Identifying correlation clusters in many-body localized systems

Kévin Hémery [1,*] Frank Pollmann,[1,2] and Adam Smith [1,3,4]

[1]*Department of Physics, TFK, Technische Universität München, James-Franck-Straße 1, 85748 Garching, Germany*
[2]*Munich Center for Quantum Science and Technology, Schellingstrasse 4, D-80799 Munich, Germany*
[3]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*
[4]*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems,*
University of Nottingham, Nottingham NG7 2RD, United Kingdom

We introduce techniques for analyzing the structure of quantum states of many-body localized (MBL) spin chains by identifying correlation clusters from pairwise correlations. These techniques proceed by interpreting pairwise correlations in the state as a weighted graph, which we analyze using an established graph theoretic clustering algorithm. We validate our approach by studying the eigenstates of a disordered XXZ spin chain across the MBL to ergodic transition, as well as the nonequilibrium dynamics in the MBL phase following a global quantum quench. We successfully reproduce theoretical predictions about the MBL transition obtained from renormalization group schemes. Furthermore, we identify a clear signature of many-body dynamics analogous to the logarithmic growth of entanglement. The techniques that we introduce are computationally inexpensive and, in combination with matrix product state methods, allow for the study of large-scale localized systems. Moreover, the correlation functions we use are directly accessible in a range of experimental settings, including cold atoms.

I. INTRODUCTION

Initiated by the seminal work of Anderson [1], many-body localization (MBL) is now understood as a dynamical quantum phase of matter [2,3], defined by the properties of its highly excited many-body eigenstates. In particular, the entanglement of eigenstates in the MBL phase has been found to obey an area law even at finite energy densities [4–6] and to violate the eigenstate thermalization hypothesis [7,8] due to the existence of quasilocal conserved quantities [4,6,9–11]. The concept of MBL has since proven central to the understanding of several aspects of nonequilibrium physics. For instance, MBL is essential to stabilize various emergent Floquet phases of matter, such as discrete time crystals [12,13].

The study of MBL has been driven by large-scale numerics and experimental advances in the control of isolated quantum systems. These efforts have identified characteristic properties of MBL, such as the unbounded logarithmic growth of entanglement following a global quench [14–18], which distinguishes it from Anderson localization (AL), where the entanglement saturates, and the presence of an eigenstates transition to an ergodic phase at finite disorder strengths [19–24]. A slow growth of entanglement-related quantities has since been observed experimentally for small system sizes in Rydberg atomic systems and in superconducting circuits [25,26]. However, extracting the entanglement entropy experimentally generically requires high-fidelity measurements of a number of nonlocal observables that scale exponentially with

system size. This makes experimental measurements of the entanglement entropy prohibitively difficult for large systems. In cold-atom setups, large systems and long times can be reached, even in two dimensions, and clear signals of MBL have been detected in local measurements [27–29].

In spite of the recent progress, the MBL transition is still not fully understood. While we have powerful numerical and analytical techniques that allow us to investigate the slightly entangled eigenstates deep in the MBL phase [9,10,30–33], the transition to the ergodic phase is much harder to study. Phenomenological renormalization group (RG) approaches have emerged as a promising theoretical description of the transition [34–39]. Although the assumptions behind the various models differ, most of them describe the MBL transition in terms of the proliferation of "thermal blocks" versus "insulating blocks," i.e., regions of the spin chain that look locally thermal and fully localized, respectively. However, the interpretation of these approaches rests on phenomenological assumptions which could bias the results. Indeed, most models assume that each of these blocks is local, although the existence of sparse thermal blocks spanning the whole chain has also been suggested [38]. These RG studies suggest that the MBL transition is of the Kosterlitz-Thouless type with a delocalization mechanism called *avalanche instability*, also sometimes referred as *quantum avalanche* [35,40].

Since RG approaches provide a clear mechanism for the transition and allow for a prediction of scaling behavior close to the transition, it is desirable to have a clear prescription in order to identify these "blocks" from states obtained numerically or experimentally. The first numerical validation of this picture in a microscopic model was provided in Ref. [41],

*kevin.hemery@tum.de

which proposed identifying these blocks by finding what the authors denoted *entanglement clusters*. These are clusters of spins that are more strongly entangled with each other than the rest of the system. A numerical investigation using exact diagonalization for small systems revealed that the average block size of these entanglement clusters is, indeed, consistent with the RG analysis of the transition. Entanglement entropy is paramount for this approach, but it is costly to calculate both numerically and from experimental measurements.

Motivated by that work, we propose an approach in which we identify these structures in MBL systems in a scalable way that is relevant for efficient matrix product state (MPS) based simulations and is accessible in experiments. We focus on the XXZ spin chain in the presence of a disordered z-directed field defined by the Hamiltonian

$$\hat{H} = -J \sum_i \left[\hat{S}_i^x \hat{S}_{i+1}^x + \hat{S}_i^y \hat{S}_{i+1}^y + \Delta \hat{S}_i^z \hat{S}_{i+1}^z + h_i \hat{S}_i^z \right]. \quad (1)$$

The disordered field h_i is sampled uniformly from the interval $[-W, W]$, where $W \geqslant 0$ controls the strength of the disorder. We consider the Anderson insulator at $\Delta = 0$ as well as the Heisenberg model at $\Delta = 1$, which is believed to exhibit an MBL transition at W_C estimated to be between 2.7 and 3.8 [19,20,24,42–44].

In this paper, we present practical tools to efficiently identify the ergodic clusters within MBL eigenstates using pairwise correlations by applying methods originally developed in the context of graph theory [45–48]. We validate our approach in two ways: First, we show that the two-site mutual information (TSMI) is a useful proxy for analyzing the structure of MBL eigenstates. Second, we demonstrate in Sec. IV that our clustering algorithm applied during time evolution, using the TSMI as well as the pairwise connected correlation functions in the σ_z basis, indicates the logarithmic spreading of entanglement.

II. FROM CORRELATIONS TO GRAPH THEORY

The quantum mutual information of two subsystems A and B is a correlation measure defined as

$$I(A; B) = S(A) + S(B) - S(A \cup B), \quad (2)$$

where $S(A) = -\text{Tr}[\rho_A \ln(\rho_A)]$ is the von Neumann entanglement entropy for subsystem A. The TSMI corresponds to the case where subsystems A and B each consist of a single site, and in this case we denote it $I(i; j)$ for sites i and j. The TSMI captures the classical and quantum correlations between two sites and has already been shown to be a relevant probe of localization [49]. In particular, spatial fluctuations in the TSMI grow logarithmically under nonequilibrium dynamics, mirroring the entanglement entropy [49]. Another useful quantity to study quantum correlations is the (connected) correlation function $C(\hat{O}_A, \hat{O}_B)$ of two operators $\hat{O}_A$ and $\hat{O}_B$:

$$C(\hat{O}_A, \hat{O}_B) = \langle \hat{O}_A \hat{O}_B \rangle - \langle \hat{O}_A \rangle \langle \hat{O}_B \rangle, \quad (3)$$

where $\langle \hat{O} \rangle$ denotes the expectation value of the operator $\hat{O}$. Although TSMI takes into account all pairwise correlations [50,51], it is less accessible in experiments than certain correlation functions. In this paper we introduce tools borrowed

from the field of graph theory to extract what we call *correlation clusters*, in analogy to Ref. [41]. This provides an efficient method for studying correlations in MBL systems. Graph theory has been used in the past to detect quantum phase transition in equilibrium settings [52–54]. Recently, another work identified the so-called ergodic bubbles (i.e., regions of space where the expectation values of local operators look thermal) using neural network techniques [55].

Our starting point is to construct a matrix M_{ij} containing the correlations between sites i and j, and to interpret it as an adjacency matrix for a weighted graph, as illustrated in Fig. 1. The vertices of this graph are the lattice sites of our system, and the bonds connecting them are weighted by the matrix element M_{ij} between that pair. Our goal to find the correlation clusters in the state translates to finding "communities" within this graph. We consider $M_{ij} = I(i; j)$ in the case of eigenstates, to which we add $M_{ij} = C(\hat{\sigma}_i^z, \hat{\sigma}_j^z)$ for dynamics.

The task of finding communities has received considerable attention in the field of graph theory [45–48]. This is usually achieved by splitting the graph into disjoint sets of vertices which we refer to as *clusters*. A given decomposition of a graph into clusters is referred to as a *clustering*. The goal is to find a clustering that is optimal by some well-defined measure. Inspired by the well-established Girvan-Newman approach [45,47], we propose the following three-step procedure, shown schematically in Fig. 1, for finding the optimal clustering from the correlation matrix M_{ij}:

1. We successively remove the weakest bonds of the graph.

2. When the removal of a bond results in two parts of the graph becoming disconnected, we store the new clustering. This clustering corresponds to a set of clusters, where a cluster contains sites that are connected to each other.

3. Repeating steps 1 and 2 appropriately, we eventually end up with a completely disconnected graph and have stored a sequence of different clusterings. For each of these stored clusterings we then compute the *modularity*:

$$Q = \frac{1}{2m} \sum_{ij} \left(M_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j), \quad (4)$$

where $k_i = \sum_j M_{ij}$ and $m = \frac{1}{2} \sum_{ij} M_{ij}$. The delta function $\delta(c_i, c_j)$ is 1 if sites i and j are connected for the given clustering and 0 if they are not. The modularity takes values $Q \in [-1/2, 1]$ and quantifies how good the clustering is, with close to 1 corresponding to a good clustering, or "community structure." We select the correct clustering as the one with the highest modularity.

The first step differs from the original Girvan-Newman approach. While in our case we are guided by the physical intuition that two correlation clusters are connected by only weak bonds, Girvan and Newman used a quantity called *edge betweenness* to assess which bonds are most likely to link separated communities [46].

III. CORRELATION CLUSTERS IN EIGENSTATES

We will now focus on the clustering in midspectrum eigenstates for the Hamiltonian (1) for different values of the disorder strength. We analyze the structure of the optimal clustering for eigenstates across the MBL-ergodic phase
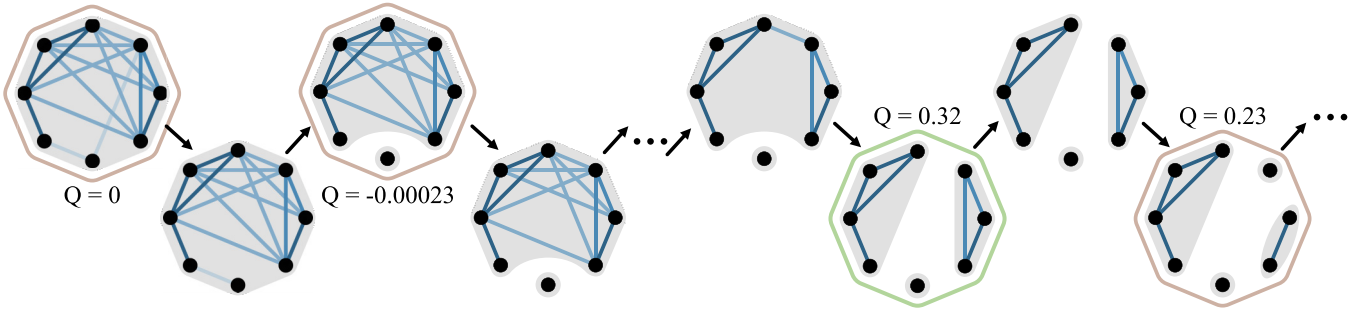
FIG. 1. Schematic description of our graph theory approach using an example of the mutual information matrix M_{ij}. The full graph has zero modularity [see Eq. (4)]. We then successively remove the weakest edges until the graph is broken into a larger number of clusters. These steps where a new clustering is obtained are saved and are indicated by a ring around them, and the value of their modularity is given. In this example we find a very low modularity of $Q = -0.00023$ for the first clustering, indicating no community structure. The next new clustering occurs after removing several bonds (not all shown) and has a relatively high modularity of 0.32, which turns out to be the highest obtained for this example. We identify this clustering as the physical one, and it is indicated by a green ring. The next steps of the decomposition yield four clusters with a modularity of $Q = 0.23$, smaller than in the last step, indicating the community structure of a lower quality. The rest of the procedure was not represented here, but the modularity decreased at each new step.

transition, using the TSMI to define the graph M_{ij}. Earlier studies showed that the number of entanglement clusters can be taken as a relevant scaling parameter for the MBL transition [41]. In order to validate our graph clustering approach, we perform a similar scaling analysis. The average number of correlation clusters n as a function of disorder is shown in Fig. 2 for different system sizes. We select 50 eigenstates from the middle of the spectrum of 700 disorder configurations and then apply the algorithm outlined in the previous section to extract the average number of clusters in the optimal clustering. The critical disorder W_c is located at the crossing of the curves at $W = 3.8$, in agreement with Ref. [41] (see the inset in Fig. 2). The data collapse convincingly with scaling
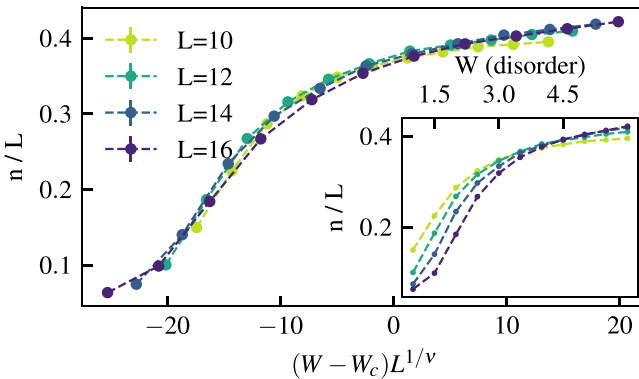

FIG. 2. Scaling collapse of the average number of clusters n divided by system size as a function of the disorder obtained using our modified Girvan-Newman approach. For the collapse, the disorder strength was rescaled to take the form $(W - W_c)L^{1/\nu}$, with $W_c = 3.8$ and $\nu = 1.26$. When the modularity was lower than $Q_{\text{th}} = \alpha(1 - a/L)$, with $\alpha = 0.3$ and $a = 3.59$, the state was considered fully ergodic and was made of only one cluster. The coefficient a has been fitted to take into account the finite-size effects at $W = 6$, according to the finite-size scaling: $Q_L = q - \frac{a}{L}$ of the average modularity Q_L obtained for system size L. Inset: average number of clusters n divided by system size as a function of disorder strength W.

$n/L = f[(W - W_c)L^{1/\nu}]$, with the parameter $\nu = 1.26$, taken from Ref. [41]. It was pointed out that this scaling is consistent with theoretical studies in which a Harris-type bound on the exponents has been derived [56]. We note that if the system is ergodic, then we would expect the mutual information to be uniform on average between all pairs of sites [49]. In this case, the optimal clustering is a single cluster containing all sites, but the algorithm as defined will instead choose a clustering with very low modularity. We therefore need to bypass the graph theory algorithm by setting a threshold Q_{th} below which the states yielding modularity $Q < Q_{\text{th}}$ are considered ergodic.

Since our numerics are performed on finite system sizes up to $L = 16$, the modularity will be affected by finite-size effects that we must take into account in Q_{th}. To understand these effects we consider states deep in the MBL phase where we can make statements about the optimal clustering. In particular, MBL eigenstates are simultaneous eigenstates of an extensive number of exponentially localized l bits with a characteristic localization length [4]. This means that the structure of the clustering should be independent of systems size, as long as it is sufficiently large compared to the localization length. As explained in Appendix E, this actually results in a system size dependence of the modularity for similar clusterings. To account for this we use the system size dependent threshold

$$Q_{\text{th}}(L) = \alpha\left(1 - \frac{a}{L}\right),\tag{5}$$

where $\alpha \in [0, 1]$. In practice we obtain the coefficient a by fitting $Q(W = 6, L)$, where $W = 6$ is the maximum disorder strength considered in our scaling analysis and is located deep within the MBL phase. In the main text, we present results for the overall cutoff parameter $\alpha = 0.3$. We show in Appendix A that as long as α gives the correct clustering behavior deep in the MBL and ergodic phases, the scaling collapse is not sensitive to the specific choice of this coefficient.

After focusing on the average number of clusters, we will now investigate the structure of individual eigenstates using the clustering algorithms. The TSMI matrix M_{ij} is shown in Fig. 3 for a single midspectrum eigenstate in an $L = 50$
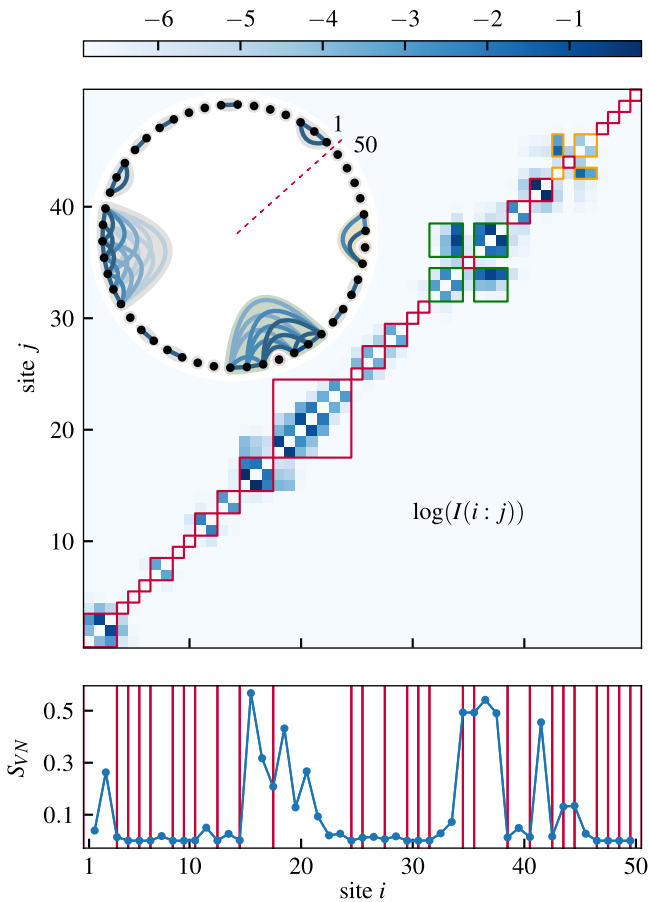
FIG. 3. Example of the mutual information matrix and the associated communities (correlation clusters) of an eigenstate of a MBL Hamiltonian obtained using the DMRG-X algorithm [30]. The disorder strength is $W = 12$, and the system size is $L = 50$. In the top panel, we plot the mutual information matrix. We draw boxes around the matrix elements belonging to the same "correlation cluster." We use a red box when a cluster is connected (i.e., no leapfrogging), while we use orange and green boxes for the two disconnected clusters. In the bottom panel, we present the bipartite entanglement entropy as a function of sites. The boundary between two clusters is signaled by a vertical red line. In the inset we show the graph corresponding to the optimal clustering. The dashed red line separates the first and last sites of the chain.

system with disorder strength $W = 12$, obtained using density matrix renormalization group excited (DMRG-X) [30], and is compared to the bipartite von Neumann entanglement entropy for cuts along different bonds. Here we can see that the localized state is decomposed into a sequence of small clusters (red boxes) and there are only weak off-diagonal (long-range) correlations in the matrix. However, we observe several examples of clusters that contain sites that are not nearest neighbors, a phenomenon which, following Ref. [41], we refer to as "leapfrogging" (green and yellow boxes). Ideally, we would like to be able to average over many eigenstates obtained by MPS methods on the MBL side of the transition and to therefore extrapolate its scaling. However, given the current state of algorithms, we find this goal impossible to achieve due to the bias in the sampling of the states [57].

A few comments are in order: First, the clustering algorithm is a numerically very inexpensive procedure which is easily scalable since only two-site correlations need to be computed, allowing us to apply it to states in the MPS form.

Second, there is clear agreement between the strong correlations and the increase in entanglement, as can be seen by comparing the TSMI matrix with the bipartite von Neumann entanglement entropy (see Fig. 3). Indeed, two *local* communities are, in general, separated by a local minimum of bipartite entanglement entropy, but not all local minima of entanglement entropy signal a separation between two communities, which is the case, for example, between sites 19 and 20 in Fig. 3. Moreover, entanglement entropy is unable to detect nonlocal clusters, i.e., leapfrogging, which we detect with our graph theory approach, for example, at site 36 in Fig. 3. Therefore, our approach give us insights into the structure of the state different from the one provided by the bipartite entanglement entropy alone. This brings us to our third point, namely, that our approach does not rest on *a priori* physical assumptions, such as the locality of the clusters. for example. Indeed, the graph theory algorithm does not know about the spatial arrangements of the sites since its only input is the TSMI matrix. However, we note that in all the cases we considered, the clusters were still relatively local and did not extend throughout the system, in accordance with the results of Ref. [41].

IV. NONEQUILIBRIUM DYNAMICS

We now turn to the behavior under nonequilibrium dynamics in the localized phase and compare AL and MBL systems. We now consider a global quantum quench protocol, starting from an initial Néel state $|\cdots \uparrow\downarrow\uparrow\downarrow \cdots\rangle$, and time evolve it using the Hamiltonian (1) with $\Delta = 1$ (MBL) or $\Delta = 0$ (AL). We can then analyze the correlations as a function of time and identify the time dependence of the correlation clusters. We compare results obtained using the TSMI, $M_{ij} = I(i:j)$, and the correlation functions, $M_{ij} = C(\hat{\sigma}_i^z, \hat{\sigma}_j^z)$.

Figures 4(a) and 4(b) show the numerical results for the average cluster length l as a function of time. When using $M_{ij} = I(i:j)$, l stays approximately constant throughout time, both in the interacting and noninteracting cases. In contrast, when using $M_{ij} = C(\hat{\sigma}_i^z, \hat{\sigma}_j^z)$, l decreases in the MBL case, while it stays constant in the AL case.

In order to understand these results better and to be able to distinguish MBL from AL further using graph theory, we show the numerical results for the average modularity as a function of time in Figs. 4(b) and 4(d). The offset of the modularity has been shifted so that the values for different system sizes coincide at short times. Indeed, Appendix E shows that the modularity scales with the system size as $Q \sim q - aL^{-1}$ for comparable clusters. The value of a is found by fitting the data at short times, and we find it to be roughly the same for both AL and MBL. In the noninteracting case, the modularity Q stays constant throughout the time evolution. On the contrary, Q decreases in the interacting case.

These observations can be explained as follows: at very short times (of the order of $\frac{1}{J}$), correlation clusters appear similarly for both AL and MBL. Due to dephasing in the MBL case, these clusters interact exponentially slowly with separation between them, leading to a slow decrease of the
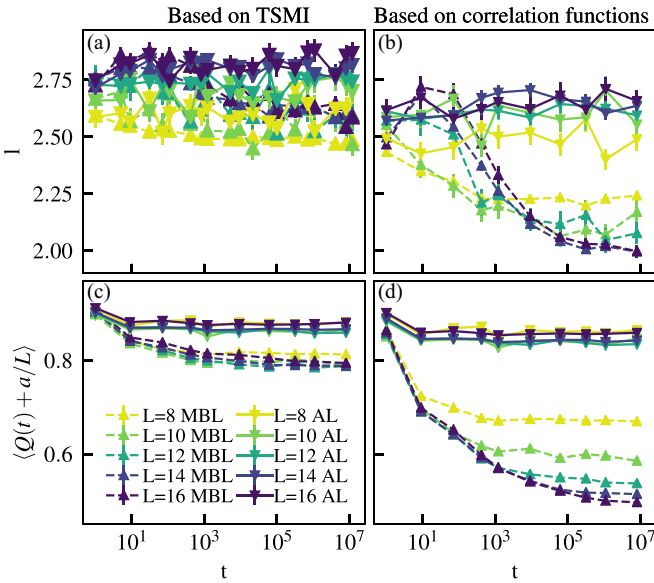
FIG. 4. (a) and (b) Average length of the clusters and (c) and (d) average modularity as a function of time for different system sizes, with disorder strength $W = 8$, for both an Anderson localized Hamiltonian (AL; $\Delta = 0$) and MBL Hamiltonian ($\Delta = 1$). We start from a Néel state, simulate a quench using exact time evolution, and apply our graph theory approach to the TSMI matrix (left panels) and to the pairwise correlation functions in the σ_z basis (right panels). The fitting parameters are as follows: in (c) $a = 3.69$ (MBL) and $a = 3.63$ (AL) and in (d) $a = 3.76$ (MBL) and $a = 3.77$ (AL).

modularity until it reaches a minimum set by the system size. Over time, the correlations induced by this long-range dephasing will build up until they are comparable to the weakest correlations within a given cluster. At this point, those most weakly correlated sites in the cluster will be excluded in favor of forming a stronger, smaller cluster, as can be seen in the example presented in Fig. 8 in Appendix D. This leads to a decrease of the average cluster length as a function of time in the interacting case, which is more pronounced with the pairwise correlations in the σ_z basis. This is consistent with the fact that at early times, the average lengths of the clusters are identical for MBL and AL, while at later times, due to the dephasing of the l bits, the average length becomes smaller for MBL. Nonetheless, due to the presence of l bits in the MBL system, for the system sizes and timescales accessible to us, these clusters are robust, as the modularity stays relatively high and the clustering at long times is still reminiscent of the structure found at early times.

The effects induced by the dephasing of the l bits are more pronounced with the correlation functions than for the TSMI. This difference in behavior stems from the lack of transport in MBL [58], which implies that off-diagonal correlation functions cannot build up beyond the localization length. Thus, for our charge-conserving model, only the σ_z component contributes to the growth of the TSMI at long times. Numerical evidence for the spreading of correlation functions in the σ_z basis is presented in Appendix F. As a consequence, when using the TSMI, the information contained in the σ_z correlation functions is washed out by all the other correlations, which necessarily decay for sufficiently large distances. This leads

to more robust clusters which interact less strongly with each other. This is in line with findings in previous works [59–61], which showed that quantities based on these correlators, in particular certain types of quantum Fisher information, can probe the logarithmic growth of entanglement in MBL systems.

These findings show that it is advantageous to consider the σ_z component in this context. In particular, the diagonal σ_z correlations are accessible in existing quantum gas microscope experiments [27,62,63], and thus, our technique can be directly applied in such settings.

V. CONCLUSION

In this work, we have shown how to efficiently investigate the structure of MBL states using pairwise correlation functions and TSMI. We focused on two applications. First, we provided numerical techniques for probing the structure of MBL eigenstates, scalable to large systems, which are particularly relevant for states obtained by MPS methods. Second, we showed that our approach can provide a characterization of dynamics in the MBL phase. We have demonstrated that our clustering procedure yields results physically consistent with previously known results. When looking at the eigenstates, the scaling of the length of the clusters found in previous works [41] was recovered. When looking at the dynamics, our results were consistent with the dephasing process between distant l bits which is observed in other quantities such as the entanglement entropy or the quantum Fisher information. Moreover, we found that the quality of the clustering at late times was still high, a fact which underlines the relevance of the clustering in the time evolution of MBL systems, for it is the persistence of these relatively well separated clusters which prevents full thermalization of the state and keeps the saturation of entanglement entropy well below the Page value.

More broadly, we have demonstrated the possibility of probing the structure of quantum states based solely on pairwise correlations using a graph theory approach. Our approach is well suited to experiments, where correlations in the diagonal σ_z basis are typically easy to measure. Furthermore, our approach is agnostic to the structure or dimensionality of the underlying Hamiltonian or dynamics. It would therefore be exciting to test this approach in the task of distinguishing MBL systems from Anderson localized systems in experiments in both one- and two-dimensional setups.
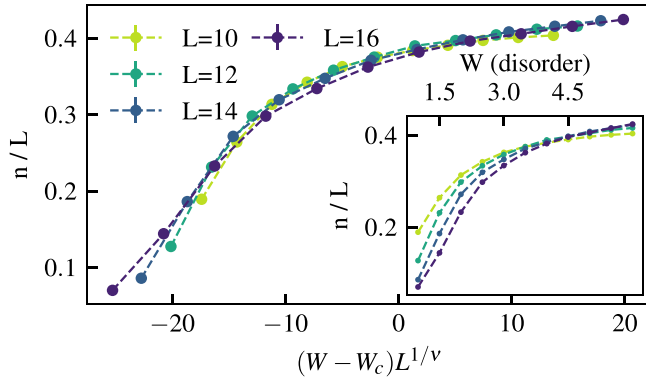
FIG. 5. Scaling collapse of the average number of clusters n divided by system size as a function of the disorder obtained using our modified Girvan-Newman approach and $\alpha = 0.25$. The parameters used for the scaling collapse are the same as in the main text. Inset: average number of clusters n divided by system size as a function of disorder strength W.

APPENDIX A: SCALING WITH DIFFERENT MODULARITY THRESHOLDS

In the main text, we presented the scaling collapse of the number of clusters divided by system size. When the modularity obtained for one clustering is smaller than $Q_{th}(L) = \alpha(1 - a/L)$, we bypass our algorithm and consider that the state is fully ergodic and therefore made of a single cluster. In Figs. 5 and 6, we show that the scaling collapse is not sensitive to the value of the coefficient α, as long as α is such that the modularity of almost all eigenstates deep in the ergodic (MBL) phase is below (above) $Q_{th}(L)$.

APPENDIX B: SCALING COLLAPSE OF THE NUMBER OF CLUSTERS USING THE PAIRWISE CORRELATION FUNCTIONS IN THE σ^z BASIS

In the main text, we presented in Fig. 2 a scaling collapse of the average number of clusters divided by system size for

FIG. 6. Scaling collapse of the average number of clusters n divided by system size as a function of the disorder obtained using our modified Girvan-Newman approach and $\alpha = 0.4$. The parameters used for the scaling collapse are the same as in the main text. Inset: average number of clusters n divided by system size as a function of disorder strength W.
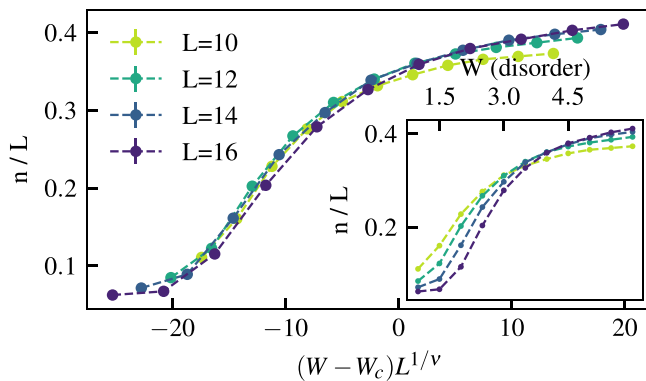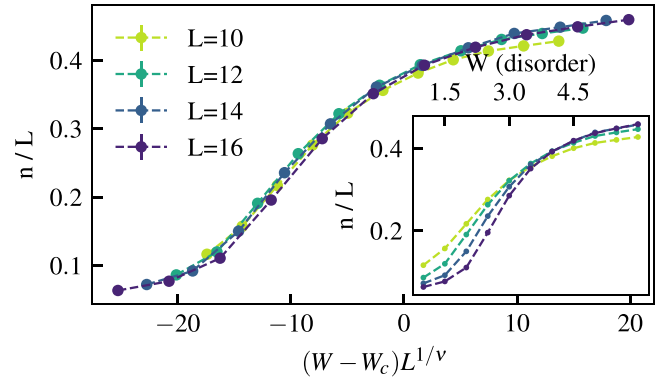

FIG. 7. Scaling collapse of the average number of clusters n divided by system size as a function of the disorder obtained using our modified Girvan-Newman approach applied to the pairwise correlation functions in the σ^z basis, $\alpha = 0.15$ and $a = 3.01$. The parameters used for the scaling collapse are the same as in the main text. Inset: average number of clusters n divided by system size as a function of disorder strength W.

which we used the TSMI of the eigenstates as the adjacency matrix in our graph theory approach. We show in Fig. 7 that the same approach using the pairwise correlation functions in the σ^z basis yields the same scaling collapse. In order to ensure that all states deep in the MBL phase are identified as such, we need to choose a smaller coefficient α for Q_{th} [see Eq. (5)]. Here we choose $\alpha = 0.15$.

APPENDIX C: DETAILS ABOUT THE MODULARITY

We would like to have a quantity that enables us to judge the quality of a graph clustering in order to compare different clusterings. In order to achieve this, a naive trial would be to evaluate the fraction of weighted edges connecting vertices belonging to the same community over the sum of all weights. If M is the adjacency matrix and c_i is the community to which vertex i belongs, this ratio can be expressed as

$$R = \frac{1}{2m} \sum_{ij} M_{ij} \delta(c_i, c_j), \qquad (C1)$$

where $m = \frac{1}{2} \sum_{ij} M_{ij}$ and $\delta(c_i, c_j)$ ensures that vertices i and j are in the same community. Although any good clustering of a given graph should yield a high value of R, this quantity is not useful for gathering information about the community structure. Indeed, considering only one community containing all the vertices would result in a maximal value of $R = 1$ [45]. In order to overcome this limitation, the formula for R was modified by following the idea that a random graph is not expected to present a community structure. A good measure of quality would be obtained when comparing the fraction of weights belonging to the same community to the one we would have if the weights had been assigned randomly. This translates to the following expression for the *modularity* [45,47,48] of a graph partition:

$$Q = \frac{1}{2m} \sum_{ij} (M_{ij} - P_{ij}) \delta(c_i, c_j), \qquad (C2)$$
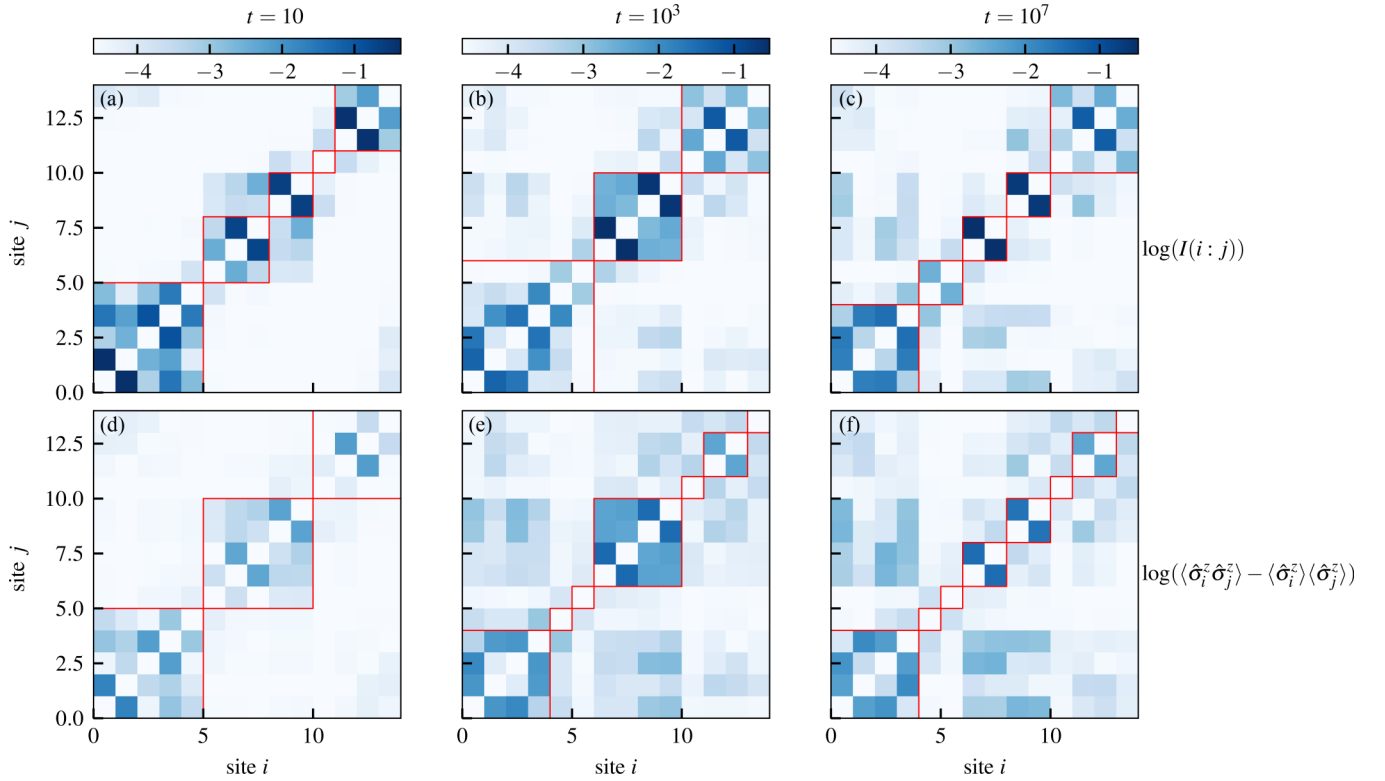
FIG. 8. Cluster decomposition for different times obtained for an initial Néel state, time evolved with the Hamiltonian (1) with periodic boundary conditions for disorder strength $W = 8$ and $L = 14$. (a) and (d) $t = 10$, (b) and (e) $t = 10^3$, and (c) and (f) $t = 10^7$. (a)–(c) $M_{ij} = I(i : j)$. (d)–(f) $M_{ij} = C(\hat{\sigma}_i^z, \hat{\sigma}_j^z)$.

where P_{ij} is the expected adjacency matrix of the random graph which has the same structural properties as the original graph of interest without presenting the same community structure. This random graph is also sometimes called the "null model." In order to determine the matrix P, we must first specify a choice for the null model. Since it has to be similar to the original graph, we impose that the vertex of the random graph has to have the same *degrees* k_i as the original one, that is to say,

$$\sum_j M_{ij} = \sum_j P_{ij} = k_i. \tag{C3}$$

In other words, every vertex of the null model shares as much weight with the rest of thesystem as the graph of interest, although the connections between vertices are assigned randomly. On average, vertices i and j will be connected by an edge of weight $P_{ij} = \frac{k_i k_j}{2m}$ [48], yielding [47]

$$Q = \frac{1}{2m} \sum_{ij} \left(M_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j). \tag{C4}$$

We can see that this measure solves the issue initially encountered since the partition containing all vertices has zero modularity. A value of modularity close to zero means that the partition is not better than a random one, while a value close to 1 indicates a strong community structure.

APPENDIX D: EXAMPLE OF THE EVOLUTION OF THE CLUSTERING IN THE DYNAMICAL CASE

Figure 8 shows the evolution of the clustering for $t = 10$, $t = 10^3$, and $t = 10^7$. At short times, correlations start to build up locally, resulting in the formation of three large clusters. At intermediary times these blocks start to break up as the correlations become more nonlocal. Intercluster correlations (corresponding to "off-diagonal" elements on the correlation
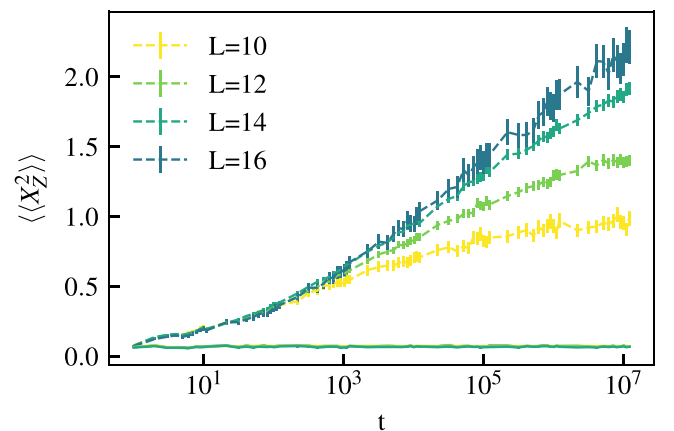

FIG. 9. $\langle\langle X_Z^2 \rangle\rangle$ for $W = 8$ for various system sizes. The MBL case is shown by dashed lines, while the AL case is plotted by solid lines. All the curves for the AL case are superimposed. This demonstrates that the pairwise correlations in the σ_z basis are sufficient to probe the logarithmic propagation of information.

matrix) are more important, resulting in a decrease in modularity. At long times, this process continues to unfold, with a further fragmentation of the cluster structure. However we note that, despite longer-range correlations, a clear cluster structure is present, and the correlations are not completely scrambled. Moreover, the intercluster interactions are more pronounced in the case of the correlation functions than in the case of the TSMI.

APPENDIX E: SCALING OF THE MODULARITY WITH SYSTEM SIZE

For a system where the optimal decomposition yields N clusters, the modularity can be written in the following way [65]:

$$Q = \sum_{i=1}^{N} \frac{e_i}{m} - \left(\frac{d_i}{2m}\right)^2, \qquad (E1)$$

where the sum runs over the clusters. In the formula above, d_i denotes the total degree of nodes in the cluster i: $d_i = \sum_j k_j \delta(c_j, c_i)$ in the notation of the main text, e_i is the number of edges in cluster i, and m is, as in the main text, the total number of edges. Defining $\langle e \rangle = \frac{1}{N} \sum_i e_i$ and $\langle d \rangle = \frac{1}{N} \sum_i d_i$, we obtain

$$Q = \sum_{i=1}^{N} \frac{\langle e \rangle}{m} - \frac{\langle d^2 \rangle}{(2m)^2}. \qquad (E2)$$

We now introduce the quantity $\langle e^{\text{out}} \rangle$, which is the average weight leaving each cluster:

$$\langle e^{\text{out}} \rangle = \frac{1}{N} \sum_i \sum_j M_{i,j} [1 - \delta(c_i, c_j)]. \qquad (E3)$$

Using the fact that $\langle d \rangle = \langle 2e \rangle + \langle e^{\text{out}} \rangle$ and $m = \frac{1}{2} N \langle d \rangle$, we obtain

$$Q = \sum_{i=1}^{N} \frac{\langle e \rangle}{\frac{N}{2}(2\langle e \rangle + \langle e^{\text{out}} \rangle)} - \frac{\langle (2e + e^{\text{out}})^2 \rangle}{N^2 (2\langle e \rangle + \langle e^{\text{out}} \rangle)^2}. \qquad (E4)$$

Finally, noting that the number of clusters N is proportional to the system size, we recover the scaling of the main text:

$$Q = \frac{\langle e \rangle}{\langle e \rangle + \langle e^{\text{out}} \rangle / 2} - \frac{1}{N} \frac{\langle (2e + e^{\text{out}})^2 \rangle}{(2\langle e \rangle + \langle e^{\text{out}} \rangle)^2}. \qquad (E5)$$

APPENDIX F: INFORMATION PROPAGATION USING PAIRWISE CORRELATION IN THE σ_z BASIS

Reference [49] showed that one can use the TSMI to detect the MBL phase. More precisely, one has to monitor the following quantity during a global quench:

$$\langle \langle X_I^2 \rangle \rangle(t) = \sum_j j^2 I_j(t) - \left(\sum_j j I_j(t) \right)^2, \qquad (F1)$$

where $I_j(t) = I(0; j)(t)$. The MBL phase is characterized by logarithmic growth of $\langle \langle X_I^2 \rangle \rangle$ since this quantity measures the spreading of information in the system. This is explained by the fact that two separate portions of the system need a time exponential with their distance to get entangled.

To demonstrate this, we perform exact time evolution with the Hamiltonian (1) with open boundary conditions and calculate the following disorder-averaged quantity:

$$\langle \langle X_Z^2 \rangle \rangle(t) = \sum_j j^2 C(\hat{\sigma}_0^z, \hat{\sigma}_j^z, t) - \left(\sum_j j C(\hat{\sigma}_0^z, \hat{\sigma}_j^z, t) \right)^2. \qquad (F2)$$

This quantity also exhibits logarithmic growth, as demonstrated on Fig. 9.

[1] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. **109**, 1492 (1958).

[2] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states, Ann. Phys. **321**, 1126 (2006).

[3] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Interacting Electrons in Disordered Wires: Anderson Localization and Low-t Transport, Phys. Rev. Lett. **95**, 206603 (2005).

[4] M. Serbyn, Z. Papić, and D. A. Abanin, Local Conservation Laws and the Structure of the Many-Body Localized States, Phys. Rev. Lett. **111**, 127201 (2013).

[5] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Many-Body Localization in a Disordered Quantum Ising Chain, Phys. Rev. Lett. **113**, 107204 (2014).

[6] B. Bauer and C. Nayak, Area laws in a many-body localized state and its implications for topological order, J. Stat. Mech. (2013) P09005.

[7] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43**, 2046 (1991).

[8] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[9] J. Z. Imbrie, On Many-Body Localization for Quantum Spin Chains, J. Stat. Phys. **163**, 998 (2016).

[10] V. Ros, M. Müller, and A. Scardicchio, Integrals of motion in the many-body localized phase, Nucl. Phys. B **891**, 420 (2015).

[11] J. Z. Imbrie, V. Ros, and A. Scardicchio, Local integrals of motion in many-body localized systems, Ann. Phys. (Berlin, Ger.) **529**, 1600278 (2017).

[12] J. Zhang, P. W. Hess, A. Kyprianidis, P. Becker, A. Lee, J. Smith, G. Pagano, I.-D. Potirniche, A. C. Potter, A. Vishwanath, N. Y. Yao, and C. Monroe, Observation of a discrete time crystal, Nature (London) **543**, 217 (2017).

[13] S. Choi, J. Choi, R. Landig, G. Kucsko, H. Zhou, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, V. Khemani, C. von Keyserlingk, N. Y. Yao, E. Demler, and M. D. Lukin, Observation of discrete time-crystalline order in a disordered dipolar many-body system, Nature (London) **543**, 221 (2017).

[14] G. De Chiara, S. Montangero, P. Calabrese, and R. Fazio, Entanglement entropy dynamics of Heisenberg chains, J. Stat. Mech. (2006) P03001.

[15] M. Žnidarič, T. Prosen, and P. Prelovšek, Many-body localization in the Heisenberg XXZ magnet in a random field, Phys. Rev. B **77**, 064426 (2008).

[16] J. H. Bardarson, F. Pollmann, and J. E. Moore, Unbounded Growth of Entanglement in Models of Many-Body Localization, Phys. Rev. Lett. **109**, 017202 (2012).

[17] A. Nanduri, H. Kim, and D. A. Huse, Entanglement spreading in a many-body localized system, Phys. Rev. B **90**, 064201 (2014).

[18] M. Serbyn, Z. Papić, and D. A. Abanin, Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems, Phys. Rev. Lett. **110**, 260601 (2013).

[19] D. J. Luitz, N. Laflorencie, and F. Alet, Many-body localization edge in the random-field Heisenberg chain, Phys. Rev. B **91**, 081103(R) (2015).

[20] A. Pal and D. A. Huse, Many-body localization phase transition, Phys. Rev. B **82**, 174411 (2010).

[21] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, Phys. Rev. B **75**, 155111 (2007).

[22] T. C. Berkelbach and D. R. Reichman, Conductivity of disordered quantum lattice models at infinite temperature: Many-body localization, Phys. Rev. B **81**, 224429 (2010).

[23] F. Pietracaprina, G. Parisi, A. Mariano, S. Pascazio, and A. Scardicchio, Entanglement critical length at the many-body localization transition, J. Stat. Mech. (2017) 113102.

[24] S. D. Geraedts, N. Regnault, and R. M. Nandkishore, Characterizing the many-body localization transition using the entanglement spectrum, New J. Phys. **19**, 113021 (2017).

[25] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Probing entanglement in a many-body–localized system, Science **364**, 256 (2019).

[26] B. Chiaro *et al.*, Growth and preservation of entanglement in a many-body localized system, arXiv:1910.06024.

[27] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Luschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Observation of many-body localization of interacting fermions in a quasirandom optical lattice, Science **349**, 842 (2015).

[28] J.-Y. Choi, S. Hild, J. Zeiher, P. Schauss, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, Exploring the many-body localization transition in two dimensions, Science **352**, 1547 (2016).

[29] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, Probing Slow Relaxation and Many-Body Localization in Two-Dimensional Quasiperiodic Systems, Phys. Rev. X **7**, 041047 (2017).

[30] V. Khemani, F. Pollmann, and S. L. Sondhi, Obtaining Highly Excited Eigenstates of Many-Body Localized Hamiltonians by the Density Matrix Renormalization Group Approach, Phys. Rev. Lett. **116**, 247204 (2016).

[31] X. Yu, D. Pekker, and B. K. Clark, Finding Matrix Product State Representations of Highly Excited Eigenstates of Many-Body Localized Hamiltonians, Phys. Rev. Lett. **118**, 017201 (2017).

[32] S. P. Lim and D. N. Sheng, Many-body localization and transition by density matrix renormalization group and exact diagonalization studies, Phys. Rev. B **94**, 045111 (2016).

[33] M. Serbyn, A. A. Michailidis, D. A. Abanin, and Z. Papić, Power-Law Entanglement Spectrum in Many-Body Localized Phases, Phys. Rev. Lett. **117**, 160601 (2016).

[34] A. C. Potter, R. Vasseur, and S. A. Parameswaran, Universal Properties of Many-Body Delocalization Transitions, Phys. Rev. X **5**, 031033 (2015).

[35] A. Morningstar and D. A. Huse, Renormalization-group study of the many-body localization transition in one dimension, Phys. Rev. B **99**, 224205 (2019).

[36] L. Zhang, B. Zhao, T. Devakul, and D. A. Huse, Many-body localization phase transition: A simplified strong-randomness approximate renormalization group, Phys. Rev. B **93**, 224201 (2016).

[37] A. Goremykina, R. Vasseur, and M. Serbyn, Analytically Solvable Renormalization Group for the Many-Body Localization Transition, Phys. Rev. Lett. **122**, 040601 (2019).

[38] V. Khemani, S. P. Lim, D. N. Sheng, and D. A. Huse, Critical Properties of the Many-Body Localization Transition, Phys. Rev. X **7**, 021013 (2017).

[39] F. Huveneers, Classical and quantum systems: transport due to rare events, Ann. Phys. (Berlin, Ger.) **529**, 1600384 (2017).

[40] P. T. Dumitrescu, A. Goremykina, S. A. Parameswaran, M. Serbyn, and R. Vasseur, Kosterlitz-Thouless scaling at many-body localization phase transitions, Phys. Rev. B **99**, 094205 (2019).

[41] L. Herviou, S. Bera, and J. H. Bardarson, Multiscale entanglement clusters at the many-body localization phase transition, Phys. Rev. B **99**, 134205 (2019).

[42] F. Pietracaprina, N. Macé, D. J. Luitz, and F. Alet, Shift-invert diagonalization of large many-body localizing spin chains, SciPost Phys. **5**, 045 (2018).

[43] M. Serbyn, Z. Papić, and D. A. Abanin, Criterion for Many-Body Localization-Delocalization Phase Transition, Phys. Rev. X **5**, 041047 (2015).

[44] A. De Luca and A. Scardicchio, Ergodicity breaking in a model showing many-body localization, Europhys. Lett. **101**, 37003 (2013).

[45] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E **69**, 026113 (2004).

[46] M. Girvan and M. E. J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).

[47] M. E. J. Newman, Analysis of weighted networks, Phys. Rev. E **70**, 056131 (2004).

[48] S. Fortunato, Community detection in graphs, Phys. Rep. **486**, 75 (2010).

[49] G. De Tomasi, S. Bera, J. H. Bardarson, and F. Pollmann, Quantum Mutual Information as a Probe for Many-Body Localization, Phys. Rev. Lett. **118**, 016804 (2017).

[50] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Area Laws in Quantum Systems: Mutual Information and Correlations, Phys. Rev. Lett. **100**, 070502 (2008).

[51] R. X. Dong and D. L. Zhou, Correlation function and mutual information, J. Phys. A **43**, 445302 (2010).

[52] M. A. Valdez, D. Jaschke, D. L. Vargas, and L. D. Carr, Quantifying Complexity in Quantum Phase Transitions via Mutual Information Complex Networks, Phys. Rev. Lett. **119**, 225301 (2017).

[53] A. A. Bagrov, M. Danilov, S. Brener, M. Harland, A. I. Lichtenstein, and M. I. Katsnelson, Detecting quantum critical points in the t-t' Fermi-Hubbard model via complex network theory, Sci. Rep. **10**, 20470 (2020).

[54] B. Sokolov, M. A. C. Rossi, G. García-Pérez, and S. Maniscalco, Emergent entanglement structures and self-similarity in quantum spin chains, arXiv:2007.06989.

[55] T. Szołdra, P. Sierant, K. Kottmann, M. Lewenstein, and J. Zakrzewski, Detecting ergodic bubbles at the crossover to many-body localization using neural networks, Phys. Rev. B **104**, L140202 (2021).

[56] A. Chandran, C. R. Laumann, and V. Oganesyan, Finite size scaling bounds on many-body localized phase transitions, arXiv:1509.04285.

[57] N. Pomata, S. Ganeshan, and T.-C. Wei, In search of a many-body mobility edge with matrix product states in a generalized Aubry-André model with interactions, arXiv:2012.09853.

[58] D. J. Luitz and Y. B. Lev, Absence of slow particle transport in the many-body localized phase, Phys. Rev. B **102**, 100202(R) (2020).

[59] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Many-body localization in a quantum simulator with programmable random disorder, Nat. Phys. **12**, 907 (2016).

[60] G. De Tomasi, F. Pollmann, and M. Heyl, Efficiently solving the dynamics of many-body localized systems at strong disorder, Phys. Rev. B **99**, 241114(R) (2019).

[61] A. Safavi-Naini, M. L. Wall, O. L. Acevedo, A. M. Rey, and R. M. Nandkishore, Quantum dynamics of disordered spin chains with power-law interactions, Phys. Rev. A **99**, 033610 (2019).

[62] W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice, Nature (London) **462**, 74 (2009).

[63] J. F. Sherson, C. Weitenberg, M. Endres, M. Cheneau, I. Bloch, and S. Kuhr, Single-atom-resolved fluorescence imaging of an atomic Mott insulator, Nature (London) **467**, 68 (2010).

[64] J. Hauschild and F. Pollmann, Efficient numerical simulations with Tensor Networks Tensor Network Python (TeNPy), SciPost Phys. Lect. Notes **5** (2018), code available from https://github.com/tenpy/tenpy.

[65] B. H. Good, Y.-A. de Montjoye, and A. Clauset, Performance of modularity maximization in practical contexts, Phys. Rev. E **81**, 046106 (2010).