


## Tower of two-dimensional scar states in a localized system

Michael Iversen,<sup>1</sup> Jens H. Bardarson<sup>1,2</sup>, and Anne E. B. Nielsen<sup>1</sup>

<sup>1</sup>*Department of Physics and Astronomy, Aarhus University, DK-8000 Aarhus C, Denmark*

<sup>2</sup>*Department of Physics, KTH Royal Institute of Technology, 106 91 Stockholm, Sweden*

 (Received 26 August 2023; revised 3 January 2024; accepted 5 January 2024; published 6 February 2024)

The eigenstate thermalization hypothesis describes how most isolated many-body quantum systems reach thermal equilibrium. However, the hypothesis is violated by phenomena such as many-body localization and quantum many-body scars. In this work, we study a finite, two-dimensional, disordered model hosting a tower of scar states. This construction is a particular instance of a general framework and we demonstrate its generality by constructing two disordered models hosting a different tower of scar states. At weak disorder, we find numerically that the spectra are nonthermal, and the scar states appear as exact eigenstates with high entropy for certain bipartitions. At strong disorder, the spectra localize and the scar states are identified as inverted scars since the scar states are embedded in a localized background as opposed to a thermal background. We argue that, for the considered type of models, the localization is stronger than what would be naively expected, and we show this explicitly for one of the models. The argument also provides guidelines for obtaining similarly strong localization in other scarred models. We study the transition from the thermal phase to localization by observing the adjacent gap ratio shifting from the Wigner surmise to the Poisson distribution with increasing disorder strength. Moreover, the entanglement entropy transitions from volume-law scaling with system size at weak disorder to area-law scaling at strong disorder. Finally, we demonstrate that localization protects scar revivals for initial states with partial support in the scar subspace.

DOI: [10.1103/PhysRevA.109.023310](https://doi.org/10.1103/PhysRevA.109.023310)

### I. INTRODUCTION

The eigenstate thermalization hypothesis (ETH) describes how isolated many-body quantum systems reach thermal equilibrium [1–3]. The hypothesis asserts that expectation values of local observables coincide with those from the microcanonical ensemble. The ETH makes predictions about generic quantum systems and has been verified for large classes of models, see Ref. [4] and references therein. However, several phenomena are known to violate the ETH.

When certain interacting many-body quantum systems are exposed to strong disorder, they transition from the thermal phase to being many-body localized (MBL) [5–7]. Localized systems conflict with ETH by, e.g., being insulating at finite temperature [8], having slow entanglement growth [9,10], subsystems retaining information about the initial state after a quench [11,12], etc. The nonthermal properties are attributed to the appearance of a complete set of quasilocal integrals of motion in the localized phase [13,14]. In MBL systems without a mobility edge, all energy eigenstates are nonthermal and MBL hence represents a strong violation of ETH. Besides existing in disordered systems, MBL may also emerge from gradient fields [15–17] or periodic driving [18–20]. Numerical studies have firmly established MBL in finite systems [7,9,11,21–26] and signatures of MBL have been observed in various experimental setups [27–31]. To what extent MBL remains stable in the thermodynamic limit is still being debated [32–37].

Quantum many-body scars (QMBSs) represent another violation of the ETH [38–40]. In scarred systems, a small number of ETH-violating eigenstates are embedded in an

otherwise thermal spectrum. Hence, a QMBSs represents a weak violation of the ETH. Contrary to thermal eigenstates, scar states display subvolume-law scaling of entanglement entropy. Furthermore, when scar states are equally spaced in energy, dynamical revivals are observed from initial states in the scar subspace. QMBSs can be traced back to the discovery of analytic excited eigenstates in the Affleck-Kennedy-Lieb-Tasaki model [41] which were later recognized as scar states [42]. Scar states have also been discovered in numerous other models [43–52] and several unifying formalisms have been developed [53–61]. QMBSs have been realized in experiments with interacting Rydberg atoms [62,63], on superconducting processors [64–66], in ultracold bosons [67] and in nitrogen-vacancy centers [68]. These experiments demonstrate that, even though QMBSs represent a vanishingly small fraction of the Hilbert space, they have a strong influence on system dynamics.

While MBL and QMBSs are independent nonthermal phenomena, several works have attempted to realize MBL and QMBSs simultaneously in one-dimensional systems [69–74]. One approach is to exploit the analytic structure of QMBSs to determine a set of local operators that annihilate the scar states. Adding these operators with random coefficients to the Hamiltonian introduces disorder into the model without disturbing the scar states. Disordered models with scars generally display features similar to MBL at strong disorder, e.g., energy levels following the Poisson distribution and area-law scaling of entanglement entropy with system size. It is unclear whether these models are truly MBL in the thermodynamic limit. But because of the similarity with MBL for finite system sizes, we refer to disordered models with scars as being

localized at strong disorder. Instead of being embedded in a thermal spectrum, the scar states reside among localized energy eigenstates. The scar states serve as “inverted scars” since they are not localized and hence represent a weak violation of localization. The interplay between localization and QMBSs generates interesting effects, such as the appearance of a mobility edge [71] or disorder stabilization of scar revivals [72]. However, demanding that a model hosts QMBSs puts constraints on the type of disorder that can be introduced. Consequently, these models may display a weaker type of localization and hence not localize in accordance with conventional MBL [72].

In this work, we construct a two-dimensional, disordered model hosting a tower of scar states based on the rainbow scar. At weak disorder, the majority of eigenstates near the center of each symmetry sector are thermal while the scar states are nonthermal. When increasing the disorder strength, the model transitions from the thermal phase to being localized. We demonstrate that the model localizes stronger than a naive prediction would suggest and that the model displays properties similar to MBL for the system sizes considered. Furthermore, we present guidelines for obtaining equally strong localization in other scarred models. The scar states display volume-law scaling of entanglement entropy for a particular bipartition while the remaining spectrum displays area-law scaling at strong disorder. Hence, the scar states represent a tower of inverted scars which remain nonlocalized even at strong disorder. We study the dynamics of initial states with partial support in the scar subspace and show that disorder enhances scar revivals. Finally, we consider a different tower of scar states and construct two disordered parent Hamiltonians. In one model, each scar state resides in a degenerate subspace for certain system sizes while, in the other model, the scar states are not degenerate with other states in the spectrum. These scar states display volume-law scaling of entanglement entropy with respect to multiple bipartitions. We demonstrate that the models localize at strong disorder and that the scar states represent a tower of inverted scars in both models.

In Sec. II A, we introduce the rainbow scar on a two-dimensional grid of spin-1/2 particles and discuss its basic properties. We also describe the projections of the rainbow scar into subspaces with definite magnetization in the  $z$  direction which will serve as a tower of inverted scars. In Sec. II B, we discuss the scaling of entanglement entropy with system size for the rainbow scar and the projections of the rainbow scar. In Sec. II C, we introduce a disordered parent Hamiltonian for the projections of the rainbow scar. We demonstrate that the model is thermal at weak disorder and that the scar states represent nonthermal outliers. In Sec. III A, we characterize the localization by first introducing the relevant concepts through a simple example. In Sec. III B, we provide a general description of the localization and discuss guidelines for ensuring similarly strong localization in other scarred models. In Sec. III C, we verify that the model localizes by observing the level spacing statistics transition from the Wigner surmise to the Poisson distribution with increasing disorder strength. In Sec. III D, we demonstrate that the entanglement entropy shifts from volume-law to area-law scaling with system size for increasing disorder strength, as expected for localization. In Sec. IV, we show that disorder stabilizes

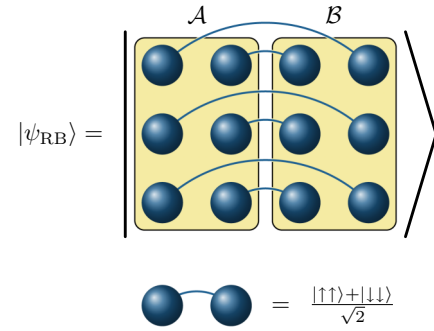


FIG. 1. The system consists of spin-1/2 particles arranged in a two-dimensional grid. The figure illustrates a system of size  $L_x \times L_y = 4 \times 3$ . The system is separated along the vertical direction into two parts  $\mathcal{A}$  and  $\mathcal{B}$  of equal size. The rainbow scar is the tensor product of Bell states between parts  $\mathcal{A}$  and  $\mathcal{B}$  as described in Eq. (4).

scar revivals from initial states with partial support in the scar subspace. In Sec. V A, we introduce a different tower of scar states. In Sec. V B, we demonstrate that the entanglement entropy of these scar states displays volume-law scaling with system size for multiple bipartitions. In Sec. V C, we introduce two disordered parent Hamiltonians for the tower of scar states. In Sec. V D, we demonstrate that the models localize at strong disorder and that the scar states represent a tower of inverted scars in both models. In Sec. VI, we summarize our results.

## II. MODEL

### A. Projections of the rainbow scar

We consider a two-dimensional grid  $\mathbf{r} = (x, y)$  of size  $L_x \times L_y$  occupied by spin-1/2 particles. We take  $L_x$  to be even throughout this work. The system is separated into two halves  $\mathcal{A}$  and  $\mathcal{B}$  according to

$$\begin{aligned} \mathcal{A} &= \left\{ (x, y) \mid x = 0, \dots, \frac{L_x}{2} - 1, y = 0, \dots, L_y - 1 \right\}, \\ \mathcal{B} &= \left\{ (x, y) \mid x = \frac{L_x}{2}, \dots, L_x - 1, y = 0, \dots, L_y - 1 \right\}. \end{aligned} \quad (1)$$

We denote the Hilbert space of part  $\mathcal{A}$  by  $\mathcal{H}_{\mathcal{A}}$  and the Hilbert space of part  $\mathcal{B}$  by  $\mathcal{H}_{\mathcal{B}}$ . However, since the two parts contain the same number of sites, we have  $\mathcal{H}_{\mathcal{A}} = \mathcal{H}_{\mathcal{B}}$ . Figure 1 illustrates the system and the corresponding partitioning for system size  $L_x \times L_y = 4 \times 3$ .

We construct a tower of scar states based on the rainbow scar [49,75,76]. Let basis( $\mathcal{H}_{\mathcal{A}}$ ) be a basis for  $\mathcal{H}_{\mathcal{A}}$ . The rainbow scar with respect to basis( $\mathcal{H}_{\mathcal{A}}$ ) is given by

$$|\psi_{\text{RB}}^{\text{general}}\rangle = \frac{1}{\sqrt{\dim(\mathcal{H}_{\mathcal{A}})}} \sum_{|\varphi\rangle \in \text{basis}(\mathcal{H}_{\mathcal{A}})} |\varphi\rangle \otimes \hat{\mathcal{M}}|\varphi\rangle, \quad (2)$$

where  $\dim(\mathcal{H}_{\mathcal{A}})$  is the dimension of  $\mathcal{H}_{\mathcal{A}}$  and  $\hat{\mathcal{M}}$  is the mirror operator. The mirror operator reflects each lattice site in the line  $x = (L_x - 1)/2$  without flipping the spin. The rainbow scar has the special property that acting with any operator  $\hat{\mathcal{O}}$  on part  $\mathcal{H}_{\mathcal{A}}$  is equivalent to acting with  $\hat{\mathcal{M}}\hat{\mathcal{O}}^T\hat{\mathcal{M}}$  on part

$\mathcal{H}_B$

$$(\hat{O} \otimes \hat{1})|\psi_{\text{RB}}^{\text{general}}\rangle = (\hat{1} \otimes \hat{M} \hat{O}^T \hat{M})|\psi_{\text{RB}}^{\text{general}}\rangle, \quad (3)$$

where  $\hat{1}$  is the identity operator and  $\hat{O}^T$  is the transpose of  $\hat{O}$  with respect to basis( $\mathcal{H}_A$ ). We discuss this property in Appendix A. For systems with two degrees of freedom on each site, Eq. (2) can be rewritten as a tensor product of Bell states,

$$|\psi_{\text{RB}}\rangle = \bigotimes_{r \in \mathcal{A}} \left( \frac{|\downarrow\rangle_r \otimes |\downarrow\rangle_{\tilde{r}} + |\uparrow\rangle_r \otimes |\uparrow\rangle_{\tilde{r}}}{\sqrt{2}} \right), \quad (4)$$

where  $\tilde{r} = (L_x - 1 - x, y)$  is the mirror image of  $r = (x, y)$ . The rainbow scar is illustrated in Fig. 1.

Following Ref. [49], we construct a tower of scar states from the rainbow scar by projecting  $|\psi_{\text{RB}}\rangle$  into subspaces with definite magnetization in the  $z$  direction. Let  $\hat{P}_M$  be the projection onto the subspace of the Hilbert space with magnetization  $M$ . The normalized projection of the rainbow scar is given by

$$|\psi_{\text{RB}}^M\rangle = \frac{\hat{P}_M |\psi_{\text{RB}}\rangle}{\sqrt{\langle \psi_{\text{RB}} | \hat{P}_M | \psi_{\text{RB}} \rangle}}. \quad (5)$$

Notice that the rainbow scar only has support in every other subspace, i.e.,  $|\psi_{\text{RB}}^M\rangle = 0$  for  $M = -L_x L_y / 2 + 1, -L_x L_y / 2 + 3, \dots, L_x L_y / 2 - 1$ . Hence, we focus on the set of scar states

$$\left\{ |\psi_{\text{RB}}^M\rangle \middle| M = -\frac{L_x L_y}{2}, -\frac{L_x L_y}{2} + 2, \dots, \frac{L_x L_y}{2} \right\}. \quad (6)$$

Alternatively, these states may be constructed from the anchor state  $|\Omega\rangle = \bigotimes_{r \in \mathcal{A} \cup \mathcal{B}} |\downarrow\rangle_r$  by acting repeatedly with the operator  $\hat{Q}^\dagger = \sum_{r \in \mathcal{A}} \hat{S}_r^+ \hat{S}_r^+$  where  $\hat{S}_r^+ = \hat{S}_r^x + i\hat{S}_r^y$ , i.e.,  $|\psi_{\text{RB}}^M\rangle \propto (\hat{Q}^\dagger)^{M/2 + L_x L_y / 4} |\Omega\rangle$ .

### B. Entanglement entropy of the scar states

Consider separating the system into two parts  $A$  and  $B$  (not necessarily equal to  $\mathcal{A}$  and  $\mathcal{B}$ ). For a state  $|\psi\rangle$ , the reduced density matrix is given by  $\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|)$ , where  $\text{Tr}_B$  is the partial trace over part  $B$ . The von Neumann entanglement entropy is given by  $S = -\text{Tr}[\rho_A \ln(\rho_A)]$ . The scaling of entanglement entropy with system size of the states  $|\psi_{\text{RB}}\rangle$  and  $|\psi_{\text{RB}}^M\rangle$  depends on the choice of partitioning. First, consider the vertical bipartition that separates the system into the left part  $A = \mathcal{A}$  and the right part  $B = \mathcal{B}$ . In this case, parts  $A$  and  $B$  contain one spin-1/2 particle from each Bell pair. Therefore, the entanglement entropy of  $|\psi_{\text{RB}}\rangle$  is  $S = L_x L_y \ln(2)/2$  and hence displays volume-law scaling with system size. For the projections of the rainbow scar  $|\psi_{\text{RB}}^M\rangle$ , the entanglement entropy also displays volume-law scaling with system size [49]. Next, let  $L_y$  be even and consider the horizontal bipartition where  $A$  consists of the spin-1/2 particles in the top half and  $B$  consists of the bottom half. In this case, each Bell pair is fully contained in either  $A$  or  $B$  and the entropy of the rainbow scar vanishes  $S = 0$ . The entropy of the states  $|\psi_{\text{RB}}^M\rangle$  is identical to the entropy of the ‘‘fine-tuned’’ cut in the corresponding one-dimensional model [49]. Therefore, the entropy of  $|\psi_{\text{RB}}^M\rangle$  displays logarithmic scaling with system size for the horizontal bipartition.

### C. Parent Hamiltonian

Consider a general Hamiltonian of the form

$$\hat{H} = \hat{H}_A \otimes \hat{1} + \hat{1} \otimes \hat{H}_B + \hat{H}_{AB} + \hat{H}_{\text{SG}}, \quad (7)$$

where  $\hat{H}_A$  acts within  $\mathcal{H}_A$  and  $\hat{H}_B$  acts within  $\mathcal{H}_B$ . The operators  $\hat{H}_{AB}$  and  $\hat{H}_{\text{SG}}$  act on degrees of freedom in both  $\mathcal{H}_A$  and  $\mathcal{H}_B$ . We choose the first term according to

$$\hat{H}_A = J \sum_{\substack{r, r' \in \mathcal{A} \\ \langle r, r' \rangle}} \mathbf{S}_r \cdot \mathbf{S}_{r'} + \sum_{r \in \mathcal{A}} h_r \hat{S}_r^z, \quad (8)$$

where the first sum is over all nearest neighbors  $\langle r, r' \rangle$  in  $\mathcal{A}$  and  $\mathbf{S}_r = (\hat{S}_r^x, \hat{S}_r^y, \hat{S}_r^z)$  are the spin-1/2 operators. The numbers  $h_r$  are drawn randomly from the uniform probability distribution across  $[-W, W]$ , where  $W$  is the disorder strength. We ensure the rainbow scar is an eigenstate of Eq. (7) by choosing the second term according to

$$\hat{H}_B = -\hat{M} \hat{H}_A \hat{M}. \quad (9)$$

Even though both  $\hat{H}_A$  and  $\hat{H}_B$  are disordered, Eq. (3) ensures that their sum annihilates the rainbow scar. Note, however, that Eq. (9) implies the random numbers  $h_r$  in parts  $\mathcal{A}$  and  $\mathcal{B}$  are not independent. The operator  $\hat{H}_{AB}$  connects parts  $\mathcal{A}$  and  $\mathcal{B}$ . We choose this operator such that the rainbow scar remains an exact eigenstate:

$$\hat{H}_{AB} = c \sum_{\substack{r \in \mathcal{A}, r' \in \mathcal{B} \\ \langle r, r' \rangle}} \mathbf{S}_r \cdot \mathbf{S}_{r'}. \quad (10)$$

The three operators  $\hat{H}_A$ ,  $\hat{H}_B$ , and  $\hat{H}_{AB}$  conserve the total magnetization in the  $z$  direction. Hence, the projection of the rainbow scar into a magnetization sector is itself an exact eigenstate. These projections are degenerate with respect to  $\hat{H}_A \otimes \hat{1} + \hat{1} \otimes \hat{H}_B + \hat{H}_{AB}$ . We lift the degeneracy by choosing the spectrum-generating term according to

$$\hat{H}_{\text{SG}} = \mu \sum_{r \in \mathcal{A} \cup \mathcal{B}} \hat{S}_r^z. \quad (11)$$

The energy of  $|\psi_{\text{RB}}^M\rangle$  is given by

$$\hat{H} |\psi_{\text{RB}}^M\rangle = \left( \frac{1}{4} c L_y + \mu M \right) |\psi_{\text{RB}}^M\rangle. \quad (12)$$

The states  $\{|\psi_{\text{RB}}^M\rangle\}$  hence represent a tower of scar states with equal energy spacing. Notice, that while the rainbow scar  $|\psi_{\text{RB}}\rangle$  is an eigenstate of the first three terms in Eq. (7), it is not an eigenstate of  $\hat{H}_{\text{SG}}$ .

Figure 2 shows the entanglement entropy of eigenstates of  $\hat{H}$  as a function of energy for several magnetization sectors. We consider parameter values  $J = \mu = c = 1$  and respectively weak and strong disorder. At weak disorder, the entropy forms a narrow arc within each symmetry sector. The scar states, on the other hand, generally appear as outliers external to each arc. This behavior signals that the scar states are nonthermal. When increasing the disorder strength, we observe that the entropy of generic eigenstates decreases. Note that the entropy of the scar states is constant as a function of disorder strength. Consequently, the entropy of the scar states is larger than that of generic eigenstates within the same symmetry sector at strong disorder. We also observe other eigenstates with similar energy to the scar states and

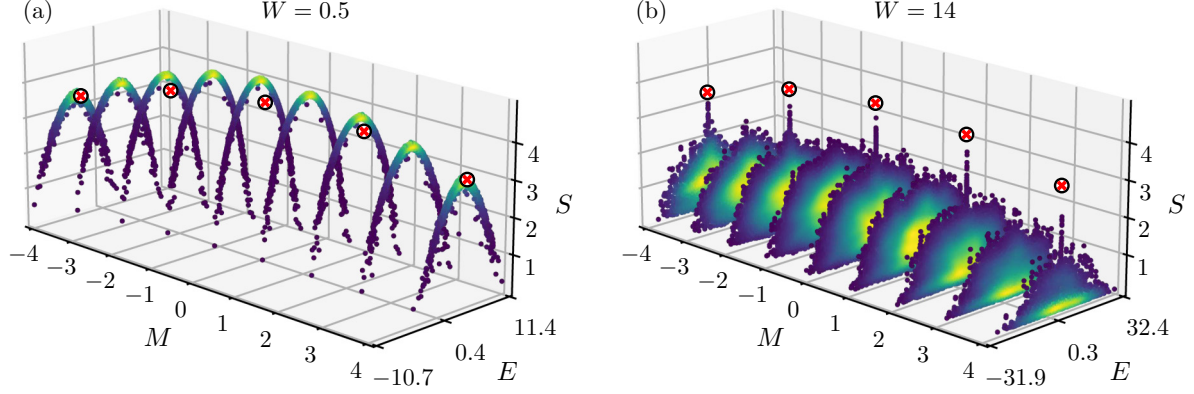


FIG. 2. The entanglement entropy of eigenstates of the Hamiltonian from Eq. (7) as a function of energy for several magnetization sectors. The entanglement entropy is computed with respect to the bipartition  $A = \mathcal{A}$  and  $B = \mathcal{B}$  and each panel displays the results for a single disorder realization. We consider system size  $L_x \times L_y = 4 \times 4$ , parameter values  $J = c = \mu = 1$  and (a) weak disorder  $W = 0.5$  and (b) strong disorder  $W = 14$ . The color intensity illustrates the density of points and lighter (darker) colors display higher (lower) density of points. The scar states are shown as crosses enclosed by a circle.

entropy at intermediate values. We describe the origin of these high-entropy eigenstates and investigate the behavior of the remaining spectrum at strong disorder in Sec. III.

Throughout this work, we generally consider parameter values  $J = \mu = c = 1$  and the largest symmetry sector  $M = 0$  unless otherwise stated. We generally reach similar results for other values  $J, \mu, c \neq 0$  and symmetry sectors.

### III. LOCALIZATION

It is not immediately obvious how the model behaves with increasing disorder strength for a fixed system size. On one hand, there exist numerous examples of spin models becoming MBL in a strongly disordered magnetic field [21,22]. In particular, a one-dimensional model similar to Eq. (7) was shown to display MBL characteristics at strong disorder [71]. On the other hand, the random magnetic fields are correlated which may represent an obstacle for localization. Furthermore, MBL is believed to be unstable for two-dimensional systems in the thermodynamic limit [77,78]. In this section, we demonstrate that the model localizes at strong disorder and that the model has characteristics similar to localized models with scar states in one dimension. Before tackling the general problem of characterizing the localization, we introduce the relevant concepts through a simple example.

#### A. Simple example

It is convenient to rewrite the Hamiltonian to highlight the correlations in the magnetic fields  $h_r$ :

$$\hat{H} = \hat{H}_0 + \sum_{r \in \mathcal{A}} h_r \hat{D}_r, \quad (13)$$

where we denote  $\hat{D}_r = \hat{S}_r^z - \hat{S}_{\tilde{r}}^z$  as “disorder operators” and  $\hat{H}_0$  is given by

$$\hat{H}_0 = J \left( \sum_{\substack{r, r' \in \mathcal{A} \\ (r, r')}} \mathbf{S}_r \cdot \mathbf{S}_{r'} - \sum_{\substack{r, r' \in \mathcal{B} \\ (r, r')}} \mathbf{S}_r \cdot \mathbf{S}_{r'} \right) + \hat{H}_{AB} + \hat{H}_{SG}. \quad (14)$$

In MBL systems, the set of disorder operators is commonly chosen to act uniquely on each state in a basis, i.e., for a set of disorder operators  $\{\hat{D}_i\}$  and a basis  $\{|\varphi\rangle\}$ , if  $\langle \varphi | \hat{D}_i | \varphi \rangle = \langle \varphi' | \hat{D}_i | \varphi' \rangle$  for all  $i$ , then  $|\varphi\rangle = |\varphi'\rangle$ . This is not true for our model in Eq. (13). Here, the disorder operators have the same action on some basis states. For instance, consider a lattice of size  $4 \times 2$ . For any disorder realization  $\{h_r | r \in \mathcal{A}\}$ , we have

$$\begin{aligned} & \left\langle \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \uparrow \\ \uparrow & \uparrow & \downarrow & \uparrow \end{array} \middle| \sum_{r \in \mathcal{A}} h_r \hat{D}_r \middle| \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \uparrow \\ \uparrow & \uparrow & \downarrow & \uparrow \end{array} \right\rangle \\ &= \left\langle \begin{array}{cccc} \downarrow & \uparrow & \uparrow & \downarrow \\ \downarrow & \uparrow & \downarrow & \uparrow \end{array} \middle| \sum_{r \in \mathcal{A}} h_r \hat{D}_r \middle| \begin{array}{cccc} \downarrow & \uparrow & \uparrow & \downarrow \\ \downarrow & \uparrow & \downarrow & \uparrow \end{array} \right\rangle. \quad (15) \end{aligned}$$

Hence, the disorder operators  $\{\hat{D}_r\}$  do not distinguish between the two basis states. Intuitively, this behavior is caused by pairs of spins being mirror symmetric. In the above example, sites  $\mathbf{r}_1 = (0, 0)$ ,  $\tilde{\mathbf{r}}_1 = (3, 0)$  and  $\mathbf{r}_2 = (1, 1)$ ,  $\tilde{\mathbf{r}}_2 = (2, 1)$  are mirror symmetric. The disorder operators act trivially on these pairs of spins

$$(h_{r_1} \hat{D}_{r_1} + h_{r_2} \hat{D}_{r_2}) \left| \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \uparrow \\ \uparrow & \uparrow & \downarrow & \uparrow \end{array} \right\rangle = 0. \quad (16)$$

The same is true for the corresponding basis state with both pairs flipped. The fact that the disorder operators do not distinguish between two basis states manifests itself in the structure of the energy eigenstates at strong disorder. We explore this structure in Sec. III B.

In the above example, the disorder operators had the same action on two different basis states within the zero magnetization sector. The discussion naturally extends to the case where more than two basis states are treated identically by the disorder operators. In this case, more than two pairs of spins are mirror symmetric and flipping any two pairs yields a new basis state which cannot be distinguished by the disorder operators.

### B. General characterization of localization

Let  $|\mathbf{D}, n\rangle$  be a simultaneous eigenket of the disorder operators  $\hat{D}_r$  where  $\mathbf{D} = \{D_r | r \in \mathcal{A}\}$  are the corresponding eigenvalues, i.e.,  $\hat{D}_r |\mathbf{D}, n\rangle = D_r |\mathbf{D}, n\rangle$  for all  $r \in \mathcal{A}$ . We refer to  $\mathbf{D}$  as the ‘‘disorder indices.’’ The disorder indices do not fully describe a state and the extra index  $n$  distinguishes between states with identical  $\mathbf{D}$ .

At strong disorder, the last term in Eq. (13) dominates and we expect each energy eigenstate to have significant support only on basis states with identical disorder indices. We refer to this behavior as eigenstates localizing within a subspace  $\mathcal{V}_{\mathbf{D}} = \text{span}(\{|\mathbf{D}, n\rangle | n = 1, 2, \dots\})$ . Note, however, that this behavior may not persist at any finite disorder strength in the thermodynamic limit. The energy eigenstates are labeled by  $\mathbf{D}$  at strong disorder. We further introduce an index  $m$  to distinguish energy eigenstates localizing within the same subspace  $\mathcal{V}_{\mathbf{D}}$ :

$$|E_{\mathbf{D},m}\rangle \approx \sum_n c_{\mathbf{D},mn} |\mathbf{D}, n\rangle. \quad (17)$$

This expression is an approximation rather than an equality due to the small support on other subspaces  $\mathcal{V}_{\mathbf{D}'}$  with  $\mathbf{D}' \neq \mathbf{D}$  which decreases with increasing disorder strength.

The special case where the disorder operators act uniquely on a basis state corresponds to  $\dim(\mathcal{V}_{\mathbf{D}}) = 1$ . At large disorder, the corresponding eigenstate has significant support on a single basis state, which we refer to as strong localization.

In the case  $\dim(\mathcal{V}_{\mathbf{D}}) > 1$ , the structure of the energy eigenstates is not immediately obvious. One might fear that all coefficients in Eq. (17) are nonzero  $c_{\mathbf{D},mn} \neq 0$  and the system ‘‘partially localizes’’ [72]. It turns out, however, that the majority of energy eigenstates localize more strongly than predicted by Eq. (17). Figure 3 illustrates the localization for system size  $4 \times 2$ . The figure demonstrates that eigenstates generally localize on a smaller subspace within each  $\mathcal{V}_{\mathbf{D}}$ . For instance, eigenstates in subspaces with  $\dim(\mathcal{V}_{\mathbf{D}}) = 2$  tend to localize on a single basis state. We find similar results for larger system sizes.

We understand the strong localization of eigenstates from degenerate perturbation theory. The central observation is that the off-diagonal matrix elements  $\langle \mathbf{D}, n | \hat{H}_0^p | \mathbf{D}, n' \rangle$  generally vanish for small powers  $p$ . In other words, the basis states with identical disorder indices are not connected directly by  $\hat{H}_0^p$ . Consequently, these states do not tend to mix and the energy eigenstates have significant support on only a few basis states. We present a detailed analysis in Appendix B.

While the eigenstates localize more strongly than predicted by Eq. (17), they might not localize on a single basis state. However, the energy difference between eigenstates with significant support within the same subspace  $\mathcal{V}_{\mathbf{D}}$  vanishes with increasing disorder strength. Consequently, the dynamics arising from such eigenstates can be made arbitrarily slow for sufficiently strong disorder.

Let  $\mathcal{D}_0$  be the disorder indices given by  $[\mathcal{D}_0]_r = 0$  for all  $r \in \mathcal{A}$ . The perturbative arguments are valid for all subspaces except  $\mathcal{V}_{\mathcal{D}_0}$ . This subspace contains the projections of the rainbow scar as well as other eigenstates. By construction, this subspace evades disorder and the entropy of eigenstates in  $\mathcal{V}_{\mathcal{D}_0}$  is typically larger than that of localized states but smaller or equal to the entropy of the scar states. The subspace  $\mathcal{V}_{\mathcal{D}_0}$  hence

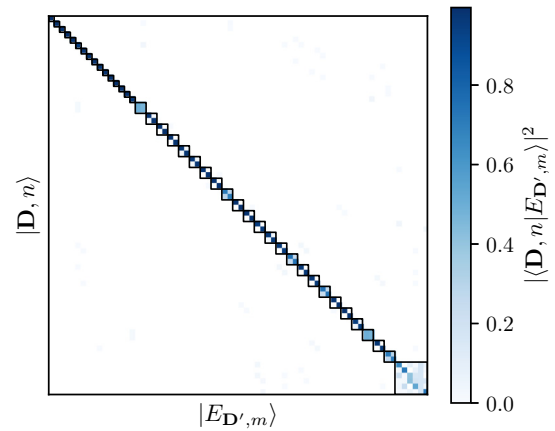


FIG. 3. The energy eigenstates of a single disorder realization for system size  $4 \times 2$ , disorder strength  $W = 14$ , parameters  $J = c = \mu = 1$ , and in the magnetization sector  $M = 0$ . Each row corresponds to a basis state  $|\mathbf{D}, n\rangle$  and each column to an eigenstate  $|E_{\mathbf{D}',m}\rangle$ . The color intensity of each pixel displays the norm squared overlap between the corresponding basis state and eigenstate  $|\langle \mathbf{D}, n | E_{\mathbf{D}',m}\rangle|^2$ . The eigenstates are rearranged to allow the diagonal shape and hence are not sorted into ascending order with respect to energy. The square boxes show the subspaces  $\mathcal{V}_{\mathbf{D}}$  where the eigenstates are expected to localize according to Eq. (17). The figure illustrates that most eigenstates localize on smaller subspaces within each  $\mathcal{V}_{\mathbf{D}}$ . Similar results are found for larger system sizes.

represents a nonlocalized subspace embedded in an otherwise localized spectrum. The subspace is visible in Fig. 2(b) as the eigenstates with energy close to a scar state and intermediate values of entropy. The part of  $\mathcal{V}_{\mathcal{D}_0}$  with magnetization  $M = 0$  is also visible in the bottom right corner of Fig. 3. In the following, we include the symmetry sector in the notation, i.e., the subspace with disorder indices  $\mathbf{D}$  and magnetization  $M$  is denoted as  $\mathcal{V}_{\mathbf{D}}^M$ . We determine the dimension of the nonlocalized subspace across all symmetry sectors  $\cup_M \mathcal{V}_{\mathcal{D}_0}^M$  by noting that it consists of all mirror-symmetric basis states. These basis states are fully determined by the spins in part  $\mathcal{A}$  and the dimension is given by  $\dim(\cup_M \mathcal{V}_{\mathcal{D}_0}^M) = 2^{L_x L_y / 2}$ . Consistent with Ref. [71], we find that the nonlocalized eigenstates represent a vanishing small part of the full Hilbert space  $\lim_{L_x, L_y \rightarrow \infty} [\dim(\cup_M \mathcal{V}_{\mathcal{D}_0}^M) / \dim(\mathcal{H})] = 0$ .

The considerations presented in this section are not specific to our model. They can be taken as general guidelines for constructing strongly localized models hosting QMBs. Strong localization is obtained by ensuring the nondisordered part of the Hamiltonian restricted to the subspace  $\mathcal{V}_{\mathbf{D}}$  is diagonal. This rule may help future work on disordered models hosting QMBs achieve stronger localization.

### C. Adjacent gap ratio

The distribution of energy levels indicates whether the system is thermal or localized. The distribution follows the Wigner surmise in the thermal phase. In particular, the distribution is given by the Gaussian orthogonal ensemble (GOE) since the Hamiltonian in Eq. (7) is invariant under time reversal. We expect the energy levels to follow the Poisson distribution at strong disorder similar to an MBL system [79].

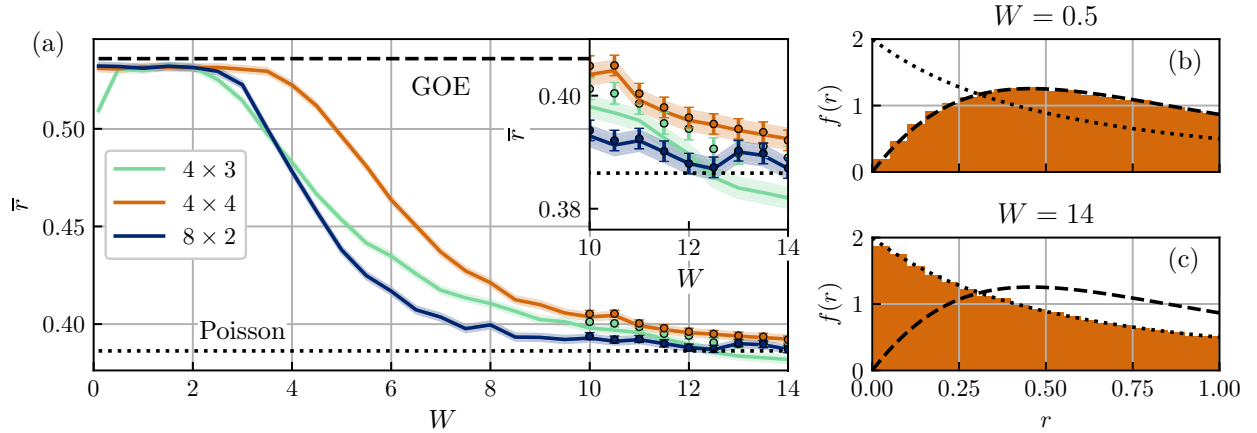


FIG. 4. (a) The adjacent gap ratio  $\bar{r}$  averaged over the  $10^2$  energies closest to  $(E_{\min} + E_{\max})/2$  in  $10^3$  disorder realizations (solid lines) as a function of disorder strength  $W$  for different system sizes  $L_x \times L_y$ . The GOE and Poisson values are illustrated as dashed and dotted horizontal lines. The shaded areas display two standard deviations on the estimate of the mean when assuming a Gaussian distribution. At weak disorder,  $\bar{r}$  coincides with GOE indicating that the system is thermal. At strong disorder,  $\bar{r}$  agrees with the Poisson distribution showing that the system is localized. The adjacent gap ratio decreases below the Poisson value for system size  $4 \times 3$  at strong disorder  $12 \lesssim W$  as illustrated in the inset. When restricting the computation of  $\bar{r}$  to eigenstates with different disorder indices (dots), the adjacent gap ratio converges to the Poisson value for all system sizes. The error bars display two standard deviations on the estimate of the mean. (b), (c) The distribution of adjacent gap ratio for system size  $4 \times 4$  at (b) weak disorder  $W = 0.5$  and (c) strong disorder  $W = 14$ . The distributions  $f_{\text{GOE}}$  and  $f_{\text{Poisson}}$  from Eq. (19) are illustrated as dashed and dotted curves. The data agree with (b)  $f_{\text{GOE}}$  at weak disorder, showing that the system is thermal, and agrees with (c)  $f_{\text{Poisson}}$  at strong disorder, indicating that the system is localized. In all panels, we consider parameters  $J = c = \mu = 1$  and the symmetry sector  $M = 0$ .

Let  $\{E_i\}$  be the energies in ascending order and let  $\delta_i = E_{i+1} - E_i > 0$  be the  $i$ th energy gap. We consider the adjacent gap ratio [7]

$$r_i = \frac{\min(\delta_i, \delta_{i+1})}{\max(\delta_i, \delta_{i+1})}. \quad (18)$$

The transition from thermal to localized behavior is identified by studying the distribution of the adjacent gap ratio. The probability density functions of  $r$  for a thermal and localized system are given by [80]

$$f_{\text{GOE}}(r) = \frac{27}{4} \frac{r(1+r)}{(1+r+r^2)^{5/2}}, \quad (19a)$$

$$f_{\text{Poisson}}(r) = \frac{2}{(1+r)^2}. \quad (19b)$$

The corresponding expectation values of the two distributions are given by  $\langle r \rangle_{\text{GOE}} = 2(2 - \sqrt{3}) \approx 0.536$  and  $\langle r \rangle_{\text{Poisson}} = 2 \ln 2 - 1 \approx 0.386$ .

We consider  $10^3$  disorder realizations and compute the adjacent gap ratio from the  $10^2$  energies closest to the center of the spectrum, i.e.,  $(E_{\min} + E_{\max})/2$  where  $E_{\min}$  and  $E_{\max}$  are respectively the smallest and largest energy in the spectrum. The adjacent gap ratio is averaged over both disorder realizations and this central part of the spectrum. Figure 4(a) illustrates the mean adjacent gap ratio  $\bar{r}$  as a function of disorder strength  $W$  for different system sizes  $L_x \times L_y$ . At weak disorder  $0.1 \lesssim W \lesssim 2$ , the mean adjacent gap ratio agrees with GOE for all system sizes. Hence, the system is thermal at weak disorder. When increasing the disorder strength  $2 \lesssim W \lesssim 12$ ,  $\bar{r}$  departs from the GOE value and approaches the

Poisson value. At strong disorder  $12 \lesssim W$ , the mean adjacent gap ratio agrees well with the Poisson value indicating that the model localizes.

For system size  $4 \times 3$  at strong disorder  $12 \lesssim W$ , the adjacent gap ratio is seen to decrease below the Poisson value. This behavior is explained by the model not localizing analogously to conventional MBL. As discussed in Sec. III B, the energy gap between eigenstates with identical disorder indices decreases with increasing disorder strength. Therefore, the distribution of energy levels only approaches the Poisson distribution when considering eigenstates with different disorder indices. With this in mind, we analyze the adjacent gap ratio more carefully by only including energy gaps  $\delta_i$  of eigenstates with different disorder indices, i.e., pairs of adjacent eigenstates  $|E_{\mathbf{D},m}\rangle$  and  $|E_{\mathbf{D}',m'}\rangle$  with  $\mathbf{D} \neq \mathbf{D}'$ . This filtering is only possible at strong disorder where all eigenstates are sufficiently localized and the disorder indices of each eigenstate can be reliably determined numerically. For each eigenstate  $|E_i\rangle$ , we determine the basis states  $|\mathbf{D}, n\rangle$  with the largest norm squared overlap and thereby deduce the disorder indices, i.e.,  $\text{argmax}_{\mathbf{D}} (\sum_n |\langle \mathbf{D}, n | E_i \rangle|^2)$ . As illustrated in Fig. 4(a), the adjacent gap ratio of all system sizes converges to the Poisson value when only considering eigenstates of different disorder indices. The two methods of calculating the adjacent gap ratio nearly coincide for system sizes  $4 \times 4$  and  $8 \times 2$ .

Finally, we illustrate the adjacent gap ratio distribution in Figs. 4(b) and 4(c) and compare it with Eq. (19). The distribution agrees with  $f_{\text{GOE}}$  at weak disorder and with  $f_{\text{Poisson}}$  at strong disorder. Thus, the system transitions from the thermal phase to being localized as disorder is introduced.

### D. Entanglement entropy

We further establish the transition from the thermal phase to localization by studying the von Neumann entropy. The system is separated into two parts  $A$  and  $B$ . Let  $\rho$  be the density matrix describing the full system and let  $\rho_A = \text{Tr}_B(\rho)$  be the reduced density matrix of part  $A$ . The von Neumann entropy is given by

$$S = -\text{Tr}[\rho_A \ln(\rho_A)]. \quad (20)$$

In the thermal phase, the entropy scales with the volume of the system. In particular, for a thermal system with tensor product structure  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ , the entropy of an infinite-temperature eigenstate is given by the Page value  $\langle S \rangle_{\text{Page}}$  [81]. The model described by Eqs. (7)–(11) conserves the total magnetization, and the Hilbert space of a generic magnetization sector  $\mathcal{H}_M$  does not have tensor product structure,  $\mathcal{H}_M \neq \mathcal{H}_A \otimes \mathcal{H}_B$ . Therefore, the Page value is not representative of the average entropy of a thermal state in this model. Instead, the expected entropy in the thermal phase is given by [82,83]

$$\begin{aligned} \langle S \rangle_{\text{therm}} = & \sum_{M_A = \max(-\frac{N_A}{2}, M - \frac{N_B}{2})}^{\min(\frac{N_A}{2}, M + \frac{N_B}{2})} \frac{\dim(\mathcal{H}_{M_A}) \dim(\mathcal{H}_{M_B})}{\dim(\mathcal{H}_M)} \\ & \times \{ \langle S \rangle_{\text{Page}} + \Psi[\dim(\mathcal{H}_M) + 1] \\ & - \Psi[\dim(\mathcal{H}_{M_A}) \dim(\mathcal{H}_{M_B}) + 1] \}. \end{aligned} \quad (21)$$

In this expression,  $M_A$  is the magnetization of part  $A$  and  $M_B = M - M_A$  is the magnetization of part  $B$ . The Hilbert space of part  $A$  ( $B$ ) with magnetization  $M_A$  ( $M_B$ ) is denoted by  $\mathcal{H}_{M_A}$  ( $\mathcal{H}_{M_B}$ ). The number of lattice sites in part  $A$  ( $B$ ) is denoted by  $N_A$  ( $N_B$ ). Finally,  $\Psi$  is the digamma function. In the MBL phase, on the other hand, the entanglement entropy is proportional to the boundary of the partition [13,84]. Furthermore, the variance of entropy displays a peak at the transition between the thermal and MBL phase [23,85]. We expect our model to display similar characteristics of entanglement entropy at strong disorder.

We consider  $2 \times 10^3$  disorder realizations and for each realization compute the entropy of the eigenstate closest in energy to  $(E_{\min} + E_{\max})/2$ . The sample average  $\bar{S}$  and sample variance  $\sigma_S^2$  are then determined. Figure 5(a) illustrates the average entropy as a function of disorder strength for different system sizes and partitions  $A, B$ . We consider three partitions described by the lines  $x = (L_x - 1)/2$ ,  $x = (L_x - 3)/2$  and  $y = \lfloor L_y/2 \rfloor - 1/2$ , where  $\lfloor \cdot \rfloor$  is the function that rounds down to the nearest integer. All sites on one side of the line represent part  $A$  and the remaining sites represent part  $B$ . The partitions are illustrated in Fig. 6 for system size  $4 \times 3$ .

Figure 5(a) illustrates the average entropy as a function of disorder strength and Fig. 5(b) shows the variance. In both figures, we display multiple system sizes and partitions. For weak disorder, the average entropy agrees with  $\langle S \rangle_{\text{therm}}$  for all system sizes and partitions. As the disorder strength is increased, the average entropy decreases and the variance displays a sharp peak. The large variance signals that the system transitions from the thermal phase to being localized. In the inset of Fig. 5(a), we display the entropy as a function of boundary size for each system size and partition at  $W = 14$ .

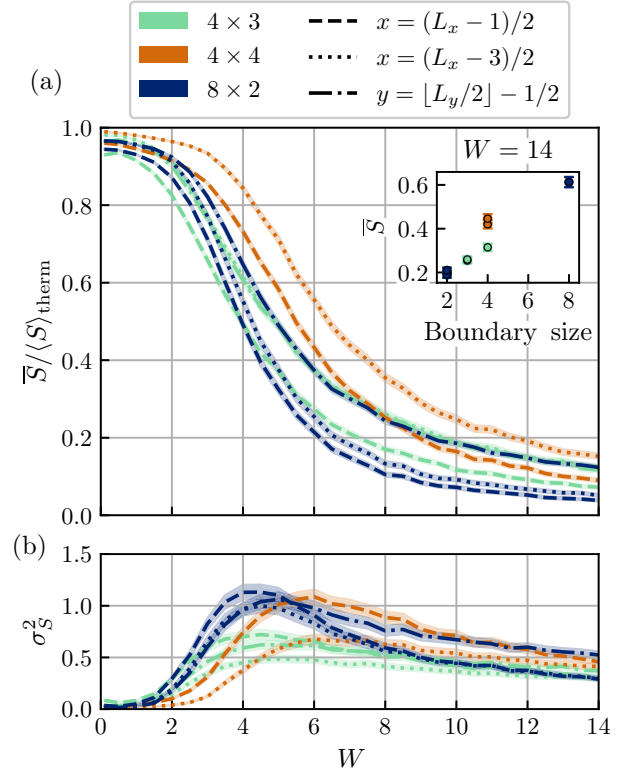


FIG. 5. (a) The average entanglement entropy  $\bar{S}$  over  $2 \times 10^3$  disorder realizations of the eigenstate closest in energy to  $(E_{\min} + E_{\max})/2$ . The entropy is plotted as a function of disorder strength  $W$  for different system sizes and partitions. The system size is indicated by color and the partition by line style. At weak disorder, the entropy for all considered system sizes and partitions agree with  $\langle S \rangle_{\text{therm}}$ . At strong disorder, the entropy departs from the thermal value and instead displays area-law scaling as illustrated in the inset. (b) The sample variance as a function of disorder strength. The variance displays a peak at intermediate disorder strength indicating that the system transitions from being thermal to localized. The shaded areas and error bars in both figures display two standard deviations on the estimate of the mean and variance. We consider parameters  $J = c = \mu = 1$  and the symmetry sector  $M = 0$ .

We observe that the entropy is proportional to the size of the boundary as expected for localization.

### IV. SCAR DYNAMICS

The presence of rainbow scars in the model causes non-thermal dynamics. We observe such dynamics by choosing the initial state with large support in the scar subspace. We consider the thermofield double state  $|\psi_{\text{TFD}}(\beta)\rangle$  at inverse temperature  $\beta$  [86]. Let  $\hat{H}_{\text{TFD}}$  be an operator acting on  $\mathcal{H}_A$ , then the thermofield double state is given by

$$|\psi_{\text{TFD}}(\beta)\rangle = \frac{1}{\sqrt{Z}} \sum_i e^{-\beta E_i/2} |E_i\rangle \otimes \hat{\mathcal{M}}|E_i\rangle, \quad (22)$$

where  $E_i$  and  $|E_i\rangle$  are the energies and eigenstates of  $\hat{H}_{\text{TFD}}$  and  $Z$  is a normalization constant. Several protocols have been proposed for realizing  $|\psi_{\text{TFD}}(\beta)\rangle$  [86–88]. The thermofield double state reduces to the rainbow scar from Eq. (2) in

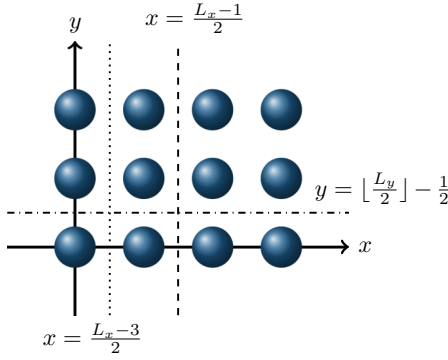


FIG. 6. The system is separated into two parts A and B. We illustrate the separation for system size  $4 \times 3$ . The boundary of the two parts is described by the lines  $x = (L_x - 1)/2$  (dashed line),  $x = (L_x - 3)/2$  (dotted line) and  $y = \lfloor \frac{L_y}{2} \rfloor - \frac{1}{2}$  (dash-dotted line).

the limit of infinite temperature  $|\psi_{\text{TFD}}(0)\rangle = |\psi_{\text{RB}}\rangle$  for any operator  $\hat{H}_{\text{TFD}}$ .

First, we consider the initial state  $|\psi(t=0)\rangle = |\psi_{\text{RB}}\rangle$  and study the system dynamics. The state is time-evolved according to  $|\psi(t)\rangle = e^{-i\hat{H}t}|\psi(0)\rangle$  and the fidelity  $F(t) = |\langle\psi(0)|\psi(t)\rangle|^2$  is determined. Figure 7(a) shows the fidelity as a function of time. As expected,  $F(t)$  displays revivals with period  $T_{\text{scar}} = \pi/\mu$  since the initial state resides fully in the scar subspace.

Next, we initialize the system only partially within the scar subspace  $|\psi(0)\rangle = |\psi_{\text{TFD}}(\beta)\rangle$  with  $\beta = 5$ . The thermofield double state is prepared with respect to the operator

$$\hat{H}_{\text{TFD}} = J \sum_{r,r' \in A} \hat{S}_r^z \hat{S}_{r'}^z + h \sum_{r \in A} \hat{S}_r^z, \quad (23)$$

with  $J = h = 1$ . Since  $|\psi_{\text{TFD}}(\beta)\rangle$  only partially resides in the scar subspace, we do not expect perfect revivals. Figure 7(b) illustrates the average fidelity over 500 disorder realizations for different disorder strengths. At weak disorder, the fidelity displays revivals with a small amplitude. The reduced amplitude is caused by the initial state having significant support outside the scar subspace. At strong disorder, the fidelity displays revivals with a larger amplitude than the thermal case. This behavior is explained by the initial state having support on basis states  $|\mathbf{D}_0, n\rangle$  with identical disorder indices to the rainbow scar. Consequently, the initial state is a linear combination of energy eigenstates which all have energy close to a scar state. Thus, the disorder enhances the scar revivals from initial states with partial support in the scar subspace.

## V. ALTERNATIVE MODEL

Inverted quantum scars are characterized by having a larger entanglement entropy than the background of localized eigenstates. The entanglement entropy of the scar states from Eq. (6) displays logarithmic scaling with system size for a particular bipartition. The entropy of the scar states for this bipartition may therefore be similar to that of localized states for small system sizes. We address this issue by constructing a different tower of scar states with larger entanglement entropy for generic bipartitions. The model follows the same theo-

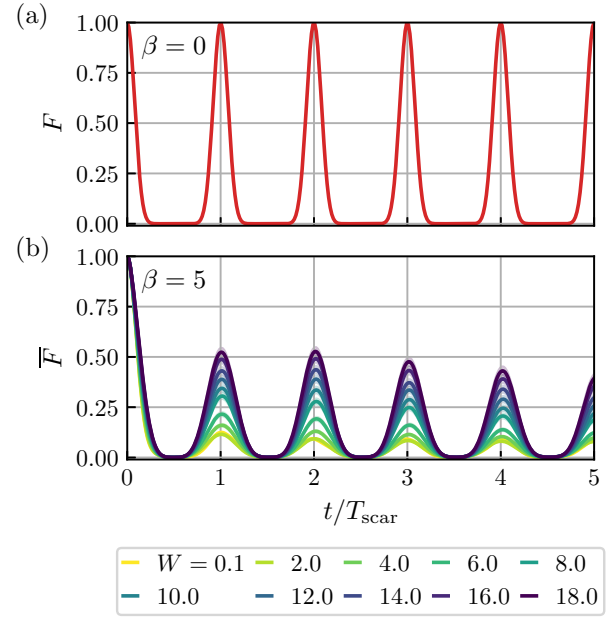


FIG. 7. (a) The fidelity  $F$  as a function of time  $t$  when initializing the system as the rainbow scar  $|\psi(0)\rangle = |\psi_{\text{RB}}\rangle$ . The fidelity displays revivals because the projections  $|\psi_{\text{RB}}^M\rangle$  have equal energy spacing. (b) The system is initialized as the thermofield double state  $|\psi(0)\rangle = |\psi_{\text{TFD}}(\beta)\rangle$  with  $\beta = 5$  and time evolved at different disorder strengths. The figure displays the average fidelity over 500 disorder realizations and the shaded areas show two standard deviations on the estimate of the mean. At weak disorder, the fidelity displays revivals with a small amplitude. The revival amplitude increases with increasing disorder strength. In both panels, we consider system size  $L_x \times L_y = 4 \times 3$  and parameters  $J = c = \mu = 1$ .

retical framework presented in Appendix A and hence also illustrates the flexibility in our construction.

### A. Projections of the scar state

We construct a scar state which, similar to the rainbow scar, is the tensor product of Bell pairs when the system consists of spin-1/2 particles. We ensure this scar state has larger entanglement entropy for generic bipartitions than the rainbow scar by increasing the distance between the two sites constituting each Bell pair. Consider the operator  $\hat{\mathcal{R}}_\pi$  which rotates all lattice sites around the center of the lattice  $\mathbf{r}_{\text{center}} = [(L_x - 1)/2, (L_y - 1)/2]$  by an angle  $\pi$  without flipping any spins. We consider the state

$$|\psi_{\text{rot}}^{\text{general}}\rangle = \frac{1}{\sqrt{\dim(\mathcal{H}_A)}} \sum_{|\varphi\rangle \in \text{basis}(\mathcal{H}_A)} |\varphi\rangle \otimes \hat{\mathcal{R}}_\pi |\varphi\rangle. \quad (24)$$

Similar to the rainbow scar, this state has a special property. For any operator  $\hat{\mathcal{O}}$  acting on  $\mathcal{H}_A$  we have

$$(\hat{\mathcal{O}} \otimes \hat{1}) |\psi_{\text{rot}}^{\text{general}}\rangle = (\hat{1} \otimes \hat{\mathcal{R}}_\pi \hat{\mathcal{O}}^T \hat{\mathcal{R}}_\pi) |\psi_{\text{rot}}^{\text{general}}\rangle. \quad (25)$$

When the system consists of spin-1/2 particles, the state may be written as a tensor product of Bell states:

$$|\psi_{\text{rot}}\rangle = \bigotimes_{r \in A} \left( \frac{|\downarrow\rangle_r \otimes |\downarrow\rangle_{\bar{r}} + |\uparrow\rangle_r \otimes |\uparrow\rangle_{\bar{r}}}{\sqrt{2}} \right), \quad (26)$$



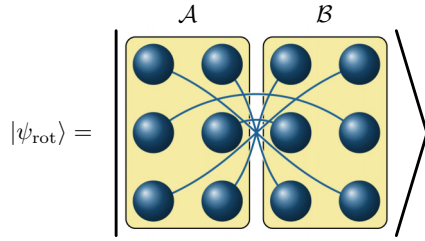


FIG. 8. Illustration of the state from Eq. (26) for a lattice of size  $4 \times 3$ . The state is the tensor product of Bell states  $(|\downarrow\rangle_r|\downarrow\rangle_{\bar{r}} + |\uparrow\rangle_r|\uparrow\rangle_{\bar{r}})/\sqrt{2}$  where  $\bar{r}$  corresponds to rotating  $r$  around the center of the lattice  $r_{\text{center}}$  by an angle  $\pi$ . Two balls connected by a line represent a Bell state.

with  $r = (x, y)$  and  $\bar{r} = (L_x - 1 - x, L_y - 1 - y)$  is obtained by rotating  $r$  around the center of the lattice  $r_{\text{center}}$  by an angle  $\pi$ . Figure 8 illustrates the state from Eq. (26).

Consider the normalized projection of Eq. (26) into the subspace with magnetization  $M$  in the  $z$  direction

$$|\psi_{\text{rot}}^M\rangle = \frac{\hat{P}_M |\psi_{\text{rot}}\rangle}{\sqrt{\langle \psi_{\text{rot}} | \hat{P}_M | \psi_{\text{rot}} \rangle}}. \quad (27)$$

We focus on the scar states

$$\left\{ |\psi_{\text{rot}}^M\rangle \middle| M = -\frac{L_x L_y}{2}, -\frac{L_x L_y}{2} + 2, \dots, \frac{L_x L_y}{2} \right\}, \quad (28)$$

since  $|\psi_{\text{rot}}^M\rangle = 0$  for  $M = -L_x L_y/2 + 1, -L_x L_y/2 + 3, \dots, L_x L_y/2 - 1$ .

### B. Entanglement entropy of the scar states

The entanglement entropy of the scar states from Eq. (28) display volume-law scaling with system size for a variety of bipartitions. In the following, we partition the system into two parts  $A_\ell$  and  $B_\ell$  using a line  $\ell$ . For a given line  $\ell$ , all lattice sites on one side of the line belong to  $A_\ell$  and the remaining sites belong to  $B_\ell$ . We focus on lines not passing through any lattice points. Consider the line  $\ell_{r_{\text{dir}}} : \{r_{\text{center}} + s r_{\text{dir}} | s \in \mathbb{R}\}$  which passes through the center of the lattice  $r_{\text{center}}$  in the direction  $r_{\text{dir}}$ . Notice that for any  $r_{\text{dir}}$ , the line  $\ell_{r_{\text{dir}}}$  separates all Bell pairs. Consequently, the entanglement entropy of  $|\psi_{\text{rot}}\rangle$  is given by  $S = L_x L_y \ln(2)/2$  and hence displays volume-law scaling for any  $r_{\text{dir}}$ . Furthermore, the entanglement entropy of the projections  $|\psi_{\text{rot}}^M\rangle$  also displays volume-law scaling for any such bipartition.

We compare the entanglement entropy of the scar states in Eq. (28) with the scar states based on the rainbow scar in Eq. (6). The distance between two lattice sites forming a Bell pair in  $|\psi_{\text{rot}}\rangle$  is equal to or larger than the corresponding Bell pair in  $|\psi_{\text{RB}}\rangle$ . Consequently, the probability that a randomly chosen bipartition separates a Bell pair is greater for  $|\psi_{\text{rot}}\rangle$  than for  $|\psi_{\text{RB}}\rangle$ . The entanglement entropy of Eq. (28) is hence larger than the entropy of Eq. (6) for a generic bipartition. This means that the scar states  $|\psi_{\text{rot}}^M\rangle$  are more easily distinguished from localized states by the entanglement entropy than  $|\psi_{\text{RB}}^M\rangle$ .

### C. Parent Hamiltonian

We construct a parent Hamiltonian for the scar states  $|\psi_{\text{rot}}^M\rangle$  by following the ideas presented in Sec. II C and Appendix A. Consider the Hamiltonian

$$\hat{H}_{\text{rot}} = \hat{H}_A \otimes \hat{1} + \hat{1} \otimes \hat{H}_B + \tilde{H}_{AB} + \hat{H}_{\text{SG}}, \quad (29)$$

where  $\hat{H}_A$  and  $\hat{H}_{\text{SG}}$  are identical to Eqs. (8) and (11). The remaining terms are given by

$$\tilde{H}_B = -\hat{R}_\pi \hat{H}_A \hat{R}_\pi, \quad (30)$$

$$\tilde{H}_{AB} = c \left[ \sum_{\substack{r \in A, r' \in B \\ (r, r') \\ y, y' \leq \frac{L_y-1}{2}}} (\mathbf{S}_r \cdot \mathbf{S}_{r'}) - \sum_{\substack{r \in A, r' \in B \\ (r, r') \\ \frac{L_y-1}{2} < y, y'}} (\mathbf{S}_r \cdot \mathbf{S}_{r'}) \right], \quad (31)$$

where  $y$  and  $y'$  are the  $y$  coordinates of  $r$  and  $r'$ , respectively. The first sum in Eq. (31) includes all nearest neighbors across the boundary between  $A$  and  $B$ , which lies in the lower part of the lattice. Similarly, the second sum includes all nearest neighbors across the boundary in the upper part of the lattice. The states  $|\psi_{\text{rot}}^M\rangle$  are eigenstates of  $\hat{H}_{\text{rot}}$  with energy depending on whether  $L_y$  is even or odd

$$\hat{H}_{\text{rot}} |\psi_{\text{rot}}^M\rangle = \mu M |\psi_{\text{rot}}^M\rangle \quad (L_y \text{ even}), \quad (32a)$$

$$\hat{H}_{\text{rot}} |\psi_{\text{rot}}^M\rangle = \left( \frac{c}{4} + \mu M \right) |\psi_{\text{rot}}^M\rangle \quad (L_y \text{ odd}). \quad (32b)$$

The scar states hence form a tower with equal energy spacing. For even  $L_y$ , a degenerate subspace of energy eigenstates resides at each scar energy  $E = \mu M$  for  $M = -L_x L_y/2, -L_x L_y/2 + 2, \dots, L_x L_y/2$ . For the system sizes considered, we numerically find that the subspace at energy  $E = \mu M$  consists of energy eigenstates with magnetization  $M$ . We further find that the dimension of the subspace at energy  $E = \mu M$  is given by  $\binom{L_x L_y/2}{M/2 + L_x L_y/4}$  for the system sizes considered.

We also study a related model hosting the same scar states without the degenerate subspaces. Consider the operator

$$\hat{H}_{\text{nnn}} = c' (\mathbf{S}_{r_1} \cdot \mathbf{S}_{\bar{r}_1} - \mathbf{S}_{r_2} \cdot \mathbf{S}_{\bar{r}_2}), \quad (33)$$

where  $r_1 = (L_x/2 - 1, L_y/2 - 1)$ ,  $r_2 = (L_x/2 - 1, L_y/2)$  are in part  $A$  and  $\bar{r}_1 = (L_x/2, L_y/2)$ ,  $\bar{r}_2 = (L_x/2, L_y/2 - 1)$  are obtained by rotating  $r_1$  and  $r_2$  around  $r_{\text{center}}$  by an angle  $\pi$ . The operator  $\hat{H}_{\text{nnn}}$  connects two next-nearest-neighbor pairs across the boundary of parts  $A$  and  $B$ . We study the model

$$\hat{H}'_{\text{rot}} = \hat{H}_{\text{rot}} + \hat{H}_{\text{nnn}} \quad (34)$$

for even  $L_y$ . The scar states are eigenstates of  $\hat{H}'_{\text{rot}}$  with energy  $\hat{H}'_{\text{rot}} |\psi_{\text{rot}}^M\rangle = \mu M |\psi_{\text{rot}}^M\rangle$ . For generic parameter values, the scar states are the only eigenstates at the energies  $E = \mu M$ .

### D. Localization

We study the two models from Eqs. (29) and (34) at different disorder strengths and investigate to what extent the models localize. We generally consider parameter values  $J = c = c' = \mu = 1$  and the largest magnetization sector  $M = 0$  but we expect similar results for other parameter values and

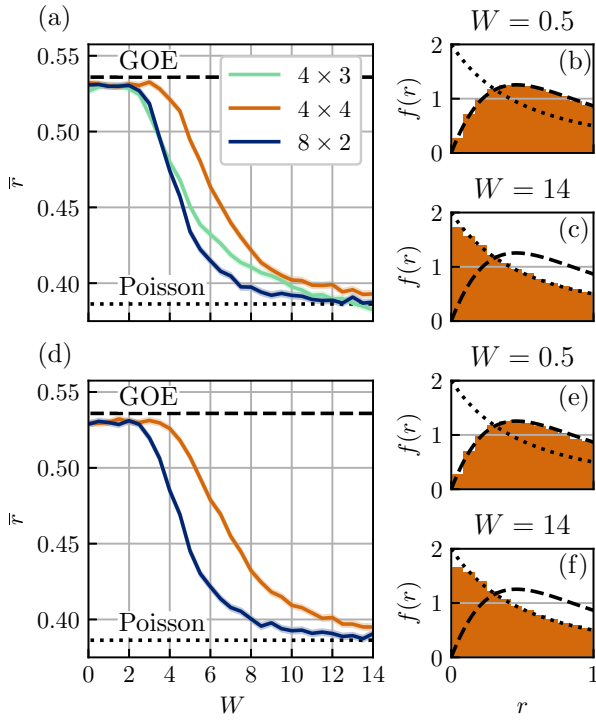


FIG. 9. Level spacing statistics for the Hamiltonians (a)–(c)  $\hat{H}_{\text{rot}}$  and (d)–(f)  $\hat{H}'_{\text{rot}}$ . We consider parameters  $J = c = c' = \mu = 1$  and the symmetry sector  $M = 0$ . Panels (a) and (d) display the mean adjacent gap ratio  $\bar{r}$  as a function of disorder strength  $W$  for different system sizes. The adjacent gap ratio is averaged over the  $10^2$  energies closest to  $(6E_{\min} + 4E_{\max})/10$  in  $10^3$  disorder realizations. The shaded areas show two standard deviations on the estimate of the mean. The expected value for GOE is shown by the dashed lines and for the Poisson distribution by the dotted lines. For both models, the average adjacent gap ratio agrees with GOE at weak disorder indicating that the models are thermal. At strong disorder, the average adjacent gap ratio agrees with the Poisson value, indicating that both models are localized. Panels (b), (c) and (e), (f) show the distribution of the adjacent gap ratio for system size  $4 \times 4$ . The disorder is weak  $W = 0.5$  in panels (b) and (e) while it is strong  $W = 14$  in panels (c) and (f). The distributions  $f_{\text{GOE}}$  and  $f_{\text{Poisson}}$  from Eq. (19) are displayed as dashed and dotted curves, respectively. In both models, the distribution of the adjacent gap ratio agrees with  $f_{\text{GOE}}$  at weak disorder and the distribution agrees with  $f_{\text{Poisson}}$  at strong disorder.

symmetry sectors. For even  $L_y$ , the Hamiltonian  $\hat{H}_{\text{rot}}$  restricted to the  $M = 0$  sector anticommutes with the rotation operator  $\hat{\mathcal{R}}_{\pi}$ . Consequently, for each eigenstate with energy  $E_i$ , there exists a different eigenstate with energy  $-E_i$  and the spectrum is hence symmetric around  $E = 0$ . The spectrum contains a degenerate subspace at energy  $E = 0$  which includes a scar state. Therefore, we generally avoid the center of the spectrum when investigating the localization of generic eigenstates. Instead, we study eigenstates with energies slightly below the center of the spectrum, i.e., eigenstates with energies close to  $E = (6E_{\min} + 4E_{\max})/10$ . We investigate how the models behave with increasing disorder strength. Figures 9(a) and 9(d) illustrate the mean adjacent gap ratio as a function of disorder strength for respectively  $\hat{H}_{\text{rot}}$  and  $\hat{H}'_{\text{rot}}$ . For each model, the adjacent gap ratio is averaged over the  $10^2$  eigenstates closest

in energy to  $(6E_{\min} + 4E_{\max})/10$  in  $10^3$  disorder realizations. At weak disorder, the mean adjacent gap ratio agrees with GOE indicating that the models are thermal. The mean adjacent gap ratio decreases with increasing disorder strength and coincides with the Poisson value at strong disorder. This behavior signals that the models localize at strong disorder. Figures 9(b), 9(c) and 9(e), 9(f) display the distribution of the adjacent gap ratio at weak and strong disorder for  $\hat{H}_{\text{rot}}$  and  $\hat{H}'_{\text{rot}}$ . For both models, the distribution agrees with GOE at weak disorder and it agrees with the Poisson distribution at strong disorder. We further investigate the transition by studying the entanglement entropy as a function of disorder strength, as illustrated in Fig. 10. We consider  $2 \times 10^3$  disorder realizations and compute the entropy of the eigenstate closest in energy to  $(6E_{\min} + 4E_{\max})/10$ . The average and variance of the entanglement entropy are then computed. Figures 10(a) and 10(c) display the results for  $\hat{H}_{\text{rot}}$  while Figs. 10(b) and 10(d) show the results for  $\hat{H}'_{\text{rot}}$ . At weak disorder, the average entanglement entropy agrees with the thermal value  $\langle S \rangle_{\text{therm}}$  from Eq. (21) for both models. As the disorder strength is increased, the average entropy decreases and the variance displays a peak. At strong disorder, the average entropy is approximately proportional to the boundary size of the partition. These results support the claim that the two models are thermal at weak disorder and become localized at strong disorder.

## VI. CONCLUSION

We studied a two-dimensional, disordered model hosting a tower of scar states based on the rainbow scar. At weak disorder, the nonthermal scar states are embedded among thermal eigenstates. The model transitions from being thermal to localized with increasing disorder strength. Using a perturbative approach, we demonstrated that the model displays strong localization at large disorder and we confirmed this statement numerically. Strong localization refers to energy eigenstates having significant support on a small number of basis states. Furthermore, we provided general guidelines for obtaining strong localization in other scarred models. We verified that the model transitions from the thermal phase to being localized with increasing disorder strength by studying the level spacing statistics and entanglement entropy. We showed that the adjacent gap ratio shifts from GOE to the Poisson distribution with increasing disorder strength. The entanglement entropy displayed volume-law scaling with system size at weak disorder and area-law scaling at strong disorder. Consequently, the scar states were identified as inverted scars at strong disorder. We studied the system dynamics from initial states with support in the scar subspace. The fidelity displayed persistent revivals when the initial state was fully embedded in the scar subspace. When the initial state had partial support outside the scar subspace, the fidelity displayed revivals with a smaller amplitude. We demonstrated that the revival amplitude increases with increasing disorder strength and we interpret this result as the localization protecting the scar revivals. Finally, we constructed two disordered models hosting a tower of scar states with larger entanglement entropy for generic bipartitions. We demonstrated that these models localize by studying the adjacent gap ratio and the scaling

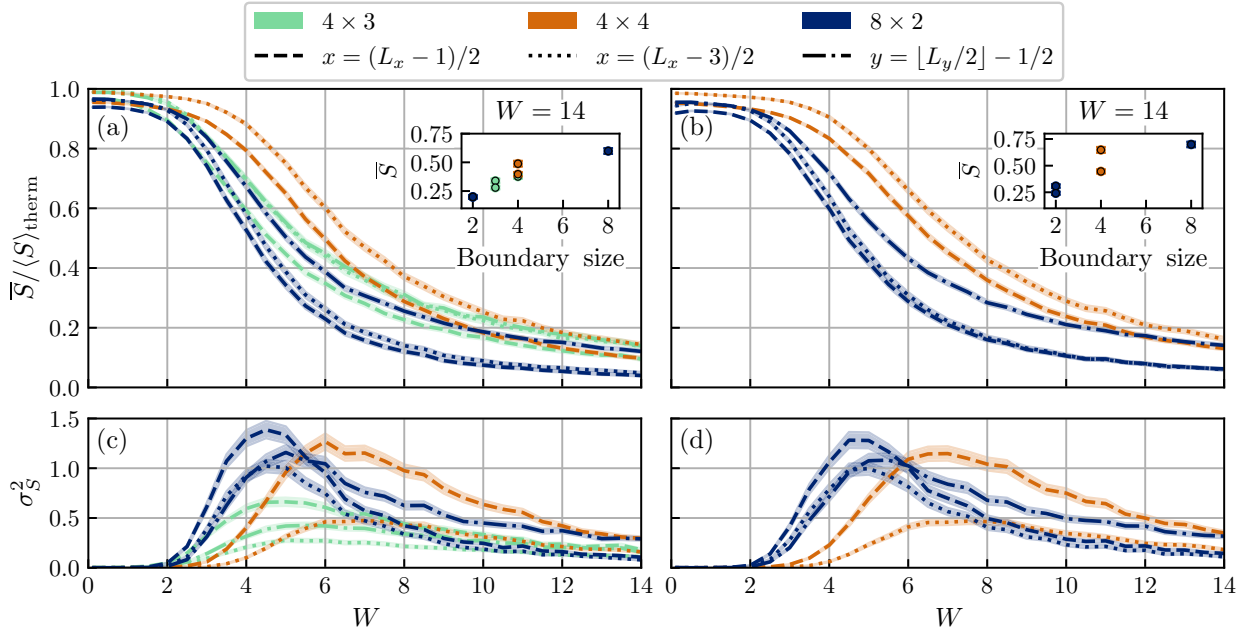


FIG. 10. (a), (b) The average entanglement entropy  $\bar{S}$  and (c), (d) the variance of entropy  $\sigma_S^2$  of the eigenstate with energy closest to  $(6E_{\min} + 4E_{\max})/10$  over  $2 \times 10^3$  disorder realizations. Panels (a) and (c) correspond to the model  $\hat{H}_{\text{rot}}$  while panels (b) and (d) correspond to  $\hat{H}'_{\text{rot}}$ . The average and variance are shown as a function of disorder strength  $W$  for different system sizes and partitions. We consider the partitions illustrated in Fig. 6. The system size is indicated by color and the partition by line style. The shaded areas display two standard deviations on the estimate of the mean and variance. We consider the parameters  $J = c = c' = \mu = 1$  and magnetization  $M = 0$ . In both models, the average entropy agrees with the thermal value  $(S)_{\text{therm}}$  from Eq. (21) at weak disorder and decreases with increasing disorder strength. The insets display the average entropy at strong disorder  $W = 14$  as a function of boundary size. The variance of the entropy displays a peak as the models transition from being thermal to localized.

of entanglement entropy as a function of disorder strength. Hence, the scar states represent a tower of inverted scars in both models.

#### ACKNOWLEDGMENTS

This work has been supported by Carlsbergfondet under Grant No. CF20-0658 and has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 101001902).

#### APPENDIX A: SCAR STATES BASED ON THE EINSTEIN-PODOLSKY-ROSEN STATE

Scar states and corresponding parent Hamiltonians may be constructed from the Einstein-Podolsky-Rosen (EPR) state. Let  $\mathcal{H}$  be a Hilbert space of the form  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$  with  $\mathcal{H}_A = \mathcal{H}_B$  and basis  $(\mathcal{H}_A)$  a basis for  $\mathcal{H}_A$ . The EPR state with respect to basis  $(\mathcal{H}_A)$  is given by

$$|\psi_{\text{EPR}}\rangle = \frac{1}{\sqrt{\dim(\mathcal{H}_A)}} \sum_{|\varphi\rangle \in \text{basis}(\mathcal{H}_A)} |\varphi\rangle \otimes |\varphi\rangle, \quad (\text{A1})$$

where  $\dim(\mathcal{H}_A)$  is the dimension of  $\mathcal{H}_A$ . For any operator  $\hat{\mathcal{O}}$  acting within  $\mathcal{H}_A$ , the EPR state has the special property

$$(\hat{\mathcal{O}} \otimes \hat{\mathbb{1}})|\psi_{\text{EPR}}\rangle = (\hat{\mathbb{1}} \otimes \hat{\mathcal{O}}^T)|\psi_{\text{EPR}}\rangle, \quad (\text{A2})$$

where  $\hat{\mathcal{O}}^T$  is the transpose of  $\hat{\mathcal{O}}$  with respect to basis  $(\mathcal{H}_A)$  [89].

Based on the EPR state, one may construct other states which inherit a property similar to Eq. (A2). Let  $\hat{U}_A$  and  $\hat{U}_B$  be unitary operators acting within  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , respectively. Consider the state

$$|\psi_{\hat{U}_A, \hat{U}_B}\rangle = (\hat{U}_A \otimes \hat{U}_B)|\psi_{\text{EPR}}\rangle. \quad (\text{A3})$$

For any operator  $\hat{\mathcal{O}}$  acting within  $\mathcal{H}_A$ , the state fulfills the following property:

$$(\hat{\mathcal{O}} \otimes \hat{\mathbb{1}})|\psi_{\hat{U}_A, \hat{U}_B}\rangle = (\hat{\mathbb{1}} \otimes \hat{\mathcal{O}}')|\psi_{\hat{U}_A, \hat{U}_B}\rangle, \quad (\text{A4a})$$

with

$$\hat{\mathcal{O}}' = \hat{U}_B \hat{U}_A^T \hat{\mathcal{O}}^T \hat{U}_A^* \hat{U}_B^\dagger \quad (\text{A4b})$$

where  $(\dots)^*$  is the complex conjugate with respect to basis  $(\mathcal{H}_A)$  and  $(\dots)^\dagger$  is the Hermitian conjugate. This property may be verified by direct calculation. Following Ref. [89], Eq. (A4) can be utilized to determine parent Hamiltonians for the state  $|\psi_{\hat{U}_A, \hat{U}_B}\rangle$ . Consider a general Hamiltonian of the form

$$\hat{H} = \hat{H}_A \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \hat{H}_B + \sum_i \lambda_i \hat{\mathcal{O}}_A^{(i)} \otimes \hat{\mathcal{O}}_B^{(i)}, \quad (\text{A5})$$

where  $\hat{H}_A, \hat{\mathcal{O}}_A^{(i)}$  act within  $\mathcal{H}_A$  while  $\hat{H}_B, \hat{\mathcal{O}}_B^{(i)}$  act within  $\mathcal{H}_B$  and  $\lambda_i \in \mathbb{R}$ . The state  $|\psi_{\hat{U}_A, \hat{U}_B}\rangle$  is an eigenstate of  $\hat{H}$  with

energy  $E$  if

$$\left[ \hat{\mathbb{1}} \otimes \left( \hat{H}'_A + \hat{H}_B + \sum_i \lambda_i \hat{O}_B^{(i)} \hat{O}_A^{(i)'} \right) \right] |\psi_{\hat{U}_A, \hat{U}_B}\rangle = E |\psi_{\hat{U}_A, \hat{U}_B}\rangle, \quad (\text{A6})$$

where  $\hat{H}'_A$  and  $\hat{O}_A^{(i)'}$  are determined from Eq. (A4b). The rainbow scar is described by this framework and corresponds to the choice  $\hat{U}_A = \hat{\mathbb{1}}$  and  $\hat{U}_B = \hat{\mathcal{M}}$ , where  $\hat{\mathcal{M}}$  is the mirror operator, i.e.,  $|\psi_{\text{RB}}\rangle = |\psi_{\hat{\mathbb{1}}, \hat{\mathcal{M}}}\rangle$ . The state from Eq. (24) corresponds to  $\hat{U}_A = \hat{\mathbb{1}}$  and  $\hat{U}_B = \hat{\mathcal{R}}_\pi$ , where  $\hat{\mathcal{R}}_\pi$  is the rotation operator around the center of the lattice by an angle  $\pi$ , i.e.,  $|\psi_{\text{rot}}\rangle = |\psi_{\hat{\mathbb{1}}, \hat{\mathcal{R}}_\pi}\rangle$ .

## APPENDIX B: A PERTURBATIVE APPROACH TO CHARACTERIZING THE LOCALIZATION

At strong disorder, the eigenstates of the Hamiltonian in Eq. (13) are expected to localize within a vector space  $\mathcal{V}_D = \text{span}\{|\mathbf{D}, n\rangle | n = 1, 2, \dots\}$ ,

$$|E_{\mathbf{D}, m}\rangle \approx \sum_n c_{\mathbf{D}, mn} |\mathbf{D}, n\rangle. \quad (\text{B1})$$

At first glance, when the sum in Eq. (B1) contains more than one term, the eigenstates seem to have significant support on many basis states. It turns out, however, that the eigenstates exhibit stronger localization than predicted by Eq. (B1). The eigenstates generally localize on small subspaces of  $\mathcal{V}_D$ . We describe the localization using degenerate Rayleigh-Schrödinger perturbation theory [90]. For completeness, we briefly review degenerate perturbation theory before applying it to our model.

### 1. Degenerate perturbation theory

We investigate the model at strong disorder. The Hamiltonian is rewritten to extract the disorder strength,

$$\hat{H} = \hat{H}_0 + W \sum_{r \in \mathcal{A}} h'_r \hat{D}_r, \quad (\text{B2})$$

where  $h'_r$  are drawn randomly with uniform probability from the interval  $[-1, 1]$ . Next, we define  $\lambda = 1/W$  and study the related Hamiltonian  $\tilde{H} = \hat{H}/W$ :

$$\tilde{H} = \lambda \hat{H}_0 + \sum_{r \in \mathcal{A}} h'_r \hat{D}_r. \quad (\text{B3})$$

The two Hamiltonians share all eigenstates  $|\tilde{E}_{\mathbf{D}, m}\rangle = |E_{\mathbf{D}, m}\rangle$  and the energies are related according to  $\tilde{E}_{\mathbf{D}, m} = E_{\mathbf{D}, m}/W$ . In the following, we focus on a set of eigenstates with identical disorder indices  $\{|\tilde{E}_{\mathbf{D}, m}\rangle | m = 1, 2, \dots\}$  at strong disorder. In the extreme limit  $\lambda = 0$ , these eigenstates are exactly degenerate. At  $0 < \lambda$ , we expand the energies and eigenstates in powers of  $\lambda$ :

$$\tilde{E}_{\mathbf{D}, m} = \sum_{\ell=0}^{\infty} \lambda^\ell \tilde{E}_{\mathbf{D}, m}^{(\ell)}, \quad (\text{B4a})$$

$$|\tilde{E}_{\mathbf{D}, m}\rangle = \sum_n c_{\mathbf{D}, mn}^{(0)} |\mathbf{D}, n\rangle + \sum_{\ell=1}^{\infty} \sum_{\mathbf{D}' \neq \mathbf{D}} \sum_n \lambda^\ell c_{\mathbf{D}', mn}^{(\ell)} |\mathbf{D}', n\rangle. \quad (\text{B4b})$$

Note that Eq. (B4b) is not normalized to simplify notation.

We investigate how the eigenstates localize within the subspaces  $\mathcal{V}_D$  by determining the zero-order coefficients  $c_{\mathbf{D}, mn}^{(0)}$ . For convenience, we collect these coefficients into vectors  $\mathbf{c}_m^{(0)}$  with  $[\mathbf{c}_m^{(0)}]_n = c_{\mathbf{D}, mn}^{(0)}$ .

We substitute Eq. (B4) into  $\tilde{H}|\tilde{E}_{\mathbf{D}, m}\rangle = \tilde{E}_{\mathbf{D}, m}|\tilde{E}_{\mathbf{D}, m}\rangle$  and compare similar powers of  $\lambda$ . This leads to the familiar result that the first-order correction to the energies  $\tilde{E}_{\mathbf{D}, m}^{(1)}$  are the eigenvalues of the matrix

$$[\hat{M}^{(1)}]_{mn'} = \langle \mathbf{D}, n | \hat{H}_0 | \mathbf{D}, n' \rangle. \quad (\text{B5})$$

The zero-order coefficients  $\mathbf{c}_m^{(0)}$  are the eigenvectors of the same matrix. If all eigenvalues of  $\hat{M}^{(1)}$  are different, then every coefficient  $c_{\mathbf{D}, mn}^{(0)}$  is uniquely determined. In the case where some eigenvalues of  $\hat{M}^{(1)}$  are degenerate, the corresponding zero-order coefficients  $\mathbf{c}_m^{(0)}$  remain undetermined. Let  $\{\mathbf{v}_i^{(1)} | i = 1, 2, \dots\}$  be eigenvectors of  $\hat{M}^{(1)}$  corresponding to the same eigenvalue. The yet undetermined coefficients are linear combinations of these eigenvectors  $\mathbf{c}_m^{(0)} \in \text{span}(\{\mathbf{v}_i^{(1)} | i = 1, 2, \dots\})$ . The correct linear combinations are found as the eigenvectors of  $\hat{M}^{(2)}$  with the corresponding eigenvalues being the second-order energy corrections  $E_{\mathbf{D}, m}^{(2)}$ . This matrix is obtained from the general expression

$$\hat{M}^{(\ell)} = \hat{P} \hat{H}_0 \hat{T}^{(\ell-1)}, \quad (\text{B6})$$

where  $\hat{P} = \sum_n |\mathbf{D}, n\rangle \langle \mathbf{D}, n|$  is the projection onto  $\mathcal{V}_D$ . The matrices  $\hat{T}^{(\ell)}$  are determined from the following recursion:

$$\hat{T}^{(1)} = \hat{\mathcal{E}} \hat{H}_0 \hat{P}, \quad (\text{B7a})$$

$$\hat{T}^{(\ell)} = \hat{\mathcal{E}} \left( \hat{H}_0 \hat{T}^{(\ell-1)} - \sum_{\ell'=1}^{\ell-1} E_n^{(\ell-\ell')} \hat{T}^{(\ell')} \right), \quad (\text{B7b})$$

with

$$\hat{\mathcal{E}} = \sum_{\mathbf{D}' \neq \mathbf{D}} \sum_{n'} \frac{1}{E_{\mathbf{D}, n}^{(0)} - E_{\mathbf{D}', n'}^{(0)}} |\mathbf{D}', n'\rangle \langle \mathbf{D}', n'|. \quad (\text{B8})$$

If the degeneracy persists to second order, we proceed to third order by diagonalizing  $\hat{M}^{(3)}$  in the relevant subspaces and so on. In this manner, all coefficients  $\mathbf{c}_m^{(0)}$  can be found by going to high enough order in perturbation theory.

### 2. Characterizing the localization

Returning to the model from Eq. (13), we study the energy eigenstates at strong disorder. We note that  $\hat{H}_0$  contains two-body kinetic terms. Furthermore, for two basis states belonging to the same magnetization sector and with identical disorder indices  $|\mathbf{D}, n\rangle$  and  $|\mathbf{D}, n'\rangle$ , the spins are different on more than two sites. This implies that the off-diagonal elements of  $M^{(1)}$  are zero for all  $\mathbf{D}$ . Hence, for each unique eigenvalue of  $M^{(1)}$ , the corresponding eigenstate localizes on a single basis state, i.e.,  $c_{\mathbf{D}, mn}^{(0)} = \delta_{mn}$ , where  $\delta_{mn}$  is the Kronecker delta. These eigenstates display strong localization. Similar arguments are valid for coefficients  $c_{\mathbf{D}, mn}^{(0)}$  determined in higher order of perturbation theory. Let  $p_D$  be the largest integer such that  $\langle \mathbf{D}, n | \hat{H}_0^{pp} | \mathbf{D}, n' \rangle = 0$  for all  $n \neq n'$ . For instance, we have  $p_D = 4$  in the example from Eq. (15). This integer is typically large since  $\hat{H}_0$  only contains local two-body terms. Inspecting Eqs. (B6)–(B8), the off-diagonal elements

of  $\hat{M}^{(\ell)}$  vanish for all  $\ell \leq p_D$ . Hence, if  $c_m^{(0)}$  is determined in  $\ell$ th order of perturbation theory with  $\ell < p_D$ , then the corresponding energy eigenstate localizes on a single basis state. From these considerations, we expect most energy eigenstates to localize on smaller subspaces than predicted by Eq. (17). We remark that the perturbative arguments are not valid for  $\mathcal{V}_{D_0}$  with  $[D_0]_r = 0$  for all  $r \in \mathcal{A}$ . This subspace is insensitive to disorder, i.e.,  $\hat{D}_r |D_0, n\rangle = 0$  for all  $r \in \mathcal{A}$ . We expect the eigenstates with significant support in  $\mathcal{V}_{D_0}$  to remain nonlocalized even at strong disorder.

While most energy eigenstates display strong localization, some eigenstates localize in a weaker sense. These eigenstates have significant support on multiple basis states at strong disorder. However, the nonlocalized nature of these eigenstates is not visible in the dynamics of observables at small times. Let  $\{|\tilde{E}_{D,m}\rangle | m = 1, 2, \dots\}$  be a set of such eigenstates and recall that the degeneracy is lifted in second or higher order of perturbation theory, i.e.,  $\tilde{E}_{D,m}^{(1)} = \tilde{E}_{D,m'}^{(1)}$  for  $m \neq m'$ . We return to the energies  $E_{D,m}$  and eigenstates  $|E_{D,m}\rangle$  of the original Hamiltonian. Using the relation  $E_{D,m} = \lambda \tilde{E}_{D,m}$ , the gap between two such energy eigenstates is given by

$$E_{D,m} - E_{D,m'} = \lambda(E_{D,m}^{(2)} - E_{D,m'}^{(2)}) + \mathcal{O}(\lambda^2), \quad (\text{B9})$$

where  $\mathcal{O}(\lambda^2)$  refer to second- or higher-order terms in  $\lambda$ . Equation (B9) shows that the energy gap vanishes in the limit of very strong disorder. The fact that these eigenstates have similar energy eliminates all dynamics at small times. To illustrate this point, consider multiple eigenstates with significant support on the same product states  $|E_{D,m}\rangle \approx \sum_n c_{D,mn} |D, n\rangle$ . The system is initialized as one of these product states  $|\psi(0)\rangle = |D, n\rangle$  and subsequently time evolved  $|\psi(t)\rangle = e^{-i\hat{H}t} |\psi(0)\rangle$ . The time evolved state only differs slightly from the initial state at small times, i.e.,  $|\langle \psi(0) | \psi(t) \rangle|^2 \approx 1$ .

### 3. Accuracy of the perturbative approach

The analysis above is, strictly speaking, only true in the limit  $\lambda \rightarrow 0$  or alternatively  $W \rightarrow \infty$ . However, we expect the results to be a good approximation for large, finite disorder strength. We verify the formulas presented in Sec. B1 by computing the zero-order coefficients  $c_{D,mn}^{(0)}$  using Eqs. (B5)–(B8). The corresponding zero-order eigenstates are denoted  $|E_{D,m}^{(0)}\rangle = \sum_n c_{D,mn}^{(0)} |D, n\rangle$ . Furthermore, we determine the true eigenstates  $|E_i\rangle$  at different disorder strengths using exact diagonalization. For each  $|E_{D,m}^{(0)}\rangle$ , we determine the most similar exact eigenstate, i.e.,  $\text{argmax}_i (|\langle E_{D,m}^{(0)} | E_i \rangle|^2)$ , and compute the “error”

$$\varepsilon_{D,m} = 1 - \max_i (|\langle E_{D,m}^{(0)} | E_i \rangle|^2). \quad (\text{B10})$$

The distribution of  $\varepsilon_{D,m}$  is determined from  $10^3$  disorder realizations. The error is displayed as a function of disorder strength in Fig. 11. At weak disorder, the error is near its maximum value  $\varepsilon_{D,m} \approx 1$  since all exact eigenstates are delocalized. The error decreases with increasing disorder strength

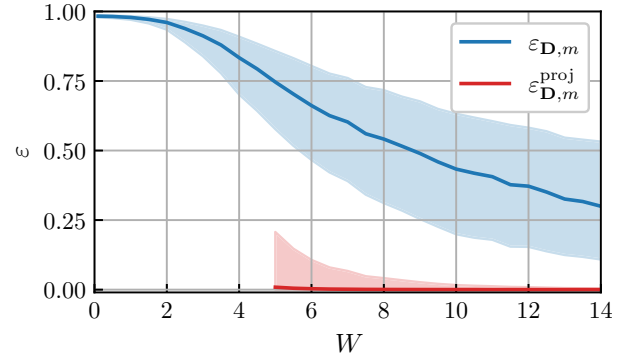


FIG. 11. The error measures  $\varepsilon_{D,m}$  (blue upper curve) and  $\varepsilon_{D,m}^{\text{proj}}$  (red lower curve) as a function of disorder strength  $W$  for system size  $L_x \times L_y = 4 \times 3$ , parameters  $J = c = \mu = 1$ , and symmetry sector  $M = 0$ . We consider the distribution of the error measures over the full spectrum in  $10^3$  disorder realizations. The figure displays the median (solid line) and the interquartile range (shaded area) of the distribution of each error measure. Note that  $\varepsilon_{D,m}^{\text{proj}}$  can only be determined at strong disorder  $5 \leq W$  where the disorder indices of energy eigenstates are well defined. The error  $\varepsilon_{D,m}$  decreases with increasing disorder strength but remains finite due to support on other subspaces  $\mathcal{V}_{D'}$  with  $D' \neq D$ . However, the error  $\varepsilon_{D,m}$  will vanish at large enough disorder. The quantity  $\varepsilon_{D,m}^{\text{proj}}$  is generally close to zero indicating that the zero-order eigenstates correctly describe the spectrum.

but remains strictly larger than zero even at very strong disorder. The error remains finite since all exact eigenstates have a small overlap with other subspaces  $\mathcal{V}_{D'}$  with  $D' \neq D$ . We expect the error to vanish for large enough disorder strength.

We further investigate the validity of our approach by considering another error measure. In the following discussion, we consider sufficiently strong disorder  $5 \leq W$  so each exact eigenstate has well-defined disorder indices  $D$ . Let  $|E_{D,m}^{\text{proj}}\rangle$  be the normalized projection of  $|E_{D,m}\rangle$  onto the subspace  $\mathcal{V}_D$ :

$$|E_{D,m}^{\text{proj}}\rangle = \frac{\hat{P}|E_{D,m}\rangle}{\sqrt{\langle E_{D,m} | \hat{P} | E_{D,m} \rangle}}, \quad (\text{B11})$$

where  $\hat{P}$  is the projection onto  $\mathcal{V}_D$ . We consider the fidelity between  $|E_{D,m}^{\text{proj}}\rangle$  and the matching zero-order eigenstate from perturbation theory  $|\langle E_{D,m}^{(0)} | E_{D,m}^{\text{proj}} \rangle|^2$ . We study the error

$$\varepsilon_{D,m}^{\text{proj}} = 1 - |\langle E_{D,m}^{(0)} | E_{D,m}^{\text{proj}} \rangle|^2. \quad (\text{B12})$$

The distribution of  $\varepsilon_{D,m}^{\text{proj}}$  is determined from  $10^3$  disorder realizations. Figure 11 illustrates this error as a function of disorder strength. The error approaches zero at large disorder indicating that the perturbative approach yields the correct zero-order eigenstates at strong, but finite, disorder strength.

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

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

- [3] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, *Nature (London)* **452**, 854 (2008).
- [4] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, *Adv. Phys.* **65**, 239 (2016).
- [5] 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).
- [6] D. Basko, I. Aleiner, and B. Altshuler, Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states, *Ann. Phys. (NY)* **321**, 1126 (2006).
- [7] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, *Phys. Rev. B* **75**, 155111 (2007).
- [8] P. Prelovšek, M. Mierzejewski, O. Barišič, and J. Herbrych, Density correlations and transport in models of many-body localization, *Ann. Phys. (Berlin, Ger.)* **529**, 1600362 (2017).
- [9] 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).
- [10] M. Serbyn, Z. Papić, and D. A. Abanin, Universal slow growth of entanglement in interacting strongly disordered systems, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [11] M. Serbyn, Z. Papić, and D. A. Abanin, Quantum quenches in the many-body localized phase, *Phys. Rev. B* **90**, 174302 (2014).
- [12] V. Ros and M. Müller, Remanent magnetization: Signature of many-body localization in quantum antiferromagnets, *Phys. Rev. Lett.* **118**, 237202 (2017).
- [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**, 127201 (2013).
- [14] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phenomenology of fully many-body-localized systems, *Phys. Rev. B* **90**, 174202 (2014).
- [15] M. Schulz, C. A. Hooley, R. Moessner, and F. Pollmann, Stark many-body localization, *Phys. Rev. Lett.* **122**, 040606 (2019).
- [16] E. van Nieuwenburg, Y. Baum, and G. Refael, From Bloch oscillations to many-body localization in clean interacting systems, *Proc. Natl. Acad. Sci. USA* **116**, 9269 (2019).
- [17] L. Zhang, Y. Ke, W. Liu, and C. Lee, Mobility edge of Stark many-body localization, *Phys. Rev. A* **103**, 023323 (2021).
- [18] E. Bairey, G. Refael, and N. H. Lindner, Driving induced many-body localization, *Phys. Rev. B* **96**, 020201(R) (2017).
- [19] S. Choi, D. A. Abanin, and M. D. Lukin, Dynamically induced many-body localization, *Phys. Rev. B* **97**, 100301(R) (2018).
- [20] D. S. Bhakuni, R. Nehra, and A. Sharma, Drive-induced many-body localization and coherent destruction of Stark many-body localization, *Phys. Rev. B* **102**, 024201 (2020).
- [21] M. Žnidarič, T. c. v. Prosen, and P. Prelovšek, Many-body localization in the Heisenberg  $XXZ$  magnet in a random field, *Phys. Rev. B* **77**, 064426 (2008).
- [22] A. Pal and D. A. Huse, Many-body localization phase transition, *Phys. Rev. B* **82**, 174411 (2010).
- [23] 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).
- [24] P. Sierant, D. Delande, and J. Zakrzewski, Many-body localization due to random interactions, *Phys. Rev. A* **95**, 021601(R) (2017).
- [25] G. Lemut, M. Mierzejewski, and J. Bonča, Complete many-body localization in the  $t$ - $J$  model caused by a random magnetic field, *Phys. Rev. Lett.* **119**, 246601 (2017).
- [26] E. Chertkov, B. Villalonga, and B. K. Clark, Numerical evidence for many-body localization in two and three dimensions, *Phys. Rev. Lett.* **126**, 180602 (2021).
- [27] 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**, 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] 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).
- [30] K. Xu, J.-J. Chen, Y. Zeng, Y.-R. Zhang, C. Song, W. Liu, Q. Guo, P. Zhang, D. Xu, H. Deng, K. Huang, H. Wang, X. Zhu, D. Zheng, and H. Fan, Emulating many-body localization with a superconducting quantum processor, *Phys. Rev. Lett.* **120**, 050507 (2018).
- [31] A. Rubio-Abadal, J. Y. Choi, J. Zeiher, S. Hollerith, J. Rui, I. Bloch, and C. Gross, Many-body delocalization in the presence of a quantum bath, *Phys. Rev. X* **9**, 041014 (2019).
- [32] J. Z. Imbrie, Diagonalization and many-body localization for a disordered quantum spin chain, *Phys. Rev. Lett.* **117**, 027201 (2016).
- [33] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Ergodicity breaking transition in finite disordered spin chains, *Phys. Rev. B* **102**, 064207 (2020).
- [34] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, *Phys. Rev. E* **102**, 062144 (2020).
- [35] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Slow delocalization of particles in many-body localized phases, *Phys. Rev. B* **103**, 024203 (2021).
- [36] D. Abanin, J. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. Parameswaran, F. Pollmann, A. Potter, M. Serbyn, and R. Vasseur, Distinguishing localization from chaos: Challenges in finite-size systems, *Ann. Phys. (NY)* **427**, 168415 (2021).
- [37] 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).
- [38] M. Serbyn, D. A. Abanin, and Z. Papić, Quantum many-body scars and weak breaking of ergodicity, *Nat. Phys.* **17**, 675 (2021).
- [39] S. Moudgalya, B. A. Bernevig, and N. Regnault, Quantum many-body scars and Hilbert space fragmentation: A review of exact results, *Rep. Prog. Phys.* **85**, 086501 (2022).
- [40] A. Chandran, T. Iadecola, V. Khemani, and R. Moessner, Quantum many-body scars: A quasiparticle perspective, *Annu. Rev. Condens. Matter Phys.* **14**, 443 (2023).
- [41] D. P. Arovas, Two exact excited states for the  $S = 1$  AKLT chain, *Phys. Lett. A* **137**, 431 (1989).

- [42] S. Moudgalya, N. Regnault, and B. A. Bernevig, Entanglement of exact excited states of Affleck-Kennedy-Lieb-Tasaki models: Exact results, many-body scars, and violation of the strong eigenstate thermalization hypothesis, *Phys. Rev. B* **98**, 235156 (2018).
- [43] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Weak ergodicity breaking from quantum many-body scars, *Nat. Phys.* **14**, 745 (2018).
- [44] M. Schecter and T. Iadecola, Weak ergodicity breaking and quantum many-body scars in spin-1  $XY$  magnets, *Phys. Rev. Lett.* **123**, 147201 (2019).
- [45] T. Iadecola and M. Schecter, Quantum many-body scar states with emergent kinetic constraints and finite-entanglement revivals, *Phys. Rev. B* **101**, 024306 (2020).
- [46] D. K. Mark and O. I. Motrunich,  $\eta$ -pairing states as true scars in an extended Hubbard model, *Phys. Rev. B* **102**, 075132 (2020).
- [47] S. Moudgalya, E. O'Brien, B. A. Bernevig, P. Fendley, and N. Regnault, Large classes of quantum scarred Hamiltonians from matrix product states, *Phys. Rev. B* **102**, 085120 (2020).
- [48] N. Shibata, N. Yoshioka, and H. Katsura, Onsager's scars in disordered spin chains, *Phys. Rev. Lett.* **124**, 180604 (2020).
- [49] C. M. Langlett, Z.-C. Yang, J. Wildeboer, A. V. Gorshkov, T. Iadecola, and S. Xu, Rainbow scars: From area to volume law, *Phys. Rev. B* **105**, L060301 (2022).
- [50] K. Lee, R. Melendrez, A. Pal, and H. J. Changlani, Exact three-colored quantum scars from geometric frustration, *Phys. Rev. B* **101**, 241111(R) (2020).
- [51] E. Chertkov and B. K. Clark, Motif magnetism and quantum many-body scars, *Phys. Rev. B* **104**, 104410 (2021).
- [52] J. Wildeboer, A. Seidel, N. S. Srivatsa, A. E. B. Nielsen, and O. Erten, Topological quantum many-body scars in quantum dimer models on the kagome lattice, *Phys. Rev. B* **104**, L121103 (2021).
- [53] N. Shiraishi and T. Mori, Systematic construction of counterexamples to the eigenstate thermalization hypothesis, *Phys. Rev. Lett.* **119**, 030601 (2017).
- [54] S. Moudgalya, N. Regnault, and B. A. Bernevig,  $\eta$ -pairing in Hubbard models: From spectrum generating algebras to quantum many-body scars, *Phys. Rev. B* **102**, 085140 (2020).
- [55] N. O'Dea, F. Burnell, A. Chandran, and V. Khemani, From tunnels to towers: Quantum scars from Lie algebras and  $q$ -deformed Lie algebras, *Phys. Rev. Res.* **2**, 043305 (2020).
- [56] K. Pakrouski, P. N. Pallegar, F. K. Popov, and I. R. Klebanov, Many-body scars as a group invariant sector of Hilbert space, *Phys. Rev. Lett.* **125**, 230602 (2020).
- [57] J. Ren, C. Liang, and C. Fang, Quasisymmetry groups and many-body scar dynamics, *Phys. Rev. Lett.* **126**, 120604 (2021).
- [58] S. Moudgalya and O. I. Motrunich, Exhaustive characterization of quantum many-body scars using commutant algebras, [arXiv:2209.03377](https://arxiv.org/abs/2209.03377).
- [59] J. Ren, C. Liang, and C. Fang, Deformed symmetry structures and quantum many-body scar subspaces, *Phys. Rev. Res.* **4**, 013155 (2022).
- [60] P.-G. Rozon and K. Agarwal, Broken unitary picture of dynamics in quantum many-body scars, [arXiv:2302.04885](https://arxiv.org/abs/2302.04885).
- [61] B. Buča, Unified theory of local quantum many-body dynamics: Eigenoperator thermalization theorems, *Phys. Rev. X* **13**, 031013 (2023).
- [62] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Probing many-body dynamics on a 51-atom quantum simulator, *Nature (London)* **551**, 579 (2017).
- [63] D. Bluvstein, A. Omran, H. Levine, A. Keesling, G. Semeghini, S. Ebadi, T. T. Wang, A. A. Michailidis, N. Maskara, W. W. Ho, S. Choi, M. Serbyn, M. Greiner, V. Vuletić, and M. D. Lukin, Controlling quantum many-body dynamics in driven Rydberg atom arrays, *Science* **371**, 1355 (2021).
- [64] I. C. Chen, B. Burdick, Y. Yao, P. P. Orth, and T. Iadecola, Error-mitigated simulation of quantum many-body scars on quantum computers with pulse-level control, *Phys. Rev. Res.* **4**, 043027 (2022).
- [65] P. Zhang, H. Dong, Y. Gao, L. Zhao, J. Hao, J.-Y. Desaulles, Q. Guo, J. Chen, J. Deng, B. Liu, W. Ren, Y. Yao, X. Zhang, S. Xu, K. Wang, F. Jin, X. Zhu, B. Zhang, H. Li, C. Song *et al.*, Many-body Hilbert space scarring on a superconducting processor, *Nat. Phys.* **19**, 120 (2023).
- [66] E. J. Gustafson, A. C. Y. Li, A. Khan, J. Kim, D. M. Kurkcuoglu, M. S. Alam, P. P. Orth, A. Rahmani, and T. Iadecola, Preparing quantum many-body scar states on quantum computers, *Quantum* **7**, 1171 (2023).
- [67] G.-X. Su, H. Sun, A. Hudomal, J.-Y. Desaulles, Z.-Y. Zhou, B. Yang, J. C. Halimeh, Z.-S. Yuan, Z. Papić, and J.-W. Pan, Observation of many-body scarring in a Bose-Hubbard quantum simulator, *Phys. Rev. Res.* **5**, 023010 (2023).
- [68] H. Zhou, H. Gao, N. T. Leitaó, O. Makarova, I. Cong, A. M. Douglas, L. S. Martin, and M. D. Lukin, Robust Hamiltonian engineering for interacting qudit systems, [arXiv:2305.09757](https://arxiv.org/abs/2305.09757).
- [69] N. S. Srivatsa, R. Moessner, and A. E. B. Nielsen, Many-body delocalization via emergent symmetry, *Phys. Rev. Lett.* **125**, 240401 (2020).
- [70] M. Iversen, N. S. Srivatsa, and A. E. B. Nielsen, Escaping many-body localization in an exact eigenstate, *Phys. Rev. B* **106**, 214201 (2022).
- [71] N. S. Srivatsa, H. Yarloo, R. Moessner, and A. E. B. Nielsen, Mobility edges through inverted quantum many-body scarring, *Phys. Rev. B* **108**, L100202 (2023).
- [72] M. Iversen and A. E. B. Nielsen, Tower of quantum scars in a partially many-body localized system, *Phys. Rev. B* **107**, 205140 (2023).
- [73] Q. Chen and Z. Zhu, Inverting multiple quantum many-body scars via disorder, [arXiv:2301.03405](https://arxiv.org/abs/2301.03405).
- [74] P. Kolb and K. Pakrouski, Stability of the many-body scars in fermionic spin-1/2 models, *PRX Quantum* **4**, 040348 (2023).
- [75] G. Ramírez, J. Rodríguez-Laguna, and G. Sierra, From conformal to volume law for the entanglement entropy in exponentially deformed critical spin 1/2 chains, *J. Stat. Mech.: Theory Exp.* (2014) P10004.
- [76] G. Vitagliano, A. Riera, and J. I. Latorre, Volume-law scaling for the entanglement entropy in spin-1/2 chains, *New J. Phys.* **12**, 113049 (2010).
- [77] W. De Roeck and F. Huveneers, Stability and instability towards delocalization in many-body localization systems, *Phys. Rev. B* **95**, 155129 (2017).
- [78] W. De Roeck and J. Z. Imbrie, Many-body localization: stability and instability, *Philos. Trans. R. Soc. London A* **375**, 20160422 (2017).
- [79] M. Serbyn and J. E. Moore, Spectral statistics across the many-body localization transition, *Phys. Rev. B* **93**, 041424(R) (2016).

- [80] 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**, 084101 (2013).
- [81] D. N. Page, Average entropy of a subsystem, *Phys. Rev. Lett.* **71**, 1291 (1993).
- [82] E. Bianchi and P. Donà, Typical entanglement entropy in the presence of a center: Page curve and its variance, *Phys. Rev. D* **100**, 105010 (2019).
- [83] E. Bianchi, L. Hackl, M. Kieburg, M. Rigol, and L. Vidmar, Volume-law entanglement entropy of typical pure quantum states, *PRX Quantum* **3**, 030201 (2022).
- [84] B. Bauer and C. Nayak, Area laws in a many-body localized state and its implications for topological order, *J. Stat. Mech.: Theory Exp.* (2013) P09005.
- [85] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Many-body localization in a disordered quantum Ising chain, *Phys. Rev. Lett.* **113**, 107204 (2014).
- [86] W. Cottrell, B. Freivogel, D. M. Hofman, and S. F. Lokhande, How to build the thermofield double state, *J. High Energy Phys.* **02** (2019) 058.
- [87] J. Wu and T. H. Hsieh, Variational thermal quantum simulation via thermofield double states, *Phys. Rev. Lett.* **123**, 220502 (2019).
- [88] D. Zhu, S. Johri, N. M. Linke, K. A. Landsman, C. H. Alderete, N. H. Nguyen, A. Y. Matsuura, T. H. Hsieh, and C. Monroe, Generation of thermofield double states and critical ground states with a quantum computer, *Proc. Natl. Acad. Sci. USA* **117**, 25402 (2020).
- [89] J. Wildeboer, C. M. Langlett, Z.-C. Yang, A. V. Gorshkov, T. Iadecola, and S. Xu, Quantum many-body scars from Einstein-Podolsky-Rosen states in bilayer systems, *Phys. Rev. B* **106**, 205142 (2022).
- [90] J. J. Sakurai, *Modern Quantum Mechanics; Rev. Ed.* (Addison-Wesley, Reading, 1994).