Replica symmetry breaking in random non-Hermitian systems

Antonio M. García-García,[1,*] Yiyang Jia (贾抑扬),[2,†] Dario Rosa,[3,4,‡] and Jacobus J. M. Verbaarschot[5,§]

[1]*Shanghai Center for Complex Physics, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 7610001, Israel*

[3]*Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon 34126, Korea*

[4]*Basic Science Program, Korea University of Science and Technology (UST), Daejeon 34113, Korea*

[5]*Center for Nuclear Theory, Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*

(Received 8 April 2022; accepted 27 May 2022; published 29 June 2022)

Recent studies have revealed intriguing similarities between the contribution of wormholes to the gravitational path integral and the phenomenon of replica symmetry breaking observed in spin glasses and other disordered systems. Interestingly, these configurations may also be important for the explanation of the information paradox of quantum black holes. Motivated by these developments, we investigate the thermodynamic properties of a *PT*-symmetric system composed of two random non-Hermitian Hamiltonians with no explicit coupling between them. After performing ensemble averaging, we identify numerically and analytically a robust first-order phase transition in the free energy of two models with quantum chaotic dynamics: the elliptic Ginibre ensemble of random matrices and a non-Hermitian Sachdev-Ye-Kitaev (SYK) model. The free energy of the Ginibre model is temperature independent in the low-temperature phase. The SYK model has similar behavior for sufficiently low temperature, and then it experiences a possible continuous phase transition to a phase with a temperature-dependent free energy before the first-order transition takes place at a higher temperature. We identify the order parameter of the first-order phase transition and obtain analytical expressions for the critical temperature. The mechanism behind the transition is the existence of replica symmetry breaking configurations coupling left and right replicas that control the low-temperature limit of the partition function. We speculate that quantum chaos may be necessary for the observed dominance of off-diagonal replica symmetry breaking configurations in the low-temperature limit.

DOI: 10.1103/PhysRevD.105.126027

I. INTRODUCTION

The replica trick [1] is a powerful tool in the study of disordered systems. It consists of replicating the action n times which facilitates the explicit calculation of the average over disorder. The resulting n-dependent action, describing the ensemble-averaged system, is then, in most cases, solved in the mean-field limit by the saddle-point method. In the last step of the calculation, the value of n is set to a value that depends on the observable of interest (typically 0 or 1).

The replica trick has been employed in a broad variety of problems in different research fields including disordered spin systems [2,3], quantum disordered conductors [4], random matrix theory [5], QCD [6,7], and the development of error correction codes [8]. For instance, in the context of disordered spin systems describing certain magnetic alloys, the replica trick plays a pivotal role in the physical description of the low-temperature spin-glass phase characterized by an energy landscape with multiple local minima and a splitting of the Gibbs measure into separate components (called pure states), which is a signature of breaking of ergodicity [9,10].

It was also found [2,3,9,10] that replica symmetry breaking solutions of the Sherrington-Kirkpatrick model [11], a model for these disordered spin systems, describe the low-temperature spin-glass region, while replica symmetric configurations are dominant for higher

*amgg@sjtu.edu.cn
†yiyang.jia@weizmann.ac.il
‡dario_rosa@ibs.re.kr
§jacobus.verbaarschot@stonybrook.edu

temperatures. Replica symmetry breaking (RSB) refers to solutions of the saddle-point equations which couple different replicas and that, superficially, should be subleading in the mean-field limit. These RSB solutions have a precise physical meaning for spin glasses [3,9]: they represent the overlap of probability among pure states which is directly related to the order parameter of the transition.

A different type of RSB is found in the context of disordered systems [4] and random matrix theory [5,12,13]. In this case, the replica symmetry between the advanced and retarded sectors of the Green's function is broken, leading to Goldstone's modes that dominate the partition function. These configurations give nonperturbative contributions to spectral correlators that provide information on the dynamics for scales on the order of the Heisenberg time. Indeed, fully accounting for all RSB solutions, it reproduces [5,12,13] the exact random matrix theory result for the two-level correlation function.

Another model that has recently been intensively studied by means of the replica trick is the Sachdev-Ye-Kitaev model [14–19]: a model describing N Majorana fermions with infinite-range random interactions in Fock space. Variants of this model with complex fermions were originally introduced [14,15,20] and studied [21–26] in the context of nuclear physics and quantum chaos over half a century ago.

The renewed interest in this model is due to intriguing similarities with Jackiw-Teitelboim (JT) gravity [27,28], a two-dimensional theory of gravity that describes almost extremal black holes in near AdS_2 (Anti-de Sitter space in two dimensions) backgrounds [19,29,30]. In the infrared limit, both models share the same action: a Schwarzian whose path integral can be evaluated exactly [31]. The resulting spectral density [32,33], which grows exponentially for excitations close to the ground state, is consistent with that of quantum black holes. The dynamics is quantum chaotic [16] with spectral correlations given by random matrix theory predictions [32,34], classified according to the global symmetries of the system [35,36]. Likewise, a weakly coupled two-site Sachdev-Ye-Kitaev (SYK) model, which is also quantum chaotic [37–40] for sufficiently high energies, reproduces the physics of the transition from a traversable wormhole to a two-black-hole configuration in near-AdS_2 backgrounds with Lorentzian signature [41,42].

On the gravity side, it may seem that disorder, and therefore any nontrivial structure in replica space, plays no role and that these similarities with the SYK model, where the replica symmetric solution is typically chosen, are unrelated to the fact that the SYK model is a disordered system. However, recent results in the gravity literature put in doubt this prediction. In a recent work by Saad *et al.* [43], it was found that the dual theory of JT gravity was exactly given by a random matrix theory in a certain scaling limit which suggests that the gravitational path integral involves an average over different theories (see also Ref. [44]). Moreover, a replica calculation [45] of the free energy in JT gravity identified a range of parameters where the contribution of RSB configurations, called replica wormholes in this context, are dominant compared to replica symmetric configurations. Similarly, the calculation [46–48] of the evolution of the von Neumann entropy in JT gravity plus additional matter, modeling the black hole evaporation process, showed that for late times the growth stops due to additional RSB saddle points in the gravitational path integral. These RSB saddle points represent wormholes connecting different copies of black holes. This behavior is in agreement with that expected for Hermitian systems [49]. The "information paradox" is therefore avoided.

However, these results also raise some fundamental issues. It seems that the gravitational path integral represents an ensemble over theories, something that is not yet well understood. Moreover, at least in field theories with a gravity dual, Euclidean wormholes raise the so-called factorization puzzle, namely, the field theory dual to wormholes connecting two boundaries should be related to a field theory partition function that does not factorize [50–56], but it is unclear how exactly to define such an object. Another problem is that these Euclidean wormholes, at least in JT gravity without additional matter, are not solutions of the classical equations of motion [43,57], so their interpretation as RSB saddle solutions is not straightforward. In the simplest case of two replicas, it was possible [58] to find wormhole solutions of the classical JT gravity equations, provided that complex sources were added. The system undergoes a first-order wormhole–black hole transition where the wormhole phase is characterized by a free energy that depends only weakly on the temperature until a possible second continuous phase transition occurs, below which the free energy becomes temperature independent.

Given these recent advances, an interesting question to ask is whether it is possible to find field theories whose dominating saddle points are RSB configurations and whether their role is qualitatively similar to that of wormholes in gravity theories. A positive answer to this question may shed some light on the factorization and information loss puzzle mentioned above and, more generally, on the role of wormholes in holography and quantum gravity. Even putting aside any gravitational interpretation, it is a problem of fundamental interest to determine the conditions for the dominance of off-diagonal replicas in disordered and strongly interacting quantum mechanical systems.

The main goal of this paper is to address this problem by studying several random non-Hermitian but PT symmetric two-site systems with no explicit coupling between them. Among others, we investigate the elliptic Ginibre ensemble of random matrices and the non-Hermitian SYK model [59]. By downgrading the Hermiticity of the SYK model to just PT symmetry [60], so that the model still has a real positive partition function, we identify RSB configurations

that control the free energy in the low-temperature limit. The restoration of replica symmetry at higher temperature triggers a first-order thermal phase transition. If the imaginary part of the SYK model is large enough, we have indications of the existence of an additional continuous phase transition at a temperature below the one at which the first-order transition takes place. Moreover, we obtain explicit expressions for the critical temperature, the ground-state energy, and the order parameter that characterizes the RSB phase. Our results are qualitatively similar to those of a gravitational system [58] and also largely universal, provided that the dynamics is quantum chaotic [59].

We note that the role of RSB configurations has already been the subject of different studies [61,62] for the SYK model with real couplings. Although there is not yet consensus in the literature, it seems that in these cases most of the features of the model, which are also present in JT gravity, do not involve any RSB.

The paper is organized as follows. In Sec. II, we qualitatively explain why we expect a universal thermal phase transition due to RSB configurations in a non-Hermitian random quantum system. This is illustrated in Sec. III by an analytical solution of a non-Hermitian random matrix model with PT symmetry, which roughly corresponds to the two-site non-Hermitian SYK model with a q-body ($q > 2$) interaction. In Sec. IV, by an explicit solution of the Schwinger-Dyson (SD) equations and also by the numerical calculation of the free energy from the eigenvalues of the SYK Hamiltonian, we show that a $q = 4$ two-site non-Hermitian SYK model with PT symmetry and no explicit coupling between the two sites also undergoes a first-order phase transition induced by RSB configurations. We close with concluding remarks and a list of topics for further research in Sec. V. Technical details are worked out in six Appendixes. Some of the results of this paper were announced in a recent letter [59].

II. REPLICA SYMMETRY BREAKING IN RANDOM NON-HERMITIAN, PT-SYMMETRIC SYSTEMS

In this section, we aim to give a qualitative argument for the existence of a rather universal phase transition for the free energy of a PT-symmetric system composed of two random disconnected non-Hermitian Hamiltonians. This can be viewed as a replicated version (with two replicas) of a single-site non-Hermitian Hamiltonian. The low-temperature phase is dominated by RSB configurations whose effect is strikingly similar to that of Euclidean wormholes in AdS_2 gravity. In later sections, we discuss examples including a two-site non-Hermitian SYK model where an explicit replica analysis is possible.

We argue below that for the two-site non-Hermitian systems we are interested in the replica trick gives correct results. We will also see that for these systems the quenched and annealed free energies are identical in the thermodynamic limit. This justifies using annealed averaging to obtain quenched free energies, which we will do for the Schwinger-Dyson calculation of the free energy.

In the second part of this section, we show that when eigenvalues have the universal characteristics of quantum chaotic systems the connected two-level correlation function corresponding to RSB configurations contributes to the free energy at leading order. Moreover, we argue that these contributions indeed control the low-temperature limit of the free energy. In Sec. IV A, an analysis of the SD equations for the one-replica SYK model will show more explicitly that RSB configurations are directly responsible for the phase transition which mimics that observed for Euclidean wormholes in JT gravity [58].

A. Quenched free energy by the replica trick

We consider the partition functions of two-site Hamiltonians of the form

$$H = H_L \otimes 1 + 1 \otimes H_R. \tag{1}$$

We are mostly interested in the case where $H_L = H_R^\dagger$, and in general H_L, H_R, H are non-Hermitian, but H is PT-symmetric [60], namely,

$$[PT, H] = 0,$$

with P a permutation matrix that interchanges the L and R Hilbert spaces and where the antiunitary operator is $T = CK \otimes CK$. Here, C is a charge conjugation matrix and K the complex conjugation operator. If the D eigenvalues of the complex $D \times D$ matrix H_L are denoted by E_k, then the D^2 eigenvalues of H are given by $E_k + E_l^*$. The eigenvalues with $k = l$ are real, while the other eigenvalues come in complex-conjugate pairs, consistent with the existence of PT symmetry.

The partition function of this Hamiltonian (before averaging over the disorder) is given by

$$Z(\beta) = \mathrm{Tr} e^{-\beta H} = Z_L Z_R = |\mathrm{Tr} e^{-\beta H_L}|^2, \tag{2}$$

where we have defined

$$Z_L \equiv \mathrm{Tr} e^{-\beta H_L}, \qquad Z_R \equiv \mathrm{Tr} e^{-\beta H_R}, \tag{3}$$

and obviously $Z_L = Z_R^*$. If $\rho(z)$ is the eigenvalue density of H_L, then

$$Z(\beta) = \int d^2 z_1 d^2 z_2 \rho(z_1)\rho(z_2) e^{-\beta(z_1 + z_2^*)}. \tag{4}$$

The quenched free energy must be computed by a quenched average $-\beta\langle F\rangle = \langle \log Z\rangle$, where β is the inverse of temperature T. A direct analytical calculation of the

quenched disorder average is in general technically demanding. The replica trick was introduced [1] to circumvent these difficulties by using that

$$\langle \log Z \rangle = \lim_{n\to 0} \frac{\langle Z^n \rangle - 1}{n} = \lim_{n\to 0} \frac{\langle (Z_L Z_L^*)^n \rangle - 1}{n}. \quad (5)$$

The average on the right-hand side is much easier to evaluate analytically by replicating n times the original action, carrying out the averages analytically and taking the limit $n \to 0$ at the end of the calculation.

However, a word of caution is in order: it is well documented that the replica trick may give incorrect results if applied naively [63,64]. An example is the Sherrington-Kirkpatrick model mentioned earlier, where the entropy is negative for sufficiently low temperature if the replica trick is naively applied [11]. A number of fixes have been introduced [3,5,12,65,66], including the supersymmetric method that avoids the replica trick altogether [67–71]. However, in many situations, there are no realistic alternatives, so it is necessary to understand under which conditions the trick is applicable. The replica trick is premised on Carlson's theorem [72], which states that if a holomorphic function $f(z)$ on $\mathrm{Re}(z) > 0$ vanishes for all positive integers n it also vanishes on the right half-plane, provided that $|f(z)| < C \exp(\pi|z|)$ on the imaginary axis and grows no faster than an exponential elsewhere on the right half-plane. For a non-Hermitian Hamiltonian such as H_L, $\log Z_L$ in general has a nonzero imaginary part, and therefore it is unclear whether the conditions of Carlson's theorem are satisfied in the low-temperature limit. Hence, if we were interested in the free energy of the one-site model, the naive replica trick

$$\langle \log Z_L \rangle = \lim_{n\to 0} \frac{\langle Z_L^n \rangle - 1}{n} \quad (6)$$

is likely to give incorrect results.

The average of the one-site free energy can be expressed as

$$\langle \log Z_L \rangle = \langle \log |Z_L| \rangle + \langle i \arg Z_L \rangle. \quad (7)$$

For a non-Hermitian Hamiltonian, the phase of Z_L is expected to oscillate rapidly so that the average of the second term vanishes. If that is the case, we have

$$\langle \log Z_L \rangle = \langle \log |Z_L| \rangle = \frac{1}{2} \langle \log Z_L Z_L^* \rangle. \quad (8)$$

This shows that the quenched average free energy is necessarily given by the quenched free energy of a replica and a conjugate replica (in the sense of the one-site model). Because $\log(Z_L Z_L^*)$ is real, we have that $\langle \exp n \log Z_L Z_L^* \rangle$ is bounded for imaginary n so that there is a chance we can apply Carlson's theorem to validate the replica trick. We thus have

$$\langle \log Z_L \rangle = \lim_{n\to 0} \frac{1}{2} \frac{\langle (Z_L Z_L^*)^n \rangle - 1}{n}. \quad (9)$$

Notice that this is exactly half of (5); therefore, the correct replica description of a non-Hermitian one-site model naturally involves the conjugate replicas. This procedure is actually well known for quenched averages (now understood as ignoring the fermion determinant) of a similar quantity, namely, the resolvent $G(z) = \mathrm{Tr}(H_L - z)^{-1}$ [6,73,74]. For a non-Hermitian Hamiltonian, the quenched resolvent is given by the replica limit,

$$G(z) = \lim_{n\to 0} \frac{1}{2n} \frac{d}{dz} \langle \det{}^n (H_L + z) \det{}^n (H_L^\dagger + z^*) \rangle, \quad (10)$$

which is sometimes referred to as Hermitization [6,73,75,76].

More importantly, we will study the two-site system using the mean-field approximation. Because $Z_L Z_L^*$ is positive definite, we expect the replica trick to work, and if the two-site system has a replica-diagonal behavior

$$\langle (Z_L Z_L^*)^n \rangle = \langle Z_L Z_L^* \rangle^n, \quad (11)$$

then the replica limit (9) is given by

$$\langle \log Z_L \rangle = \lim_{n\to 0} \frac{1}{2} \frac{\langle (Z_L Z_L^*)^n \rangle - 1}{n} = \lim_{n\to 0} \frac{1}{2} \frac{\langle Z_L Z_L^* \rangle^n - 1}{n} = \frac{1}{2} \log \langle Z_L Z_L^* \rangle. \quad (12)$$

We conclude that in the thermodynamic limit the quenched free energy of Z_L is given by half the annealed free energy of $Z_L Z_L^*$. The latter, for non-Hermitian theories, is generally different from the annealed free energy of Z_L.

At this point, it is useful to clarify a potentially confusing semantic point of our notion of RSB which is different from that in spin glasses. It is reminiscent of RSB in disordered systems where RSB happens between n retarded and n advanced Green's functions, i.e., $U(2n) \to U(n) \times U(n)$. In the present case, we have RSB between replicas and conjugate replicas of the partition function. In the replica symmetric phase, the replicas remain uncoupled after averaging so that

$$\langle (Z_L Z_L^*)^n \rangle = \langle Z_L \rangle^n \langle Z_L^* \rangle^n. \quad (13)$$

When replica symmetry is broken, this factorization no longer holds,

$$\langle (Z_L Z_L^*)^n \rangle \neq \langle Z_L \rangle^n \langle Z_L^* \rangle^n, \quad (14)$$

but we still have that

$$\langle (Z_L Z_L^*)^n \rangle = \langle Z_L Z_L^* \rangle^n. \tag{15}$$

So, from the two-site model perspective, $\langle Z_L Z_L^* \rangle$ is the one-replica partition function of the two-site Hamiltonian H, which is not expected to bring about any further RSB. On the other hand, this can be viewed as the two-replica partition function of the single-site Hamiltonian H_L. In that case, one can legitimately talk about RSB. However, the two perspectives are mathematically equivalent.

For a characterization of the conditions to observe dominant RSB configurations is important to split the partition function into a connected and a disconnected piece:

$$\langle Z \rangle = \langle Z_L Z_L^* \rangle = \langle Z_L Z_L^* \rangle_c + \langle Z_L \rangle \langle Z_L^* \rangle. \tag{16}$$

The first term receives contributions from the connected two-point function, while the second term is determined by the one-point function. Because of the non-Hermiticity, $\langle Z_L(\beta) \rangle$ may actually be exponentially suppressed so that the connected part of the partition function may become dominant. As discussed in the previous paragraph, we will refer to this situation as RSB. In a field theory formulation of the partition function, the corresponding saddle-point configuration of the action connects *different* replicas. In this paper, we will not pursue an explicit gravitational interpretation of these results. However, the analogy with gravity is evident: RSB configurations are the field theory analog of Euclidean wormhole solutions in gravity, which are tunneling geometries connecting two or more otherwise disconnected space-time regions. However, we note that if this analogy applies the relation with a gravitational system is not at the level of the microscopic Hamiltonian but rather at the level of the effective action resulting from the replica trick after ensemble averaging.

The remainder of this section is devoted to a better understanding of both the circumstances for which the connected part becomes relevant so that RSB configurations control the free energy of the system and the effect of RSB on the thermodynamic properties of the system.

B. Existence of a phase transition induced by RSB configurations

For the random Hamiltonian (1), we calculate the expectation value of the partition function as

$$\langle Z(\beta) \rangle = \int d^2 z_1 d^2 z_2 \int \langle \rho(z_1) \rho(z_2) \rangle e^{-\beta(z_1 + z_2^*)}, \tag{17}$$

where $\beta \equiv \frac{1}{T}$ is the inverse temperature and the level density is given by

$$\rho(z) = \sum_k \delta^2(z - z_k). \tag{18}$$

The two-point correlation function can be decomposed as

$$\langle \rho(z_1) \rho(z_2) \rangle = \langle \rho(z_1) \rangle \langle \rho(z_2) \rangle + \delta^2(z_1 - z_2) \langle \rho(z_1) \rangle + R_{2,c}(z_1, z_2) \tag{19}$$

with

$$R_{2,c}(z_1, z_2) \equiv \sum_{k \neq l} [\langle \delta^2(z_1 - z_k) \delta^2(z_2 - z_l) \rangle - \langle \delta^2(z_1 - z_k) \rangle \langle \delta^2(z_2 - z_l) \rangle], \tag{20}$$

which is the two-point correlator without self-correlations. Because of the normalization of the level density, we have the sum rule obtained by integrating (19) over z_1:

$$\int dz_1 R_{2,c}(z_1, z_2) = -\langle \rho(z_2) \rangle. \tag{21}$$

The decomposition of the partition function corresponding to (19) is

$$\langle Z(\beta) \rangle = |\langle Z_L(\beta) \rangle|^2 + \int d^2 z \rho(z) e^{-\beta(z + z^*)} + \int d^2 z_1 d^2 z_2 R_{2,c}(z_1, z_2) e^{-\beta(z_1 + z_2^*)}. \tag{22}$$

Notice that, from (22) on, we will no longer make a notational distinction between $\rho(z)$ and $\langle \rho(z) \rangle$ when the context is free of confusion. Because of the sum rule (21), the second term in (22) cancels the third term for $\beta = 0$. Therefore, there is no RSB in the infinite temperature limit. We shall see that for sufficiently low temperature the situation is different.

To simplify the argument, for now we assume a rotationally invariant eigenvalue distribution, so that

$$\int d^2 z \rho(z) e^{-\beta z} = D \tag{23}$$

with D the number of eigenvalues of the one-site Hamiltonian. We note that this is a realistic situation which can occur for instance for H given by two copies of the Ginibre ensemble [77] of complex random matrices.

We next estimate the connected part of the partition function. To do that, we need to make three assumptions about the two-point correlations of the eigenvalues:

(1) The correlations are isotropic and only depend on the distance of the eigenvalues so that

$$R_{2,c}(z_1, z_2) = -\rho(\bar{z})^2 F(|z_1 - z_2|/\lambda), \tag{24}$$

where $\bar{z}$ is the center of mass coordinate $\bar{z} = (z_1 + z_2)/2$.

(2) The correlation length is a function of $\rho(\bar{z})$ only.

(3) The average spectral density $\rho(\bar{z})$ does not appreciably vary on the scale of the correlation length.

These assumptions are expected to hold for the universal correlations of non-Hermitian quantum chaotic systems. In terms of the integral over the center of mass [i.e., $(z_1+z_2)/2$] and the differences of the eigenvalues (i.e., z_1-z_2), the sum rule reads

$$-\rho(\bar{z})^2\int F(|z_1-z_2|/\lambda)d^2(z_1-z_2)=-\rho(\bar{z}). \tag{25}$$

This requires that $\lambda\sim 1/\sqrt{\rho(\bar{z})}$ and

$$\rho(\bar{z})\int d^2(z_1-z_2)F(|z_1-z_2|\sqrt{\rho(\bar{z})})=1. \tag{26}$$

Since $\rho(\bar{z})\sim D$, we have that the length scale of the eigenvalue correlations is $1/\sqrt{D}$.

For the connected part of the partition function, we then obtain

$$Z_{2,c}=-\int d^2\bar{z}\int d^2(z_1-z_2)\rho(\bar{z})^2F(|z_1-z_2|\sqrt{\rho(\bar{z})})\times e^{-\beta[\bar{z}+\bar{z}^*+i\mathrm{Im}(z_1-z_2)]}+\int d^2\bar{z}\rho(\bar{z})e^{-\beta(\bar{z}+\bar{z}^*)}. \tag{27}$$

Since the correlations are short ranged, we can Taylor expand $\exp(i\mathrm{Im}(z_1-z_2))$. The first term in the Taylor expansion cancels with the second integral in Eq. (27) because of the sum rule, and the second term of the Taylor expansion vanishes because the correlations are even in y_1-y_2. So, to leading nonvanishing order in $1/\rho(\bar{z})$, we obtain

$$Z_{2,c}=\frac{\beta^2}{2}\int d^2\bar{z}\int d^2(z_1-z_2)\rho(\bar{z})^2F(|z_1-z_2|\sqrt{\rho(\bar{z})})\times(\mathrm{Im}(z_1-z_2))^2e^{-\beta(\bar{z}+\bar{z}^*)}. \tag{28}$$

We can scale $\rho(\bar{z})$ out of the z_1-z_2 integrations. Then, the integral factorizes, and the integral over the difference is just a constant which we will denote by $\langle\zeta^2\rangle$. The connected part of the partition function is thus given by

$$Z_{2,c}=\frac{\beta^2\langle\zeta^2\rangle}{2}\int_{|\bar{z}|<E_0}d^2\bar{z}e^{-\beta(\bar{z}+\bar{z}^*)}=\frac{\pi}{2}\beta E_0\langle\zeta^2\rangle I_1(2\beta E_0), \tag{29}$$

where E_0 is the radius of the support of $\rho(\bar{z})$, I_1 is a modified Bessel function of the first kind, and

$$\langle\zeta^2\rangle=\int d^2sF(|s|)(\mathrm{Im}s)^2, \tag{30}$$

with $s=\sqrt{\rho(\bar{z})}(z_1-z_2)$. The annealed free energy is thus given by

$$-\beta F=\log[c\beta I_1(2\beta E_0)+D^2]. \tag{31}$$

Since $E_0>0$, the first term becomes dominant at low temperatures. More specifically, a genuine phase transition at finite temperature can occur, provided that $\log D$ and E_0 scale linearly with the system size N. This is indeed the case for interacting fermionic systems such as the (two-site) SYK model, where the ground-state energy $E_0=Ne_0/2$ (each site has $N/2$ Majoranas) with e_0 the ground-state energy per particle that does not depend on N and $D=2^{N/4}$. In the thermodynamic limit, we then can use the asymptotic limit of the modified Bessel function, while the prefactors are irrelevant. The free energy is given by

$$F/N=-\theta(T_c-T)e_0-\theta(T-T_c)T\log 2/2 \tag{32}$$

with

$$T_c=\frac{2e_0}{\log 2}. \tag{33}$$

This argument for the transition is based on the existence of the above large-N scaling behavior and the scaling properties of the two-point correlation function. In the next section, we will study an example where these conditions are met. The rotational invariance of the spectrum is in fact not essential. In particular, from the universality arguments and Eqs. (28)–(30), we see that the connected part of the partition function depends on $\rho(z)$ only through the shape of its support. Moreover, the exponential dependence on $2\beta E_0$ is largely independent of the shape of the support of $\rho(z)$; consider the integral in Eq. (29) with a generic support, namely,

$$\int_{\mathrm{supp}(\rho)}d^2\bar{z}e^{-\beta(\bar{z}+\bar{z}^*)}=\int_{\mathrm{supp}(\rho)}e^{-2\beta x}dxdy. \tag{34}$$

If the support of ρ on the real axis has a projection $[-E_0,E_{\max}]$, and for each x the slice of the support has a length $L(x)$ along the y direction, the integral becomes

$$\int_{-E_0}^{E_{\max}}e^{-2\beta x}L(x)dx. \tag{35}$$

Since E_0 is a large parameter ($E_0\sim N$), the integral on the right-hand side localizes at the maximum of the integrand, namely, at $\tilde{x}=-E_0$ as long as $L(x)\ll e^x$. This establishes that $Z_{2,c}\sim e^{2\beta E_0}$.

III. FREE ENERGY FOR THE ELLIPTIC GINIBRE MODEL

In this section, we evaluate the free energy for the elliptic Ginibre model. In the large D limit, the elliptic Ginibre model has a constant level density inside an ellipse, while the Ginibre model has a constant level density inside a circle. This model can be seen as a representative of the universality class of non-Hermitian Hamiltonians which do not necessarily have a rotationally invariant level density. Our calculations will be framed in general terms and can be applied to the non-Hermitian SYK model with slight modifications. Using the general arguments of Sec. II A, the quenched free energy of the elliptic Ginibre model is equal to the annealed free energy of the model with one replica and one conjugate replica. In the case of the ordinary Ginibre model, the two-site partition function can be evaluated analytically for finite D (see Appendix A). In that case, the one-site partition function does not depend on β. This is related directly to the fact that the eigenvalue density is rotationally invariant. We now show that the derivation of the previous section is also valid for the universality class of the elliptic Ginibre ensemble.

The Hamiltonian of the elliptic Ginibre ensemble [78,79] is given by

$$H_L = \frac{H_1 + ikH_2}{\sigma(k)}; \tag{36}$$

here, $\sigma(k)$ is a scale factor which can be k dependent. In this model, H_1 and H_2 are independent matrices extracted from the Gaussian Unitary Ensemble according to the probability distribution

$$e^{-\frac{2N}{\sigma_0^2}(\mathrm{Tr}H_1^2+\mathrm{Tr}H_2^2)}. \tag{37}$$

Note that σ_0 is a parameter that is unrelated to $\sigma(k)$. In most of this paper, we choose $\sigma(k) = 1$, but we point out another interesting possibility $\sigma(k) = \sqrt{2(1-k^2)}$ such that the variance of the real part of the eigenvalues is independent of k. This guarantees, as we will see below, that the disconnected partition function, dominant in the high-temperature phase, is k independent. This could be of interest for quantitative comparisons with the gravity picture [58] where the high-temperature phase corresponds to two decoupled black holes.

We can choose (1) as the definition of our total Hamiltonian where $H_L = H_R^\dagger$, but unlike the SYK model, this would not give a PT-symmetric Hamiltonian. We can instead let $H_L = H_R^*$; then, we have a Hamiltonian which is PT symmetric. In any case, both choices give the same spectrum and hence the same partition function. We next evaluate the two-site partition function with H_L given by (36).

The calculation of the partition function can actually be carried out at finite D by integrating the expressions for the spectral density and two-level correlation functions of the elliptic Ginibre ensemble which were first obtained in Ref. [78]. Here, we are interested in the large D limit and only need the asymptotic form of the spectral density which is constant inside an ellipse with long axis length $2E_0$ and short axis length $2y_0$,

$$\rho(z) = \frac{D}{\pi E_0 y_0}\theta\left(1 - \frac{x^2}{E_0^2} - \frac{y^2}{y_0^2}\right) \tag{38}$$

with

$$E_0 = \frac{\sigma_0}{\sqrt{1+k^2\sigma(k)}}, \qquad y_0 = \frac{\sigma_0 k^2}{\sqrt{1+k^2\sigma(k)}}, \tag{39}$$

and $z = x + iy$. We do not need the specific form of the two-point correlations other than that they are isotropic and short range with a range that scales as $1/\sqrt{D}$ as given by the general form (24).

To obtain the disconnected part of the partition function, we need to evaluate the one-site partition function, which can be easily computed by the following parametrization of the energy integration variable:

$$z = E_0 r\cos\phi + iy_0 r\sin\phi \tag{40}$$

with $r \in [0,1]$ and $\phi \in [-\pi,\pi]$. Using that the Jacobian of this transformation is rE_0y_0, we find the partition function

$$\begin{aligned} Z_L(\beta) &= \int d^2z e^{-\beta z}\rho(z) \\ &= \frac{D}{\pi}\int_0^1 rdr\int_{-\pi}^{\pi} d\phi e^{-\beta r(E_0\cos\phi + iy_0\sin\phi)} \\ &= \frac{2D}{\beta\sqrt{E_0^2-y_0^2}} I_1\left(\beta\sqrt{E_0^2-y_0^2}\right). \end{aligned} \tag{41}$$

Using that $E_0^2 - y_0^2 = \sigma_0^2(1-k^2)/\sigma^2(k)$, we obtain the disconnected piece of the two-site partition function

$$Z_{2,\mathrm{dis}}(\beta) = \frac{4D^2\sigma^2(k)}{\beta^2\sigma_0^2(1-k^2)}\left[I_1(\beta\sigma_0\sqrt{1-k^2}/\sigma(k))\right]^2. \tag{42}$$

When we have $\sqrt{1-k^2}/\sigma(k) \to 0$ for $k \to 1$, as is the case for our generic choice of $\sigma(k) = 1$, the disconnected contribution to the free energy becomes temperature independent for $k = 1$. The connected part of the partition function is given by

$$Z_{2,c} = \int d^2z_1 d^2z_2 \rho_{2c}(z_1,z_2) e^{-\beta(z_1+z_2^*)}, \tag{43}$$

where the connected two-point correlation function $\rho_{2c}(z_1, z_2)$ is given by the sum

$$\rho_{2,c}(z_1, z_2) = R_{2,c}(z_1, z_2) + \delta^2(z_1 - z_2)\rho(z_2). \quad (44)$$

The first term represents the true two-point correlations involving two different eigenvalues, while the second term is due to self-correlations. To evaluate the connected partition function, we can use the general argument given in Sec. II B, but now we have an explicit expression for the two-point correlation function which satisfies the conditions used in that section. In particular, the two-point correlation function is given by

$$R_{2,c}(z_1, z_2) = -\rho(\bar{z})^2 F_{\text{unv}}\left(\sqrt{\rho(\bar{z})}|z_1 - z_2|\right), \quad (45)$$

where $\bar{z} = (z_1 + z_2)/2$ and F_{unv} is a universal function that is given by the large D result for the Ginibre ensemble [77]:

$$F_{\text{unv}}(s) = e^{-\pi s^2}. \quad (46)$$

It satisfies the sum rule

$$\int dz_1 R_{2,c}(z_1, z_2) = -\rho(z_2) \quad (47)$$

and that the eigenvalue correlations are short ranged on the scale of $1/\sqrt{D}$. The two-point function (46) can also be derived rigorously for the elliptic Ginibre ensemble [78].

We can now proceed in exactly the same way as in Sec. II B but with an explicit expression for the two-point correlation function. Let us expand the Boltzmann factor in powers of $y_1 - y_2$, the imaginary part of $z_1 - z_2$,

$$e^{-\beta(z_1+z_2^*)} = e^{-\beta(\bar{z}+\bar{z}^*)}\left(1 - i\beta(y_1 - y_2) - \frac{1}{2}\beta^2(y_1 - y_2)^2 + \cdots\right), \quad (48)$$

where $\bar{z} = (z_1 + z_2)/2$. As mentioned in Sec. II B, the contribution due to the first term vanishes because of the sum rule (47), and the second term does not contribute because the integral is even in $y_1 - y_2$. We thus find

$$Z_{2,c} = \frac{\beta^2}{2}\langle(y_1 - y_2)^2\rangle \int d^2\bar{z} e^{-\beta(\bar{z}+\bar{z}^*)}\theta\left(1 - \frac{\bar{x}^2}{E_0^2} - \frac{\bar{y}^2}{y_0^2}\right), \quad (49)$$

where

$$\langle(y_1 - y_2)^2\rangle = \int (y_1 - y_2)^2 \rho(\bar{z})^2 e^{-\pi|z_1 - z_2|^2\rho(\bar{z})} d^2(z_1 - z_2) = \frac{1}{2\pi}. \quad (50)$$

Using the parametrization (40), we obtain

$$Z_{2,c} = \frac{\beta y_0}{4} I_1(2\beta E_0). \quad (51)$$

To derive this result, we have interchanged the large D limit and the integrations over the spectral density and spectral correlations. In Appendix A, we show that this misses additional corrections which change the prefactor in (51). Since these corrections do not change the exponential D-dependence of the contribution, they do not affect the free energy in the thermodynamic limit.

The total partition function is given by

$$Z = Z_{2,\text{dis}} + Z_{2,c}. \quad (52)$$

Taking only the leading nonvanishing terms in the thermodynamic limit, we simplify the free energy to

$$F(T) = -T \log\left(e^{2\frac{E_0}{T}} + D^2 e^{\frac{2}{T}\sqrt{E_0^2 - y_0^2}}\right). \quad (53)$$

In order to mimic forthcoming results for the two-site SYK model and more generically of interacting fermionic systems, we set $D = 2^{N/4}$ and $\sigma_0 = e_0 N/2$, where e_0 is a size-independent microscopic energy scale, and we stress that N is not the number of eigenvalues of the Ginibre Hamiltonian. With these choices, the free energy per particle can be written as

$$\frac{F}{N} = -\theta(T_c - T)\frac{e_0}{\sqrt{1 + k^2\sigma(k)}} - \theta(T - T_c)\left(T\frac{\log 2}{2} + \frac{e_0\sqrt{1 - k^4}}{\sqrt{1 + k^2\sigma(k)}}\right), \quad (54)$$

where the critical temperature of the first-order phase transition is given by

$$T_c = \frac{2e_0}{\sigma(k)\sqrt{1 + k^2}\log 2}\left(1 - \sqrt{1 - k^4}\right), \quad (55)$$

which for small k scales as k^4. These results are fully consistent with the universal expression (32). Indeed, the free energy for the elliptic Ginibre model is qualitatively similar to those of the ordinary (circular) Ginibre case; in both cases, there is a first-order phase transition separating a low-temperature region where the free energy is dominated by RSB configurations. We now explore whether this first-order transition is a feature of more realistic fermionic systems such as a non-Hermitian SYK model, where there are N Majoranas in zero spatial dimensions with infinite-range interactions and random complex couplings. Therefore, we do not expect that any artificial choice of scaling is necessary to observe the transition. For $k = 1$, the elliptic Ginibre model reduces to the ordinary circular

Ginibre model. Its partition function can be evaluated exactly at finite D, and up to a prefactor, the large D limit of this result is in agreement with the results derived in this section. The details are worked out in Appendix A.

IV. FREE ENERGY AND RSB FOR THE *PT*-SYMMETRIC SYK MODEL

We now turn to the study of the Hamiltonian (1) with H_L and H_R given by a $q = 4$ SYK model with complex couplings,

$$H = \sum_{i<j<k<l}^{N/2} (J_{ijkl} + ikM_{ijkl})\psi_L^i\psi_L^j\psi_L^k\psi_L^l + \sum_{i<j<k<l}^{N/2} (J_{ijkl} - ikM_{ijkl})\psi_R^i\psi_R^j\psi_R^k\psi_R^l, \quad (56)$$

where the variances of the couplings are

$$\langle (J_{i_1\cdots i_q})^2\rangle = \langle (M_{i_1\cdots i_q})^2\rangle = \frac{2^{q-1}(q-1)!}{q(N/2)^{q-1}}v^2 \quad (57)$$

and v sets the physical scale. The strength of the complex deformation resulting in a non-Hermitian Hamiltonian is controlled by the parameter k. The Majorana fermions satisfy the Clifford algebra

$$\{\psi_L^i, \psi_L^j\} = \{\psi_R^i, \psi_R^j\} = \delta^{ij}, \qquad \{\psi_L^i, \psi_R^j\} = 0. \quad (58)$$

We have also studied variations of this non-Hermitian SYK model, for example, a model where the couplings are not complex conjugated. However, the partition function of this model is not positive definite. A more interesting possibility is to include an explicit coupling term between the two sites. This model has the remarkable property that all eigenvalues become real beyond a critical value of the coupling. Below, we will see that we will have to add an infinitesimal explicit coupling term to break the symmetry between the left and right replicas. The effect of a finite coupling will be studied in detail in Ref. [80].

This section is divided into two parts. First, we provide theoretical arguments, supported by numerical results obtained by exact diagonalization of the Hamiltonian, which show that the free energy of this SYK model is quantitatively similar to that of the elliptic Ginibre model. In the second part, we confirm this conclusion by explicitly calculating the free energy from the solution of the SD equations. These equations are the saddle-point equations derived for one replica and one conjugate replica and give the large N limit of the free energy [18]. We will see the free energy obtained from the SD equations agrees with the Ginibre prediction.

For the numerical calculations, we diagonalize the *one-site* Hamiltonian with up to $N/2 = 34$ Majoranas. In this case, we can directly calculate the quenched free energy, which is equal to half the free energy of the two-site model [see Eq. (8)], and there is no need to use the replica trick. We have found that the ensemble fluctuations of $\log|Z_L|^2$ are small and this quantity seems to be self-averaging for large N. The annealed average $\langle|Z_L|^2\rangle$, which corresponds to one replica and one conjugate replica, shows much stronger fluctuations, and it is not clear if it is self-averaging. For comparison with theoretical predictions, it is necessary to eliminate the fluctuations by averaging about many disorder realizations. We shall see that indeed, after averaging, annealed and quenched averages lead to similar results by comparing the quenched free energy from exact diagonalization with that obtained from the solution of the Schwinger-Dyson equations that assumes annealed averages in its derivation.

We start with the analysis of the distribution of complex eigenvalues. In Fig. 1, we depict the distribution of the eigenvalues of a one-site SYK model with $N/2 = 30$ Majoranas and compare it with the ellipse (red curve) given by

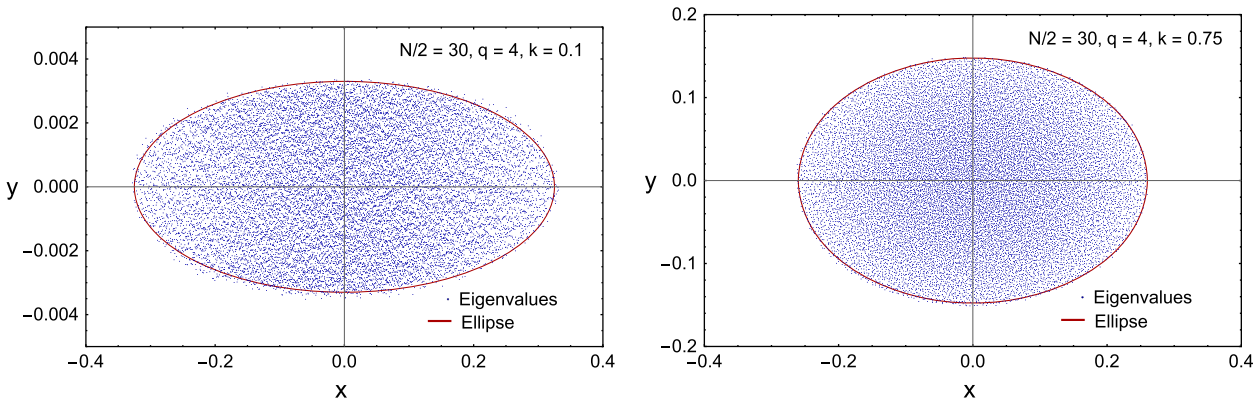

FIG. 1. Eigenvalue distributions of a single realization of the non-Hermitian one-site SYK model compared to the elliptical eigenvalue distributions with non-Hermiticity parameters $k = 0.1$ (left) and $k = 0.75$ (right). Note the different scales of the y axes in the two plots.

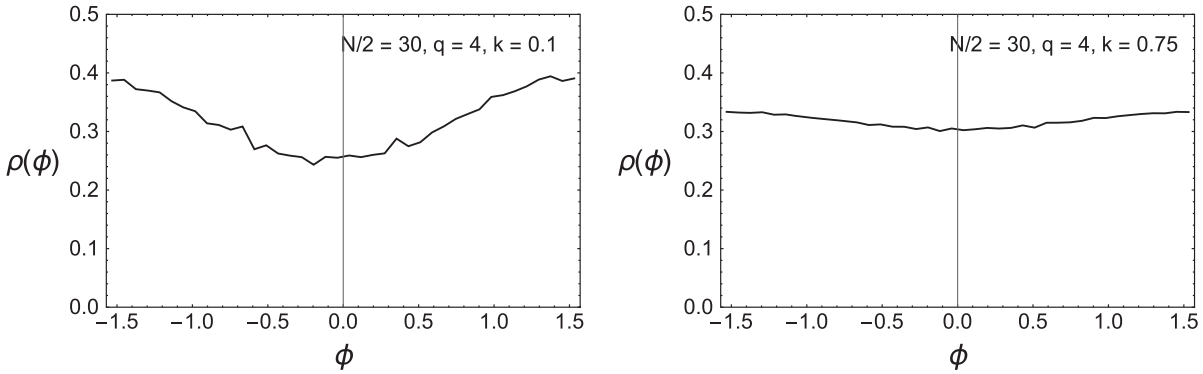

FIG. 2. Distribution of the phases of the eigenvalues after linearly rescaling the spectrum from an elliptical shape to a circular shape. This distribution is uniform for $k = 1$ and becomes less uniform for smaller values of k. This is illustrated for $k = 0.1$ (left) and for $k = 0.75$ (right).

$$\frac{x^2}{E_0^2(k)} + \frac{y^2}{y_0^2(k)} = 1 \tag{59}$$

with $E_0(k)$ and $y_0(k)$ fitting parameters. The quality of the fit of the support of the spectrum is comparable to that of the elliptic Ginibre model, but contrary to the elliptic Ginibre model, the eigenvalue distribution is not completely uniform. In Fig. 2, we plot the distribution of the phase ϕ_n of the rescaled eigenvalues

$$|\tilde{E}_n| e^{i\phi_n} = \frac{\mathrm{Re}(E_n)}{E_0(k)} + i\frac{\mathrm{Im}(E_n)}{y_0(k)} \tag{60}$$

for $k = 0.1$ (left) and $k = 0.75$ (right). For $k = 1$, the distribution of the phase is uniform but becomes less uniform for smaller values of k. However, the deviation from uniformity is well fitted by a $\cos 2\phi$-dependence.

Finally, in Fig. 3, we compare the fitted values of $E_0(k)$ and $y_0(k)$ to the analytical functional dependence obtained for the elliptic Ginibre model, namely,

$$E_0(k) = \frac{N}{2}\frac{e_0}{\sqrt{1+k^2}}, \tag{61}$$

$$y_0(k) = \frac{N}{2}\frac{e_0 k^2}{\sqrt{1+k^2}}, \tag{62}$$

with e_0 the ground-state energy per particle for $k = 0$ [for $N/2 = 30$, we obtain $e_0 \approx 0.011$; see Eq. (66)]. The agreement is excellent, which strongly suggests that indeed the two models have very similar spectral properties. We now turn to the study of the free energy.

We recall two of the main features of the free energy of the elliptic Ginibre model. First, because of the non-Hermiticity, the disconnected part of the partition function is exponentially suppressed for $k \neq 0$, which makes it possible for its magnitude to be comparable to that of the connected part. Second, because the spectral correlations are short range, the details of these correlations are irrelevant. As a consequence of a spectral sum rule, both the leading contribution due to the self-correlations and those due to the genuine two-point correlations are of the same magnitude but with an opposite sign and cancel at leading

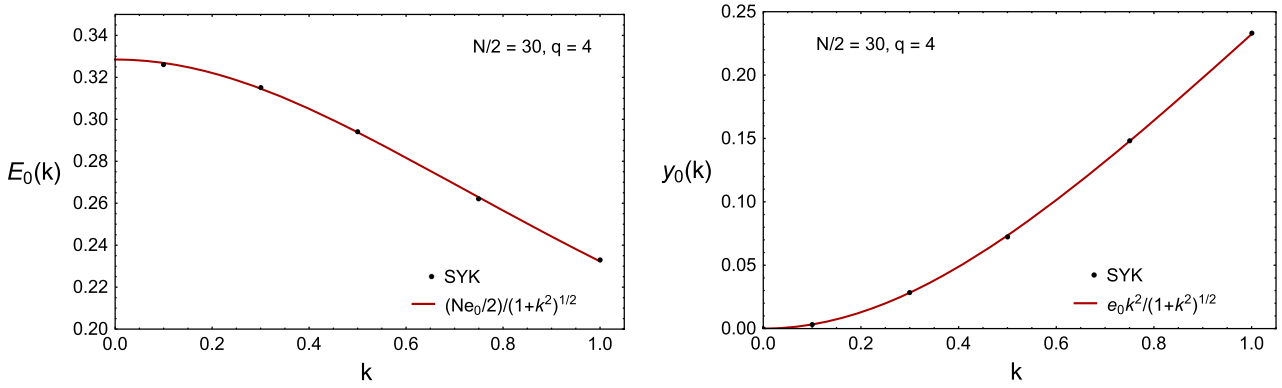

FIG. 3. The k-dependence of the long semi-axis, the ground-state energy $E_0(k)$, and the short semi-axis $y_0(k)$ that define the elliptical support of the eigenvalues of a one-site non-Hermitian SYK model with $N/2 = 30$. The k-dependence is compared to the prediction for the elliptic Ginibre model.

order for large N. A phase transition induced by RSB can only occur if, after this leading-order cancellation, the remaining two-point piece is comparable with the disconnected part. We can get an estimate of the critical temperature by assuming that the eigenvalues are distributed uniformly inside an ellipse with long axes $2E_0$ and short axis $2y_0$; in other words, they are given by the distribution of the elliptic Ginibre ensemble. Using the results of the previous section, we find the critical temperature

$$T_c = \frac{E_0(k) - \sqrt{E_0^2(k) - y_0^2(k)}}{\log 2} \frac{4}{N} \tag{63}$$

with the k-dependence of $E_0(k)$ and $y_0(k)$ given by the results for the elliptic Ginibre model (62). This gives a critical temperature,

$$T_c = \frac{2e_0}{\log 2} \frac{1 - \sqrt{1-k^4}}{\sqrt{1+k^2}}, \tag{64}$$

which behaves as k^4 for small k. The free energy in terms of e_0 is given by

$$-\frac{F}{N/2} = \theta(T_c - T) \frac{2e_0}{\sqrt{1+k^2}} + \theta(T - T_c)\left(T \log 2 + 2e_0\sqrt{1-k^2}\right). \tag{65}$$

The energy e_0, the ground-state energy per particle for the $k = 0$ SYK model, is given by [32,33]

$$(e_0 N/2)^2 = \frac{4M_2}{1-\eta}, \tag{66}$$

where M_2 is our choice for the second moment of the one-site $k = 0$ SYK model

$$M_2 = \binom{N/2}{4} \langle J_{ijkl}^2 \rangle = \binom{N/2}{4} \frac{1}{6N^3}, \tag{67}$$

and

$$\eta = \binom{N/2}{q}^{-1} \sum_{m=0}^{q} (-1)^m \binom{q}{m} \binom{N/2 - q}{q - m} \sim 1 - \frac{4q^2}{N}, \tag{68}$$

where we have chosen $q = 4$ and $v = \frac{1}{24}$ in (57). This choice is the one employed in the numerical calculations. For $N/2 = 30$, we find $e_0 N/2 = 0.332$, while from Fig. 3, we can read off a value of 0.328 which is only slightly lower. The analytical result for the critical temperature using Eq. (63) and $e_0 N/2 = 0.332$ is equal to $T_c = 0.0225$ for $k = 1$, which is also close to the result from exact diagonalization, which is approximately 0.022. As will be discussed later in this section, an independent calculation of T_c and E_0 by exact diagonalization is in agreement with these results.

The analytical results for the Ginibre model are largely based on the uniformity of the distribution of the eigenvalues. However, in the SYK case, the phase is only uniform for $k = 1$, while the radial distribution is never uniform. For $k < 1$, we have found that the ϕ-dependence of the spectral density is well fitted by (see Fig. 2)

$$\rho(\phi) = \frac{1}{2\pi}(1 - \alpha \cos 2\phi). \tag{69}$$

The angular integral of the disconnected part of the partition function then becomes

$$2I_0\left(\sqrt{2}\beta r\sqrt{1-1/k^2}\right) - 2\alpha\frac{k^4+1}{k^4-1} I_2\left(\sqrt{2}\beta r\sqrt{1-1/k^2}\right). \tag{70}$$

Therefore, the leading exponent is not affected. The same argument can be made for the self-correlations and genuine two-point correlations. The deviation of the radial distribution from uniformity also does not change the leading exponent. This implies that, even for $k < 1$, we expect the same results as for the elliptic Ginibre model; namely, in the large N limit, there is a k-dependent first-order phase transition.

So far, we have restricted our analysis to the $k \leq 1$ region. It is easy to see that for $k > 1$ the partition function is equivalent to that resulting from the transformation $k \to 1/k$ and $\beta \to i\beta$. The calculation of the free energy can be carried out along the line of the $k < 1$ calculation. Details are worked out in Appendix D.

Another interesting question is whether the eigenvalue distribution can be related to that of the SYK with real couplings. We have found, see Appendix C for details, that indeed the real and imaginary parts of the eigenvalues of the non-Hermitian SYK are still well described by the Q-Hermite prediction [33,81], though the fitted value of η is no longer given by the analytical estimate (68). Notably, for $k < 0.3$, the distribution of the imaginary part of the eigenvalues is very close to semicircular. These facts are not directly related to the physics of the RSB but illustrate the rather deep connections between the models we are considering.

A. Free energy, ground-state energy, and gap of the SYK model from the SD equations

We now compare the predictions of the Ginibre model with a calculation of the free energy from the solution of the SD equations for the same two-site $q = 4$ non-Hermitian SYK model in the ΣG representation [18,82,83]. This formulation is based on the replica trick for the quenched partition function, $\langle \log |Z_L|^2 \rangle$. However, we assume that

the mean-field calculation does not break the replica symmetry so that the free energy can be obtained from the one-replica calculation, i.e., from the annealed partition function $\langle |Z_L|^2 \rangle$; see the end of Sec. II A where this terminology is introduced. However, we do have RSB between Z_L and Z_L^*, which are only coupled by the disorder. We will see that, when the temperature is sufficiently low, the dominant solutions of the saddle-point equations couple a replica and a conjugate replica. We refer to Sec. II for a justification of both the correctness of the replica trick in this non-Hermitian case and the equivalence of quenched and annealed averages. This is important as by design the free energy from the SD equations involves an annealed average but we are interested only in quenched averages.

Following the standard procedure [16,18], we obtain the SYK action in Euclidean time as a simple variation of the action considered in Ref. [41],

$$-\frac{2S_E}{N} = \log \mathrm{Pf}(\partial_t \delta_{ab} - \Sigma_{ab}) - \frac{1}{2}\int d\tau_1 d\tau_2 \sum_{a,b}\left[\Sigma_{ab}(\tau_1,\tau_2)G_{ab}(\tau_1,\tau_2) - s_{ab}\frac{\mathcal{J}_{ab}^2}{2q^2}[2G_{ab}(\tau_1,\tau_2)]^q\right] - \frac{i}{2}\epsilon\int d\tau (G_{LR}(\tau,\tau) - G_{RL}(\tau,\tau)), \tag{71}$$

where the indices a and b can be equal to L or R. The function s_{ab} takes the values $s_{LL} = s_{RR} = 1$ and $s_{LR} = s_{RL} = (-1)^{q/2}$, and the couplings $\mathcal{J}_{ab}$ are taken to be $\mathcal{J}$ when $a = b$ and $\tilde{\mathcal{J}}$ when $a \neq b$. They are related to the variance and covariance of the random L and R couplings by

$$\begin{aligned}\langle (J^L_{ijkl})^2\rangle &= \langle (J^R_{ijkl})^2\rangle = \frac{2^{q-1}(q-1)!}{q(N/2)^{q-1}}(1-k^2)v^2 \\ &\equiv \frac{2^{q-1}(q-1)!}{q(N/2)^{q-1}}\mathcal{J}^2, \\ \langle J^L_{ijkl}J^R_{ijkl}\rangle &= \frac{2^{q-1}(q-1)!}{q(N/2)^{q-1}}(1+k^2)v^2 \\ &\equiv \frac{2^{q-1}(q-1)!}{q(N/2)^{q-1}}\tilde{\mathcal{J}}^2,\end{aligned} \tag{72}$$

where the left and right couplings are related to the couplings of the Hamiltonian (56) by

$$\begin{aligned}J^L_{ijkl} &\equiv J_{ijkl} + ikM_{ijkl}, \\ J^R_{ijkl} &\equiv J_{ijkl} - ikM_{ijkl}.\end{aligned} \tag{73}$$

This gives

$$\mathcal{J}^2 \equiv (1-k^2)v^2, \qquad \tilde{\mathcal{J}}^2 \equiv (1+k^2)v^2. \tag{74}$$

Finally, G_{ab} denotes the fermion bilinear defined via the equations

$$G_{ab}(\tau_1,\tau_2) \equiv \frac{2}{N}\sum_{i=1}^{N/2}\psi^i_a(\tau_1)\psi^i_b(\tau_2) \tag{75}$$

(we are assuming $0 < \tau < \beta$), while Σ_{ab} are the Lagrange multipliers that implement this constraint. They can also be interpreted as the self-energies of the fermions, and the expectation value of $G_{ab}(\tau_1,\tau_2)$ is the Green's function. Note the $i\epsilon$ term in the action (71) would have corresponded to a term

$$i\epsilon\sum_i \psi^i_L\psi^i_R \tag{76}$$

in the Hamiltonian, which was not present in the original Hamiltonian (56). However, as we will see in the next section, we will need this infinitesimal term added to detect the symmetry breaking whose order parameter is G_{LR}.

1. Symmetries of the Green's functions

From the definition (75), we then obtain the symmetry relations

$$G_{ab}(\tau_1,\tau_2) = -G_{ba}(\tau_2,\tau_1). \tag{77}$$

Assuming translational invariance as we will do in the remainder of this section, we have

$$G_{ab}(\tau_1,\tau_2) \to G_{ab}(\tau_1-\tau_2), \qquad \Sigma_{ab}(\tau_1,\tau_2) \to \Sigma_{ab}(\tau_1-\tau_2). \tag{78}$$

This results in

$$G_{LR}(\tau) = -G_{RL}(-\tau), \qquad G_{LL}(\tau) = -G_{LL}(-\tau), \qquad G_{RR}(\tau) = -G_{RR}(-\tau). \tag{79}$$

For $\epsilon = 0$, the action is invariant under

$$\psi_L(\tau) \to -\psi_L(\tau), \qquad \psi_R(\tau) \to \psi_R(\tau). \tag{80}$$

This symmetry can be implemented by the operator $\prod_{i=1}^{N/2}\psi^i_L$, when $N/2$ is even. Therefore, we have that

$$G_{LR}(\tau) = \langle \psi_L(\tau)\psi_R(0)\rangle_\beta = 0, \tag{81}$$

where $\langle\cdot\rangle_\beta$ denotes a thermal expectation value. We take $\tau > 0$ to avoid using the time-ordering symbol. This symmetry is broken by a nonzero value of ϵ. In the large N limit, we shall see it is broken spontaneously for $\epsilon \to 0$, at sufficiently low temperature. Below, we always consider the limit

$$\lim_{\epsilon\to 0}\lim_{N\to\infty} G_{LR}(\tau). \tag{82}$$

Also, in terms of eigenvectors and eigenfunctions of the SYK model, $G_{LR}(\tau)$ vanishes identically without the presence of ϵ term as one can easily check numerically for small values of N.

Next, we consider an antiunitary PT operation,

$$PT_1 : \psi_L^i(\tau) \to \psi_R^i(\tau), \quad \psi_R^i(\tau) \to -\psi_L^i(\tau), \quad i \to -i. \tag{83}$$

This operation can be implemented by

$$P = \exp\left(-\frac{\pi}{2}\sum_j \psi_L^j\psi_R^j\right) = \prod_j \frac{1}{\sqrt{2}}(1 - 2\psi_L^j\psi_R^j),$$
$$T_1 = T_L \otimes T_R, \tag{84}$$

where T_L, T_R are just the conventional time reversals of a single-site SYK that leaves the fermions invariant and takes i to $-i$. We note that PT_1 is a symmetry of the two-site Hamiltonian without the $i\epsilon$ term (76) but gets explicitly broken by the $i\epsilon$ term. Fortunately, if we compose PT_1 with the symmetry of equation (80), we get another antiunitary operation that is a symmetry even in the presence of the $i\epsilon$ term,

$$PT_2 : \psi_L^i(\tau) \to \psi_R^i(\tau), \quad \psi_R^i(\tau) \to \psi_L^i(\tau), \quad i \to -i, \tag{85}$$

implemented by the same P and

$$T_2 = T_1 \prod_{i=1}^{N/2} \psi_L^i. \tag{86}$$

This PT_2 symmetry ensures the reality of the partition function even in the presence of the $i\epsilon$ term. The PT_2 symmetry results in the identities

$$G_{LR}(\tau) = \langle \psi_L(\tau)\psi_R(0)\rangle_\beta = \langle \psi_R(\tau)\psi_L(0)\rangle_\beta^* = G_{RL}^*(\tau), \tag{87}$$

$$G_{LL}(\tau) = \langle \psi_L(\tau)\psi_L(0)\rangle_\beta = \langle \psi_R(\tau)\psi_R(0)\rangle_\beta^* = G_{RR}^*(\tau), \tag{88}$$

where we have used the antiunitarity of PT_2. Note that, although the P operator alone is not a symmetry, it satisfies

$$PHP^{-1} = H^\dagger \tag{89}$$

in the presence of the $i\epsilon$ term. This means

$$\psi_L(\tau)^\dagger = -P\psi_R(-\tau)P^{-1}, \quad \psi_R(\tau)^\dagger = P\psi_L(-\tau)P^{-1}. \tag{90}$$

Hence, we have[1]

$$\begin{aligned} G_{LR}^*(\tau) &= (\mathrm{Tr} e^{-\beta H}\psi_L(\tau)\psi_R(0))^* \\ &= \mathrm{Tr}(\psi_R(0)^\dagger\psi_L(\tau)^\dagger e^{-\beta H^\dagger}) \\ &= -\mathrm{Tr}(\psi_L(0)\psi_R(-\tau)e^{-\beta H}) \\ &= -\mathrm{Tr}(e^{-\beta H}\psi_L(\tau)\psi_R(0)) = -G_{LR}(\tau). \end{aligned} \tag{91}$$

This is to say $G_{LR}(\tau)$ is purely imaginary. Together with the symmetries (79) and (87), this gives

$$G_{LR}(\tau) = G_{LR}(-\tau). \tag{92}$$

Since $G_{LR}(\tau+\beta) = -G_{LR}(\tau)$, we also find

$$G_{LR}(\beta - \tau) = -G_{LR}(\tau). \tag{93}$$

Using the symmetries (79) and the antiperiodicity of $G_{LL}(\tau)$, we obtain

$$G_{LL}(\beta - \tau) = G_{LL}(\tau). \tag{94}$$

That is, $G_{LR}(\tau)$ is odd about $\beta/2$, whereas $G_{LL}(\tau)$ is even about $\beta/2$.

So far, all the symmetry relations we worked out are true for each independent realization of the random couplings. For the Hermitian Maldacena-Qi SYK model [41], there is one more relation that holds:

$$G_{RR}(\tau) = G_{LL}(\tau), \qquad G_{LL}(\tau) = G_{LL}^*(\tau). \tag{95}$$

In our non-Hermitian model, there is not enough symmetry for the above to hold realization by realization. Indeed, as we can numerically verify, for a generic realization, G_{LL} and G_{RR} are complex and only $G_{LL} = G_{RR}^*$ [Eq. (88)] holds. However, if we perform the ensemble averaging, we would expect Eq. (95) to hold for the non-Hermitian model because

$$H(J, M, \epsilon)^\dagger = H(J, -M, \epsilon) \tag{96}$$

[1]We omit the normalization factor $\mathrm{Tr}(e^{-\beta H})$ in the denominator for this derivation since it is real and does not affect the reality property of Green's functions. Also note that τ is the Euclidean time.

and the distribution of the disorder M is an even function. Let us summarize all the symmetry relations of the Green's functions in one place:

$$\begin{aligned} G_{ab}(\tau) &= -G_{ba}(-\tau), \\ G_{LR}(\tau) &= G_{LR}(-\tau) = -G^*_{LR}(\tau), \\ G_{LL}(\beta-\tau) &= G_{LL}(\tau), \qquad G_{LR}(\beta-\tau) = -G_{LR}(\tau), \\ G_{LR}(\tau) &= G^*_{RL}(\tau), \qquad G_{LL}(\tau) = G^*_{RR}(\tau), \\ \langle G_{RR}(\tau)\rangle &= \langle G_{LL}(\tau)\rangle, \qquad \langle G_{LL}(\tau)\rangle = \langle G^*_{LL}(\tau)\rangle. \end{aligned} \tag{97}$$

We stress all the above equations except the last line hold for each realization of the ensemble. It is also useful to note that the LL Green's function satisfies

$$G_{LL}(0) = \frac{\langle \frac{2}{N}\mathrm{Tr}\sum_k (\psi^k_L)^2 e^{-\beta H}\rangle}{\langle \mathrm{Tr} e^{-\beta H}\rangle} = \frac{1}{2}. \tag{98}$$

The symmetries of G_{ab} are inherited by Σ_{ab}. This also follows from the Schwinger-Dyson equations, which will be discussed in the next subsection.

2. Schwinger-Dyson equations

Starting from the action (71), the stationarity of Σ_{ab} gives the following set of saddle-point equations for the Fourier components of Σ_{ab} and G_{ab}[2]:

$$\begin{pmatrix} i\omega_n + \Sigma_{LL}(\omega_n) & \Sigma_{LR}(\omega_n) \\ \Sigma_{RL}(\omega_n) & i\omega_n + \Sigma_{RR}(\omega_n) \end{pmatrix} \times \begin{pmatrix} G_{LL}(-\omega_n) & G_{RL}(-\omega_n) \\ G_{LR}(-\omega_n) & G_{RR}(-\omega_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{99}$$

Using that $G_{RR}(\omega_n) = G_{LL}(\omega_n) = -G_{LL}(-\omega_n)$ and $G_{LR}(\omega_n) = G_{LR}(-\omega_n) = -G_{RL}(\omega)$, the saddle-point equations can be simplified to

$$\begin{aligned} -(i\omega_n + \Sigma_{LL}(\omega_n))G_{LL}(\omega_n) + \Sigma_{LR}(\omega_n)G_{LR}(\omega_n) &= 1, \\ (i\omega_n + \Sigma_{LL}(\omega_n))G_{LR}(\omega_n) + \Sigma_{LR}(\omega_n)G_{LL}(\omega_n) &= 0. \end{aligned} \tag{100}$$

In (100), we have introduced the fermionic Matsubara frequencies

$$\omega_n \equiv \frac{2\pi}{\beta}\left(n + \frac{1}{2}\right), \tag{101}$$

with β being the inverse temperature. At the stationary points of the G integral, the SD equations for self-energies Σ_{LL} and Σ_{LR} in the time domain take the form of

$$\begin{aligned} \Sigma_{LL}(\tau) &= \frac{\mathcal{J}^2}{q}(2G_{LL}(\tau))^{q-1}, \\ \Sigma_{LR}(\tau) &= (-1)^{q/2}\frac{\tilde{\mathcal{J}}^2}{q}(2G_{LR}(\tau))^{q-1} - i\epsilon\delta(\tau), \\ \Sigma_{RR}(\tau) &= \frac{\mathcal{J}^2}{q}(2G_{RR}(\tau))^{q-1}, \\ \Sigma_{RL}(\tau) &= (-1)^{q/2}\frac{\tilde{\mathcal{J}}^2}{q}(2G_{RL}(\tau))^{q-1} + i\epsilon\delta(\tau). \end{aligned} \tag{102}$$

They can be rewritten as integral equations for the Fourier components of G_{ab} and Σ_{ab}, so that the SD equations constitute a set of coupled integral equations.

Using the symmetry properties of G_{ab} in Eq. (97), we obtain the following relations (for even q):

$$\begin{aligned} \Sigma_{LL}(\tau) &= \Sigma_{RR}(\tau), \qquad \Sigma_{RL}(\tau) = -\Sigma_{LR}(\tau), \\ \Sigma_{LL}(-\tau) &= -\Sigma_{LL}(\tau), \qquad \Sigma_{LR}(-\tau) = \Sigma_{LR}(\tau). \end{aligned} \tag{103}$$

This again leads to the evenness of $\Sigma_{LL}(\tau)$ the oddness of $\Sigma_{LR}(\tau)$ about $\beta/2$.

Using the symmetry properties, the SD equations (100) can be conveniently rewritten in the following form:

$$\begin{aligned} G_{LL}(\omega_n) &= -\frac{i\omega_n + \Sigma_{LL}(\omega_n)}{(i\omega_n + \Sigma_{LL}(\omega_n))^2 + \Sigma^2_{LR}(\omega_n)}, \\ G_{LR}(\omega_n) &= \frac{\Sigma_{LR}(\omega_n)}{(i\omega_n + \Sigma_{LL}(\omega_n))^2 + \Sigma^2_{LR}(\omega_n)}. \end{aligned} \tag{104}$$

Except for the special case $q = 2$, for which the self-energies are linear in $G_{ab}(\tau)$ and can be easily transformed to the frequency space, the exact solutions of the SD equations above have to be obtained numerically. Details of the numerical procedure are given in Appendix B.

Depending on the temperature, the saddle-point equations may have more than one solution. The physical solution is the one with the lowest free energy subject to the condition that the steepest descent manifold (Lefschetz thimble) it lies on can be continuously deformed into the original integration manifold. This can be worked out explicitly for $q = 2$, where it turns out that the saddle point with the lowest free energy is not always the one that determines the physical free energy [84]. The phase transition occurs when the solution with the lower free energy switches to a different solution at the critical temperature. The free energy depicted in the left panel of Fig. 4 clearly illustrates this hysteresis mechanism.

In our case, we shall see that at low temperatures the system is dominated by the solution with a nonzero G_{LR} while at higher temperatures the replica symmetric solution with a vanishing G_{LR} dominates. The two solutions intersect at a point where the system undergoes a first-order phase transition. Interestingly, we see that in the RSB phase the free energy is almost constant. This is suggestive

[2]The omitted correlators G_{RL} and G_{RR} are easy to obtain from G_{LL} and G_{LR}, thanks to the symmetry properties (97).

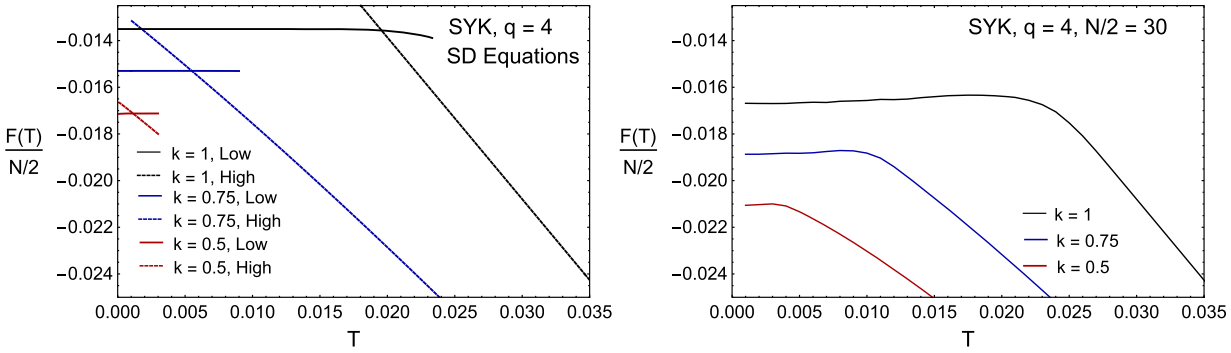

FIG. 4. Left: annealed free energy for different values of k from the solutions of the SD equations. Right: quenched free energy per particle from exact diagonalization of the two-site SYK model for different values of k and $N/2 = 30$.

of the existence of a finite gap between the ground state and the excited spectrum of the effective theory similar to the wormhole phase in the Maldacena-Qi model [37,41]. A comment is in order: from the gravity perspective, it may seem strange that the high-temperature phase depends on the strength k of the imaginary part of the coupling. However, note that this k-dependence can be eliminated by an overall rescaling of the Hamiltonian by a function of k as we did for the Ginibre case. For the sake of simplicity, we stick with the Hamiltonian (56).

Finally, we study the differences between the annealed free energy, obtained from the solution of the saddle-point SD equations (see Fig. 4, left) and the quenched free energy (see Fig. 4, right, and Fig. 5), which is accessible by an exact diagonalization of the Hamiltonian. For the latter, due to technical limitations, we have considered $N/2 \leq 34$. The number of disorder realizations is such that for any given N and k at least 10^6 eigenvalues are obtained. It is well known [32,34] that the SYK model requires a relatively large

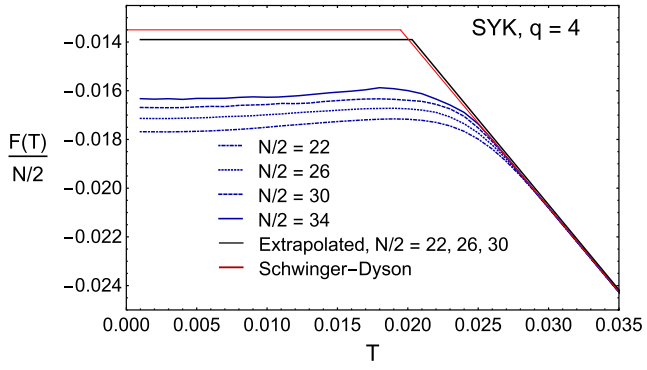

FIG. 5. Quenched free energy per particle obtained from exact diagonalization of the two-site SYK model for $k = 1$ and different values of N. We also include the extrapolated, $N/2 \to \infty$, free energy resulting from a finite size scaling analysis of the numerical results (black curve) together with the large N prediction from the solution of the SD saddle-point equations (red curve). Agreement between the two results is excellent. See the main text and Fig. 6 for details and limitations of the finite size scaling analysis.

number of fermions, $N \geq 30$ in most cases, to approach the thermodynamic limit. For that reason, we have also carried out a finite size scaling analysis with a fitting function (for $k = 1$),

$$F(\varepsilon, T_c, w) = \frac{\int_0^\infty ds e^{-(s-T)^2/2w^2}[\theta(T_c - s)\varepsilon - \theta(s - T_c)s\log 2]}{\int_0^\infty ds e^{-(s-T)^2/2w^2}}, \tag{105}$$

which provides an excellent global fit for $N \leq 30$. The accuracy of the free energy data for $N = 34$ reduces the quality of the fit, and we did not include it in the $N \to \infty$ extrapolation. However, the $N = 34$ data are within fluctuations of the extrapolation from $N = 22$, $N = 26$, and $N = 30$. In Fig. 6, we show the dependence of ε (left), T_c (middle), and w (right) on $1/N$.

The large deviation from linearity for $N = 34$ are due to fluctuations close to the critical temperature, which can only be suppressed by increasing the size of the ensemble well beyond our computational resources. The extrapolated free energy agrees well with the result from solving the Schwinger-Dyson equations (compare the red and the black curves in Fig. 5). Also, the extrapolated values of the physical parameters $\varepsilon(k = 1) = -0.0139$ and $T_c = 0.0203$ are in agreement with results from the Schwinger-Dyson equation where $\varepsilon(k = 1) = -0.0135$ and $T_c = 0.0197$. In the thermodynamic limit at fixed $q = 4$, the values of these parameters are equal to $\varepsilon(k = 1) = -0.0147$ and $T_c = 0.0225$. The parameter $\varepsilon(k)$ is well determined by the global fit and can also be obtained from extrapolating at a single temperature well below T_c, where we can also include the $N = 34$ data. The critical temperature then follows from the intersection point with the high-temperature curve, which gives

$$T_c = \frac{\gamma(k) - \varepsilon(k)}{S_0}, \tag{106}$$

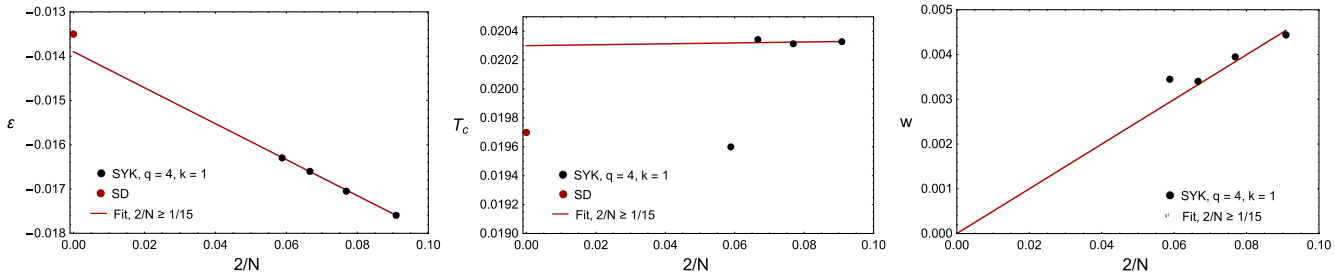

FIG. 6. The fitting parameters $\varepsilon = \varepsilon(k)$ (left), T_c (middle), and w (right) as a function of $2/N$ for $N = 22$, 26, 30, and 34. The linear dependence works well for all N for ε, but in case of T_c and w, the $N = 34$ values are aberrant. This is due to the limited size of the ensemble (40 realizations) in this case. The result obtained from the SD equations is indicated by the red dot at $2/N \to 0$.

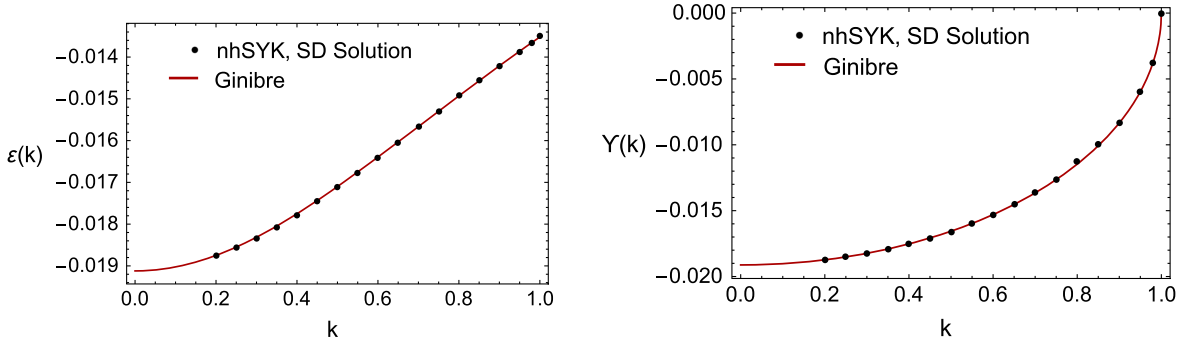

FIG. 7. The intercepts of the free energy with the $T = 0$ axis for the low-temperature phase (left) and the high-temperature phase (right) both obtained from solving the Schwinger-Dyson equations of the non-Hermitian SYK model. The results are compared to analytical formulas for the Ginibre ensemble, where $\varepsilon(k)$ and $\gamma(k)$ are half of the long axis and the focal length of the ellipse containing the eigenvalues of the Hamiltonian. The curves are obtained without fitting with $\varepsilon(0)$ as the only parameter.

with $-S_0$ the slope of the high-temperature curve and $\gamma(k)$ and $\epsilon(k)$ the intercepts of the high-temperature curve and the low-temperature curve with the $T = 0$ vertical axis, respectively. Within the accuracy of our calculations, this finite size scaling analysis gives the same result as obtained from using the finite size scaling form (105).

B. Critical temperature and the ground-state energy

Since the system develops a first-order phase transition, we can study the free energy of both phases separately and use this to determine the critical temperature $T_c(k)$ as a function of k. We start with the low-temperature phase. From Fig. 5, it is clear that the free energy is close to being temperature independent. The intercept of the free energy of the low-temperature phase with the $T = 0$ axis is well determined. In Fig. 7, we show the intercept $\varepsilon(k)$ versus k (black points) and compare it to the k-dependence obtained for the elliptic Ginibre model

$$\varepsilon(k) = \frac{2}{N} E_0(k) = \frac{e_0}{\sqrt{1+k^2}} \tag{107}$$

with e_0 determined by the ground-state energy of the $k = 0$ SYK model. From elementary considerations, it is clear that $-\varepsilon(k)N/2$ is equal to the smallest real part of the eigenvalues. The excellent agreement of the k-dependence shows that the two-site non-Hermitian SYK model is in the universality class of the elliptic Ginibre model.

Next, we consider the free energy of the high-temperature phase. Its intercept with the $T = 0$ axis $\gamma(k)$ is compared to the focal point of the ellipse containing the eigenvalues obtained for the elliptic Ginibre model

$$\gamma(k) = -\frac{2}{N}\sqrt{E_0^2(k) - y_0^2(k)}. \tag{108}$$

Again, the agreement is excellent without any fitting [see Fig. 7 (right)]. The free energy of the high-temperature phase is approximately linear in T, but the entropy per particle [see Fig. 8 (left)] is only equal to $\log 2$ for $k = 1$, contrary to the expectation from the Ginibre model. For $k < 1$, the zero-temperature entropy of the high-temperature phase is a constant equal to the $k = 0$ value of $C/\pi + \frac{1}{4}\log 2$ (red curve in Fig. 8, left) where C is the Catalan constant,[3] but jumps to $\log 2$ close to $k = 1$.

[3]This comes from the zero-temperature entropy density formula [16,18] $S_0 = \frac{1}{2}\log 2 - \int_0^{1/q} \pi(\frac{1}{2} - x)\tan(\pi x)dx$ for the one-site Hermitian SYK model.

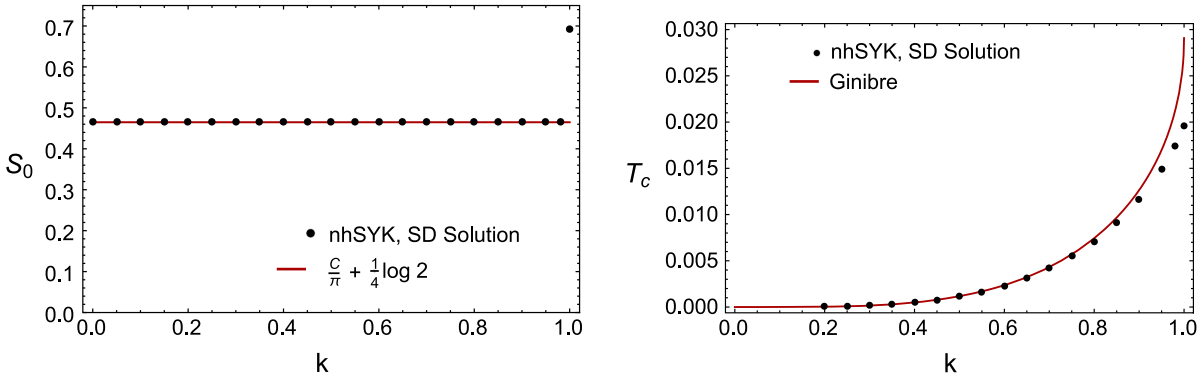

FIG. 8. The zero-temperature entropy per particle $S_0(k)$ vs k for the black hole phase calculated from the slope of the free energy (left), compared to $C/\pi + \frac{1}{4}\log 2$ with C the Catalan constant (0.915966). The right figure compares the critical temperature obtained from the solutions of the Schwinger-Dyson equations to the formula from the elliptic Ginibre ensemble with the high-temperature entropy substituted by the entropy of the curve shown in the left figure. If we would have used the actual values of the entropy for $k = 1$ (which is $\log 2$), the result of the SD equations would have agreed with the result from the Ginibre ensemble.

Even for $k = 0.98$, the zero-temperature entropy is very close to the $k = 0$ value.

The critical temperature is determined by the intersection of the free energy of the low- and high-temperature phases. The results are given in Fig. 8 (right). We also show the result for the Ginibre ensemble (blue curve) that can be obtained from $\epsilon(k)$ and $\gamma(k)$. However, it is clear that this cannot work because the slope of the free energy of the high-temperature phase is always $\log 2$ for the elliptic Ginibre model. If we substitute $\log 2$ by the actual slope for $k < 1$, $S_0(k) = 0.464848$, we obtain

$$T_c(k) = \frac{e_0}{S_0(k)}\left(1 - \frac{\sqrt{1-k^4}}{\sqrt{1+k^2}}\right), \tag{109}$$

which is depicted by the red curve in Fig. 8. If we use the actual value of $S_0(k)$ for $k = 1$, we also find agreement with the result for the Ginibre ensemble. The agreement with the Ginibre ensemble provides strong support to the physical picture of RSB configurations dominating the free energy in the low-temperature limit and inducing a first-order phase transition.

We already have seen that the entropy of the two-black-hole phase is not given by the Ginibre ensemble. Also, for the low-temperature phase, we observe deviations from the Ginibre ensemble, which gives a vanishing entropy. Indeed, if we plot the entropy per particle

$$S(T) = -\frac{2}{N}\frac{dF}{dT} \tag{110}$$

on a log-log scale (see Fig. 9), we find a clear temperature dependence. For each of the four k values, $k = 0.4$, $k = 0.7$, $k = 0.9$, and $k = 1$, we observe a strong first-order phase transition at T_c, discussed earlier in this section. At this point, the entropy per particle jumps from a small positive value to a value in the range $[0.5, \log 2]$ (see the caption of Fig. 9). For $k = 1$, we observe a second critical temperature, T_0, below which the entropy vanishes to the accuracy of the calculation. Between this temperature and T_c, the entropy becomes a small nonzero positive number after first becoming negative. Changing the discretization steps by a factor 2, or even a factor 1000, does not change this picture for $k = 1$. Note that the apparent jump between T_0 and T_c is due to plotting $|S(T)|$ on a log-log scale. For $k = 0.4$, $k = 0.7$, and $k = 0.9$ (in fact for $k < 0.95$), the entropy remains positive, and we plot $S(T)$ rather than $|S(T)|$.

The discretization error is expected to be of second order in the discretization step $\Delta t = \beta/M$. Indeed, at low temperatures, the entropy behaves as $1/T^2$ (see Fig. 9). After Richardson extrapolation (see the black points in Fig. 9),

$$S_{\text{extrapolated}}(T) = 2S(M,T) - S(M/2,T), \tag{111}$$

the leading-order dependence on the step size is canceled, and the discretization error is of order $(\Delta t)^3$. Indeed, the extrapolated result (black dots in Fig. 9) behave as $1/T^3$ for temperatures below the kink (see green lines). We conclude that for temperatures below the kink the nonzero value of the entropy is due to finite size effects. We expect that in the continuum limit the entropy will also vanish in this region for $k < 1$. Because G_{LR} is continuous at $t = 0$ and $t = \beta$, the finite size effects for contributions involving G_{LR} are very small. On the other hand, G_{LL} is discontinuous at $t = 0$ and $t = \beta$, which results in large finite size effects of contributions of G_{LL} to the free energy. This explains why the entropy for $k = 1$, which depends only on G_{LR}, does not depend on the discretization step for a large range of M values. We also notice that by choosing half-integer discretization points the discretization errors are reduced by an order of magnitude with respect to choosing integer discretization points.

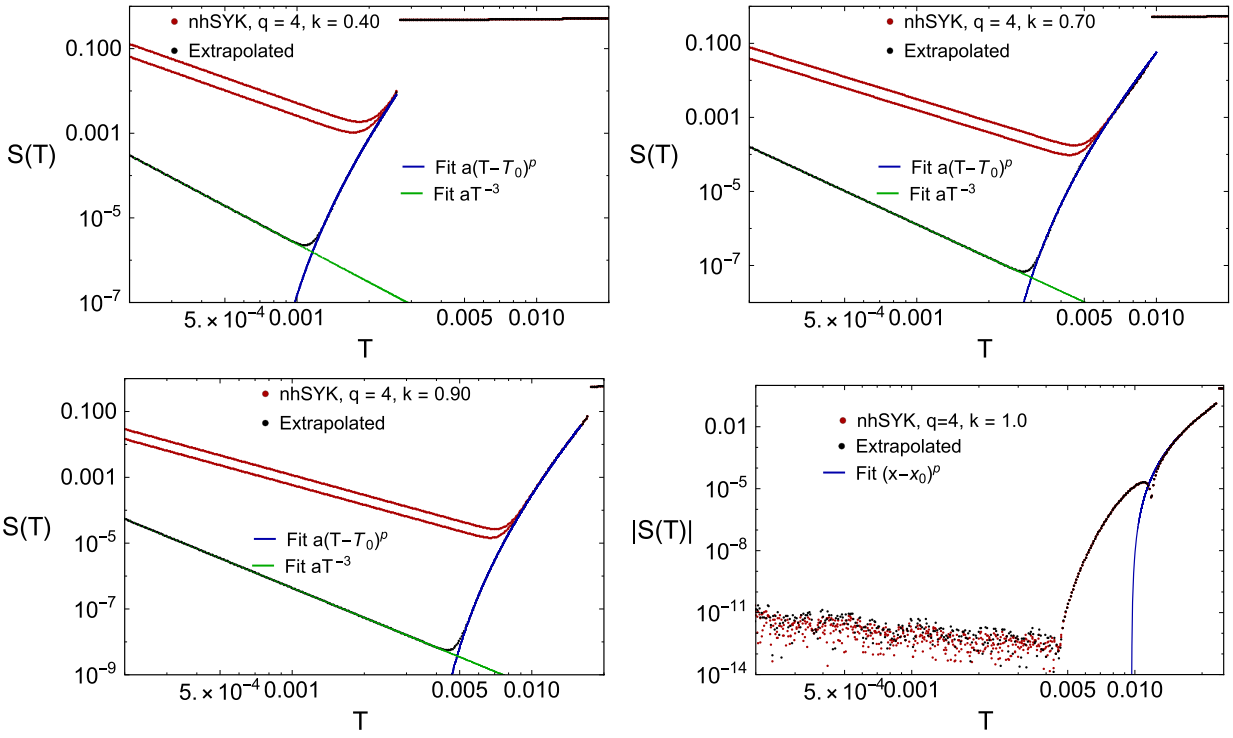

FIG. 9. Log-log plot of the entropy of the two-site non-Hermitian $q = 4$ SYK model as a function of the temperature for $k = 0.4$ (upper left), $k = 0.7$ (upper right), $k = 0.9$ (lower left), and $k = 1$ (lower right). The red dots show results obtained from a numerical solution of the Schwinger-Dyson equations for 10^5 discretization points on $[0, \beta]$, and the black dots are a Richardson extrapolation [85] from the solutions with 5×10^4 and 10^5 discretization points. At the first-order phase transition point, the entropy jumps from 0.010 to 0.474, from 0.032 to 0.506, from 0.070 to 0.569, and from 0.132 to 0.693 (log 2) for $k = 0.4$, $k = 0.7$, $k = 0.9$, and $k = 1$, respectively. For $k = 1$, the entropy becomes negative on the interval [0.0046, 0.0118] and vanishes up to the accuracy of the calculation for $T < 0.0046$. For $k = 1$, the results do not depend on the discretization step until it is increased by a factor of about 1000. The blue curve for $k = 1$ is a fit for $T > 0.0118$ until the first-order transition point.

The entropy is also given by [18] (i.e., using $F = U - TS$)

$$\frac{dF}{dT} = \frac{F}{T} + \mathcal{J}^2\beta^2 \frac{1}{M}\sum_{n=1}^{M} G_{LL}\left(n - \frac{1}{2}\right) + \tilde{\mathcal{J}}^2\beta^2 \frac{1}{M}\sum_{n=1}^{M} G_{LR}\left(n - \frac{1}{2}\right). \tag{112}$$

Each of the terms can be calculated separately from the free energy and the Green's functions. This identity (which is valid at finite M) is satisfied numerically to three or four significant digits (except when the entropy is very small and large cancellations occur in the right-hand side). We could not fully explain why this identity is not satisfied with greater accuracy, but it could be due large finite size effects in G_{LL} for t close to 0 or β. For $k > 0.95$, the entropy becomes negative. However, its magnitude is very small—the monotony of the free energy is only violated by about 10^{-6} of its value. Within a wide range of the parameters, it also does not depend on the size of the discretization step and the convergence criterion. However, we cannot exclude that the negativity of the entropy may be an artifact of the algorithm.

To identify the value of the second critical point, we fit the logarithm of $a(T - T_0)^p$ to the logarithm of the extrapolated entropy between T_0 and T_c well away from the end points to reduce finite size effects. The results are shown in Fig. 10. For $k = 0.98$ and $k = 1$, we only fit in the region where the entropy is positive. Examples of the fitted curves are shown in Fig. 9 (blue curves). Taking these fits at face value would indicate a high-order continuous phase transition. Because of the smallness of the entropy and the magnitude of the finite size effects for $k < 1$, we cannot exclude that such conclusion is an artifact of the algorithm we are using.

C. Decay of $G_{LR}(\tau)$ and the gap

In order to understand the nature of the RSB configurations, we investigate the behavior of $G_{LR}(\tau)$ in more detail.

We are particularly interested in its exponential decay rate with τ, which, for traversable wormholes or weakly coupled two-site SYK models with real coupling (Maldacena-Qi model [41]), is directly related to a gap E_g in the spectrum. Contrary to the Maldacena-Qi model,

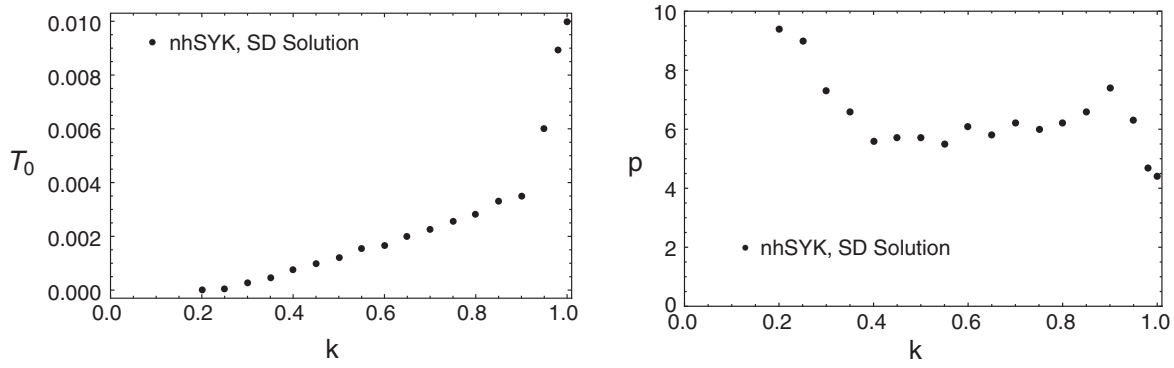

FIG. 10. The fitting parameters T_0 and p as a function of k.

where the gap E_g is equal to the energy difference of the first excited state and the ground state, the non-Hermitian SYK Hamiltonian does not have a genuine gap, but as we will see below, $G_{LR}(\tau)$ still decreases exponentially for large τ.

The gap (decay rate) is computed by fitting the long-time behavior of the propagator $G_{LR}(\tau)$, or $G_{LL}(\tau)$, for sufficiently low temperature with an exponential ansatz. More specifically, taking into account the symmetries of the solution (see Sec. IV A 1 and Appendix B), we employ the ansatz

$$G_{LR}(\tau) \sim \sinh\left(E_g(\beta/2-\tau)\right),$$
$$G_{LR}(\tau) \sim \cosh\left(E_g(\beta/2-\tau)\right), \tag{113}$$

where the gap E_g is a fitting parameter. A feature to note is that these nontrivial solutions for G_{LR} continue to exist for a range of temperatures when they no longer minimize the free energy (see Fig. 4). The two-black-hole solutions exist for all temperatures. We have checked that at small temperatures this ansatz reproduces well the behavior of the propagator for $0 \ll |\tau-\beta/2| \ll \beta/2$ (for small k, it also agrees well for τ around $\beta/2$; see Fig. 11. Close to the end points, we see significant deviations from the free propagator in particular for $G_{LR}(\tau)$.

The value of the decay rate E_g as a function of the coupling k in the low-temperature limit is shown in Fig. 12 (left). For small k, it depends quadratically on k, but there are significant deviations for $k > 0.8$.

Except for very low temperatures, the value of $G_{LR}(0)$ is almost constant as a function of the temperature in the RSB phase and vanishes beyond the critical temperature; see Fig. 13. Therefore, $G_{LR}(0)$ can be considered as the order parameter of a first-order phase transition. The thermodynamic limit of this order parameter will be analyzed in detail in the next subsection. The wormhole solution continues to exist for $T > T_c$ until $T \approx 0.006$. The k-dependence of $G_{LR}(0)$ is shown in the right panel of Fig. 12. The small k behavior can be fitted by $\sim(k-b)^{2/3}$. The value of $b = 0.076$ is consistent with the value of k below which we can no longer obtain a wormhole solution from the SD equations.

Qualitatively, the observation of an exponential decay, see Fig. 13 (left), with a decay rate E_g gives further support to the physical picture of RSB configurations as tunneling events connecting different replicas. This is also the interpretation of wormholes in the gravity partition function. We note that the use of the term "gap" for the decay rate E_g is more by analogy with the traversable wormhole case where it can be demonstrated rigorously that E_g is the difference between the ground and the first excited state.

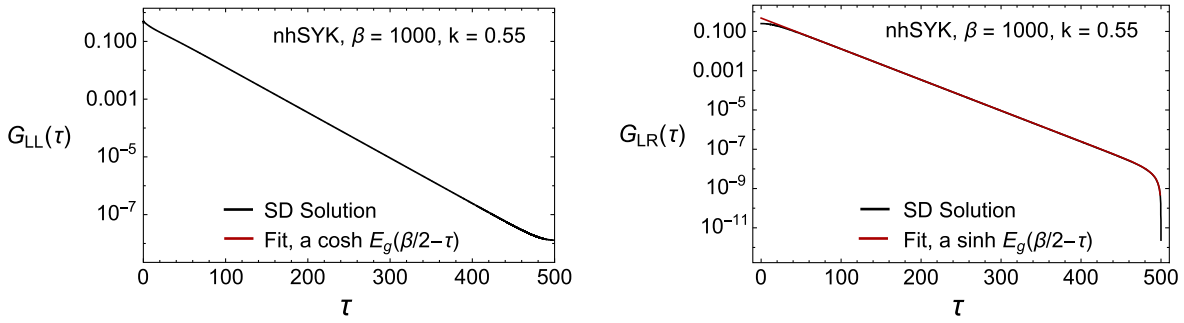

FIG. 11. The nontrivial solutions of the SD equations for G_{LL} (left) and G_{LR} (right), for $T = 0.001$ and $k = 0.55$. The SD solutions (black curve) are compared to the free Green's functions (red curves) with a fitting parameter that sets the scale (the gap E_g) and an overall constant. The exponents of G_{LL} and G_{LR} agree to six significant digits.

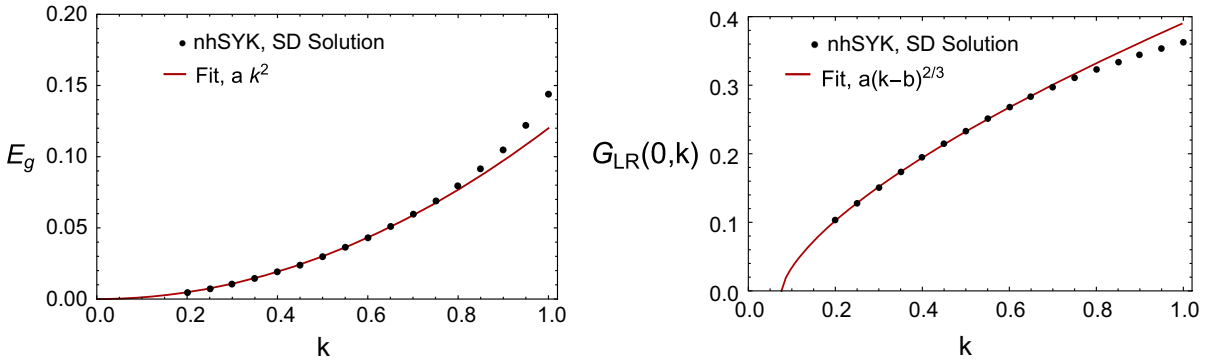

FIG. 12. Top: t energy gap E_g, namely, the rate of exponential decay of $G_{LR}(\tau)$ (left) and $G_{LR}(0)$ (right) as a function of the k calculated from the propagator for $T = 0.0005$. For small k, we observe a quadratic dependence of E_g on k. The exponent of the k-dependence of $G_{LR}(0)$ is not well determined, but an $a(k-b)^{2/3}$ gives a reasonable fit.

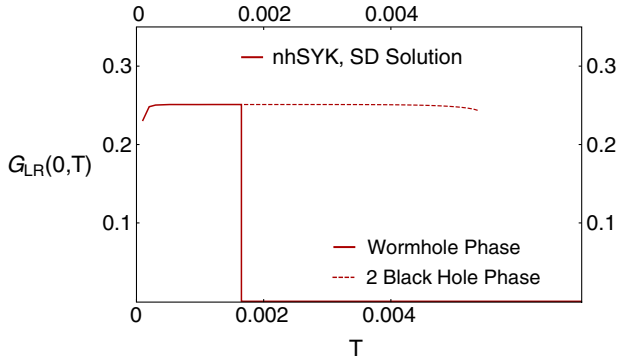

FIG. 13. The behavior of the $G_{LR}(0,T)$, as a function of the temperature for $k = 0.55$. As expected in a first-order transition, it drops to zero abruptly at the transition, which suggests that it can be considered an order parameter of the transition. The dashed red curve indicates the value of $G_{LR}(0,T)$ for $T > T_c$ [41].

In our case, the Hamiltonian is non-Hermitian, and its spectrum does not have a gap. Therefore, this interpretation can only be applicable to the associated replica field theory.

D. Order parameter for the phase transition

In this subsection, we study the behavior of the mixed propagator at the origin, $G_{LR}(0)$. We will show that it is an order parameter of the first-order phase transition discussed before. Since the eigenstates of the Hamiltonian are degenerate for mod $(N/2, 8) \neq 0$, the value of $G_{LR}(0)$ is basis dependent. For example, if the eigenstates are chosen to be also eigenstates of the chirality operators $\gamma_5^L = i^{N/4(N/2-1)} \prod_j^{N/2} \psi_L^j$ and $\gamma_5^R = i^{N/4(N/2-1)} \prod_j^{N/2} \psi_R^j$, the Green's function $G_{LR}(0)$ vanishes identically.

At finite temperature, $G_{LR}(0)$ is given by

$$G_{LR}(0) = \frac{1}{Z}\frac{1}{N/2}\left\langle \mathrm{Tr}\left[e^{-\beta H}\sum_k \psi_k^L \psi_k^R\right]\right\rangle. \quad (114)$$

As discussed before, this quantity vanishes for $\epsilon = 0$ in the action (71) and is purely imaginary for nonzero values of ϵ. The zero-temperature limit is given by the ground-state expectation value

$$\langle 0|S|0\rangle \equiv \langle 0|\frac{1}{N/2}\sum_k \psi_k^L \psi_k^R|0\rangle. \quad (115)$$

For mod $(N/2, 4) = 2$ (Gaussian Unitary Ensemble universality class), we have four degenerate ground states which can be characterized by the chirality of the L (left) and R (right) SYK models (see Appendix E). Eigenstates with these quantum numbers are obtained by adding an infinitesimal term $\sim\gamma_{5,L}\gamma_{5,R}$ to the Hamiltonian. In this basis, the spin operator with basis states $|++\rangle, |--\rangle, |+-\rangle, |-+\rangle$ is given by

$$\langle\chi_L\chi_R|iS|\chi_L\chi_R\rangle = \begin{pmatrix} 0 & i\alpha & 0 & 0 \\ -i\alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & i\beta \\ 0 & 0 & -i\beta & 0 \end{pmatrix}, \quad (116)$$

where $\chi_L, \chi_R = \pm$ are the possible chiralities of the left and right ground states. The antisymmetry follows from the Hermiticity and the anticommutation properties of the gamma matrices. Using the representation

$$\psi_L^k = \gamma_k \otimes 1,$$
$$\psi_R^k = \gamma_5 \otimes \gamma_k, \quad (117)$$

we can write the constants α and β as

$$\alpha = \sum_k \langle +|\gamma_k|-\rangle^2 = 0$$
$$\beta = -\sum_k |\langle -|\gamma_k|+\rangle|^2 \neq 0, \quad (118)$$

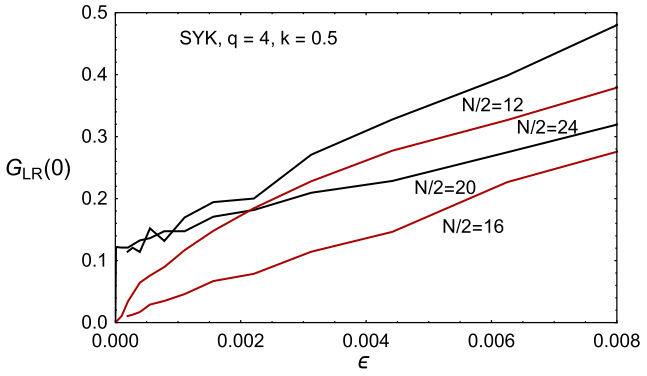

FIG. 14. The ground-state expectation value of S at zero temperature, namely, $G_{LR}(0)$, as a function of ϵ for $k = 0.5$.

and the minus sign is due to the chirality. For the γ_k in (117), we use the representation

$$\gamma_{2k-1} = 2^{-1/2}\overbrace{\sigma_3 \otimes \cdots \otimes \sigma_3}^{k} \otimes \sigma_1 \otimes \overbrace{\sigma_0 \otimes \cdots \otimes \sigma_0}^{\frac{N}{2}-k-1},$$
$$\gamma_{2k} = 2^{-1/2}\overbrace{\sigma_3 \otimes \cdots \otimes \sigma_3}^{k} \otimes \sigma_2 \otimes \overbrace{\sigma_0 \otimes \cdots \otimes \sigma_0}^{\frac{N}{2}-k-1}. \tag{119}$$

For each term contributing to the sum in (118) containing a σ_1, there is a corresponding gamma matrix with a σ_2 at the same position in the tensor product. The nonvanishing matrix elements differ by $\pm i$ so the sum over the squares of the matrix elements, which gives α, vanishes. Since β is equal to the sum of the absolute value of the matrix elements, it does not vanish. The ground state is thus given by

$$|G\rangle = \frac{1}{\sqrt{2}}(|+-\rangle \pm i|-+\rangle). \tag{120}$$

We thus find that for $\epsilon \neq 0$ the ground-state expectation value of $G_{LR}(0) = \mp i\beta$ is nonvanishing. Note that the sign of $G_{LR}(0)$ is determined by the sign of ϵ.

Next, we consider the case mod $(N/2, 8) = 0$. Then, the ground state of the single-site SYK is unique and can be in each of the two chirality sectors. Therefore, the ground state of the two-site SYK model has either $|++\rangle$ or $|--\rangle$ as ground state in terms of the chiralities of each SYK. This means that the expectation value of S vanishes. For a nonzero value of ϵ, a perturbative calculation yields

$$G_{LR}(0) = \frac{\epsilon}{E_{++} - E_{--}} \tag{121}$$

with E_{++} and E_{--} the lowest energy with both chiralities positive or negative, respectively. In the thermodynamic limit, the spacing of $E_{++} - E_{--} \sim \exp[-\frac{1}{2}NS_0]$ with S_0 the zero-temperature entropy density, and a finite value is possible if the thermodynamic limit is taken before the limit $\epsilon \to 0$.

Finally, we consider the case mod $(N/2, 8) = 4$ (Gaussian Symplectic Ensemble universality class). In this case, the levels of each SYK are doubly degenerate (see Appendix E). Therefore, we have four degenerate ground states. Since the charge conjugation matrix is the product of an even number of gamma matrices, both states of a Kramer's degenerate pair have the same chirality. We conclude that all four ground states have either the chirality $|++\rangle$ or $|--\rangle$ so that the ground-state expectation value of S vanishes. For small ϵ, the component of the wave function with the opposite chiralities is again given by first-order perturbation theory, and $G_{LR}(0)$ is again of the form (121). We expect to obtain a finite value if the thermodynamic limit is taken before the $\epsilon \to 0$. These arguments also apply to the original Maldacena-Qi (MQ) model with real couplings.

In Fig. 14 (left), we show the ϵ-dependence of $G_{LR}(0)$ for different values of N. As was expected, only for $N/2 = 12$ or $N/2 = 24$, we observe spontaneous symmetry breaking.

The theoretical arguments given in this section are consistent with the numerical calculation of $G_{LR}(0)$ from the solution of the SD equations (see Fig. 15), which identifies $G_{LR}(0)$ as the order parameter of a first-order phase transition.

We now show that an exact finite N calculation of the order parameter $G_{LR}(0)$ based on the exact diagonalization of the Hamiltonian yields similar results. Since a

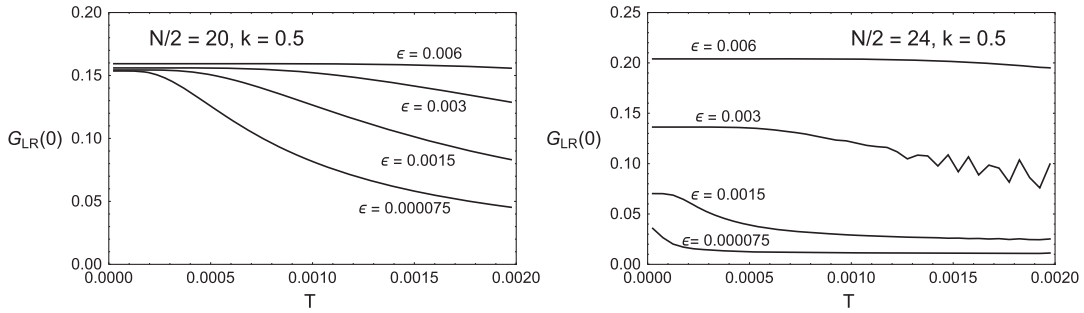

FIG. 15. The temperature dependence of $G_{LR}(0)$ as a function of the temperature T for $k = 0.5$ and $N = 20$ (left) or $N = 24$ (right). The values of ϵ are indicated in the figures.

non-Hermitian matrix can be diagonalized by a similarity transformation, $e = V^{-1}HV$, we have that

$$G_{LR}(0) = \frac{1}{Z}\frac{2}{N}\mathrm{Tr}[e^{-\beta e}V^{-1}SV]. \tag{122}$$

In Fig. 14, we show the zero-temperature limit of this quantity as a function of ϵ for $N/2 = 12$, $N/2 = 16$, $N/2 = 20$, and $N/2 = 24$, all for $k = 0.5$. The random matrix theory universality class of a single SYK is Gaussian Unitary Ensemble (GUE), Gaussian Orthogonal Ensemble (GOE), GUE and Gaussian Symplectic Ensemble (GSE), in this order. In agreement with the above arguments, in the GUE class, the symmetry is spontaneously broken, but in the GOE and the GSE classes, it is not clear whether a finite result can be obtained for large N and small ϵ. The temperature dependence of $G_{LR}(0)$ for $N = 20$ and $N = 24$ is shown in Fig. 15. Again, we observe that the GUE universality class (for $N = 20$) and the GSE universality class (for $N = 24$) behave qualitatively different. In the GUE universality class (left), the low-temperature limit of $G_{LR}(0)$ saturates to a finite value, while the pseudocritical temperature seems to be proportional to ϵ. For the GSE universality class, the zero-temperature value is proportional to ϵ, while the critical temperature seems to scale as ϵ^2. Also, the numerical value of the zero-temperature limit of $G_{LR}(0)$ is below the result obtained from the SD equations, which may be due to the slow convergence of the large N limit.

V. CONCLUSIONS

By downgrading the condition of Hermiticity to only PT symmetry in random quantum systems, the saddle-point equations have RSB solutions connecting different replicas with lower energy than the replica-symmetric ones. The nature of these solutions is strikingly similar to that of wormholes in JT gravity. With the free energy as observable, we have identified a first-order transition in two examples where the dynamics is quantum chaotic, the two-site Ginibre model and the two-site non-Hermitian SYK model.

The free energy of the two-site non-Hermitian SYK model was calculated in two ways: by explicit diagonalization of the Hamiltonian at finite N and by solving the Schwinger-Dyson equations in the thermodynamic limit. A strong first-order transition at finite T separates a low-temperature phase where the free energy is dominated by RSB configurations (the wormhole phase) from the high-temperature phase controlled by replica symmetric configurations (the two-black-hole phase). Although we did not present explicit results relating the infrared limit of this SYK model with a gravity theory, this transition is reminiscent of an Euclidean wormhole-to-black-hole transition [58]. The solutions of the Schwinger-Dyson equations also indicate that there is a second phase transition at a lower temperature, which is continuous, below which the free energy becomes strictly constant. In between these two phase transitions, the free energy increases rapidly until it jumps to the value of the two-black-hole phase at the first-order phase transition temperature. Because of substantial finite size effects, only for $k = 1$, the existence of a phase with a constant free energy is established unambiguously. For $k < 1$, we cannot exclude that the derivative of the free energy remains nonvanishing all the way to zero temperature. Another remarkable observation is that the zero-temperature entropy of the black hole phase does not depend on the degree of non-Hermiticity as long as $k < 1$ but jumps to the high-temperature value[4] at $k = 1$. We have no good explanation for what causes the discontinuity of the k-dependence of the entropy.

Although we have restricted our analysis to $q = 4$, we expect that this transition is universal, provided that $q > 2$ when the dynamics is quantum chaotic, and therefore its spectral correlations are expected to be those of the Ginibre ensemble. This expectation is based on the following facts:

(i) We have found excellent agreement between the two-site elliptic Ginibre ensemble and the $q = 4$ two-site non-Hermitian SYK model—in some sense, the Ginibre model is an SYK model with $q \sim N$.

(ii) For real couplings, spectral correlation does not depend qualitatively on the value of $q > 2$.

An important conceptual question arises: for this first-order phase transition to happen, is it a requirement that the dynamics is quantum chaotic? Indeed, for the two-site $q = 2$ SYK model which is integrable [18,84], we only have second-order phase transitions [84]. However, as is shown in Appendix F, for uniform uncorrelated random eigenvalues inside the complex unit disk, we do find a first-order phase transition, but in the low-temperature phase, the free energy depends linearly on the temperature.

The results of this paper open several promising research avenues. First of all, it remains to establish unambiguously the existence of the second continuous phase transition mentioned above. Because of the smallness of the entropy in the wormhole phase, this requires a new algorithm for solving the SD equations that greatly reduces the finite size effects, allowing us to study the details of this phase transition. An improved algorithm will also help us to analyze the discontinuity of the k-dependence of the entropy.

It would also be interesting to explore whether a generalized Schwarzian is still the effective description of the infrared limit of both a perturbed JT gravity theory (related to Euclidean wormholes) and the two-site non-Hermitian SYK investigated in this paper. If this is the case, like for traversable wormholes [41], it would be a strong indication that RSB configurations are the field theory

[4]The high-temperature value of the free energy is given by $F = -\frac{N}{2}T\log 2$ resulting in a high-temperature entropy per particle [defined in Eq. (110)] of $\log 2$.

equivalent of Euclidean wormholes which may be relevant in the solution of the factorization problem in holography. It would also be interesting to include in our model an explicit Maldacena-Qi coupling in order to study transition/crossover from Euclidean to traversable wormholes [80].

In light of our results, and recent developments in the resolution of the information paradox [86,87], an interesting research direction is to investigate the growth of entanglement entropy in a setting based on the non-Hermitian SYK. Of special interest is the contribution of multireplica wormholes in the late stages of the time evolution. Another problem that deserves further attention is a more exact delimitation of the conditions to observe RSB configurations even within PT-symmetric systems. Is quantum chaos always a necessary and sufficient condition, beyond the $q = 2$ example discussed? If not so, is it possible to characterize the existence of RSB configurations as a function of the range of interactions? Are many-body correlations important, or can similar results be obtained in single-particle, noninteracting two-site disordered systems? Is the extension of these results to higher spatial dimensions straightforward? If so, does the existence of RSB configurations depend on the strength of disorder or the hopping range in real space? Can they occur in the presence of Anderson or many-body localization? We plan to address some of these questions in the near future.

ACKNOWLEDGMENTS

J. J. V. and Y. J. acknowledge partial support from U.S. DOE Grant No. DE-FAG-88FR40388. Y. J. is also partly funded by an Israel Science Foundation center for excellence grant (Grant No. 2289/18), by Grant No. 2018068 from the United States-Israel Binational Science Foundation (BSF), by the Minerva foundation with funding from the Federal German Ministry for Education and Research, by the German Research Foundation through a German-Israeli Project Cooperation (DIP) grant "Holography and the Swampland," by a Koshland fellowship, and by a research grant from Martin Eisenstein. A. M. G. was partially supported by the National Natural Science Foundation of China (NSFC) (Grant No. 11874259) and by the National Key R&D Program of China (Project No. 2019YFA0308603) and also acknowledges financial support from a Shanghai talent program. D. R. acknowledges the support by the Institute for Basic Science in Korea (Grant No. IBS-R024-D1). A. M. G. acknowledges illuminating correspondence with Victor Godet, Zhenbin Yang, Juan Diego Urbina, and Klaus Richter.

APPENDIX A: CALCULATION OF THE FREE ENERGY FOR THE TWO-SITE GINIBRE ENSEMBLE

In this Appendix, we evaluate the partition function of the two-site Ginibre model. For a Ginibre ensemble of $D \times D$ matrices, the eigenvalue kernel is given by [77,88]

$$K(z_1, z_2) = \frac{e^{-z_1 z_2^*}}{\pi} \sum_{k=0}^{D-1} \frac{(z_1 z_2^*)^k}{k!}. \tag{A1}$$

The eigenvalue density reads

$$\rho(z) = K(z, z), \tag{A2}$$

and the connected two-point correlation function is equal to

$$\rho_{2,c}(z_1, z_2) = -K(z_1, z_2)K(z_2, z_1) + \delta(z_1 - z_2)K(z_1, z_1), \tag{A3}$$

where the second term is due to the self-correlations. The spectral density is normalized to D, and the eigenvalues are located in a circle of radius $\sqrt{D}$. To be able to adjust the overall scale of the eigenvalues, we include a factor σ in the definition of the partition function

$$Z_2(\beta) = \int d^2 z_1 d^2 z_2 \rho_{2,c}(z_1, z_2) e^{-\beta(z_1 + z_2^*)/\sigma} + |Z_1(\beta)|^2. \tag{A4}$$

The second contribution is due to the disconnected part of the two-point function with Z_1 given by

$$Z_1(\beta) = \int d^2 z \rho(z) e^{-\beta z/\sigma}. \tag{A5}$$

We first calculate the disconnected contribution. The one-site partition function requires the integral

$$Z_1(\beta) = \int \frac{d^2 z}{\pi} e^{-\beta z/\sigma} e^{-|z|^2} \sum_{k=0}^{D-1} \frac{(zz^*)^k}{k!}. \tag{A6}$$

Only the first term of the Taylor expansion of $\exp(-\beta z)$ gives a nonvanishing result, and after changing to polar coordinates, we obtain

$$Z_1(\beta) = \int_0^\infty ds e^{-s} \sum_{k=0}^{D-1} \frac{s^k}{k!} = D. \tag{A7}$$

Next, we calculate the contribution due to self-correlations. It is given by

$$\begin{aligned} Z_{\text{self}}(\beta) &= \frac{1}{\pi} \int d^2 z e^{-\beta(z+z^*)/\sigma} e^{-|z|^2} \sum_{m=0}^{D-1} \frac{(zz^*)^m}{m!} \\ &= \sum_{m=0}^{D-1} \sum_{l=0}^{\infty} \left(\frac{\beta}{\sigma}\right)^{2l} \frac{(m+l)!}{m!l!l!} \\ &= \sum_{m=0}^{D-1} e^{\beta^2/\sigma^2} L_m^0(-\beta^2/\sigma^2) \\ &= e^{\beta^2/\sigma^2} L_{D-1}^1(-\beta^2/\sigma^2), \end{aligned} \tag{A8}$$

where we have used a summation formula for associated Laguerre polynomials:

$$\sum_{m=0}^{n} L_m^a(x) = L_n^{a+1}(x). \tag{A9}$$

The asymptotic behavior of the Laguerre polynomials is given by

$$L_n^\alpha(-x) \sim n^{\alpha/2}(-x)^{-\alpha/2} e^{-x/2} J_\alpha(2i\sqrt{nx}). \tag{A10}$$

For $\alpha = 1$ and $x = \beta^2/N$, this gives

$$L_D^1(-\beta^2/\sigma^2) \sim \frac{\sigma}{\beta}\sqrt{D} e^{-\beta^2/2\sigma^2} I_1(2\beta\sqrt{D}/\sigma). \tag{A11}$$

It is instructive to calculate the large D asymptotics by expressing the sum in the first line of (A8) as an incomplete Γ-function:

$$Z_{\text{self}}(\beta) = 2\int_0^\infty r dr \frac{\Gamma(D, r^2)}{\Gamma(D)} I_0(2\beta r/\sigma). \tag{A12}$$

For large D, the incomplete Γ-function can be approximated by

$$\frac{\Gamma(D, r^2)}{\Gamma(D)} \approx \frac{1}{2}\text{erfc}((r - \sqrt{D})\sqrt{2}). \tag{A13}$$

Inserting this in Eq. (A12), we obtain after a partial integration

$$\begin{aligned} Z_{\text{self}}(\beta) &= \frac{\sigma}{\beta}\sqrt{\frac{2}{\pi}}\int_0^\infty r dr e^{-2(r-\sqrt{D})^2} I_1(2\beta r/\sigma) \\ &\approx \frac{\sigma}{\beta}\sqrt{\frac{D}{2}} e^{\beta^2/2\sigma^2} I_1(2\beta\sqrt{D}/\sigma). \end{aligned} \tag{A14}$$

For large D, we have that $\frac{1}{2}\text{erfc}((r-\sqrt{D})\sqrt{2}) \to \theta(\sqrt{D}-r)$, but with this asymptotic, we would have missed the $\exp(\beta^2/2\sigma^2)$ factor.

The contribution of the genuine two-point correlations can be worked out in the same way. We obtain

$$\begin{aligned} &-\frac{1}{\pi^2}\int d^2z_1 d^2z_2 e^{-(\beta/\sigma)(z_1+z_2^*)} e^{-|z_1|^2-|z_2|^2} \\ &\times \sum_{m=0}^{D-1}\frac{(z_1 z_2^*)^m}{m!}\sum_{n=0}^{D-1}\frac{(z_2 z_1^*)^n}{n!} \\ &= -\sum_{n=0}^{D-1}\sum_{m=0}^{n}\binom{n}{m}\frac{1}{m!}\left(\frac{\beta^2}{\sigma^2}\right)^m \\ &= -\sum_{n=0}^{D-1} L_n^0(-\beta^2/\sigma^2) \\ &= -L_{D-1}^1(-\beta^2/\sigma^2), \end{aligned} \tag{A15}$$

where we have again used the summation formula (A9). We are interested in a scaling limit where $\beta/\sigma \ll 1$. Then, we can expand the exponent in (A8). The first term in the expansion is canceled by the two-point correlations (A15). For the partition function, we then obtain the result

$$Z_2(\beta) = D^2 + \frac{\beta^2}{\sigma^2} L_{D-1}^1(-\beta^2/\sigma^2). \tag{A16}$$

To obtain a free energy density $\log Z/N$ that is stable in the large N limit, we choose $D = 2^{N/2}$. Inserting on the right-hand side the large D limit of the Laguerre polynomial given in (A11), we find

$$Z_2(\beta) = D^2 + \sqrt{D}\frac{\beta}{\sigma} I_1(2\beta\sqrt{D}/\sigma). \tag{A17}$$

This contribution scales in the same way with N as the disconnected part if we choose $\sigma \sim \sqrt{D}/N$. Ignoring logarithmic corrections, we obtain the free energy

$$\begin{aligned} \frac{F(T)}{N} &= -\frac{T}{N}\log\left(D^2 + I_1(2\beta\sqrt{D}/\sigma)\right) \\ &\approx -\frac{T}{N}\log\left(D^2 + e^{2\beta\sqrt{D}/\sigma}\right). \end{aligned} \tag{A18}$$

The two exponents are equal at

$$T_c = \frac{2\sqrt{D}/\sigma}{2\log D}. \tag{A19}$$

Choosing $\sigma = \sqrt{D}/N$, we obtain $T_c = 2/\log 2$, and the free energy is equal to

$$\frac{F(T)}{N} = -\frac{T}{\log 2}\theta(T - T_c) - 2\theta(T_c - T), \tag{A20}$$

in agreement with the $k \to 1$ limit of (56).

The integrals for the contributions of the self-correlations (A8) and genuine two-point correlations (A15) generally cannot be calculated exactly, and we must rely on an approximate evaluation. To do that, we assume that the eigenvalue correlations are in the universality class of the Ginibre ensemble. Noting that the integral over z_1 and z_2 can be written as an integral over the center of mass and the difference of z_1 and z_2, we assume that the integral of $z_1 - z_2$ can be extended to the entire complex plane, while the integral over $(z_1 + z_2)/2$ is replaced by the large D limit of the spectral density. Since the eigenvalue correlations are short ranged, we expect that the extension of the integral over $z_1 - z_2$ to the entire complex plane and using the large D limit of the two-point function is a good approximation. However, we have seen earlier in this section that this approximation does not determine the exponent α of the prefactor $\exp[\alpha\beta^2/\sigma^2]$. In particular,

contributions from the boundary of the eigenvalue disk may change the value of α, but this does not affect the free energy in the thermodynamic limit.

In the last part of this section, we compare the exact and approximate expressions for the replica breaking part of the partition function. For the approximate calculation of the contribution of the self-correlations to the partition function, we replace the spectral density by its large D limit

$$\frac{1}{\pi}e^{-|z|^2}\sum_{k=0}^{D-1}\frac{(zz^*)^k}{k!}\to\frac{1}{\pi}\theta(\sqrt{D}-|z|), \tag{A21}$$

which after changing to polar coordinates results in

$$\begin{aligned}Z_{\text{self,app}}&=\frac{1}{\pi}\int_0^{\sqrt{D}}rdr\int_0^{2\pi}d\phi e^{-\frac{2r\beta}{\sigma}\cos\phi}\\&=2\int_0^{\sqrt{D}}rdrI_0(2r\beta/\sigma)=\frac{\sigma\sqrt{D}}{\beta}I_1(2\beta\sqrt{D}/\sigma).\end{aligned} \tag{A22}$$

This differs by factor $\exp(\beta^2/2\sigma^2)$ from the exact result (A14), and we have seen earlier in this section that this is due to contributions from the boundary of the eigenvalue region.

The approximate result for the contribution of the genuine two-point correlations is given by

$$\begin{aligned}Z_{c,\text{app}}=&-\frac{1}{\pi^2}\int_{|z|<\sqrt{D}}d^2\bar{z}\\&\times\int d^2\eta e^{-(\beta/\sigma)(\bar{z}+\bar{z}^*+i\text{Im}(z_1-z_2))}e^{-|z_1-z_2|^2},\end{aligned} \tag{A23}$$

where $\bar{z}=(z_1+z_2)/2$ and $\eta=z_1-z_2$, and the two-point correlation function is replaced by its large D limit. The integral over the center of mass is replaced by an integral over a disk with constant density which is justified in the large D limit. The integral over $\text{Im}(z_1-z_2)$ can be performed by completing squares, while the integral over $\text{Re}(z_1-z_2)$ is a simple Gaussian. This results in

$$Z_{c,\text{app}}=-\frac{e^{-\beta^2/4\sigma^2}}{\pi}\int_{|z|<\sqrt{D}}d^2\bar{z}e^{-(\beta/\sigma)(\bar{z}+\bar{z}^*)}. \tag{A24}$$

The integral over the center of mass is the same as the integral that enters in the calculation of the contribution of the self-correlations. We thus find

$$Z_{c,\text{app}}=-e^{-\beta^2/4\sigma^2}\frac{\sigma\sqrt{D}}{\beta}I_1(2\beta\sqrt{D}/\sigma), \tag{A25}$$

which also does not reproduce the prefactor $\exp[-\beta^2/\sigma^2]$ of the large D limit of the exact calculation (A15). The sum of the approximate result for the contribution of the self-correlation and the contribution of the genuine two-point correlations is given by

$$Z_{\text{self,app}}+Z_{c,\text{app}}=\frac{\beta^2}{4\sigma^2}\frac{\sigma\sqrt{D}}{\beta}I_1(2\beta\sqrt{D}/\sigma), \tag{A26}$$

which differs by a factor 4 from the exact result. However, this prefactor does not contribute to the large D limit of the free energy.

In the main text, we used the same approximation to calculate the partition function for the elliptic Ginibre ensemble, and we expect that also in that case only constant prefactor is affected by our approximations.

APPENDIX B: NUMERICAL SOLUTION OF THE SD EQUATIONS FOR THE NON-HERMITIAN SYK

We proceed iteratively as follows:

(i) Start from an appropriate initial ansatz for $G_{LL}(\omega_n)$ and $G_{LR}(\omega_n)$. To be explicit, we found that the disconnected solutions, dominant at large temperatures, could be easily reached by starting with two copies of the free ansatz, used to solve the standard SYK model. On the other hand, RSB solutions can be found starting from the solution for the two-site model in the presence of an explicit coupling and sending the coupling to zero.

(ii) Using a fast Fourier algorithm, one computes the corresponding correlator in the time domain.

(iii) The self-energies are calculated according to (102), and then $\Sigma_{LL}(\omega_n)$ and $\Sigma_{LR}(\omega_n)$ are obtained by an inverse Fourier transform.

(iv) The propagators in frequency space are updated using the weighted rule

$$\begin{aligned}G_{LL}^{\text{new}}(\omega_n)&=(1-x)G_{LL}^{\text{old}}(\omega_n)-x\frac{i\omega_n+\Sigma_{LL}(\omega_n)}{(i\omega_n+\Sigma_{LL}(\omega_n))^2+\Sigma_{LR}^2(\omega_n)},\\G_{LR}^{\text{new}}(\omega_n)&=(1-y)G_{LR}^{\text{old}}(\omega_n)+y\frac{\Sigma_{LR}(\omega_n)}{(i\omega_n+\Sigma_{LL}(\omega_n))^2+\Sigma_{LR}^2(\omega_n)},\end{aligned} \tag{B1}$$

where x and y are real parameters between 0 and 1 introduced to prevent over-relaxation. In practice, we initially fix them at 0.5 and further reduce them if the updated propagators start running away.

(v) The procedure is repeated until we reach convergence when difference of the absolute value of the Fourier coefficients becomes less than ϵ. In most calculations, we take $\epsilon = 10^{-10}$. We have checked the convergence by taking ϵ as small as 10^{-14}.

The time domain $[0,\beta]$ is discretized as $\tau_k = \beta(k-1/2)/M$, $k = 1,\ldots,M$, where M is chosen to be 2×10^5 and 10^5, which allows us to extrapolate to the continuum limit. The corresponding Matsubara frequencies are equal to

$$\omega_n = \frac{2\pi(n+\frac{1}{2})}{\beta} \tag{B2}$$

with $n = -M/2, -M/2+1, \cdots M/2-1$.

1. Symmetry properties of the correlators

As we have shown in the main text, the propagators $G_{LL}(\tau)$ and $G_{LR}(\tau)$ satisfy the following symmetry properties:

(i) $G_{LL}(\tau)$ is real and is symmetric about $\beta/2$.
(ii) $G_{LR}(\tau)$ is purely imaginary and is antisymmetric about $\beta/2$.

Let us analyze the implications of the above conditions on the Fourier components of $G_{LR}(\tau)$ since they are relevant for the calculation of the order parameter. The Fourier decomposition of the propagator given by

$$G_{LR}(\tau) = \frac{1}{\beta}\sum_{n=-M/2}^{M/2-1} \exp\left(-i\frac{2\pi}{\beta}\left(n+\frac{1}{2}\right)\tau\right) G_{LR}^{(n)} \tag{B3}$$

is purely imaginary if

$$\sum_{n=-M/2}^{M/2-1} \exp\left(-i\frac{2\pi}{\beta}\left(n+\frac{1}{2}\right)\tau\right) G_{LR}^{(n)} = -\sum_{m=-M/2}^{M/2-1} \exp\left(i\frac{2\pi}{\beta}\left(m+\frac{1}{2}\right)\tau\right)\overline{G_{LR}^{(m)}}. \tag{B4}$$

This requires that the Fourier coefficients satisfy

$$G_{LR}^{(m)} = -\overline{G_{LR}^{(-m-1)}}. \tag{B5}$$

Similarly, we can impose the condition that G_{LR} be antisymmetric about $\beta/2$. This results in the requirement that the coefficients $G_{LR}^{(m)}$ are *purely imaginary*

$$G_{LR}^{(m)} = -\overline{G_{LR}^{(m)}}. \tag{B6}$$

APPENDIX C: SPECTRAL DENSITY OF THE NON-HERMITIAN SYK MODEL AND Q-HERMITE POLYNOMIALS

In this Appendix, we discuss the distribution of the real and imaginary parts of the eigenvalues. In Fig. 16, we show results for $k = 0.3$, $k = 0.75$ and $k = 1$ (see caption). For $k = 1$, the distribution of the real part (left) is the same as the distribution of the imaginary part (right). The real and imaginary parts of the eigenvalues have been normalized by the length of the long and short axes of the ellipses containing the eigenvalues. The eigenvalue distribution is fitted to the Q-Hermite spectral density using the Q-parameter η as a fitting parameter. For $k = 0$, this parameter is given by

$$\eta = \binom{N/2}{4}^{-1}\sum_{m=0}^{4}(-1)^m\binom{4}{m}\binom{N/2-4}{4-m}, \tag{C1}$$

which is equal to $\eta = 0.233$ for $N/2 = 30$. The values we obtain by fitting are lower, in particular for the distribution of the imaginary part; see the legend of Fig. 16. We also point out that for $k = 0.3$ the distribution of the imaginary parts of the eigenvalues is very close to semicircular. We expect that this also will be the case for $k < 0.3$.

APPENDIX D: PARTITION FUNCTION FOR THE SYK MODEL AT $k > 1$

In this Appendix, we relate the free energy for $k > 1$ to the free energy for $k < 1$. The single-site Hamiltonian for $k > 1$ can be written as

$$H_L(k) = H_1 + ikH_2 = ik\left(H_2 - \frac{i}{k}H_1\right). \tag{D1}$$

Since the probability distribution of the Hamiltonian is invariant for $H_1 \to -H_2$ and $H_2 \to H_1$, as far as ensemble averaged observables are concerned, the Hamiltonian of this ensemble can be written as

$$ik\left(H_1 + \frac{i}{k}H_2\right) = ikH_L(1/k). \tag{D2}$$

Note, however, the above change of variables produces an overall minus sign for H_R; therefore, the relation for the two-site Hamiltonian reads

$$H(k) = ik[H_L(1/k) - H_R(1/k)]. \tag{D3}$$

The one-site partition function is equal to

$$Z_L(\beta) = \langle \mathrm{Tr} e^{-\beta H_L(k)}\rangle = \langle \mathrm{Tr} e^{-i\beta k H_L(1/k)}\rangle, \tag{D4}$$

so it is equivalent to the partition function for non-Hermiticity parameter $1/k$ at an imaginary inverse

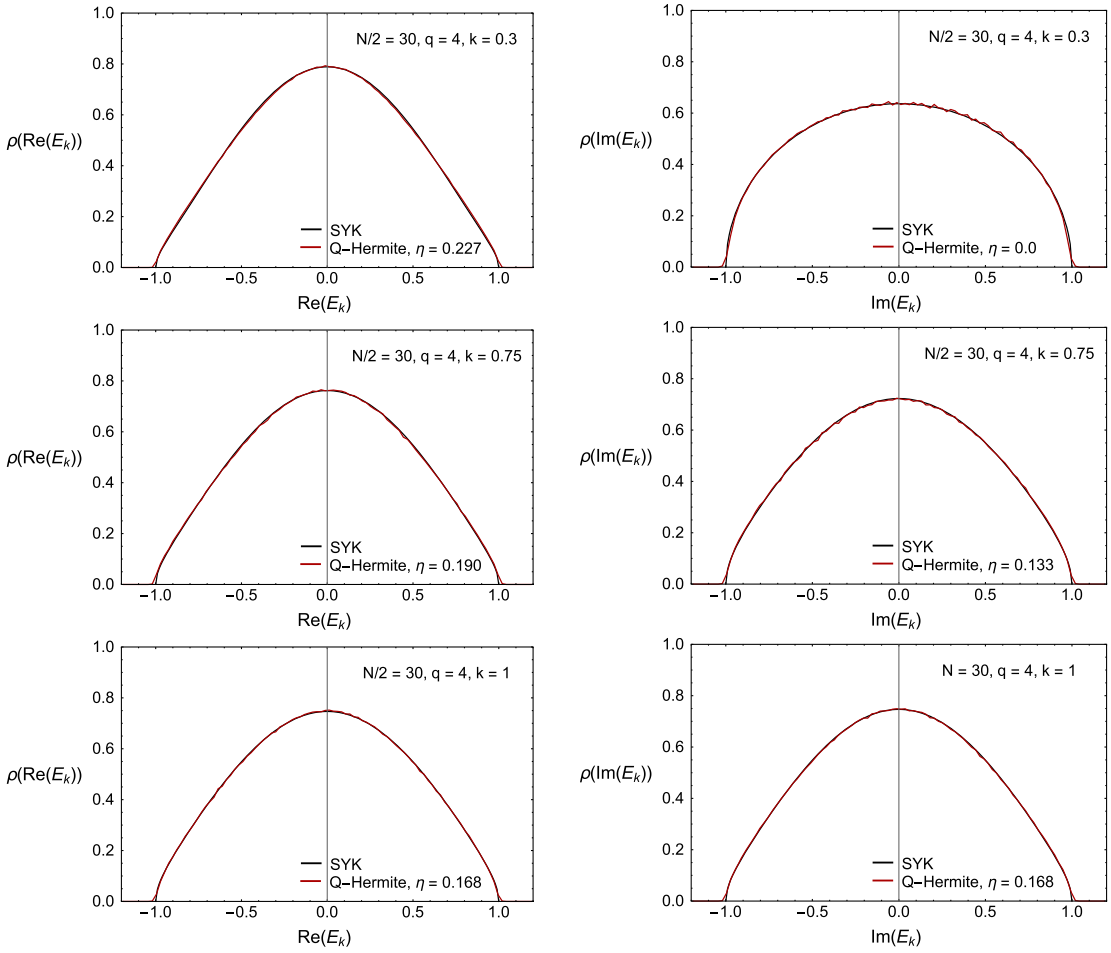

FIG. 16. The distribution of the real (left) and imaginary parts of the eigenvalues of the non-Hermitian one-site SYK model with $N/2 = 30$, $q = 4$ and k as given in the legend of the figure (black curve). The red curves represent a fit of the Q-Hermite density function with a value of η given in the legend of the figure.

temperature $ik\beta$. For $k \to \infty$, we see from Eq. (D3) that the two-site partition function becomes the spectral form factor of a Hermitian SYK model. For $\beta \to \infty$, the form factor (or the partition function) will be dominated by the self-correlations and is thus given by

$$Z(\beta) = Z_L Z_R = \sum_{mn} e^{i\beta k(E_m - E_n)} \approx D = 2^{N/4}, \qquad \text{(D5)}$$

resulting in a free energy of $F(T)/(N/2) = -T/2 \log 2$ as compared to the high-temperature limit of the free energy given by $F(T)/(N/2) = -T \log 2$. In Fig. 17, we show the free energy as a function of the temperature for $k = 1$ red and $k = 2$ (black). The ratio of the intercepts with the y axis is 1.578, while from the Ginibre ensemble, we get $\sqrt{5/2} = 1.58114$. If we assume that for $k > 1$ the zero-temperature slope of the free energy does not depend on k as is the case for $k < 1$, an estimate for the critical temperature can be obtained by equating $-T/2 \log 2 = E_0 = -0.009849$. This gives $T_c = 0.0284$, which is in

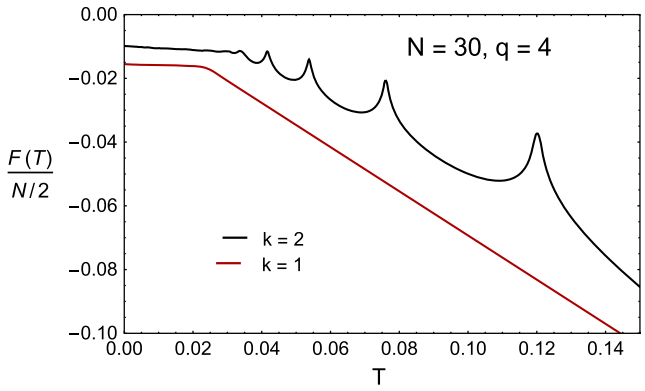

FIG. 17. The quenched free energy per particle as a function of the temperature for the $q = 4$ non-Hermitian SYK model with $N = 30$ and the non-Hermiticity parameters $k = 1$ (red) and $k = 2$ (black). The positions of the peaks are related to the zeros of $J_1(x)$ with the leftmost peak corresponding to the smallest nontrivial zero of $J_1(x)$.

good agreement with Fig. 17. The positions of the peaks are approximately inversely proportional the positions of the zeros of $J_1(x)$ with the leftmost peak corresponding to the smallest nontrivial zero, which will be explained in the next paragraph.

We can work out the $k > 1$ partition function in more detail for the elliptic Ginibre ensemble (which provides a good approximation of the spectra of the non-Hermitian SYK model). Using the scaling (D4), the eigenvalues of this ensemble are distributed homogeneously inside an ellipse given by

$$x = E_0 \cos\phi, \qquad y = y_0 \sin\phi \tag{D6}$$

with E_0 and y_0 given by Eq. (39) with $k \to 1/k$, namely,

$$E_0 = \frac{k\sigma_0}{\sqrt{1+k^2\sigma(1/k)}}, \qquad y_0 = \frac{\sigma_0/k}{\sqrt{1+k^2\sigma(1/k)}}. \tag{D7}$$

The one-site partition function is then given by

$$\begin{aligned} Z_L(\beta) &= \frac{2D}{ik\beta\sqrt{E_0^2-y_0^2}} I_1\left(ik\beta\sqrt{E_0^2-y_0^2}\right) \\ &= \frac{2D}{k\beta\sqrt{E_0^2-y_0^2}} J_1\left(k\beta\sqrt{E_0^2-y_0^2}\right), \end{aligned} \tag{D8}$$

which vanishes at the zeros of the Bessel function. If the zeros of J_1 are given by z_n, the partition function vanishes when

$$\beta\sqrt{E_0^2-y_0^2} = z_n. \tag{D9}$$

The free energy is singular at the critical temperatures

$$T_n = \frac{\sqrt{E_0^2-y_0^2}}{z_n}, \tag{D10}$$

which are of the order of the system size in the normalization that E_0 scales linearly with the number of particles.

Since the exponent of the disconnected part is purely imaginary, it does not contribute to the free energy in the thermodynamic limit. The disconnected part of the two-site partition function is thus given by

$$Z^{k>1}_{\text{disconnected}} \sim D^2. \tag{D11}$$

Using the arguments leading to (29), the connected part of the partition function is equal to

$$Z^{k>1}_{\text{connected}} \sim e^{2\beta|E_0|}. \tag{D12}$$

Equating the two gives the critical temperature

$$T_c(k>1) = \frac{2|E_0(k)|}{\log D^2} = \frac{|E_0(k)|}{N/2\log 2}. \tag{D13}$$

This is the result for the Ginibre ensemble. The temperature dependence of the free energy of the high-temperature phase of the SYK model is not linear in T for $k \neq 1$ (apart from the peaks due to the zeros of the Bessel function). As is the case for $k < 1$, we expect that the zero-temperature slope is smaller than $\log 2$. Earlier in this Appendix, we have argued that it is equal to $\frac{1}{2}\log 2$ resulting in a critical temperature of

$$T_c(k>1) = \frac{2|E_0(k)|}{N/2\log 2}, \tag{D14}$$

which is in good agreement with Fig. 17.

APPENDIX E: DEGENERACIES OF THE NON-HERMITIAN SYK MODELS

Numerical diagonalization reveals that the non-Hermitian one-site SYK model has exactly the same energy level degeneracies as the Hermitian model:

(i) $N/2 \bmod 8 = 0$, $T^2 = 1$, no degeneracy.
(ii) $N/2 \bmod 8 = 2, 6$, $T^2 = \pm 1$, twofold degeneracy.
(iii) $N/2 \bmod 8 = 4$, $T^2 = -1$, twofold degeneracy.

Here, T is the time reversal operator with

$$T = CK \tag{E1}$$

with C the product of the odd or even gamma matrices and K the complex conjugation operator. In the Hermitian case, a proof can be found in Appendix A of Ref. [34]. In the first case, the T operator commutes with the chirality matrix γ_c and with the product of four gamma matrices in the Hamiltonian. This does not impose any conditions on the eigenvalues of the two blocks also in the non-Hermitian case so that the eigenvalues are nondegenerate. In the latter two cases, the degeneracy proof of Ref. [34] needs to be modified for the non-Hermitian case.

1. $N/2 \bmod 8 = 2, 6$

In this case, the degeneracy can be shown by a light modification of the proof of Ref. [34]. The time reversal operator anticommutes with the chirality matrix γ_c, and hence, in the chiral basis, it has the form

$$T = \begin{pmatrix} 0 & cK \\ c^*K & 0 \end{pmatrix}, \tag{E2}$$

where K is the complex conjugation and $cc^* = \pm 1$. Since γ_c commutes with the Hamiltonian, we have

$$H = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{E3}$$

In the non-Hermitian model, time reversal transforms the Hamiltonian as

$$T^{-1}HT = H^\dagger. \tag{E4}$$

Then, Eqs. (E2) and (E3) imply

$$c^{-1}Ac = B^T, \tag{E5}$$

and hence we conclude A and B have the same eigenvalues, so H is generically twofold degenerate. Notice that the degeneracy comes form two different chirality sectors, so $\langle\Omega|\psi_L\psi_R|\Omega\rangle$ in the corresponding two-site model is not necessarily zero if we take $|\Omega\rangle$ to be a linear combination of two different chiralities.

2. $N/2$ mod 8 = 4

This would be the GSE case for the Hermitian SYK model. The twofold degeneracy is a Kramers degeneracy, and the proof is simple for the Hermitian case, which we now briefly recapitulate. Since T is a symmetry in the Hermitian SYK, if $|v\rangle$ is an eigenstate, then $|Tv\rangle$ is also an eigenstate. To prove $|Tv\rangle$ is linearly independent from $|v\rangle$, we use the fact that T is antiunitary and $T^2 = -1$:

$$\langle v|Tv\rangle = \langle v|T^{-1}T^2v\rangle = \langle Tv|T^2v\rangle^* = -\langle Tv|v\rangle^* = -\langle v|Tv\rangle. \tag{E6}$$

Hence, $|Tv\rangle$ is orthogonal to $|v\rangle$, and we conclude twofold degeneracy.

In the non-Hermitian case, T is no longer a symmetry, and we only have

$$T^{-1}HT = H^\dagger, \tag{E7}$$

so $|Tv\rangle$ is not an eigenstate of H even if $|v\rangle$ is. However, the spectrum is still twofold degenerate, and the proof for Kramers degeneracy needs modification. The gist is a proof by contradiction: if there exists an eigenvalue of H that is not degenerate, then T cannot be invertible and hence contradicts its antiunitarity.

Suppose H is a $D \times D$ matrix and the complete set of eigenvectors of H is $\{|v_1\rangle, |v_2\rangle, \ldots, |v_D\rangle\}$. Suppose $|v_1\rangle$ is a vector with a nondegenerate eigenvalue λ_1,

$$H|v_1\rangle = \lambda_1|v_1\rangle, \tag{E8}$$

$$H^\dagger|T^{-1}v_1\rangle = \lambda_1^*|T^{-1}v_1\rangle, \tag{E9}$$

where the second equality follows from $T^{-1}HT = H^\dagger$. We still have $\langle v_1|Tv_1\rangle = 0$ for the same reason as in the Hermitian case. For any other eigenvector $|v_i\rangle$ $(i \neq 1)$, we have that

$$H|v_i\rangle = \lambda_i|v_i\rangle \Rightarrow \langle v_i|H^\dagger = \langle v_i|\lambda_i^*. \tag{E10}$$

Combining the above equations, we can deduce that for $i \neq 1$

$$\lambda_i^*\langle v_i|T^{-1}v_1\rangle = \langle v_i|H^\dagger T^{-1}v_1\rangle = \lambda_1^*\langle v_i|T^{-1}v_1\rangle, \tag{E11}$$

and since $\lambda_1 \neq \lambda_i$ by assumption, we conclude $\langle Tv_i|v_1\rangle = \langle v_i|T^{-1}v_1\rangle^* = 0$.

We have now proven that if $|v_1\rangle$ has a nondegenerate eigenvalue then $|v_1\rangle$ is orthogonal to all $T|v_i\rangle$ $(i = 1, 2, \ldots, D)$; namely, T brings the full Hilbert space into the orthogonal complement of $|v_1\rangle$, and this contradicts the fact that T is an invertible operator. Hence, by contradiction, we have proven that every eigenvalue is at least twofold degenerate. Given that there is no symmetry mechanism to enforce an even higher degeneracy, we will see twofold degeneracy for a generic realization of the ensemble.

It is worth noting that the above proof applies to each chiral block Hamiltonian because both H and T commute with γ_c so in the chiral basis they are simultaneously block-diagonal. So, degenerate eigenstates always have the same chirality, and $\langle 0|\psi_L\psi_R|0\rangle$ in the corresponding two-site model (without the $i\epsilon$ symmetry breaking term) must vanish, regardless of which linear combination one takes.

APPENDIX F: RANDOM EIGENVALUES IN A DISK

In this Appendix, we discuss the partition function of uniform uncorrelated random eigenvalues in a disk with level density given by

$$\rho(z) = \frac{D}{\pi R^2}\theta(R - |z|). \tag{F1}$$

The radius R will be adjusted in order to get a stable large D limit.

The one-site annealed partition function is the same as the one for the Ginibre model:

$$Z_1(\beta) = \int \rho(z)e^{-\beta z}d^2z = D. \tag{F2}$$

The two-site partition function corresponding to one replica and one conjugate replica is given by

$$\begin{aligned} Z_2(\beta) &= \left\langle \sum_{kl} e^{-\beta(E_k + E_l^*)} \right\rangle \\ &= \left\langle \sum_{k\neq l} e^{-\beta(E_k + E_l^*)} \right\rangle + \left\langle \sum_{k} e^{-\beta(E_k + E_k^*)} \right\rangle. \end{aligned} \tag{F3}$$

Since different eigenvalues are uncorrelated, and all eigenvalues have the same distribution, we obtain in the large D limit

$$Z_2(\beta) = \frac{D(D-1)}{\pi^2 R^4}\int_{\mathcal{D}} d^2z e^{-\beta z}\int_{\mathcal{D}} d^2z e^{-\beta z^*} + \frac{D}{\pi R^2}\int_{\mathcal{D}} d^2z e^{-\beta(z+z^*)} = D(D-1) + \frac{D}{R\beta} I_1(2R\beta). \tag{F4}$$

The large D limit of the free energy density is given by

$$\frac{F(T)}{\log D} = -\lim_{D\to\infty}\frac{T}{\log D}\log(Z_1^2(\beta) + Z_2(\beta)) = -\lim_{D\to\infty}\frac{T}{\log D}\log\left(e^{2\log D} + \frac{D}{R\beta} I_1(2R\beta)\right). \tag{F5}$$

A nontrivial large D limit is obtained if we scale R as $\log D$, and we will choose $R = \log D$. At the critical temperature T_c, the two leading exponents are equal so that

$$2\log D = \log D + 2\beta_c \log D, \tag{F6}$$

resulting in

$$T_c = 2, \tag{F7}$$

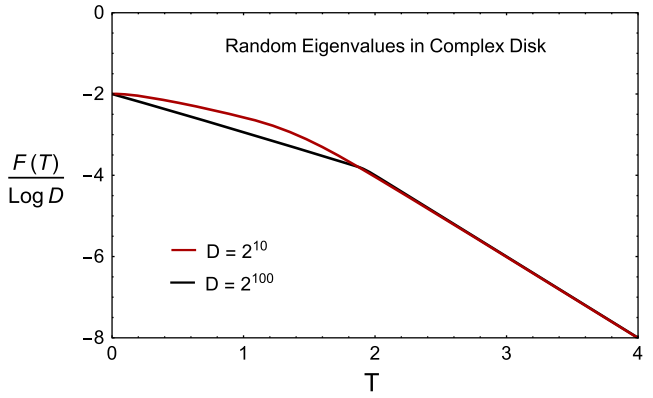

FIG. 18. Annealed two-site free energy of uniformly distributed random eigenvalues in a complex disk of radius $\log D$. Results are plotted for $D = 2^{10}$ (red) and 2^{100} (black).

and the free energy is given by

$$\frac{F(T)}{\log D} = (-2-T)\theta(T_c - T) - 2T\theta(T - T_c). \tag{F8}$$

In Fig. 18, we show the free energy (F5) versus the temperature. The convergence to the asymptotic result (F8) is slow, but for $D = 2^{100}$ (black curve), a kink becomes visible.

[1] S. F. Edwards and P. W. Anderson, Theory of spin glasses, J. Phys. F **5**, 965 (1975).
[2] G. Parisi, Toward a mean field theory for spin glasses, Phys. Lett. **73A**, 203 (1979).
[3] G. Parisi, Order Parameter for Spin-Glasses, Phys. Rev. Lett. **50**, 1946 (1983).
[4] F. Wegner, The mobility edge problem: Continuous symmetry and a conjecture, Z. Phys. B Condens. Matter **35**, 207 (1979).
[5] A. Kamenev and M. Mezard, Level correlations in disordered metals: The replica sigma model, Phys. Rev. B **60**, 3944 (1999).
[6] M. A. Stephanov, Random Matrix Model of QCD at Finite Density and the Nature of the Quenched Limit, Phys. Rev. Lett. **76**, 4472 (1996).
[7] G. Akemann, J. C. Osborn, K. Splittorff, and J. J. M. Verbaarschot, Unquenched QCD Dirac operator spectra at nonzero baryon chemical potential, Nucl. Phys. **B712**, 287 (2005).
[8] H. Nishimori and K. Y. M. Wong, Statistical mechanics of image restoration and error-correcting codes, Phys. Rev. E **60**, 132 (1999).
[9] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro, Nature of the Spin-Glass Phase, Phys. Rev. Lett. **52**, 1156 (1984).
[10] M. Mézard, G. Parisi, and M. A. Virasoro, Random free energies in spin glasses, J. Phys. Lett. **46**, 217 (1985).
[11] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin-Glass, Phys. Rev. Lett. **35**, 1792 (1975).
[12] E. Kanzieper, Replica Field Theories, Painleve Transcendents and Exact Correlation Functions, Phys. Rev. Lett. **89**, 250201 (2002).
[13] S. M. Nishigaki and A. Kamenev, Replica treatment of non-hermitian disordered hamiltonians, J. Phys. A **35**, 4571 (2002).
[14] J. B. French and S. S. M. Wong, Validity of random matrix theories for many-particle systems, Phys. Lett. **33B**, 449 (1970).
[15] O. Bohigas and J. Flores, Two-body random hamiltonian and level density, Phys. Lett. **34B**, 261 (1971).
[16] A. Kitaev, A simple model of quantum holography, KITP strings seminar and Entanglement 2015 program, 12 February, 7 April and 27 May 2015, http://online.kitp.ucsb.edu/online/entangled15/ (2015).

[17] S. Sachdev and J. Ye, Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet, Phys. Rev. Lett. **70**, 3339 (1993).
[18] J. Maldacena and D. Stanford, Remarks on the Sachdev-Ye-Kitaev model, Phys. Rev. D **94**, 106002 (2016).
[19] K. Jensen, Chaos in AdS_2 Holography, Phys. Rev. Lett. **117**, 111601 (2016).
[20] J. B. French and S. S. M. Wong, Some random-matrix level and spacing distributions for fixed-particle-rank interactions, Phys. Lett. **35B**, 5 (1971).
[21] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Random-matrix physics: Spectrum and strength fluctuations, Rev. Mod. Phys. **53**, 385 (1981).
[22] J. J. M. Verbaarschot, H. A. Weidenmuller, and M. R. Zirnbauer, Grassmann integration in stochastic quantum physics: The case of compound nucleus scattering, Phys. Rep. **129**, 367 (1985).
[23] J. Flores, M. Horoi, M. Muller, and T. H. Seligman, Spectral statistics of the two-body random ensemble revisited, Phys. Rev. E **63**, 026204 (2001).
[24] L. Benet and H. A. Weidenmuller, Review of the k body embedded ensembles of Gaussian random matrices, J. Phys. A **36**, 3569 (2003).
[25] V. Zelevinsky and A. Volya, Nuclear structure, random interactions and mesoscopic physics, Phys. Rep. **391**, 311 (2004).
[26] V. K. B. Kota and M. Vyas, Statistical nuclear spectroscopy with q-normal and bivariate q-normal distributions and q-Hermite polynomials, arXiv:2201.12509.
[27] C. Teitelboim, Gravitation and Hamiltonian structure in two spacetime dimensions, Phys. Lett. **126B**, 41 (1983).
[28] R. Jackiw, Lower dimensional gravity, Nucl. Phys. **B252**, 343 (1985).
[29] J. Maldacena, D. Stanford, and Z. Yang, Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space, Prog. Theor. Exp. Phys. **2016**, 12C104 (2016).
[30] J. Engelsöy, T. G. Mertens, and H. Verlinde, An investigation of ads2 backreaction and holography, J. High Energy Phys. 07 (2016) 139.
[31] D. Stanford and E. Witten, Fermionic localization of the Schwarzian theory, J. High Energy Phys. 10 (2017) 008.
[32] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, Black holes and random matrices, J. High Energy Phys. 05 (2017) 118.
[33] A. M. García-García and J. J. M. Verbaarschot, Analytical spectral density of the sachdev-ye-kitaev model at finite n, Phys. Rev. D **96**, 066012 (2017).
[34] A. M. García-García and J. J. M. Verbaarschot, Spectral and thermodynamic properties of the sachdev-ye-kitaev model, Phys. Rev. D **94**, 126010 (2016).
[35] Y.-Z. You, A. W. W. Ludwig, and C. Xu, Sachdev-ye-kitaev model and thermalization on the boundary of many-body localized fermionic symmetry-protected topological states, Phys. Rev. B **95**, 115150 (2017).
[36] A. M. García-García, Y. Jia, and J. J. M. Verbaarschot, Universality and thouless energy in the supersymmetric sachdev-ye-kitaev model, Phys. Rev. D **97**, 106003 (2018).
[37] A. M. García-García, T. Nosaka, D. Rosa, and J. J. M. Verbaarschot, Quantum chaos transition in a two-site Sachdev-Ye-Kitaev model dual to an eternal traversable wormhole, Phys. Rev. D **100**, 026002 (2019).
[38] M. Fremling, M. Haque, and L. Fritz, A bipartite Sachdev-Ye-Kitaev model: Conformal limit and level statistics, Phys. Rev. D **105**, 066017 (2022).
[39] Z. Cao, Z. Xu, and A. del Campo, Probing quantum chaos in multipartite systems, arXiv:2111.12475.
[40] E. Caceres, A. Misobuchi, and R. Pimentel, Sparse SYK and traversable wormholes, J. High Energy Phys. 11 (2021) 015.
[41] J. Maldacena and X.-L. Qi, Eternal traversable wormhole, arXiv:1804.00491.
[42] P. Gao, D. L. Jafferis, and A. Wall, Traversable wormholes via a double trace deformation, J. High Energy Phys. 12 (2017) 151.
[43] P. Saad, S. H. Shenker, and D. Stanford, JT gravity as a matrix integral, arXiv:1903.11115.
[44] J. J. Heckman, A. P. Turner, and X. Yu, Disorder averaging and its UV discontents, Phys. Rev. D **105**, 086021 (2022).
[45] N. Engelhardt, S. Fischetti, and A. Maloney, Free energy from replica wormholes, Phys. Rev. D **103**, 046021 (2021).
[46] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, Replica wormholes and the entropy of Hawking radiation, J. High Energy Phys. 05 (2020) 013.
[47] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, The entropy of Hawking radiation, Rev. Mod. Phys. **93**, 035002 (2021).
[48] G. Penington, Entanglement wedge reconstruction and the information paradox, J. High Energy Phys. 09 (2020) 002.
[49] D. N. Page, Information in Black Hole Radiation, Phys. Rev. Lett. **71**, 3743 (1993).
[50] J. Maldacena and L. Maoz, Wormholes in AdS, J. High Energy Phys. 02 (2004) 053.
[51] P. Saad, S. Shenker, and S. Yao, Comments on wormholes and factorization, arXiv:2107.13130.
[52] A. Belin, J. de Boer, P. Nayak, and J. Sonner, Generalized spectral form factors and the statistics of heavy operators, arXiv:2111.06373.
[53] C. V. Johnson, The microstate physics of JT gravity and supergravity, arXiv:2201.11942.
[54] M. Berkooz, N. Brukner, S. F. Ross, and M. Watanabe, Going beyond ER = EPR in the SYK model, arXiv:2202.11381.
[55] J.-M. Schlenker and E. Witten, No ensemble averaging below the black hole threshold, arXiv:2202.01372.
[56] K. Goto, K. Suzuki, and T. Ugajin, Factorizing wormholes in a partially disorder-averaged SYK model, arXiv:2111.11705.
[57] P. Gao and L. Lamprou, Seeing behind black hole horizons in SYK, arXiv:2111.14010.
[58] A. M. García-García and V. Godet, Euclidean wormhole in the Sachdev-Ye-Kitaev model, Phys. Rev. D **103**, 046014 (2021).
[59] A. M. García-García, Y. Jia, D. Rosa, and J. J. M. Verbaarschot, Dominance of Replica Off-Diagonal Configurations and Phase Transitions in a PT Symmetric Sachdev-Ye-Kitaev Model, Phys. Rev. Lett. **128**, 081601 (2022).
[60] C. M. Bender and S. Boettcher, Real Spectra in nonHermitian Hamiltonians Having PT Symmetry, Phys. Rev. Lett. **80**, 5243 (1998).

[61] I. Aref'eva, M. Khramtsov, M. Tikhanovskaya, and I. Volovich, Replica-nondiagonal solutions in the SYK model, J. High Energy Phys. 07 (2019) 113.
[62] H. Wang, D. Bagrets, A. L. Chudnovskiy, and A. Kamenev, On the replica structure of Sachdev-Ye-Kitaev model, J. High Energy Phys. 09 (2019) 057.
[63] J. J. M. Verbaarschot and M. R. Zirnbauer, Critique of the replica trick, J. Phys. A **18**, 1093 (1985).
[64] M. R. Zirnbauer, Another critique of the replica trick, arXiv: cond-mat/9903338.
[65] K. Splittorff and J. J. M. Verbaarschot, Factorization of correlation functions and the replica limit of the Toda lattice equation, Nucl. Phys. **B683**, 467 (2004).
[66] T. A. Sedrakyan, Toda lattice representation for random matrix model with logarithmic confinement, Nucl. Phys. **B729**, 526 (2005).
[67] K. Efetov, Supersymmetry and theory of disordered metals, Adv. Phys. **32**, 53 (1983).
[68] F. Wegner, Exact density of states for lowest Landau level in white noise potential superfield representation for interacting systems, Z. Phys. B Condens. Matter **51**, 279 (1983).
[69] J. Verbaarschot and M. Zirnbauer, Replica variables, loop expansion, and spectral rigidity of random-matrix ensembles, Ann. Phys. (N.Y.) **158**, 78 (1984).
[70] J. J. M. Verbaarschot, H. A. Weidenmüller, and M. R. Zirnbauer, Grassmann integration and the theory of compound-nucleus reactions, Phys. Lett. **149B**, 263 (1984).
[71] T. A. Sedrakyan and K. B. Efetov, Supersymmetry method for interacting chaotic and disordered systems: The Sachdev-Ye-Kitaev model, Phys. Rev. B **102**, 075146 (2020).
[72] R. Carlson, Sur une classe de séries de Taylor, Dissertation, Uppsala, Sweden (1914).
[73] V. L. Girko, *Theory of Random Determinants* (Kluwer Academic Publishers, Dordrecht, Netherlands, 2012), Vol. 45.
[74] S. M. Nishigaki and A. Kamenev, Replica treatment of non-Hermitian disordered Hamiltonians, J. Phys. A **35**, 4571 (2002).
[75] J. Feinberg and A. Zee, NonGaussian nonHermitian random matrix theory: Phase transition and addition formalism, Nucl. Phys. **B501**, 643 (1997).
[76] R. A. Janik, M. A. Nowak, G. Papp, and I. Zahed, Non-Hermitian random matrix models. 1., Nucl. Phys. **B501**, 603 (1997).
[77] J. Ginibre, Statistical ensembles of complex, quaternion, and real matrices, J. Math. Phys. (N.Y.) **6**, 440 (1965).
[78] Y. V. Fyodorov, B. A. Khoruzhenko, and H.-J. Sommers, Almost Hermitian Random Matrices: Crossover from Wigner-Dyson to Ginibre Eigenvalue Statistics, Phys. Rev. Lett. **79**, 557 (1997).
[79] Y. V. Fyodorov and H.-J. Sommers, Random matrices close to hermitian or unitary: Overview of methods and results, J. Phys. A **36**, 3303 (2003).
[80] Y. Can, A. M. García-García, V. Godet, and J. P. Zheng, Euclidean-to-Lorentzian wormhole transition and gravitational symmetry breaking in the Sachdev-Ye-Kitaev model, arXiv:2204.08558.
[81] L. Erdős and D. Schröder, Phase transition in the density of states of quantum spin glasses, Math. Phys. Anal. Geom. **17**, 441 (2014).
[82] D. Bagrets, A. Altland, and A. Kamenev, Sachdev–Ye–Kitaev model as Liouville quantum mechanics, Nucl. Phys. **B911**, 191 (2016).
[83] D. Bagrets, A. Altland, and A. Kamenev, Power-law out of time order correlation functions in the SYK model, Nucl. Phys. **B921**, 727 (2017).
[84] Y. Jia, D. Rosa, and J. J. M. Verbaarschot, Replica symmetry breaking for the integrable two-site Sachdev-Ye-Kitaev model, arXiv:2201.05952.
[85] L. F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam, Phil. Trans. R. Soc. A **210 (459–470)**, 307 (1911).
[86] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, Replica wormholes and the entropy of hawking radiation, J. High Energy Phys. 05 (2020) 013.
[87] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, Replica wormholes and the black hole interior, J. High Energy Phys. 03 (2020) 205.
[88] M. L. Mehta, *Random Matrices* (Academic Press, New York, 2004).