Sachdev-Ye-Kitaev model: Non-self-averaging properties of the energy spectrum

Richard Berkovits
Department of Physics, Jack and Pearl Resnick Institute, Bar-Ilan University, Ramat-Gan 52900, Israel

(Received 29 August 2022; accepted 18 January 2023; published 24 January 2023)

The short time (large energy) behavior of the Sachdev-Ye-Kitaev model (SYK) is one of the main reasons for the growing interest garnered by this model. True chaotic behavior sets in at the Thouless time, which can be extracted from the energy spectrum. In order to do so, it is necessary to unfold the spectrum, i.e., to filter out global tendencies. Using a simple ensemble average for unfolding results in a parametically low estimation of the Thouless energy. By examining the behavior of the spectrum as the distribution of the matrix elements is changed into a log-normal distribution, it is shown that the sample-to-sample level spacing variance determines this estimation of the Thouless energy. Using the singular value decomposition method, which filters out these global sample-to-sample fluctuations, the Thouless energy becomes parametrically much larger, essentially of the order of the band width. It is shown that the SYK model is non-self-averaging even in the thermodynamic limit which must be taken into account in considering its short time properties.

DOI: 10.1103/PhysRevB.107.035141

I. INTRODUCTION

The interplay between disorder and interactions in quantum systems has been a central theme in condensed-matter physics for the last half century. Recently the Sachdev-Ye-Kitaev (SYK) model [1,2] has garnered much interest in the fields of quantum gravity and quantum field theory [3,4]. A key feature of the model is that it follows random matrix theory (RMT) behavior, as is manifested in the chaotic behavior of its energy spectrum.

Two-body random interaction models have a long history [5]. The SYK model first appeared in the context of spin liquids [1] and then in string theory [6] and quantum gravity [4]. The SYK model is known to be chaotic [7–9], showing a Wigner-like behavior of the energy spectra. Once the SYK model is perturbed by a single-body random term which mimics diagonal disorder in the Anderson model [10–16], or several SYK dots are coupled by single-body random terms [17–21], a transition from metallic (chaotic) to insulating behavior occurs which leads to a Wigner-to-Poisson crossover of the statistical properties of the spectra.

While studying nuclear and condensed-matter systems it became clear that many physical systems exhibit universal behavior at long times (short energy scales) [22–27]. Nevertheless, universality may break at shorter times (large energy scales) for which a particle has no time to sample the entire phase space of the system and its behavior depends on local nonuniversal features [28]. Thus, in the context of disordered metals the scale for which the metallic spectrum deviates from the universal behavior is known as the Thouless energy, $E_{\rm Th} = \hbar D/\tilde{L}^2 = g\Delta$ (D is the diffusion constant, $\tilde{L}$ is the linear dimension, g is the dimensionless conductance, and Δ is the average level spacing), and the Thouless time, $t_{\rm Th} = \hbar/E_{\rm Th} = \tilde{L}^2/D$.

The question whether an analog of the Thouless energy manifests itself in the SYK model has surfaced in Ref. [7], where García-García and Verbaarschot have studied (among other things) the variance of the number of levels as a function of the size of an energy window E, denoted by $\langle\delta^2 n(E)\rangle$ (where $\langle\ldots\rangle$ represents an ensemble average and $n(E)$ is the number of levels within E). A departure from the RMT behavior is apparent above a certain value of E, which quite naturally was identified with the Thouless energy. Moreover, at larger energy windows, the number variance adopts a quadratic form, $\langle\delta^2 n(E)\rangle = a + b\langle n(E)\rangle^2$. Evidence for the Thouless energy has also been seen for other measures such as off-diagonal expectation values [29] and the spectral form factor [30]. For the latter the unfolding procedure has influence on determining the time for which the spectral form factor transits into a linear increase and then into a plateau determining the transition into RMT behavior. In Ref. [20] the origin of the Thouless energy for the SYK model was identified as the relaxation of modes prevailing at shorter scale. For disordered metals these modes are know as the diffusion modes. For the SYK model it was suggested that similar modes can be constructed, where the number of such modes is connected to the number of independent interaction terms in the SYK model, which is much smaller than the size of the Hilbert space. This leads to an energy scale which determines the Thouless energy. Thus, the energy scale for the departure of the SYK model level number variance from the RMT prediction is determined by the scale of the sample-to-sample fluctuations of the ensemble.

In the SYK model all the terms have just a single variance scale, which is well behaved (box or Gaussian) and therefore the origin of the additional energy scale determining the Thouless energy is not obvious. As noted recently by Jia and Verbaarschot [31] this deviation from RMT behavior has its root in large-scale sample-to-sample fluctuations of the spectrum. They attribute these fluctuations to the relatively small number of independent random variables contributing to the SYK Hamiltonian. The basic premises of self-averaging

quantities is that averaging a quantity over an ensemble of realizations or averaging over a single large realization will lead to the same result. If not, the system is not self-averaging. For detecting the universal behavior of the fluctuations of the energy spectrum, one must remove the nonuniversal model-dependent band structure. For a self-averaging model the procedure is straight forward, just average the nearest-neighbor level spacing around a particular level over all the realizations in the ensemble. For the SYK model, comparing the averaged level spacing to the level spacing of a particular sample shows (see, e.g., Fig 4) an almost constant deviation between them which persists for large portions of the energy spectrum. Parametrizing these long-range deviations between the ensemble average and particular realizations by terms of Q-Hermite orthogonal polynomials and removing these deviations from the spectrum results in a pure RMT behavior which is retained up to a very large energy scale.

The influence could be quantified [31] by estimating the energy scale for which the small number of independent random variables will change the level number variance. Using the notation for the complex SYK (CSYK) half-filled model defined in the next section, where L is the number of sites, the size of the Hilbert space is $\binom{L}{L/2} \sim 2^L$. This is equivalent to the usual formulation of the SYK model with $2L$ Majorana fermions with four-body interactions. The number of independent variables is $\binom{2L}{4} \sim L^4$, leading to a variance of L^{-2} in any observable. Thus, one expects the number of levels n to deviate significantly $O(1)$ from RMT predictions on a scale of $n \sim L^2$. A similar result emerges from the calculation in Ref. [20] where the coefficient b in the number variance was estimated as $b \sim L^{-4}$, thus becoming significant at $n \sim L^2$.

Although the deviation from RMT is shifted to larger n as the system size L increases, the proportion of levels following RMT predictions out of the total number of states goes to zero as $L^2/2^L$. On the other hand, one would expect that after filtering out the sample-specific large-scale deviation from the ensemble-averaged density of states (level spacings) the RMT behavior will be followed for a finite portion of the spectrum. Thus, one expects the Thouless energy to crucially depend on whether one simply averages over an ensemble or takes into account the sample-specific large-scale deviations. This is a hallmark of a non-self-averaging system [32].

Hence, the energy scale for which the spectrum departs from the RMT predictions crucially depends on the unfolding, i.e., the method by which the average over the density of states is performed. Estimating the local density of states by a simple ensemble average will give a different value than an unfolding method that is able to take into account sample-specific global behavior of the spectrum. In recent studies it has been shown [33–39] that these sample-specific long-ranged features of the spectrum can be identified by the singular value decomposition (SVD) procedure. The basic idea behind the procedure is to arrange the energy spectrum of the ensemble of different realizations as a matrix, where each row is P consecutive eigenvalues of a given realization, and the number of rows is the number of realizations M. As detailed in Sec. V, SVD is essentially a procedure by which the $M \times P$ matrix is written as a sum over a series of amplitudes multiplied by $M \times P$ matrices which are constructed by an outer product of two different vectors of sizes P and M. The vector of size P is in essence the averaged correction to the spectra of all realizations for this term in the series, while the M terms of the second vector attenuate this averaged correction for each of the M different realizations, and the amplitude gives the overall weight of this term in the series. Thus, each matrix in the series is a product of $P + M$ numbers, compared to the PM numbers composing the original data. Summing up m terms in the series gives the matrix closest to the original one for the summation of matrices constructed by outer products of two vectors. Generally, plotting the amplitudes from large to small (known as a Scree plot) shows that the largest amplitudes [usually $O(1)$ modes] are orders of magnitude larger than the rest, while the following amplitudes obey a power law. The largest terms depict the global behavior of the energy spectrum, while the terms whose amplitudes follow a power law capture the shorter-range properties. Thus SVD is a very natural method to examine the SYK spectrum behavior, and uncover the universal properties of the spectrum, obscured by realization-specific deviations.

One may conclude that there are two possible energy scales for the deviation of the spectrum of the SYK model from RMT predictions. The first is the energy scale for which the number variance of the spectrum unfolded by the ensemble-averaged level spacing deviates from the RMT prediction, which for self-averaging systems such as disordered metals is the Thouless energy. The second energy scale corresponds to the energy for which the number variance diverges from RMT predictions where the unfolding is performed with methods such as the Q-Hermite orthogonal polynomials [31] or by SVD [33–39], or probably by other methods which factor in global sample-specific fluctuations. For a non-self-averaging system such as the SYK model, these two energy scales are not equal, and for clarity we retain the notation of Thouless energy, E_{Th}, for the case where the spectrum is unfolded by an ensemble average over all realizations, while the second scale is denoted by E_{Th^*}.

Here we show that one can tweak the behavior of the CSYK model to a more non-self-averaging behavior by changing the distribution of the off-diagonal terms to a log-normal distribution. Thus, it is possible to enhance the non-self-averaging behavior and study its influence on E_{Th} and E_{Th^*}. Such a wide distribution was not previously considered for the SYK model and should help to clarify the divergence of E_{Th^*} from E_{Th}.

The paper is arranged as follows: the CSYK model is defined in Sec. II. Corroborating the expected behavior for short-range energy scales is performed in Sec. III. The universal statistics fourfold symmetry as a function of the system size is observed. In Sec. IV long-range energy scales are probed by the number variance. The spectrum is unfolded by using the local ensemble-averaged level spacing. RMT predictions hold only up to E_{Th}, which becomes smaller as the log-normal distribution acquires a thicker tail towards larger values. Above E_{Th} the variance increases quadratically as a function of the number of levels in the energy window. In Sec. V we switch to the SVD analysis. This analysis reveals that each realization has a level spacing structure related to the ensemble-averaged level spacing with sample-specific

adjustments. Thus, one should adapt the unfolding for each realization. After a realization-adapted unfolding, the number variation follows the RMT prediction for much larger energy scales, i.e., $E_{\mathrm{Th}^*} \gg E_{\mathrm{Th}}$. Actually, E_{Th} could be estimated from the sample-to-sample variance in the level spacing. In Sec. VI these results are discussed in the limit of large CSYK systems, showing that in the thermodynamic limit the SYK model is non-self-averaging.

II. COMPLEX SYK MODEL

Here we use the complex spinless fermion version of the SYK model given by the following Hamiltonian:

$$\hat{H} = \sum_{i>j>k>l}^{L} V_{i,j,k,l}\hat{c}_i^\dagger \hat{c}_j^\dagger \hat{c}_k \hat{c}_l, \tag{1}$$

the couplings $V_{i,j,k,l}$ are complex numbers, where the real and imaginary components are independently drawn from an identical distribution. We study here two different distributions: The first is a box distribution between $-L^{-3/2}/2 \ldots L^{-3/2}/2$, where L is the number of sites. The second distribution is the log-normal distribution

$$P(V) = (A/|V|)e^{\frac{-\ln^2(|V|/V_{\mathrm{typ}})}{2p\ln(V_{\mathrm{typ}}^{-1})}}, \tag{2}$$

with $V_{\mathrm{typ}} \sim K^{-\gamma/2}$, K is the size of the Hilbert space, A is a normalization, and γ and p are parameters. For simplicity, here we set $p = 0.5$, while γ is varied [40]. Thus, the log-normal distribution becomes wider and more skewed as γ increases. The number of fermions is conserved and we considered the $N = L/2$ sector for even L and the $N = (L+1)/2$ sector for odd L, resulting in a Hilbert space size of $K = \binom{L}{N}$ and a matrix size of $K \times K$.

III. SHORT ENERGY SCALES

As a first step we would like to probe the nearest-neighbor level spacing statistics of the CSYK box distribution in order to establish the extended regime of this model. One expects that in the extended-regime short energy scales (long times) follow the Wigner statistics. The short energy scale statistics is revealed by the ratio statistics, defined as

$$r = \left\langle \min\left(r_n, r_n^{-1}\right)\right\rangle, \quad r_n = \frac{E_n - E_{n-1}}{E_{n+1} - E_n}, \tag{3}$$

where E_n is the nth eigenvalue of the Hamiltonian and $\langle \ldots \rangle$ is an average over an ensemble of different realizations of disorder and half of the eigenvalues around the middle of the band. For the Wigner distribution $r_s \cong 0.5307$ for the GOE symmetry, $r_s \cong 0.5996$ for the GUE symmetry, and $r_s \cong 0.6744$ for the GSE symmetry [41].

An interesting behavior emerges for the CSYK model. It is known that as a consequence of the symmetries of the SYK model, the spectrum of Eq. (1) shows statistics which depend on L [7–9]. For $L \mod 4 = 0$ the statistics are GOE, For $L \mod 4 = 2$ the statistics are GSE, and for $L \mod 4 = 1$ and 3 the statistics are GUE. This is indeed seen in Fig. 1 where r_s averaged over the middle half of the energy spectrum, for different sizes $L = 8, 9, 10, 11, 12, 13, 14, 15, 16$, and 17 and

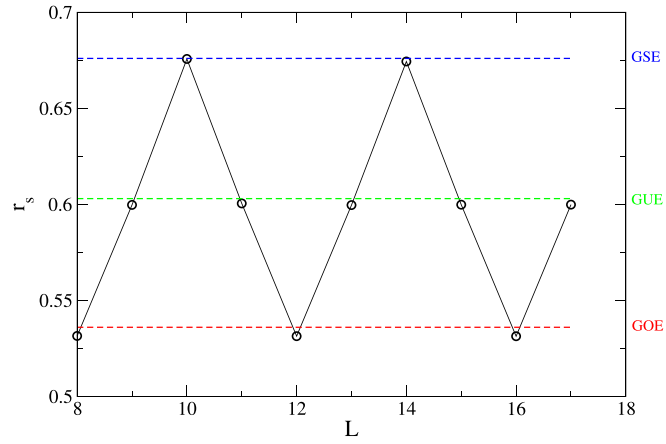

FIG. 1. The nearest-neighbor level spacing statistics as manifested in the behavior of the ratio statistics r_s for different system sizes L of the CSYK model with a box distribution are indicated by the black circles. The number of fermions is $N = L/2$ for even L and $N = (L+1)/2$ for odd L. The r_s values expected for GOE (dashed red), GUE (dashed green), and GSE (dashed blue) are marked. The expected fourfold symmetry is seen.

numbers of fermions $N = 4$, 5, 5, 6, 6, 7, 7, 8, 8, and 9, is plotted. In all cases we exactly diagonalize the corresponding $K \times K$ matrix and average over 3000 realizations of disorder (except for the largest size for which only 100 realizations were computed). It can be seen that the expected fourfold symmetry is followed.

Concentrating on the $L = 16$ with $N = 8$ systems, we investigate the role played by changing the distribution from the box distribution to a log-normal distribution. Setting $p = 0.5$ and increasing γ we sweep through the values $\gamma = 1$, 1.5, 2, 2.5, 3, 4, and 5. For all these values $r_s \cong 0.5307 \pm 0.001$. Thus, GOE universal behavior on short energy scales is perfectly followed.

IV. LOCAL ENSEMBLE UNFOLDING

As discussed, the practice of determining the Thouless energy is fraught with difficulties. In order to compare any spectrum to RMT predictions, one must recast the spectrum such that it will exhibit an averaged constant density of states, i.e., a constant level spacing throughout the region examined. What is the averaging procedure? Usually, one averages the level spacing over an ensemble of disordered realizations in a given energy region and then reconstructs a particular spectrum, such that the level spacing is on the average equal to 1 everywhere. Specifically, the averaged level spacing for the ith level is $\Delta_i = \langle E_{i+p} - E_{i-p}\rangle/2p$ [where p is $O(1)$, here chosen as $p = 5$], and the unfolded spectrum for the jth realization is $\varepsilon_i^j = \varepsilon_{i-1}^j + (E_i^j - E_{i-1}^j)/\Delta_i$. For brevity we call this unfolding procedure *local ensemble unfolding*.

Implementation of this local unfolding procedure for the level number variance of the CSYK model with a box distribution and for log-normal distributions with $\gamma = 1$ and 1.5 results in the behavior depicted in the inset of Fig. 2, which is in agreement with the behavior observed in Refs. [7,31]. Here the number variance begins by following the GOE predictions

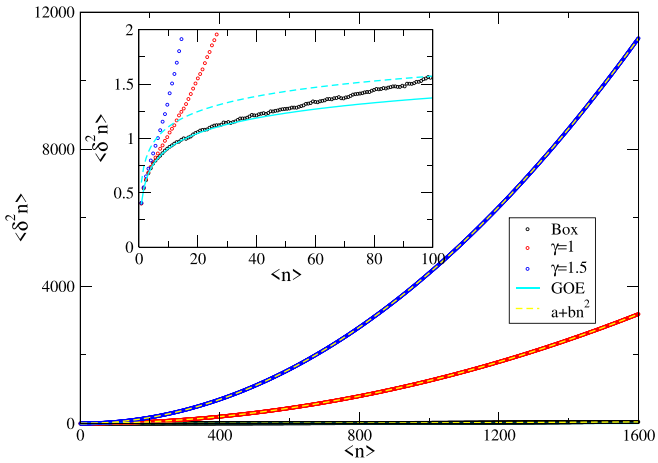

FIG. 2. The level number variance, $\langle \delta^2 n(E) \rangle$, as a function of the energy window E. Symbols (black, box; red, $\gamma = 1$; blue, $\gamma = 1.5$) represent the variance with local ensemble unfolding, fitted for larger values by $a + b\langle n(E) \rangle^2$ (where $b = 1.78 \times 10^{-5}$ for the box distribution, $b = 1.25 \times 10^{-3}$ for $\gamma = 1$, and $b = 4.4 \times 10^{-3}$ for $\gamma = 1.5$). The cyan line is the GOE prediction $\langle \delta^2 n(E) \rangle = (2/\pi^2) \ln(\langle n(E) \rangle) + 0.44$. Insert: Zoom into smaller values of $\langle n \rangle$. Deviations from the GOE behavior are observed. A curve depicting GOE plus a constant of 0.2 corresponds to the dashed cyan line. Using the intersection between the variance and the dashed cyan line to determine the Thouless energy results in $E_{\rm Th} \sim 95\Delta$ for the box distribution, $E_{\rm Th} \sim 13\Delta$ for $\gamma = 1$, and $E_{\rm Th} \sim 7\Delta$ for $\gamma = 1.5$.

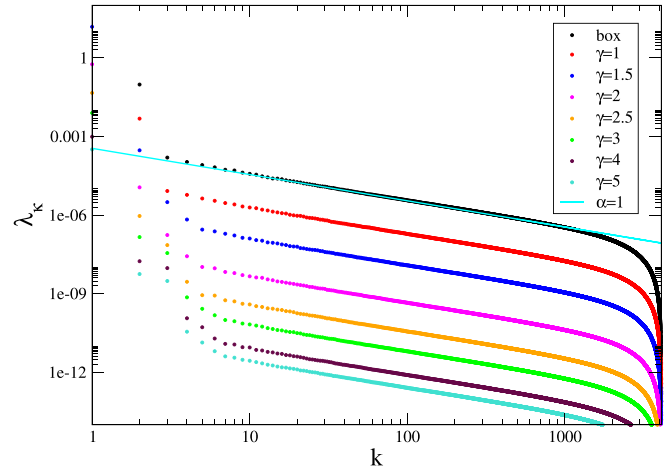

FIG. 3. The scree plot of the singular values for the CSYK model with box and log-normal distributions with $p = 0.5$ and $\gamma = 1$, 1.5, 2, 2.5, 3, 4, and 5 for $L = 16$ and $N = 8$, with $M = 4096$ realizations and $P = 4096$ eigenvalues around the middle of the band. The square amplitudes of the singular values λ_k are indicated by the symbols. The cyan line corresponds to $\lambda_k \sim k^{-\alpha}$, with $\alpha = 1$, as expected for a spectrum which follows Wigner-Dyson statistics. The lower k modes which capture the nonuniversal global structure of the spectrum deviate from Wigner-Dyson statistics.

and grows quadratically for larger $\langle n \rangle$. The Thouless energy corresponds to the point where the number variance deviates from GOE predictions. As an estimate of the Thouless energy we chose the point for which the variance deviates by an arbitrary amount (set as 0.2), resulting in $E_{\rm Th} \sim 95\Delta$ for the box distribution, $E_{\rm Th} \sim 13\Delta$ for $\gamma = 1$, and $E_{\rm Th} \sim 7\Delta$ for $\gamma = 1.5$, where $\Delta = \langle \Delta_i \rangle$. This will naturally lead to the conclusion that, as γ increases, $E_{\rm Th}$ strongly decreases. For larger values of γ the variance departs from the GOE predictions close to Δ and those values were not plotted to avoid cluttering the figure at small $\langle n \rangle$.

As emphasized by Jia and Verbaarschot [31], since unlike typical RMT models for which all nondiagonal terms are random, SYK models have just $\binom{2L}{4} \ll \binom{L}{N}$ independent nondiagonal terms, which leads to significant sample-to-sample fluctuations within the ensemble, and using the average level spacing obtained by an ensemble average may significantly skew the number variance. Thus, we need a way to characterize the level spacing for a specific sample more accurately. As previously discussed, in Ref. [31] this was achieved by parametrizing the spectrum using Q-Hermite orthogonal polynomials. As described in the next section, here the SVD method is used.

V. SINGULAR VALUE DECOMPOSITION

SVD can be used to characterize the features of the spectrum; for example, on what scale does it follow RMT predictions [33–39]? For the analysis, one tabulates P eigenvalues around the center of the band of M realizations of disorder as a matrix X of size $M \times P$, where X_{mp} is the pth level of the mth realization. The matrix X is decomposed to $X = U\Sigma V^T$, where U and V are $M \times M$ and $P \times P$ matrices, correspondingly, and Σ is a *diagonal* matrix of size $M \times P$ and rank $r = \min(M, P)$. The r diagonal elements of Σ are the singular values amplitudes (SV) σ_k of X. All σ_k are positive and therefore may be ordered by their size $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \sigma_r$. The Hilbert-Schmidt norm of the matrix $||X||_{\rm HS} = \sqrt{{\rm Tr} X^\dagger X} = \sum_k \lambda_k$ (where $\lambda_k = \sigma_k^2$). The matrix X could be written as a series composed of matrices $X^{(k)}$, where $X^{(k)}_{ij} = U_{ik} V^T_{jk}$ and $X_{ij} = \sum_k \sigma_k X^{(k)}_{ij}$. This series is an approximation of matrix X, where the sum of the first m modes gives a matrix $\tilde{X} = \sum_{k=1}^{m} \sigma_k X^{(k)}$, for which $||X||_{\rm HS} - ||\tilde{X}||_{\rm HS}$ is minimal, i.e., the minimal departure between the approximate energy spectrum of all realizations, $\tilde{X}$, obtained using $m(M + P)$ independent variables compared to the full energy spectrum which requires MP variables. For the local ensemble average one unfolds using an averaged spectrum, thus employing P independent values. Thus, if only the first mode of the SVD is taken into account ($m = 1$), the approximation of the energy spectrum has M additional parameters. Writing down explicitly the $k = 1$ term of the spectrum for the ith realization results in $X^{(k=1)}_{ij} = U_{i1} V^T_{j1}$. Thus, the P values, common to all realizations V^T_{j1}, are multiplied by a single parameter U_{i1} unique to each of the M realizations. In other words, the ensemble-averaged spectrum V^T_{j1} is scaled by a single parameter U_{i1} for each realization.

The idea is that low modes capture the global behavior of the spectrum common to all realizations, while higher modes sample the fluctuations on smaller energy scales. If a distinct pattern of behavior of the amplitude as a function of the mode number can be seen for a particular range of k, it is meaningful to discuss different behaviors at different energy scales. Indeed, as can be seen in Fig. 3, examining

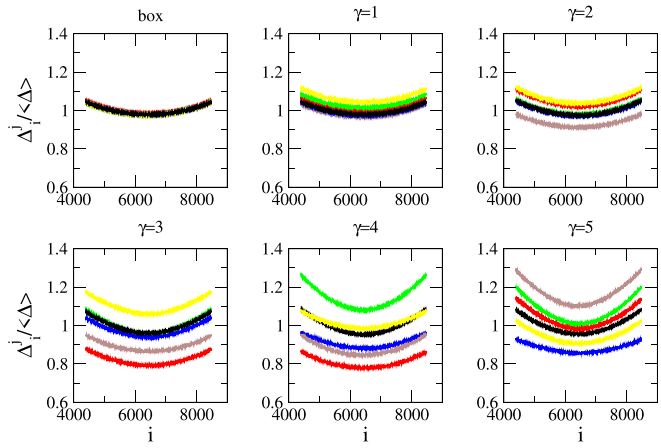

FIG. 4. The level spacing for the CSYK model with box distribution and different log-normal distribution with $\gamma = 1, 2, 3, 4$, and 5 for $L = 16$ and $N = 8$, for $M = 4096$ realizations and $P = 4096$ eigenvalues around the middle of the band. The ensemble-averaged level spacing, Δ_i, for the local unfolding method is the same for all samples and is depicted by the black curve. The SVD unfolded level spacing, Δ_i^j, is realization dependent. Five individual realizations for each distribution are shown (red, green, blue, yellow, and brown curves). As γ increases the sample-to-sample fluctuation increases. It is clear that the long-range correlations within a sample are very significant.

the scree plot of the singular values λ_k for the box distribution and the log-normal distribution with different values of γ and a fixed $p = 0.5$, for $L = 16$, $N = 8$, with $M = 4096$ realizations and $P = 4096$ eigenvalues around the middle of the band, one sees two distinct regions. The lowest modes' ($k = 1$ and 2 for the box distribution and $\gamma = 1$; $k = 1$, 2, 3, and 4 for $1.5 \leqslant \gamma \leqslant 2.5$; and $k = 1$, 2, 3, 4, 5, and 6 for $3 \leqslant \gamma \leqslant 5$) amplitudes are much larger than those of all the other modes, and they determine the very large scale, nonuniversal, behavior of the spectrum. The bulk of the modes for larger k follow a power-law behavior ($\lambda_k \sim k^{-\alpha}$) with $\alpha = 1$, as expected for Wigner-Dyson statistics [33–35].

In order to illustrate the difference between the local ensemble unfolding and unfolding using the lower modes of the SVD, it is useful to examine the difference in the behavior of level spacing of the ith level, Δ_i, obtained by each method. While for the local ensemble unfolding the level spacing Δ_i is averaged over all realizations and therefore is not realization dependent, for the SVD unfolding the ith level spacing of the jth realization $\Delta_i^j = (\tilde{\varepsilon}_{i+p}^j - \tilde{\varepsilon}_{i-p}^j)/2p$, where $\tilde{\varepsilon}_i^j = \sum_{k=1}^2 \sigma_k U_{ik} V_{jk}^T$ is realization dependent. As can be seen in Fig. 4 there is a noticeable difference between the realization-specific level spacing Δ_i^j and the ensemble average Δ_i. This difference becomes larger as γ increases, i.e., as the width of the log-normal distribution increases. Moreover, the difference, $\Delta_i^j - \Delta_i$, is long-range correlated for a given realization across thousands of levels. Thus, it makes sense to define the typical spacing difference between realizations $\delta\Delta = \sqrt{\langle(\Delta_i^j - \Delta_i)^2\rangle_{i,j}}$. For $L = 16$ and $N = 8$ with $M = 4096$ realizations and $P = 4096$ levels around the center of the band, one gets $\delta\Delta = 4.44 \times 10^{-3}\Delta$ for the box distribution, $\delta\Delta = 3.52 \times 10^{-2}\Delta$ for $\gamma = 1$, and $\delta\Delta = 6.61 \times 10^{-2}\Delta$ for $\gamma = 1.5$.

This long-range sample-to-sample level fluctuation can explain the behavior of the level number variance of the local ensemble unfolding seen in Fig. 3. Essentially, for larger values of $\langle n\rangle$, $\langle\delta^2 n\rangle = a + b\langle n\rangle^2$, with $b = 1.78 \times 10^{-5}$ for the box distribution, $b = 1.25 \times 10^{-3}$ for $\gamma = 1$, and $b = 4.4 \times 10^{-3}$ for $\gamma = 1.5$. The quadratic behavior could be understood as the consequence of realization-specific long-range fluctuation of the level spacing. Calculations show $\langle\delta^2 n\rangle = \langle(n - \langle n\rangle)^2\rangle$, taking into account that after unfolding $\langle n\rangle = n$ and for a typical realization $n \sim n(1 + \delta\Delta/\Delta)$. Thus, $\langle\delta^2 n\rangle \sim (n\delta\Delta/\Delta)^2$, and after plugging in the above-mentioned values of $\delta\Delta$ one obtains $\langle\delta^2 n\rangle \sim 1.97 \times 10^{-5} n^2$ for the box distribution, $\langle\delta^2 n\rangle \sim 1.24 \times 10^{-3} n^2$ for $\gamma = 1$, and $\langle\delta^2 n\rangle \sim 4.36 \times 10^{-5} n^2$ for $\gamma = 1.5$, in good agreement with the numerical values quoted above. Moreover, using our previous definition of the Thouless energy as the energy for which the deviation from RMT results becomes larger than some threshold, $b(E_{\rm Th}/\Delta)^2 = 0.2$, and using $b = (\delta\Delta/\Delta)^2$, one gets $E_{\rm Th} = 0.44\Delta^2/\delta\Delta$, resulting in $E_{\rm Th} = 100\Delta$, 18.5Δ, and 6.7Δ for box and log-normal distributions with $\gamma = 1$ and 1.5, correspondingly, in good agreement with the results in Fig. 3.

One concludes that the main contribution to the local ensemble-averaged Thouless energy $E_{\rm Th}$ comes from the sample-to-sample long-range fluctuations which can be quantified by $\delta\Delta$, the typical level spacing difference between the different realizations in the ensemble.

Thus, using the sample-specific level spacing from the SVD (i.e., Δ_i^j) for unfolding will eliminate the sample-to-sample fluctuations' contribution to the number variance. Indeed, using the SVD unfolded spectrum of the jth realization $\tilde{\varepsilon}_i^j$ defined by $\varepsilon_i^j = \varepsilon_{i-1}^j + (\tilde{\varepsilon}_i^j - \tilde{\varepsilon}_{i-1}^j)/\Delta_i^j$, for the calculation of the realization-specific level number variance one obtains the values depicted in Fig. 5. The number variance fits well with GOE expectations up to $E_{\rm Th^*} \sim 800\Delta$ for the box and $\gamma = 1$ distributions, $E_{\rm Th^*} \sim 500\Delta$ for $1.5 \leqslant \gamma \leqslant 2.5$ distributions, and $E_{\rm Th^*} \sim 300\Delta$ for $3 \leqslant \gamma \leqslant 5$ distributions. The deviation is towards lower variance than predicted by GOE, similar to the behavior of the number variance after unfolding using Q-Hermite orthogonal polynomials [31]. As shown in Ref. [38] the energy of deviation from the universal behavior can be read off the scree plot by defining the mode for which the behavior deviates from the k^{-1} power law as $k_{\rm Th^*}$. Thus, as can be garnered from Fig. 3, $k_{\rm Th^*} \sim 2$ for the box and $\gamma = 1$ distributions, $k_{\rm Th^*} \sim 4$ for $1.5 \leqslant \gamma \leqslant 2.5$ distributions, and $k_{\rm Th^*} \sim 6$ for $3 \leqslant \gamma \leqslant 5$ distributions. Estimating the Thouless energy by $E_{\rm Th^*} = r\Delta/2k_{\rm Th^*}$ [38] results in $E_{\rm Th^*} \sim 1000\Delta$, $E_{\rm Th^*} \sim 500\Delta$, and $E_{\rm Th^*} \sim 300\Delta$, not to far from the estimations obtained in Fig. 5.

Up to now we have mainly presented results for $L = 16$, which is the largest system for which we have ample statistics. Nevertheless, as can be seen in Fig. 1, for the CSYK model the statistics change for different values of L. In order to verify that our conclusions regarding the non-self-averaging behavior of the model are relevant for any symmetry, we calculated $\langle\delta^2 n(E)\rangle$ for the box distribution using both local ensemble unfolding and SVD unfolding for $L = 15$, with

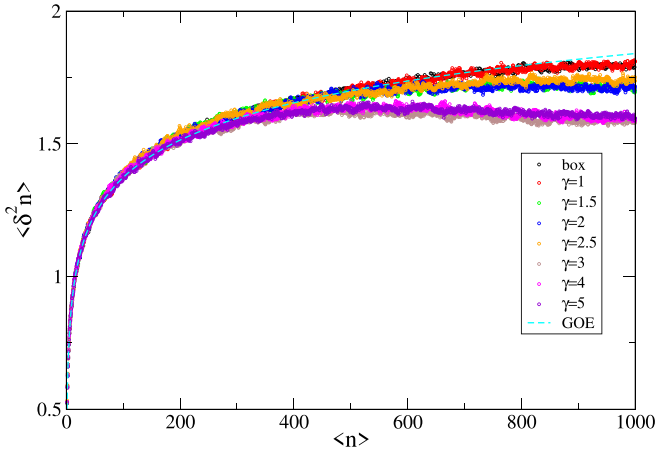

FIG. 5. The SVD unfolded level number variance, $\langle\delta^2 n(E)\rangle$, as a function of the energy window E for box and log-normal $\gamma = 1$, 1.5, 2, 2.5, 3, 4, and 5 distributions, represented by the symbols. The cyan dashed line is the GOE prediction $\langle\delta^2 n(E)\rangle = (2/\pi^2)\ln(\langle n(E)\rangle) + 0.44$. The SVD unfolded number variance departs from GOE predictions at $E_{\mathrm{Th}^*} \sim 800\Delta$ for the box and $\gamma = 1$ distributions, at $E_{\mathrm{Th}^*} \sim 500\Delta$ for $1.5 \leqslant \gamma \leqslant 2.5$ distributions, and at $E_{\mathrm{Th}^*} \sim 300\Delta$ for $3 \leqslant \gamma \leqslant 5$ distributions.

$N = 8$ particles. $P = 3000$ eigenvalues around the middle of the band were taken for $M = 3000$ realizations. For this case, the GUE behavior $\langle\delta^2 n(E)\rangle = (1/\pi^2)\ln(\langle n(E)\rangle) + 0.35$, is expected. Indeed, as can be seen in Fig. 6, both unfolding methods fit the GUE expectation for small $\langle n\rangle$. On larger scales the same difference that was seen for $L = 16$ GOE is seen for $L = 15$ GUE systems. The local ensemble unfolding results in $E_{\mathrm{Th}} \sim 90\Delta$, and a fit to a quadratic behavior of the form $a + bn^2$ leads to $b = 2.48 \times 10^{-5}$, while for the SVD

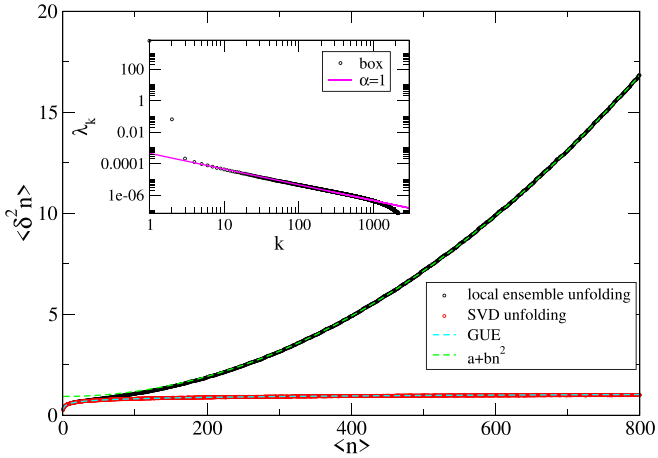

FIG. 6. The level number variance, $\langle\delta^2 n(E)\rangle$, as a function of the energy window E for the box distribution using local ensemble unfolding (black symbols) and SVD unfolding (red symbols) for $L = 15$, $N = 8$, $P = 3000$, and $M = 3000$. The cyan dashed line is the GUE prediction $\langle\delta^2 n(E)\rangle = (1/\pi^2)\ln(\langle n(E)\rangle) + 0.35$, while the green dashed line follows $a + bn^2$, with $b = 2.48 \times 10^{-5}$. The different behavior for the two unfolding methods is clear. Inset: The scree plot of the singular values λ_k indicated by the symbols as functions of k. The purple line corresponds to $\lambda_k \sim k^{-\alpha}$, with $\alpha = 1$.

unfolding $E_{\mathrm{Th}^*} \sim 800\Delta$. Comparing with the typical spacing difference between realizations for the box distribution which for $L = 15$ equals to $\delta\Delta = 5.17 \times 10^{-3}$, one would estimate $E_{\mathrm{Th}} = 0.44\Delta^2/\delta\Delta \sim 85\Delta$ and $b = (\delta\Delta/\Delta)^2 \sim 2.67 \times 10^{-5}$, both in good agreement with the numerical results. As explained above $E_{\mathrm{Th}^*} = r\Delta/2k_{\mathrm{Th}^*}$, here $r = 3000$ and $k_{\mathrm{Th}^*} \sim 2$, resulting in $E_{\mathrm{Th}^*} \sim 750\Delta$, again in good agreement with the results.

VI. DISCUSSION

Much of the fascination with the SYK model is connected to its chaotic behavior. Since short time scales are associated with large energy scales, deviation from GOE behavior of the spectra on large energy scales indicates nonchaotic behavior on short time scales. Estimating the time for which the chaotic behavior emerges is relevant to the estimation of the scrambling time of the SYK models, which motivates its application to studies of quantum gravity in black holes [42]. Ensemble averaging also plays an important role in the duality between the SYK model and classical general relativity.

For self-averaging systems there is no difference between averaging over the ensemble or averaging over a large single realization. As discussed, for finite-size SYK samples, there is a huge difference between the Thouless time for SVD unfolding compared to the ensemble-averaged unfolding, a difference which is larger as the distribution of the interaction elements is more skewed.

Nevertheless, it is relevant to extrapolate to infinite systems ($L \to \infty$) in order to see whether this difference persists. The SYK model has three timescales [43]: (i) band structure time t_B associated with the band width B; (ii) Thouless time, t_{Th} or t_{Th^*}, on which much of the paper concentrated, and (iii) Heisenberg time t_{H}. Since the band width depends linearly on L, $t_B = \hbar/B \sim \hbar/L$. Thus, the Heisenberg time $t_{\mathrm{H}} = \hbar/\Delta$, where $\Delta \sim B/2^L$, is $t_{\mathrm{H}} = \hbar\, 2^L/L$. The Thouless time for the ensemble-averaged unfolding is $t_{\mathrm{Th}} = \hbar/E_{\mathrm{Th}}$, where $E_{\mathrm{Th}} \sim \Delta L^2$ and resulting in $t_{\mathrm{Th}} = \hbar\, 2^L/L^3$. For the SVD unfolding, $E_{\mathrm{Th}^*} \sim 2^L \Delta B/k_{\mathrm{Th}^*}$, with $k_{\mathrm{Th}^*} \sim O(1)$, and thus, $t_{\mathrm{Th}^*} \sim k_{\mathrm{Th}^*}\hbar/B \sim k_{\mathrm{Th}^*}/L$. Hence, $t_{\mathrm{H}} > t_{\mathrm{Th}} \gg t_{\mathrm{Th}^*} \geqslant t_B$, and while the realization-adjusted Thouless time $\lim_{L\to\infty} t_{\mathrm{Th}^*} \to 0$, the ensemble-averaged Thouless time $\lim_{L\to\infty} t_{\mathrm{Th}} \to \infty$. The difference between t_{Th} and t_{Th^*} increases as the distribution is more skewed (Figs. 2 and 5). Other estimates of the Thouless energy using the spectral form factor coupled with a way to factor in the level density [30] results in $t_{\mathrm{Th}^*} \sim \ln(L)$. For small values of L this result is quite close to t_B and the Thouless time we obtained from the number variance using SVD unfolding. It would be impossible to compute large L, and therefore an exact dependence of t_{Th^*} on L is hard to establish. Nevertheless, it would be interesting to study the spectral factor using SVD unfolding in order to compare both measures directly.

This leads to the conclusion that one cannot determine the behavior of the SYK model by ensemble average for times shorter than t_{Th}, since shorter times are non-self-averaging. Moreover, these times ($t_{\mathrm{Th}^*} \sim t_B < t_{\mathrm{Th}} < t_{\mathrm{H}}$) correspond to a parametically large span of times. Therefore, when one wishes to study the transition from chaotic to localized behavior

in modified SYK models [10–21] for which the number of independent random variables remains low, sample-to-sample fluctuations are expected to remain important and non-self-averaging at short times should be considered. In principal, although the transition occurs at long times, nevertheless these effects could obscure the transition point and influence the perceived nature of the metallic regime. This will be considered in future studies.

[1] S. Sachdev and J. Ye, Phys. Rev. Lett. **70**, 3339 (1993).
[2] A. Kitaev, Talks at the KITP on April 7th and May 27th, 2015, https://online.kitp.ucsb.edu/online/entangled15/kitaev/ and https://online.kitp.ucsb.edu/online/entangled15/kitaev2/.
[3] S. Sachdev, Phys. Rev. X **5**, 041025 (2015).
[4] J. Maldacena and D. Stanford, Phys. Rev. D **94**, 106002 (2016).
[5] O. Bohigas and J. Flores, Phys. Lett. B **34**, 261 (1971).
[6] J. Maldacena, Int. J. Theor. Phys. **38**, 1113 (1999).
[7] A. M. García-García and J. J. M. Verbaarschot, Phys. Rev. D **94**, 126010 (2016).
[8] Y.-Z. You, A. W. W. Ludwig, and C. Xu, Phys. Rev. B **95**, 115150 (2017).
[9] T. Li, J. Liu, Y. Xin, and Y. Zhou, J. High Energy Phys. 06 (2017) 111.
[10] A. M. García-García, Y. Jia, and J. J. M. Verbaarschot, Phys. Rev. D **97**, 106003 (2018).
[11] A. M. García-García, B. Loureiro, A. Romero-Bermúdez, and M. Tezuka, Phys. Rev. Lett. **120**, 241603 (2018).
[12] T. Nosaka, D. Rosa, and J. Yoon, J. High Energy Phys. 09 (2018) 041.
[13] A. M. García-García and M. Tezuka, Phys. Rev. B **99**, 054202 (2019).
[14] T. Micklitz, F. Monteiro, and A. Altland, Phys. Rev. Lett. **123**, 125701 (2019).
[15] F. Monteiro, M. Tezuka, A. Altland, D. A. Huse, and T. Micklitz, Phys. Rev. Lett. **127**, 030601 (2021).
[16] F. Monteiro, T. Micklitz, M. Tezuka, and A. Altland, Phys. Rev. Res. **3**, 013023 (2021).
[17] C.-M. Jian, Z. Bi, and C. Xu, Phys. Rev. B **96**, 115122 (2017).
[18] S.-K. Jian and H. Yao, Phys. Rev. Lett. **119**, 206602 (2017).
[19] X. Chen, R. Fan, Y. Chen, H. Zhai, and P. Zhang, Phys. Rev. Lett. **119**, 207603 (2017).
[20] A. Altland, D. Bagrets, and A. Kamenev, Phys. Rev. Lett. **123**, 106601 (2019).
[21] D. K. Nandy, T. Čadež, B. Dietz, A. Andreanov, and D. Rosa, Phys. Rev. B **106**, 245147 (2022).
[22] M. L. Mehta, *Random Matrices*, 2nd ed. (Academic, New York, 1991).
[23] B. I. Shklovskii, B. Shapiro, B. R. Sears, P. Lambrianides, and H. B. Shore, Phys. Rev. B **47**, 11487 (1993).
[24] T. Guhr, A. Muller-Groeling, and H. A. Weidenmuller, Phys. Rep. **299**, 189 (1998).
[25] Y. Alhassid, Rev. Mod. Phys. **72**, 895 (2000).
[26] A. D. Mirlin, Phys. Rep. **326**, 259 (2000).
[27] F. Evers and A. D. Mirlin, Rev. Mod. Phys. **80**, 1355 (2008).
[28] B. Altshuler and B. Shklovskii, Sov. Phys. JETP **64**, 127 (1986) [Zh. Eksp. Teor. Fiz. **91**, 220 (1986)].
[29] J. Sonner and M. Vielma, J. High Energy Phys. 11 (2017) 149.
[30] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, J. High Energy Phys. 07 (2018) 124.
[31] Y. Jia and J. J. M. Verbaarschot, J. High Energy Phys. 07 (2020) 193.
[32] A. Aharony and A. B. Harris, Phys. Rev. Lett. **77**, 3700 (1996).
[33] R. Fossion, G. Torres-Vargas, and J. C. López-Vieyra, Phys. Rev. E **88**, 060902(R) (2013).
[34] G. Torres-Vargas, R. Fossion, C. Tapia-Ignacio, and J. C. López-Vieyra, Phys. Rev. E **96**, 012110 (2017).
[35] G. Torres-Vargas, J. A. Méndez-Bermúdez, J. C. López-Vieyra and R. Fossion, Phys. Rev. E **98**, 022110 (2018).
[36] R. Berkovits, Phys. Rev. B **102**, 165140 (2020).
[37] R. Berkovits, Phys. Rev. B **104**, 054207 (2021).
[38] R. Berkovits, Phys. Rev. B **105**, 104203 (2022).
[39] W.-J. Rao, Phys. Rev. B **105**, 054207 (2022).
[40] I. M. Khaymovich, V. E. Kravtsov, B. L. Altshuler, and L. B. Ioffe, Phys. Rev. Res. **2**, 043346 (2020).
[41] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Phys. Rev. Lett. **110**, 084101 (2013).
[42] S. Sachdev, arXiv:2205.02285.
[43] A. Altland, D. Bagrets, P. Nayak, J. Sonner, and M. Vielma, Phys. Rev. Res. **3**, 033259 (2021).