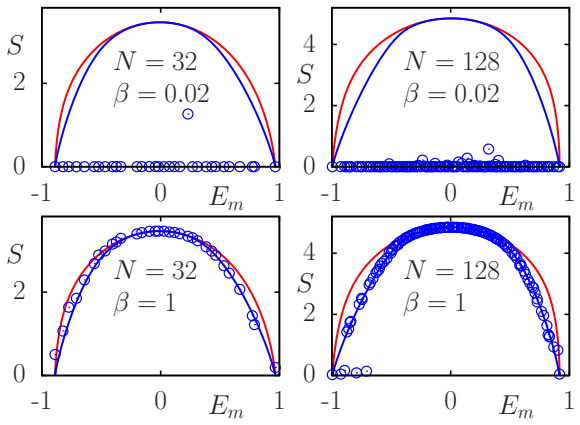

FIG. S2: As Fig. 1 for one RMT realisation, $\beta = 0.02, 1$, $N = 32, 128$, ρ_m obtained by the time average for $2^{21} \le t \le 2^{22}$ for $N = 128$ and $2^{26} \le t \le 2^{27}$ for $N = 32$ (blue $\circ$, all panels). The curves represent the theoretical $S(E)$-curves from the EQ (blue) and BE (red) approaches using the exact spectrum of the used RMT realisation. The data point with $S > 1$ for $N = 32$, $\beta = 0.02$ is not saturated and still increasing at the given maximal time $t = 2^{27}$. The data for $\beta = 2$ and $N = 128$ coincide very well with the EQ ansatz. Also for $N = 32$ the EQ ansatz is more appropriate. Here the small differences to the theoretical EQ-curve are due to the fact that on the x-axis the initial energy E_m is used and not the averaged linear energy $\langle E\rangle = \sum_{m'} E_{m'}\rho_{m'}$ using the long time average $\rho_{m'}$ and which is slightly different from E_m due to the nonlinear term. Using $\langle E\rangle$ the data points (for the cases with good thermalization) fall nearly exactly on the theoretical curve.

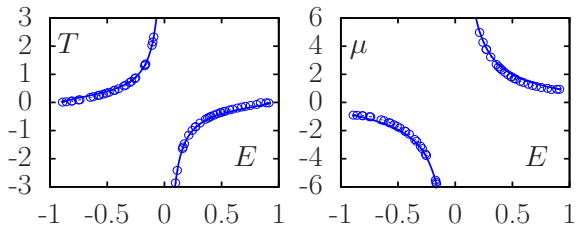

FIG. S3: As Fig. 2 but for the BE case. Here the data points for T and μ, computed from the numerical data of S, show some small deviations from the theoretical curves which are visible in the figure and significantly larger than in Fig. 2 where no deviations for the EQ case (on graphical precision) are visible. Furthermore, in comparison to the EQ case of Fig. 2 the typical values of T and μ for the BE case are considerably larger.

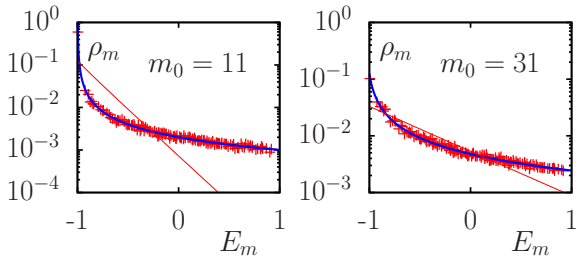

FIG. S4: Dependence of ρ_m on E_m for two states with initial state $m_0 = 11, 31$ for $\beta = 1$, $N = 128$ obtained by an time average in the interval $2^{21} \le t \le 2^{22}$. The blue curve shows the theoretical values based on EQ with $\rho_{\rm EQ}(E) = T/(E-\mu)$ and $T = 0.002007, 0.004986$, $\mu = -0.9995, -1.041$ for $m_0 = 11, 31$. T and μ were determined from the solution of the equations $1 = \sum_m \rho_{\rm EQ}(E_m)$ and $\langle E\rangle = \sum_m E_m\rho_{\rm EQ}(E_m)$ with $\langle E\rangle = \sum_m E_m\rho_m \approx E_{m_0}$. The red line shows the theoretical values based on BE with $\rho_{\rm BE}(E) = 1/(\exp[(E-\mu)/T] - 1)$, $T = 0.1983, 0.5562$, $\mu = -1.434, -2.914$ for $m_0 = 11, 31$ and T, μ determined from the solution of the equations $1 = \sum_m \rho_{\rm BE}(E_m)$ and $\langle E\rangle = \sum_m E_m\rho_{\rm BE}(E_m)$. The thermalization of both states according to the EQ theory is very good despite the shorter averaging time as compared to Fig. 3 indicating that for larger values of N the thermalization time scale is reduced. *Complementary information for* Fig. 3: The T and μ values for the EQ ansatz and the four states $m_0 = 3, 11, 30, 56$ with $N = 64$ and $\beta = 1$ of Fig. 3 are $T = 0.001372, 0.005984, 0.07585, -0.004538$, $\mu = -0.8964, -0.9178, -4.892, 0.9293$ and the corresponding values for the sames states and the BE ansatz are $T = 0.112, 0.3581, 4.913, -0.2649$, $\mu = -1.062$, -1.794, -20.52, 1.496.

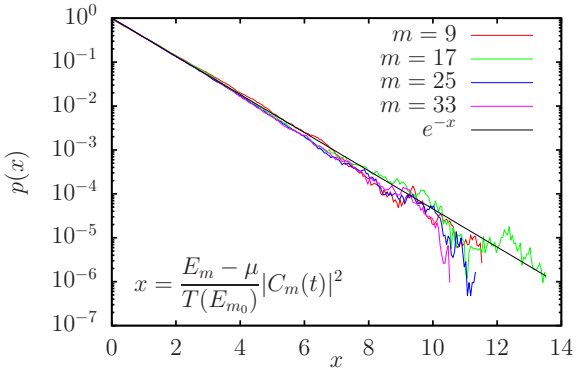

FIG. S5: Statistical distribution of the rescaled variable $x = (E_m - \mu)|C_m(t)|^2/T(E_{m_0})$ for $\beta = 1$, $N = 64$, $m_0 = 9$, $m = 9, 17, 25, 33$ using the time values in the interval $2^{23} \leq t \leq 2^{24}$ and a histogram of bin width 0.05. The thin black line shows the theoretical distribution $p(x) = e^{-x}$ according to the EQ approach. The numerical distributions follow the theoretical distribution for values up to $x \approx$ 8-10 providing an additional confirmation for the validity of the EQ ansatz.

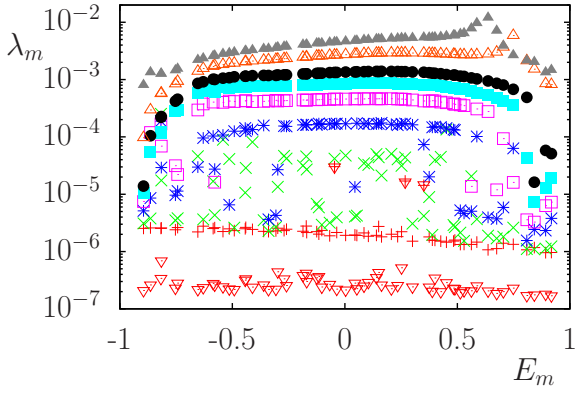

FIG. S6: As Fig. 5 but with additional β values : Lyapunov exponent λ_m dependence on E_m with m being the index of the initial state for $N = 64$. λ_m has been determined from the fit $\ln \|\Delta\psi(t)\| = a + b\ln(t) + \lambda_m t$ for $t \leq 2^{22}$ and $\beta = 2$ (grey ▲; top), $\beta = 1.5$ (orange △), $\beta = 1$ (black •), $\beta = 0.75$ (cyan ■), $\beta = 0.5$ (pink □), $\beta = 0.25$ (blue ∗), $\beta = 0.1$ (green ×), $\beta = 0.02$ (red +), $\beta = 0.02$ for $t \leq 2^{27}$ (red ▽; bottom). The numerical data suggests that most λ_m for $\beta = 0.02$ decay as $\lambda_m \sim 1/\sqrt{t}$ for $t \geq 10^7$ (see Fig. S10 below). However, three λ_m values for $\beta = 0.02$ do not decay with time (data points with same red + and ▽; e.g. $m = 45$ and $E_m \approx 0.35$) and have significantly larger values $\lambda_m > 10^{-5}$ indicating a trajectory in a chaotic region while other initial conditions correspond to trajectories in bounded KAM regions. These cases are also visible in Fig. 1 (a) with entropy values slightly above 0.

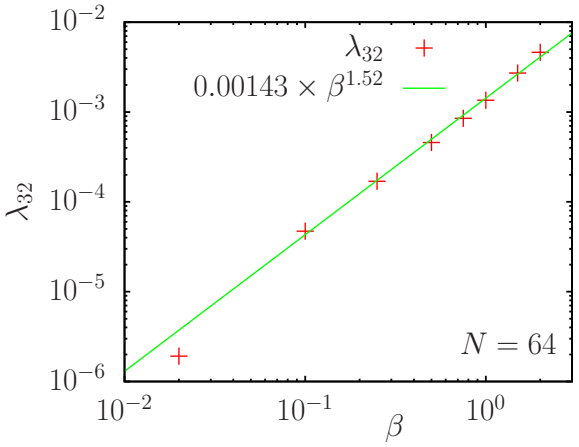

FIG. S7: Dependence of Lyapunov exponent λ_{32} in the band center on β for $N = 64$ and computation time $t = 2^{22}$. The straight green line shows the power law fit $\lambda_{32} = a\beta^{\eta}$ with $a = 0.00143 \pm 0.00005$ and $\eta = 1.52 \pm 0.03$. For this fit the smallest data point at $\beta = 0.02$ was not used since for this value the Lyapunov exponent continues to decrease with increasing computation time t and it is most likely below the chaos border.

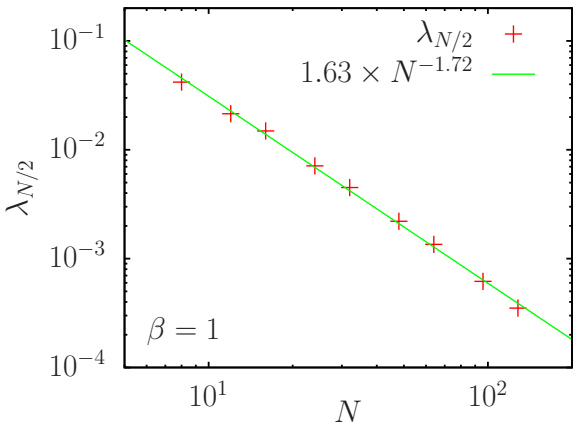

FIG. S8: Dependence of Lyapunov exponent $\lambda_{N/2}$ in the band center on N for $\beta = 1$ and computation time $t = 2^{22}$. The straight green line shows the power law fit $\lambda_{N/2} = aN^{-\nu}$ with $a = 1.63 \pm 0.16$ and $\nu = 1.72 \pm 0.03$.

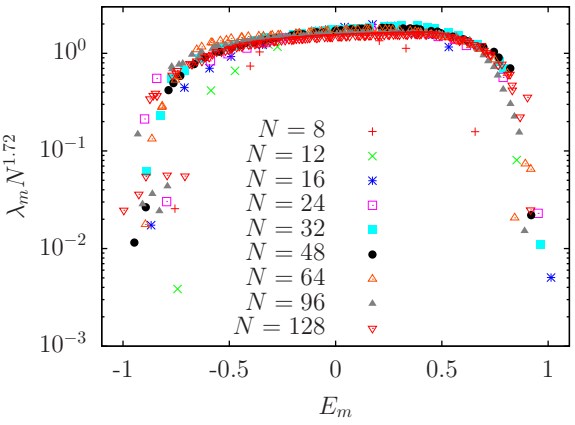

FIG. S9: Dependence of rescaled Lyapunov exponent $\lambda_m N^\nu$ on the initial energy E_m using the exponent $\nu \approx 1.72$ found in the fit of Fig. S8 for different values $8 \leq N \leq 128$.

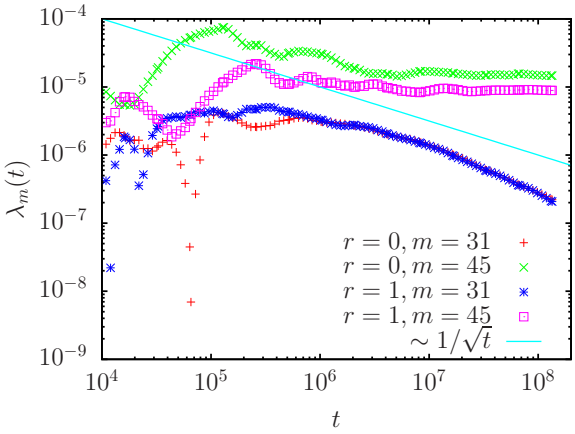

FIG. S10: Dependence of Lyapunov exponent $\lambda_m(t)$ on iteration time t for the case $\beta = 0.02$, $N = 64$, two initial states in the band center, $m = 31, 45$, and two RMT realisations $r = 0$ (same realisation as for most main and SupMat figures concerning the RMT case) and $r = 1$. The cyan full line shows $10^{-2}/\sqrt{t}$ to indicate a $t^{-1/2}$ power law. $\lambda_m(t)$ has been obtained by the fit $\ln \|\Delta\psi(\tau)\| = a + b \ln(\tau) + \lambda_m \tau$ for $0 \leq \tau \leq t$ and for values $t \leq 2^{27}$ where $\Delta\psi(\tau) = \tilde{\psi}(\tau) - \psi(\tau)$ is the difference vector between two close initial conditions with $\psi_n(0) = \phi_m(n)$, $\tilde{\psi}(0) = \psi(0) + \Delta\psi(0)$ and $\Delta\psi(0)$ being a random vector with initial norm $\|\Delta\psi(0)\| = 10^{-12}$. During the computation the difference vector $\Delta\psi(\tau)$ is regularly renormalized to the norm 10^{-12} when its norm has become larger than 10^{-10} such that both trajectories stay close and the logarithm of the renormalization factor is added to a special variable measuring the quantity $\|\Delta\psi(\tau)\|$ which is used for the computation of the Lyapunov exponent. The rescaled logarithmic growth $(\ln \|\Delta\psi(t)\|)/t$ shows roughly the same behavior as $\lambda_m(t)$, with a final slope somewhat closer to the exponent $-1/2$ than for $\lambda_m(t)$ (in logarithmic representation and for $t \geq 10^7$). *Discussion:* The two cases at $m = 31$, $r = 0, 1$ indicate a vanishing Lyapunov exponent in the limit $t \to \infty$ and a trajectory in a bounded KAM region. The Lyapunov exponent for the other two cases at $m = 45$ (with $E_{45}(r = 0) \approx 0.351$ and $E_{45}(r = 1) \approx 0.310$) saturate to the values $\lambda_{45}(r = 0) \approx 1.47 \times 10^{-5}$ and $\lambda_{45}(r = 1) \approx 8.89 \times 10^{-6}$ in the limit $t \to \infty$ indicating a trajectory in a chaotic region probably due to the effect of a near triple quasi-resonance for the given RMT realisation. For the first realisation $r = 0$ there are three cases like this as can be seen in Figs. 5 and S6 (see also caption therein). The observation that for both realisations there are saturated Lyapunov values at the same index $m = 45$ is a coincidence and for example for $m = 32$ (not shown in the figure) there is a saturated Lyapunov exponent for $r = 1$ but not for $r = 0$.

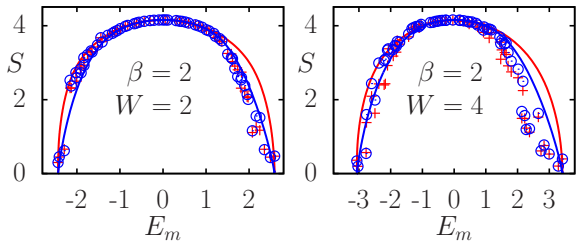

FIG. S11: As Fig. 1 but for one realisation of the DANSE model of [29] at disorder strength $W = 2$ and $W = 4$ for $\beta = 2$ and $N = 64$. The data points correspond to the averaging time $2^{23} \leq t \leq 2^{24}$ (blue $\circ$) and $2^{20} \leq t \leq 2^{21}$ (red $+$; similar t values as in [29]).

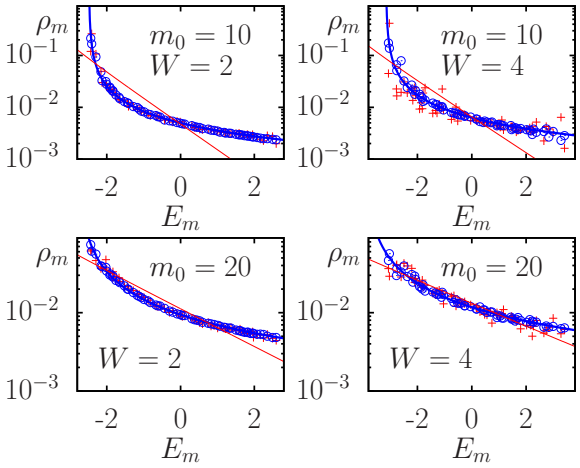

FIG. S12: As Figs. 3 and S4 but for one realisation of the DANSE model of [29] at disorder strengths $W = 2$ and $W = 4$ for $\beta = 2$, $N = 64$ and two initial states with $m_0 = 10, 20$. The data points correspond to ρ_m obtained by the averaging time $2^{23} \leq t \leq 2^{24}$ (blue $\circ$) and $2^{20} \leq t \leq 2^{21}$ (red $+$; similar t values as in [29]). The values of T and μ for the EQ approach are $T = 0.0124, 0.02626, 0.01997, 0.04636$ and $\mu = -2.484, -2.784, -3.157, -3.952$ for $W = 2$ with $m_0 = 10, 20$ and $W = 4$ with $m_0 = 10, 20$. The values of T and μ for the BE approach are $T = 0.8794, 1.815, 1.26, 3.011$ and $\mu = -4.709, -8.171, -6.346, -13.07$ for the same states. For $W = 2$ both states are well thermalized according to the EQ case. For $W = 4$ the thermalization also corresponds to the EQ case but there are still stronger fluctuations, especially for the data with shorter averaging time (and corresponding to the data of [29]).

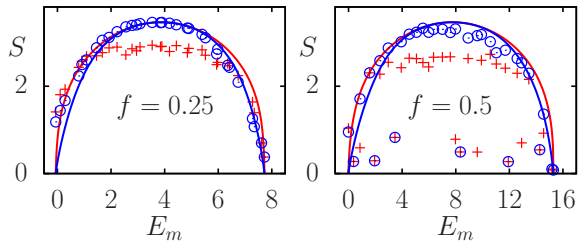

FIG. S13: As Fig. 1 for the case of an RMT plus extra diagonal matrix elements fn with parameter $f = 0.25$ or $f = 0.5$ for $\beta = 1$ and $N = 32$. The data points correspond to the averaging time $2^{26} \leq t \leq 2^{27}$ (blue $\circ$ for $f = 0.25, 0.5$), $2^{19} \leq t \leq 2^{20}$ (red $+$ for $f = 0.5$) and $2^{14} \leq t \leq 2^{15}$ (red $+$ for $f = 0.25$).

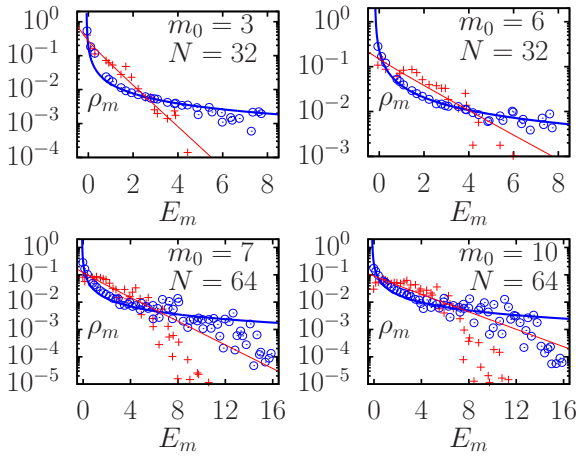

FIG. S14: As Fig. 3 for the case of an RMT plus extra diagonal matrix elements fn with parameter $f = 0.25$ for $\beta = 1$, $m_0 = 3, 6$ ($N = 32$) or $m_0 = 7, 10$ ($N = 64$). The data points correspond to ρ_m obtained by the averaging time $2^{26} \leq t \leq 2^{27}$ (blue $\circ$), $2^{19} \leq t \leq 2^{20}$ (red $+$; for $N = 64$) and $2^{14} \leq t \leq 2^{15}$ (red $+$; for $N = 32$). At longer times $t = 2^{27}$ the states are (quite) well thermalized according to the EQ case (with somewhat stronger flucutations for $N = 64$). However, at the intermediate time scale the values of ρ_m are closer to the BE line thus explaining that the corresponding entropy values are also closer to the BE curve.