Supplemental Material for "Thermalization slowing-down in multidimensional Josephson junction networks"

Gabriel M. Lando and Sergej Flach
Center for Theoretical Physics of Complex Systems, Institute for Basic Science, Daejeon 34126, Korea

The objective of this supplement is to provide a comprehensive account of all numerical and analytical details omitted from the main text, together with stringent tests for the convergence of our results. In the following we use the conventions adopted in the main text, namely that H and Φ are the Hamiltonian and its flow and $\mathbf{z} = (\mathbf{p}, \mathbf{q})$ denotes a $2N$-dimensional phase space vector. 1-, 2- and 3-dimensional networks are denoted by writing their number of sites as a power of dimensionality, and here the $N = 12^2$ network is used for the most delicate numerical tests. We will also refer to rescaled Lyapunov spectra as LS, dropping the "rescaled".

I. RESCALING TRANSFORMATIONS

The relevant control parameter of our system is E_J/h. Let $\mathbf{P}(t)$ and $\mathbf{Q}(t)$ be the projections of the flow for some value of E_J/h and initial condition $(\mathbf{p}, \mathbf{q})$ into momentum and position subspaces, *i.e.* $\mathbf{P} = \pi_1 \circ \Phi$, $\mathbf{Q} = \pi_2 \circ \Phi$, where $\pi_1(\mathbf{p}, \mathbf{q}) = \mathbf{p}$ and $\pi_2(\mathbf{p}, \mathbf{q}) = \mathbf{q}$ are the canonical projections. Any $\mathbf{Q}(t)$, $\mathbf{P}(t)$ can be transformed into $\tilde{\mathbf{Q}}(\tau) = \mathbf{Q}(\mu t)$ and $\tilde{\mathbf{P}}(\tau) = \mathbf{P}(\mu t)/\mu$, $\tau = \mu t$ and $\tilde{E}_J = E_J/\mu^2$ for an arbitrary real and positive choice of μ. The transformed trajectory evolves at an energy density $\tilde{h} = h/\mu^2$ such that $\tilde{E}_J/\tilde{h} = E_J/h$. For any practical purposes we perform computations for $E_J \leq h$ at $h = 1$ and running time t. Instead for $E_J > h$ we fix $E_J = 1$ and rescale time accordingly. This is done in order to avoid the numerical issues of dealing with large h or E_J values, which cause either the kinetic of potential terms to blow up. The above transformations avoid this problem by allowing us to deal exclusively with decreasing either E_J or h while holding the other parameter fixed. For the LS, for example, we connect the different regimes with

$$\Lambda_{E_J}(z) = \frac{\Lambda_h(z)}{\sqrt{h}} \quad , \tag{1}$$

where the subscript denotes which quantity is being varied.

II. NETWORKS

In the LRN regime $h/E_J \ll 1$ we can assume $E_J = 1$ and $|p|, |q| \ll 1$. At the integrable limit we expand the cosine interaction potential to 2nd order, $\cos(x) \approx 1 - x^2/2$. In proximity to that limit the leading order nonintegrable perturbation results from the next quartic term in the Taylor expansion of the potential, namely $\cos(x) \approx 1 - x^2/2 + x^4/24$. For simplicity we stay in dimension 1 (extension to higher dimensions is straightforward). The resulting Hamiltonian is a variant of the celebrated β-Fermi-Pasta-Ulam-Tsingou model

$$H = \sum_l \frac{p_l^2}{2} + \frac{E_J}{2}(q_l - q_{l-1})^2 - \frac{E_J}{24}(q_l - q_{l-1})^4 \ . \tag{2}$$

We use the canonical transformation to normal mode momenta and coordinates $\{P_k, Q_k\}$

$$\begin{pmatrix} P_k \\ Q_k \end{pmatrix} = \sqrt{\frac{2}{N+1}} \sum_{n=1}^{N} \begin{pmatrix} p_n \\ q_n \end{pmatrix} \sin\left(\frac{\pi n k}{N+1}\right) \tag{3}$$

for $k = 1, \ldots, N$. This transformation diagonalizes the integrable quadratic Hamiltonian part $H_0 = \sum_{k=1}^{N} E_k$ in Eq.(**??**), where the normal mode energies E_k are

$$E_k = \frac{P_k^2 + \Omega_k^2 Q_k^2}{2} \ , \qquad \Omega_k = 2\sqrt{E_J} \sin\left(\frac{\pi k}{2(N+1)}\right) \ . \tag{4}$$

The equations of motion in the normal mode coordinates (**??**) then read

$$\ddot{Q}_k + \Omega_k^2 Q_k = \frac{E_J}{12(N+1)} \sum_{l_1,l_2,l_3} \Omega_k \Omega_{l_1} \Omega_{l_2} \Omega_{l_3} A_{k,l_1,l_2,l_3} \times Q_{l_1} Q_{l_2} Q_{l_3} \tag{5}$$

where

$$\begin{aligned} A_{k,l_1,l_2,l_3} = \ & \delta_{k-l_1+l_2-l_3,0} + \delta_{k-l_1-l_2+l_3,0} \\ & - \delta_{k+l_1+l_2-l_3,0} - \delta_{k+l_1-l_2+l_3,0} \end{aligned} \tag{6}$$

represents the coupling between the Fourier coordinates Q_k. Using the canonical transformation

$$Q_k = \sqrt{2J_k} \sin\theta_k \qquad P_k = \Omega_k \sqrt{2J_k} \cos\theta_k \tag{7}$$

it follows that

$$\dot{J}_k = \sum_{l_1,l_2,l_3} \mathcal{A}_{k,l_1,l_2,l_3} \sqrt{J_k J_{l_1} J_{l_2} J_{l_3}} \tag{8}$$

where the coefficients $\mathcal{A}_{k,l_1,l_2,l_3}$ depend on the angles $\{\theta_k\}_k$:

$$\mathcal{A}_{k,l_1,l_2,l_3} = \frac{E_J}{6} \frac{A_{k,l_1,l_2,l_3}}{2(N+1)} \Omega_{l_1} \Omega_{l_2} \Omega_{l_3} \cos\theta_k \sin\theta_{l_1} \sin\theta_{l_2} \sin\theta_{l_3} \ . \tag{9}$$

Therefore each action J_k is interacting with all other actions via the proliferating number of $\sim N^2$ nonlinear terms on the RHS of (**??**). This constitutes the LRN imposed by the nonintegrable perturbation onto the set of actions, and the number of terms on the RHS readily generalizes to $N^2 d$ for higher dimensions.

In the SRN regime $E_J/h \ll 1$ the integrable limit model is the case of free rotors $E_J = 0$. The coordinates $p_l = J_l$ and $q_l = \theta_l$ are already the proper action-angle choice. The nonintegrable perturbation is the entire Josephson coupling potential term in the original Hamiltonian, therefore imposing nearest neighbour couplings between the actions. This constitutes the SRN, where the number of actions a given reference action is interacting with is finite, and not scaling with the system size. The corresponding trivial metric allows to introduce a distance between actions distinguishing large from small distances.

III. RESONANCE PROBABILITIES

In the LRN regime a normal mode interacts with N^{2d} additive quadruplets of other modes (see Sec. **??**). The probability p of resonance for one quadruplet is obtained from first-order perturbation theory in the nonintegrable perturbation $E_J(q_{\sigma_1} - q_{\sigma_2})^4$ and is easily evaluated to be $p \sim (h/E_J)^2$. The probability for one mode to be resonant with at least one quadruplet is then obtained from the complementary probability not to be in resonance with any of them, which reads $p_R = 1 - (1-p)^{N^2} \approx 1 - e^{-pN^2}$. For macroscopic systems it follows that $p_R = 1$.

In the SRN regime the first-order perturbation correction to a free rotor dynamics $p^{(0)}_{\sigma_1} = \text{const}$ reads $p^{(1)}_{\sigma_1} \sim E_J/(p^{(0)}_{\sigma_1} - p^{(0)}_{\sigma_2})$. The probability to be in resonance is then $p_R \sim 2d(E_J/h)$. These sparse resonances are at an average distance $p_R^{-1/d}$ from each other, with this distance increasing upon approaching the integrable limit.

IV. SAMPLING OF INITIAL CONDITIONS

For all dimensionalities considered in the main text, the total momentum of the system is conserved. Without loss of generality we sample initial conditions according to the following recipe:

1. Set $\mathbf{q}(0) = 0$;
2. Sample the components of the initial momentum according to a χ_2 distribution for the 1- and 2-dimensional models and χ_3 for the 3-dimensional one;
3. Transform each component of momentum to $p_i(0) \longmapsto p_i(0) - \langle \mathbf{p}(0) \rangle$, where $\langle \mathbf{p}(0) \rangle$ is the mean initial momentum.

χ_s is the distribution with probability density function

$$f(x;s) = \begin{cases} \dfrac{x^{s-1}e^{x^2/2}}{2^{s/2-1}\Gamma(s/2)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} , \qquad (10)$$

from which the Maxwell-Boltzmann distribution is recovered for $s = 3$. The procedure above results in zero initial momentum, which should be approximately conserved by symplectic integration alongside the energy. The energy density of the system is then fixed and equal to

$$h = \frac{1}{N}\sum_{i=1}^{N} \frac{p_i(0)^2}{2} , \qquad (11)$$

since $\mathbf{q}(0) = 0$. Through a further rescaling of the initial momentum as $\mathbf{p}(0) \longmapsto \mathbf{p}(0)/\sqrt{k}$ one is effectively dividing the energy density by k, such that by increasing k while keeping E_J fixed we can access the LRN regime. The SRN regime, on the other hand, is accessed by keeping $k = 1$ and decreasing E_J, as discussed in Sec. **??**.

For dynamics in the tangent space, the initial monodromy matrix to be propagated is chosen as

$$\mathcal{M}(0) = \begin{pmatrix} I & 0 \\ 0 & \widetilde{I} \end{pmatrix} , \quad I_{ij} = \delta_{i,j} , \quad \widetilde{I}_{ij} = \delta_{i,2N-j+1} \quad , \qquad (12)$$

which departs from the established choice of just propagating the identity matrix. Nevertheless, this simple permutation of rows guarantees that the sum of positive and negative exponents is conserved in time and machine-equal to zero – an important property of time-independent Hamiltonian systems [**?**].

V. CONSERVED QUANTITIES

Symplectic integrators (SIs), which are the method we chose to compute the flow in phase space and tangent space, come in a variety of flavours with respect to accuracy and stability [**?** **?**]. Typically, these integrators are chosen to be of high-order, with their large number of iterations per step compensated by choosing big time steps [**?** **?** **?** **?**]. However, the QR-decomposition method used to compute the LS works best with small time steps [**?**], which would render high-order integrators prohibitively slow. We have indeed consistently observed that LS converge more smoothly when employing a low-order integrator with a smaller time step than a high-order one with a large time step. This favors the choice of a low-order integrator, despite the increase in computational cost. In order to avoid triggering numerical degeneracies in $\mathcal{M}$ as it evolves in time, we have also chosen to apply the QR-decomposition method at every time step instead of waiting, which slows down computations but is necessary due to several LEs in the spectra being extremely small and easily lost [**?** **?**].

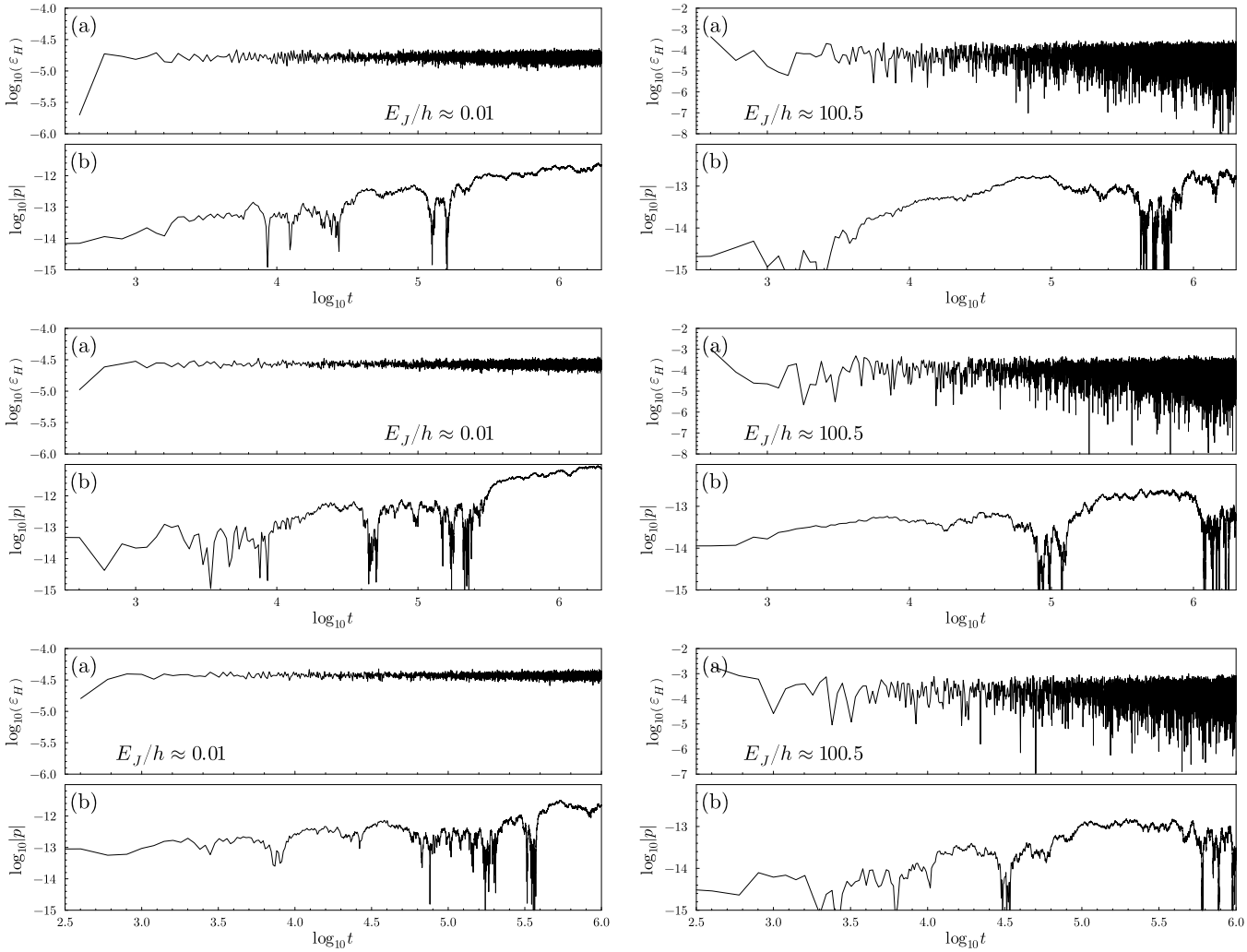

FIG. 1. (Left panels) Relative energy (a) and total momentum (b) errors for $N = 140$ (upper panels), $N = 12^2$ (middle panels) and $N = 5^3$ (bottom panels) in the SRN regime. The final times are the same as the ones used to calculate the spectra. All errors remain bounded below 10^{-4}, but become more stable as the dimensionality is increased. (Right panels) Same but for the LRN regime.

Nevertheless, SIs do not exactly preserve conserved quantities, and their accuracy is usually measured in terms of "drifts" in things algebraically known to be constant. In strongly chaotic systems like the ones we deal with, this is in fact the only way one can check the quality of an integrator, since the trajectories for one particular choice of dt will be completely different than for another due to exponential sensitivity – thus, trajectory errors are essentially useless. Note, however, that LS computed with different dts should be very close to each other due to the shadowing lemma [**?**] (see Sec. **??**).

In the left panels of Fig. **??** we display the drifts (*i.e.* the errors) in the deep SRN regime for all dimensionalities discussed in the main paper, and three choices of N (the number of sites). The [relative] energy drift is defined as

$$\varepsilon_H = \left| 1 - \frac{(H \circ \Phi)(\mathbf{z}; t)}{H(\mathbf{z})} \right| \quad , \tag{13}$$

while the momentum error is just the norm of the momentum vector, which due to our initial conditions should be conserved and equal to zero. Drifts are shown as a function of time, starting from 10^2 up to the final times used for computing the LS. The same is done in the right panel of Fig. **??**, but this time for the LRN regime. Note that the errors are larger in LRN than in SRN, being bounded by 10^{-3} and 10^{-4}, respectively, and also less stable for the former than the latter. This is an indication that the LRN regime is numerically harder to deal with than the SRN one, which can also be noted from the fact that their LEs are much smaller than their SRN counterparts.

We only display the errors near integrability because in

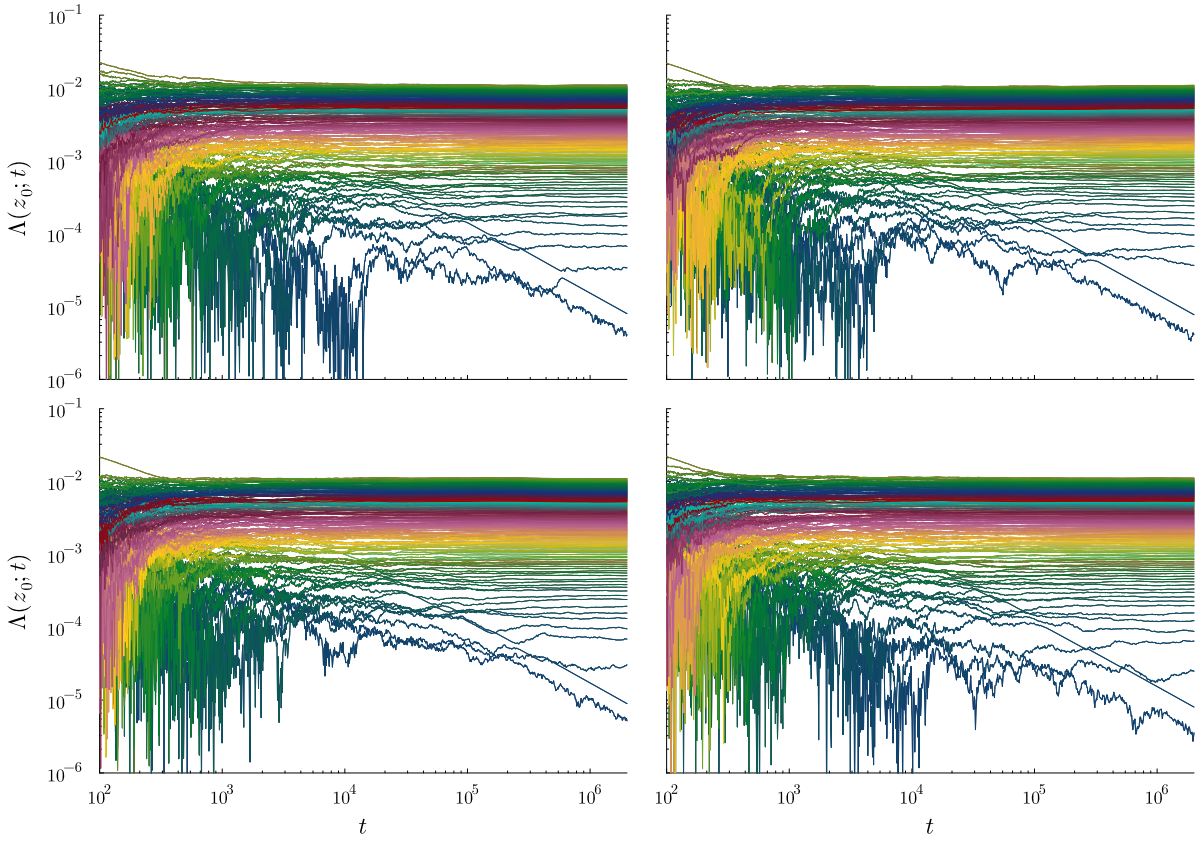

FIG. 2. Unrescaled LS for the $N = 12^2$ network using four different initial conditions, z_0, as a function of time, with $t_f = 2\times10^6$. Here, we are far from both the LRN and the SRN regimes, with $\log_{10}(E_J/h) \approx 0.5$.

the intermediate regime, say $|\log_{10}(E_J/h)| < 0.5$, achieving convergence in the spectra is very easy and, in principle, one does not even need to go to very long times – although, for consistency, we did use the same final times to cover the whole E_J/h axis.

VI. CONVERGENCE OF LS

We now move on to quantifying the errors in the LS as a function of time. The easiest regime to deal with is surely $|\log_{10}(E_J/h)| < 0.5$, as stated earlier. As as example, in Fig. **??** we display the time-dependent LS for 4 different initial conditions, obtained following the sampling procedure described in Sec. **??**, using the $N = 12^2$ system as an example. Since there are 144 LEs in the LS, it is essentially impossible to see them simultaneously, such that we color them based on their values to facilitate visualization. This figure shows several important trends:

1. We are able to resolve the whole spectrum, including the two zeros, which are the two lowest LEs decaying as $1/t$;
2. All runs produce essentially the same final spectrum at $t = t_f$;
3. Almost all LEs, including the small non-zero ones, have remained stable since $t \approx 10^5$.

In the above, item (1) is a proof that we have achieved thermalization/equipartition in our numerical simulations for fixed initial conditions, but (2) shows that all initial conditions lead to the same result – *i.e.* we have achieved true numerical ergodicity. Point (3) shows that LS are properly converged, and one must not be fooled by the logarithm scales in the time axis: Convergence between 10^5 and 2×10^6 means that the LS were stable for 95% of the simulation time.

What we see in Fig. **??** is the perfect scenario one can achieve when computing LS. As we approach an integrable limit, the situation will be different and the LS will display both finite-time and finite-size effects. In Fig. **??** we present the equivalent of Fig. **??**, but now deep in the LRN and SRN regimes with $\log_{10}(E_J/h) \approx 2.2$ and $\log_{10}(E_J/h) \approx -2.2$, respectively. In the LRN it is clear we were able to resolve the whole finite component of the LS, although to resolve the zeros we would need even longer times. Nevertheless, the finite LEs have clearly remained stable for the same interval as the ones in Fig. **??**, and all runs show very little dependence on

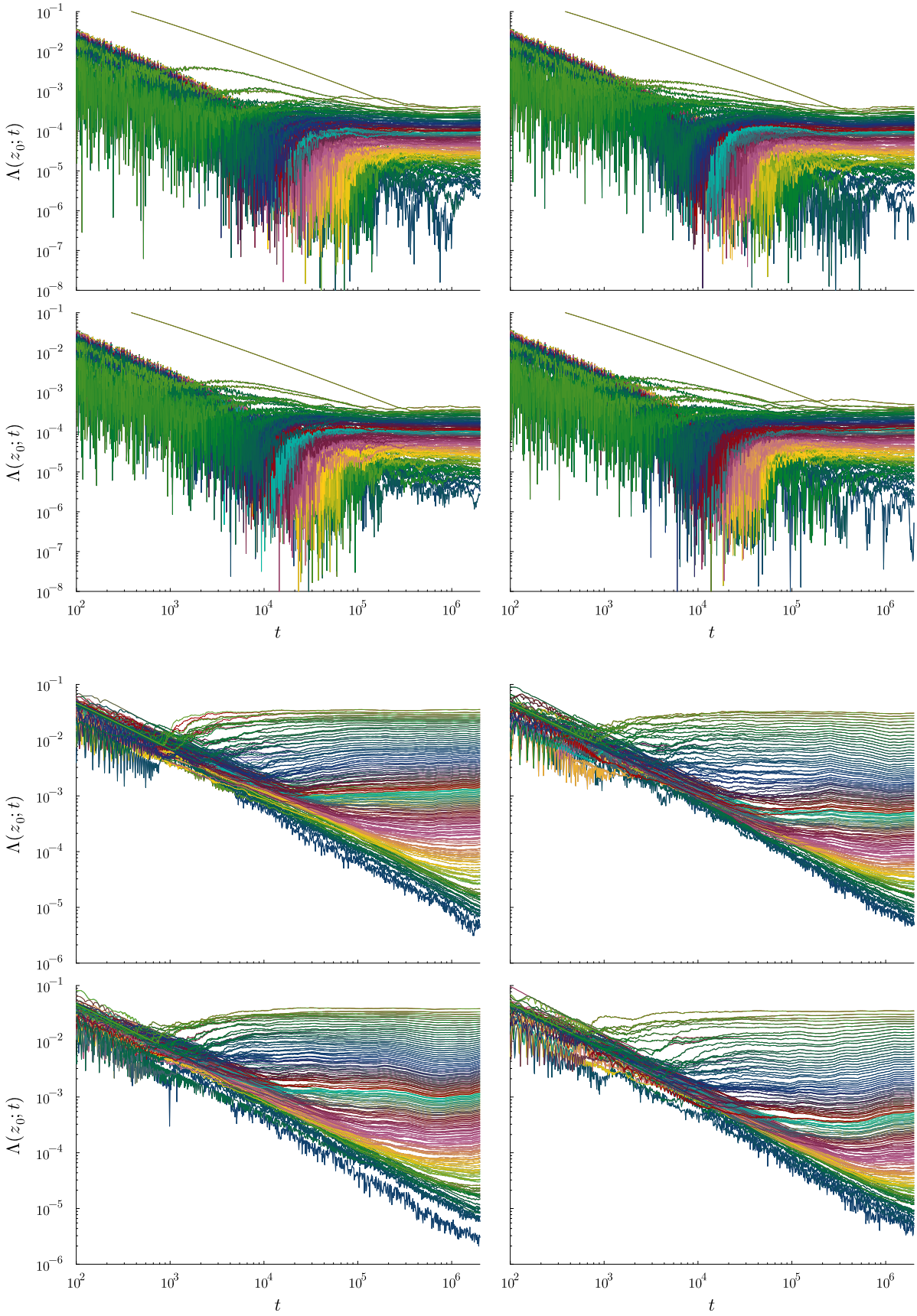

FIG. 3. Same as Fig. **??**, but now in the deep LRN regime with $\log_{10}(E_J/h) \approx 2.2$ (top 4 panels) and in the deep SRN regime with $\log_{10}(E_J/h) \approx -2.2$ (bottom 4 panels).

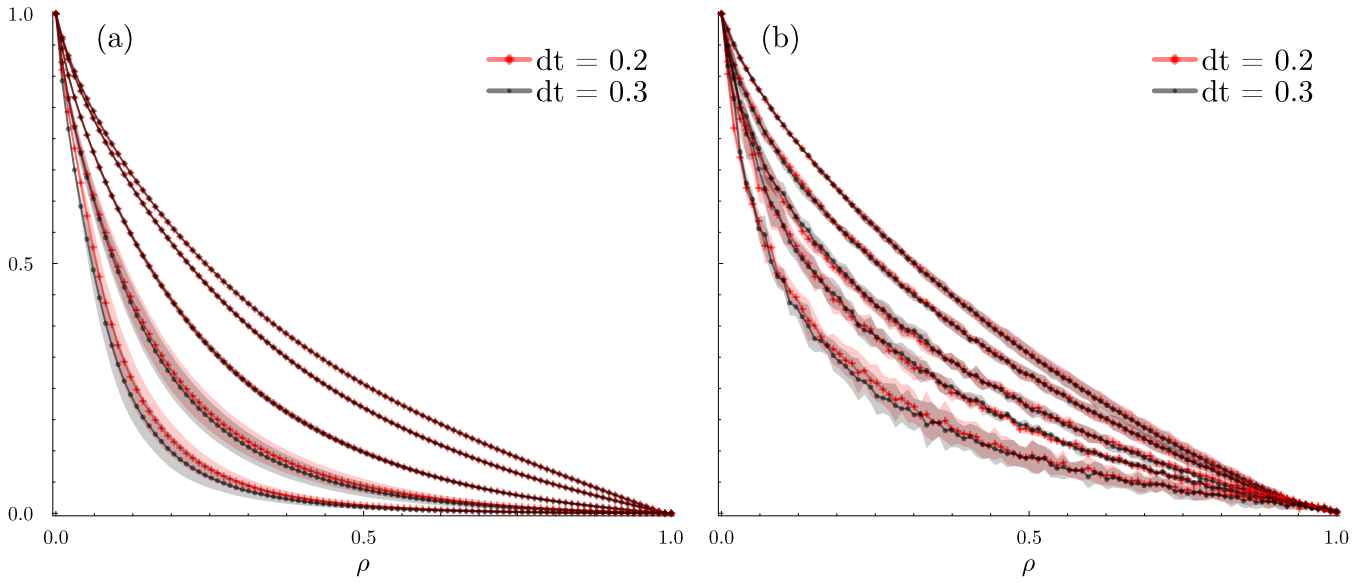

FIG. 4. Qualitative comparison between LS obtained with two different step sizes for $N = 12^2$. Panel (a) shows LS in the SRN regime, while (b) is similar but in the LRN regime.

the initial condition. In the SRN case, however, one can see that there are visible differences between the different runs due to the some weak dependence on initial conditions, which is a manifestation of finite size effects. It is also clear that we are not able to resolve the full finite component of the LS, since we know that only two exponents truly decay as $1/t$, while in the figures we can see that a handful of them are still behaving as numerical zeros and have not been resolved. We consider that being able to resolve around 97% of the LEs is a remarkable achievement in terms of numerical and algorithmic stability, and the exponents that we are unable to catch are essentially invisible in the final spectrum – they are around ten thousand times smaller than the mLE.

Resolving the full LS becomes a much harder task in the 1d systems due to the reasons discussed in the main paper, such that we are able to resolve around 30 or 40 out of 100 exponents for the $N = 100$ case. However, almost all of the finite LEs in the deep SRN regime are several orders of magnitude smaller than the mLE, such that the resolution we achieve is more than enough to characterize the 1d system as well. As a further confirmation, we also employ two system sizes in all of our computations, such that very strict numerical convergence is required given that the results behave as simarly as seen in Figs. 2 and 3 in the main paper.

Now that we have shown how the LS behave as a function of time and initial condition, we can define how we quantify errors in the spectra. For each initial condition, we define the final LS corresponding to it as being the time-average of the time-dependent LS for half the simulation time. Then, the final LS is given by the average of all LS computed with the same parameters, but different initial conditions, and the error is defined as their standard deviation. This is done for every value of E_J/h. We run 7 different initial conditions for the 1d systems because they are less stable, but for the 2d and 3d systems 3 or 4 runs are enough for accurate error estimates. The time needed to obtain a single LS varies between three and fifteen days in core-hours. From then on, we perform all the procedures (spectral fits, critical exponent fits, etc) using the mean spectra. The uncertainty in the critical exponents is obtained as the margin error at 5% significance [**?**], and the two critical exponents obtained for different system sizes agree within their error bounds.

VII. NUMERICAL ROBUSTNESS

Stiffness is a fundamental effect to be taken into account when employing numerical methods to solve ordinary differential equations. Loosely speaking, a system is stiff if the time-step required to enforce error boundedness tends to zero, despite the solution curve being well-behaved (*i.e.* smooth or C^r for large r) [**?**]. Although a precise definition of stiffness is lacking, its impact is easily identifiable when employing adaptive step-size methods due to anomalous behavior in the step sizes. Naturally, this behavior will quantitatively depend on the step-size controller, but stiff methods are usually assumed to trigger anomalies independently of how the step sizes as chosen.

The Josephson-junction networks considered here are very likely non-stiff systems, although technically this could only be pinpointed by employing an adaptive-step method and analyzing step-size behavior for a given tolerance. The well-behaved drifts in Fig. **??** alone are not enough to determine numerical robustness when employ-

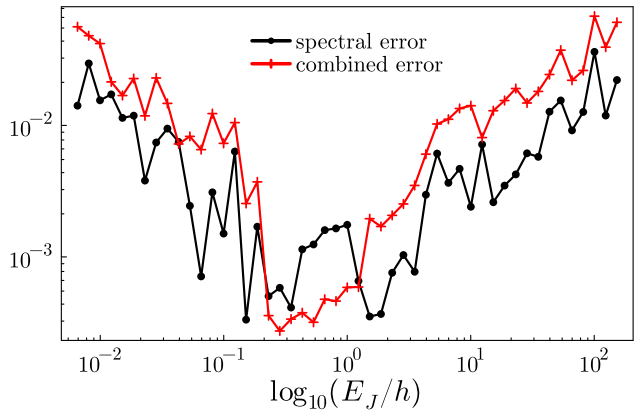

FIG. 5. Spectral error (**??**) versus combined error (**??**) as a function of E_J/h.

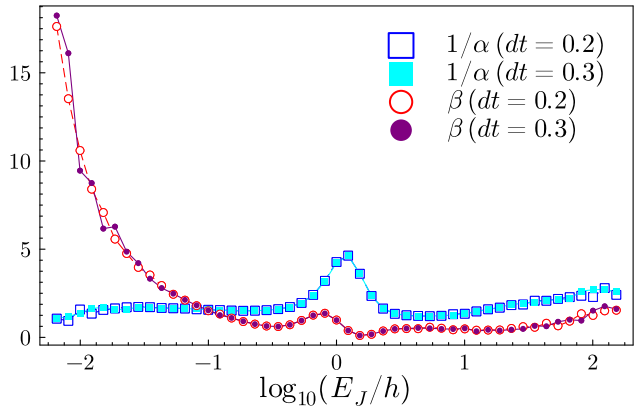

FIG. 6. Fit coefficients for the ansatz in the main text for two different step sizes as a function of E_J/h.

ing SIs [**?** **?**]. It is, therefore, fundamental that results obtained with SIs be tested in order to exclude method- and step size-dependent effects. One option is to rerun a set of LS using a different (higher-order) SI, and another is fixing the algorithm and modifying the time step. Due to the extreme numerical sensitivity present in numerical simulations of strongly chaotic systems, this will generate a completely different trajectory even if the same initial conditions are used. Nevertheless, we must guarantee that the spectra at the end of the run are the same within the errors devised in the earlier section.

We therefore retain the optimized 2nd method, but *increase* the step size from $\mathrm{d}t = 0.2$ to $\mathrm{d}t = 0.3$. This is a particularly stringent test because we are essentially showing that the step size used in the simulations was more than enough to achieve robust results. A qualitative comparison between a set of mean LS in the SRN and LRN regimes obtained using different step sizes can be seen in Fig. **??**. As is clearly visible, LS coincide within their error bounds for all values of E_J/h. For a quantitative comparison, we define the *spectral error* for each E_J/h value as the root mean square

$$\varepsilon_1(E_J/h) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N^2}\left(\overline{\Lambda_i'} - \overline{\Lambda_i}\right)^2} \quad , \tag{14}$$

where the prime denotes a spectrum computed using a different step size. We then compare this error measure with the integrated *combined error* for each curve, *i.e.*

$$\varepsilon_2(E_J/h) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N^2}\left[\varepsilon(\overline{\Lambda_i'})^2 + \varepsilon(\overline{\Lambda_i})^2\right]} \quad , \tag{15}$$

where $\varepsilon(\overline{\Lambda_i'})$ and $\varepsilon(\overline{\Lambda_i})$ are the point-errors in LS computed using $\mathrm{d}t = 0.3$ and $\mathrm{d}t = 0.2$. Stating that $\varepsilon_1 < \varepsilon_2$ means that changing the step size generates the same LS from the point of view of numerical accuracy, and in Fig. **??** one can see this is precisely the case except for the strongly chaotic neighborhood of $E_J/h \approx 1$. However, the distance between the errors is smaller than 10^{-3}, such that even here we can consider runs with different step sizes to have produced the same LS.

As a last comparison, in Fig. **??** we display the fit coefficients for the ansatz used in the main text for the LS computed with different step sizes. It is clear that the power-law behavior for $\mathrm{d}t = 0.2$ is the same as for $\mathrm{d}t = 0.3$, and the critical exponents for both β curves are the same within their error bounds.

[] R. I. McLachlan and G. Quispel, What kinds of dynamics are there? Lie pseudogroups, dynamical systems and geometric integration, Nonlinearity **14**, 1689 (2001).

[] H. Yoshida, Construction of higher-order symplectic integrators, Phys. Lett. A **150**, 262 (1990).

[] R. I. McLachlan and P. Atela, The accuracy of symplectic integrators, Nonlinearity **5**, 541 (1992).

[] H. Kinoshita, H. Yoshida, and H. Nakai, Symplectic integrators and their application to dynamical astronomy, Celest. Mech. Dyn. Astron. **50**, 59 (1991).

[] S. Choi and J. Vaníček, High-order geometric integrators for representation-free ehrenfest dynamics, The Journal of Chemical Physics **155**, 124104 (2021).

[] J. Laskar and P. Robutel, High order symplectic integrators for perturbed Hamiltonian systems, Celest. Mech. Dyn. Astron. **80**, 39 (2001).

[] F. Mogavero, N. H. Hoang, and J. Laskar, Timescales of chaos in the inner solar system: Lyapunov spectrum and quasi-integrals of motion, Phys. Rev. X **13**, 021018 (2023).

[] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Lyapunov characteristic exponents for smooth dynami-

cal systems and for Hamiltonian systems: a method for computing all of them. Parts 1 and 2, Meccanica **15**, 9 (1980).

[] K. Geist, U. Parlitz, and W. Lauterborn, Comparison of different methods for computing Lyapunov exponents, Prog. Theor. Phys. **83**, 875 (1990).

[] K. J. Palmer, Exponential dichotomies, the shadowing lemma and transversal homoclinic points, Dynamics reported , 265 (1988).

[] LsqFit.jl: Least–Squares fitting in Julia, `https://github.com/JuliaNLSolvers/LsqFit.jl` (2019).

[] G. Wanner and E. Hairer, *Solving ordinary differential equations II*, Vol. 375 (Springer Berlin Heidelberg New York, 1996).

[] G. Benettin and A. Giorgilli, On the Hamiltonian interpolation of near-to-the identity symplectic mappings with application to symplectic integration algorithms, Journal of Statistical Physics **74**, 1117 (1994).

[] E. Hairer and C. Lubich, The life-span of backward error analysis for numerical integrators, Numerische Mathematik **76**, 441 (1997).