



stances for more than five decades [4–6]. In recent years, with the advances of numerical techniques, computational capacities, and experimental facilities, the interaction effect within many-body disordered systems has become of great interest in light of the novel competition between many-body interactions and quenched disorder. One would like to thoroughly explore the effects of disorder on the fundamental properties of many-body interacting systems associated with ergodicity breaking, eigenstate thermalization, and quantum criticality. Particularly, many-body localization (MBL) is expected to be intrinsically different from Anderson localization [7–10].

In the absence of disorder, a nonintegrable interacting system is usually considered to satisfy the so-called eigenstate thermalization hypothesis (ETH) [11, 12], which assumes that the ergodicity of many-body states exists, implying that eigenstate fluctuations in the expectation of a quantum mechanical observable is about the same order of magnitude as the statistical fluctuations in a microcanonical ensemble controlled by an appropriate energy density [7]. When sufficiently strong disorder is introduced into such many-body interacting systems, an ergodic many-body state may turn into an MBL state, in which local integrals of motion (LIOMs) emerge [13–15] such that the integrability is considerably restored, resulting in the Poisson distribution of the energy levels, rather than the Wigner–Dyson surmise for the typical ETH phases as is derived from the random matrix theory [7, 16].

Phenomenology of MBL can be relevant for many different but related settings, such as a Floquet disordered system [17–20], for which the driven system is considered to implement the so-called discrete time crystal [21, 22], the quantum frustrated Heisenberg spin chains [23], and the one-dimensional (1D) QED lattice version of the Schwinger model with a disordered gauge [24], as well as in several open quantum setups [25–30]. Physical properties of MBL have been widely studied numerically for 1D Heisenberg or similarly interacting hard-core boson systems with randomness [23, 31–36]. One finds that by exploiting entanglement entropy, a dynamical transition beyond the quantum phase transition occurs between the ETH and MBL states that is driven by disorder strength when varying the energy density [23, 31–34]. While an ETH state has an extended thermal nature, an MBL state is featured by its insulating nature, being close to a product of localized states. Correspondingly, entanglement entropy possesses a volume law in the ETH phase [23] but an area law in the MBL phase. Moreover, a slowly dephasing process is unveiled in MBL phase, engendering a logarithmical entropy increase upon the time evolution of a many-body state [37, 38].

Many experimental efforts are cast into the measurements of disorder effects for interacting many-body

systems [21, 22, 36, 39–42]. Specifically, Lukin *et al.* [39] succeeded in realizing a 1D interacting Aubry–André model for bosons. By probing particle number entropy and configuration entropy, they detected that the signature of quantum thermalization, the particle’s finite localization length, and an area-law scaling of the particle number entropy can help to distinguish the MBL phase from the ergodic ETH phase. Moreover, they also found a slowly growing behavior of the configuration entropy that ultimately results in a volume-law scaling, demonstrating that the MBL state is qualitatively different from a noninteracting Anderson insulating state. Curiously, the key difference between the Bose model and the Fermi (or spin) model is that Bose statistics can result in the clustering of multiple particles, potentially giving rise to the novel phenomena of Bose type MBL. This distinctive character is missing in Refs. [39, 43].

In this paper, we extensively explore the nature of MBL phase in 1D Bose–Hubbard model with disorder by exact diagonalization (ED) [44] and several complementary techniques [45, 46]. For the present system, numerical calculations are limited to small sizes since the occupancy number of bosons can be any integer, depending on the filling factor under consideration. However, since Hamiltonian (1) below commutes with total number of particles, the ED calculation can be performed in the subspace with a given number of particles. For a chain with  $L$  sites and  $N$  bosons, the dimension for the matrix diagonalization is  $(L + N - 1)!/[N!(L - 1)!]$  [44]. By implementing the separation of von Neumann entanglement entropy, one can further compute the particle number entropy and configuration entropy to access the so-called symmetry-resolved entropy. In our numerical calculations, we exclusively focus on the filling factor  $1/2$ . Considering the rapid growth of the Hilbert-space dimension, it is practically feasible that 10000 samples were taken for the averages with  $L = 8, 5000$  with  $L = 10, 3000$  with  $L = 12$ , and for  $L = 14, 200$  samples are used for weak disorder, 400 for medium disorder, and 600 for strong disorder. We impose open boundary conditions in the calculations to compute the two-body density-density correlation function and periodic boundary conditions for other quantities. To access larger system sizes, we also employ time-evolving block decimation (TEBD) [45, 46] method to evaluate the time evolution of the two-body correlation function. Errors resulting from sample averages are given explicitly for most cases and are invisible otherwise.

In the following, we first introduce the disordered Bose–Hubbard (dBH) model and briefly discuss current status regarding the ground-state studies on this model in Section 2. In Section 3, we focus on the phase transition from ETH to MBL; at the beginning of this section, a detailed outline of its content is provided. In Section 4, we analyze different MBL phases within dBH model through the properties of its eigenstates; this section also



starts with a detailed content overview. A summary and pertinent discussions are finally given in Section 5. For completeness, several corresponding supplementary results are included in Appendixes A–D.

## 2 Model

The 1D dBH model can be written as

$$\hat{H} = -J \sum_i (\hat{a}_i^\dagger \hat{a}_{i+1} + \text{H.c.}) + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) + \mu \sum_i \delta_i \hat{n}_i, \quad (1)$$

where  $\hat{a}_i^\dagger$  and  $\hat{a}_i$  are particle creation and annihilation operators, respectively, and  $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$  is the particle occupation number operator at site  $i$ . Disorder is introduced through a random number  $\delta_i$ , which is uniformly distributed within  $[-1, 1]$  for all sites, thus chemical potential  $\mu$  becomes a tuning parameter for the disorder strength. For convenience, we set  $J = 1$  as the energy-scale unit hereafter.

For Hamiltonian (1), we use the ED technique to find all its eigenvalues  $E_i$  with an order of  $E_1 \leq E_2 \leq \dots \leq E_D$  and  $D$  is the dimension of the Hamiltonian matrix. The energy density  $\varepsilon$  used throughout this work is then defined by

$$\varepsilon = \frac{E_i - E_1}{E_D - E_1}. \quad (2)$$

At zero temperature, the BH model, in the absence of disorder, undergoes a quantum phase transition from a superfluid phase to the Mott-insulator state with an integer filling factor as the interaction strength  $U$  increases [47]. When disorder is introduced, the nature of the ground state is anticipated to alter [48–51]. For a given diagonal disorder sample  $\{\delta_i\}$  with  $i \in [1, L]$ , one finds that the system changes from the superfluid state first to the Bose-glass state and then to the Mott-insulator state under the rise of  $U$  [49, 50]. In the case of an off-diagonal disorder configuration assigned to  $J$ , a phase diagram has also been established in terms of the interaction strength  $U$  and the randomness strength in the hopping amplitude  $J$  [49–51].

In this work, we fix  $U = 3J$  and always maintain the filling factor to be one-half per site, i.e.,  $N = \sum_i \hat{n}_i = \frac{L}{2}$ , for which one has a superfluid ground state in the case of a clean chain [52].

## 3 Eigenstate phase transition

A plethora of physical quantities have been used to numerically detect the MBL phase transition, but most are experimentally inaccessible. Among these, the von Neumann entropy  $S_{\text{vN}}$  is a common choice, yet remains

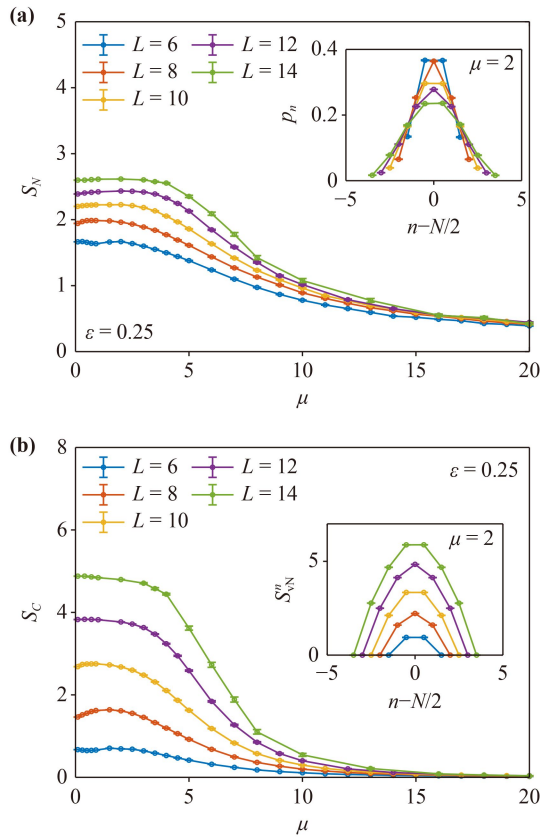
challenging to be directly measured. Conceiving the experimentally accessible probes to assess the MBL transition therefore comprises one of the main goals of the present work.

In Section 3.1, based on the  $U(1)$  symmetry, we decompose  $S_{\text{vN}}$  into the particle number entropy  $S_N$  (readily measurable) and the configuration entropy  $S_C$  (not currently measurable). We first observe that the  $S_C$  of the eigenstate nears zero in the localized phase, implying that  $S_C$  may largely be phase transition irrelevant (or at least less relevant). Motivated by this, in Section 3.2, we introduce a new quantity that quantifies the deviation of  $S_N$  from the ideal thermal distribution. Subsequent finite-size analysis shows that it shares the same phase transition point with  $S_{\text{vN}}$  but bears a greater critical exponent, indicating that it is  $S_N$  that primarily drives the phase transition. Most importantly, this deviation of  $S_N$  can be preliminarily measured in the laboratory, as has already been demonstrated by recent Harvard experiments [39, 43] (especially in boson systems). Admittedly, here we have also asserted the assumption that the energy-resolved eigenstate measurements we consider can be qualitatively executed through the careful combinations of disorder realizations and initial state preparations for using the dynamics, as was successfully implemented in the recent programmable-superconducting-processor-based experiment on MBL [41]. Recent debates about whether the MBL transition is a true phase transition or a crossover [53, 54] prompt us to seriously inspect such experimental measurability, and these considerations can contribute to the settling of these ongoing controversies (see comments at the end of this subsection). Finally, in Section 3.3, to gain an overview of the eigenstate-transition scenario, we sketch the finite-system dynamical phase diagram obtained from the large-scale numerical simulations.

### 3.1 Entanglement entropy and particle number fluctuation

In Appendix A, we detail the decomposition of the total entanglement entropy  $S_{\text{vN}}$  into the particle number entropy  $S_N$  and the configuration entropy  $S_C$  when the total number of particles  $N$  in the system is conserved and discuss how to calculate the configuration entropy in an efficient way.

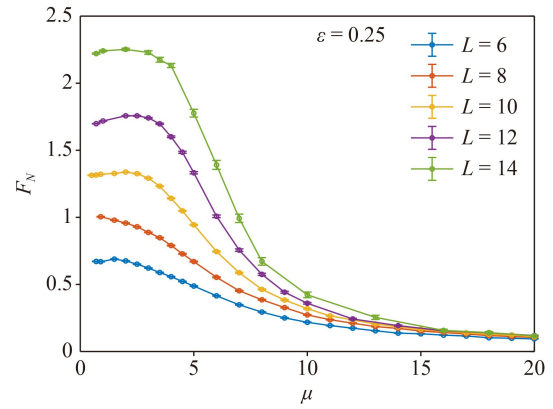
Now we mainly focus on the obtained numerical results for these two quantities, the particle number and configuration entropies. Figure 1(a) shows the size dependence of  $S_N$ . For each length  $L$ ,  $S_N$  decreases monotonically with the increasing disorder strength  $\mu$ . Likewise, for  $S_C$ , we can see from Fig. 1(b) that it lessens as disorder enhances, and approaches zero once the disorder strength  $\mu$  becomes significant. This is because if disorder is strong enough, the system is expected to be localized in several particular configura-



**Fig. 1** (a) Particle number entropy  $S_N$  versus disorder intensity  $\mu$  for different sizes  $L$  at an energy density  $\varepsilon = 0.25$ . (b) Results for configuration entropy  $S_C$  at the same energy density. Companion results for  $p_n$  and  $S_{VN}^n$  with respect to differing sizes are shown in the insets of (a) and (b) for  $\mu = 2$ .

tions. The configuration entropy due to the superposition of these configurations thereby tends to near zero. Under the circumstance of larger  $\mu$ , the particle number entropy  $S_N$  appears to dominate the evolution of the total entropy. Besides, both entropies exhibit a volume law at weak disorder and an area law at strong disorder, consistent with the general trend of the total entropy found in previous MBL studies [23, 33].

As pointed out in Appendix A,  $S_N$  and  $S_C$  can be expressed in terms of  $p_n$  and  $S_{VN}^n$ . Here,  $n$  denotes the number of particles in the left half chain, marking the chunks of the reduced density matrix which we call channels. The value of  $n$  ranges from 0 to  $N$  (there are a total of  $N + 1$  options for  $n$ ). In the inset of Fig. 1(a), we give the distribution of  $p_n$  as a function of  $n - N/2$  with respect to different chain lengths for a small disorder strength  $\mu = 2$ . It can be seen that since  $p_n$  needs to be normalized and there is no difference between tracing out the left or the right portion of the system at ETH, it is a symmetric distribution of the Gaussian lineshape. In the inset of Fig. 1(b), the results of  $S_{VN}^n$  are given, again symmetric. Overall, as the chain length increases, the dimension of the channel grows, so does the corresponding



**Fig. 2** Half-chain particle number fluctuation  $F_N$  versus  $\mu$  at an energy density  $\varepsilon = 0.25$  for various lengths  $L$ .

entropy.

$S_N$  arises from the superposition of states with varying particle numbers in half-chain subsystems, generated by particle motion across the boundary of the bipartite ( $A$  and  $B$ ). This fluctuation and its measurability can both be understood in terms of the particle number variances for part  $A$ , which is typically defined as follows:

$$\begin{aligned}
 F_N &= \left\{ \langle \hat{N}_A^2 \rangle - \langle \hat{N}_A \rangle^2 \right\}_{\text{av}} \\
 &= - \left\{ \sum_{i < \frac{L}{2}} \sum_{j > \frac{L}{2}} [\langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle] \right\}_{\text{av}}, \quad (3)
 \end{aligned}$$

where  $\hat{N}_A$  is the total particle number operator for part  $A$ .  $F_N$  accounts for the correlation between the two-particle densities on the two sites located inside portions  $A$  and  $B$ , respectively, reflecting the particle-number-fluctuation induced entanglement between the bipartite.

Figure 2 shows the results of  $F_N$  at  $\varepsilon = 0.25$ . One can see that it indeed behaves analogously to  $S_N$ . This implies that it also has a volume law for small  $\mu$  in the ETH phase and tends toward the area law at large  $\mu$  for the MBL phase. Analyzing the last expression of (3) when subject to the sufficiently large  $\mu$ , one may expect that the correlations would become exponentially small with increasing  $|i - j|$ . It turns out that in the summation of (3), those terms with small  $|i - j|$  do dominate, i.e., only the contributions of those correlations at the interface between parts  $A$  and  $B$  remain finite for the MBL phase, thereby leading to an area law for  $F_N$ . This also reinforces the experimental observation on the area-law scaling of  $S_N$  in the MBL phase as was reported in the Harvard experiment [39].

### 3.2 Number entropy dominates the MBL transition

As seen in the previous subsection, in the localized phase, the  $S_C$  of the eigenstate nears zero and consequently  $S_N$  dominates the total entropy. This implies that  $S_C$  may

be much less relevant to the phase transition and the transition is chiefly governed by  $S_N$ . Next, we will show numerical evidence to substantiate this claim.

Firstly, we calculate the total von Neumann entropy  $S_{vN}$  for a given  $\varepsilon = 0.25$  using the various sizes  $L$ , which is expected to capture the finer signal of the transition point according to the general scaling theory of the critical phenomena [55, 56]. As per the finite-size scaling theory, the phase transition point and the critical exponent thereof can be obtained when all data curves are collapsed into a single line describable by the following scaling form ansatz [23, 57],

$$Q(L, \mu) = g(L)f((\mu - \mu_c)L^{1/\nu}), \quad (4)$$

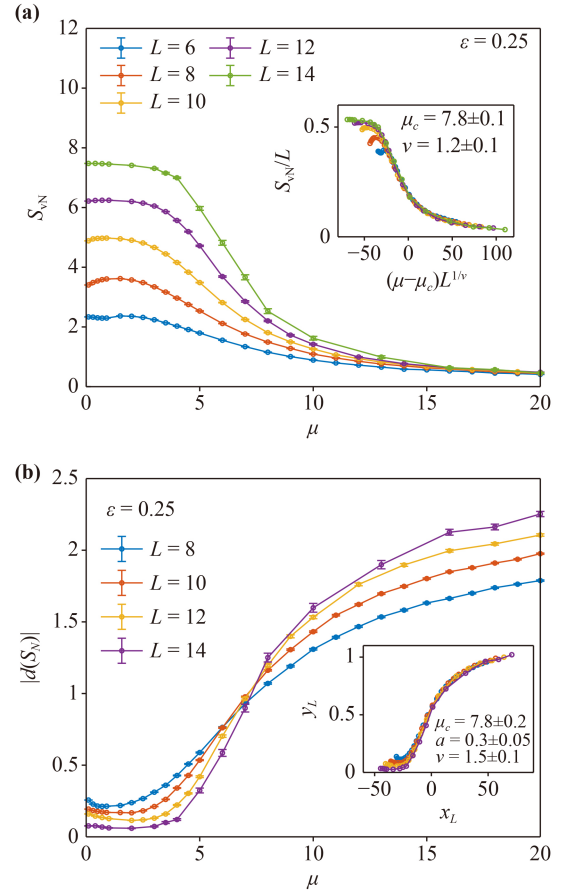
which gives rise to a collapsed scaling relation of  $y_L = Q(L, \mu)/g(L)$  versus  $x_L = (\mu - \mu_c)L^{1/\nu}$  and the critical value  $\mu_c$  as well as the critical exponent  $\nu$  for the transition.  $g(L)$  is generally a polynomial of  $L$ . Here, we set  $g(L) = L$ . In Fig. 3(a), we show a size dependence of  $S_{vN}$  at  $\varepsilon = 0.25$  for  $L = 6, 8, 10, 12, 14$ . Using the scaling relation mentioned above, we plot  $y_L = S_{vN}/L$  in the inset of Fig. 3(a), where we can collapse the data into almost a single curve except for those points at the very low  $\mu$ . We are thus able to obtain a critical value  $\mu_c \approx 7.8$  and the associated critical exponent  $\nu \approx 1.2$ , namely a phase transition point at  $\varepsilon = 0.25$ .

Secondly, for an ideally thermalized system where the system has an equal probability to access various configurations, a classical probability can be obtained by dividing the dimension of each symmetry-channel (or symmetry-sector) block by the total dimension of the Hilbert space. For the present model, the total Hilbert space of a bosonic system with the chain length  $L$  and the number of particles  $N$  is  $D_N^L = \frac{(L+N-1)!}{N!(L-1)!}$ , and the dimension of the Hilbert subspace with particle number  $n$  on the left half chain is  $D_n = D_n^{L/2} D_{N-n}^{L/2}$ , so  $p_n^{\text{ideal}} = D_n/D_N^L$ , and we can get  $S_N^{\text{ideal}} = -\sum_{n=0}^N p_n^{\text{ideal}} \log_2 p_n^{\text{ideal}}$ . To characterize the extent to which the actual probability distribution deviates from this ideal thermalization situation, we next define the following quantity so as to potentially locate the phase transition more accurately than the bare  $S_N$ ,

$$|d(S_N)| = |\{S_N\}_{\text{av}} - S_N^{\text{ideal}}|. \quad (5)$$

As shown in Fig. 3(b), we take  $g(L) = L^a$  for the case at hand, and one can see that the phase transition point thus obtained from  $|d(S_N)|$  agrees with the phase transition point obtained from the total entropy  $S_{vN}$  shown in Fig. 3(a).

More encouragingly, the critical exponent obtained from  $|d(S_N)|$ ,  $\nu \approx 1.5$ , is closer to satisfy the Harris bound ( $\nu \geq 2$ ) [58] compared to the critical exponent  $\nu \approx 1.2$  obtained from  $S_{vN}$ . Previous numeric evaluations of various physical quantities have yielded critical exponents all around  $\nu \approx 1.0$  [59], similar to that obtained from  $S_{vN}$ . This suggests that the inclusion of  $S_C$  may weaken the



**Fig. 3** (a) Variation of the total entropy  $S_{vN}$  with the disorder intensity for different sizes; the inset gives the result of its data collapse, yielding the estimate of the phase transition point at  $\mu_c \approx 7.8$ . (b) is the corresponding result of  $|d(S_N)|$  for different chain lengths when disorder increases; its data collapse is given also in the inset, showing the same phase transition point at around  $\mu_c \approx 7.8$ . The energy density here is always fixed to be  $\varepsilon = 0.25$ .

critical phenomena in general and enhance the finite-size effects in particular. This could partially lead to the pronounced drifts in the finite-size calculations, making it challenging to distinguish between a phase transition and a crossover. Furthermore, it is noteworthy that we define the deviation  $|d(S_N)|$  using  $S_N$ , and  $S_N$  itself is computed by decomposing the total entropy as per the  $U(1)$  symmetry of the system. It is this symmetry-based decomposition that allows us to separate out the phase-transition-less-relevant quantity  $S_C$ . This indicates that symmetry plays a crucial role in the MBL phase transition. Finally, the deviation  $|d(S_N)|$  we defined can be accessed directly through the  $S_N$  measurements, and recent experiments (especially in boson systems [39, 43]) have shown that  $S_N$  might be available.

Currently, there has been substantial literature debating the existence of the MBL phase and whether the MBL transition is really a phase transition or merely a

crossover [53, 54, 60]. We do not assess their conclusions. However, based on the above analyses, the devised physical quantities in these works may have the following limitations.

(i) The models they studied generally have the  $U(1)$  symmetry, but the computed physical quantities do not incorporate this symmetry thoroughly. Even worse is the case that the computational scheme they used may explicitly break those important internal symmetries of the system, such as the open-system approach used in [54] which spoils the  $U(1)$  symmetry of the Heisenberg chain. This lack of the symmetry embedding, particularly in a way parallel to the entropy decomposition already implemented here in Appendix A, may result in the dangerous inclusion of the phase-transition-irrelevant components or contributions, which probably distorts the revealing of the underlying critical phenomena.

(ii) Most quantities designed might not be experimentally measurable.

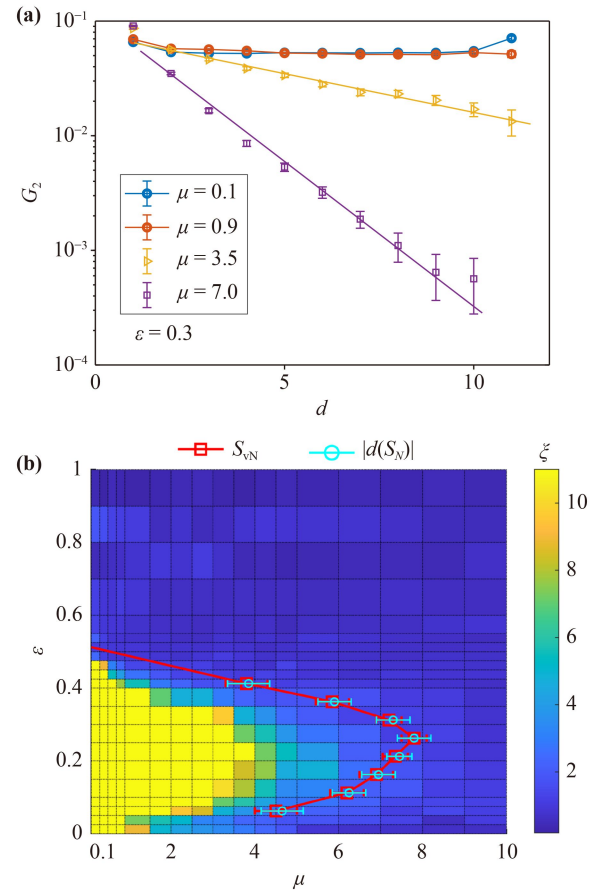
(iii) Various transport calculations require to involve the information of the whole energy spectrum. As will be illustrated shortly, the higher- and lower-energy behaviors of the generic bosonic systems can be notably distinct. This means that these regimes need to be considered and examined carefully and separately. For instance, when calculating the transport properties of a bosonic chain at infinite temperature, blindly mixing all eigenstates equally may pose a problem.

Following the recent theoretical framework proposed by Ref. [54], the observed MBL phase is likely akin to what they call the finite-size MBL regime, with the phase transition point corresponding to the landmark  $\mathcal{L}_r$  within the finite-size accuracy. In essence, while debates over the MBL phase and transition persist, we provide an experimentally practical quantity to aid the investigation of these critical issues.

### 3.3 Finite-system dynamical phase diagram

Based on the scaling analyses discussed in the previous subsection, we analogously extracted the phase transition points for a spectrum of other energy densities, ultimately producing the eigenstate phase transition diagram for the dBH model, as shown by Figs. 4(b) and 5. It can be seen that  $S_{vN}$  and  $|d(S_N)|$  yield nearly identical estimates for the phase boundaries.

Inspired by the experimental pertinence, we may use the measurable localization length to illustrate that the system's thermalization states predominantly reside in the low-energy-density region. Firstly, the two-body density-density correlation function  $G_2(d)$  is introduced to measure the correlation of the particles located at the two sites separated by a distance  $d$  in an eigenstate with the eigenenergy closest to a given energy density  $\varepsilon$ . Specifically, we borrow the definition from the Harvard experiment [39]



**Fig. 4** (a) Numerical results on the two-body density-density correlation function  $G_2$  for  $L = 12$  from the ED calculations.  $G_2$  versus  $d$  shows an exponential decay for  $\mu = 3.5$  and  $7.0$  at  $\varepsilon = 0.3$ , from which the localization lengths might be extracted. (b) Phase diagram on the plane of energy density  $\varepsilon$  and disorder strength  $\mu$  is constructed based on the color contour of the localization length  $\xi$  and consists of the ETH and MBL regimes, separated by a line derived from the scaling analyses of the total entropy  $S_{vN}$  and  $|d(S_N)|$ .

$$G_2(d) = - \left\{ \frac{\sum_{i+d \leq L} [\langle \hat{n}_i \hat{n}_{i+d} \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_{i+d} \rangle]}{N_{i+d \leq L}} \right\}_{av}, \quad (6)$$

where  $N_{i+d \leq L}$  indicates the number of pairs in the summation over  $i$  and the given  $d$ , and  $\{\dots\}_{av}$  denotes an average over the disorder samples with a given set of  $\mu$  and  $\varepsilon$ . Since the total particle number is conserved for the present dBH model, one generally has  $G_2(d) \geq 0$ . Further, one expects that  $G_2(d)$  decreases exponentially with  $d$  when the disorder strength is large enough. The localization length can thereby be extracted from this exponentially decaying behavior of  $G_2(d)$ , alternatively providing a distinguishable signal between the ETH and MBL phases.

In Fig. 4(a), the two-body density-density correlation function at  $\varepsilon = 0.3$  is shown for the four disorder strengths. In the region of small disorder strength, the

correlation function decays very slowly with the distance  $d$ . It is nearly a constant, indicating that all the system's parts are highly entangled and the many-body eigenstates are well thermalized. As the disorder strength increases, an exponential decay behavior becomes sharper and sharper so that it can be fitted by a formula of the form  $G_2(d) = Ae^{-d/\xi}$  and the localization length  $\xi$  can be extracted thereof. (Note, in comparison, the localization length cannot be gained via fitting to an exponential form when the disorder is weak and the energy density is small.) Accordingly, the emergence and the change of the localization length tunable by  $\mu$  reveal an eigenstate phase transition from the ETH state to the MBL state.

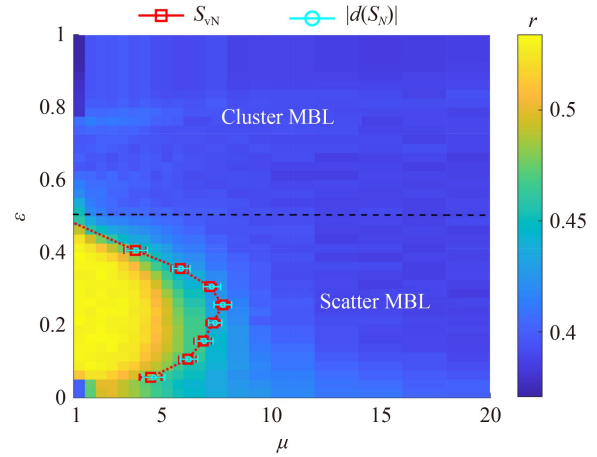
As illustrated in Fig. 4(b), by exploiting the extracted localization lengths, for the 1D dBH model with  $U/J = 3$  and the half-filling, we find an ETH phase in the bottom-left corner (yellow color) and the MBL phase (blue color) in the rest of the  $\mu$ - $\varepsilon$  parameter plane. As already stressed, the thermalized states predominantly distribute within the low-energy, weakly disordered region of the phase diagram.

Moreover, as an aside, when comparing the phase transition boundary to the color map of the localization length, one notices that the localization lengths are generally much greater than several lattice spacings in the ETH phase. Nonetheless, we would like to remind the reader the procedure that for a thin area of the very weak disorder strengths in the ETH phase, the values of these large localization lengths that cannot be determined by using the exponential fit of  $G_2$  have all been truncated to the  $L - 1$  lattice spacings, the maximal possible value of  $d$ , so that one can still represent the ETH phase in Fig. 4(b) in a coherent fashion.

To further confirm the construct of the phase diagram, it is instructive to study the complementary energy-resolved adjacent gap ratio defined by [61, 62, 63]

$$r_n = \frac{\min[\delta E_n, \delta E_{n+1}]}{\max[\delta E_n, \delta E_{n+1}]}, \quad (7)$$

where  $\delta E_n = E_n - E_{n-1}$  is the difference between the two successive eigenenergies arranged in an ascending order. For each disorder sample, we calculate the 16 adjacent  $r_n$  to obtain the averaged value  $\bar{r}$  subject to a given energy density  $\varepsilon$ , and  $r = \{\bar{r}\}_{\text{av}}$  denotes the average over the disorder samples. The ratio  $r$  reflects how the disorder strength  $\mu$  determines the nature of the level statistics at a specific energy density  $\varepsilon$ . Figure 5 shows the landscape of  $r$  versus the energy density  $\varepsilon$  and the disorder strength  $\mu$  with  $L = 12$ . One can see that the ETH state is located in the yellow region with  $r$  approaching the Gaussian orthogonal ensemble (GOE) limit  $r_{\text{GOE}} \approx 0.536$ . By contrast, the MBL state, represented in blue, tends to reach the Poisson limit  $r_{\text{Poi.}} \approx 0.386$  [64]. The results of the phase boundaries, obtained from finite-size analyses of  $S_{\text{vN}}$  and  $|d(S_N)|$ , are also overlaid in this phase diagram for comparison. From Fig. 5, it is apparent that there



**Fig. 5** Landscape of the phase diagram from the adjacent gap ratio  $r$ , using a dBH chain of  $L = 12$ . The right color bar labels  $r$  approximately between the GOE limit  $r_{\text{GOE}} \approx 0.536$  and the Poisson limit  $r_{\text{Poi.}} \approx 0.386$ . A blue square in the bottom-left corner is an artifact that may occur from an insufficient number of eigenstates, particularly in such an area where both  $\varepsilon$  and  $\mu$  are small, i.e., essentially originating from the finite-size effect. In addition, schematics of the different MBL regions are indicated: scatter MBL and cluster MBL. The red (from  $S_{\text{vN}}$ ) and cyan (from  $|d(S_N)|$ ) squared symbols with the horizontal error bars for  $\mu$  mark the position estimates of the phase boundaries at  $\varepsilon = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4$ .

appears to be an intermediate region between the yellow-region boundary and the phase-transition boundary. In Appendix B, we explore this low-energy and medium- $\mu$  region further from the perspective of the local compressibility distribution.

In the region where the energy density is less than 0.5, the overall phase transition resembles that in the full spin or fermion models [32–34]. It is anticipated that when the energy density is relatively low, it is not easy to form the high-energy situation where multiple particles occupy the same position. While, in the high-energy-density region, particles can and do form clusters due to interaction and bosonic statistics, yet possibly with very different behaviors between the limits of small and large disorder strengths.

As shown in Figs. 4(b) and 5, there is basically no significant change in the color of the localized part of the diagram. But we are now equipped to elucidate the properties and distinctions of these phase regions in some detail by scrutinizing the eigenstate properties and the quench dynamics in the high- and low-energy portions of the phase diagram. These peculiarities of the dBH model, as will be described next in Section 4, exemplify the intricate interplay of interaction, quantum statistics, and disorder, and give the schematic delineation of the phase regions in Fig. 5 as per the broad classification of scatter and cluster MBLs.

## 4 Many-body localized eigenstates

This section focuses on the properties of MBL eigenstates in the dBH chain. Our investigations reveal the following findings. i) In Section 4.1, we point out that the low-energy-density portion of the eigenspectrum of the dBH model closely resembles the entire eigenspectrum of the fermionic (or spin) models [32–34], featured by a transition from the thermalized to localized states. In view of the fact that the low-energy mobile identities are scatter particles, we name the resulting localized state the scatter MBL phase. In Section 4.1, we analyze the eigenstate particle number distribution for scatter MBL and summarize the associated quench dynamics characteristics. ii) Owing to the bosonic statistics, the high-energy-density portion of the eigenspectrum allows the formation of the distinct boson clusters in a random potential background. This situation is unique to the dBH-type model, and we name the resulting localized state the cluster MBL phase. After examining the eigenstate particle number distribution and the quench dynamics features, we show that the cluster MBL may differ in significant aspects from the scatter MBL. iii) Regarding the localized eigenstate properties, the traditional measures might not be directly sensitive to the differences between the high- and low-energy-density sections of the eigenspectrum. In Section 4.3, by exploring the structure of the symmetry-resolved entanglement entropy  $S_{\text{VN}}^n$ , we introduce another measure to help discriminate between the eigenstate phases emerging at the respective low and high energies.

### 4.1 Scatter MBL in the low-energy-density spectrum

First, we define the following two states that will be used below. Note, the subscripts in Eqs. (8) and (9) label the site indices.

(i) A line-shape initial product state, called the  $l$ -state, consists of one boson on each site of the left half chain. Pictorially, it is given by

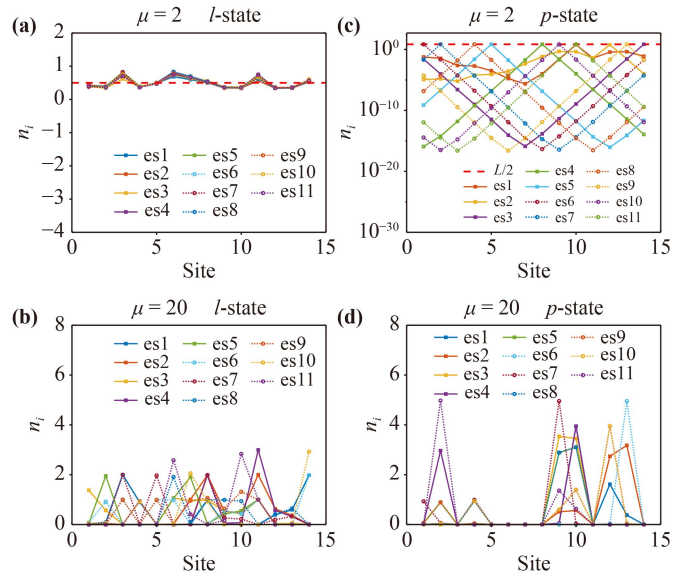
$$|l\rangle = |1_1, 1_2, \dots, 1_{\frac{L}{2}}, 0_{\frac{L}{2}+1}, \dots, 0_L\rangle. \quad (8)$$

(ii) A point-shape initial product state, called the  $p$ -state, is formed by putting all  $N$  bosons onto the leftmost site. It is given by

$$|p\rangle = |N_1, 0_2, \dots, 0_L\rangle. \quad (9)$$

Their corresponding energy densities are shown in Appendix C. From there, it is clear that as disorder increases, the  $l$ -state always stays within the low-energy-density region, while the  $p$ -state, in comparison, always stays within the high-energy-density region.

Selecting a random sample, we first compute the energy of the  $l$ -state as per  $E = \langle l|H|l\rangle_{\mu=2}$ . Next, we identify the 11 nearest eigenstates (5 above and 5 below the



**Fig. 6** Particle number distributions of the 11 eigenstates, 5 above and 5 below the eigenstate closest in energy to the given initial state. (a) corresponds to the target energy  $E = \langle l|H|l\rangle_{\mu=2}$  realized in the region of low energy density and small disorder strength. (b) corresponds to  $E = \langle l|H|l\rangle_{\mu=20}$  within the low-energy and strong-disorder region. (c) gives the result of  $E = \langle p|H|p\rangle_{\mu=2}$  for the high energy density and weak disorder. (d) targets  $E = \langle p|H|p\rangle_{\mu=20}$  upon the high energy density and large disorder strength. The chain size here is fixed to be  $L = 14$ .

target eigenstate whose eigenenergy is closest to  $E$  given above). Subsequently, we calculate the particle number distributions for these eigenstates at each lattice site. In Fig. 6(a), the particle numbers at weak disorder ( $\mu = 2$ ) are predominantly distributed around 0.5. This is reasonable because the system is in a thermalized state whose particle numbers are distributed uniformly across the lattice sites. Consider that we are dealing with a half-filled system, then the particle number should indeed be at about 0.5. In Ref. [65], we observe that for the weak disorder at  $\mu = 2$ , starting from the  $l$ -state,  $S_N$  grows with time as a function of  $\ln t$ , and its steady-state value grows with size as per  $\ln L$ . On the other hand, the companion  $S_C$  grows with time linearly, and exhibits a volume-law scaling behavior at the long-time steady-state limit.

We repeat the same calculation for the  $l$ -state but under the condition of strong disorder, i.e.,  $\mu = 20$ . As shown by Fig. 6(b), in this scenario, some lattice sites exhibit particle clustering, while others remain nearly empty, indicating the particle localization within the chain. It is worth noticing that the maximum particle occupancy per site in Fig. 6(b) remains low, and the situation where the majority of the particles occupy a single lattice site does not occur. We call this low-energy MBL phenomenon the scatter MBL. In Ref. [65], we show that for this strong-disorder condition at  $\mu = 20$ , if

starting from the  $l$ -state, the growth of  $S_N$  slows down to  $\ln \ln t$ , and the scaling of its steady-state value obeys the area law. Correspondingly, the companion  $S_C$  in this situation grows with time as per a function of  $\ln t$ , and its steady-state value fulfills a volume law.

Incidentally, with the channel-resolved analysis, we show that this slow  $\ln \ln t$  growth of  $S_N$  is mainly owing to the localization phenomenon and its final steady-state scaling remains to be an area law [66]. These observations are in contradiction with the prediction of the eventual thermalization for the similar situation in the disordered spin-chain model as was claimed in recent literature [67].

In the recent seminal experimental work on the dBH model [39], it was observed that within the MBL phase,  $S_N$  follows an area-law scaling behavior, while concurrently it was implicit that  $S_C$  follows a volume-law scaling behavior ( $S_C$  cannot be measured directly, so a closely related correlator was measured experimentally as an alternative). This suggests that the low-energy scatter MBL behavior has been experimentally detectable, but to our knowledge, the subsequent discussion regarding the high-energy cluster MBL behaviors has never been reported before.

#### 4.2 Cluster MBL in the high-energy-density spectrum

We do the same calculation by switching the initial state from the low-energy  $l$ -state to the high-energy  $p$ -state. For weak disorder ( $\mu = 2$ ), as depicted in Fig. 6(c), we observe that the 11 nearest eigenstates at the target energy  $E = \langle p|H|p \rangle_{\mu=2}$  have a common feature in their particle number distributions. Concretely, almost all the particles are concentrated to occupy a single lattice site, while the particle numbers at other sites decay exponentially. This signifies a strong degree of localization within the system; we call it the hard-type cluster MBL. This is because, in the high-energy-density regime, Bose statistics allows for the formation of the multi-particle clusters. Without disorder, these clusters can superpose to form the cat like states, but disorder breaks the translational symmetry, thus obstructing their coherent superposition and consequently giving rise to the localized, isolated clusters, ultimately resulting in the robust cluster MBL.

In Ref. [65], we find that, for the situation of low disorder ( $\mu = 2$ ), starting from the  $p$ -state, both  $S_N$  and  $S_C$  exhibit almost no growth over time, with the steady-state values following the area law. Loosely speaking, these behaviors resemble the Anderson localization in some aspects.

Similar calculations upon the  $p$ -state are performed under the strong-disorder condition ( $\mu = 20$ ) as well. It can be seen from Fig. 6(d) that, at this circumstance, the system also exhibits the distinct clustering behavior. But it is no longer as extreme as in Fig. 6(c), where nearly all particles are localized onto a single point, nor

is it close to Fig. 6(b), where particles are localized to only reach the relatively low local concentrations. Therefore, we refer to this situation as the soft-type cluster MBL.

In Ref. [65], we demonstrate that for this strong disorder at  $\mu = 20$ , starting from the  $p$ -state,  $S_N$  exhibits the diminished time-dependent growth, with the steady-state values following an area law. By contrast, the companion  $S_C$  keeps increasing slowly over time in a tentative  $\ln \ln t$  format, but its steady-state values seem to still obey the area law.

#### 4.3 Fine structure of $S_{\text{VN}}$

Curiously, is there a single quantity that can both reveal the thermalized-state distribution and differentiate between the behaviors within the low- and high-energy-density sections of the phase diagram? Through exploiting the channel reflection symmetry of the set  $\{S_{\text{VN}}^n\}$ , we introduce a new physical quantity that effectively addresses this question. Complementarily, this might provide a different means to probe the phase transition zone from the pure quantum entanglement  $\{S_{\text{VN}}^n\}$  via its intrinsic structure.

As claimed in Appendix A, for the symmetry-resolved entropy  $\{S_{\text{VN}}^n\}$ , the relation  $S_{\text{VN}}^{n_A=n} = S_{\text{VN}}^{n_B=N-n}$  always holds. In the absence of disorder, the preserved spatial reflection symmetry guarantees the general satisfaction of  $S_{\text{VN}}^{n_A=n} = S_{\text{VN}}^{n_B=N-n}$ , which generates a channel reflection symmetry of  $\{S_{\text{VN}}^n\}$ , as already illustrated in the inset of Fig. 1(b) above. It turns out that we can determine the degree of deviation from this symmetry by the summed difference of the corresponding pairs of channels, which is equivalent to determining the degree of deviation from the ideal thermalization. Thus, we try to define the pertinent physical quantity as follows:

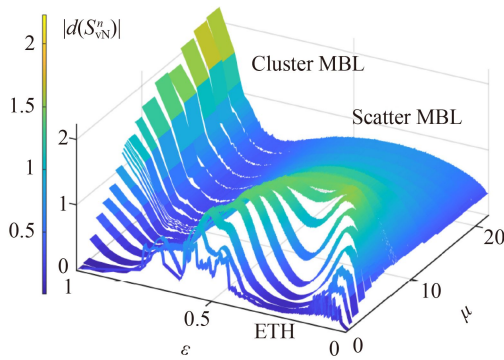
$$|d(S_{\text{VN}}^n)| = \left\{ \sum_{n=1}^M |S_{\text{VN}}^n - S_{\text{VN}}^{N-n}| \right\}_{\text{av}}, \quad (10)$$

where

$$M = \begin{cases} (N-2)/2 & \text{for } N = \text{even} \\ (N-1)/2 & \text{for } N = \text{odd} \end{cases}. \quad (11)$$

Since  $S_{\text{VN}}^{n_A=0} = S_{\text{VN}}^{n_B=N} = 0$ , the summation starts from  $n = 1$ . In the ETH region, the left and right parts are approximately symmetric, giving rise to a small value of  $|d(S_{\text{VN}}^n)|$ . In the scatter MBL region, although the relative difference is nonzero, the value of  $|d(S_{\text{VN}}^n)|$  remains small because of the absolute smallness of the values of  $\{S_{\text{VN}}^n\}$  themselves. Taken together, these trends produce a peak in the middle, corresponding to the critical regime of the underlying phase transition.

Further, from Fig. 7, we can see that the ETH region corresponds to the valley in the 3D figure of  $|d(S_{\text{VN}}^n)|$ . The closer to the bottom of the valley, the lower the degree



**Fig. 7** 3D plot of  $|d(S_{\nu N}^n)|$  at  $L = 12$ , where the phase transition boundary and the three different phases can be perceived.

of the channel-reflection-symmetry breaking is and the more thorough the thermalization of the system assumes. There are two peaks at the valley's boundaries, corresponding to the location of the energy-resolved phase transition. When the two peaks merge and disappear,  $|d(S_{\nu N}^n)|$  becomes flattened, signalling that the system begins to enter the scatter MBL region. Between the high- and low-energy-density sections, a distinctive dip effectively separates the cluster MBL at high energies and the scatter MBL at low energies. In this regard,  $|d(S_{\nu N}^n)|$  may possess the capacity to distinguish between these MBL phases and aptly identify the thermalized region at the low energies.

The above content indicates that the internal structure of the entanglement entropy effectively distinguishes between different phase regions, upon resort to the concept of symmetry decomposition. Once again, one can appreciate the importance of “symmetry combined with entanglement” in the study and elucidation of the MBL phases and transitions.

## 5 Summary and discussion

This paper provides a comprehensive study on the eigenstate properties of the dBH model in one dimension.

Firstly, we explore the transition from the thermalized state to the MBL state at low energies. The von Neumann entropy,  $S_{\nu N}$ , is often used to detect the MBL transition, even though it cannot be directly measured. Utilizing  $U(1)$  symmetry, the von Neumann entropy is decomposed into the particle number entropy,  $S_N$ , and the configurational entropy,  $S_C$ . Notably,  $S_C$  approaches zero within the localized phase. We take advantage of this trend and define  $|d(S_N)|$  to quantify the deviation of  $S_N$  from the ideal thermal distribution. Finite-size scaling analysis shows that  $|d(S_N)|$  shares the same critical point with  $S_{\nu N}$  but exhibits a larger critical exponent, suggesting that the transition is primarily set by  $S_N$  and its fluctuations. It is worth noting that  $|d(S_N)|$  might be measured via protocols already used by recent experiments, partic-

ularly in the case of bosonic systems. Recent debates about whether MBL transitions are true phase transitions or crossovers prompt us to seek such an experimentally measurable quantity that can contribute to addressing these concerns.

Secondly, the paper contrasts the distinctive behaviors occurring in the high- and low-energy-density regimes of the phase diagram for the dBH model. We find that: i) The thermalizing states predominantly occupy the low-energy spectrum, as evidenced by energy density distribution, localization lengths, and gap ratios. ii) The low-energy spectrum closely resembles the fermionic (or spin) models, where the transition from the thermalizing state to the scatter MBL phase is visible. We analyze the characteristics of these low-energy localized eigenstates, and their dynamics are also summarized. iii) Bose statistics leads to the formation of distinct clusters in the high-energy spectrum, resulting in the unique cluster MBL phase that significantly differs from the scattering MBL phase in low energies, as reflected through the eigenstate and dynamics properties. iv) Traditional measures such as localization length and gap ratios cannot directly reveal the differences in the localization phenomena between the high- and low-energy sections. We define a new quantity,  $|d(S_{\nu N}^n)|$ , from the internal structure of  $S_{\nu N}$ , which effectively captures the distinctions between phase regions in high and low energies.

The two quantities we defined based on the decomposition of entropy by symmetry,  $|d(S_N)|$  and  $|d(S_{\nu N}^n)|$ , prove to be useful in analyzing critical points and distinguishing between different phase regions. This shows that symmetry plays a key role in MBL phases and phase transitions. Currently, a grand portion of the MBL research has not given due attention to this aspect, and the physical quantities employed to analyze MBL phases and transitions have not properly embedded the symmetry information, particularly in those contentious issues.

As can be noticed, there needs a further study on whether the MBL regions are separated by phase transitions or crossovers. Moreover, the bimodal compressibility distribution structure seen in Appendix B seems to predict the existence of a critical regime arising from the Griffiths effect [42, 56, 68–70]. Different implementations of the randomized samples, e.g., quasiperiodic fields, may give rise to qualitatively different results [58, 71, 72]. Although the numerical resource to precisely settle the localization problem for the dBH model is far beyond the current computational capacity, one might still imagine to resort to the available finite-size numerics to search for the emergent new physics when the filling factor and the Hubbard interaction are tuned within the wider parameter ranges.

**Acknowledgements** We thank Zi Cai, Changfeng Chen, Shijie Hu, Haiqing Lin, Mingpu Qin, Jie Ren, and Wei Su for fruitful discussions. This work is supported by the Ministry of Science and Technology of



China (Grant No. 2016YFA0300500) and the National Natural Science Foundation of China (No. 11974244). C. C. acknowledges the support from the SJTU start-up fund and the sponsorship from Yangyang Development fund.

## Appendix A: Decompose and compute the entanglement entropy

Consider that the system consists of two parts,  $A$  and  $B$ . For convenience, we choose the system size  $L$  to contain an even number of sites, and the sites for both parts  $A$  and  $B$  are equal, i.e.,  $L_A = L_B = L/2$  in the present work. For a given many-body wavefunction  $|\psi(i_A, i_B)\rangle$ , one may construct a reduced density matrix  $\rho$  for part  $A$  by tracing out those bases of part  $B$ , which can be written formally as  $\rho(i_A, j_A) = \sum_{i_B}^{D_B} |\psi(i_A, i_B)\rangle \langle \psi(j_A, i_B)|$  where  $D_B$  is the dimension of the Hilbert space for part  $B$ , and  $i_A$  and  $i_B$  are the indices to denote the bases of parts  $A$  and  $B$ , respectively. The von Neumann entanglement entropy is thus defined for this system as  $S_{\text{vN}} = -\text{Tr} \rho \log_2 \rho$ , which is widely adapted to explore the natures of both the ETH and the MBL phases, caused by many-body interactions and disorder. One finds that  $S_{\text{vN}}$  can be written as a sum of a particle number entropy  $S_N$  and a configuration entropy  $S_C$ , which account for the entanglements of the particle number fluctuations and the correlations of particles between the two parts  $A$  and  $B$ , respectively. Although  $S_{\text{vN}}$  is not yet directly accessible in experiments, the recent Harvard experiment has successfully measured the particle number entropy  $S_N$  and the configuration correlators that are somehow equivalent to the configuration entropy  $S_C$ , and found the area law of  $S_N$  and the volume law of the configuration correlators in time evolving a state within the MBL phase [39].

For the Hamiltonian (1), since  $[\hat{H}, \hat{N}] = 0$  where  $\hat{N} = \sum_i^L \hat{n}_i$ , then  $[\rho, N_A] = 0$  where  $N_A$  is the matrix representation of  $\hat{N}_A = \sum_i^{L_A} \hat{n}_i$ . This implies that the reduced density matrix is block diagonal, which can be written as a direct summation of the form  $\rho = \bigoplus_{n=0}^N \rho^{(n)}$ . Each  $\rho^{(n)}$  has a dimension of  $D_n^A$ , representing the number of those bases having  $n$  particles in the part  $A$ . It turns out that one can naturally explore the relative entropy for each given  $n$ , i.e., the particle number entropy, if one introduces the probability  $p_n$  to sum up the amount of those states with  $n$  particles, as given by  $p_n = \text{Tr} \rho^{(n)}$ . Subsequently, one may define a proper reduced density matrix  $\rho_n = \rho^{(n)}/p_n$  for  $n$ , which obeys the compulsory normalization condition  $\text{Tr} \rho_n = 1$ . The reduced density is thus rewritten as  $\rho = \bigoplus_{n=0}^N p_n \rho_n$ .

In terms of  $p_n$  and  $\rho_n$ , the von Neumann entropy can be expressed as

$$S_{\text{vN}} = - \sum_{n=0}^N [p_n \rho_n] \log_2 [p_n \rho_n] \quad (\text{A1})$$

and can be further decoupled into a sum of the two parts  $S_N$  and  $S_C$  [39], which are given as follows:

$$S_N = - \sum_{n=0}^N p_n \log_2 p_n, \quad (\text{A2})$$

$$S_C = - \sum_{n=0}^N p_n S_{\text{vN}}^n, \quad (\text{A3})$$

where  $S_{\text{vN}}^n$  is called the symmetry-resolved entropy, and  $S_{\text{vN}}^n = - \sum_{i=1}^{D_n} \rho_n(i, i) \log_2 \rho_n(i, i)$ , where  $\rho_n(i, i)$  is the normalized eigenvalue of  $\rho^{(n)}$ .

$S_{\text{vN}}^n$  is a partial configuration entropy that measures the entanglement between the two parts  $A$  and  $B$ , resulting from the dynamics of particles' virtual superpositions and hoppings without changing  $n$ . Thus, each partial configuration entropy  $S_{\text{vN}}^n$  for that  $n$  carries a weight  $p_n$ , the sum of which forms the total configuration entropy  $S_C$ .  $S_N$ , on the other hand, stands for the entanglement of the two parts  $A$  and  $B$  associated to the particle number fluctuations encoded in the given  $|\psi(A, B)\rangle$  under the constraint that the filling factor of the system is fixed.

Next, let us discuss an optimization issue. We found that it is practically significant to construct the bases optimally for both parts  $A$  and  $B$ , which enables us to treat each  $\rho^{(n)}$  directly so as to access larger-size systems in the most economic usage of the memory and the CPU resource. For this purpose, the Hilbert space for part  $A$  is naturally assigned into the sequential subspaces, with an ascending order of the particle number from 0 to  $N$ , i.e., the  $(n+1)$ -subspace contains  $n$  particles with a dimension of  $D_n^A$ . The Hilbert space for part  $B$  is also constructed similarly for consistency.

Based on the above construction of the bases, one can consider  $|\psi(i_A, i_B)\rangle$  as a rank-2 tensor with indices  $i_A$  and  $i_B$  which denote the allocation site of each basis in the Hilbert spaces of the parts  $A$  and  $B$ , respectively, with dimensions  $D_A = \sum_{n=0}^N D_n^A$  and  $D_B = \sum_{n=0}^N D_n^B$ . Consequently, this tensor can contain rectangular sub-tensors  $|\varphi(i_A, i_B)\rangle$  with the dimensions  $D_n^A$  and  $D_{N-n}^B$ , respectively. We note that  $|\varphi(i_A, i_B)\rangle$  is null whenever  $n_A + n_B \neq N$ , since the bases are allocated in the optimal manner described above. Therefore, once the relative position of  $|\varphi(i_A, i_B)\rangle$  is settled, one can easily construct  $\rho^{(n)}$  so that  $\rho_n$  and  $p_n$  follow successively.

Since  $S_{\text{vN}}^{n_A=n} = S_{\text{vN}}^{n_B=N-n}$ , in the actual calculation, the smaller side of  $n_A$  and  $n_B$  can be chosen such that the dimension of the obtained matrix  $\rho^{(n_A)}$  or  $\rho^{(n_B)}$  is smaller and easier to be diagonalized numerically. In particular, for the two channels with  $n=0$  and  $n=N$ , there is always one side with a dimension 1, so  $S_{\text{vN}}^{n_A=0} = S_{\text{vN}}^{n_B=N} = 0$  always holds. Moreover, since the relative position and the dimension of each subspace of the parts  $A$  and  $B$  are independent of disorder, one only needs to compute  $|\varphi(i_A, i_B)\rangle$  for each sample with a given disorder strength.

It is clear that the above treatment greatly enhances the computational performance by saving both the CPU hours and the memories, especially for large chain lengths.

### Appendix B: Distributions of local compressibility and density of states

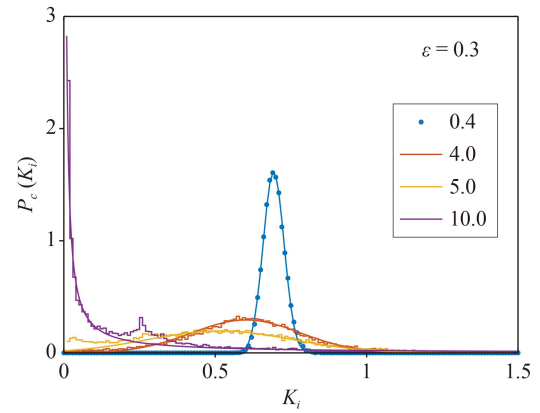
For the clean BH model, the compressibility is relevant for describing the transition between the superfluid and the Mott insulating phases. In the presence of disorder, this quantity can be appropriately replaced by the local compressibility  $K_i$  at a given energy density  $\varepsilon$  as defined by [48]

$$K_i = \{ \langle \hat{n}_i^2 \rangle - \langle \hat{n}_i \rangle^2 \}_{\text{av}}, \quad (\text{A4})$$

which essentially displays how the fluctuation of the particle number at each site is impacted by the disorder for a given  $\varepsilon$ . In Eq. (A4),  $\{ \dots \}_{\text{av}}$  stands for the average over all available disorder samples for  $\varepsilon$  and  $\mu$ . For a given  $\mu$ , the eigenvalue closest to  $\varepsilon$  may depend on the different samples. To enhance the statistical efficiency for the distribution of  $K_i$ , denoted as  $P_c(K_i)$ , instead of the closest eigenstate, we calculate  $K_i$  for the 8 eigenstates adjacent in eigenenergy to  $\varepsilon$  for each sample. An additional average is then made over these 8 eigenstates with a proper normalization for  $P_c(K_i)$ . Consequently, in view of the fact that  $K_i$  measures the variance of  $\hat{n}_i$  caused by the randomness  $\mu$  at a given  $\varepsilon$ , one anticipates that  $P_c(K_i)$  might serve as a statistics device capable of capturing the fundamentally different characteristics in the distribution of the set  $\{K_i\}$  between the ETH and the MBL phases.

Figure A1 shows  $P_c(K_i)$  for  $\mu = 0.4, 4, 5, 10$  at  $\varepsilon = 0.3$ . The distribution changes from a Gaussian distribution for  $\mu = 0.4$  and 4 to a power-law distribution in the large  $\mu$  limit. The Gaussian distribution corresponds to the well-defined superfluid states. In general, when  $\mu$  increases, the height of the curve decreases and its width gets broadened. However, as shown by  $\mu = 5$ , while the peak is greatly broadened, a warp emerges at the small  $K_i$ . The warp grows rapidly with further increasing  $\mu$ . Eventually, it behaves as a power law shown for  $\mu = 10$ , where there is still a small satellite peak at  $K_i \approx 0.25$ , which would disappear when  $\mu$  is sufficiently larger. Remarkably, one may find that a relatively sudden change of  $P_c(K_i)$  happens between  $\mu = 4$  and 5, and the disorder greatly suppresses  $K_i$  in MBL.

From the above, it is clear that in the ETH phase, the distribution is Gaussian, while in the MBL phase, it exhibits a power-law decay with a satellite peak. Notably, at  $\mu = 5$ , the right side conforms to a Gaussian distribution, one characteristic of ETH, while the left side shows an upturn near zero and a satellite peak around 0.25, indicative of the onset of MBL. This

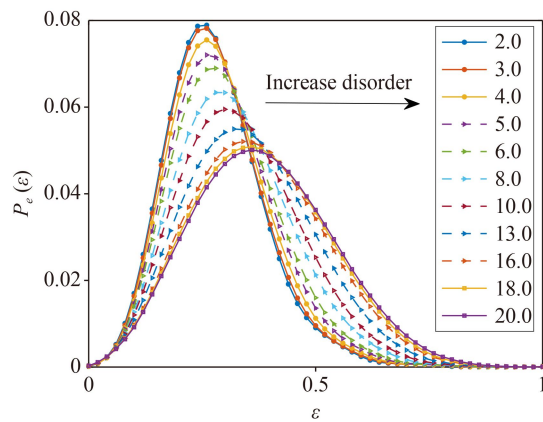


**Fig. A1** The distributions of the local compressibility  $K_i$  at  $\varepsilon = 0.3$  for several disorder strengths with  $L = 14$  and  $K_i \in [0, 1.5]$ . The histograms are obtained using ED and are accompanied by the fitting solid curves. For  $\mu = 0.4, 4.0$ , the distributions are well fitted by a Gaussian function, and a two-peak structure arises at  $\mu = 5.0$ , where the main peak can still be fitted by a Gaussian function. When  $\mu = 10.0$ , however, it fulfills a power law within the range of small  $K_i$ , meanwhile exhibiting a satellite peak at  $K_i \approx 0.25$ .

implies that the system experiences a breakdown in thermalization at this point, with a partial MBL component, but has not yet undergone a phase transition. The phase transition appears to occur around  $\mu = 7.8$ , as shown in Fig. 5. This implies the presence of an intermediate region in the system, being prior to the MBL phase transition. In both Figs. 5 and 4(b), an intermediate region is observed between the yellow thermalization region’s boundary and the phase transition boundary.

Research indicates that the distribution of thermalized states is positively correlated with the density of states (also called energy density distribution) [73], with a higher density of states corresponding to a greater presence of the thermalization states. For typical spin systems, the density of states is usually approximated by a Gaussian function [74].

Figure A2 illustrates the disorder effects on the energy density distribution  $P_c(\varepsilon)$  for a wide range of the disorder strengths. It is remarkable that for the Hamiltonian (1), its eigenstates are mainly concentrated on the range of energy densities  $\varepsilon < 0.5$  for the weak disorder, where only a tiny tail shows up for  $\varepsilon > 0.5$ . However, the distribution becomes broadened with the increasing disorder strength  $\mu$ . Its peak position starts to slowly shift toward  $\varepsilon = 0.5$  when  $\mu$  is sufficiently large, where a reasonably large tail is also developed for  $\varepsilon > 0.5$ . Indeed, under weak disorder, the thermalization states concentrate in the low-energy region, resulting in the trend of energy density distribution mainly centered around  $\varepsilon < 0.5$ . Unlike the spin model, we find that these  $P_c(\varepsilon)$  curves can be fitted by the summation of two Gaussian functions.



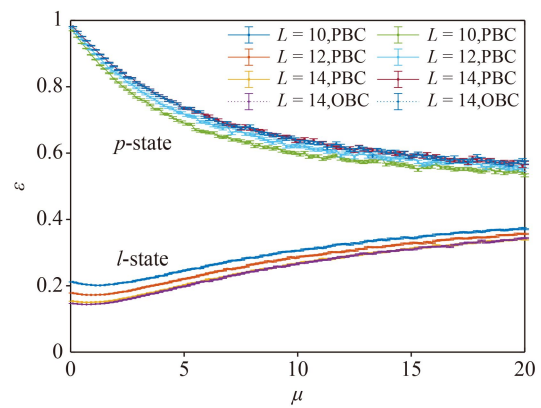
**Fig. A2** Energy density distributions for various disorder strengths  $\mu \in [2, 20]$ , obtained from ED using a chain of length  $L = 12$ .

### Appendix C: Averaged energy density of initial state

Figure A3 shows how the averaged energy densities of low-energy  $l$ -state and high-energy  $p$ -state change as a function of the disorder strength for different chain lengths. For the  $p$ -state, the energy density decreases as the disorder increases. Because the system realizes cluster MBL at small disorder, the clustering tendency enhances the Hubbard interaction, leading to the high energy densities. As the disorder increases, the degree of the clustering of particles reduces, and the effect of the interaction  $U$  term weakens; thus, the overall energy density decreases. Note that as the system size increases, the overall energy density also slightly grows. By contrast, it can be seen that the energy density of the  $l$ -state changes slowly with the increasing disorder strength and decreases with the increasing system size. Comparing to the energy density distribution shown in Fig. A2, one can see that these initial states correspond to somewhat different densities of energy eigenstates. Compared with the phase diagram Fig. 5, the energy-density trajectories of these two states evolve under the change of the disorder strength, passing across several different phase regions, which explains why they have very different time-evolutionary behaviors. Using these two kinds of initial states, we present our numerical results on the respective time evolutions of the two-body density–density correlation function and the two types of entropies.

### Appendix D: Localization length from steady state

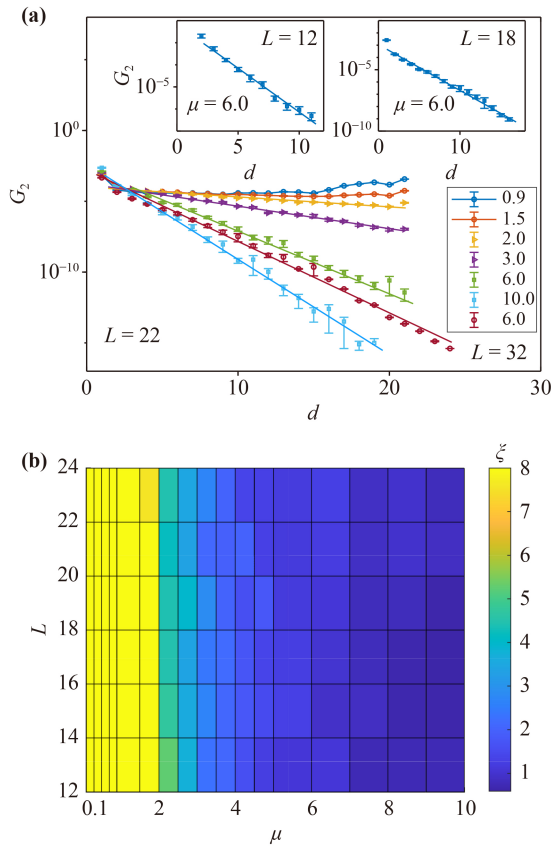
The time evolutions of  $G_2$  are calculated at  $t = 100\tau$  with  $\tau = 0.1$  [39, 69] for several system sizes  $L$ . Specifically, we perform the computations at  $L = 12$  using ED [44], and at  $L = 14, 16, 18$  using the global Krylov method [45], and



**Fig. A3** The average energy densities for the two kinds of initial states as a function of  $\mu$  for different chain lengths. In the present study, the  $p$ -state corresponds to an initial state where all  $L/2$  bosons are loaded at the first site and zero occupancy for the remainder of the system. At the same time, one particle occupies each site of the left part of the system, and no particle is in all sites of the right part, comprising the  $l$ -state.

at  $L = 20, 22, 32$  using TEBD [46], for which the 2nd-order Trotter decomposition is implemented to suppress the systematical errors. For the TEBD calculation, 600 samples are taken for the disorder averages. To study the single-particle localization length of cluster MBL, we take the high-energy  $p$ -state, i.e., all  $L/2$  bosons loaded at the first site and zero occupancy for the rest of the system, which is a high-energy-density state; as the disorder strength increases, its energy density maintains in the cluster MBL region, as can be seen from Figs. A3 and 5. Since the first site has the highest occupancy, the occupancy number cannot be truncated in the TEBD simulations. The local physical dimension on each site is thus  $L/2 + 1$ . We note that evaluating the time evolution of  $G_2$  in such a fledged Bose system without truncating the single site occupancy costs huge computational resource for moderate to large values of  $L$ .

Figure A4(a) shows the disorder effects on the time-evolved two-body density–density correlation function at  $t = 100\tau$  on a  $L = 22$  chain for several values of the disorder strength. The two insets therein display the results of  $G_2(d)$  at  $\mu = 6.0$  for  $L = 12$  and 18, respectively. Upon the weak disorder, i.e., in the ETH region,  $G_2$  has relatively small values for the moderate values of  $d$  and becomes large when  $d \sim 1$  or  $L$ . Here, the open boundary conditions are imposed. As the disorder strength gets further increased,  $G_2(d)$  starts to display an exponential decrease with respect to  $d$ . This trend can still be visible even down to the order of  $10^{-15}$ , which potentially allows for the extracting of the localization length for the various cases we are interested in. In principle, using the TEBD technique, we can calculate  $G_2(d)$  for relatively large sizes up to  $L = 32$  within our current computational capacities. This has been illustrated for the case of



**Fig. A4** (a) The correlation function  $G_2(d)$  for  $L = 22$  with  $\mu = 0.9, 1.5, 2.0, 3.0, 6.0, 10.0$ . The result for  $L = 32$  and  $\mu = 6.0$  is also included for comparison; the insets of (a) are for  $G_2$  at  $\mu = 6.0$ , left with  $L = 12$  and right with  $L = 18$ . (b) The localization length  $\xi$  for  $L = 12, 14, 16, 18, 20, 22$  at different disorder strengths. The data are extracted from  $G_2(d)$  evaluated in the many-body states which have been evolved up to  $t = 100\tau$ . Here, the initial product state is chosen to be the  $p$ -state.

$\mu = 6.0$  in Fig. A4(a). Note that the results up to  $L = 22$  have been sufficient to elucidate the time evolution of  $G_2(d)$  at  $t = 100\tau$ . We plot the localization lengths at  $L = 12, 14, 16, 18, 20, 22$  for  $\mu \in [0.1, 10.0]$  in Fig. A4(b). From the color contrast of different sizes, one can see that the extracted localization lengths might be insensitive to the change in  $L$  for these values of  $\mu$  we consider.

For cluster MBL at small disorder, multiple particles of the Bose system can be localized at the same point. The steady-state  $S_N$  and  $S_C$  are thus obedient to the area law, and neither grows with time, indicating that it is a more robustly localized phenomenon [65]. However, as seen from Fig. A4, the cluster MBL has a greater single-particle localization length, which seems contradictory. One possible scenario for this might be the effect of energy mismatch. Concretely, on the one hand, when part of the cluster could move, the overall probability for such processes remains low due to the creation of the significant energy mismatch. Meanwhile, on the other

hand, these disentangled single particles might conversely gain the potential to hop far away in view of the fact that the disorder and the hopping generated energy offsets are barely small changes relative to the existing energy gap. Collectively, these effects give rise to a larger localization length for the single-particle process under the small disorder strengths. To some extent, the above reasoning is consistent with our explanations on the channel reflection asymmetry of  $\{S_N^n\}$  for the initial  $p$ -state as already detailed in Ref. [65].

## References

1. P. W. Anderson, Absence of diffusion in certain random lattices, *Phys. Rev.* 109(5), 1492 (1958)
2. E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Scaling theory of localization: Absence of quantum diffusion in two dimensions, *Phys. Rev. Lett.* 42(10), 673 (1979)
3. D. Vollhardt and P. Wölfle, Diagrammatic, self-consistent treatment of the Anderson localization problem in  $d \leq 2$  dimensions, *Phys. Rev. B* 22(10), 4666 (1980)
4. S. John, Electromagnetic absorption in a disordered medium near a photon mobility edge, *Phys. Rev. Lett.* 53(22), 2169 (1984)
5. K. Arya, Z. B. Su, and J. L. Birman, Localization of the surface plasmon polariton caused by random roughness and its role in surface-enhanced optical phenomena, *Phys. Rev. Lett.* 54(14), 1559 (1985)
6. Q. J. Chu and Z. Q. Zhang, Localization of phonons in mixed crystals, *Phys. Rev. B* 38(7), 4906 (1988)
7. R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, *Annu. Rev. Condens. Matter Phys.* 6(1), 15 (2015)
8. 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(5), 1126 (2006)
9. V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, *Phys. Rev. B* 75(15), 155111 (2007)
10. M. Žnidarič, T. Prosen, and P. Prelovšek, Many-body localization in the Heisenberg  $XXZ$  magnet in a random field, *Phys. Rev. B* 77(6), 064426 (2008)
11. J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* 43(4), 2046 (1991)
12. M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* 50(2), 888 (1994)
13. M. Serbyn, Z. Papić, and D. A. Abanin, Local conservation laws and the structure of the many-body localized states, *Phys. Rev. Lett.* 111(12), 127201 (2013)
14. A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, Constructing local integrals of motion in the many-body localized phase, *Phys. Rev. B* 91(8), 085425 (2015)
15. S. D. Geraedts, R. N. Bhatt, and R. Nandkishore, Emergent local integrals of motion without a complete set of localized eigenstates, *Phys. Rev. B* 95(6), 064204 (2017)
16. L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol,

- From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, *Adv. Phys.* 65(3), 239 (2016)
17. A. Lazarides, A. Das, and R. Moessner, Fate of many-body localization under periodic driving, *Phys. Rev. Lett.* 115(3), 030402 (2015)
  18. P. Ponte, Z. Papić, F. Huveneers, and D. A. Abanin, Many-body localization in periodically driven systems, *Phys. Rev. Lett.* 114(14), 140401 (2015)
  19. D. V. Else, B. Bauer, and C. Nayak, Floquet time crystals, *Phys. Rev. Lett.* 117(9), 090402 (2016)
  20. N. Y. Yao, A. C. Potter, I. D. Potirniche, and A. Vishwanath, Discrete time crystals: Rigidity, criticality, and realizations, *Phys. Rev. Lett.* 118(3), 030401 (2017)
  21. 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* 543(7644), 217 (2017)
  22. 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* 543(7644), 221 (2017)
  23. J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Many-body localization in a disordered quantum Ising chain, *Phys. Rev. Lett.* 113(10), 107204 (2014)
  24. M. Brenes, M. Dalmonte, M. Heyl, and A. Scardicchio, Many-body localization dynamics from gauge invariance, *Phys. Rev. Lett.* 120(3), 030601 (2018)
  25. E. Levi, M. Heyl, I. Lesanovsky, and J. P. Garrahan, Robustness of many-body localization in the presence of dissipation, *Phys. Rev. Lett.* 116(23), 237203 (2016)
  26. M. H. Fischer, M. Maksymenko, and E. Altman, Dynamics of a many-body-localized system coupled to a bath, *Phys. Rev. Lett.* 116(16), 160401 (2016)
  27. H. P. Lüschen, P. Bordia, S. S. Hodgman, M. Schreiber, S. Sarkar, A. J. Daley, M. H. Fischer, E. Altman, I. Bloch, and U. Schneider, Signatures of many-body localization in a controlled open quantum system, *Phys. Rev. X* 7(1), 011034 (2017)
  28. J. Ren, Q. Li, W. Li, Z. Cai, and X. Wang, Noise-driven universal dynamics towards an infinite temperature state, *Phys. Rev. Lett.* 124(13), 130602 (2020)
  29. M. Žnidarič, A. Scardicchio, and V. K. Varma, Diffusive and subdiffusive spin transport in the ergodic phase of a many-body localizable system, *Phys. Rev. Lett.* 117(4), 040601 (2016)
  30. S. P. Kelly, R. Nandkishore, and J. Marino, Exploring many-body localization in quantum systems coupled to an environment via Wegner–Wilson flows, *Nucl. Phys. B* 951, 114886 (2020)
  31. C. Chamon, A. Hamma, and E. R. Mucciolo, Emergent irreversibility and entanglement spectrum statistics, *Phys. Rev. Lett.* 112(24), 240501 (2014)
  32. C. R. Laumann, A. Pal, and A. Scardicchio, Many-body mobility edge in a mean-field quantum spin glass, *Phys. Rev. Lett.* 113(20), 200405 (2014)
  33. 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)
  34. I. Mondragon-Shem, A. Pal, T. L. Hughes, and C. R. Laumann, Many-body mobility edge due to symmetry-constrained dynamics and strong interactions, *Phys. Rev. B* 92(6), 064203 (2015)
  35. R. Mondaini and Z. Cai, Many-body self-localization in a translation-invariant Hamiltonian, *Phys. Rev. B* 96(3), 035153 (2017)
  36. M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, 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(6250), 842 (2015)
  37. J. H. Bardarson, F. Pollmann, and J. E. Moore, Unbounded growth of entanglement in models of many-body localization, *Phys. Rev. Lett.* 109(1), 017202 (2012)
  38. M. Serbyn, Z. Papić, and D. A. Abanin, Universal slow growth of entanglement in interacting strongly disordered systems, *Phys. Rev. Lett.* 110(26), 260601 (2013)
  39. 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(6437), 256 (2019)
  40. J. 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(6293), 1547 (2016)
  41. Q. Guo, C. Cheng, Z. H. Sun, Z. Song, H. Li, Z. Wang, W. Ren, H. Dong, D. Zheng, Y. R. Zhang, R. Mondaini, H. Fan, and H. Wang, Observation of energy-resolved many-body localization, *Nat. Phys.* 17(2), 234 (2021)
  42. M. Rispoli, A. Lukin, R. Schittko, S. Kim, M. E. Tai, J. Léonard, and M. Greiner, Quantum critical behaviour at the many-body localization transition, *Nature* 573(7774), 385 (2019)
  43. J. Léonard, S. Kim, M. Rispoli, A. Lukin, R. Schittko, J. Kwan, E. Demler, D. Sels, and M. Greiner, Probing the onset of quantum avalanches in a many-body localized system, *Nat. Phys.* 19(4), 481 (2023)
  44. J. M. Zhang, and R. X. Dong, Exact diagonalization: The Bose–Hubbard model as an example, *Eur. J. Phys.* 31(3), 591 (2010)
  45. S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Time-evolution methods for matrix-product states, *Ann. Phys.* 411, 167998 (2019)
  46. G. Vidal, Efficient simulation of one-dimensional quantum many-body systems, *Phys. Rev. Lett.* 93(4), 040502 (2004)
  47. M. P. Kennett, Out-of-equilibrium dynamics of the Bose–Hubbard model, *ISRN Cond. Matter Phys.* 393616, 39 (2013)
  48. P. Sengupta, and S. Haas, Quantum glass phases in the disordered Bose–Hubbard model, *Phys. Rev. Lett.* 99(5), 050403 (2007)
  49. V. Gurarie, L. Pollet, N. V. Prokof'ev, B. V. Svistunov, and M. Troyer, Phase diagram of the disordered Bose–Hubbard model, *Phys. Rev. B* 80(21), 214519 (2009)
  50. L. Pollet, N. V. Prokof'ev, B. V. Svistunov, and M. Troyer, Absence of a direct superfluid to Mott insulator transition in disordered Bose systems, *Phys. Rev. Lett.*

- 103(14), 140402 (2009)
51. M. Gerster, M. Rizzi, F. Tschirsich, P. Silvi, R. Fazio, and S. Montangero, Superfluid density and quasi-long-range order in the one-dimensional disordered Bose–Hubbard model, *New J. Phys.* 18(1), 015015 (2016)
  52. S. Hu, Y. Wen, Y. Yu, B. Normand, and X. Wang, Quantized squeezing and even–odd asymmetry of trapped bosons, *Phys. Rev. A* 80(6), 063624 (2009)
  53. J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, *Phys. Rev. E* 102(6), 062144 (2020)
  54. A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, Avalanches and many-body resonances in many-body localized systems, *Phys. Rev. B* 105(17), 174205 (2022)
  55. B. Widom, Surface tension and molecular correlations near the critical point, *J. Chem. Phys.* 43(11), 3892 (1965)
  56. T. Vojta, Phases and phase transitions in disordered quantum systems, *AIP Conf. Proc.* 1550, 188 (2013)
  57. T. Orell, A. A. Michailidis, M. Serbyn, and M. Silveri, Probing the many-body localization phase transition with superconducting circuits, *Phys. Rev. B* 100(13), 134504 (2019)
  58. S. X. Zhang and H. Yao, Universal properties of many body localization transitions in quasiperiodic systems, *Phys. Rev. Lett.* 121(20), 206601 (2018)
  59. V. Khemani, D. N. Sheng, and D. A. Huse, Two universality classes for the many-body localization transition, *Phys. Rev. Lett.* 119(7), 075702 (2017)
  60. J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Ergodicity breaking transition in finite disordered spin chains, *Phys. Rev. B* 102(6), 064207 (2020)
  61. P. Sierant, D. Delande, and J. Zakrzewski, Many-body localization for randomly interacting bosons, *Acta Phys. Pol. A* 132(6), 1707 (2017)
  62. P. Sierant, D. Delande, and J. Zakrzewski, Many-body localization due to random interactions, *Phys. Rev. A* 95, 021601(R) (2017)
  63. T. Orell, A. A. Michailidis, M. Serbyn, and M. Silveri, Probing the many-body localization phase transition with superconducting circuits, *Phys. Rev. B* 100(13), 134504 (2019)
  64. Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Distribution of the ratio of consecutive level spacings in random matrix ensembles, *Phys. Rev. Lett.* 110(8), 084101 (2013)
  65. J. Chen, C. Chen, and X. Wang, Energy- and symmetry-resolved entanglement dynamics in disordered Bose–Hubbard chain, arXiv: 2303.14825 (2023)
  66. 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)
  67. M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Evidence for unbounded growth of the number entropy in many-body localized phases, *Phys. Rev. Lett.* 124(24), 243601 (2020)
  68. D. J. Luitz, N. Laflorencie, and F. Alet, Extended slow dynamical regime close to the many-body localization transition, *Phys. Rev. B* 93, 060201(R) (2016)
  69. T. Kohler, S. Scherg, X. Li, H. P. Lüschen, S. Das Sarma, I. Bloch, and M. Aidelsburger, Observation of many-body localization in a one-dimensional system with a single-particle mobility edge, *Phys. Rev. Lett.* 122(17), 170403 (2019)
  70. K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, Rareregion effects and dynamics near the many-body localization transition, *Ann. Phys.* 529(7), 1600326 (2017)
  71. V. Khemani, D. N. Sheng, and D. A. Huse, Two universality classes for the many-body localization transition, *Phys. Rev. Lett.* 119(7), 075702 (2017)
  72. V. Khemani, S. P. Lim, D. N. Sheng, and D. A. Huse, Critical properties of the many-body localization transition, *Phys. Rev. X* 7(2), 021013 (2017)
  73. S. Schierenberg, F. Bruckmann, and T. Wettig, Wigner surmise for mixed symmetry classes in random matrix theory, *Phys. Rev. E* 85(6), 061130 (2012)
  74. S. D. Geraedts, R. Nandkishore, and N. Regnault, Many-body localization and thermalization: Insights from the entanglement spectrum, *Phys. Rev. B* 93(17), 174202 (2016)