

## Variational wave functions for Sachdev-Ye-Kitaev models

Arijit Haldar<sup>1</sup>,\* Omid Tavakol<sup>1</sup>, and Thomas Scaffidi

*Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7*



(Received 18 September 2020; accepted 9 February 2021; published 7 April 2021)

Given a class of  $q$ -local Hamiltonians, is it possible to find a simple variational state whose energy is a finite fraction of the ground-state energy in the thermodynamic limit? Whereas product states often provide an affirmative answer in the case of bosonic (or qubit) models, we show that Gaussian states fail dramatically in the fermionic case, such as for the Sachdev-Ye-Kitaev (SYK) models. This prompts us to propose a new class of wave functions for SYK models inspired by the variational coupled cluster algorithm. We introduce a static “(0+0)-dimensional” large- $N$  field theory to study the energy, two-point correlators, and entanglement properties of these states. Most importantly, we demonstrate a finite disorder-averaged approximation ratio of  $r \approx 0.62$  between the variational and ground-state energy of the SYK model for  $q = 4$ . Moreover, the variational states provide an exact description of spontaneous symmetry breaking in a related two-flavor SYK model.

DOI: [10.1103/PhysRevResearch.3.023020](https://doi.org/10.1103/PhysRevResearch.3.023020)

### I. INTRODUCTION

Variational wave functions are at the heart of our understanding of a variety of condensed matter systems such as quantum Hall systems [1], superconductors [2], and correlated metals [3]. These wave functions provide an intuitive description of these phases and are often useful for numerics. Working with pure states also makes it possible to study entanglement, a property which has been crucial to characterize exotic phases of matter [4]. Furthermore, with the advent of quantum simulators [5,6], and, in particular, of hybrid quantum-classical variational algorithms [7–9], it is desirable to find preparable states that can reach low-energy regimes of strongly correlated Hamiltonians.

A related topic of recent interest is Hamiltonian complexity [10], which studies the computational complexity of approximating the ground state of certain classes of Hamiltonians. These problems belong to the quantum Merlin-Arthur (QMA) class since a verifier can check a solution (i.e., a quantum state) efficiently on a quantum computer by measuring its energy [11–13]. Whereas approximating the ground-state energy within a small additive error was shown to be QMA complete for a wide range of Hamiltonians, the complexity of approximating the ground-state energy density within finite relative error is still undecided and is closely related to the quantum probabilistically checkable proof (PCP) [14–18] and no low-energy trivial states (NLTS) conjectures [19]. Proving these conjectures would, roughly speaking, require finding classes of Hamiltonians for which not only the ground state

but also all states below a finite energy density are impossible to reach with a simple ansatz.

Given a class of traceless Hamiltonians and a class of ansatz wave functions, one can define a figure of merit called the approximation ratio, given by  $r_\psi \equiv E_\psi/E_{\text{GS}}$ , where  $E_{\text{GS}}$  is the energy of the ground state and  $E_\psi = \min_\psi \langle \psi | H | \psi \rangle$ , where  $\psi$  belongs to the class of ansatz wave functions. For nontrivial Hamiltonians, simple wave functions (e.g., product states) are of course not expected to reach an approximation ratio very close to 1. The question we aim to answer instead is whether they can at least achieve  $r_\psi > 0$  in the thermodynamic limit. Remarkably, the answer can be shown to be affirmative for a variety of bosonic (or qubit) models [14,16,20–23]. For example, for traceless two-local qubit Hamiltonians of the type

$$H = \sum_{i,j=1}^N \sum_{\mu,\nu=1}^3 J_{i,j}^{\mu,\nu} \sigma_i^\mu \sigma_j^\nu, \quad (1)$$

where  $\sigma_j^\nu$  are Pauli matrices, Lieb showed that the approximation ratio of product states has a lower bound:  $r_{\text{prod}} \geq 1/9$  [20,22].

Our work is motivated by the following question: Can similar results be obtained for  $q$ -local *fermionic* Hamiltonians [22]? For fermionic systems, Gaussian states are a natural analog of product states and include the Slater determinants calculated with the Hartree-Fock method. However, for  $q > 2$ , we will provide strong evidence that the approximation ratio of Gaussian states goes to 0 in the thermodynamic limit:  $r_{\text{Gauss}} \rightarrow 0$  for  $N \rightarrow \infty$ . This highlights a fundamental difference between the bosonic and fermionic case. It also motivates the following question: If Gaussian states are not up to the task, is there any other class of tractable wave functions that could provide a finite approximation ratio?

Rather than trying to make statements about all problems in the class, we study instances of  $q$ -local fermionic Hamiltonians that are typical for a natural measure, which

\*arijit.haldar@utoronto.ca

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

enables us to establish relations with the statistical mechanics of disordered quantum systems. Namely, we will focus on a paradigmatic example of disordered fermionic systems, the Sachdev-Ye-Kitaev (SYK) models [24–26]. The model has become a primary platform for studying non-Fermi-liquid regimes [27–37], quantum many-body chaos and operator complexity [26,29,38,39], thermalization [40–42], and dualities between quantum-field theory and gravity [25,43–47]. Whereas a lot is known already about thermal ensembles in SYK models, less is known about wave functions of typical low-energy states. In fact, existing work on pure states in SYK models has relied on thermal states in disguise, such as thermofield double states [48,49] and Kourkoulou-Maldacena (KM) states [50,51], and thus requires computation in a thermal-field theory. We will propose instead a class of variational wave functions for which equal time observables can be computed within a static (0+0)-dimensional (“0+0D”) field theory.

This paper is organized as follows. In Sec. II, we formally define the  $q$ -SYK model and construct a variational ansatz for the model. In Sec. III, we show that the energy and particle density of the ansatz can be evaluated exactly in the large- $N$  limit. In the same section, we compare the analytical predictions for the ansatz with those obtained using the thermal-field theory of the SYK model. We discuss the nature of entanglement for the ansatz in Sec. IV by computing the scaling of the second Rényi entropy with subsystem size. Finally, we provide a discussion of our findings in Sec. V. Additional details about various results are provided in the Appendixes and referred to in the main text.

## II. MODEL AND ANSATZ

The  $q$ -SYK model is defined as

$$H_{\text{SYK}} = g \sum_{\substack{1 \leq i_1 < \dots < i_{(q/2)} \leq N, \\ 1 \leq j_1 < \dots < j_{(q/2)} \leq N}} \times J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}} \hat{c}_{i_1}^\dagger \dots \hat{c}_{i_{(q/2)}}^\dagger \hat{c}_{j_1} \dots \hat{c}_{j_{(q/2)}}, \quad (2)$$

with  $i, j \in [1, N]$  and with  $g = (q/2)! / \sqrt{(q/2)} \binom{N}{2}^{\frac{q}{2} - \frac{1}{2}}$ . The symbols  $\hat{c}_i^\dagger$  and  $\hat{c}_i$  denote fermionic creation and annihilation operators, respectively. The couplings  $J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}}$  are Gaussian random numbers which satisfy appropriate symmetrization conditions [52]. The variance is represented as  $J^2$  and will be set to 1 except when written explicitly. This Hamiltonian has an extensive energy bandwidth which is symmetric around zero due to particle-hole symmetry [53].

The simplest variational wave functions for a fermionic model are Gaussian states (which include Slater determinants), and the corresponding optimization procedure is the celebrated Hartree-Fock method [54]. In quantum chemistry, this technique typically recovers 99% of the electronic energy and is the basis for a variety of more sophisticated approaches. By contrast, for SYK models with  $q > 2$ , an elementary calculation (see Appendix A) shows that the energy bandwidth of Gaussian states (which is centered around zero) scales subextensively with  $N$ . In the large- $N$  limit, Gaussian states therefore only reach a vanishing fraction of the full many-body bandwidth of the SYK model, and their

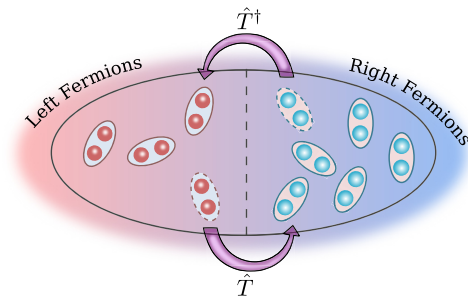


FIG. 1. Constructing the wave function: The  $N$  orbitals are partitioned into left and right subsystems. The operator  $\hat{T}^\dagger$  moves  $q/2$  fermions (with  $q = 4$  in the figure) at a time from the right side to the left side, with the same amplitude  $J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}}$  as the corresponding term in the Hamiltonian. Starting from a state  $|\tilde{0}\rangle$  in which the right side is filled with fermions and the left side is empty, the variational wave function is constructed by repeated applications of  $\hat{T}^\dagger$ .

disorder-averaged approximation ratio goes to zero. This is a strong indication that the “worst-case” approximation ratio of Gaussian states for  $q$ -local fermionic Hamiltonians with  $q > 2$  goes to 0 in the large- $N$  limit, in contradiction to the conjecture found in Ref. [22]. Intuitively, this happens since minimizing the energy requires optimizing over the value of  $q$ -point correlators, but these correlators are overconstrained for a Gaussian state: Due to Wick’s theorem, all higher-order correlators are simple functions of two-point correlators.

Since the Hartree-Fock method does not produce any useful result, we take a different approach: Let us look for a subset of terms in  $H$  which commute with each other and for which the energy can be minimized easily. The selected subset of terms should be extensive in order for the state to have a finite approximation ratio; that is, it should contain a number of terms which scales as  $N^q$ . We propose to construct such a set by partitioning the system into two subsystems (see Fig. 1), with  $N_L$  sites on the left and  $N_R = N - N_L$  sites on the right, and by keeping only terms with creation operators on the left side and annihilation operators on the right side:

$$\hat{T}^\dagger = g \sum_{\substack{i_1 < \dots < i_{(q/2)} \in L, \\ j_1 < \dots < j_{(q/2)} \in R}} \times J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}} \hat{c}_{i_1}^\dagger \dots \hat{c}_{i_{(q/2)}}^\dagger \hat{c}_{j_1} \dots \hat{c}_{j_{(q/2)}}, \quad (3)$$

where  $L = [1, \dots, N_L]$  and  $R = [N_L + 1, \dots, N]$ . The parameter  $p = N_R/N_L$  can be tuned at will, but we will focus on  $p = 1$  for now. It will be useful to define the partitioned-SYK Hamiltonian,

$$H_{\text{pSYK}} = \hat{T} + \hat{T}^\dagger, \quad (4)$$

which contains an extensive subset of the terms of  $H_{\text{SYK}}$  and which is an example of the systems studied in Ref. [36].

Using this notation, the ansatz wave function is defined as

$$|\psi(a)\rangle = \frac{1}{\sqrt{\mathcal{N}}} \exp(-a\hat{T}^\dagger)|\tilde{0}\rangle, \quad (5)$$

where  $|\tilde{0}\rangle$  is the state for which all states on the right (left) are full (empty),  $a$  is a real variational parameter, and the nor-

malization is given by  $\mathcal{N}(a) = \langle \tilde{0} | \exp(-a\hat{T}) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle$ . The intuition behind this state is as follows: Starting from a state that is empty on the left and fully occupied on the right, we create a population of particles on the left and holes on the right by applying the corresponding terms from the Hamiltonian.

Interestingly, this wave function belongs to the class of variational coupled cluster (VCC) states developed for quantum chemistry [55–58]. This algorithm has the advantage of being variational (as opposed to regular coupled cluster [59,60]) but is usually limited to a very small number of orbitals due to the factorial complexity of the method. By contrast, we were able to perform VCC directly in the large- $N$  limit for a class of SYK models.

The disorder-averaged energy density for the state is given by  $E(a) = \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle}$  and can be calculated (see Appendix B for details) using

$$E(a) = \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle} = -\frac{1}{N} \frac{\partial \overline{\ln(\mathcal{N})}}{\partial a}, \quad (6)$$

where we used the fact that the expectation value of the terms which are present in  $H_{\text{SYK}}$  but not in  $H_{\text{pSYK}}$  vanishes after disorder averaging.

### III. LARGE- $N$ THEORY

To enable the computation of  $\overline{\ln(\mathcal{N})}$ , we introduce a field-theoretic approach similar to the fermionic path integral (see Appendix B for details). First, we perform a particle-hole transformation on the right side, whereby  $\hat{c}_{i \in R} = \hat{h}_{i \in R}^\dagger$  and  $\hat{c}_{i \in L}^\dagger = \hat{h}_{i \in L}$ . We then define the fermionic-coherent states  $|c_{i \in L}\rangle, |h_{i \in R}\rangle$  for left and right, characterized by the Grassmann numbers  $c_i, \bar{c}_i$  and  $h_i, \bar{h}_i$ , respectively, such that  $\langle c_i | \hat{c}_i^\dagger = \langle c_i | \bar{c}_i$ ,  $\langle h_i | \hat{h}_i^\dagger = \langle h_i | \bar{h}_i$ . The disorder averaging is implemented using the replica trick  $\overline{\ln(\mathcal{N})} = \lim_{\mathcal{R} \rightarrow 0} [\overline{\mathcal{N}^{\mathcal{R}}} - 1]/\mathcal{R}$ . This results in a “static” action involving Grassmann numbers  $c_i, h_i$  with no imaginary-time dynamics. Introducing the static Green’s functions

$$G_c = -N_L^{-1} \sum_{i \in L} \langle c_i \bar{c}_i \rangle, \quad G_h = -N_R^{-1} \sum_{i \in R} \langle h_i \bar{h}_i \rangle, \quad (7)$$

along with the self-energies  $\Sigma_c, \Sigma_h$ , into the action allows us to integrate the fermions  $c_i, h_i$ . The particle densities in the left and right subsystems are simply given by  $\rho_L = N_L^{-1} \sum_{i \in L} \langle \hat{c}_i^\dagger \hat{c}_i \rangle = 1 + G_c$  and  $\rho_R = N_R^{-1} \sum_{i \in R} \langle \hat{h}_i \hat{h}_i^\dagger \rangle = -G_h$ , respectively. For  $p = 1$ , particle conservation implies  $\rho_L + \rho_R = 1$ , and thus  $G_c = G_h \equiv G$  and  $\Sigma_c = \Sigma_h = \Sigma$ . At the saddle point, one finds

$$\begin{aligned} -G^{-1} &= 1 + \Sigma, \\ \Sigma &= -a^2 J^2 G^{q-1}, \end{aligned} \quad (8)$$

which are polynomial equations for  $G(a)$  and  $\Sigma(a)$  that can easily be solved numerically. These relations derive from the generating function  $\overline{\ln(\mathcal{N})}$ , which takes the form

$$-\overline{\ln(\mathcal{N})} = -N \left[ \ln(1 + \Sigma) + \Sigma G + \frac{a^2 J^2}{q} G^q \right], \quad (9)$$

at the large- $N$  saddle point. Interestingly, this generating functional can be interpreted as a static limit of the free energy for

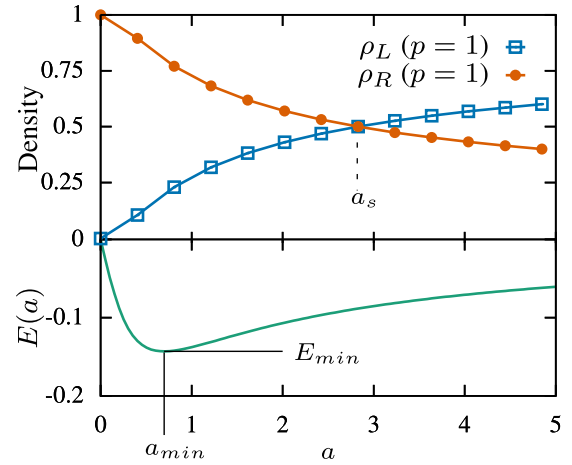


FIG. 2. Top: Particle densities on the left and right sides ( $\rho_L$  and  $\rho_R$ , respectively). The densities are equal at  $a_s$ . Bottom: Variational energy, with a minimum at  $a_{\min}$ .

the SYK model [26,29], given by

$$\begin{aligned} F_{\text{SYK}} &= -N \left[ T \ln \det(\partial_\tau + \Sigma) + \int d\tau \Sigma(\tau) G(\beta - \tau) \right. \\ &\quad \left. + (J^2/q) \int d\tau G(\tau)^{(q/2)} G(\beta - \tau)^{(q/2)} \right], \end{aligned} \quad (10)$$

where  $\tau \in [0, \beta]$  denotes the imaginary-time variable and  $\beta = T^{-1}$  is the inverse temperature. Indeed, if the imaginary-time dynamics is eliminated by substituting  $\partial_\tau \rightarrow 1$ ,  $G(\tau) \rightarrow G$ , and  $\Sigma(\tau) \rightarrow \Sigma$ , an expression similar to  $-\overline{\ln(\mathcal{N})}$  in Eq. (9) is recovered.

The energy density  $E(a)$  is calculated using Eq. (6) to give

$$E(a) = -\frac{2}{q} a J^2 G^q, \quad (11)$$

where  $G$  is obtained by solving the saddle-point equations (see Fig. 2). The most important point is that  $E(a)$  does not decay with  $N$ , which means that the variational states have an extensive bandwidth and thus a finite approximation ratio in the large- $N$  limit.

The variational energy has a single minimum as a function of  $a$  (see Fig. 2, bottom), with the following properties:

$$a_{\min} = \frac{1}{J} \frac{(q+1)^{\frac{q-1}{2}}}{q^{\frac{q}{2}}}, \quad (12)$$

$$E_{\min} = -J \frac{2}{q} \frac{q^{\frac{q}{2}}}{(q+1)^{\frac{q+1}{2}}}. \quad (13)$$

We can now compare  $E_{\min}$  with the energy density of the ground state of the SYK model ( $E_{\text{SYK}}$ ). The latter can be obtained by taking the zero-temperature limit of  $F_{\text{SYK}}$  [see Eq. (10)]. We give a comparison as a function of  $q$  in Fig. 3(a). For example, for  $q = 4$ , we find  $E_{\min} = 8/25\sqrt{5} \approx -0.143$  and  $E_{\text{SYK}} \approx -0.2295$ . Since we expect both  $E_{\min}$  and  $E_{\text{SYK}}$  to be self-averaging, we define the disorder-averaged approximation ratio as  $r_\psi = E_{\min}/E_{\text{SYK}}$ . We thus find  $r_\psi \approx 0.62$  for  $q = 4$ . To put things into perspective, we have calculated that  $E_{\min}$  has the same energy density as the thermal ensemble of the SYK model at temperature  $T/J \approx 0.455$ .

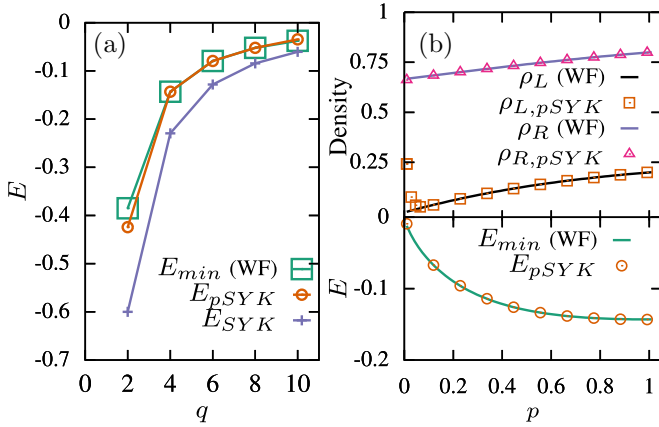


FIG. 3. (a) Comparison between variational energy  $E_{\min}$  and exact ground-state energies of  $H_{\text{SYK}}$  and  $H_{p\text{SYK}}$ , for  $p = 1$  and varying  $q$ . We find  $E_{\min}$  and  $E_{p\text{SYK}}$  to be equal within our numerical accuracy for  $q \geq 4$ . (b) Comparison between variational wave function (WF; lines) and exact ground state of  $H_{p\text{SYK}}$  (symbols), when  $q = 4$ , for the energy and the particle densities in the left and right subsystems.

A peculiarity of  $|\psi(a)\rangle$  is that the particle densities on the left and right depend on  $a$  and are only equal for  $a = a_s = 2^{(q/2)-1/2}$  (see Fig. 2) [61]. Since  $a_s \neq a_{\min}$ , the variational state with the lowest energy has an asymmetric particle density between left and right, in contrast to the ground state of the original SYK model, for which all orbitals are at half filling. This discrepancy arises from the fact that our construction aims at minimizing  $H_{p\text{SYK}}$ , which contains only a subset of the terms in  $H_{\text{SYK}}$ , and creates an artificial distinction between the two subsystems. The Hamiltonian  $H_{p\text{SYK}}$  is actually interesting in its own right as it can be understood as an example of two-flavor SYK models, in which two SYK quantum dots are coupled by  $q$ -body interactions, as studied in Ref. [36]. For  $q \geq 4$ ,  $H_{p\text{SYK}}$  was shown to have a low-temperature phase which exhibits phase separation: One subsystem (say, the one on the left) has density  $1/(q+1)$ , and the other one has density  $q/(q+1)$ . This phase has a gap to single-particle excitations and spontaneously breaks particle-hole symmetry and left-right interchange symmetry, but it conserves their product.

Interestingly,  $|\psi(a_{\min})\rangle$  reproduces this density imbalance perfectly: We find  $\rho_{L,\min} = 1 - \rho_{R,\min} = 1/(q+1)$ . Furthermore, we find  $E_{\min}$  to be equal to the ground-state energy of  $H_{p\text{SYK}}$  (which can be obtained in a similar fashion as  $E_{\text{SYK}}$ ; see Appendix C) within numerical accuracy for  $q \geq 4$ . We checked that this agreement even extends to the asymmetric case of  $p = N_R/N_L \neq 1$ . In the context of two-flavor SYK models, this ratio gives the relative size of the two dots [36]. The comparison for  $\rho_L$ ,  $\rho_R$ , and  $E_{\min}$  with the exact values is shown in Fig. 3(b), for  $q = 4$ . The only discrepancy appears as  $p \rightarrow 0$ , which is expected since  $H_{p\text{SYK}}$  undergoes an additional phase transition to a gapless phase at  $p_c \approx 0.072$  [36]. Another discrepancy appears for  $q = 2$ , in which case the variational wave function fails to describe the Fermi liquid phase of  $H_{p\text{SYK}}$  which survives down to  $T = 0$  (see Appendix D for more details).

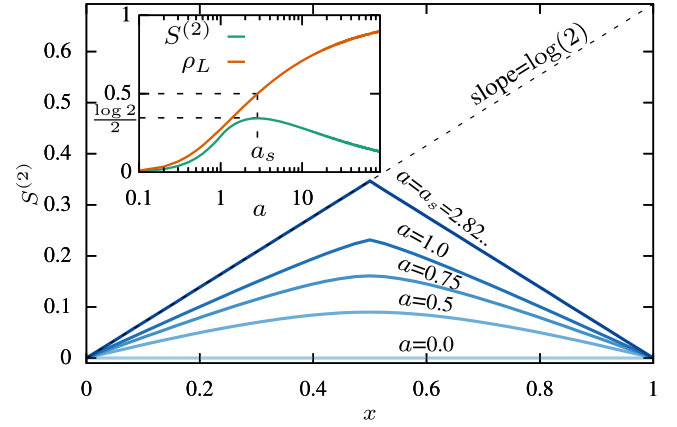


FIG. 4. Second Rényi entropy  $S^{(2)}$  of the state  $|\psi(a)\rangle$  as a function of partition size  $x$  for multiple values of  $a$  and for  $p = 1$  and  $q = 4$ . Inset:  $S^{(2)}(x=0.5)$  and  $\rho_L$  as a function of  $a$ . The state is maximally entangled at the left-right symmetric value of  $a = a_s$ . The entanglement decays monotonically with  $a$  beyond that value.

#### IV. ENTANGLEMENT

The entanglement properties of  $|\psi(a)\rangle$  can also be calculated using a recently developed formalism [37]. Some of the earlier studies on entanglement in the SYK model can be found in Refs. [62–65]. We focus on the second Rényi entropy,  $S^{(2)} = -N^{-1} \ln \text{Tr}[\hat{\rho}_A^2]$ , for a bipartition of the system into regions  $A$  and  $B$ , and where  $\hat{\rho}_A = \text{Tr}_B|\psi(a)\rangle\langle\psi(a)|$  is the reduced-density matrix. The partition is parametrized by  $x \in [0, 1]$ , which gives the proportion of orbitals in  $A$ . For  $x \leq 0.5$  we take region  $A$  to be entirely composed of the left-side fermions, while  $x > 0.5$  also includes a portion  $(x - 0.5)$  of the right-side fermions. The large- $N$  limit for  $S^{(2)}$  is obtained using an approach similar to calculating  $\ln(N)$  (see Appendix E).

The results for  $S^{(2)}$  are shown in Fig. 4 for  $q = 4$ . The  $x$  dependence of  $S^{(2)}$  resembles the one obtained for KM states in the SYK model [51], with a small- $x$  linear behavior indicative of a volume law of entanglement and a maximum at  $x = 0.5$ . Starting from 0 at  $a = 0$ , the entanglement grows until the left-right symmetric point  $a = a_s$  is reached, after which it decays monotonically (see Fig. 4 inset). Remarkably, we find  $S^{(2)}(x) = \min(x, 1-x) \ln(2)$  at  $a = a_s$ , which means  $|\psi(a_s)\rangle$  is maximally entangled between the left and right subsystems.

#### V. DISCUSSION

In this paper, we have highlighted a fundamental difference between bosonic (or qubit) and fermionic  $q$ -local Hamiltonians, as regards the complexity of finding wave functions with a finite approximation ratio ( $E_\psi/E_{\text{GS}} > 0$ ) in the thermodynamic limit. We showed that for a prototypical fermionic model, the SYK model, the bandwidth of Gaussian states scales subextensively with system size, leaving a parametrically large gap between the ground state and Gaussian states. This raises the question of whether other classes of tractable wave functions could (partially) fill this gap. We took a step in that direction by proposing a wave function inspired by the

variational coupled cluster algorithm with a disorder-averaged approximation ratio of  $r \approx 0.62$ .

From a physical perspective, this wave function is easily tractable, since it is described by a static large- $N$  field theory for which saddle-point equations are simply given by polynomial equations. It remains, however, unpractical from a computational point of view since a “brute-force” calculation of its properties would have factorial complexity on a classical computer. Furthermore, to the best of our knowledge, there is no efficient algorithm to prepare a VCC state on a quantum computer. It is therefore desirable to find other classes of wave functions with  $r > 0$  which could efficiently be studied with a classical or quantum computer. Unitary coupled cluster states are particularly promising regarding the latter possibility [7,8,66,67] and could be studied by extending the techniques developed here.

Moreover, our approach of focusing on a subset of terms in the SYK Hamiltonian could be transposed to other versions of SYK models with a reduced number of terms, such as low-rank SYK [68] and sparse SYK [69,70]. More generally, we surmise that large- $N$  techniques and SYK models could prove a useful tool in the search for new variational wave functions.

#### ACKNOWLEDGMENTS

We would like to acknowledge helpful discussions with Ehud Altman, Xiangyu Cao, Matthias Degroote, Tim Hsieh, Bryce Kobrin, Sumilan Banerjee, and Arun Paramakanti. This research was enabled in part by support provided by Compute Canada [71]. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), in particular, the Discovery Grant (RGPIN-2020-05842), the Accelerator Supplement (RGPAS-2020-00060), and the Discovery Launch Supplement (DGECR-2020-00222). T.S. contributed to this work prior to joining Amazon.

#### APPENDIX A: SUBEXTENSIVE SCALING OF ENERGY FOR GAUSSIAN STATES

In this Appendix we show that the energy bandwidth of Gaussian states is subextensive for the SYK model with  $q > 2$ . We start with a derivation specific to  $q = 4$  and then treat the more general case by mapping it to a classical spin-glass model.

*Case of  $q = 4$ .* We use the Majorana version of the SYK model for convenience, written as

$$H = -\frac{1}{N^{3/2}} \sum_{ijkl} J_{ijkl} \gamma_i \gamma_j \gamma_k \gamma_l, \quad (\text{A1})$$

where  $\gamma_i$  represent the Majorana fermions with  $\{\gamma_i, \gamma_j\} = \delta_{ij}$ . Using Wick’s theorem and the permutation properties of  $J_{ijkl}$ , the expectation value of the Hamiltonian for an arbitrary Gaussian state can be written as

$$\begin{aligned} \langle H \rangle &= -\frac{3}{N^{3/2}} \sum_{ijkl} J_{ijkl} \langle \gamma_i \gamma_j \rangle \langle \gamma_k \gamma_l \rangle \\ &= -12 \frac{1}{N^{3/2}} \sum_{i < j, k < l} J_{ijkl} \langle \gamma_i \gamma_j \rangle \langle \gamma_k \gamma_l \rangle. \end{aligned} \quad (\text{A2})$$

Interpreting  $J_{i < j, k < l}$  as a real symmetric matrix and  $L_{i < j} \equiv i \langle \gamma_i \gamma_j \rangle$  as a vector, we go to the eigenbasis of  $J$ , leading to

$$\langle H \rangle = 12 \frac{1}{N^{3/2}} \sum_{i < j, k < l} J_{ijkl} L_{ij} L_{kl} = 12 \frac{1}{N^{3/2}} \sum_{\mu} \lambda_{\mu} L_{\mu}^2, \quad (\text{A3})$$

where  $\lambda_{\mu}$  are the eigenvalues of  $J_{i < j, k < l}$  and  $L_{\mu}$  are its eigenvectors. Minimizing  $\langle H \rangle$  now amounts to minimizing this quadratic form, but with an extensive number of constraints on the values of  $L_{\mu}$  in order for them to be consistent with a Gaussian state. In order to obtain a nontrivial bound on  $\langle H \rangle$ , it is sufficient to take into account the simplest of such constraints, which sets the norm of the vector  $L$ :

$$\sum_{\mu} L_{\mu}^2 = \sum_{i < j} L_{ij}^2 = N/8. \quad (\text{A4})$$

Minimizing the quadratic form under this single constraint is straightforward and leads to the following bound:

$$\langle H \rangle \geq 12 \frac{1}{N^{3/2}} \frac{N}{8} \lambda_{\min}, \quad (\text{A5})$$

where  $\lambda_{\min}$  is the smallest eigenvalue of  $J$ .

We now need to find the scaling of  $\lambda_{\min}$ . In the large- $N$  limit, we expect the matrix  $J_{i < j, k < l}$  to behave as a random matrix of dimension  $O(N^2) \times O(N^2)$  and thus to have a semi-circle distribution of eigenvalues with radius  $O(N)$  (this was verified numerically for  $N$  up to 200). We therefore expect  $\lambda_{\min}$  to be a negative number of order  $N$ . From Eq. (A5), this means that the bandwidth of Gaussian states scales at most as  $\sqrt{N}$  (whereas the full bandwidth scales as  $N$  since it is extensive).

*General case.* We can show that the above subextensive scaling also holds when  $q > 4$  by mapping the problem to the  $p$ -spin spherical spin-glass model [72]. To do this, we start from the following Hamiltonian:

$$H = \frac{i^{q/2}}{N^{(q-1)/2}} \sum_{i_1 \dots i_q} J_{i_1 \dots i_q} \gamma_{i_1} \dots \gamma_{i_q}. \quad (\text{A6})$$

We compute the expectation value for a Gaussian state in a similar way as above, leading to

$$\begin{aligned} \langle H \rangle &= \frac{i^{q/2}}{N^{(q-1)/2}} \frac{q!}{(q/2)!} \\ &\times \sum_{(i_1 < i_2), (i_3 < i_4), \dots, (i_{q-1} < i_q)} J_{i_1, \dots, i_q} \langle \gamma_{i_1} \gamma_{i_2} \rangle \dots \langle \gamma_{i_{q-1}} \gamma_{i_q} \rangle. \end{aligned}$$

Denoting  $a_1 = (i_1 < i_2)$  and similarly for the other indices, we rewrite the expectation value as

$$\langle H \rangle = \frac{1}{N^{(q-1)/2}} \frac{q!}{(q/2)!} \sum_{a_1, \dots, a_{q/2}} J_{a_1 \dots a_{q/2}} L_{a_1} \dots L_{a_{q/2}}, \quad (\text{A7})$$

where  $L_a$  is again understood as an  $[N(N-1)/2]$ -dimensional vector. Even though there exists a large number of constraints on the vector  $L$ , we find again that it is sufficient to impose the simplest one ( $\sum_a L_a^2 = N/8$ ) to obtain a nontrivial bound. This will provide the spherical constraint for the mapping to the spherical  $p$ -spin model. The  $p$ -spin spherical

model is defined as [72]

$$H_{p\text{-spin}} = \frac{1}{M^{(p-1)/2}} \sum_{a_1 \dots a_p=1}^M J_{a_1 \dots a_p} s_{a_1} \dots s_{a_p}, \quad (\text{A8})$$

with  $a \in 1, \dots, M$  and  $\sum_a s_a^2 = M$ , and where  $J_{a_1 \dots a_p}$  are Gaussian-distributed random numbers. This model is extensive: Its bandwidth scales as  $M$ , the number of classical spins.

We can now make the identifications

$$p = q/2, \quad (\text{A9})$$

$$M = \frac{N(N-1)}{2}, \quad (\text{A10})$$

$$s_a = 2\sqrt{N-1}L_a \quad (\text{A11})$$

in order to relate the two models. This finally leads to

$$\langle H \rangle \geq \frac{1}{N^{(q-1)/2}} \frac{q!}{(q/2)!} \frac{1}{(2\sqrt{N-1})^{q/2}} M^{(p-1)/2} E_{p\text{-spin}}, \quad (\text{A12})$$

where  $E_{p\text{-spin}}$  is the ground-state energy of an instance of the spherical  $p$ -spin model for which the couplings  $J_{a_1, \dots, a_p}$  are given by the corresponding  $J_{(i_1 < i_2), \dots, (i_{q-1}, i_q)}$  of the SYK Hamiltonian. We now make the assumption that these instances of the  $p$ -spin spherical model are typical, or in other words that the correlations present in  $J_{(i_1 < i_2), \dots, (i_{q-1}, i_q)}$  due to permutation symmetries can be neglected. If that is the case, we can use the fact that the spherical  $p$ -spin model is extensive to deduce

that  $E_{p\text{-spin}}$  scales as  $M \sim O(N^2)$ . By using this relation, the right-hand side of Eq. (A12) can be shown to scale as  $N^{\frac{3}{2}-\frac{q}{4}}$ . The bandwidth of Gaussian states therefore scales at most as  $N^{\frac{3}{2}-\frac{q}{4}}$ , which is subextensive for  $q > 2$ . Setting  $q = 2$ , we find a Gaussian state bandwidth which is extensive, as expected since in that case the ground state is a Gaussian state. For  $q = 4$ , we find  $\sqrt{N}$  as previously shown. For larger  $q$ , the Gaussian states' bandwidth gets narrower and narrower.

### APPENDIX B: LARGE- $N$ ANALYSIS OF THE VARIATIONAL WAVE FUNCTION

In this Appendix, we discuss the details pertaining to the computation of  $\overline{\ln(\mathcal{N})}$  [see Eq. (9)] in the large- $N$  limit. As stated in the main text, the said quantity works as a generating functional for computing observables and correlation functions for the variational wave function. Since calculating the disorder average of the  $\ln$  term directly is hard, we use the replica trick to represent the term as

$$\overline{\ln(\mathcal{N})} = \lim_{\mathcal{R} \rightarrow 0} \frac{\overline{\mathcal{N}^{\mathcal{R}} - 1}}{\mathcal{R}}, \quad (\text{B1})$$

where  $\mathcal{R}$  denotes the number of replicas. The normalization  $\mathcal{N}(a) = \langle \tilde{0} | \exp(-a\hat{T}) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle$  [see Eq. (3) for the definition of  $\hat{T}$ ] can be written as an integral over the fermionic-coherent states  $|c_{i \in L}\rangle, |h_{j \in R}\rangle$ , representing the particles and holes, such that

$$\begin{aligned} \mathcal{N}^{\mathcal{R}} &= \left( \int \mathcal{D}[c, h] \langle \tilde{0} | \exp(-a\hat{T}) | c_i, h_j \rangle \langle c_i, h_i | \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle \right)^{\mathcal{R}} \\ &= \int \mathcal{D}[c, h] \exp \left[ \sum_{r=1}^{\mathcal{R}} \left( - \sum_{i \in L} \bar{c}_{i,r} c_{i,r} - \sum_{j \in R} \bar{h}_{j,r} h_{j,r} \right) + \sum_r (-a)g \right. \\ &\quad \left. \times \sum_{\substack{i_1 < \dots < i_{(q/2)} \in L, \\ j_1 < \dots < j_{(q/2)} \in R}} (J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}} \bar{c}_{i_1, r} \dots \bar{c}_{i_{(q/2)}, r} \bar{h}_{j_{(q/2)}, r} \dots \bar{h}_{j_1, r} + J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}}^* h_{j_1, r} \dots c_{i_1, r}) \right], \quad (\text{B2}) \end{aligned}$$

where the Grassmann numbers  $\bar{c}_{ir}, c_{ir}, \bar{h}_{ir}, h_{ir}$  are indexed by the replica index  $r$  and the site index  $i$ . Contrary to usual thermal-field theory, the Grassmann numbers do not require an imaginary-time  $\tau$  index since the terms in the cluster operator  $\hat{T}$  commute. Disorder-averaging Eq. (B2) over all possible realizations of  $J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}$  gives us

$$\overline{\mathcal{N}^{\mathcal{R}}} = \int \mathcal{D}[c, h] \exp \left[ \sum_r \left( - \sum_{i \in L} \bar{c}_{ir} c_{ir} - \sum_{j \in R} \bar{h}_{jr} h_{jr} \right) + \frac{2a^2 J^2}{q(\sqrt{N_L N_R})^{q-1}} \sum_{r_1, r_2} \left( \sum_{i \in L} \bar{c}_{i, r_1} c_{i, r_2} \right)^{q/2} \left( \sum_{j \in R} \bar{h}_{j, r_1} h_{j, r_2} \right)^{q/2} \right]. \quad (\text{B3})$$

To obtain the large- $N$  limit of the above integral, we introduce the static Green's functions  $G_c, G_h$  and demand that they must satisfy

$$\begin{aligned} G_c(r_1, r_2) &= -\frac{1}{N_L} \sum_{i \in L} \langle c_{ir_1} \bar{c}_{ir_2} \rangle, \\ G_h(r_1, r_2) &= -\frac{1}{N_R} \sum_{j \in R} \langle h_{jr_1} \bar{h}_{jr_2} \rangle \end{aligned} \quad (\text{B4})$$

at the large- $N$  saddle point. The above constraints can be incorporated into Eq. (B3) using the static self-energies  $\Sigma_c$ ,  $\Sigma_h$  such that

$$\begin{aligned} \overline{\mathcal{N}^{\mathcal{R}}} &= \int \mathcal{D}[c, h] \mathcal{D}[G, \Sigma] \\ &\times \exp \left[ \sum_{r_1, r_2, i} -\bar{c}_{ir_1} [\delta_{r_1, r_2} + \Sigma_c(r_1, r_2)] c_{ir_2} \right. \\ &\quad \left. - \sum_{r_1, r_2, j} \bar{h}_{jr_1} [\delta_{r_1, r_2} + \Sigma_h(r_1, r_2)] h_{jr_2} \right] \\ &\times \exp \left[ \sum_{r_1, r_2} N_L \Sigma_c(r_1, r_2) G_c(r_2, r_1) \right. \\ &\quad + N_R \Sigma_h(r_1, r_2) G_c(r_2, r_1) \\ &\quad \left. + \frac{2a^2 J^2 \sqrt{N_L N_R}}{q} G_c(r_1, r_2)^{q/2} G_h(r_1, r_2)^{q/2} \right], \quad (\text{B5}) \end{aligned}$$

where  $\Sigma_{c,h}$  act as Lagrange multipliers. We integrate out the fermions from the above to get

$$\begin{aligned} \overline{\mathcal{N}^{\mathcal{R}}} &= \int \mathcal{D}[G, \Sigma] \exp(-S[G, \Sigma]), \\ S[G, \Sigma] &= -N_L \ln[\det(\mathbf{1} + \Sigma_c)] - N_R \ln[\det(\mathbf{1} + \Sigma_h)] \\ &\quad - \sum_{r_1, r_2} \left( N_L \Sigma_c(r_1, r_2) G_c(r_2, r_1) \right. \\ &\quad + N_R \Sigma_h(r_1, r_2) G_h(r_2, r_1) \\ &\quad \left. + \frac{2a^2 J^2 \sqrt{N_L N_R}}{q} G_c(r_1, r_2)^{q/2} G_h(r_1, r_2)^{q/2} \right), \quad (\text{B6}) \end{aligned}$$

where we have introduced the effective action  $S[G, \Sigma]$  and  $\mathbf{1}$  represents the identity matrix in the replica space. We evaluate the integral in Eq. (B6) at the saddle point for the action  $S$ . Furthermore, we shall consider a replica-diagonal ansatz for  $G_{c,h}$ ,  $\Sigma_{c,h}$ , i.e.,  $G_{c,h}(r_1, r_2) = \delta_{r_1, r_2} G_{c,h}$  and  $\Sigma_{c,h}(r_1, r_2) = \delta_{r_1, r_2} \Sigma_{c,h}$ . This results in the following simplified form for the effective action:

$$\begin{aligned} S[G, \Sigma] &= -\mathcal{R} N (1+p)^{-1} \left( \ln(1 + \Sigma_c) + p \ln(1 + \Sigma_h) \right. \\ &\quad \left. + \Sigma_c G_c + p \Sigma_h G_h + \frac{2a^2 J^2 \sqrt{p}}{q} G_c^{q/2} G_h^{q/2} \right), \quad (\text{B7}) \end{aligned}$$

where we have used the site ratio  $p = N_R/N_L$  and the total number of sites  $N = N_R + N_L$ . Minimizing the above replica-diagonal action with respect to  $G_{c,h}$ ,  $\Sigma_{c,h}$ , we get the saddle-point conditions

$$\begin{aligned} (1 + \Sigma_{c,d})^{-1} &= -G_{c,d}, \\ \Sigma_c &= -\sqrt{p} a^2 J^2 G_c^{q/2-1} G_d^{q/2}, \\ \Sigma_d &= -a^2 (1/\sqrt{p}) J^2 G_d^{q/2-1} G_c^{q/2}. \quad (\text{B8}) \end{aligned}$$

The value for  $\overline{\ln(\overline{\mathcal{N}})}$  at the saddle point is given by

$$\begin{aligned} \overline{\ln(\overline{\mathcal{N}})} &= \lim_{\mathcal{R} \rightarrow 0} \frac{\exp^{-S[G, \Sigma]} - 1}{\mathcal{R}} \\ &= N(1+p)^{-1} \left( \ln(1 + \Sigma_c) + p \ln(1 + \Sigma_h) \right. \\ &\quad \left. + \Sigma_c G_c + p \Sigma_h G_h + \frac{2a^2 J^2 \sqrt{p}}{q} G_c^{q/2} G_h^{q/2} \right). \quad (\text{B9}) \end{aligned}$$

The expression for  $\overline{\ln(\overline{\mathcal{N}})}$  given in Eq. (9) of the main text is then obtained by setting  $p = 1$  and  $G_c = G_h = G$ ,  $\Sigma_c = \Sigma_h = \Sigma$ , so that

$$\overline{\ln(\overline{\mathcal{N}})} = N \left[ \ln(1 + \Sigma) + \Sigma G + \frac{a^2 J^2}{q} G^q \right]. \quad (\text{B10})$$

Similarly, the saddle-point conditions in Eq. (B8) take the form

$$\begin{aligned} -G^{-1} &= 1 + \Sigma, \\ \Sigma &= -a^2 J^2 G^{q-1}, \quad (\text{B11}) \end{aligned}$$

as reported in Eq. (8) of the main text.

*Energy density.* Having obtained the saddle-point solutions  $G_c$ ,  $G_h$ , we can now calculate the energy density for the ansatz with respect to the full SYK Hamiltonian, which can be written as  $H_{\text{SYK}} = H_{\text{pSYK}} + H_{\text{other}}$ . The  $H_{\text{pSYK}}$  [defined in Eq. (4)] encodes the scattering of  $q/2$  fermions from left (L) to right (R) and vice versa, whereas  $H_{\text{other}}$  denotes the other scattering processes not accounted for by  $H_{\text{pSYK}}$ . It is easy to show that after disorder averaging over  $J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}}$ ,

$$\begin{aligned} E(a) &= \frac{1}{N} \overline{\langle \psi(a) | H_{\text{SYK}} | \psi(a) \rangle} \\ &= \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle} + \frac{1}{N} \underbrace{\overline{\langle \psi(a) | H_{\text{other}} | \psi(a) \rangle}}_{=0} \\ &= \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle}, \quad (\text{B12}) \end{aligned}$$

since the  $J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}}$  in  $H_{\text{other}}$  will appear an odd number of times and average out to zero. We now work with  $\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle$ . Since  $H_{\text{pSYK}} = \hat{T} + \hat{T}^\dagger$  [see Eq. (4)], we have

$$\begin{aligned} -\partial_a \ln(\overline{\mathcal{N}}) &= -\frac{1}{N} \partial_a \ln(\langle \tilde{0} | \exp(a\hat{T}) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle) \\ &= \frac{1}{N} \frac{\langle \tilde{0} | \exp(a\hat{T}) (\hat{T} + \hat{T}^\dagger) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle}{\overline{\mathcal{N}}} \\ &= \langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle, \quad (\text{B13}) \end{aligned}$$

i.e.,  $E(a) = \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle} = -(1/N) \partial_a \overline{\ln(\overline{\mathcal{N}})}$ , which was reported in Eq. (6) of the main text. Proceeding forward, we can calculate  $E(a)$  from  $G_c$ ,  $G_h$  as shown below:

$$\begin{aligned} E(a) &= \frac{1}{N} \overline{\langle \psi(a) | H_{\text{pSYK}} | \psi(a) \rangle} = -\partial_a \frac{\overline{\ln \overline{\mathcal{N}}}}{N_L + N_R} \\ &= -|a| \frac{4J^2}{q} \frac{\sqrt{p}}{1+p} G_c^{q/2} G_h^{q/2}, \quad (\text{B14}) \end{aligned}$$

where we have used Eq. (B9) to take the derivative.

*Density of particles and holes.* The density of particles, say, for the left-side fermions, is obtained by calculating the expectation value

$$\rho_L = \frac{\langle \tilde{0} | \exp(-a\hat{T}) \hat{c}_i^\dagger \hat{c}_i \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle}{\langle \tilde{0} | \exp(-a\hat{T}) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle} \quad (\text{B15})$$

in the large- $N$  limit. Instead of evaluating the above expression directly, we use a chemical-potential-like source term  $\mu$ , such that

$$\rho_L = \partial_{\mu \rightarrow 0} \overline{\ln \langle \tilde{0} | \exp(-a\hat{T}) \exp(\mu \hat{c}_i^\dagger \hat{c}_i) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle}. \quad (\text{B16})$$

The advantage of using a source term is that we can repeat the same analysis used for computing the energy earlier [Eq. (B14)] in this case as well. At the end of this, we get the following saddle-point equations:

$$\begin{aligned} [1 + (1 + \mu)\Sigma_c]^{-1}(1 + \mu) &= -G_c, \\ (1 + \Sigma_h)^{-1} &= -G_h, \\ \Sigma_c &= -a^2 J^2 G_c^{q/2-1} G_h^{q/2}, \\ \Sigma_h &= -a^2 J^2 G_h^{q/2-1} G_c^{q/2}, \end{aligned} \quad (\text{B17})$$

which give back the saddle-point conditions of Eq. (B8) in the  $\mu \rightarrow 0$  limit. The corresponding replica-diagonal action for the  $\ln(\dots)$  term in Eq. (B16) is found to be

$$\begin{aligned} S_{\rho_L}(\mu) &= -\mathcal{R} \left[ N_L \ln \det [1 + (1 + \mu)\Sigma_c] \right. \\ &\quad + N_R \ln \det [1 + \Sigma_h] + \frac{2a^2 J^2 \sqrt{N_L N_R}}{q} G_c^{q/2} G_h^{q/2} \\ &\quad \left. + N_L \Sigma_c G_c + N_R \Sigma_h G_h \right]. \end{aligned} \quad (\text{B18})$$

Using the fact that  $\ln \langle \tilde{0} | \exp(-a\hat{T}) \exp(\mu \hat{c}_i^\dagger \hat{c}_i) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle = S_{\rho_L}(\mu)/\mathcal{R}$ , we compute the derivative of  $S_{\rho_L}(\mu)/\mathcal{R}$  with respect to  $\mu$  as shown below:

$$\begin{aligned} \partial_{\mu \rightarrow 0} \overline{\ln \langle \tilde{0} | \exp(-a\hat{T}) \exp(\mu \hat{c}_i^\dagger \hat{c}_i) \exp(-a\hat{T}^\dagger) | \tilde{0} \rangle} \\ &= \frac{N}{2} [1 + (1 + \mu)\Sigma_c]^{-1} \Sigma_c \\ &= \frac{N}{2} \frac{\Sigma_c}{(1 + \Sigma_c)} = \frac{N}{2} [1 - (1 + \Sigma_c)^{-1}] = \frac{N}{2} [1 + G_c], \end{aligned}$$

which according to Eq. (B16) gives us the density of particles on the left side

$$\rho_L = 1 + G_c. \quad (\text{B19})$$

Similarly, the density of holes on the right side, i.e.,  $\langle \hat{h}_i^\dagger \hat{h}_i \rangle$ , can be calculated by using the source term  $\exp(\mu \hat{h}_i^\dagger \hat{h}_i)$  in place of  $\exp(\mu \hat{c}_i^\dagger \hat{c}_i)$  in Eq. (B16) to get  $\langle \hat{h}_i^\dagger \hat{h}_i \rangle = 1 + G_h$ , from which the density of right-side fermions (particles) can be determined to be

$$\rho_R = 1 - \langle \hat{h}_i^\dagger \hat{h}_i \rangle = -G_h. \quad (\text{B20})$$

### APPENDIX C: THERMAL-FIELD THEORY FOR THE PARTITIONED-SYK MODEL

We now discuss the the thermal-field theory for the partitioned-SYK model. This will allow us to compute the exact properties for the ground state when the temperature  $T$  is extrapolated to *zero*. We reiterate the Hamiltonian for the partitioned-SYK model for ease of access:

$$\begin{aligned} H_{\text{pSYK}} &= g \sum_{\substack{i_1 < \dots < i_{(q/2)} \in L, \\ j_1 < \dots < j_{(q/2)} \in R}} \\ &\quad \times J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}} \hat{c}_{i_1}^\dagger \dots \hat{c}_{i_{(q/2)}}^\dagger c_{j_1} \dots c_{j_{(q/2)}} \\ &\quad + \text{H.c.}, \end{aligned} \quad (\text{C1})$$

where  $g = (q/2)!/\sqrt{q/2}(\sqrt{N_L N_R})^{q/2}$ . The partition function  $Z = \text{Tr}[\exp(-\beta H_{\text{pSYK}})]$ , where  $\beta = T^{-1}$ , can be written as a path integral

$$\begin{aligned} Z &= \int \mathcal{D}[\bar{c}, c] \exp(-S[\bar{c}, c]), \\ S[\bar{c}, c] &= \int_0^\beta d\tau \left[ \sum_i \bar{c}_i(\tau) \partial_\tau c_i(\tau) + g \sum_{\substack{i_1 < \dots < i_{(q/2)} \in L, \\ j_1 < \dots < j_{(q/2)} \in R}} \right. \\ &\quad \left. \times (J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}} \bar{c}_{i_1} \dots \bar{c}_{i_{(q/2)}} c_{j_1} \dots c_{j_{(q/2)}} + \text{H.c.}) \right], \end{aligned} \quad (\text{C2})$$

using the fermionic-coherent states  $|c_{i=1 \dots N}\rangle$  [36] described by antiperiodic Grassmann fields  $\bar{c}_i(\tau)$ ,  $c_i(\tau)$  living on the imaginary-time interval  $\tau \in [0, \beta]$ . In the above equation, we have defined the action  $S[\bar{c}, c]$  whose saddle point would give us access to the large- $N$  limit. Since we are interested in the disorder-averaged free energy

$$F = -T \overline{\ln Z}, \quad (\text{C3})$$

we use the replica trick,  $\overline{\ln Z} = \lim_{\mathcal{R} \rightarrow 0} (\overline{Z^{\mathcal{R}}} - 1)/\mathcal{R}$ , yet again, to perform the averaging over  $J_{i_1 \dots i_{(q/2); j_1 \dots j_{(q/2)}}$ . The replica-partition function  $\overline{Z^{\mathcal{R}}}$  is found to be

$$\begin{aligned} \overline{Z^{\mathcal{R}}} &= \int \mathcal{D}[\bar{c}, c] \exp(-S_{\mathcal{R}}[\bar{c}, c]), \\ S_{\mathcal{R}} &= \int_0^\beta d\tau_{1,2} \sum_{r_1, r_2=1}^{\mathcal{R}} \left[ \sum_{i=1}^N \bar{c}_{ir_1}(\tau_1) \delta_{r_1, r_2} \delta(\tau_1 - \tau_2) \partial_{\tau_1} c_{ir_1}(\tau_2) \right. \\ &\quad \left. - \frac{2J^2}{q(\sqrt{N_L N_R})^{q-1}} \left( \sum_{i \in L} \bar{c}_{i, r_1}(\tau_1) c_{i, r_2}(\tau_2) \right)^{q/2} \right. \\ &\quad \left. \times \left( \sum_{j \in R} c_{j, r_1}(\tau_1) \bar{c}_{j, r_2}(\tau_2) \right)^{q/2} \right], \end{aligned} \quad (\text{C4})$$



where  $S_{\mathcal{R}}$  denotes a new action over replicas and the Grassmann fields  $\bar{c}_{i,r_1}(\tau)$ ,  $c_{i,r_2}(\tau)$  have picked up the additional replica indices  $r_1, r_2$ . Introducing the large- $N$  Green's functions

$$\begin{aligned} G_L^{(r_1,r_2)}(\tau_1, \tau_2) &= -\frac{1}{N_L} \sum_{i \in L} \langle c_{i,r_1}(\tau_1) \bar{c}_{i,r_2}(\tau_2) \rangle, \\ G_R^{(r_1,r_2)}(\tau_1, \tau_2) &= -\frac{1}{N_R} \sum_{j \in R} \langle c_{i,r_1}(\tau_1) \bar{c}_{i,r_2}(\tau_2) \rangle \end{aligned} \quad (\text{C5})$$

for the left-side and right-side fermions along with their associated self-energies  $\Sigma_L^{(r_1,r_2)}(\tau_1, \tau_2)$  and  $\Sigma_R^{(r_1,r_2)}(\tau_1, \tau_2)$  and subsequently integrating out the fermionic fields  $\bar{c}_i, c_i$ , etc., we arrive at the following action:

$$\begin{aligned} S_{\mathcal{R}}[G, \Sigma] &= -N_L \ln \det [\delta_{r_1,r_2} \delta(\tau_1 - \tau_2) \partial_{\tau_1} + \Sigma_L] - N_R \ln \det [\delta_{r_1,r_2} \delta(\tau_1 - \tau_2) \partial_{\tau_1} + \Sigma_R] \\ &\quad - \frac{2J^2 \sqrt{N_L N_R} (-1)^{q/2}}{q} \int_0^\beta d\tau_{1,2} \sum_{r_1, r_2=1}^{\mathcal{R}} G_L^{(r_1,r_2)}(\tau_1, \tau_2)^{q/2} G_R^{(r_2,r_1)}(\tau_2, \tau_1)^{q/2} \\ &\quad - \int_0^\beta d\tau_{1,2} \sum_{r_1, r_2=1}^{\mathcal{R}} [N_L \Sigma_L^{(r_1,r_2)}(\tau_1, \tau_2) G_L^{(r_2,r_1)}(\tau_2, \tau_1) + N_R \Sigma_R^{(r_1,r_2)}(\tau_1, \tau_2) G_R^{(r_2,r_1)}(\tau_2, \tau_1)], \end{aligned} \quad (\text{C6})$$

such that  $\overline{Z^{\mathcal{R}}} = \int \mathcal{D}[G, \Sigma] \exp(-S_{\mathcal{R}}[G, \Sigma])$ . Assuming time-translational invariance and a replica-diagonal ansatz for the saddle point, i.e.,  $G_{L,R}^{(r_1,r_2)}(\tau_1, \tau_2) \propto \delta_{r_1,r_2} G_{L,R}(\tau_1 - \tau_2)$  and the same for  $\Sigma_{L,R}^{(r_1,r_2)}(\tau_1, \tau_2)$ , we obtain a simplified form for the replica action

$$\begin{aligned} S_{\mathcal{R}} &= \mathcal{R} N S \\ &= N(1+p)^{-1} \left[ -\ln \det [\delta(\tau_1 - \tau_2) \partial_{\tau_1} + \Sigma_L] \right. \\ &\quad - p \ln \det [\delta(\tau_1 - \tau_2) \partial_{\tau_1} + \Sigma_R] \\ &\quad - \frac{2J^2 \sqrt{p} (-1)^{q/2}}{q} \beta \int_0^\beta d\tau G_L(\tau)^{q/2} G_R(-\tau)^{q/2} \\ &\quad \left. - \beta \int_0^\beta d\tau [\Sigma_L(\tau) G_L(-\tau) + p \Sigma_R(\tau) G_R(-\tau)] \right], \end{aligned} \quad (\text{C7})$$

where we have used the site ratio  $p = N_R/N_L$  and defined the action per replica  $S$ . The saddle-point conditions are found to be

$$\begin{aligned} G_{L,R} &= -[\delta(\tau_1 - \tau_2) \partial_{\tau_1} + \Sigma_{L,R}]^{-1}, \\ \Sigma_L(\tau) &= (-1)^{q/2+1} J^2 \sqrt{p} G_R(\tau)^{q/2} G_L(-\tau)^{q/2-1}, \\ \Sigma_R(\tau) &= (-1)^{q/2+1} \frac{J^2}{\sqrt{p}} G_L(\tau)^{q/2} G_R(-\tau)^{q/2-1}, \end{aligned} \quad (\text{C8})$$

by minimizing  $S$  with respect to  $G, \Sigma$ . The above equations were solved iteratively [36], for a given value of  $T$  and  $p$ , after discretizing the imaginary-time interval  $[0, \beta]$ . The free energy can then be calculated by plugging the solutions of Eq. (C8) into Eq. (C7) and using

$$F = -T \ln(Z) = -T \lim_{\mathcal{R} \rightarrow 0} \frac{\exp^{-S[G, \Sigma]} - 1}{\mathcal{R}} = T N S. \quad (\text{C9})$$

The ground-state energy density  $E_{\text{pSYK}}$  is calculated using the thermodynamic relation

$$E_{\text{pSYK}} = (F/N) + T \mathbf{s}, \quad (\text{C10})$$

where  $\mathbf{s} = -(1/N) \partial_T F$  is the entropy density obtained from  $F$  via numerical differentiation. The densities for the left- and right-side fermions are obtained using

$$\rho_{L,R} = G_{L,R}(\tau = 0^-), \quad (\text{C11})$$

which follows from the usual definition of the two-point Green's functions. We access the energy density and particle density for the ground state by numerically extrapolating the values for small but finite  $T$  to  $T \rightarrow 0$ .

#### APPENDIX D: THE NONINTERACTING ( $q = 2$ ) PARTITIONED-SYK MODEL

In this Appendix, we study the noninteracting limit of the partitioned-SYK model on a system of  $2N$  sites. The Hamiltonian for the model is obtained by setting  $q = 2$  in Eq. (3), which is

$$H_{\text{pSYK}}(q = 2) = \hat{T} + \hat{T}^\dagger, \quad (\text{D1})$$

where

$$\begin{aligned} \hat{T}^\dagger &= \frac{1}{\sqrt{N}} \sum_{ij} J_{ij} c_i^\dagger d_j, \\ \hat{T} &= \frac{1}{\sqrt{N}} \sum_{ij} J_{ij}^* d_j^\dagger c_i. \end{aligned} \quad (\text{D2})$$

The single-particle spectrum of the above model was checked to be gapless via exact diagonalization of the Hamiltonian. The energy for our variational ansatz was also calculated using the large- $N$  approach discussed in the main text (see Appendix B), and the minimized energy was found to be  $E_{\text{min}} = -0.3849 \dots$ . Interestingly, due to the noninteracting nature of the Hamiltonian, we can calculate this value for energy analytically. We now discuss the analytical approach. For simplicity, let us take  $J_{ij}$  to be a real  $N \times N$  symmetric

matrix, with Gaussian matrix elements having variance 1. We can then diagonalize  $\hat{T}^\dagger$ , leading to

$$\hat{T}^\dagger = \sum_{\mu} \epsilon_{\mu} C_{\mu}^{\dagger} D_{\mu}, \quad (\text{D3})$$

with  $\epsilon_{\mu} \sim O(1)$  distributed according to the semicircle law. The operators  $C_{\mu}^{\dagger}$ ,  $C_{\mu}$  and  $D_{\mu}^{\dagger}$ ,  $D_{\mu}$  represent the single-particle eigenstates (orbitals) obtained after diagonalization and obey fermionic anticommutation relations. We express the variational ansatz in the following way:

$$\begin{aligned} |\psi(a)\rangle &= e^{-a\hat{T}^\dagger} |0\rangle \\ &= \bigotimes_{\mu} (|01\rangle_{\mu} + (-a)\epsilon_{\mu}|10\rangle_{\mu}), \end{aligned} \quad (\text{D4})$$

where  $\bigotimes_{\mu}$  represents the direct product operation and  $|01\rangle_{\mu}$  denotes the state where the  $C_{\mu}$  orbital is occupied and the  $D_{\mu}$  orbital is empty, while the reverse is true for the state  $|10\rangle_{\mu}$ . The energy for the above ansatz is then obtained as

$$\begin{aligned} E/2 &= \frac{\langle \psi(a) | \hat{T} + \hat{T}^\dagger | \psi(a) \rangle}{\langle \psi | \psi \rangle} \\ &= - \sum_{\mu} \epsilon_{\mu} \frac{a\epsilon_{\mu}}{1 + a^2\epsilon_{\mu}^2} \\ &= -\frac{1}{a} \text{Tr} \left[ \frac{a^2 \mathcal{J}^2}{1 + a^2 \mathcal{J}^2} \right] \\ &= -\frac{1}{a} \sum_{n \geq 1} (-1)^{n+1} a^{2n} \text{Tr}[\mathcal{J}^{2n}] \\ &= +\frac{1}{a} \sum_{n \geq 1} (-a^2)^n \text{Tr}[\mathcal{J}^{2n}], \end{aligned} \quad (\text{D5})$$

where we have defined the matrix  $\mathcal{J} = \frac{1}{\sqrt{N}} J_{ij}$ . We can now take the disorder average using random matrix theory

$$\overline{\text{Tr}[\mathcal{J}^{2n}]} = N C_n, \quad (\text{D6})$$

where  $C_n$  are the Catalan numbers. Using the disorder average, we find

$$\overline{E}/2N = -\frac{1}{a} [F(-a^2) - 1], \quad (\text{D7})$$

where  $F(x)$  is the ordinary generating function of the Catalan numbers, given by

$$F(x) = \sum_{n \geq 0} C_n x^n = \frac{1 - \sqrt{1 - 4x}}{2x}. \quad (\text{D8})$$

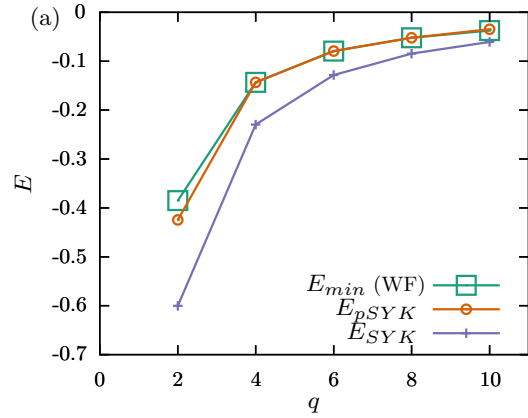
This leads to an expression for the disorder-averaged energy

$$\overline{E}/2N = -\frac{1 + 2a^2 - \sqrt{1 + 4a^2}}{2a^3}. \quad (\text{D9})$$

The minimum of  $\overline{E}/2N$  occurs at  $a = \sqrt{3}/2$ , and the minimum value is  $E_{\min} = \overline{E}/2N = -2/(3\sqrt{3}) \simeq -0.38$ , which is equal to the value reported at the beginning of this Appendix.

*Comparison with exact ground state.* The exact ground state (GS) at half filling is given by

$$|\text{GS}\rangle = \bigotimes_{\mu} \left( \frac{1}{\sqrt{2}} |01\rangle_{\mu} - \frac{1}{\sqrt{2}} \text{sgn}(\epsilon_{\mu}) |10\rangle_{\mu} \right), \quad (\text{D10})$$



(b)

$q$	$E_{\text{SYK}}$	$E_{\min}$	$r = E_{\min}/E_{\text{SYK}}$
2	-0.600	-0.385	0.641
4	-0.230	-0.143	0.623
6	-0.128	-0.079	0.618
8	-0.084	-0.052	0.616
10	-0.061	-0.037	0.619

FIG. 5. (a) Comparison of the exact ground-state energy density  $E_{\text{pSYK}}$  (circles), obtained using thermal-field theory by extrapolating  $T \rightarrow 0$ , with the prediction from the variational ansatz [Eq. (5)]  $E_{\min}$  (squares). The match is excellent for  $q \geq 4$  since the partitioned-SYK model breaks particle-hole symmetry and develops a gap in the single-particle spectrum. However, when  $q = 2$ , i.e., the non-interacting limit of the partitioned-SYK model, the single-particle spectrum is gapless, and the prediction from the ansatz deviates from the exact value. The ground-state energy  $E_{\text{SYK}}$  (denoted with crosses) for the full  $q$ -body SYK model is shown for comparison. (b) Tabular description of the data in (a) listing the numerical values for the approximation ratio  $r = E_{\min}/E_{\text{SYK}}$ .

where the states  $|01\rangle_{\mu}$ ,  $|10\rangle_{\mu}$  have the same meaning as described below Eq. (D4). The ground-state energy density is given by

$$E_{\text{pSYK}} = -\frac{1}{2N} \sum_{\mu} |\epsilon_{\mu}|. \quad (\text{D11})$$

Since  $\epsilon_{\mu}$  are distributed according to the semicircle law, taking the disorder average leads to

$$\begin{aligned} \overline{E}_{\text{pSYK}} &= -\frac{1}{2} \frac{1}{2\pi} \int_{-2}^2 d\epsilon \sqrt{4 - \epsilon^2} |\epsilon| \\ &= -\frac{4}{3\pi} \simeq -0.42. \end{aligned} \quad (\text{D12})$$

Comparing the energy density of the wave function ( $E_{\min} \approx -0.38$ ) with  $E_{\text{pSYK}}$ , we find  $E_{\min} > E_{\text{pSYK}}$  (only slightly). More importantly, we see that, unlike in the  $q \geq 4$  case, the wave function does not predict the energy exactly when the ground state is gapless; see Fig. 5.

#### APPENDIX E: SCALING OF SECOND RÉNYI ENTROPY WITHIN THE VARIATIONAL WAVE FUNCTION

In order to estimate entanglement within the variational wave function, we divide the system into parts  $A$  (the

subsystem) and  $B$  (the rest) and compute the reduced-density matrix  $\hat{\rho}_A$  from the full density matrix

$$\hat{\rho} = |\psi(a)\rangle\langle\psi(a)| = \frac{1}{\mathcal{N}} \exp(-a\hat{T}^\dagger)|\tilde{0}\rangle\langle\tilde{0}| \exp(-a\hat{T}). \quad (\text{E1})$$

We denote the fraction of sites in  $A$  as  $x$ . We demonstrate the computation of Rényi entropy for  $x = 0.5$ , i.e.,  $A$  is composed of the left-side fermions, and report the result for arbitrary  $x$  at the end. Additionally, we also set the site ratio  $p = 1.0$ . The reduced-density matrix  $\hat{\rho}_A$ , when  $x = 0.5$ , is found to be

$$\begin{aligned} \hat{\rho}_A &= \text{Tr}_B[\hat{\rho}] = \int \prod_{i \in R} d\bar{h}_i dh_i \exp\left(-\sum_i \bar{h}_i h_i\right) \langle -h_i | \hat{\rho} | h_i \rangle \\ &= \frac{1}{\mathcal{N}} \int \prod_i d^2 h_i \exp\left(-\sum_i \bar{h}_i h_i\right) \\ &\quad \times \exp\left(-(-1)^{(q/2)} agJ_{ij} c_i^\dagger \cdots c_{i(q/2)}^\dagger \bar{h}_{j_1} \cdots \bar{h}_{j(q/2)}\right) \\ &\quad \times |\tilde{0}\rangle_c \langle \tilde{0}|_c \exp\left(-agJ_{ij}^* h_{j(q/2)} \cdots h_{j_1} \hat{c}_{i(q/2)} \cdots \hat{c}_i\right), \quad (\text{E2}) \end{aligned}$$

where we have used the fermionic-coherent state  $|h_{i \in R}\rangle$  for the holes and their corresponding Grassmann numbers  $\bar{h}_i, h_i$ . The symbol  $|\tilde{0}\rangle_c$  denotes the vacuum for the  $\hat{c}_{i \in L}$  fermions. Since, the second Rényi entropy is related to the second moment of the reduced-density matrix, i.e.,

$$S^{(2)} = -\overline{\ln \text{Tr}[\hat{\rho}_A^2]}, \quad (\text{E3})$$

we represent  $\hat{\rho}_A^2$  as an integral over Grassmann variables

$$\begin{aligned} \hat{\rho}_A^2 &= \frac{1}{\mathcal{N}^2} \left[ \int d^2 h_{J1} \exp(-\bar{h}_{j_1} h_{j_1}) \right. \\ &\quad \times \exp(-(-1)^{(q/2)} gaJ_{IJ} \hat{c}_I^\dagger \bar{h}_{J1}) |\tilde{0}\rangle_c \langle \tilde{0}|_c \\ &\quad \times \exp(-gaJ_{IJ}^* h_{J1} \hat{c}_I) \left. \right] \left[ \int d^2 h_{J2} \exp(-\bar{h}_{j_2} h_{j_2}) \right. \\ &\quad \times \exp(-(-1)^{(q/2)} gaJ_{IJ} \hat{c}_I^\dagger \bar{h}_{J2}) \\ &\quad \times |\tilde{0}\rangle_c \langle \tilde{0}|_c \exp(-gaJ_{IJ}^* h_{J2} \hat{c}_I) \left. \right], \quad (\text{E4}) \end{aligned}$$

where we have introduced the shorthand notation  $I = \{i_1, \dots, i_{(q/2)}\}$ ,  $J = \{j_1, \dots, j_{(q/2)}\}$ ,  $c_I = c_{i_1} \cdots c_{i_{(q/2)}}$ , etc., with sum over repeated indices ( $i, j, I, J$ ) implied. We can evaluate the trace of  $\hat{\rho}_A^2$  by introducing the Grassmann numbers  $\bar{c}_{i1}, c_{i1}$ , to get

$$\begin{aligned} \text{Tr}_A(\hat{\rho}_A^2) &= \frac{1}{\mathcal{N}^2} \int d^2 c_{I1} \exp(-\bar{c}_{i1} c_{i1}) \\ &\quad \times \left[ \int d^2 h_{J1} \exp(-\bar{h}_{j_1} h_{j_1}) \exp((-1)^{q+1} gaJ_{IJ} \right. \\ &\quad \times \bar{c}_{I1} \bar{h}_{J1}) \langle \tilde{0}|_c \exp(-gaJ_{IJ}^* h_{J1} \hat{c}_I) \left. \right] \\ &\quad \times \left[ \int d^2 h_{J2} \exp(-\bar{h}_{j_2} h_{j_2}) \exp(-(-1)^{(q/2)} agJ_{IJ} \right. \\ &\quad \times \hat{c}_I^\dagger \bar{h}_{J2}) \langle \tilde{0}|_c \exp(-agJ_{IJ}^* h_{J2} c_{I1}) \left. \right]. \quad (\text{E5}) \end{aligned}$$

The expectation value appearing inside the trace can be evaluated by introducing  $\bar{c}_{i2}, c_{i2}$ , as shown below:

$$\begin{aligned} &\langle \tilde{0}|_c \exp(-agJ_{IJ}^* h_{J1} \hat{c}_I) \exp((-1)^{1+(q/2)} agJ_{IJ} \hat{c}_I^\dagger \bar{h}_{J2}) |\tilde{0}\rangle_c \\ &= \int d^2 c_{I2} \exp(-\bar{c}_{i2} c_{i2}) \langle \tilde{0}|_c \exp(-agJ_{IJ}^* h_{J1} \hat{c}_I) |c_{I2}\rangle \\ &\quad \times \langle c_{I2}| \exp(-(-1)^{(q/2)} agJ_{IJ} \hat{c}_I^\dagger \bar{h}_{J2}) |\tilde{0}\rangle_c \\ &= \int d^2 c_{I2} \exp(-\bar{c}_{i2} c_{i2}) \exp(-gaJ_{IJ}^* h_{J1} c_{I2}) \\ &\quad \times \exp(-(-1)^{(q/2)} agJ_{IJ} \bar{c}_{I2} \bar{h}_{J2}). \quad (\text{E6}) \end{aligned}$$

Substituting the above expression into Eq. (E5), we get an expression for  $\text{Tr}_A(\hat{\rho}_A^2)$  involving *only* Grassmann variables, i.e.,

$$\text{Tr}_A[\hat{\rho}_A^2] = \frac{1}{\mathcal{N}^2} \int d^2 c_{I1, J2} d^2 h_{J1, J2} \exp(-S), \quad (\text{E7})$$

where the action

$$\begin{aligned} S &= \sum_{\alpha=1}^2 (-\bar{c}_{i\alpha} c_{i\alpha} - \bar{h}_{j\alpha} h_{j\alpha}) \\ &\quad + ga \sum_{I, J} [J_{IJ} \bar{c}_{I1} \bar{h}_{J1} + J_{IJ}^* h_{J1} c_{I2} \\ &\quad + (-1)^{(q/2)} J_{IJ} \bar{c}_{I2} h_{J2} + J_{IJ}^* h_{J2} c_{I1}]. \quad (\text{E8}) \end{aligned}$$

A useful point to note here is that the Grassmann variables are now indexed by an additional number 1 or 2, since we are dealing with the square of the density matrix  $\hat{\rho}_A$ . The disorder averaging of  $\ln \text{Tr}[\hat{\rho}_A^2]$  [see Eq. (E3)] over  $J_{i_1 \dots i_{(q/2)}; j_1 \dots j_{(q/2)}}$  can be implemented in the same way, using the replica trick, as was done for  $\ln \mathcal{N}$  (see Appendix B). Subsequently, the large- $N$  limit of the resulting replica action, as in the  $\ln \mathcal{N}$  case, can also be obtained by introducing the static Green's functions  $[G_{c,h}]$ , see Eq. (B4) and self-energies  $(\Sigma_{c,h})$ , except that this time they are  $2 \times 2$  matrices. Therefore we have

$$\begin{aligned} G_c &= N_L^{-1} \begin{bmatrix} \sum_{i \in L} \langle \bar{c}_{i1} c_{i1} \rangle & \sum_{i \in L} \langle \bar{c}_{i1} c_{i2} \rangle \\ \sum_{i \in L} \langle \bar{c}_{i2} c_{i1} \rangle & \sum_{i \in L} \langle \bar{c}_{i2} c_{i2} \rangle \end{bmatrix}, \\ G_h &= N_R^{-1} \begin{bmatrix} \sum_{i \in R} \langle \bar{h}_{i1} h_{i1} \rangle & \sum_{i \in R} \langle \bar{h}_{i1} h_{i2} \rangle \\ \sum_{i \in R} \langle \bar{h}_{i2} h_{i1} \rangle & \sum_{i \in R} \langle \bar{h}_{i2} h_{i2} \rangle \end{bmatrix} \quad (\text{E9}) \end{aligned}$$

and similar definitions for the self-energies. At the large- $N$  saddle point, the second Rényi entropy is found to be

$$S^{(2)} = F_1 + 2\overline{\ln(\mathcal{N})}, \quad (\text{E10})$$

where

$$\begin{aligned} F_1 &= -\frac{N}{2} [\ln[\det(\mathbf{1} + \Sigma_c)] + \ln[\det(\mathbf{1} + \Sigma_h)] \\ &\quad + \text{Tr}[\Sigma_c G_c] + \text{Tr}[\Sigma_h G_h]] \\ &\quad - \frac{d^2 N}{q} [G_c(2, 1)^{(q/2)} G_h(1, 1)^{(q/2)} \\ &\quad + (-1)^{(q/2)} G_c(2, 2)^{(q/2)} G_h(1, 2)^{(q/2)} \\ &\quad + G_c(1, 1)^{(q/2)} G_h(2, 1)^{(q/2)} \\ &\quad + (-1)^{(q/2)} G_c(1, 2)^{(q/2)} G_h(2, 2)^{(q/2)}], \quad (\text{E11}) \end{aligned}$$

where  $\mathbf{1}$  represents the  $2 \times 2$  identity matrix and  $\ln(\mathcal{N})$  is obtained from Eq. (B10). Also, we have set  $J = 1$ . The  $2 \times 2$  matrices  $G$ ,  $\Sigma$ , appearing above, are found by solving the saddle-point conditions

$$\begin{aligned} (\mathbf{1} + \Sigma_{c,h}) &= -G_{c,h}^{-1}, \\ \Sigma_c(1, 1) &= -a^2 G_h(2, 1)^{(q/2)} G_c(1, 1)^{(q/2)-1}, \\ \Sigma_c(1, 2) &= -a^2 G_h(1, 1)^{(q/2)} G_c(2, 1)^{(q/2)-1}, \\ \Sigma_c(2, 1) &= -a^2 (-1)^{(q/2)} G_h(2, 2)^{(q/2)} G_c(1, 2)^{(q/2)-1}, \\ \Sigma_c(2, 2) &= -a^2 (-1)^{(q/2)} G_h(1, 2)^{(q/2)} G_c(2, 2)^{(q/2)-1}, \end{aligned} \tag{E12}$$

with the equations for the components of  $\Sigma_h$  obtained by interchanging  $c \leftrightarrow h$  in the subscript. Similarly, the result for arbitrary subsystem sizes  $x$  can be found to be

$$S^{(2)}(x) = 2\overline{\ln(\mathcal{N})} + \begin{cases} F(x) & 0 \leq x \leq 0.5 \\ F(1-x) & 0.5 \leq x \leq 1, \end{cases} \tag{E13}$$

where

$$\begin{aligned} F(x) &= -\frac{N}{2} [2x \ln[\det(\mathbf{1} + \Sigma_A)] \\ &+ (1-2x) \ln[\det(\mathbf{1} + \tilde{\Sigma}_B)] + \ln[\det(\mathbf{1} + \Sigma_B)] \\ &+ 2x \text{Tr}[\Sigma_A G_A] + (1-2x) \text{Tr}[\tilde{\Sigma}_B \tilde{G}_B] + \text{Tr}[\Sigma_B G_B]] \\ &- \frac{a^2 N}{q} [(2x G_A(2, 1) + (1-2x) \tilde{G}_B(1, 1))^{\frac{q}{2}} G_B(1, 1)^{\frac{q}{2}} \\ &+ (-1)^{\frac{q}{2}} (2x G_A(2, 2) - (1-2x) \tilde{G}_B(1, 2))^{\frac{q}{2}} G_B(1, 2)^{\frac{q}{2}} \\ &+ (2x G_A(1, 1) + (1-2x) \tilde{G}_B(2, 1))^{\frac{q}{2}} G_B(2, 1)^{\frac{q}{2}} \\ &+ (-1)^{\frac{q}{2}} (2x G_A(1, 2) - (1-2x) \tilde{G}_B(2, 2))^{\frac{q}{2}} G_B(2, 2)^{\frac{q}{2}}], \end{aligned} \tag{E14}$$

and the saddle-point conditions are given by

$$\begin{aligned} G_{A,B} &= -(\mathbf{1} + \Sigma_{A,B})^{-1}, \\ \tilde{G}_B &= -(\mathbf{1} + \tilde{\Sigma}_B)^{-1}, \end{aligned}$$

$$\begin{aligned} \Sigma_A^{11} &= -a^2 G_B(2, 1)^{\tilde{q}} [2x G_A(1, 1) \\ &+ (1-2x) \tilde{G}_B(2, 1)]^{\tilde{q}-1}, \\ \Sigma_A^{12} &= -a^2 G_B(1, 1)^{\tilde{q}} [2x G_A(2, 1) \\ &+ (1-2x) \tilde{G}_B(1, 1)]^{\tilde{q}-1}, \\ \Sigma_A^{21} &= -a^2 [-G_B(2, 2)]^{\tilde{q}} [2x G_A(1, 2) \\ &- (1-2x) \tilde{G}_B(2, 2)]^{\tilde{q}-1}, \\ \Sigma_A^{22} &= -a^2 [-G_B(1, 2)]^{\tilde{q}} [2x G_A(2, 2) \\ &- (1-2x) \tilde{G}_B(1, 2)]^{\tilde{q}-1}, \\ \tilde{\Sigma}_B^{11} &= -a^2 G_B(1, 1)^{\tilde{q}} [2x G_A(2, 1) \\ &+ (1-2x) \tilde{G}_B(1, 1)]^{\tilde{q}-1}, \\ \tilde{\Sigma}_B^{12} &= -a^2 G_B(2, 1)^{\tilde{q}} [2x G_A(1, 1) \\ &+ (1-2x) \tilde{G}_B(2, 1)]^{\tilde{q}-1}, \\ \tilde{\Sigma}_B^{21} &= -a^2 [-G_B(1, 2)]^{\tilde{q}} [2x G_A(2, 2) \\ &- (1-2x) \tilde{G}_B(1, 2)]^{\tilde{q}-1}, \\ \tilde{\Sigma}_B^{22} &= -a^2 [-G_B(2, 2)]^{\tilde{q}} [2x G_A(1, 2) \\ &- (1-2x) \tilde{G}_B(2, 2)]^{\tilde{q}-1}, \\ \Sigma_B^{11} &= -a^2 G_B(1, 1)^{\tilde{q}-1} [2x G_A(2, 1) \\ &+ (1-2x) \tilde{G}_B(1, 1)]^{\tilde{q}}, \\ \Sigma_B^{12} &= -a^2 G_B(2, 1)^{\tilde{q}-1} [2x G_A(1, 1) \\ &+ (1-2x) \tilde{G}_B(2, 1)]^{\tilde{q}}, \\ \Sigma_B^{21} &= -a^2 [-G_B(1, 2)]^{\tilde{q}-1} [2x G_A(2, 2) \\ &- (1-2x) \tilde{G}_B(1, 2)]^{\tilde{q}}, \\ \Sigma_B^{22} &= -a^2 [-G_B(2, 2)]^{\tilde{q}-1} [2x G_A(1, 2) \\ &- (1-2x) \tilde{G}_B(2, 2)]^{\tilde{q}}, \end{aligned} \tag{E15}$$

where we have defined  $\tilde{q} = (q/2)$ . Here,  $G_A$  ( $\tilde{G}_B$ ) represents the Green's function for the left-side fermions in  $A$  ( $B$ ), and  $G_B$  represents the Green's function for the right-side fermions in  $B$ . The same convention applies for the self-energies as well. When we substitute  $x = 0.5$  into Eqs. (E14) and (E15),  $G$ ,  $\Sigma_{A,B} \rightarrow G$ ,  $\Sigma_{c,h}$  while the terms involving  $\tilde{G}_B$  drop out, and we recover Eqs. (E11) and (E12), respectively.

- 
- [1] R. B. Laughlin, Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations, *Phys. Rev. Lett.* **50**, 1395 (1983).
  - [2] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, *Phys. Rev.* **108**, 1175 (1957).
  - [3] M. C. Gutzwiller, Effect of Correlation on the Ferromagnetism of Transition Metals, *Phys. Rev. Lett.* **10**, 159 (1963).
  - [4] N. Laflorencie, Quantum entanglement in condensed matter systems, *Phys. Rep.* **646**, 1 (2016).
  - [5] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
  - [6] S. Lloyd, Universal quantum simulators, *Science* **273**, 1073 (1996).
  - [7] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
  - [8] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
  - [9] F. Arute *et al.*, Google AI Quantum and Collaborators, Hartree-Fock on a superconducting qubit quantum computer, *Science* **369**, 1084 (2020).
  - [10] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, Quantum Hamiltonian complexity, *Found. Trends Theor. Comput. Sci.* **10**, 159 (2015).
  - [11] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics Vol. 47 (American Mathematical Society, Providence, 2002).
  - [12] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, in *FSTTCS 2004: Foundations of Software Technology and Theoretical Computer Science*, edited

- by K. Lodaya and M. Mahajan (Springer, New York, 2005), pp. 372–383.
- [13] J. Watrous, Quantum computational complexity, [arXiv:0804.3401](https://arxiv.org/abs/0804.3401).
- [14] S. Gharibian and J. Kempe, Approximation algorithms for QMA-complete problems, *SIAM J. Comput.* **41**, 1028 (2012).
- [15] D. Aharonov, I. Arad, and T. Vidick, Guest column: the quantum PCP conjecture, *ACM SIGACT News* **44**, 47 (2013).
- [16] F. G. S. L. Brandão and A. W. Harrow, Product-state approximations to quantum states, *Commun. Math. Phys.* **342**, 47 (2016).
- [17] L. Eldar and A. W. Harrow, Local Hamiltonians whose ground states are hard to approximate, in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, 2017), pp. 427–438.
- [18] C. Nirkhe, U. Vazirani, and H. Yuen, Approximate low-weight check codes and circuit lower bounds for noisy ground states, in *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, *Leibniz International Proceedings in Informatics (LIPIcs)*, Vol. 107, edited by I. Chatzigiannakis, C. Kaklamanis, D. Marx, and D. Sannella (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2018), pp. 91:1–91:11.
- [19] M. H. Freedman and M. B. Hastings, Quantum systems on non- $k$ -hyperfinite complexes: A generalization of classical statistical mechanics on expander graphs, [arXiv:1301.1363](https://arxiv.org/abs/1301.1363).
- [20] E. H. Lieb, The classical limit of quantum spin systems, *Commun. Math. Phys.* **31**, 327 (1973).
- [21] B. Hsu, C. R. Laumann, A. M. Läuchli, R. Moessner, and S. L. Sondhi, Approximating random quantum optimization problems, *Phys. Rev. A* **87**, 062334 (2013).
- [22] S. Bravyi, D. Gosset, R. König, and K. Temme, Approximation algorithms for quantum many-body problems, *J. Math. Phys. (Melville, NY)* **60**, 032203 (2019).
- [23] S. Gharibian and O. Parekh, Almost optimal classical approximation algorithms for a quantum generalization of Max-Cut, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, *Leibniz International Proceedings in Informatics (LIPIcs)*, edited by D. Achlioptas and L. A. Végh, Vol. 145 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019), pp. 31:1–31:17.
- [24] S. Sachdev and J. Ye, Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet, *Phys. Rev. Lett.* **70**, 3339 (1993).
- [25] A. Kitaev, A simple model of quantum holography, in *Entanglement in Strongly-Correlated Quantum Matter* (April 6 to July 2, 2015), Talks at KITP, University of California, Santa Barbara, <https://online.kitp.ucsb.edu/online/entangled15/>.
- [26] J. Maldacena and D. Stanford, Remarks on the Sachdev-Ye-Kitaev model, *Phys. Rev. D* **94**, 106002 (2016).
- [27] O. Parcollet and A. Georges, Transition from Overscreening to Underscreening in the Multichannel Kondo Model: Exact Solution at Large  $N$ , *Phys. Rev. Lett.* **79**, 4665 (1997).
- [28] Y. Gu, A. Lucas, and X.-L. Qi, Spread of entanglement in a Sachdev-Ye-Kitaev chain, *J. High Energy Phys.* **09** (2017) 120.
- [29] S. Banerjee and E. Altman, Solvable model for a dynamical quantum phase transition from fast to slow scrambling, *Phys. Rev. B* **95**, 134302 (2017).
- [30] S.-K. Jian and H. Yao, Solvable Sachdev-Ye-Kitaev Models in Higher Dimensions: From Diffusion to Many-Body Localization, *Phys. Rev. Lett.* **119**, 206602 (2017).
- [31] X.-Y. Song, C.-M. Jian, and L. Balents, Strongly Correlated Metal Built from Sachdev-Ye-Kitaev Models, *Phys. Rev. Lett.* **119**, 216601 (2017).
- [32] R. A. Davison, W. Fu, A. Georges, Y. Gu, K. Jensen, and S. Sachdev, Thermoelectric transport in disordered metals without quasiparticles: The Sachdev-Ye-Kitaev models and holography, *Phys. Rev. B* **95**, 155131 (2017).
- [33] A. Haldar, S. Banerjee, and V. B. Shenoy, Higher-dimensional Sachdev-Ye-Kitaev non-Fermi liquids at Lifshitz transitions, *Phys. Rev. B* **97**, 241106(R) (2018).
- [34] A. A. Patel, J. McGreevy, D. P. Arovas, and S. Sachdev, Magnetotransport in a Model of a Disordered Strange Metal, *Phys. Rev. X* **8**, 021049 (2018).
- [35] D. Chowdhury, Y. Werman, E. Berg, and T. Senthil, Translationally Invariant Non-Fermi-Liquid Metals with Critical Fermi Surfaces: Solvable Models, *Phys. Rev. X* **8**, 031024 (2018).
- [36] A. Haldar and V. B. Shenoy, Strange half-metals and Mott insulators in Sachdev-Ye-Kitaev models, *Phys. Rev. B* **98**, 165135 (2018).
- [37] A. Haldar, S. Bera, and S. Banerjee, Rényi entanglement entropy of Fermi liquids and non-Fermi liquids: Sachdev-Ye-Kitaev model and dynamical mean field theories, *Phys. Rev. Research* **2**, 033505 (2020).
- [38] T. Scaffidi and E. Altman, Chaos in a classical limit of the Sachdev-Ye-Kitaev model, *Phys. Rev. B* **100**, 155128 (2019).
- [39] D. E. Parker, X. Cao, A. Avdoshkin, T. Scaffidi, and E. Altman, A Universal Operator Growth Hypothesis, *Phys. Rev. X* **9**, 041017 (2019).
- [40] A. Eberlein, V. Kasper, S. Sachdev, and J. Steinberg, Quantum quench of the Sachdev-Ye-Kitaev model, *Phys. Rev. B* **96**, 205123 (2017).
- [41] A. Haldar, P. Haldar, S. Bera, I. Mandal, and S. Banerjee, Quench, thermalization, and residual entropy across a non-Fermi liquid to Fermi liquid transition, *Phys. Rev. Research* **2**, 013307 (2020).
- [42] A. Almheiri, A. Milekhin, and B. Swingle, Universal constraints on energy flow and SYK thermalization, [arXiv:1912.04912](https://arxiv.org/abs/1912.04912).
- [43] S. Sachdev, Holographic Metals and the Fractionalized Fermi Liquid, *Phys. Rev. Lett.* **105**, 151602 (2010).
- [44] S. Sachdev, Bekenstein-Hawking Entropy and Strange Metals, *Phys. Rev. X* **5**, 041025 (2015).
- [45] J. Maldacena and X.-L. Qi, Eternal traversable wormhole, [arXiv:1804.00491](https://arxiv.org/abs/1804.00491).
- [46] A. Kitaev and S. J. Suh, The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual, *J. High Energy Phys.* **05** (2018) 183.
- [47] X.-L. Qi and P. Zhang, The coupled SYK model at finite temperature, *J. High Energy Phys.* **05** (2020) 129.
- [48] P. Gao, D. L. Jafferis, and A. C. Wall, Traversable wormholes via a double trace deformation, *J. High Energy Phys.* **12** (2017) 151.
- [49] J. Maldacena, D. Stanford, and Z. Yang, Diving into traversable wormholes, *Fortschr. Phys.* **65**, 1700034 (2017).
- [50] I. Kourkoulou and J. Maldacena, Pure states in the SYK model and nearly- $AdS_2$  gravity, [arXiv:1707.02325](https://arxiv.org/abs/1707.02325).

- [51] P. Zhang, Entanglement entropy and its quench dynamics for pure states of the Sachdev-Ye-Kitaev model, *J. High Energy Phys.* **06** (2020) 143.
- [52] For example, when  $q = 4$ , we have  $J_{i_1, i_2; j_1, j_2} = -J_{i_2, i_1; j_1, j_2} = J_{i_2, i_1; j_2, j_1}$  due to fermion anticommutation relations, and  $J_{i_1, i_2; j_1, j_2} = J_{j_2, j_1; i_2, i_1}^*$  for maintaining Hermiticity.
- [53] Strictly speaking, particle-hole symmetry is only present in the  $N \rightarrow \infty$  limit, but all of our results are obtained in that limit anyway.
- [54] S. M. Blinder, Introduction to the Hartree-Fock method, in *Mathematical Physics in Theoretical Chemistry*, Developments in Physical & Theoretical Chemistry, edited by S. M. Blinder and J. E. House (Elsevier, New York, 2019), Chap. 1, pp. 1–30.
- [55] R. J. Bartlett and J. Noga, The expectation value coupled-cluster method and analytical energy derivatives, *Chem. Phys. Lett.* **150**, 29 (1988).
- [56] P. G. Szalay, M. Nooijen, and R. J. Bartlett, Alternative ansätze in single reference coupled-cluster theory. III. A critical analysis of different methods, *J. Chem. Phys.* **103**, 281 (1995).
- [57] T. Van Voorhis and M. Head-Gordon, Benchmark variational coupled cluster doubles results, *J. Chem. Phys.* **113**, 8873 (2000).
- [58] A. Bastianello and S. Sotiriadis, Cluster expansion for ground states of local Hamiltonians, *Nucl. Phys. B* **909**, 1020 (2016).
- [59] F. Coester and H. Kümmel, Short-range correlations in nuclear wave functions, *Nucl. Phys.* **17**, 477 (1960).
- [60] J. Čížek, On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods, *J. Chem. Phys.* **45**, 4256 (1966).
- [61] In usual applications of coupled cluster theory, the partitioning of the system is decided by the Hartree-Fock method, which separates the orbitals that are occupied in the Hartree-Fock state from the others. In that setting, the left-right asymmetry is natural and measures how many particle-hole excitations from the reference Hartree-Fock state are created in order to accommodate the interaction terms in the Hamiltonian. By contrast, in our case the partitioning between left and right is artificial, and the fact that  $|\psi(a_{\min})\rangle$  is unbalanced is an artifact of the technique that should not be present in the true eigenstates of  $H_{\text{SYK}}$ . However, when studying  $H_{\text{pSYK}}$ , this balance is actually physical and is a manifestation of spontaneous symmetry breaking.
- [62] W. Fu and S. Sachdev, Numerical study of fermion and boson models with infinite-range random interactions, *Phys. Rev. B* **94**, 035135 (2016).
- [63] C. Liu, X. Chen, and L. Balents, Quantum entanglement of the Sachdev-Ye-Kitaev models, *Phys. Rev. B* **97**, 245126 (2018).
- [64] A. Goel, H. T. Lam, G. J. Turiaci, and H. Verlinde, Expanding the black hole interior: partially entangled thermal states in SYK, *J. High Energy Phys.* **02** (2019) 156.
- [65] Y. Huang and Y. Gu, Eigenstate entanglement in the Sachdev-Ye-Kitaev model, *Phys. Rev. D* **100**, 041901(R) (2019).
- [66] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus *et al.*, Scalable Quantum Simulation of Molecular Energies, *Phys. Rev. X* **6**, 031007 (2016).
- [67] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure, *Phys. Rev. A* **95**, 020501(R) (2017).
- [68] J. Kim, X. Cao, and E. Altman, Low-rank Sachdev-Ye-Kitaev models, *Phys. Rev. B* **101**, 125112 (2020).
- [69] A. M. García-García, Y. Jia, D. Rosa, and J. J. M. Verbaarschot, Sparse Sachdev-Ye-Kitaev model, quantum chaos and gravity duals, [arXiv:2007.13837](https://arxiv.org/abs/2007.13837).
- [70] S. Xu, L. Susskind, Y. Su, and B. Swingle, A sparse model of quantum holography, [arXiv:2008.02303](https://arxiv.org/abs/2008.02303).
- [71] <https://www.computecanada.ca/>.
- [72] A. Crisanti and H. J. Sommers, The spherical  $p$ -spin interaction spin glass model: the statics, *Z. Phys. B: Condens. Matter* **87**, 341 (1992).