

# Phantom relaxation rate of the average purity evolution in random circuits due to Jordan non-Hermitian skin effect and magic sums

Marko Žnidarič

*Physics Department, Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana, Slovenia*



(Received 19 June 2023; accepted 8 August 2023; published 1 September 2023)

Phantom relaxation is relaxation with a rate that is not given by a finite spectral gap. Studying the average purity dynamics in a staircase random Haar circuit and the spectral decomposition of a nonsymmetric matrix describing the underlying Markovian evolution, we explain how that can arise out of an ordinary-looking spectrum. Crucial are alternating expansion coefficients that diverge in the thermodynamic limit due to the non-Hermitian skin effect in the matrix describing the average purity dynamics under an overall unitary evolution. The mysterious phantom relaxation emerges out of localized generalized eigenvectors describing the Jordan normal form kernel, and, independently, also out of interesting trigonometric sums due to localized true eigenvectors. All this shows that when dealing with non-Hermitian matrices it can happen that the spectrum is not the relevant object; rather, it is the pseudospectrum or, equivalently, a delicate cancellation enabled by localized eigenvectors.

DOI: [10.1103/PhysRevResearch.5.033145](https://doi.org/10.1103/PhysRevResearch.5.033145)

## I. INTRODUCTION

Quantum evolution under Hermitian generators evolves states via unitary propagators, thereby preserving the norm and the inner products between states. Similarly, classical evolution in phase space is governed by a unitary Koopman propagator [1]. Our physics intuition is therefore often built on properties of Hermitian or unitary matrices. An example is a focus on the real spectrum of Hermitian matrices, for instance in low-temperature condensed matter [2], or using various spectral objects as indicators of quantum chaos [3]. However, real systems are never really closed, naturally leading to non-Hermiticity, an example being the Lindblad master equation [4]. Perhaps even more importantly, in a many-body system we can never control and observe all exponentially many observables. In such a situation one invariably either traces out some degrees of freedom or does a coarse graining, again ending up with a non-Hermitian description. A paradigmatic example is the second law of thermodynamics where the “arrow of time” arises out of unitary dynamics.

The question one can then ask is how good our intuition gained from Hermitian linear algebra works on non-Hermitian matrices? In the present work we shall demonstrate that it can be completely wrong. In particular, the spectrum of a non-Hermitian matrix is not necessarily the object determining physics—rather, it is the eigenvectors. Because these eigenvectors need not be orthogonal they can lead to large spectral expansion coefficients, in turn resulting in a delicate cancellation leading to slow phantom relaxation (which would be

impossible according to Hermitian intuition). Another twist can be added by a nondiagonalizable extensively large Jordan normal form block, i.e., an exceptional point, with localized generalized eigenvectors.

Let us consider a process whose dynamics is described by a matrix iteration,  $\mathbf{I}(t+1) = A\mathbf{I}(t)$ , where  $\mathbf{I}(t)$  is some state vector at time  $t$  and  $A$  is a time-independent  $n \times n$  matrix. Let the matrix  $A$  have the largest eigenvalue  $\lambda_0 = 1$ , with the corresponding right eigenvector  $\mathbf{I}_\infty = \mathbf{I}(\infty)$  describing a steady state,  $A\mathbf{I}_\infty = \mathbf{I}_\infty$ , while all other eigenvalues  $\lambda_j$  are strictly smaller in modulus,  $|\lambda_j| < 1$ . One would expect that if the gap  $1 - |\lambda_1|$  is finite in the thermodynamic limit (TDL)  $n \rightarrow \infty$ , then the relaxation rate towards the steady state  $\mathbf{I}(t) - \mathbf{I}_\infty$  will be given by that second largest eigenvalue  $\lambda_1$ .

We explain why such a statement can be wrong if  $A$  is non-Hermitian (nonsymmetric) and how a so-called phantom relaxation  $|\mathbf{I}(t) - \mathbf{I}_\infty| \sim \lambda_{\text{ph}}^t$ , where the rate is given by a nonexistent phantom “eigenvalue”  $\lambda_{\text{ph}}$  such that  $|\lambda_1| < \lambda_{\text{ph}} < 1$ , can arise out of spectral decomposition in the TDL, i.e., for a many-body system. The phenomenon has been observed for purity in several previous works [5–7], as well as in out-of-time-ordered correlation (OTOC) functions [8] and in nonrandom Floquet systems [9]. While our explicit example is for  $A$  being the average purity propagator, similar mathematical properties should be possible in other situations involving non-Hermitian matrices, for instance transfer matrices [10,11]. Indeed, it has been observed that the gap does not necessarily give the correct relaxation rate in Lindblad master equations [12–16] and that the emergence of true relaxation rate can be due to many contributing eigencomponents [17].

Focusing on an example where  $A$  describes the average purity evolution under staircase random circuit we use the spectral decomposition of  $A$  to show how  $\lambda_{\text{ph}}$ , that is not visible anywhere in the spectrum, somehow magically emerges out of localized left and right eigenvectors, also known as

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

a non-Hermitian skin effect [18–23]. At short but extensive times one also needs to take into account an extensively large Jordan normal form kernel, which also shows a non-Hermitian skin effect with localized generalized eigenvectors. We stress that the non-Hermitian skin effect in our case does not arise out of an explicit non-Hermitian generator, or an outside noise, but emerges naturally out of a nonsymmetric matrix that describes the reduced dynamics after tracing out part of the system. The Jordan kernel is also instrumental for a rapid emergence of the pseudospectrum, i.e., the spectrum of a slightly perturbed matrix, which is an alternative way to get  $\lambda_{\text{ph}}$  [6]. There are other recent observations that the pseudospectrum has physical relevance [24–27].

### A. Summary of results

We study purity evolution under unitary random circuit starting with an  $n$ -qudit product initial state  $|\psi(0)\rangle$ , with local qudit Hilbert space dimension denoted by  $d$ . The state at time  $t$  is  $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ , where  $U(t)$  is a unitary that is a product of random two-qudit gates (explicit details will be given later). The purity  $I_k(t)$  that measures bipartite entanglement between the first  $k$  and the last  $n - k$  qudits is

$$I_k(t) = \text{tr}(\rho_k^2(t)), \quad \rho_k(t) = \text{tr}_{k+1, \dots, n} |\psi(t)\rangle \langle \psi(t)|. \quad (1)$$

The purity is initially  $I_k(0) = 1$  and decays towards its random state value [28]  $[\mathbf{I}_\infty]_k = I_k(t \rightarrow \infty)$  at long times. Note that the purity of the subsystem, i.e., of the first  $k$  qudits, decreases because we are tracing out a part of the system—the evolution on the whole system of  $n$  qudits is unitary. Under such unitary evolution the purity measures genuine bipartite entanglement [29]. Relaxation of purity is then studied by looking at  $\Delta I_k(t)$ ,

$$\Delta I_k(t) = I_k(t) - [\mathbf{I}_\infty]_k, \quad [\mathbf{I}_\infty]_k = \frac{d^k + d^{n-k}}{1 + d^n}, \quad (2)$$

where we use square brackets to denote a component of a (bold) vector. One can show that the average evolution of quantities quadratic in the evolved state is Markovian [30]. Compactly writing purities  $I_k(t)$  for all  $k$  as components of a vector [31]  $\mathbf{I}$ , evolution of average purities (averaged over random circuit randomness) is governed by a simple matrix iteration [32],

$$\mathbf{I}(t+1) = A \mathbf{I}(t). \quad (3)$$

There are several equivalent ways of writing down such a matrix iteration [6,30,32,33]; in the most compact one [6], that we also use,  $A$  is an  $n \times n$ -dimensional real matrix.

Instead of purity one often studies also the second Rényi entropy  $S_2(t) = -\ln I_k(t)$ . In principle  $S_2$  can behave differently than the purity (the average of the logarithm vs the logarithm of the average); however, for our specific Haar random circuit one would observe the phantom relaxation also in the second Rényi entropy, as well as in other Rényi entropies; for numerical data in slightly different circuits see Refs. [5,9].

We are going to use the spectral decomposition of  $A$ . In fact,  $A$  has an  $\frac{n}{2} - 1$ -dimensional kernel  $A_{\text{ker}}$  composed of a

single Jordan block, so that we can write

$$A = A_\lambda + A_{\text{ker}}, \quad (4)$$

where  $A_\lambda$  is diagonalizable and contains all nonzero eigenvalues and can be spectrally decomposed as

$$A_\lambda = \sum_j \lambda_j |\mathbf{R}_j\rangle \langle \mathbf{L}_j|, \quad (5)$$

with biorthogonal left and right eigenvectors. The kernel  $A_{\text{ker}}$  contributes to the purity  $I_k(t)$  only for  $t \leq t_K$ , where

$$t_K = \left\lfloor \frac{t_c}{2} \right\rfloor, \quad t_c = n - k - 1. \quad (6)$$

and can therefore be ignored for  $t > t_K$ . Writing out localized left and right eigenvectors of  $A_\lambda$  (5), projecting the initial vector of purities  $\mathbf{I}(0)$  on the left eigenvectors, one gets a spectral resolution of  $\Delta I_k(t)$ ,

$$\Delta I_k(t) = \sum_{j=1}^{n/2-1} c_j \lambda_j^t, \quad \lambda_j = (2\alpha)^2 \cos^2 \left( \frac{\pi j}{n} \right), \quad (7)$$

in terms of non-steady-state eigenvalues  $\lambda_j$  that are all nondegenerate and expansion coefficients  $c_j$  and  $\alpha = d/(d^2 + 1)$ . In particular, the largest eigenvalue in the sum is in the TDL

$$\lambda_1 = (2\alpha)^2. \quad (8)$$

Using special properties of certain (magic) trigonometric sums involved in  $c_j$  we show that one has an exact phantom decay  $\Delta I_k(t) = \lambda_{\text{ph}}^t$  for  $t < t_c$  (6), with

$$\lambda_{\text{ph}} = \frac{\alpha}{1 - \alpha} = \sum_{k=1}^{\infty} \alpha^k. \quad (9)$$

In Fig. 1(a) we show an example of purity decay, illustrating phantom decay for  $t < t_c$ . As long as  $n - k$  is extensive ( $n - k \propto n$ ) this time grows with the system size, and therefore  $\lambda_{\text{ph}}$ , and not  $\lambda_1$ , is the relevant decay in the TDL [34]. We find that  $\lambda_{\text{ph}}$ , instead of  $\lambda_1$ , comes into play due to  $\mathbf{I}(0)$  having an extensive number of nonzero components, i.e., a nonzero overlap with localized  $|\mathbf{L}_j\rangle$ .

Because  $0 < \lambda_j < 1$ , the only way the slower decay with  $\lambda_{\text{ph}} > \lambda_1$  can emerge is if  $c_j$  (7) get large with  $n$  and if they alternate in sign ( $c_j$  are real). Namely, if  $c_j$  would be of the same sign, then one could bound  $\Delta I_k(t+1) \leq \Delta I_k(t) \lambda_1$  because  $c_j \lambda_j^{t+1} \leq c_j \lambda_j^t \lambda_1$ , and therefore the decay could not be slower than  $\lambda_1$ . The origin of the slower phantom decay is in a delicate cancellation of many diverging expansion coefficients  $c_j$ , see Fig. 2. In addition to growing with  $n$  they also grow rapidly with the eigenvalue index  $j$ , i.e.,  $c_j$  corresponding to smaller eigenvalues  $\lambda_j$  are much larger than those corresponding to larger ones. This means that, contrary to intuition, in the spectral sum (7) one cannot neglect smaller  $\lambda_j$ . In fact, quite the opposite. Up to some time (that can scale as  $\propto n$ ) terms with smaller  $\lambda_j$  give in modulus a larger contribution. All this is due to a non-Hermitian skin effect, i.e., exponential localization of eigenvectors of  $A_\lambda$ , causing the expansion coefficients  $c_j$  to blow up with  $n$  (and  $j$ ).

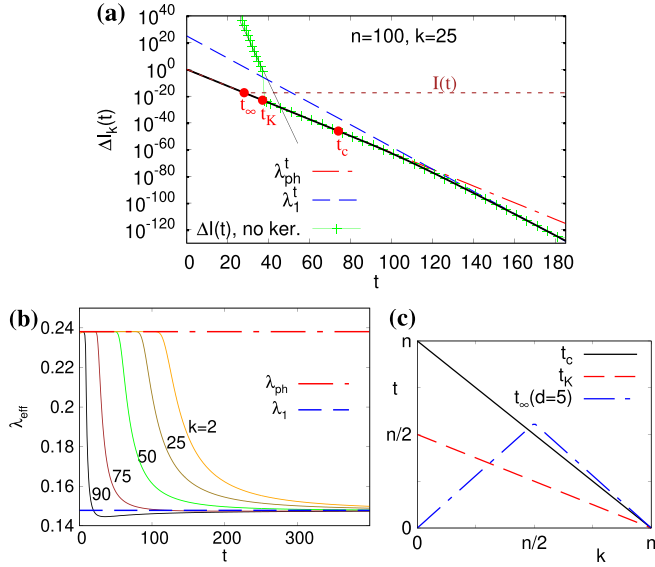


FIG. 1. Relaxation of purity to its asymptotic value. (a) Full black line shows exact  $\Delta I_k(t)$ , which initially decays with a phantom rate as  $\lambda_{\text{ph}}^t$  (red chain line), after which it eventually transitions into  $\lambda_1^t$  (blue dashed line). Brown curve that saturates shows  $I_k(t)$ . Red points mark theoretical purity saturation time  $t_\infty$  (20), time  $t_K$  (6) until the kernel contributes, and  $t_c$  (6) when the effective rate starts to deviate from  $\lambda_{\text{ph}}$ . Green pluses show purity evolution neglecting the kernel [full thin black line overlapping with diverging green pluses is Eq. (39)]. (b) Instantaneous relaxation rate  $\lambda_{\text{eff}}$  (defined via  $\Delta I_k(t) \sim \lambda_{\text{eff}}^t$  [Eq. (40)]), from which one can more clearly see the transition, for the same  $n = 100$  and  $k = 25$  as in (a), as well as for bipartitions with  $k = 2, 50, 75, 90$ . (c) Dependence of three timescales marked in (a) on the bipartition index  $k$ .

For  $t > t_c$  the effective decay  $\lambda_{\text{eff}}$  defined by  $\Delta I_k(t) \sim \lambda_{\text{eff}}^t$  starts to transition towards  $\lambda_1$ . The transition is not sharp, see Fig. 1(b), with the shape being described in the TDL by Jacobi theta functions. Using appropriate asymptotics we show that  $\lambda_{\text{eff}}(t) - \lambda_1$  reaches any  $n$ -independent value at a time scaling as  $\sim n$ , while it becomes small of the order  $\sim 1/n^2$  only at a longer time scaling as  $\sim n^2$ .

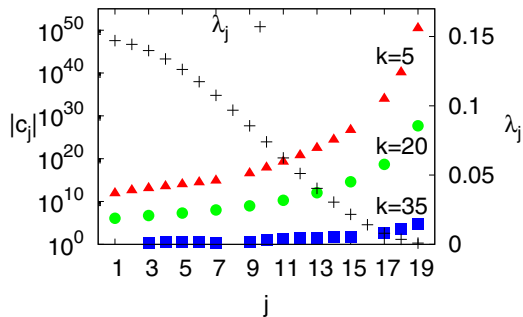


FIG. 2. Absolute values of eigenvalue expansion coefficients  $c_j$  of purity  $I_k$  (7). They alternate in signs and grow rapidly with eigenvalue index  $j$ . Shown are data for  $n = 40$  and bipartitions with  $k = 5, 20, 35$  (full red, green, and blue points). With black pluses are shown the corresponding eigenvalues  $\lambda_j$  (right axis). Missing points, e.g., at even  $j$  for  $k = 20$ , or  $j = 8, 16$  for  $k = 5, 35$ , are where  $c_j = 0$ .

The kernel can be written in terms of generalized eigenvectors [35] as

$$A_{\text{ker}} = \sum_{k=1}^{n/2-2} |r_k\rangle \langle l_{k+1}|, \quad (10)$$

where  $\langle r_k | l_j \rangle = \delta_{k,j}$ . Due to the shift property of generalized eigenvectors,  $A |r_k\rangle = |r_{k-1}\rangle$ , and  $A^T |l_k\rangle = |l_{k+1}\rangle$ , we can also immediately write powers,

$$A_{\text{ker}}^t = \sum_{k=1}^{n/2-t-1} |r_k\rangle \langle l_{k+t}|. \quad (11)$$

As mentioned, the kernel  $A_{\text{ker}}$  is important only for  $t \leq t_K$  (6) where, however,  $A_{\text{ker}}$  is crucial to get the correct phantom decay. Despite no explicit time dependence, except in the size of the kernel and vector pairings (11), writing down the contribution of  $A_{\text{ker}}$  to  $\Delta I_k(t)$  we find that this gets reflected in the time dependence that is such as to exactly cancel a diverging contribution from  $A_\lambda$  [green pluses for  $t \leq t_K$  in Fig. 1(a)]. In a way, similarly to  $t > t_K$ , where localization of eigenvectors was crucial, to get the slow decay with  $\lambda_{\text{ph}} > \lambda_1$  for  $t \leq t_K$  it is the non-Hermitian skin effect (localization) of generalized eigenvectors that is important.

Therefore, we learn that in non-Hermitian matrices it can happen that the spectrum is not the relevant object determining relaxation. Rather, it is the localization (non-Hermitian skin effects) of eigenvectors, or, as we find for  $t \leq t_K$ , also of generalized eigenvectors of the Jordan normal form structure of the kernel.

It has been shown in Ref. [6] that  $\lambda_{\text{ph}}$  is equal to the norm of the pseudospectrum—the spectrum of a slightly perturbed matrix  $A$ . While the pseudospectrum is independent of the perturbation strength  $\varepsilon$  in the TDL [36], we show that the speed with which it is established in a finite system crucially depends on the extensive size of the Jordan kernel. Namely, the effect of perturbation scales in a nonlinear way as  $\varepsilon^{1/(n/2-1)}$ , meaning that in the TDL limit even exponentially small perturbation will completely change the spectrum of  $A$ . The spectrum of  $A$  is therefore very sensitive, while the pseudospectrum is a robust object [36].

## II. PURITY DECAY AND ITS MATRIX DESCRIPTION

We are interested in the purity evolution under a random circuit. The random circuit we shall study is composed of random independent two-qudit gates drawn according to unitarily invariant Haar measure [37]. The unitary propagator for one unit of time  $U(1)$  is a product of all nearest-neighbor gates applied in a staircase order (Fig. 3),

$$U(1) = U_{n,n-1} U_{n-1,n-2} \cdots U_{1,2}, \quad (12)$$

where  $U_{i,i+1}$  are all independent and distributed according to the Haar measure; in other words, they are random  $U(d^2)$  unitaries acting on  $d$ -dimensional local space of qudits. The unitary propagator for time  $t$  is a product of  $t$  such one-step propagator  $U(1)$ , with all gates being independent and identically distributed. Random circuits have a long history, going back to Ref. [38] and more recent experiments [39,40], as well as numerous theoretical works exploring their statistical and

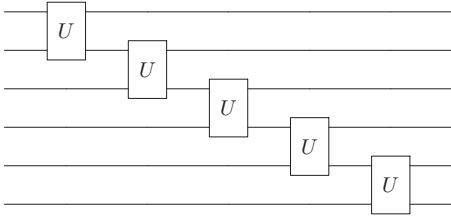


FIG. 3. Staircase configuration of gates on  $n = 6$  qudits. Time runs from left to right, and the first qudit is the topmost. All nearest-neighbor gates  $U$  are random independent Haar distributed unitaries.

dynamical properties, including Refs. [41–48]. The dynamics of the average purity is a Markovian process [30,32,33] allowing one to write the time evolution of the average purity for all  $2^n$  different bipartitions as a simple matrix iteration. The matrix involved is exponentially large in  $n$ , however, for the staircase circuit with open boundary conditions and bipartitions with a single boundary (single-cut) between the two subsystems, i.e.,  $I_k$  denotes purity for a bipartition into first  $k$  qudits and the rest, one can write the evolution for those bipartitions in terms of a matrix of size  $n$  [6]. Defining an  $n$ -component vector of purities  $\mathbf{I} = (1, I_2, I_3, \dots, I_{n-1}, 1)$  [49] one has [6]

$$\mathbf{I}(t+1) = A\mathbf{I}(t), \quad (13)$$

where the  $n \times n$  matrix  $A$  can be written in terms of two  $(n-2)$ -dimensional column vectors  $\mathbf{a}_{1,2}$  and a  $(n-2)$ -dimensional square Toeplitz matrix  $T$ ,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{a}_1 & T & \mathbf{a}_2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{aligned} \mathbf{a}_1 &= (\alpha^2, \alpha^3, \dots, \alpha^{n-1}) \\ \mathbf{a}_2 &= (0, \dots, 0, \alpha) \end{aligned} \quad (14)$$

where  $T$  is

$$T = \begin{pmatrix} \alpha^2 & \alpha & 0 & \dots & 0 \\ \alpha^3 & \alpha^2 & \alpha & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \alpha^{n-2} & \alpha^{n-3} & \ddots & \ddots & \alpha \\ \alpha^{n-1} & \alpha^{n-2} & \dots & \dots & \alpha^2 \end{pmatrix}. \quad (15)$$

Dependence on the local Hilbert space dimension  $d$  is contained in the only parameter  $\alpha$ ,

$$\alpha = \frac{d}{d^2 + 1}. \quad (16)$$

All the interesting properties that we discuss will follow from this simple matrix  $T$ . Its most important property is that it is not symmetric, which will lead to a non-Hermitian skin effect. Note that this non-Hermiticity is not put in “by hand” like in many non-Hermitian skin effect studies but emerges automatically from the underlying unitary evolution.

Simplification of the average purity evolution from an exponentially large matrix to one that is only polynomial in  $n$  is possible also for multicut bipartitions [7], with the matrix though being much more complicated. In this work we stick to the single-cut bipartitions where the simple form (14) will allow for a full spectral decomposition. We should also mention that while the exact matrix iteration (13) results only after averaging over Haar random gates, the phenomenon of

phantom relaxation is due to self-averaging observed also on a single-realization basis and for other Rényi entropies [5], as well as in some Floquet systems without any randomness [9].

Starting with a fully separable initial state all initial purities are 1, and therefore the initial vector is

$$\mathbf{I}(0) = (1, \dots, 1). \quad (17)$$

Purity for bipartition  $k$  is then simply

$$I_k(t) = \langle \mathbf{e}_k | A^t \mathbf{I}(0) \rangle, \quad (18)$$

where  $\mathbf{e}_k$  is the  $k$ th basis vector, i.e.,  $[\mathbf{e}_k]_p = \delta_{p,k}$ . Iteration in Eq. (13) with matrix  $A$  (14) is the starting point of our work. For details on its derivation see Ref. [6].

Because we want to focus on the relaxation process towards a state reached at long times, and in a finite system purity at long times is not zero, we shall subtract from  $I_k(t)$  its long-time random state value  $I_k(\infty)$  and study  $\Delta I_k(t)$  (2). In Ref. [6] it has been shown by directly solving recursive relations that the initial decay is  $\Delta I_k(t) = \lambda_{\text{ph}}^t$  in the leading order in  $n$ , with the phantom decay rate given by

$$\lambda_{\text{ph}} = \frac{\alpha}{1 - \alpha} = \sum_{k=1}^{\infty} \alpha^k. \quad (19)$$

Independently, it has been also observed [6] that this  $\lambda_{\text{ph}}$  is exactly equal to the norm of the pseudospectrum of  $A$ , which can be exactly calculated for Toeplitz matrices (the pseudospectrum is given by the so-called symbol of the Toeplitz matrix [36], which in physics language is nothing but the Bloch Hamiltonian, i.e., Fourier transformation of values along the antidiagonal).

Knowing  $\lambda_{\text{ph}}$  we can estimate the time  $t_{\infty}$  when the purity becomes close to  $I_k(\infty)$ ; see Fig. 1. Equating  $\lambda_{\text{ph}}^{t_{\infty}} = I_k(\infty)$  gives

$$t_{\infty} = \min(k, n-k) \frac{\ln d}{-\ln \lambda_{\text{ph}}}. \quad (20)$$

Before diving into spectral decomposition of  $A$  let us comment on the form of  $A$ . The main crux will be obtaining and handling spectral properties of the matrix  $T$ . Namely, eigenvectors of  $A$  will be readily constructed from those of  $T$ , while the spectrum of  $A$  is a union of the spectrum of  $T$  and two steady-state eigenvalues  $\lambda_0 = 1$ .

Matrix  $T$  is rather interesting: It is one of the simplest matrices that has an extensively large kernel and at the same time a nontrivial spectrum and eigenvectors. Some simple deformations of  $T$  result in simpler spectral properties, e.g., (i) setting the superdiagonal to 0 results in a lower triangular matrix that has only a  $(n-2)$ -dimensional Jordan block corresponding to the eigenvalue  $\alpha^2$  (and no nonzero eigenvalues); (ii) deleting the lower triangle, i.e., keeping the diagonal and superdiagonal, would result in the same; and (iii) keeping only the superdiagonal would again result in a single Jordan block with eigenvalue 0. Knowing the spectral properties of  $T$  could therefore be of interest in itself in linear algebra, for instance for analysis of numerical algorithms, where special badly conditioned matrices are used as test cases for numerical linear algebra algorithms [50]. In fact, replacing the first column in  $T$  with zeros and using  $\alpha = \frac{1}{2}$  gives the iteration matrix of a Gauss-Seidel iterative method of solving a set



of linear equations [36], where the speed of the algorithm directly relates to the pseudospectrum of the matrix. Another case is taking  $-\alpha$  instead of  $\alpha$  in the superdiagonal and setting  $\alpha = 1$ , which results in a so-called Grcar matrix [51].

### III. SPECTRAL DECOMPOSITION

In this section we are going to write the spectral decomposition of  $A_\lambda$  and get the corresponding spectral sum in Eq. (7). This will get us correct purity relaxation for  $t > t_K$  (6), including  $\lambda_{\text{ph}}$  and the transition behavior from  $\lambda_{\text{ph}}$  to  $\lambda_1$ . In the subsequent section we are then going to study  $A_{\text{ker}}$ , which will also explain the behavior for  $t \leq t_K$ .

$A_\lambda$  can be decomposed as

$$A_\lambda = |\mathbf{R}_0\rangle\langle\mathbf{L}_0| + |\mathbf{R}'_0\rangle\langle\mathbf{L}'_0| + \sum_{j=1}^{n/2-1} \lambda_j |\mathbf{R}_j\rangle\langle\mathbf{L}_j|, \quad (21)$$

where we use the bra-ket notation for left or right eigenvectors that satisfy biorthogonality,

$$\langle\mathbf{L}_k|\mathbf{R}_p\rangle = \delta_{k,p}. \quad (22)$$

In addition to biorthogonality (22) one can also enforce normalization of one of them (but not both), e.g.,  $\langle\mathbf{R}_k|\mathbf{R}_k\rangle = 1$  or  $\langle\mathbf{L}_k|\mathbf{L}_k\rangle = 1$ . The first two projectors in Eq. (21) correspond to two steady states, where as we will see only one is physically relevant.

We shall first write down eigenvectors of  $T$  and then, due to simple connection between  $T$  and  $A$ , also those of  $A$ . The left and right eigenvectors and  $\lambda_j$  of  $T$  have already been obtained in Refs. [6,52]; we write them down again for completeness. The eigenvalues of  $T$  are

$$\lambda_j = (2\alpha \cos \varphi_j)^2, \quad \varphi_j = \frac{\pi j}{n}, \quad j = 1, \dots, \frac{n}{2} - 1. \quad (23)$$

Throughout the paper we for simplicity assume that  $n$  is even so that  $n/2$  is an integer. Denoting by  $\tilde{\mathbf{L}}_j$  and  $\tilde{\mathbf{R}}_j$  the corresponding unnormalized [not satisfying Eq. (22)] eigenvectors of  $T$ , their components are

$$[\tilde{\mathbf{R}}_j]_k = (2\alpha \cos \varphi_j)^{k-2} \frac{\sin[(k+1)\varphi_j]}{\sin \varphi_j}, \quad (24)$$

$$[\tilde{\mathbf{L}}_j]_k = (2\alpha \cos \varphi_j)^{n-3-k} \frac{\sin[(n-k)\varphi_j]}{\sin \varphi_j}, \quad (25)$$

where  $k = 1, \dots, n-2$ , and  $\varphi_j$  used throughout the paper is given in Eq. (23). Note that left eigenvectors are simply the reflected right ones, i.e., can be obtained by replacing the component index  $k$  by  $n-k-1$ . Right eigenvectors are exponentially localized at the left edge (small indices  $k$ ), while the left are localized at the right edge, with the localization length being  $-1/\ln(2\alpha \cos \varphi_j)$  which is upper-bounded by an  $n$ -independent  $-1/\ln(2\alpha)$ . The inner product between left and right eigenvectors can be simplified using  $2 \sin \alpha \sin \beta = \cos(\alpha - \beta) - \cos(\alpha + \beta)$ , resulting in

$$N_j = \sum_{k=1}^{n-2} [\tilde{\mathbf{R}}_j]_k [\tilde{\mathbf{L}}_j]_k = (-1)^{j+1} (2\alpha \cos \varphi_j)^{n-5} \frac{n \cos \varphi_j}{2 \sin^2 \varphi_j}. \quad (26)$$

This normalization is exponentially small in  $n$  [remember that (16) implies  $\alpha < 1$  for  $d > 1$ ]. When enforcing

biorthogonality (22), which means we will have to divide one of them by  $N_j$ , this will cause exponentially large expansion coefficients in the spectral decomposition.

From eigenvectors of  $T$  we can construct those of  $A$ ,

$$\begin{aligned} \mathbf{R}_j &= \frac{1}{N_j} (0, \tilde{\mathbf{R}}_j, 0), \\ \mathbf{L}_j &= \left( \frac{\langle \tilde{\mathbf{L}}_j | \mathbf{a}_1 \rangle}{\lambda_j - 1}, \tilde{\mathbf{L}}_j, \frac{\langle \tilde{\mathbf{L}}_j | \mathbf{a}_2 \rangle}{\lambda_j - 1} \right), \end{aligned} \quad (27)$$

that is, components  $k = 2, \dots, n-1$  of  $\mathbf{L}_j$  and  $\mathbf{R}_j$  are exactly equal to  $n-2$  components of  $\tilde{\mathbf{L}}_j$  and  $\tilde{\mathbf{R}}_j$ , respectively, while the first and the last one are as written in Eq. (27). Note that those eigenvectors already satisfy biorthogonality (22) due to dividing  $\tilde{\mathbf{R}}_j$  by  $N_j$ . Two steady states (21) corresponding to  $\lambda_0 = 1$  are

$$\begin{aligned} \mathbf{R}_0 &= (1, \mathbf{I}_\infty, 1), \\ \mathbf{L}_0 &= (\tfrac{1}{2}, 0, \dots, 0, \tfrac{1}{2}), \end{aligned} \quad (28)$$

where  $\mathbf{I}_\infty$  is a vector of purity saturation values (2) for  $k = 2, \dots, n-1$ . The second steady state is

$$\mathbf{L}'_0 = (\tfrac{1}{2}, 0, \dots, 0, -\tfrac{1}{2}), \quad (29)$$

while the right one  $\mathbf{R}'_0$ , which we do not need, is given in Ref. [53]. Because the purity vector  $\mathbf{I}(t)$  (13) always has the first and last component equal to 1, including at  $t = 0$  (and irrespective of the initial state being fully separable), one has  $\langle \mathbf{L}'_0 | \mathbf{I}(0) \rangle = 0$  and the second steady state (29) does not contribute to purity evolution. The steady state  $\mathbf{R}_0$  (28) gives exactly the asymptotic purity (as it should be), and therefore the contribution of nonzero eigenvalues, i.e.,  $A_\lambda$ , to purity relaxation is given by Eq. (7) with the expansion coefficients,

$$c_j = \langle \mathbf{e}_k | \mathbf{R}_j \rangle \langle \mathbf{L}_j | \mathbf{I}(0) \rangle = [\mathbf{R}_j]_k \sum_{p=1}^n [\mathbf{L}_j]_p, \quad (30)$$

where we plugged in  $\mathbf{I}(0) = (1, 1, \dots, 1)$  for a fully separable initial state.

#### A. Expansion coefficients

As we have explained a necessary condition to observe phantom decay is that  $c_j$  are not all of the same sign. From Eqs. (30) and (27) we immediately see that the alternating sign of  $c_j$  will come from normalization  $N_j$  (26). Due to the reflection symmetry between the left and right eigenvectors the  $k$  dependence in the exponentially decaying prefactors cancels, and the nontrivial part of the overlap comes from  $\sin[(k+1)\varphi_j] \sin[(n-k)\varphi_j]$ . Because  $\varphi_j = j\pi/n$  those sines make  $j$  half oscillations as  $k$  runs over its values, i.e., the situation is the same as for eigenfunctions of an oscillating string fixed at boundaries. With respect to the reflection around the midpoint eigenfunctions are alternatingly even and odd. This even or odd symmetry of the oscillating part of the eigenvectors (25) is then responsible for the oscillating sign of  $N_j$  and in turn of  $c_j$ .

The sum of components of  $\mathbf{L}_j$  needed in  $c_j$  (30) has two parts, one coming from the first and the last component of

$\mathbf{L}_j$ , and the other being equal to the sum of components of eigenvector  $\tilde{\mathbf{L}}_j$  (25) of matrix  $T$ . The latter is

$$\sum_{p=1}^{n-2} [\tilde{\mathbf{L}}_j]_p = \frac{1 - \lambda_{\text{ph}}(-1)^j \lambda_j^{n/2-1}}{\lambda_{\text{ph}} - \lambda_j} \approx \frac{1}{\lambda_{\text{ph}} - \lambda_j}, \quad (31)$$

where one writes  $\sin[(n-k)\varphi_j]$  as an exponential, resulting in a simple geometrical sum, and where in the approximate sign we neglected  $\lambda_j^{n/2-1}$  that is exponentially small in  $n$ . The

first and last components on the other hand give

$$\frac{\langle \tilde{\mathbf{L}}_j | \mathbf{a}_1 + \mathbf{a}_2 \rangle}{\lambda_j - 1} = \frac{-(-1)^j \lambda_j^{n/2-1} + 1}{\lambda_j - 1} \approx -\frac{1}{1 - \lambda_j}, \quad (32)$$

where again in the approximate sign we dropped exponentially small  $\lambda_j^{n/2-1}$ . All together this gives

$$c_j \approx \frac{[\tilde{\mathbf{R}}_j]_{k-1}}{N_j} \left( \frac{1}{\lambda_{\text{ph}} - \lambda_j} - \frac{1}{1 - \lambda_j} \right), \quad (33)$$

and in turn the nonzero eigenvalues contribution to the purity

$$\Delta I_k(t) = -\frac{2(2\alpha)^{2t+k-n+2}}{n} \sum_{j=1}^{n/2-1} (-1)^j (\cos \varphi_j)^{2t+k-n+1} \sin \varphi_j \sin(k\varphi_j) \left( \frac{1}{\lambda_{\text{ph}} - \lambda_j} - \frac{1}{1 - \lambda_j} \right) = a(t) - b(t), \quad (34)$$

where  $a(t)$  is the term with  $\frac{1}{\lambda_{\text{ph}} - \lambda_j}$  and  $b(t)$  the one with  $\frac{1}{1 - \lambda_j}$ . Neglecting  $A_{\text{ker}}$  this expression is exact (up to exponentially small terms we dropped); however, it is not very handy for analysis so we are going to rewrite it.

Because  $\lambda_{\text{ph}} > \lambda_j$  we can expand  $\frac{1}{\lambda_{\text{ph}} - \lambda_j} = \frac{1}{\lambda_{\text{ph}}} \sum_{r=0}^{\infty} (\lambda_j/\lambda_{\text{ph}})^r$  and interchange the order of the two summations, obtaining

$$a(t) = \frac{2}{n} \sum_{r=0}^{\infty} \frac{(2\alpha)^{2t+2r+k-n+2}}{\lambda_{\text{ph}}^{r+1}} f_k(2t+2r+k-n+1), \quad (35)$$

where

$$f_k(p) := - \sum_{j=1}^{n/2-1} (-1)^j \cos^p \varphi_j \sin \varphi_j \sin(k\varphi_j). \quad (36)$$

A similar expression with  $\lambda_{\text{ph}}$  replaced by 1 is obtain for  $b(t)$ . The purity can therefore also be written as

$$\Delta I_k(t) = \frac{4\alpha}{n} \sum_{r=r_{\min}}^{\infty} (2\alpha)^p f_k(p) \left( \frac{1}{\lambda_{\text{ph}}^{r+1}} - 1 \right), \quad (37)$$

where  $p = 2t + 2r + k - n + 1$  and  $r_{\min} = 0$ . This is the main result that we shall repeatedly use.

First, let us look at how it looks. In Fig. 1(a) the green pluses are Eq. (37), and we can see that for  $t > t_K$ , when the kernel  $A_{\text{ker}}$  no longer contributes (we shall see that in Sec. IV), it gives the exact purity, while for  $t \leq t_K$  it gives an incorrect very fast exponential decay. One can also observe (data not shown) that the rate of this decay increases with increasing  $n$ . Let us explain it.

We can see from Eq. (37) that for small  $t$  and  $r$  the power  $p$  can be negative, and, because  $2\alpha < 1$ , this gives very large terms that in fact diverge as  $n \rightarrow \infty$  ( $f_k(p)$  is also large for negative  $p$ ; see Appendix A). Diverging  $\Delta I_k(t)$  for  $t \leq t_K$  is therefore due to terms with nonpositive  $p$ . Those terms are obtained for  $r = 0, 1, \dots, r_{\max}$ , with  $r_{\max} = \lfloor \frac{n-1-k}{2} \rfloor - t = t_K - t$  (at  $r = r_{\max}$  one has  $p = -1$  for odd  $n - k - 1$ , and  $p = 0$  for even  $n - k - 1$ ). We shall also see in Sec. IV that the contribution of the kernel  $A_{\text{ker}}^t$  is exactly equal to the negative of the terms with  $p \leq 0$ . Therefore, the kernel part will exactly

cancel those diverging terms, so that the full exact expression for the purity [up to exponentially small terms neglected in Eqs. (31) and (32)] is given by Eq. (37) taking

$$r_{\min} = \max(0, t_K - t + 1) \quad (38)$$

instead of  $r_{\min} = 0$ . Kernel, therefore, just renormalizes the sum (37) so as to cancel diverging terms.

The dominant diverging term for  $t \leq t_K$  coming from  $A_{\lambda}$  can be identified from Eq. (34) and is the  $j = n/2 - 1$  term for which  $\cos \varphi_j$  is the smallest. Collecting the  $t$ -dependent part we get the leading diverging term in the TDL,

$$\Delta I_k(t) \propto \left( \frac{4\pi\alpha}{n} \right)^{2t} = \left[ \frac{4\pi\lambda_{\text{ph}}}{(1 + \lambda_{\text{ph}})n} \right]^{2t}, \quad (39)$$

where we left out an  $n$ -dependent prefactor. This is indeed an exponential decay in  $t$  [full thin steep black line overlapping with pluses in Fig. 1(a)] but with a rate that increases logarithmically with  $n$ .

## B. Effective decay rate

Let us now focus on behavior for  $t > t_K$  and in particular on an instantaneous effective decay rate  $\lambda_{\text{eff}}$  defined as  $\Delta I(t) \sim \lambda_{\text{eff}}^t$ ; in other words,

$$\ln \lambda_{\text{eff}}(t) := \frac{d \ln \Delta I(t)}{dt}. \quad (40)$$

We have seen in Fig. 1(b) that  $\lambda_{\text{eff}}$  is initially equal to  $\lambda_{\text{ph}}$ , while at late times it approaches the second-largest eigenvalue  $\lambda_1$ . In this subsection we are going to show that this is indeed the case and find the transition time  $t_c$  when the phantom decay stops.

First, in the parentheses in Eq. (37) one can neglect 1 compared to  $1/\lambda_{\text{ph}}^{r+1}$  (because  $\lambda_{\text{ph}}$  is smaller than 1, e.g., for  $d = 2$ , when it is the largest, it is  $\lambda_{\text{ph}} = 2/3$ ; for  $d = 5$  it is  $\lambda_{\text{ph}} = 5/21 \approx 0.24$ ; and for large  $d$  it scales as  $\lambda_{\text{ph}} \propto 1/d$ ), therefore approximating  $\Delta I_k(t) \approx a(t)$ . Crucial for understanding phantom decay is the function  $f_k(p)$ . One can show (Appendix A) that  $f_k(p)$  is exactly zero for a bunch of small positive  $p$ . Specifically, if  $k$  is even, then  $f_k(p) = 0$  for odd  $p = 1, 3, \dots, n - k - 3$  [note that for even  $k$  the arguments  $p$  in Eq. (37) are odd], while if  $k$  is odd one has  $f_k(p) = 0$

for even  $p = 2, 4, \dots, n - k - 3$ . Because nonpositive arguments  $p$  are excluded from the summation due to (38) as they are canceled by the kernel contribution (Sec. IV), the first nonzero term in  $\Delta I_k(t)$  (37) is when  $p = n - k - 1$ . If we denote

$$t_c = n - k - 1, \quad (41)$$

then this nonzero term is obtained for  $r = r_1 = t_c - t$ . If  $t < t_c$ , then this first  $r_1$  is positive, and so the sum in Eq. (37) does not start at  $r = 0$ ; rather, it starts as  $\frac{n}{4\alpha} \Delta I_k(t < t_c) = (2\alpha)^{t_c} f_k(t_c) / \lambda_{\text{ph}}^{t_c-t} + (2\alpha)^{t_c+2} f_k(t_c + 2) / \lambda_{\text{ph}}^{t_c-t+1} + \dots$ . Importantly, time dependence does not appear in the argument of  $f_k(p)$  but only as a common factor  $\lambda_{\text{ph}}^t$  in the powers of  $\lambda_{\text{ph}}$ . Therefore one immediately and exactly gets

$$\lambda_{\text{eff}} \approx \frac{a(t+1)}{a(t)} = \lambda_{\text{ph}}, \quad t < t_c. \quad (42)$$

We have shown that one has a phantom decay with  $\lambda_{\text{ph}} > \lambda_1$  for  $t < t_c$ . The transition time  $t_c$  is extensive, i.e., proportional to the system size  $n$  as long as  $n - k$  is extensive (symmetry between bipartitions  $k$  and  $n - k$  is broken by the directedness of the staircase configuration; Fig. 3). We can also comment on the neglected 1 in Eq. (37) and how much this would cause deviation from  $\lambda_{\text{ph}}$ —the neglected relative correction is of order  $\lambda_{\text{ph}}^{t_c-t}$  and is therefore small for large  $t_c - t$ , i.e., as long as  $t$  is far from the extensive ending point  $t_c$  of phantom relaxation.

How about  $\lambda_{\text{eff}}$  at long times after  $t_c$ ? Let us approximate  $a(t)$  only by its first term. While this might not be a very good approximation immediately after  $t_c$ , we will see that it nevertheless soon gives a good approximation and will furthermore allow us to express the shape of the transition from  $\lambda_{\text{ph}}$  to  $\lambda_1$  in terms of Jacobi theta functions. We therefore approximate

$$\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \frac{f_k(2t + k - n + 3)}{f_k(2t + k - n + 1)}. \quad (43)$$

From Eq. (43) we can also easily get the asymptotic value of  $\lambda_{\text{eff}}$  for  $t \gg t_c$ . For very long times one can approximate  $f_k(p)$  by its first  $j = 1$  term in Eq. (36)—as  $t$  grows at fixed  $n$  the argument  $p$  increases and  $\cos^p \varphi_j$  get very small, with the largest term being the one with  $\varphi_1$ . This gives

$$\lambda_{\text{eff}} \longrightarrow (2\alpha)^2 \cos^2 \varphi_1 = \lambda_1. \quad (44)$$

That is, we correctly reproduce the asymptotic purity relaxation  $\Delta I_k(t) \asymp \lambda_1^t$ .

To get the transition shape [Fig. 1(b)] we are going to study

$$\Delta \lambda_{\text{eff}}(t) := \lambda_{\text{eff}}(t) - \lambda_1. \quad (45)$$

As we shall show in next subsections,  $\Delta \lambda_{\text{eff}}(t)$  has a scaling form for extensive  $n - k$ ,

$$\Delta \lambda_{\text{eff}}(t) \approx \frac{1}{n^2} g(t/n^2), \quad (46)$$

with the scaling function  $g(x)$  that depends on  $k$  but has a short- and long-time forms,

$$g(x \ll 1) \sim \frac{1}{x^2}, \quad g(x) \asymp e^{-x}. \quad (47)$$

This means that on short times ( $t/n \ll n$ ) the scaling variable is  $t/n$ , i.e.,  $\Delta \lambda_{\text{eff}}$  will reach any fixed value in a time that scales

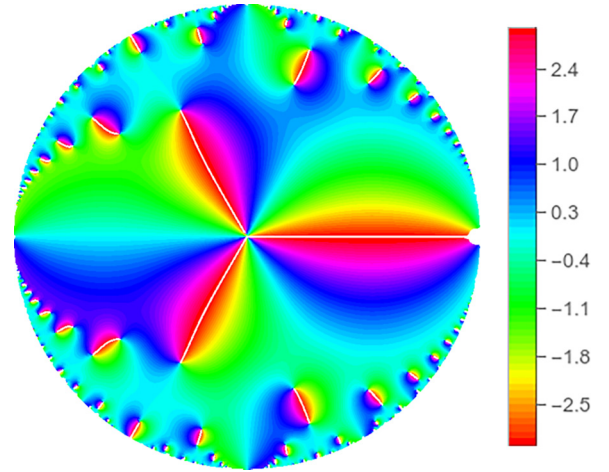


FIG. 4. The phase of  $q\vartheta_4''(q)/\vartheta_4'(q)$  (51) within the unit circle of complex  $|q| < 1$ . The transition in  $\Delta \lambda_{\text{eff}}$  for bipartition with  $k = 2$  is given by this function evaluated at real values of  $q \in [0, 1]$  (visible as a white horizontal line in figure).

as  $t \sim n$ , whereas on long times ( $t \sim n^2$ ) the scaling variable is  $t/n^2$ . This in particular means that  $\Delta \lambda_{\text{eff}}(t)$  reaches level  $\sim 1/n^2$  on a timescale  $t \sim n^2$ , explaining a rather slow-looking convergence towards  $\lambda_1$  in Fig. 1(b). In the next subsections we are going to calculate the explicit form of the scaling function  $g(x)$  for  $k = 2$ , and for  $k = n/2$ .

### C. Transition for bipartition with $k = 2$

To describe how  $\lambda_{\text{eff}}$  transitions from  $\lambda_{\text{ph}}$  to  $\lambda_1$  we need to understand  $f_2(p)$  (36),

$$f_2(p) = -2 \sum_{j=1}^{n/2-1} (-1)^j \cos^{p+1} \varphi_j \sin^2 \varphi_j. \quad (48)$$

We show in Appendix B that one can approximate  $f_2(p)$  by the Jacobi theta function [54]  $\vartheta_4(z, q)$ ,

$$\vartheta_4(z, q) = 1 + 2 \sum_{j=1}^{\infty} (-1)^j q^{j^2} \cos(2jz), \quad (49)$$

or, more precisely, its derivatives with respect to  $q$  evaluated at  $z = 0$ ,

$$\vartheta_4'(q) = \frac{d\vartheta_4(z=0, q)}{dq}, \quad \vartheta_4''(q) = \frac{d^2\vartheta_4(q)}{dq^2}, \quad (50)$$

as

$$\Delta \lambda_{\text{eff}}(t) \approx -(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 c^{2t-n+6} \frac{\vartheta_4''(c^{2t-n+4})}{\vartheta_4'(c^{2t-n+4})}, \quad (51)$$

where we abbreviated  $c := \cos \varphi_1 = \cos(\pi/n)$  and the approximation should be valid in the TDL. The transition from the phantom decay to the asymptotic decay is up to prefactors given by  $q\vartheta_4''(q)/\vartheta_4'(q)$  with real  $q = (\cos \pi/n)^{2t-n+4}$ . Theta functions are beautiful objects with an intricate structure in the complex plane, as illustrated in Fig. 4.

In Fig. 5 we compare the approximation of  $\Delta \lambda_{\text{eff}}(t)$  given by Eq. (51) with exact values. We can see that until  $t_c$  the effective decay is equal to  $\lambda_{\text{ph}}$ , after which  $\Delta \lambda_{\text{eff}}$  starts to

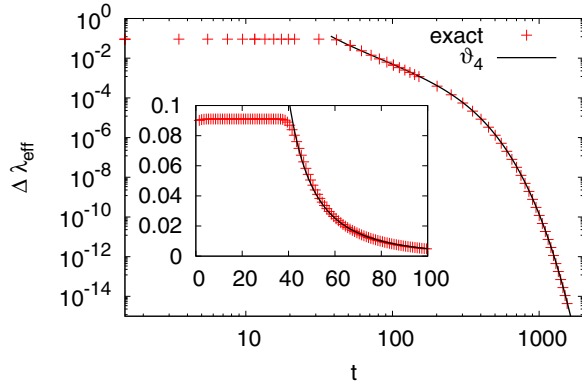


FIG. 5. The theoretical decay rate  $\Delta\lambda_{\text{eff}}$  (thin black curve) given by Eq. (51) agrees well with the exact result (symbols obtained by matrix iteration of  $A$ ) already for  $n = 40$ . Bipartition with  $k = 2$  and local dimension  $d = 5$ .

decrease. Already close to  $t_c$  approximation by Eq. (51) is good even for not very large  $n = 40$ .

One can also simplify Eq. (51) and get its short- and long-time forms. At short times  $t \sim n$  we get (Appendix B)

$$\Delta\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \frac{1}{4} \left(\frac{n}{t}\right)^2. \quad (52)$$

At long times on the scale  $t \sim n^2$  one, on the other hand, gets (Appendix B)

$$\begin{aligned} \Delta\lambda_{\text{eff}}(t) &\asymp 12(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left(\cos \frac{\pi}{n}\right)^{6t} \\ &\approx 12(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \exp(-3\pi^2 t/n^2). \end{aligned} \quad (53)$$

We compare in Fig. 6 these two expressions, seeing that they agree very well with the exact  $\Delta\lambda_{\text{eff}}(t)$ . Looking more carefully at the convergence with  $n$  (see Appendix B), one can estimate that the short-time power-law approximation (52), where the scaling variable is  $t/n$ , holds for  $5 \lesssim t/n \lesssim 0.02n$ , whereas the long-time exponential (53), where the scaling variable is instead  $t/n^2$ , holds for  $t \gtrsim 0.1n^2$ .

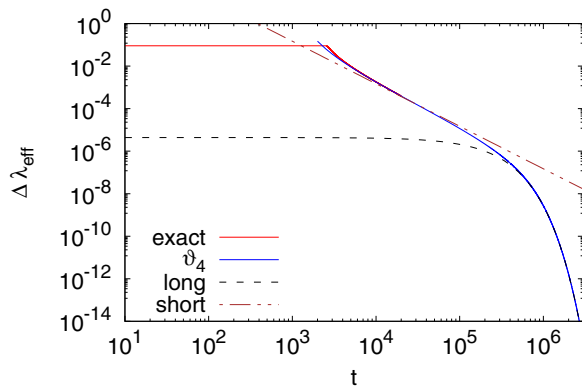


FIG. 6. The theoretical decay rate  $\Delta\lambda_{\text{eff}}$  for  $n = 2000$ ,  $k = 2$ , and  $d = 5$ . Blue full curve is expression in terms of theta function, Eq. (51), and we also show the long-time asymptotic exponential form Eq. (53) (dashed curve) and the short-time power law of Eq. (52) (chain curve).

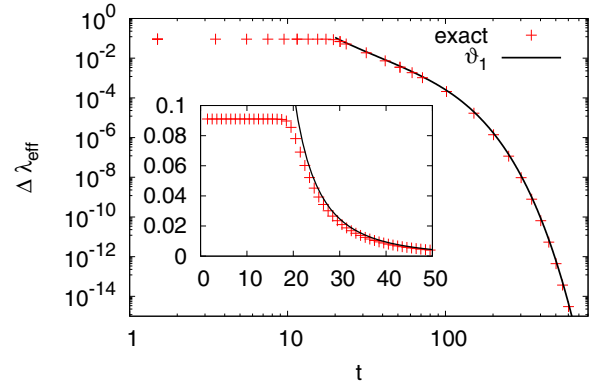


FIG. 7. Theoretical decay rate  $\Delta\lambda_{\text{eff}}$  in Eq. (56) (full black curve) agrees with the exact result (symbols, obtained by matrix iteration of  $A$ ) already for  $n = 40$ . Bipartition with  $k = 20$  and  $d = 5$ .

#### D. Transition for bipartition with $k = n/2$

For half-half bipartition we need  $f_{n/2}(p)$  (36), which is

$$f_{n/2}(p) = - \sum_{j=1}^{n/2-1} (-1)^j \cos^p \varphi_j \sin \varphi_j \sin \left( \frac{\pi}{2} j \right). \quad (54)$$

Using similar techniques as for  $k = 2$  (see Appendix C), we again get the Jacobi theta function, this time though  $\vartheta_1(z, q)$ ,

$$\begin{aligned} \vartheta_1(z, q) &= 2q^{1/4} \sum_{k=0}^{\infty} (-1)^k q^{k(k+1)} \sin[(2k+1)z], \\ \vartheta_1'(z, q) &= d\vartheta_1(z, q)/dq. \end{aligned} \quad (55)$$

The effective decay is for large  $n$ ,

$$\Delta\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left[ 1 - 4c^{x_0} \frac{\vartheta_1'(\frac{\pi}{n}, c^{x_0})}{\vartheta_1(\frac{\pi}{n}, c^{x_0})} \right], \quad (56)$$

with  $x_0 = 8t - 2n + 4$  and  $c = \cos(\pi/n)$ . In Fig. 7 we can see that this approximation agrees with the exact result very well.

Simplifying for short times, one gets (Appendix C)

$$\Delta\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \left( \frac{1}{24} + \frac{1}{8\pi^2} \right) \left(\frac{n}{t}\right)^2, \quad (57)$$

again showing that the scaling variable is  $t/n$  with the scaling function being  $g(x) \sim 1/x^2$ . At long times one instead has

$$\begin{aligned} \Delta\lambda_{\text{eff}}(t) &\asymp 24(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left(\cos \frac{\pi}{n}\right)^{16t} \\ &\approx 24(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \exp(-8\pi^2 t/n^2), \end{aligned} \quad (58)$$

and so the scaling variable is  $t/n^2$ . In Fig. 8 we see that both approximations agree well with the exact results (no fitting parameters). The short-time expression (57) holds in a range  $2 \lesssim t/n \lesssim 0.01n$ , while the long-time expression (58) for  $t \gtrsim 0.05n^2$ .

#### IV. KERNEL

Here we shall study the kernel part  $A_{\text{ker}}$  of  $A$  (4), specifically the  $(n/2 - 1)$ -dimensional kernel of  $T$  which will then



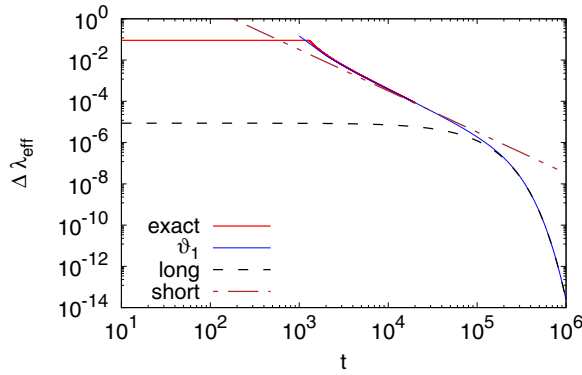


FIG. 8. Decay rate  $\Delta\lambda_{\text{eff}}$  for  $n = 2000$  and bipartition with  $k = 1000, d = 5$ . Shown is expression in terms of theta function (blue curve), Eq. (56), as well as the long-time asymptotic exponential form Eq. (58) (black dashed curve), and the short-time power law (brown chain line) of Eq. (57).

immediately give us also the  $n/2 - 1$ -dimensional kernel of  $A$ . First, we are going to write down the Jordan normal form of  $A_{\text{ker}}$  and study how it contributes to  $\Delta I_k(t)$ . In the second subsection we are then going to study how a delicate instability of the kernel under small perturbations causes the emergence of the pseudospectrum of  $A$ , which is one alternative [6] of obtaining the phantom relaxation rate  $\lambda_{\text{ph}}$ .

### A. Kernel Jordan normal form

The kernel is of size  $n/2 - 1$ , i.e., half the size of matrix  $T$ , and has a single Jordan block—the geometric multiplicity is 1 and the algebraic  $n/2 - 1$ . The spectral decomposition can be written as [35]

$$A_{\text{ker}} = \sum_{k=1}^{n/2-2} |\mathbf{r}_k\rangle\langle\mathbf{l}_{k+1}|, \quad (59)$$

where we have biorthonormality  $\langle\mathbf{r}_k|\mathbf{l}_j\rangle = \delta_{k,j}$ ,  $j, k = 1, \dots, n/2 - 1$ , and left and right vectors satisfy the shift property,  $A|\mathbf{r}_k\rangle = |\mathbf{r}_{k-1}\rangle$  and  $A^T|\mathbf{l}_k\rangle = |\mathbf{l}_{k+1}\rangle$ . In other words,  $|\mathbf{r}_k\rangle$  is in the kernel of  $A^p$  for  $p \geq k$  (but not for  $p < k$ ), while  $|\mathbf{l}_k\rangle$  is in the kernel of  $(A^T)^p$  for  $p \geq n/2 - k$  (but not for  $p < n/2 - k$ ). Such vectors are called generalized eigenvectors. If we put  $\mathbf{r}_k$  as a  $k$ th column of a  $n \times (n/2 - 1)$  rectangular matrix  $U$ , and  $\mathbf{l}_k$  as columns of a  $n \times (n/2 - 1)$  rectangular matrix  $V$ , biorthonormality becomes  $U^T V = \mathbb{I}$ , and  $V A_{\text{ker}} U^T = J$ , where  $J$  is a Jordan matrix of size  $(n/2 - 1) \times (n/2 - 1)$ , i.e., a matrix with 1 in the first superdiagonal and 0 everywhere else.

Due to the shift property the powers of  $A_{\text{ker}}$  are simple, namely

$$A_{\text{ker}}^t = \sum_{k=1}^{n/2-1-t} |\mathbf{r}_k\rangle\langle\mathbf{l}_{k+t}|, \quad t \leq n/2 - 2, \quad (60)$$

while  $A_{\text{ker}}^{n/2-1} = 0$ . Writing  $A = A_{\text{ker}} + A_\lambda$  in terms of the kernel part and the rest, and because  $A_{\text{ker}} A_\lambda = A_\lambda A_{\text{ker}} = 0$ , we have  $A^t = A_{\text{ker}}^t + A_\lambda^t$ , and we immediately know that the kernel does not play any role for  $t \geq n/2 - 1$  (a specific form

of generalized eigenvectors will actually result in even tighter condition).

Denoting by  $\tilde{\mathbf{r}}_k$  and  $\tilde{\mathbf{l}}_k$ ,  $k = 1, \dots, n/2 - 1$ , the  $(n - 2)$ -component kernel vectors of matrix  $T$  (15),

$$T_{\text{ker}} = \sum_{k=1}^{n/2-2} |\tilde{\mathbf{r}}_k\rangle\langle\tilde{\mathbf{l}}_{k+1}|, \quad (61)$$

we can immediately construct the generalized kernel eigenvectors (which have  $n$  components) of  $A$  as

$$\mathbf{r}_k = (0, \tilde{\mathbf{r}}_k, 0), \quad \mathbf{l}_k = (0, \tilde{\mathbf{l}}_k, b_k). \quad (62)$$

The scalar  $b_k$  satisfies the recursion  $b_k = b_{k+1} - \mathbf{a}_2 \cdot \tilde{\mathbf{l}}_k$ , with a starting  $b_{n/2} = 0$ . Dependence on  $\alpha$  is rather trivial, namely by a simple diagonal similarity transformation  $T$  can be transformed to a matrix with  $\alpha = 1$ , see Ref. [6], and so one can get  $\tilde{\mathbf{r}}_k$  from those for  $\alpha = 1$  by

$$[\tilde{\mathbf{r}}_k]_j = [\tilde{\mathbf{r}}_k^{(1)}]_j \alpha^{j-2k}, \quad [\tilde{\mathbf{l}}_k]_j = [\tilde{\mathbf{l}}_k^{(1)}]_j \alpha^{2k-j}, \quad (63)$$

where we have denoted by  $\tilde{\mathbf{r}}_k^{(1)}$  and  $\tilde{\mathbf{l}}_k^{(1)}$  the generalized eigenvectors for  $\alpha = 1$ .

### 1. Right vectors

Let us now calculate right vectors  $\tilde{\mathbf{r}}_k^{(1)}$  for  $T(\alpha = 1)$ . Looking at the shift property  $T|\tilde{\mathbf{r}}_k^{(1)}\rangle = |\tilde{\mathbf{r}}_{k+1}^{(1)}\rangle$  with the specific form of  $T$  (15) we can see that the recursion  $[\tilde{\mathbf{r}}_k^{(1)}]_j = [\tilde{\mathbf{r}}_{k-1}^{(1)}]_{j-1} - [\tilde{\mathbf{r}}_{k-1}^{(1)}]_{j-2}$  holds. Starting with an appropriate  $\tilde{\mathbf{r}}_1^{(1)}$  we get

$$\begin{aligned} \tilde{\mathbf{r}}_1^{(1)} &= (-1, 1, 0, \dots, 0) \\ \tilde{\mathbf{r}}_2^{(1)} &= (-1, 0, 2, -1, 0, \dots, 0) \\ \tilde{\mathbf{r}}_3^{(1)} &= (-1, 0, 1, 2, -3, 1, 0, \dots, 0), \\ \tilde{\mathbf{r}}_4^{(1)} &= (-1, 0, 1, 1, 1, -5, 4, -1, 0, \dots, 0), \end{aligned} \quad (64)$$

and so on. Recursion can in fact be solved in a closed form, with  $\tilde{\mathbf{r}}_k^{(1)}$  having nonzero only the first  $2k$  components, which are

$$[\tilde{\mathbf{r}}_k^{(1)}]_{2k-p} = -(-1)^{k+p} \sum_{r=0}^{[p/2]} (-1)^r \binom{k-r}{p-2r}, \quad (65)$$

where  $p = 0, \dots, 2k - 1$  and  $[x]$  is the largest integer that is smaller or equal than  $x$ . We for instance see that the last nonzero component is always  $[\tilde{\mathbf{r}}_k^{(1)}]_{2k} = -(-1)^k$ , the second last is  $[\tilde{\mathbf{r}}_k^{(1)}]_{2k-1} = (-1)^k k$ , the third last is quadratic in  $k$ , and so on. The modulus of  $[\tilde{\mathbf{r}}_k^{(1)}]_j$  increases with decreasing  $j$ , with the maximum reached around index  $j \approx 2k - k/2$ . At large  $k$  vectors  $[\tilde{\mathbf{r}}_k^{(1)}]_j$  are therefore peaked close to the right edge (Fig. 9); however, including the dependence on  $\alpha$  (63) will cause an exponential decay with  $j$  and therefore  $\tilde{\mathbf{r}}_k$  themselves are localized at the left edge (Fig. 10). Because  $[\tilde{\mathbf{r}}_p]_k$  is nonzero only for  $k \leq 2p$ , and  $[\mathbf{r}_p]_k = [\tilde{\mathbf{r}}_p]_{k-1}$ , we see that we will have a nonzero contribution to the purity  $I_k(t)$ , Eq. (68), only if  $k - 1 \leq 2(\frac{n}{2} - t - 1)$ , i.e.,

$$t \leq t_K = \left\lfloor \frac{n - k - 1}{2} \right\rfloor. \quad (66)$$

For bipartition  $k$  the kernel does not contribute for  $t > t_K$ .

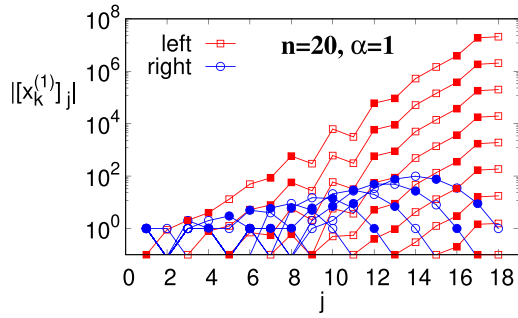


FIG. 9. Absolute values of coefficients of left  $[\tilde{\mathbf{l}}_k^{(1)}]_j$  [Eq. (67), red squares,  $k$  increases from top to bottom] and right  $[\tilde{\mathbf{r}}_k^{(1)}]_j$  [Eq. (65), blue circles,  $k$  decreases from right to left] kernel vectors of  $T$  for  $n = 20$ , where one has  $n/2 - 1 = 9$  vectors, each having  $n - 2 = 18$  components (full symbols indicate negative values).

## 2. Left vectors

We have seen that all  $\mathbf{r}_k$  are  $n$  independent with a compact explicit form. The left vectors  $\tilde{\mathbf{l}}_k$  are, on the other hand, more complicated. Because  $T^T [\tilde{\mathbf{l}}_{n/2-1}^{(1)}] = 0$  the recursion starts with the last index  $k = n/2 - 1$ , for which we have  $\tilde{\mathbf{l}}_{n/2-1}^{(1)} = (0, \dots, 0, -\frac{2}{n}, \frac{2}{n})$ , i.e., nonzero are only the last two components whose value  $\pm \frac{2}{n}$  is fixed by normalization  $\langle \tilde{\mathbf{r}}_{n/2-1}^{(1)} | \tilde{\mathbf{l}}_{n/2-1}^{(1)} \rangle = 1$ . This normalization is what brings in the  $n$  dependence [due to the  $n$ -dependent second-to-last component  $[\tilde{\mathbf{r}}_{n/2-1}^{(1)}]_{n-3} = \pm(n/2 - 1)$ ], which, as we will see, then trickles into an increasingly more complicated  $n$  dependence of subsequent left vectors  $\tilde{\mathbf{l}}_{n/2-k}^{(1)}$ . Due to the structure of  $T^T$  they have nonzero values only in the last  $2k$  components, and they satisfy the recursion  $[\tilde{\mathbf{l}}_{k-1}^{(1)}]_j = [\tilde{\mathbf{l}}_k^{(1)}]_{j+1} - [\tilde{\mathbf{l}}_k^{(1)}]_{j+2}$ . This recursion is enough to get all components of  $\tilde{\mathbf{l}}_{k-1}^{(1)}$  apart from the last two ones, say,  $a$  and  $b$ , which must instead be determined from the biorthogonality  $\langle \tilde{\mathbf{r}}_k^{(1)} | \tilde{\mathbf{l}}_{k-1}^{(1)} \rangle = 0$ , and  $T^T [\tilde{\mathbf{l}}_{k-1}^{(1)}] = [\tilde{\mathbf{l}}_k^{(1)}]$  which for the last component results in  $a + b = [\tilde{\mathbf{l}}_k^{(1)}]_{n-2}$ . This for instance gives the last two vectors,

$$\begin{aligned} \tilde{\mathbf{l}}_{n/2-1}^{(1)} &= (-1)^{n/2} \left( 0, \dots, 0, -\frac{2}{n}, \frac{2}{n} \right) \\ \tilde{\mathbf{l}}_{n/2-2}^{(1)} &= (-1)^{n/2} \left( 0, \dots, 0, \frac{2}{n}, -\frac{4}{n}, -\frac{n^2 - 52}{12n}, \frac{n^2 - 28}{12n} \right). \end{aligned} \quad (67)$$

Polynomials in the numerators of the last two components of  $\tilde{\mathbf{l}}_{n/2-k}^{(1)}$  are of order  $k$  and get more complicated for increasing  $k$ . We were not able to obtain any nice closed form for them, so we just show them in Fig. 9. We can see that they are all localized at the right edge, similarly as  $\alpha$ -dependent vectors  $\mathbf{l}_k$  (Fig. 10).

Once we have left and right kernel vectors we can use  $A_{\text{ker}}^t$  (60) to write down the contribution  $I_k^{\text{ker}}$  to the purity coming from  $A_{\text{ker}}$ . Using a product initial state  $\mathbf{I}(0)$  (17) and Eq. (18) we have

$$I_k^{\text{ker}}(t) = \sum_{p=1}^{n/2-t-1} [\mathbf{r}_p]_k \sum_{j=1}^n [\mathbf{l}_{p+t}]_j. \quad (68)$$

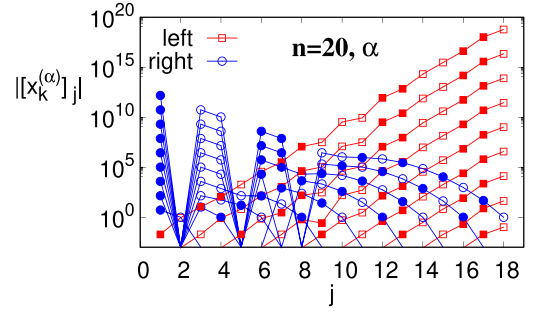


FIG. 10. The same vectors as in Fig. 9, this time for nonzero  $\alpha = 5/26$  ( $d = 5$ ,  $n = 20$ ), showing the non-Hermitian skin effect of Jordan normal form kernel generalized eigenvectors (61). Left are localized at the right edge (red squares) and right (blue circles) at the left edge (for squares  $k$  increases from top to bottom, for circles  $k$  increases from bottom to top).

Observe that, as opposed to  $A_\lambda$ , the time dependence comes in here only implicitly via the number of terms we have in the sum, and in which left vector is multiplied with which right vector. Nevertheless, as we will see, because of extensive size of the kernel, time dependence of  $I_k^{\text{ker}}(t)$  is rather interesting. The sum of components of  $\mathbf{l}_p$  become increasingly more complicated for decreasing  $p$ , see Eq. (67). The first two are for instance  $\langle \mathbf{l}_{n/2-1} | \mathbf{I}(0) \rangle = (-1)^{n/2} (\frac{2}{n} - \frac{4\alpha}{n})$ , and  $\langle \mathbf{l}_{n/2-2} | \mathbf{I}(0) \rangle = (-1)^{n/2} [-\frac{4}{n} - \frac{n^2 - 52}{12n\alpha} + \frac{n^2 - 28}{12n\alpha^2} (1 - \alpha)]$ . Using this two explicit forms we can write down an exact contribution for the two largest times when the kernel still contributes, i.e.,  $I_k^{\text{ker}}(\frac{n}{2} - 2) = [\mathbf{r}_1]_k \langle \mathbf{l}_{n/2-1} | \mathbf{I}(0) \rangle$ , and  $I_k^{\text{ker}}(\frac{n}{2} - 3) = [\mathbf{r}_1]_k \langle \mathbf{l}_{n/2-2} | \mathbf{I}(0) \rangle + [\mathbf{r}_2]_k \langle \mathbf{l}_{n/2-1} | \mathbf{I}(0) \rangle$ . At time  $n/2 - 2$  only bipartitions with  $k = 2, 3$  get the kernel contribution (66), while at  $t = n/2 - 3$  bipartitions  $k = 2, \dots, 5$  do. We do not write all expressions out but to nevertheless give a flavor we get for  $t = n/2 - 2$

$$I_2^{\text{ker}}\left(\frac{n}{2} - 2\right) = \frac{2}{n} (-1)^{n/2} \left(1 - \frac{1}{\lambda_{\text{ph}}}\right), \quad (69)$$

$$I_3^{\text{ker}}\left(\frac{n}{2} - 2\right) = -\frac{2\alpha}{n} (-1)^{n/2} \left(1 - \frac{1}{\lambda_{\text{ph}}}\right). \quad (70)$$

Expressions are relatively simple and similar looking to the spectral contribution in Eq. (37). In fact, evaluating terms in Eq. (37) for nonpositive argument  $p = -1$ , for  $t = n/2 - 2$  there is only one term in the sum having  $r = 0$ ; we get  $\Delta I_2(n/2 - 2) = \frac{4\alpha}{n} (2\alpha)^{-1} f_2(-1) (1/\lambda_{\text{ph}} - 1)$ . Using the value of  $f_2(-1)$  (A15) one gets exactly  $-I_2^{\text{ker}}(\frac{n}{2} - 2)$  (69). A diverging term coming from the spectrum of  $A_\lambda$  for a nonpositive argument  $p = -1$  exactly cancels the kernel contribution. Similar cancellation is observed also for  $\Delta I_3(n/2 - 2)$ , as well as for  $I_{k=2,3,4,5}(n/2 - 3)$ . We have also checked symbolically some other values of  $t$  and  $k$ , as well as numerically all possible  $k$  and  $t$  for a range of system sizes, and we always find an exact cancellation. We conjecture that the kernel contribution to the purity (60) is due to the specific localized form of left and right generalized eigenvectors always exactly equal to the negative of the spectral contribution from  $A_\lambda$  (37) with nonpositive  $p$ . Those terms are obtained for  $r$  in Eq. (37) in

the range

$$r = 0, 1, \dots, r_{\max} = t_K - t. \quad (71)$$

In other words, the kernel contribution is automatically accounted for if we take the summation in Eq. (37) not from  $r = 0$  but rather from  $r_{\min} = r_{\max} + 1 = t_K - t + 1$  (38) such that arguments of  $f_k(p)$  are positive.

We have seen that without taking into account the kernel, we do not get the correct decay for  $t < t_K$  [green pluses in Fig. 1(a)]. To correctly describe the phantom decay at those times one needs the kernel. One can say that the phantom decay  $\lambda_{\text{ph}}$  is somehow encoded in the structure of localized vectors of kernel Jordan normal form. At the same time we have also seen that  $\lambda_{\text{ph}}$  is also encoded in the nonzero spectrum part—for  $t > t_K$  we can get  $\lambda_{\text{ph}}$  also out of  $A_\lambda$  alone. Simple-looking matrix  $A$ , or equivalently  $T$  (15), whose spectrum does not contain any signs of  $\lambda_{\text{ph}}$  that governs the correct asymptotic relaxation rate, has  $\lambda_{\text{ph}}$  encoded in the structure of  $n/2 - 1$  eigenvectors corresponding to nonzero eigenvalues or, independently, also in the structure of the Jordan normal form kernel of size  $n/2 - 1$ . In short, it is not (just) the eigenvalues but rather eigenvectors that are crucial.

### B. Kernel breakdown and the pseudospectrum

In the previous section we have seen how  $A_{\text{ker}}$  and  $A_\lambda$  are both crucial to correctly describe phantom decay for  $t < t_K$ , while at longer times only  $A_\lambda$  is required to recover  $\lambda_{\text{ph}}$ . In both regimes the eigenvectors (not just the eigenvalues) are the ones that determine the relaxation rate.

In Ref. [6] it has been observed that the phantom decay  $\lambda_{\text{ph}}$  is exactly equal to the norm of the pseudospectrum. One definition of a pseudospectrum [36], more precisely  $\varepsilon$ -pseudospectrum, of a matrix  $T$  is that it is equal to the spectrum of a perturbed matrix (supremum of spectra for all  $\|E\| \leq 1$ ),

$$T(\varepsilon) = T + \varepsilon \cdot E. \quad (72)$$

Fixing small  $\varepsilon$  the pseudospectrum is well defined and independent of  $\varepsilon$  in the TDL. For normal matrices the pseudospectrum is equal to the spectrum; however, for non-Hermitian ones the two can be different. How does the pseudospectrum, and therefore  $\lambda_{\text{ph}}$ , emerge from the spectrum of  $A$  under small perturbation?

For a start let us remind ourselves of the standard perturbation theory [55]. A nondegenerate eigenvalue  $\lambda_j$  and eigenvector change as

$$\lambda_j(\varepsilon) = \lambda_j + \varepsilon \cdot \langle \tilde{\mathbf{L}}_j | E | \tilde{\mathbf{R}}_j \rangle + \dots, \quad (73)$$

$$|\tilde{\mathbf{R}}_j(\varepsilon)\rangle = |\tilde{\mathbf{R}}_j\rangle + \varepsilon \sum_{k \neq j} \frac{\langle \tilde{\mathbf{L}}_k | E | \tilde{\mathbf{R}}_j \rangle}{\lambda_j - \lambda_k} |\tilde{\mathbf{R}}_k\rangle + \dots \quad (74)$$

One can bound the eigenvector difference by (assuming  $\|E\| = 1$ )

$$\| |\tilde{\mathbf{R}}_j(\varepsilon)\rangle - |\tilde{\mathbf{R}}_j\rangle \| \leq \varepsilon \cdot \sum_{k \neq j} \frac{\| \tilde{\mathbf{L}}_k \| \cdot \| \tilde{\mathbf{R}}_j \|}{|\lambda_j - \lambda_k|} \| \tilde{\mathbf{R}}_k \|. \quad (75)$$

For Hermitian matrices left and right eigenvectors are the same and normalized and therefore a large change can be

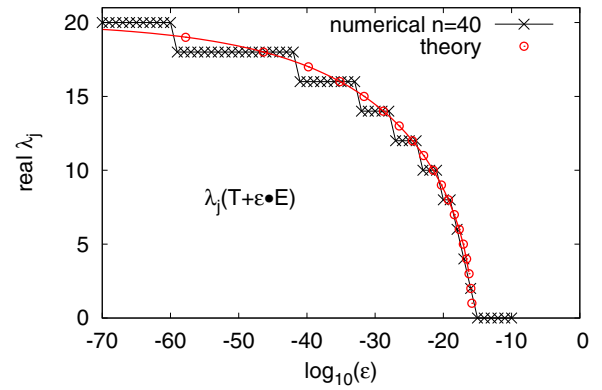


FIG. 11. Number of real eigenvalues of  $T + \varepsilon E$  for a single realization of  $E$  and  $n = 40, d = 5$ . Red curve (circles) is theory given by Eq. (79) without any fitting parameters.

achieved only if some other eigenvalue is close to  $\lambda_j$ —that is why we often focus on spectral properties like gaps. However, we can see that for non-Hermitian matrices there is another option, namely the norms of eigenvectors can diverge and this is exactly what is happening for our  $T$ . However, there is another twist in which the kernel, due to its extensive size  $n/2 - 1$ , plays a crucial role. Namely, while the standard perturbation theory (74) predicts that the effect of perturbation is linear in  $\varepsilon$ , albeit can be large due to localized eigenvectors, a degenerate perturbation theory says that a degenerate eigenvalue  $\tilde{\lambda}_j$  of algebraic multiplicity  $p$  behaves under perturbation in a nonlinear way (Poiseux series) [36,56],

$$\tilde{\lambda}_j(\varepsilon) = \tilde{\lambda}_j + b \varepsilon^{1/p} e^{i2\pi j/p} + \dots \quad (76)$$

In our case the degenerate eigenvalue in question is the zero eigenvalue of the kernel,  $\tilde{\lambda}_j = 0$  with  $j = 1, \dots, p$  and  $p = n/2 - 1$ . Because of this nonlinearity shifts of zero eigenvalues scale as  $\varepsilon^{1/(n/2-1)}$  and will be for large  $n$  much larger than the shift of nondegenerate nonzero eigenvalues (74). Already under very small perturbation the kernel will rapidly explode in  $n/2 - 1$  rotationally symmetric rays (76). The expanding front of eigenvalues will then hit the nonzero  $\lambda_j$ , causing an exceptional point [20], after which the real eigenvalues are ejected into a complex plane, eventually resulting in the pseudospectrum. Namely, for our matrix the pseudospectrum is an oval-shaped curve in the complex plane, e.g., seen as blue points in Fig. 12(a) or in Ref. [6]. It is therefore the nonlinear instability of the kernel under perturbation (76) that will govern the spectrum of  $T(\varepsilon)$ .

We can estimate the perturbation strength  $\varepsilon$  when the  $\lambda_j$  will be ejected from the real axis by

$$\lambda_j = \left( 2\alpha \cos \frac{\pi j}{n} \right)^2 = \varepsilon^{1/(n/2-1)}. \quad (77)$$

The pseudospectrum is therefore expected to be established roughly when

$$\varepsilon \sim (2\alpha)^{n-2}, \quad (78)$$

which is exponentially small in  $n$ . In Fig. 11 we compare numerically computed [57] number of real eigenvalues at

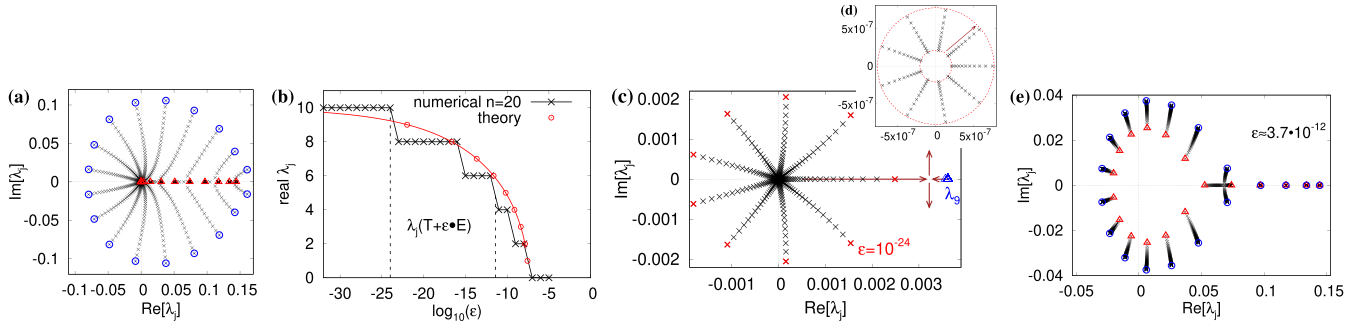


FIG. 12. Spectrum evolution of  $T + \varepsilon \cdot E$  (single instance of  $E$ ,  $n = 20$ ,  $d = 5$ ). (a) Spectrum for  $\varepsilon = 10^{-x}$  and  $x = 5, 5.5, 6, \dots, 60$  (black crosses, with the largest  $\varepsilon = 10^{-5}$  with blue circles and smallest  $\varepsilon = 10^{-60}$  with red triangles). (b) Similarly to Fig. 11 but with  $n = 20$ . Vertical dashed lines mark perturbation  $\varepsilon = 10^{-24}$  when the first collision shown in (c) happens, and  $\varepsilon = 3.7 \times 10^{-12}$  of frame (e). (c) Exploding kernel eigenvalues for small  $\varepsilon$  until just before the first collision with the smallest nonzero  $\lambda_9 = (2\alpha \sin \frac{\pi}{n})^2$  happens (brown arrows). (d) Zoom-in on  $\varepsilon = 10^{-x}$  for  $x = 55, \dots, 60$ , two red circles are  $\varepsilon^{1/9}$  for the smallest and the largest shown perturbation (77). (e) Collision of two eigenvalues around  $\varepsilon \approx 3.7 \times 10^{-12}$ . Shown are spectra for  $\varepsilon = y \times 10^{-13}$  with  $y = 1, 2, \dots, 100$  (dense black crosses, from red triangles for  $y = 1$  to blue circles for  $y = 100$ ). The remaining four real eigenvalues barely move in this range of  $\varepsilon$ .

different  $\varepsilon$  for a single realization of  $T(\varepsilon)$  (72), with  $E$  being a matrix of real Gaussian numbers of zero mean and unit variance, with theory from (77), i.e.,

$$j = \frac{n}{\pi} \arccos \left[ \frac{\varepsilon^{1/(n-2)}}{2\alpha} \right], \quad (79)$$

that predicts the number of real eigenvalues as a function of perturbation strength  $\varepsilon$ . We can observe very good agreement.

In Fig. 12 we illustrate the emergence of the pseudospectrum, that is, follow the changing spectrum of  $T(\varepsilon)$  as  $\varepsilon$  varies, again for a single realization of a Gaussian perturbation matrix  $E$ . We can see that the scenario of expanding kernel eigenvalues described by Eq. (76), taking simply  $b = 1$ , describes collisions between real eigenvalues (77) rather well.

## V. CONCLUSION

We studied the average purity evolution in a random staircase circuit with open boundary conditions. The process can be described by an iteration of a nonsymmetric matrix whose size is equal to the number of qudits  $n$ . Contrary to intuition, the relaxation rate in the thermodynamic limit is not given by the finite gap but rather is slower, that is, slower than any non-steady-state eigenvalue would suggest. Using an exact spectral decomposition we explain how such phantom relaxation rate emerges (i) from a nondegenerate spectrum and in particular due to localized eigenvectors (non-Hermitian skin effect) and (ii) how extensively large Jordan normal form of the kernel, where the generalized eigenvectors are also localized, likewise results in phantom relaxation at short times. In both cases the non-Hermitian skin effect manifests itself in exponentially large expansion coefficients that alternate in signs, almost mutually canceling, and together resulting in a slow decay.

This shows that the time dependence under non-Hermitian matrices can, up to times extensive in  $n$ , result in a behavior that would not be possible under Hermitian dynamics. Such behavior can be caused either by a completely nondegenerate

ordinarily looking spectrum or, alternatively, also by a large nondiagonalizable Jordan block. In both cases it is due to properties of (generalized) eigenvectors. In a non-Hermitian system it is the eigenvectors and a delicate cancellation and not eigenvalues that are important.

A similar conclusion is reached also from another viewpoint. Namely, it is well known in mathematics [36] that the spectrum of a non-Hermitian matrix can be very sensitive to small perturbations. Because this sensitivity can be exponentially large in  $n$ , it is the pseudospectrum rather than the spectrum that one needs to look at. Because of an extensively large Jordan normal form the spectrum depends nonlinearly on the perturbation strength  $\varepsilon$  as  $\varepsilon^{1/(n/2-1)}$ , so that the spectrum changes to the pseudospectrum already at an exponentially small perturbation strength.

It would be interesting to study such phantom relaxation in higher dimensions or in the presence of longer-range couplings where entanglement dynamics can be different. In addition, we expect similar behavior also for quantities different than purity. For instance, OTOC relaxation under the brickwall random circuit with periodic boundary conditions results in a nonsymmetric tridiagonal matrix [58] having similar underlying mathematics. While our exact solution pertains to the specific purity setting and relies on the simplicity of the resulting Toeplitz matrix (15), it should be of use in other non-Hermitian situations. The non-Hermitian skin effect arises because of the asymmetry in  $T$  (15) coming in turn from the asymmetry in relaxation between different bipartitions  $k$  in  $I_k$ : Purities with smaller  $k' < k$  feed into  $I_k$  with the amplitude decaying with  $k - k'$  [see also Fig. 1(b)]. There is an emergent directionality in the relaxation process (i.e., an “arrow of time”). An interesting question is if the specific behavior identified here is more generic in relaxation where one also expects an effective breaking of time reversibility.

## ACKNOWLEDGMENT

I acknowledge support by Grants No. J1-4385 and No. P1-0402 from the Slovenian Research Agency.



### APPENDIX A: PROPERTIES OF $f_k(p)$

We show that (assuming even  $n$ , positive integers  $k$  and  $r$ , and  $\varphi_j = \frac{\pi j}{n}$ )

$$f_k(p) = - \sum_{j=1}^{n/2-1} (-1)^j \cos^p \varphi_j \sin \varphi_j \sin(k\varphi_j),$$

$$f_{2r}(2p-1) = 0 \quad \text{for } p = 1, 2, \dots, \frac{n}{2} - r - 1, \quad (\text{A1})$$

$$f_{2r-1}(2p) = 0 \quad \text{for } p = 1, 2, \dots, \frac{n}{2} - r - 1. \quad (\text{A2})$$

First, we rewrite a product of two sines in terms of cosines, obtaining

$$2f_k(p) = h_{k+1}(p) - h_{k-1}(p),$$

$$h_r(p) = \sum_{j=1}^{n/2-1} (-1)^j \cos^p \left( \frac{\pi j}{n} \right) \cos \left( \frac{r\pi j}{n} \right). \quad (\text{A3})$$

The goal is to show that  $h_k(p)$  is constant in an appropriate range of  $p$  and  $r$ , directly implying Eqs. (A1) and (A2).

#### 1. Even bipartition $k$

Let us consider even  $k = 2r$ , for which we need  $h_{2r\pm 1}(p)$  with odd  $p$ , that is,  $h_{2r-1}(2p-1) = \sum_{j=1}^{n/2-1} (-1)^j b_j(n)$ , with

$$b_j(n) = \cos^{2p-1} \left( \frac{\pi j}{n} \right) \cos \left[ \frac{(2r-1)\pi j}{n} \right]. \quad (\text{A4})$$

One has  $b_{n/2}(n) = 0$  and  $b_{n-j}(n) = b_j(n)$ , and therefore we can extend the sum over  $j$  to  $n-1$ , getting

$$2h_{2r-1}(2p-1) = \sum_{j=1}^{n-1} (-1)^j b_j(n). \quad (\text{A5})$$

Separately summing all even  $j$ , and all odd  $j$ , we can write

$$h_{2r-1}(2p-1) = B_{\frac{n}{2}} - \frac{1}{2} B_n,$$

$$B_m = \sum_{j=1}^{m-1} \cos^{2p-1} \left( \frac{\pi j}{m} \right) \cos \left[ \frac{(2r-1)\pi j}{m} \right]. \quad (\text{A6})$$

Let us consider  $1 + B_m = b_0(m) + B_m$ . The benefit of that is that the sum over  $j = 0, \dots, m-1$  of  $\exp(i\frac{\pi}{m} j 2p)$  is zero for all integer  $p$  except for integer multiples of  $m$ ,

$$w(p) = \sum_{j=0}^{m-1} e^{i\frac{\pi}{m} j 2p} = e^{ip\frac{m-1}{m}} \frac{\sin(p\pi)}{\sin\frac{p\pi}{m}}$$

$$w(p \in \mathbb{Z}) = \begin{cases} m & p = ml, l \in \mathbb{Z} \\ 0 & \text{else} \end{cases}. \quad (\text{A7})$$

Taking  $b_j(m)$  and expressing cosines in terms of an exponential, and using binomial expansion for the  $(2p-1)$ -th power, we have

$$1 + B_m = \frac{1}{2^{2p}} \sum_{t=0}^{2p-1} \binom{2p-1}{t} \{w(t+r-p) + w(t-r-p+1)\}. \quad (\text{A8})$$

Because both arguments of  $w$  are integers, we will have a nonzero term only if one of the two arguments is a multiple of  $m$ . For positive  $r$  and  $p < m-r$  the two arguments can be only 0, but not, e.g.,  $-m$ . This means that either  $t = p-r$  or  $t = r+p-1$ . Taking into account that one needs also to have  $2p-1 \geq t \geq 0$ , which gives  $p \geq r$ , finally results in

$$1 + B_m = \begin{cases} 0, & \text{for } p = 1, \dots, r-1, \\ \frac{m}{2^{2p-1}} \binom{2p-1}{p-r}, & \text{for } p = r, \dots, m-r. \end{cases} \quad (\text{A9})$$

For larger  $p$  one would have a sum of more binomial symbols. The precise form of  $B_m$  is actually not important; what matters is that for  $p = 1, 2, \dots, n-r$  one has  $B_m \propto m$ , from which one immediately gets

$$h_{2r-1}(2p-1) = -\frac{1}{2}, \quad \text{for } p = 1, 2, \dots, \frac{n}{2} - r. \quad (\text{A10})$$

Using Eq. (A3) this proves Eq. (A1).

#### 2. Odd bipartition $k$

Let us write  $k = 2r-1$ , and we need  $h_{2r}(2p)$  in Eq. (A3),

$$h_{2r}(2p) = \sum_{j=1}^{n/2-1} (-1)^j \cos^{2p} \left( \frac{\pi j}{n} \right) \cos \left( \frac{2r\pi j}{n} \right). \quad (\text{A11})$$

By a similar procedure as before we get

$$h_{2r}(2p) = C_{\frac{n}{2}} - \frac{1}{2} C_n,$$

$$C_m = \sum_{j=1}^{m-1} \cos^{2p} \left( \frac{\pi j}{m} \right) \cos \left( \frac{2r\pi j}{m} \right). \quad (\text{A12})$$

A completely analogous procedure as for the even bipartition in this case leads to

$$1 + C_m = \begin{cases} 0, & \text{for } p = 1, \dots, r-1, \\ \frac{m}{2^{2p}} \binom{2p}{p-r}, & \text{for } p = r, \dots, m-r-1, \end{cases} \quad (\text{A13})$$

and, as a consequence,

$$h_{2r}(2p) = -\frac{1}{2}, \quad \text{for } p = 1, 2, \dots, \frac{n}{2} - r - 1, \quad (\text{A14})$$

which proves Eq. (A2).

We have therefore proved that  $f_k(p)$  is for even  $k$  zero at odd  $1 \leq p \leq n-k-3$  (A1), while for odd  $k$  it is zero for even  $2 \leq p \leq n-k-3$  (A2). For every other power  $p$  the sum is nonzero. We have shown that such alternating sums of products of powers of roots of unity have rather magic properties.

All this can be seen in Fig. 13. We can also see that  $f_k(p)$  has very different behavior at negative  $p$ , where it is large and positive. This is exactly the regime that gives a diverging contribution to purity from  $A_\lambda$ , as discussed around Eq. (38), and is precisely canceled by the kernel contribution (Sec. IV). In that context nonpositive values of  $p$  are also of interest. We are not going to study those in any detail, let us just state few values that will be used for the kernel contribution to purity at times  $t = n/2 - 2$  and  $t = n/2 - 3$ , where the arguments are  $p = -3, -2, -1, 0$ . One gets

$$f_2(-1) = (-1)^{n/2}, \quad f_2(-3) = (-1)^{n/2} \frac{n^2 - 4}{6}, \quad (\text{A15})$$

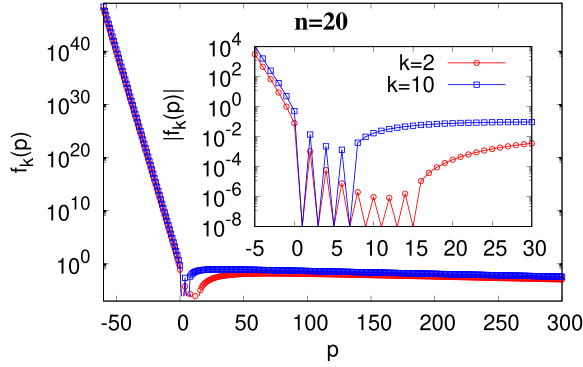


FIG. 13. Values of  $f_k(p)$  for  $n = 20$  and different  $p$ . In the inset one can see that  $f_k(p)$  is zero (dips) for odd  $p = 1, 3, \dots, 15$  at  $k = 2$ , and for  $p = 1, 3, 5, 7$  at  $k = 10$ , as stated in Eq. (A1).

holding for  $n \geq 2$ , then

$$f_3(0) = -\frac{1}{2}(-1)^{n/2}, \quad f_3(-2) = -(-1)^{n/2} \frac{n^2 - 28}{12}, \quad (\text{A16})$$

holding for  $n \geq 6$ , while  $f_4(-1) = -2(-1)^{n/2}$  for  $n \geq 6$ , and  $f_5(0) = \frac{1}{2}(-1)^{n/2}$  for  $n \geq 8$ .

## APPENDIX B: SCALING FUNCTION FOR $k = 2$

To express  $f_2(p)$  in terms of known sums we first write  $f_2(p) = h(p+3) - h(p+1)$ , where  $h(p) = 2 \sum_{j=1}^{n/2-1} (-1)^j \cos^p \varphi_j$ . For large  $p$  the high powers of  $\cos \varphi_j$  will be very small, and one can approximate

$$\cos^p \varphi_j \approx \exp[-(\pi j/n)^2 p/2]. \quad (\text{B1})$$

On top of that we can extend the sum over  $j$  in the definition of  $h(p)$  to infinity because terms with  $j \geq n/2$  are negligible for large  $p$ . This results in a sum that has a form of the Jacobi theta function (49), more precisely

$$h(p) = \vartheta_4(\cos^p \varphi_1) - 1. \quad (\text{B2})$$

In the above expression we have also at the end replaced back a Gaussian approximation of the cosine with a cosine (otherwise one would not get the correct asymptotic limit of  $\lambda_1$ ). We therefore have

$$\begin{aligned} f_2(p) &= \vartheta_4(\cos^{p+3} \varphi_1) - \vartheta_4(\cos^{p+1} \varphi_1) \\ &\approx -\frac{\pi^2}{n^2} \cos^{p+1} \varphi_1 \vartheta_4'(\cos^{p+1} \varphi_1). \end{aligned} \quad (\text{B3})$$

Plugging that into Eq. (43) gets us

$$\Delta\lambda_{\text{eff}} = (2\alpha)^2 c^2 \left[ \frac{\vartheta_4'(c^{2t-n+6})}{\vartheta_4'(c^{2t-n+4})} - 1 \right] \quad (\text{B4})$$

$$\approx -(2\alpha)^2 \left( \frac{\pi}{n} \right)^2 c^{2t-n+6} \frac{\vartheta_4''(c^{2t-n+4})}{\vartheta_4'(c^{2t-n+4})}, \quad (\text{B5})$$

where we abbreviated  $c := \cos \varphi_1 = \cos(\pi/n)$ , and the approximation leading to the second line holds for large  $n$  [due to Taylor expanding  $\vartheta_4'(c^{2t-n+4}c^2)$  and the resulting  $\cos^2 \varphi_1 - 1$ ].

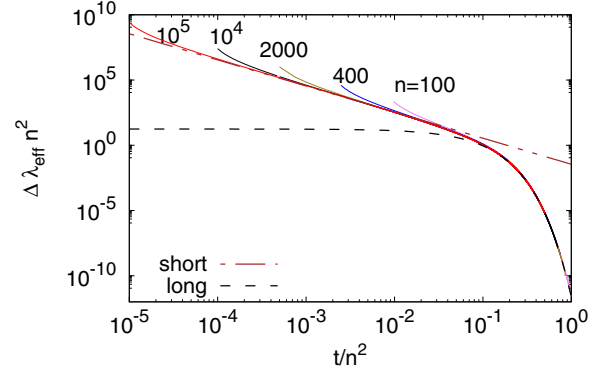


FIG. 14. Long-time scaling of the decay rate  $\Delta\lambda_{\text{eff}}$ , where the scaling variable is  $t/n^2$ . Shown is expression in terms of theta functions, (B5), and the long-time exponential decay of Eq. (B7), both without any fitting parameters. Bipartition  $k = 2, d = 5$ .

Derivatives can be also written out explicitly, resulting in

$$\Delta\lambda_{\text{eff}}(t) \approx -(2\alpha)^2 \left( \frac{\pi}{n} \right)^2 c^2 \frac{\sum_{j=1}^{\infty} (-1)^j j^2 (j^2 - 1) c^{j^2 p}}{\sum_{j=1}^{\infty} (-1)^j j^2 c^{j^2 p}}, \quad (\text{B6})$$

where  $p := 2t - n + 4$ . This form is handy to get the asymptotic form of  $\Delta\lambda_{\text{eff}}(t)$  on the scale  $t \sim n^2$ : Taking the first nonzero term in the numerator, and likewise in the denominator, one gets

$$\Delta\lambda_{\text{eff}}(t) \asymp 12(2\alpha)^2 \left( \frac{\pi}{n} \right)^2 \left( \cos \frac{\pi}{n} \right)^{6t}. \quad (\text{B7})$$

In Fig. 14 we can see that this long-time approximation works well from around  $t/n^2 \gtrsim 0.1$ .

To get the short-time expansion we approach differently. What we need is expansion of  $[\ln \vartheta_4'(q)]'$ , which we can write as  $[\ln \vartheta_4'(q)]' = (\vartheta_4''/\vartheta_4' - \vartheta_4'/\vartheta_4) + \vartheta_4'/\vartheta_4$ . It turns out that for large  $n$  and small  $t$  (but larger than  $n$ ) the first term is much smaller than the second, and we keep only the second one. Using a product form of  $\vartheta_4$

$$\vartheta_4(q) = \prod_{k=1}^{\infty} (1 - q^{2k}) \prod_{j=1}^{\infty} (1 - q^{2j-1})^2, \quad (\text{B8})$$

we can get by explicit derivatives

$$\begin{aligned} -\frac{\vartheta_4'(q)}{\vartheta_4(q)} &= \sum_{j=1}^{\infty} \frac{2jq^{2j-1}}{1 - q^{2j}} + \frac{2(2j-1)q^{2j-1}}{1 - q^{2j-1}} \\ &= \frac{\psi_q^{(1)}(1)}{2q \ln^2 q} + \frac{\psi_q^{(1)}(1) - \psi_q^{(1)}(1 - i\pi/\ln q)}{\ln^2 q}. \end{aligned} \quad (\text{B9})$$

where  $\psi_q^{(1)}(z) = d\psi_q(z)/dz$  and  $\psi_q(z)$  is a  $q$ -digamma function, specifically

$$\psi_q^{(1)}(z) = \ln^2 q \sum_{j=1}^{\infty} \frac{q^{j+z}}{1 - q^{j+z}} + \frac{q^{2j+2z}}{(1 - q^{j+z})^2}. \quad (\text{B10})$$

Remembering that we need Eq. (B9) for  $q = (\cos \frac{\pi}{n})^{2t-n+4}$ , which is for small  $t$  close to 1, we expand denominators in

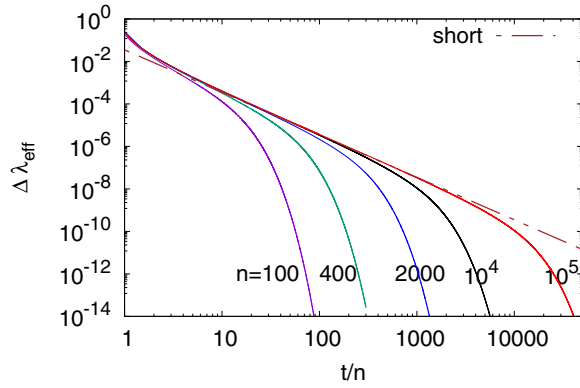


FIG. 15. Short-time scaling of the decay rate  $\Delta\lambda_{\text{eff}}$ , where the scaling variable is  $t/n$ . Shown is expression in terms of theta functions, (B5), and the short-time power law of Eq. (B11) without any fitting parameters. The power law holds from about  $t \approx 5n$ . Bipartition  $k = 2, d = 5$ .

Eq. (B10), finally getting

$$\Delta\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \frac{1}{4} \left(\frac{n}{t}\right)^2. \quad (\text{B11})$$

In Fig. 15 we can see the short-time approximation holds up to larger values of  $t/n$  as one increases  $n$ . One can estimate that it is a good approximation from  $t/n \approx 5$  until almost the time when the long-time approximation starts to hold.

Note that derivatives of theta function that we use (B5) are with respect to  $q$  and not with respect to the more usual  $z$ . However, one can also express everything in terms of derivatives with respect to  $z$ ,

$$\vartheta_4^{(1)}(q, z) := d\vartheta_4(q, z)/dz. \quad (\text{B12})$$

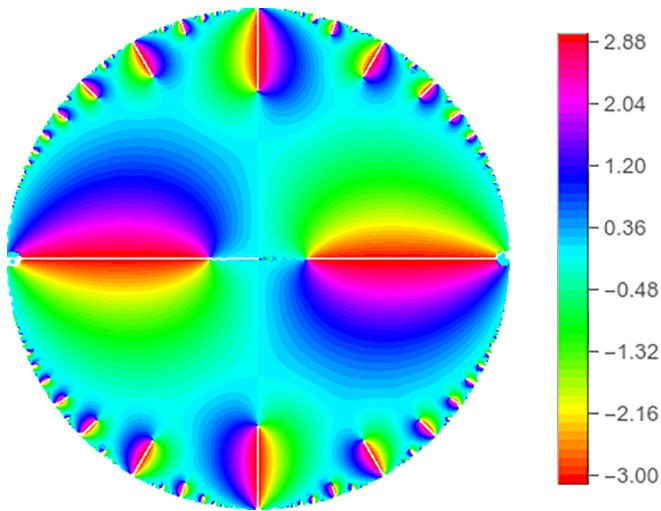


FIG. 16. The phase of  $q\vartheta'_1(z, q)/\vartheta_1(z, q)$  within a unit circle of complex  $|q| < 1$  at  $z = \frac{\pi}{1000}$ . Transition  $\Delta\lambda_{\text{eff}}$  for bipartition with  $k = n/2$  is given by such expression (C2) evaluated at real values of  $q \in [0, 1]$  (visible as white line in figure).

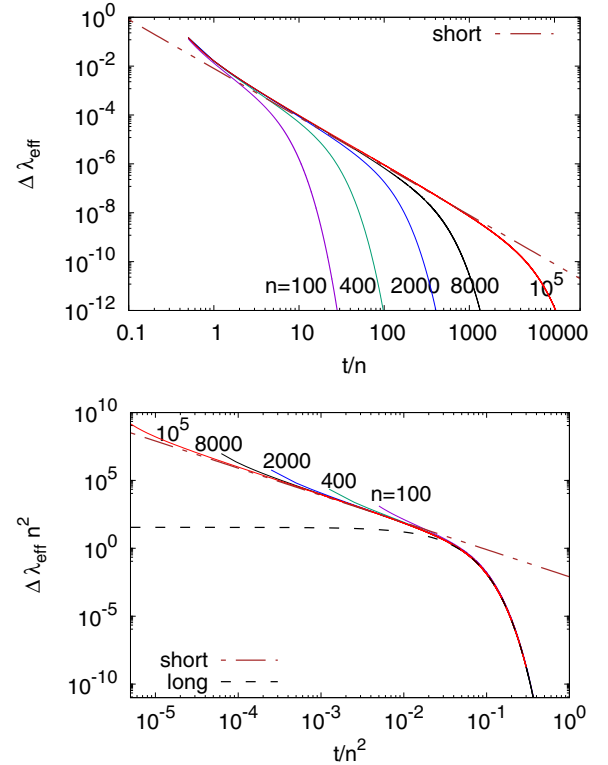


FIG. 17. Scaling functions for  $k = n/2$ , and  $d = 4$ . Top: Short-time scaling of the decay rate  $\Delta\lambda_{\text{eff}}$ , where the scaling variable is  $t/n$ . Shown is expression in terms of theta functions, Eq. (56), and the short-time power law of Eq. (C6). The power law holds from about  $t \approx 2n$  until increasingly larger  $t/n \sim 0.01n$  as one increases  $n$ . Bottom: Long-time scaling where the scaling variable is  $t/n^2$ . Shown is expression in terms of theta functions, Eq. (56), and the long-time exponential decay of Eq. (C5).

Namely, every theta function satisfies “diffusion” equation with  $q$  being an imaginary time,

$$-\frac{1}{4q}\vartheta_4^{(2)} = \vartheta_4'. \quad (\text{B13})$$

Using that Eq. (B5) can be equivalently written as

$$\Delta\lambda_{\text{eff}}(t) = (2\alpha)^2 \left(\frac{\pi}{n}\right)^2 c^2 \left[ 1 + \frac{\vartheta_4^{(4)}(c^{2t-n+4})}{\vartheta_4^{(2)}(c^{2t-n+4})} \right]. \quad (\text{B14})$$

### APPENDIX C: SCALING FUNCTION FOR $k = n/2$

We first observe that in the sum (54) only terms with an odd  $j = 2k + 1$  are nonzero, allowing us to write  $f_{n/2}(p) = \sum_{k=0}^{n/4-1} (-1)^k (\cos \varphi_{2k+1})^p \sin \varphi_{2k+1}$ . Again approximating high power of the cosine with a Gaussian, and letting the sum run to infinity, we get a sum that has a form of the Jacobi theta function [54]  $\vartheta_1(z, q)$  (55), giving

$$f_{n/2}(p) \approx \frac{1}{2} \vartheta_1\left(\frac{\pi}{n}, \left(\cos \frac{\pi}{n}\right)^{4p}\right). \quad (\text{C1})$$

Using Eq. (43) and expanding for large  $n$  results in

$$\begin{aligned}\Delta\lambda_{\text{eff}}(t) &\approx (2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left[ 1 - 4c^{x_0} \frac{\vartheta_1'(\frac{\pi}{n}, c^{x_0})}{\vartheta_1(\frac{\pi}{n}, c^{x_0})} \right] \\ &= (2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left[ 1 + \frac{\vartheta_1^{(2)}(\frac{\pi}{n}, c^{x_0})}{\vartheta_1(\frac{\pi}{n}, c^{x_0})} \right].\end{aligned}\quad (\text{C2})$$

with  $x_0 := 8t - 2n + 4$  and  $c := \cos(\pi/n)$ , and in the second line we also expressed it in terms of derivatives with respect to  $z$ ,  $\vartheta_1^{(2)} := d^2\vartheta_1/dz^2$ . See Fig. 16 for an illustration of the Jacobi theta function in complex  $q$  plane.

By using a product form of  $\vartheta_1$ ,

$$\vartheta_1(z, q) = 2q^{1/4} \sin z \prod_{k=1}^{\infty} (1 - q^{2k}) \prod_{j=1}^{\infty} (1 - 2q^{2j} \cos 2z + q^{4j}), \quad (\text{C3})$$

we can explicitly calculate the required logarithmic derivative, obtaining

$$\left[ 1 + \frac{\vartheta_1^{(2)}(z, q)}{\vartheta_1(z, q)} \right] = \sum_{j=1}^{\infty} \frac{16q^{2j}}{(1 - q^{2j})^2} \cos(2zj) + \frac{8jq^{2j}}{1 - q^{2j}}. \quad (\text{C4})$$

The benefit if this expression, compared to just plugging definition  $\vartheta_1$  (55) into Eq. (C2), is that we have a single

sum instead of a ratio of two sums, making extracting the asymptotic behavior simpler. Namely, for  $t \sim n^2$  the argument  $q$  is very small, and the whole sum can be approximated by its first term, which is  $\approx 24q^2$ , resulting in the long-time approximation

$$\Delta\lambda_{\text{eff}}(t) \asymp 24(2\alpha)^2 \left(\frac{\pi}{n}\right)^2 \left(\cos \frac{\pi}{n}\right)^{16t}. \quad (\text{C5})$$

At short times, that is, on a scale  $t \sim n$ ,  $q$  is, on the other hand, close to 1, and one can expand  $1 - q^{2j} \approx (\pi/n)^2 8tj$ . The second sum in Eq. (C4) turns into a simple geometric sum, evaluating to  $q^2 n^4 / (8\pi^4 t^2) \approx n^4 / (8\pi^4 t^2)$ , while the first sum is of type  $\sum_j q^{2j} / j^2$ , summing to  $n^4 / (4\pi^4 t^2) \text{Li}_2(q^2) \approx n^4 / (4\pi^4 t^2) \pi^2 / 6$ , where  $\text{Li}_n(z) = \sum_{j=1}^{\infty} z^j / j^n$  is a polylogarithm. The specific dilogarithm is  $\text{Li}_2(1) = \zeta(2) = \pi^2 / 6$ , giving short-time form

$$\Delta\lambda_{\text{eff}}(t) \approx (2\alpha)^2 \left( \frac{1}{24} + \frac{1}{8\pi^2} \right) \left( \frac{n}{t} \right)^2. \quad (\text{C6})$$

We again see that the short-time scaling variable is  $t/n$  (C6), while the long-time scaling variable is  $t/n^2$  (C5). Figure 17 demonstrated that both short- and long-time approximations work well.

- 
- [1] D. Braun, *Dissipative Quantum Chaos and Decoherence* (Springer, Berlin, 2001).
  - [2] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, UK, 2011).
  - [3] F. Haake, *Quantum Signatures of Chaos* (Springer, Berlin, 2010).
  - [4] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).
  - [5] J. Bensa and M. Žnidarič, Fastest Local Entanglement Scrambler, Multistage Thermalization, and a Non-Hermitian Phantom, *Phys. Rev. X* **11**, 031019 (2021).
  - [6] M. Žnidarič, Solvable non-Hermitian skin effect in many-body unitary dynamics, *Phys. Rev. Res.* **4**, 033041 (2022).
  - [7] J. Bensa and M. Žnidarič, Purity decay rate in random circuits with different configurations of gates, *Phys. Rev. A* **107**, 022604 (2023).
  - [8] J. Bensa and M. Žnidarič, Two-step phantom relaxation of out-of-time-ordered correlations in random circuits, *Phys. Rev. Res.* **4**, 013228 (2022).
  - [9] M. Žnidarič, Two-step relaxation in local many-body Floquet-systems, *arXiv:2301.06395*.
  - [10] P. W. Claeys, J. Herzog-Arbeitman, and A. Lamacraft, Correlations and commuting transfer matrices in integrable unitary circuits, *SciPost Phys.* **12**, 007 (2022).
  - [11] M. A. Rampp, R. Moessner, and P. W. Claeys, From Dual Unitarity to Generic Quantum Operator Spreading, *Phys. Rev. Lett.* **130**, 130402 (2023).
  - [12] F. Song, S. Yao, and Z. Wang, Non-Hermitian Skin Effect and Chiral Damping in Open Quantum Systems, *Phys. Rev. Lett.* **123**, 170401 (2019).
  - [13] T. Mori and T. Shirai, Resolving a Discrepancy between Liouvillian Gap and Relaxation Time in Boundary-Dissipated Quantum Many-Body Systems, *Phys. Rev. Lett.* **125**, 230604 (2020).
  - [14] T. Mori, Metastability associated with many-body explosion of eigenmode expansion coefficients, *Phys. Rev. Res.* **3**, 043137 (2021).
  - [15] T. Haga, M. Nakagawa, R. Hamazaki, and M. Ueda, Liouvillian Skin Effect: Slowing Down of Relaxation Processes without Gap Closing, *Phys. Rev. Lett.* **127**, 070402 (2021).
  - [16] T. Mori and T. Shirai, Symmetrized Liouvillian Gap in Markovian Open Quantum Systems, *Phys. Rev. Lett.* **130**, 230404 (2023).
  - [17] G. Lee, A. McDonald, and A. Clerk, Anomalous large relaxation times in dissipative lattice models beyond the non-Hermitian skin effect, *Phys. Rev. B* **108**, 064311 (2023).
  - [18] S. Yao and Z. Wang, Edge States and Topological Invariants of Non-Hermitian Systems, *Phys. Rev. Lett.* **121**, 086803 (2018).
  - [19] E. J. Bergholtz, J. C. Budich, and F. K. Kunst, Exceptional topology of non-Hermitian systems, *Rev. Mod. Phys.* **93**, 015005 (2021).
  - [20] Y. Ashida, Z. Gong, and M. Ueda, Non-Hermitian physics, *Adv. Phys.* **69**, 249 (2020).
  - [21] L. E. F. Foa Torres, Perspective on topological states of non-Hermitian lattices, *J. Phys.: Mater.* **3**, 014002 (2020).
  - [22] N. Okuma and M. Sato, Non-Hermitian topological phenomena: A review, *Ann. Rev. Condens. Matter Phys.* **14**, 83 (2023).
  - [23] R. Lin, T. Tai, L. Li, and C. H. Lee, Topological non-Hermitian skin effect, *Front. Phys.* **18**, 53605 (2023).



- [24] N. Okuma and M. Sato, Hermitian zero modes protected by nonnormality: Application of pseudospectra, *Phys. Rev. B* **102**, 014203 (2020).
- [25] N. Okuma and M. Sato, Quantum anomaly, non-Hermitian skin effects, and entanglement entropy in open systems, *Phys. Rev. B* **103**, 085428 (2021).
- [26] T. Yoshida, Real-space dynamical mean field theory study of non-Hermitian skin effect for correlated systems: Analysis based on pseudospectrum, *Phys. Rev. B* **103**, 125145 (2021).
- [27] V. P. Flynn, E. Cobanera, and L. Viola, Topology by Dissipation: Majorana Bosons in Metastable Quadratic Markovian Dynamics, *Phys. Rev. Lett.* **127**, 245701 (2021).
- [28] E. Lubkin, Entropy of an  $n$ -system from its correlation with a  $k$ -reservoir, *J. Math. Phys.* **19**, 1028 (1978).
- [29] If one would have some external noise, like, e.g., measurements, then the purity could decrease even without any partial trace. Such situation is not studied in the present work.
- [30] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, Generic Entanglement can be Generated Efficiently, *Phys. Rev. Lett.* **98**, 130502 (2007).
- [31] We shall use a bold notation  $\mathbf{x}$  for vectors,  $[\mathbf{x}]_k$  for the  $k$ th component in a given basis, and  $|\mathbf{x}\rangle$  for a basis-independent notation.
- [32] W.-T. Kuo, A. A. Akhtar, D. P. Arovas, and Y. Z. You, Markovian entanglement dynamics under locally scrambled quantum evolution, *Phys. Rev. B* **101**, 224202 (2020).
- [33] M. Žnidarič, Exact convergence times for generation of random bipartite entanglement, *Phys. Rev. A* **78**, 032324 (2008).
- [34] Note that the correct thermodynamic limit means one needs to first take the system size to  $\infty$ , only then  $t \rightarrow \infty$ .
- [35] S. H. Weintraub, *Jordan Canonical Form: Theory and Practice* (Morgan & Claypool, 2009).
- [36] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra* (Princeton University Press, Princeton, NJ, 2005).
- [37] K. Życzkowski and M. Kuś, Random unitary matrices, *J. Phys. A* **27**, 4235 (1994).
- [38] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, *Science* **302**, 2098 (2003).
- [39] X. Mi *et al.*, Information scrambling in quantum circuits, *Science* **374**, 1479 (2021).
- [40] T. Liu *et al.*, Observation of entanglement transition of pseudo-random mixed states, *Nat. Commun.* **14**, 1971 (2023).
- [41] A. W. Harrow and R. Low, Random quantum circuits are approximate 2-designs, *Commun. Math. Phys.* **291**, 257 (2009).
- [42] A. Nahum, S. Vijay, and J. Haah, Operator Spreading in Random Unitary Circuits, *Phys. Rev. X* **8**, 021014 (2018).
- [43] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator Hydrodynamics, OTOC, and Entanglement Growth in Systems without Conservation Laws, *Phys. Rev. X* **8**, 031058 (2018).
- [44] A. Chan, A. De Luca, and J. T. Chalker, Solution of a Minimal Model for Many-Body Quantum Chaos, *Phys. Rev. X* **8**, 041019 (2018).
- [45] V. Khemani, A. Vishwanath, and D. A. Huse, Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws, *Phys. Rev. X* **8**, 031057 (2018).
- [46] T. Zhou and A. Nahum, Emergent statistical mechanics of entanglement in random unitary circuits, *Phys. Rev. B* **99**, 174205 (2019).
- [47] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anti-concentrate in log depth, *PRX Quant.* **3**, 010333 (2022).
- [48] A. Nahum, S. Roy, S. Vijay, and T. Zhou, Real-time correlators in chaotic quantum many-body systems, *Phys. Rev. B* **106**, 224310 (2022).
- [49] The last component of vector  $\mathbf{I}$  corresponds to  $I_n(t)$  which is always 1,  $I_n(t) = 1$ , while the first component corresponds to  $I_0(t)$ , which is likewise  $I_0(t) = 1$ . Purity  $I_1(t)$  on the other hand does not couple back to  $I_k(t)$  with  $k = 2, \dots, n-1$ , and can be obtained from a stand-alone recursion  $I_1(t+1) = \alpha + \alpha I_2(t)$  [6].
- [50] Matrix Market, <https://math.nist.gov/MatrixMarket/>.
- [51] J. F. Grcar, *Operator Coefficient Methods for Linear eEquations*, Technical Report SAND89-8691, Sandia National Labs (1989).
- [52] G. Fairweather, On the eigenvalues and eigenvectors of a class of Hessenberg matrices, *SIAM Rev.* **13**, 220 (1971).
- [53] Components of right steady state with odd symmetry are  $[\mathbf{R}'_0]_{n/2-k} = \alpha^{n/2-k} p_{2k}/p_n$  for  $k = 1, \dots, n/2-2$ ,  $[\mathbf{R}'_0]_{n/2+k} = -\alpha^{n/2-k} p_{2k}/p_n$  for  $k = 1, \dots, n/2-1$ , and  $[\mathbf{R}'_0]_{n/2} = 0$ ,  $[\mathbf{R}'_0]_1 = 1$ ,  $[\mathbf{R}'_0]_n = -1$ , and where  $p_n = \sum_{k=0}^{\lceil \frac{n}{2} \rceil - 1} (-1)^k \alpha^{2k} \binom{\frac{n}{2}-k-1}{k}$ .
- [54] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
- [55] J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).
- [56] T. Kato, *Perturbation Theory of Linear Operators* (Springer Verlag, Berlin, 1995).
- [57] In order to get accurate numerical spectra for small  $\varepsilon$  one should avoid using numerical diagonalization with floating-point numbers. What we do is first write matrix elements of  $T(\varepsilon)$  as rational numbers (with huge numerators and denominators to be accurate with precision much better than  $\varepsilon$ ) and then calculate the spectrum for this rational matrix via a characteristic polynomial. While such calculation becomes costly for larger  $n$ , it is still feasible for, e.g.,  $n = 20$  or 40.
- [58] J. Bensa and M. Žnidarič (unpublished).