Local integrals of motion detection of localization-protected topological order

Thorsten B. Wahl [1], Florian Venn,[1] and Benjamin Béri[1,2]

[1]*DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*
[2]*T.C.M. Group, Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*

Many-body-localized (MBL) phases can be topologically distinct, but distinguishing these phases using order parameters can be challenging. Here we show how topologically distinct local integrals of motion, variationally parametrized by quantum circuits, can be used to numerically demonstrate the topological inequivalence of MBL phases. We illustrate our approach on a fermionic chain where both topologically distinct MBL phases and benchmark comparisons to order parameters are possible. We also use our approach, augmented by the excited state density-matrix renormalization group algorithm, to extract high-energy topological doublets. We describe applying our methods to higher dimensions to identify MBL topological order and topological multiplets hidden by the dense many-body spectrum.

I. INTRODUCTION

Many-body localization (MBL) [1–12] has attracted a wealth of interest in the last 15 years. One of the most striking features of MBL is the violation of the eigenstate thermalization hypothesis [13,14]: MBL systems do not thermalize, but instead retain some memory of their initial state. As a result, some MBL systems are able to protect quantum information [15–18].

MBL systems are characterized by local integrals of motion (LIOMs) [12,19–22]: exponentially localized operators commuting with the Hamiltonian and each other. As a result, all eigenstates of MBL systems obey the entanglement area law [23]. Therefore, topological order [24–26], normally present only in ground states, can also occur in high-energy MBL eigenstates [15–17,27–32]. Nonetheless, owing to the lack of local order parameter, and due to the overlap (in energy) of topological multiplets away from the strongly MBL regime or beyond one dimension (1D), numerically detecting the topology of MBL phases remains challenging for reasons beyond the mere exponential scaling of the Hilbert space [33].

In this work, we show how a topological LIOM framework [32] (with LIOMs and tLIOMS for topologically trivial and nontrivial cases, respectively), combined with quantum circuits for MBL [34], can be used to numerically detect topological MBL. Furthermore, as we also show, when used in conjunction with the excited-state density-matrix renormalization group (DMRG-X) [35,36], this approach can also identify topological multiplets provided the system is deep in a topological MBL phase.

Due to its polynomial scaling with system size, and the generality of tLIOMs, our approach is a general way to capture topological MBL, including beyond 1D. Nonetheless, to demonstrate its use, we focus on 1D: we study the disordered interacting Kitaev chain, displaying two topologically distinct MBL phases [16,37]. Studying this system is useful not only

due to the exact diagonalization (ED) benchmark available in 1D, but also due to a local order parameter benchmark available thanks to a duality to a system displaying conventional symmetry breaking.

II. (TOPOLOGICAL) LIOMS FROM QUANTUM CIRCUITS

LIOMs are typically assumed to be related to Pauli-z operators σ_i^z acting on site i via a local unitary transformation U, $\tau_i^z = U\sigma_i^z U^\dagger$. The τ_i^z are thus exponentially localized. They define a complete set of quantum numbers since $[H, \tau_i^z] = [\tau_i^z, \tau_j^z] = 0 \ \forall i, j = 1, \ldots, N$, with N the system size. The locality of U implies that the eigenstates of H are local-unitary related to local product states: H cannot display topological order [38,39]. Hence, for topological MBL systems the notion of LIOMs has to be extended [32]: One must use tLIOMs, given by $\tau_i = U S_i U^\dagger$, where U is again a local unitary, but $\{S_i\}$ is now a set of mutually commuting local stabilizers [40,41] whose common eigenstates all display (the same) topological order. On a topologically nontrivial manifold, to get a complete set of quantum numbers one augments tLIOMs by the nonlocal $\tau_i^{nl} = U S_i^{nl} U^\dagger$, where S_i^{nl} are noncontractible Wilson loops (i.e., logical operators) [32].

To utilize tLIOMs numerically, we use that owing to its locality, we can efficiently approximate U by a fixed-depth quantum circuit U_{SR} [34]. We employ U_{SR} variationally: we aim for $[H, \tau_i^z] = 0$ by minimizing

$$f = \frac{1}{2} \sum_{i=1}^{N} \text{tr}([\tau_i, H][\tau_i, H]^\dagger), \qquad (1)$$

where $\tau_i = U_{SR} S_i U_{SR}^\dagger$ and S_i are topological stabilizers for a tLIOM *Ansatz*, while $S_i = \sigma_i^z$ for a conventional LIOM *Ansatz*. The lower f, the better the approximate (t)LIOMs describe the system. In the topological case, the nonlocal τ_i^{nl} enter only for multiplet splittings. Hence, they almost
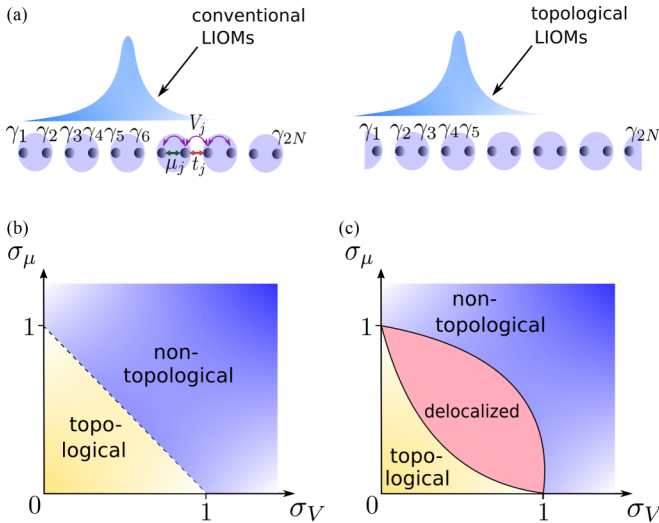
FIG. 1. (a) Couplings in the Hamiltonian, Eq. (2), with (t)LIOMs. Black dots denote Majorana fermions and blue disks the stabilizers underlying the two types of LIOMs. The stabilizers for conventional LIOMs involve Majorana pairs on physical sites; those for topological LIOMs have Majorana pairs straddling physical sites. (b),(c) Possible phase diagrams.

commute with H if all tLIOMs are optimized and thus can be omitted in Eq. (1). To minimize f, we proceed similarly to Ref. [34]: We expand H as a sum of local terms and identify which combination of local terms contributes to f (which is quadratic in H). Each of the contributions is then represented as a tensor network contraction, which we evaluate efficiently using common numerical methods [34,42,43]. While developed for nontopological systems, this approach directly applies also to the topological case: by using topological S_i (which act on multiple sites) we merely increase the number of tensors in each contraction. In the nontopological phase, we expect the conventional LIOM *Ansatz* to perform better. In the topological phase, however, the tLIOM *Ansatz* will minimize f.

III. MODEL

The disordered interacting Kitaev chain [37] is a system of Majorana fermions $\gamma_n = \gamma_n^\dagger$ with $\{\gamma_m, \gamma_n\} = 2\delta_{mn}$; the Hamiltonian is [cf. Fig. 1(a)]

$$H = \sum_{j=1}^{N-1} it_j \gamma_{2j} \gamma_{2j+1} + \sum_{j=1}^{N} i\mu_j \gamma_{2j-1} \gamma_{2j}$$
$$+ \sum_{j=1}^{N-1} V_j \gamma_{2j-1} \gamma_{2j} \gamma_{2j+1} \gamma_{2j+2}, \quad (2)$$

where the tunnel amplitude t_j, the on-site potential μ_j, and the interaction strength V_j are Gaussian distributed random variables with zero mean and standard deviation $\sigma_t = 1$, σ_μ, and σ_V, respectively. The Hamiltonian commutes with the fermion parity operator $Z = \prod_{n=1}^{2N} \gamma_n$, which splits the Hilbert space into two parity sectors.

For $\sigma_\mu, \sigma_V \gg 1$, the system is in the trivial phase with commuting projector representative

$$H^{\text{triv}} = \sum_{j=1}^{N} i\mu_j \gamma_{2j-1} \gamma_{2j} + \sum_{j=1}^{N-1} V_j \gamma_{2j-1} \gamma_{2j} \gamma_{2j+1} \gamma_{2j+2}. \quad (3)$$

H^{triv} may be expressed in terms of the stabilizers $S_j^{\text{triv}} = i\gamma_{2j-1} \gamma_{2j}$, $j = 1, \ldots, N$; consequently, its eigenstates correspond to the occupation of fermionic modes associated to physical sites [cf. Fig. 1(a)].

For $\sigma_\mu = \sigma_V = 0$ the system is in the topological phase.

$$H^{\text{topo}} = \sum_{j=1}^{N-1} it_j \gamma_{2j} \gamma_{2j+1}, \quad (4)$$

with $S_j^{\text{topo}} = i\gamma_{2j} \gamma_{2j+1}$, $j = 1, \ldots, N-1$, gives the commuting projector representative of this phase. The eigenstates again correspond to the occupation of fermionic modes, but now they come from pairs of Majorana modes straddling physical sites [cf. Fig. 1(a)]. $S^{\text{nl}} = i\gamma_{2N} \gamma_1$ forms a zero-energy fermion mode, resulting in a twofold degenerate energy spectrum.

The phases are also known along the axes of the phase diagram: For $\sigma_V = 0$ the model is noninteracting and is in the topological (trivial) phase for $\sigma_\mu < 1$ ($\sigma_\mu > 1$) [44–46]. Similarly, for $\sigma_\mu = 0$, there is a phase transition point at $\sigma_V = 1$ [47,48]. The two MBL phases can either be separated by a critical line or a delocalized phase [30,49] [cf. Figs. 1(b) and 1(c)].

IV. ED AND ORDER PARAMETER BENCHMARKS

To provide a benchmark for our subsequent tLIOM analysis, we use ED to locate the MBL phases. We first analyze the level spacing for $N = 14$. In Fig. 2(a) we show the gap ratio $r_n = \min(s_{n-1}, s_n)/ \max(s_{n-1}, s_n)$ with $s_n = E_n - E_{n-1}$ the level spacing *in a given parity sector*. For each data point we average over the mid-third of energies (as those best reflect whether the system is thermalizing [27,49,50]) and 100 disorder realizations. In an MBL phase we expect $r_P = 0.386$ [5] (Poisson distribution) due to the lack of level repulsion; in a thermal phase we expect $r_{\text{WD}} = 0.530$ (Wigner-Dyson distribution). Figure 2(a) indicates that there are two MBL phases (with $r \approx r_P$): one for $\sigma_\mu + \sigma_V \lesssim 1$ and another one for $\sigma_\mu \gtrsim 1$ or $\sigma_V \gtrsim 1$. Although the gap ratio never gets as large as r_{WD}, between these regions our results are consistent with an extended delocalized phase [49,51,52].

Topological properties can be detected by the spin-glass order parameter [16,27,53] $\chi_n^{\text{SG}} = \frac{1}{N} \sum_{i,j=1}^{N} \langle n|\sigma_i^x \sigma_j^x|n\rangle^2$ in eigenstates $|n\rangle$ of the quantum Ising chain linked to our system via Jordan-Wigner transformation (cf. Appendix A). While this is a two-point correlator of local operators in the spin language, it is a nonlocal order parameter (with γ_j strings), as befits one detecting topological features, for the fermionic system. We also use a "dual order parameter" $\chi_n^{\text{dSG}} = \frac{1}{N} \sum_{i,j=1}^{N} \langle n|\widetilde{\sigma}_i^x \widetilde{\sigma}_j^x|n\rangle^2$ where $\widetilde{\sigma}_j^x$ are analogous to disorder operators [54] in the Ising chain (cf. Appendix A). $\chi_n^{\text{SG}} \propto N$ for $N \gg 1$ in the topological phase, while $\chi_n^{\text{SG}} \to 1$ in the trivial phase. Conversely, $\chi_n^{\text{dSG}} \propto N$ for $N \gg 1$ in the trivial phase, while $\chi_n^{\text{dSG}} \to 1$ in the topological phase. Our
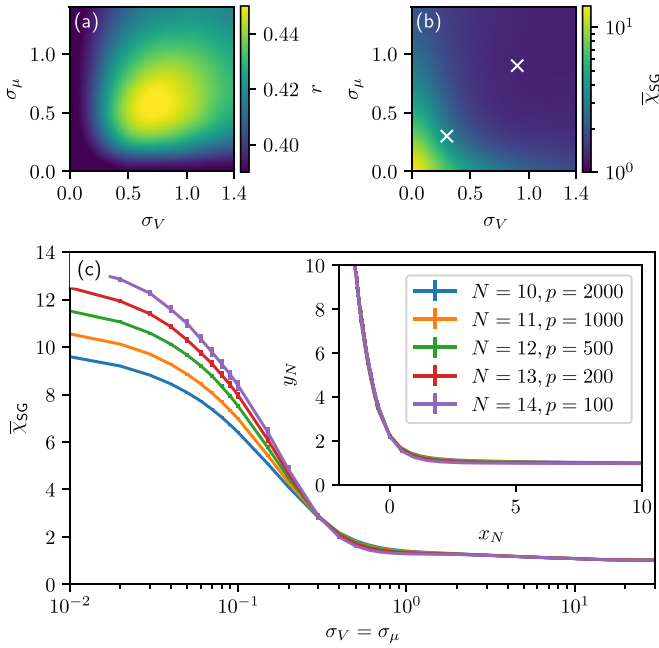
FIG. 2. Exact diagonalization results. In all panels, we show averages over p disorder realizations, both parity sectors, and the mid-third energies. (a) Gap ratio and (b) average spin-glass order parameter $\overline{\chi}^{SG}$, both for $p = 100$ and system size $N = 14$. The maximal standard error of the mean (taken across disorder realizations in all panels) for any value of σ_V, σ_μ is 0.005 in (a) and 0.22 in (b). (c) $\overline{\chi}^{SG}$ for different N and p along the diagonal $\sigma_V = \sigma_\mu$. The error bars mark the standard error of the mean. A scaling collapse [27] with $x_N = (\sigma_\mu - \sigma_c)N^{0.6}$, $y_N = \overline{\chi}^{SG}/N^{0.1}$ is shown in the inset. The extracted critical point $\sigma_c = 0.3$ and the one for the dual order parameter, $\sigma_c' = 0.9$ (see Appendix B), are marked by crosses in (b).

results for $\chi_n^{(d)SG}$ [see Figs. 2(b) and 2(c) and the Appendix B] suggest that the $\sigma_\mu + \sigma_V \lesssim 1$ phase is topological, while the other MBL phase is trivial.

V. PHASE DIAGRAM FROM tLIOMs

We now apply the tLIOM approach to the model. We optimize, using the algorithm of Ref. [34], the quantum circuits U_{SR} and $U_{SR}^{(t)}$ (for approximate LIOMs and tLIOMs, respectively) over the space of fermion parity conserving unitaries. We use two-layer quantum circuits with gates acting on ℓ sites each. In Fig. 3(a), we show the resulting normalized figure of merit f ($f^{(t)}$) for conventional (topological) LIOMs, for system size $N = 48$, gate lengths $\ell = 2, 4, 6$, and focusing on $\sigma_V = \sigma_\mu$. We also show $f^{(t)}/f$ for $\ell = 6$ in the entire two-parameter phase diagram. As a function of ℓ, we see a roughly exponential improvement deep in the MBL phases for the corresponding set of LIOMs.

Another way to map out the phase diagram using (t)LIOMs is entanglement entropy fluctuations [27,34]. We cut the approximate matrix product eigenstates, obtained from the optimized (t)LIOM *Ansätze*, at a point at least $3\ell/2$ sites away from the boundary and apply the algorithm of Ref. [34] to compute the average entanglement entropy. (Here, we averaged over all approximate eigenstates, as their entropies depend only on the expectation values of the tLIOMs near
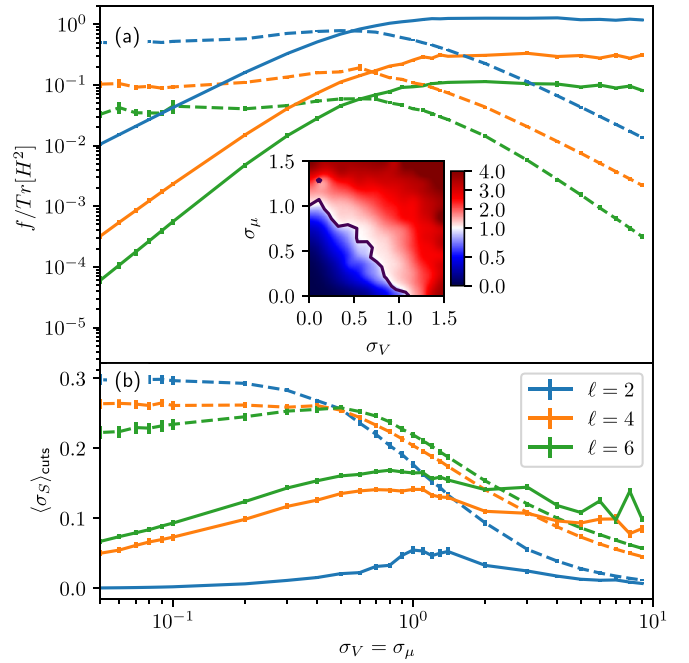

FIG. 3. (a) Figure of merit $f/\text{tr}(H^2)$ for topological (conventional) LIOMs, shown with solid (dotted) lines, for $N = 48$, averaged over 40 disorder realizations. The error bars mark the standard error of the average across the different disorder realizations. The inset shows the ratio $f^{(t)}/f$ for $N = 48$, averaged over 20 disorder realizations and using $\ell = 6$. The contour ($f^{(t)} = f$) separates the topological and trivial MBL phase, but is agnostic about the extent of the delocalized phase. (b) Standard deviation σ_S of the approximate-eigenstate-averaged entanglement entropy with respect to disorder realizations and subsequent averaging over entanglement cuts [34] using the same data as in (a). The error bars mark the standard error of $\langle\sigma_S\rangle_{\text{cuts}}$ across the cut positions.

the cut [34], which are not directly linked to overall energies.) We repeat this step for all disorder realizations and calculate the standard deviation σ_S of these averages. Finally, to reduce statistical fluctuations, we average σ_S over all cut positions to obtain $\langle\sigma_S\rangle_{\text{cuts}}$.

We expect $\langle\sigma_S\rangle_{\text{cuts}} \to 0$ for $\sigma_V = \sigma_\mu \to \infty$, because there the LIOMs are simply the trivial stabilizers (exact LIOM *Ansatz* with $U_{SR} = \mathbb{1}$): the eigenstates are independently populated local fermion modes. Similarly, $\sigma_V = \sigma_\mu = 0$ is the topological MBL limit with tLIOMs simply the topological stabilizers (exact tLIOM *Ansatz* with $U_{SR}^{(t)} = \mathbb{1}$), independently of the disorder realizations, which again implies $\langle\sigma_S\rangle_{\text{cuts}} = 0$. Away from these limits, but still in an MBL phase, i.e., with all eigenstates MBL, the entanglement and its fluctuation are low due to the area law. In an ergodic phase, where all eigenstates are volume-law entangled, the equivalence of spectral and ensemble (i.e., over disorder) averages [55] implies suppressed entanglement fluctuations. In a thermal but not ergodic phase, with a $0 < \nu < 1$ fraction of volume-law states, we expect entanglement fluctuations to diverge in the thermodynamic limit due to ν being sensitive to small changes in the disorder [27].

The maximum amount of entanglement allowed by our approximation is proportional to ℓ; hence, we expect $\langle\sigma_S\rangle_{\text{cuts}}$

to acquire a maximum, increasing with ℓ [34], where entanglement fluctuations would diverge. (This is analogous to the finite-size scaling in Ref. [27].) This is consistent with Fig. 3 if one combines the LIOM and tLIOM data in the regimes where they are reliable.

Based on the comparison between the optimized figures of merit for $\sigma_\mu = \sigma_V$, we expect the topological MBL phase to extend at most up to $\sigma_\mu = \sigma_V \lesssim 0.5$ and the trivial MBL phase to start above $\sigma_\mu = \sigma_V \gtrsim 0.5$. But the broad maxima of the entanglement entropy fluctuation of $U_{SR}^{(t)}$ at $\sigma_\mu = \sigma_V \approx 0.5$ can indicate an interstitial thermal phase.[1] This is also consistent with conventional and topological LIOMs performing comparably in a window around $\sigma_\mu = \sigma_V \approx 0.5$. Outside of this window, however, the performance difference between conventional and topological LIOM *Ansätze* allows us to clearly detect the topology of the MBL phase.

The picture obtained from our (t)LIOM approach is thus consistent with the (dual) spin-glass order parameter and the gap ratio from ED. While meeting this benchmark, using (t)LIOMs we could probe system sizes well beyond the reach of ED. This suggests that our method can generalize well for detecting topological MBL phases beyond 1D.

VI. TOPOLOGICAL DOUBLETS

Another topological MBL feature is the presence of topological multiplets across the entire energy spectrum [15,16]. (The multiplet splitting is exponentially small in system size.) In the Kitaev chain, one has topological doublets [16,37]. Here we show how these can be detected using tLIOMs. A key ingredient is the access to the approximate nonlocal $\tau_i^{nl} = U_{SR}^{(t)} S_i^{nl} U_{SR}^{(t)\dagger}$ and the corresponding flip operator (an MBL strong zero mode [56]) $\tau^{nl,x} = U_{SR}^{(t)} \gamma_1 U_{SR}^{(t)\dagger}$: the approximate eigenstate $|\psi_\alpha\rangle$ forms a doublet with $\tau^{nl,x}|\psi_\alpha\rangle$. The goal is to show that these two states have approximately the same energy. To enhance energy accuracy, and to get a method that can access spatial features beyond ℓ (which is required for studying multiplet splittings in large systems; see Appendix C), we couple the tLIOM *Ansatz* with DMRG-X [35]: we write $|\psi_\alpha\rangle$ and $\tau^{nl,x}|\psi_\alpha\rangle$ as matrix product states to use as inputs to DMRG-X. We denote the resulting outputs by $|\Psi_\alpha\rangle$ and $|\Psi_\alpha^\tau\rangle$, respectively. As DMRG-X is biased towards low-entanglement states [57], we must prevent convergence to $|\Psi_\alpha^{(\tau)}\rangle$ with different tLIOM expectation values than $|\psi_\alpha\rangle$ but with lower entanglement. To control for this, a convenient proxy is to require high $\sqrt[N]{|\langle\Psi_\alpha|\psi_\alpha\rangle|}$ and $\sqrt[N]{|\langle\Psi_\alpha^\tau|\tau^{nl,x}|\psi_\alpha\rangle|}$ which for tensor network states are broadly interpretable as geometric mean "per-site overlaps."

We now turn to demonstrating this approach. To be able to compare to ED, we work with $N = 15$ and deep in the topological MBL phase so that ED can assign almost all doublets correctly due to their small level splitting. We optimize tLIOMs for ten disorder realizations and for each realization
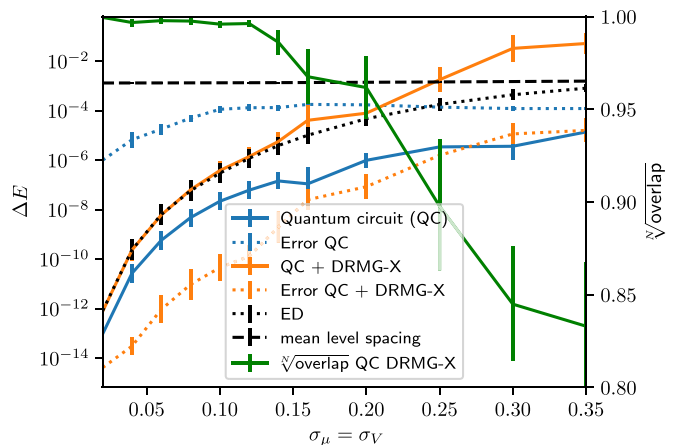

FIG. 4. Doublet splittings from DMRG-X-augmented tLIOMs, tLIOMs [quantum circuit (QC)], and ED for $N = 15$ and ten disorder realizations. The errors refer to the energy differences from the nearest ED energies. The tLIOM + DMRG-X results are averaged over ten randomly chosen approximate eigenstates, while all other data are averaged over all eigenstates, for each disorder realization. The green line shows the average per-site overlap between corresponding QC and DMRG-X-optimized states. Apart from the mean level spacing (dashed), all averages here use geometric means; this is due to splittings varying by orders of magnitude between realizations, and due to each per-site overlap being interpretable as a geometric mean. The error bars mark the standard error of the geometric mean.

we randomly choose ten approximate eigenstates and their doublet partners. In DMRG-X, we use bond dimensions up to 85. Our DMRG-X-augmented tLIOM results, direct tLIOM, and ED data are shown in Fig. 4. The DMRG-X-augmented tLIOM splittings are in quantitative agreement with ED up to $\sigma_V = \sigma_\mu \approx 0.13$. Beyond this, the breakdown of the method is signaled by the deteriorating per-site overlaps. (While we demonstrated its use for system sizes where a comparison to ED is available, the per-site overlap is expected to be a useful and feasibly computable indicator for system sizes beyond the reach of ED.)

Deep in the topological MBL phase, we expect this method to correctly predict the splittings for N beyond the reach of ED. The required bond dimensions will be larger, but still scale polynomially with N due to having area-law eigenstates. Beyond 1D, one could similarly take an approximate eigenstate obtained from the tLIOM approach [58,59] in the topological phase as a starting point for an approximation using projected entangled pair states (PEPS) [60] methods [61]. Although the mean level spacing is exponentially smaller than the multiplet splitting [33], the locality of PEPS will allow one to ensure convergence to the correct eigenstates.

VII. CONCLUSIONS

We showed how tLIOMs can detect topological MBL in numerical simulations. Using (t)LIOM numerics for the interacting disordered Kitaev chain (up to $N = 48$), we found two topologically distinct MBL phases and features consistent with an intervening thermal phase. We also showed how to detect topological multiplets, using tLIOMs which, for strong MBL and when augmented by DMRG-X, reached very high

---

[1]To definitively establish whether this interstitial thermal phase is present, one may seek clearer maxima using larger ℓ. However, this is beyond the scope of this work; our focus is on detecting topological MBL.

energy accuracy. It would be interesting to use tLIOMs above 1D (e.g., for the toric code) to detect MBL topological order and to detect topological multiplets, which here have splitting decaying slower with system size than the mean level spacing even for strong MBL.

APPENDIX A: ISING CHAINS VIA JORDAN-WIGNER TRANSFORMATIONS

We rewrite the Hamiltonian H using a Jordan-Wigner transformation, $\gamma_{2j-1} = \left(\prod_{k=1}^{j-1} \sigma_k^z\right)\sigma_j^x$, $\gamma_{2j} = -\left(\prod_{k=1}^{j-1} \sigma_k^z\right)\sigma_j^y$, $j = 1, 2, \ldots, N$,

$$H = \sum_{j=1}^{N-1} t_j \sigma_j^x \sigma_{j+1}^x + \sum_{j=1}^{N} \mu_j \sigma_j^z - \sum_{j=1}^{N-1} V_j \sigma_j^z \sigma_{j+1}^z, \quad \text{(A1)}$$

with Pauli operator σ_i^α, $\alpha = x, y, z$ acting on site i. Due to fermion parity conservation, the Hamiltonian is invariant under the parity operator $Z = \prod_{k=1}^{N} \sigma_k^z$. This generalized quantum Ising Hamiltonian is the same as the "hJJ' model" studied in Ref. [53]. Its authors introduced a real-space renormalization group method for exited states, predicting two distinct MBL phases separated by a transition visible in the entire energy spectrum.

An alternative Jordan-Wigner transformation $\gamma_{2j} = \left(\prod_{k=1}^{j-1} \tilde{\sigma}_k^z\right)\tilde{\sigma}_j^x$, $\gamma_{2j+1} = -\left(\prod_{k=1}^{j-1} \tilde{\sigma}_k^z\right)\tilde{\sigma}_j^y$ [$\gamma_1 = -\left(\prod_{k=1}^{N-1} \tilde{\sigma}_k^z\right)\tilde{\sigma}_N^y$] results in Kramers-Wannier-like duality, such that

$$H = \sum_{j=1}^{N-1} t_j \tilde{\sigma}_j^z + \sum_{j=2}^{N} \mu_j \tilde{\sigma}_{j-1}^x \tilde{\sigma}_j^x - \sum_{j=2}^{N-1} V_j \tilde{\sigma}_{j-1}^x \tilde{\sigma}_{j+1}^x$$
$$- \mu_1 \tilde{Z}\tilde{\sigma}_N^x \tilde{\sigma}_1^x + V_1 \tilde{Z}\tilde{\sigma}_N^x \tilde{\sigma}_2^x \quad \text{(A2)}$$

with $\tilde{Z} = \prod_{j=1}^{N} \tilde{\sigma}_j^z = -Z$ and $\tilde{\sigma}_j^\alpha$ likewise Pauli operators, but now acting on links. The two sets of Pauli operators are related via the transformation

$$\sigma_j^x = \begin{cases} -\left(\prod_{k=1}^{N-1} \tilde{\sigma}_k^z\right)\tilde{\sigma}_N^y & \text{if } j < N \\ -\tilde{\sigma}_N^y & \text{if } j = N, \end{cases} \quad \text{(A3)}$$

$$\sigma_j^z = \begin{cases} \tilde{\sigma}_{j-1}^x \tilde{\sigma}_j^x & \text{if } j > 1 \\ -\tilde{Z}\tilde{\sigma}_1^x \tilde{\sigma}_N^x & \text{if } j = 1, \end{cases} \quad \text{(A4)}$$

and inverse transformation

$$\tilde{\sigma}_j^x = \begin{cases} -\sigma_1^y\left(\prod_{k=2}^{j} \sigma_k^z\right) & \text{if } j > 1 \\ -\sigma_1^y & \text{if } j = 1, \end{cases} \quad \text{(A5)}$$

$$\tilde{\sigma}_j^z = \begin{cases} \sigma_j^x \sigma_{j+1}^x & \text{if } j < N \\ -Z\sigma_1^x \sigma_N^x & \text{if } j = N. \end{cases} \quad \text{(A6)}$$
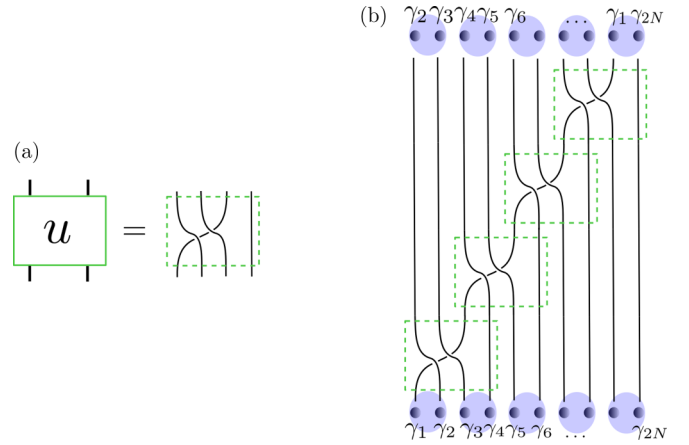

FIG. 5. Unitary mapping between the bases $\{\sigma_i^\alpha\}$ and $\{\tilde{\sigma}_i^\alpha\}$. (a) Constituting unitaries u acting on two sites, i.e., four Majorana fermions. (b) The tensor network of the overall unitary is constructed as a "staircase" of the small unitaries u (denoted by dashed boxes), giving rise to the shown Majorana braid.

[An actual Kramers-Wannier transformation, where the $-\sigma_1^y$ in Eq. (A5) are replaced by σ_1^z and the second line in Eq. (A6) by σ_N^x, changes only the last two terms of Eq. (A2), which become local boundary terms.] In each symmetry sector of $\tilde{Z}$ (even/odd parity) Eq. (A2) is a generalization of the quantum Ising model with next-nearest neighbor coupling considered in Ref. [27] (after an on-site unitary transformation sending $\sigma_i^x \leftrightarrow \sigma_i^z$). There, the authors likewise found two distinct MBL phases as evidenced by spin-glass order. We note that for $\mu_j = 0 \; \forall j = 1, \ldots, N$ the Hamiltonian in Eq. (A2) decouples *in each parity sector* into two noninteracting chains [cf. the structure of Eq. (A1) for $\sigma_V = 0$ [47,48]]. Hence, there is a phase transition point at $(\sigma_V, \sigma_\mu) = (1, 0)$.

The bases $\{\sigma_j^\alpha\}$ and $\{\tilde{\sigma}_j^\alpha\}$ are related via a deep quantum circuit, namely, the one made up of a "staircase" of unitaries acting on nearest neighbors. In the fermionic picture these unitaries permute the Majorana fermions as $(\gamma_1, \gamma_2, \gamma_3, \gamma_4) \rightarrow (\gamma_2, \gamma_3, \gamma_1, \gamma_4)$ (see Fig. 5).

APPENDIX B: ORDER PARAMETERS

The spin-glass order parameter [16,27,53] measures the presence of regions with fixed σ^x magnetization separated by domain walls and thus detects the spin-glass phase. It is given by

$$\chi_n^{\text{SG}} = \frac{1}{N} \sum_{j,k=1}^{N} \langle n|\sigma_j^x \sigma_k^x|n\rangle^2 \quad \text{(B1)}$$

for the eigenstate $|n\rangle$. In terms of Majorana fermions,

$$\chi_n^{\text{SG}} = 1 + \frac{2}{N} \sum_{k>j}^{N} (-1)^{k-j} \langle n|\prod_{i=j}^{k-1} \gamma_{2i}\gamma_{2i+1}|n\rangle^2 \quad \text{(B2)}$$

involves topological Majorana strings, i.e., contiguous products of topological stabilizers $S_j = i\gamma_{2j}\gamma_{2j+1}$ [cf. Fig. 1(a) of
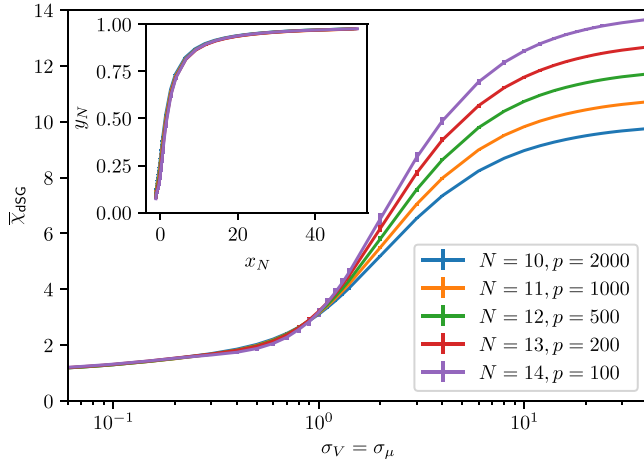
FIG. 6. ED results for $\overline{\chi}^{\text{dSG}}$ for different system sizes N averaged over p disorder realizations along the diagonal $\sigma_V = \sigma_\mu$. We show a scaling collapse with $x_N = (\sigma_\mu - \sigma_c')N^{0.1}$ and $y_N = \overline{\chi}^{\text{dSG}}/N^{1.0}$ in the inset. The extracted critical point $\sigma_c' = 0.9$ is marked by a cross in Fig. 2(b). The error bars mark the standard error of the mean.

the main text]. The dual order parameter

$$\chi_n^{\text{dSG}} = \frac{1}{N} \sum_{j,k=1}^{N} \langle n|\widetilde{\sigma}_j^x \widetilde{\sigma}_k^x|n\rangle^2 \qquad \text{(B3)}$$

detects spin-glass order in terms of the dual variables $\widetilde{\sigma}_i^x$ in Eq. (A2). In terms of Majorana fermions

$$\chi_n^{\text{dSG}} = 1 + \frac{2}{N} \sum_{k>j}^{N} (-1)^{k-j} \langle n| \prod_{i=j+1}^{k} \gamma_{2i-1} \gamma_{2i}|n\rangle^2 \qquad \text{(B4)}$$

involves strings of trivial stabilizers $S_j = i\gamma_{2j-1}\gamma_{2j}$ [cf. Fig. 1(a) of the main text]. Since the bases $\{\sigma_j^\alpha\}$ and $\{\widetilde{\sigma}_j^\alpha\}$ are related by a unitary transformation exchanging stabilizer sets of complementary topological character (see Fig. 5), while $\chi_n^{\text{SG}} \propto N$ (for $N \gg 1$) for MBL with topological order (and $\chi_n^{\text{SG}} \to 1$ otherwise), we have $\chi_n^{\text{dSG}} \propto N$ (for $N \gg 1$) for topologically trivial MBL (and $\chi_n^{\text{dSG}} \to 1$ otherwise). This behavior can be seen in Fig. 6, where we plot the dual order parameter as a function of $\sigma_\mu = \sigma_V$.

APPENDIX C: QUANTUM CIRCUIT SPLITTING FOR LARGE SYSTEMS

In this section we demonstrate that the splittings corresponding to the quantum circuit *Ansatz* are zero for sufficiently large systems; however, we note that the QC states are a good starting point for the DMRG-X calculation regardless.

The eigenstates that the quantum circuit approximation of tLIOMs predicts are given by

$$|\psi_\pm\rangle = U_{\text{SR}}|\mathbf{t}_\pm\rangle, \qquad \text{(C1)}$$

where $|\mathbf{t}_\pm\rangle$ is an eigenstate of the topological stabilizers $S_j^{\text{topo}} = i\gamma_{2j}\gamma_{2j+1}$ with eigenvalues $t_j = \pm 1$ and a ± 1 eigenstate of the bilocal stabilizer $S^{\text{nl}} = i\gamma_{2N}\gamma_1$. It may be written
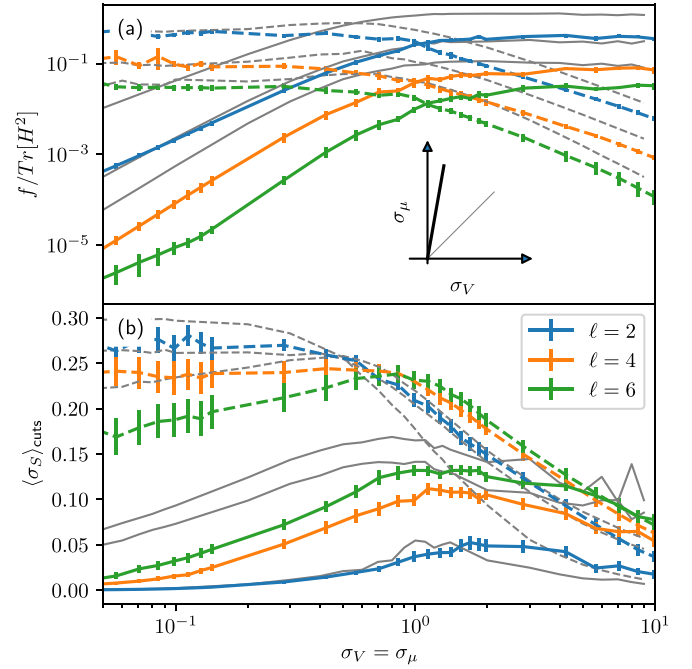

FIG. 7. (a) Figure of merit $f/\text{tr}(H^2)$ and (b) standard deviation of the entanglement entropy for topological (conventional) LIOMs, shown with solid (dotted) lines, for ten disorder samples and $N = 48$ evaluated along a line away from the diagonal in the phase diagram as indicated in the inset. For comparison we added the data from Fig. 3 in gray. In (a) the error bars show the standard error of the average across the different disorder realizations and in (b) the error bars show the standard error of $\langle\sigma_S\rangle_{\text{cuts}}$ across the cut positions.

in terms of stabilizers as

$$|\mathbf{t}_\pm\rangle\langle\mathbf{t}_\pm| = \left(\frac{1}{2} \pm \frac{1}{2}S^{\text{nl}}\right) \prod_{j=1}^{N-1} \left(\frac{1}{2} + \frac{1}{2}t_j S_j^{\text{topo}}\right). \qquad \text{(C2)}$$

Using this representation, the expectation value for the energy can be expressed as

$$\langle\psi_\pm|H|\psi_\pm\rangle$$

$$= \text{Tr}\left[\left(\frac{1}{2} \pm \frac{1}{2}S^{\text{nl}}\right) \prod_{j=1}^{N-1} \left(\frac{1}{2} + \frac{1}{2}t_j S_j^{\text{topo}}\right) U_{\text{SR}}^\dagger H U_{\text{SR}}\right] \qquad \text{(C3)}$$

and the corresponding splitting is given by

$$\Delta E = \langle\psi_+|H|\psi_+\rangle - \langle\psi_-|H|\psi_-\rangle$$

$$= \text{Tr}\left[S^{\text{nl}} \prod_{j=1}^{N-1} \left(\frac{1}{2} + \frac{1}{2}t_j S_j^{\text{topo}}\right) U_{\text{SR}}^\dagger H U_{\text{SR}}\right]. \qquad \text{(C4)}$$

To further analyze the expression, we can expand the Hamiltonian in local terms

$$H = \sum_{k=1}^{N} h_k; \qquad \text{(C5)}$$

with that the splitting is given by

$$\Delta E = \sum_{k=1}^{N} \text{Tr} \left[S^{\text{nl}} \prod_{j=1}^{N-1} \left(\frac{1}{2} + \frac{1}{2} t_j S_j^{\text{topo}} \right) U_{\text{SR}}^\dagger h_k U_{\text{SR}} \right]. \quad \text{(C6)}$$

Since the quantum circuit acts locally, the causal cones [34] of each term h_k in the Hamiltonian will, for a sufficiently large system size N, not cover both of the Majorana modes making up S^{nl}. Hence, either γ_1 or γ_N (or both) will factor from the trace, leading to $\Delta E = 0$.

The true splitting is of the order $\exp(-N/\xi_L)$, i.e., it is due to the exponentially decaying coupling between the end modes in the bilocal integral of motion τ^{nl}. It is thus not surprising that the QC approximation alone, which uses strictly short-range gates, cannot reproduce the splitting. We can capture ΔE better once the accuracy of the approximation has been improved, e.g., by feeding the QC approximate eigenstates into the DMRG-X algorithm. As pairs of QC initial states have the same configuration of approximate tLIOMs (other than for τ^{nl}) and the DMRG-X algorithm tries to maintain a high overlap with the initial states, it will build up additional entanglement around these and thus improve their accuracy. This keeps the configuration of tLIOMs fixed for sufficiently large bond dimensions. Since the DMRG-X algorithm is based on matrix product states, it is able to build entanglement across the entire chain and will thus reproduce the energy splittings, including their dependence on system size, even for large systems.

APPENDIX D: SENSITIVITY TO THE THERMAL PHASE

To further illustrate the predictive power of the tLIOM approach, we have evaluated the (t)LIOMs along the $\sigma_V / \sigma_\mu = \tan(10°)$ line, where we expect a smaller extent of the thermal phase [cf. Fig 1(c)]. Along this line the figure of merit [Fig. 7(a)] has a smaller transition region in which both the LIOMs and tLIOMs poorly describe the system, consistent with a larger regime being a part of an MBL phase. Furthermore, focusing on the entanglement entropy fluctuations [Fig. 7(b)], we find that the maxima they assume are smaller. Since the thermal phase is responsible for generating those maxima, this is also consistent with a smaller extent.

[1] L. Fleishman and P. W. Anderson, Phys. Rev. B **21**, 2366 (1980).

[2] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. **95**, 206603 (2005).

[3] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Ann. Phys. **321**, 1126 (2006).

[4] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).

[5] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).

[6] J. H. Bardarson, F. Pollmann, and J. E. Moore, Phys. Rev. Lett. **109**, 017202 (2012).

[7] J. Z. Imbrie, J. Stat. Phys. **163**, 998 (2016).

[8] R. Nandkishore and D. A. Huse, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[9] E. Altman and R. Vosk, Annu. Rev. Condens. Matter Phys. **6**, 383 (2015).

[10] D. A. Abanin and Z. Papić, Ann. Phys. **529**, 1700169 (2017).

[11] F. Alet and N. Laflorencie, C. R. Phys. **19**, 498 (2018).

[12] J. Z. Imbrie, V. Ros, and A. Scardicchio, Ann. Phys. **529**, 1600278 (2017).

[13] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).

[14] M. Srednicki, Phys. Rev. E **50**, 888 (1994).

[15] B. Bauer and C. Nayak, J. Stat. Mech. (2013) P09005.

[16] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, Phys. Rev. B **88**, 014206 (2013).

[17] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, Nat. Commun. **6**, 7341 (2015).

[18] M. Goihl, C. Krumnow, M. Gluza, J. Eisert, and N. Tarantino, SciPost Phys. **6**, 072 (2019).

[19] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, Phys. Rev. B **91**, 085425 (2015).

[20] L. Rademaker and M. Ortuño, Phys. Rev. Lett. **116**, 010404 (2016).

[21] A. K. Kulshreshtha, A. Pal, T. B. Wahl, and S. H. Simon, Phys. Rev. B **98**, 184201 (2018).

[22] M. Goihl, M. Gluza, C. Krumnow, and J. Eisert, Phys. Rev. B **97**, 134202 (2018).

[23] M. Friesdorf, A. H. Werner, W. Brown, V. B. Scholz, and J. Eisert, Phys. Rev. Lett. **114**, 170505 (2015).

[24] A. Kitaev, Ann. Phys. (NY) **321**, 2 (2006).

[25] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).

[26] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Rev. Mod. Phys. **80**, 1083 (2008).

[27] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Phys. Rev. Lett. **113**, 107204 (2014).

[28] K. Slagle, Z. Bi, Y.-Z. You, and C. Xu, arXiv:1505.05147.

[29] T. B. Wahl, Phys. Rev. B **98**, 054204 (2018).

[30] A. Chan and T. B. Wahl, J. Phys.: Condens. Matter **32**, 305601 (2020).

[31] J. Li, A. Chan, and T. B. Wahl, Phys. Rev. B **102**, 014205 (2020).

[32] T. B. Wahl and B. Béri, Phys. Rev. Research **2**, 033099 (2020).

[33] S. A. Parameswaran and R. Vasseur, Rep. Prog. Phys. **81**, 082501 (2018).

[34] T. B. Wahl, A. Pal, and S. H. Simon, Phys. Rev. X **7**, 021018 (2017).

[35] V. Khemani, F. Pollmann, and S. L. Sondhi, Phys. Rev. Lett. **116**, 247204 (2016).

[36] X. Yu, D. Pekker, and B. K. Clark, Phys. Rev. Lett. **118**, 017201 (2017).

[37] A. Y. Kitaev, Phys.-Usp. **44**, 131 (2001).

[38] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).

[39] X. Chen, Z.-C. Gu, and X.-G. Wen, Phys. Rev. B **82**, 155138 (2010). M. B. Hastings, Phys. Rev. Lett. **107**, 210501 (2011).

[40] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, California Institute of Technology, 1997.

[41] M. A. Nielsen and I. L. Chuang, *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, UK, 2000).

[42] D. G. A. Smith and J. Gray, Journal of Open Source Softw. **3**, 753 (2018).

[43] S.-J. Ran, E. Tirrito, C. Peng, X. C. Tagliacozzo, G. Su, and M. Lewenstein, *Tensor Network Contractions* (Springer Link, 2020).

[44] D. S. Fisher, Phys. Rev. Lett. **69**, 534 (1992).

[45] R. Shankar and G. Murthy, Phys. Rev. B **36**, 536 (1987).

[46] D. S. Fisher, Phys. Rev. B **51**, 6411 (1995).

[47] J.-J. Miao, H.-K. Jin, F.-C. Zhang, and Y. Zhou, Phys. Rev. Lett. **118**, 267701 (2017).

[48] M. McGinley, J. Knolle, and A. Nunnenkamp, Phys. Rev. B **96**, 241113(R) (2017).

[49] W. P. Sze, T. K. Ng, and K. T. Law, arXiv:2005.11812.

[50] D. J. Luitz, N. Laflorencie, and F. Alet, Phys. Rev. B **91**, 081103(R) (2015).

[51] S. Moudgalya, D. A. Huse, and V. Khemani, arXiv:2008.09113.

[52] R. Sahay, F. Machado, B. Ye, C. R. Laumann, and N. Y. Yao, Phys. Rev. Lett. **126**, 100604 (2021).

[53] D. Pekker, G. Refael, E. Altman, E. Demler, and V. Oganesyan, Phys. Rev. X **4**, 011052 (2014).

[54] E. Fradkin and L. Susskind, Phys. Rev. D **17**, 2637 (1978).

[55] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, Rev. Mod. Phys. **53**, 385 (1981).

[56] P. Fendley, J. Phys. A: Math. Theor. **49**, 30LT01 (2016).

[57] T. Devakul, V. Khemani, F. Pollmann, D. A. Huse, and S. L. Sondhi, Philos. Trans. R. Soc. A **375**, 20160431 (2017).

[58] T. B. Wahl, A. Pal, and S. H. Simon, Nat. Phys. **15**, 164 (2019).

[59] J. Li, A. Chan, and T. B. Wahl, arXiv:2108.08268.

[60] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.

[61] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).

[62] See, www.csd3.cam.ac.uk.

[63] See, www.dirac.ac.uk.