

PHD DISSERTATION

Ergodicity breaking in quantum systems: from exact time evolution to machine learning

Tomasz Szołdra



Supervisor: prof. dr hab. Jakub Zakrzewski

FACULTY OF PHYSICS, ASTRONOMY AND APPLIED COMPUTER SCIENCE
JAGIELLONIAN UNIVERSITY
KRAKÓW, POLAND

SEPTEMBER 2024

ABSTRACT

Generic quantum many-body systems have thermal eigenstates, as postulated by the Eigenstate Thermalization Hypothesis (ETH). Because of that, long-time values of the observables in a system initialized at a finite energy density are described by a microcanonical ensemble and depend only on the energy, while the memory of the details of the initial state is lost. However, some systems violate the ETH and do not thermalize. Examples include integrable, many-body localized (MBL), and quantum many-body-scarred (QMBS) models, as well as models with a fragmented Hilbert space. This work proposes new numerical tools useful for studying ergodicity breaking in quantum systems.

The transition between the MBL and ergodic phases is hypothesized to be governed by the so-called quantum avalanche mechanism. An initial ergodic bubble, caused by an accidental fluctuation in the system properties, can interact with its surroundings, expand and eventually thermalize everything around. We perform a detailed study of the quantum avalanche mechanism in the disordered XXZ spin-1/2 model. We propose a machine learning algorithm based on autoencoder neural networks that allows us to detect ergodic bubbles with a spatiotemporal resolution using an experimentally measurable two-body correlation function. With this detection mechanism, we study the distribution of bubble sizes in different regimes of the model, and identify rare thermalization events. In a further work, we investigate the influence of ergodic bubbles on the MBL phase by deliberately planting an ergodic bubble on one side of the system. Utilizing two-body correlation functions, we show that the observation of the termination of quantum avalanches requires a fine-tuning of the initial state, even in the deeply localized regime of the model. We also confirm that an ergodic bubble modeled by a Hamiltonian from the Gaussian Orthogonal Ensemble of random matrices, as often assumed in the theory of avalanches, has a similar effect on the system as a real bubble. The third work concentrates on the proposition of a new probe of quantum information propagation. We employ time evolutions of an autocorrelation function for a single lattice site and look at the difference between its measurement in the system of size L and its restriction to $\ell < L$ sites. The probe is benchmarked in ergodic, Anderson localized, and MBL regimes of the disordered XXZ model, and compared to other existing tools.

In quantum many-body scarred systems, almost all eigenstates are thermal, but there exists a small number of special, low-entanglement eigenstates across the entire spectrum. The Hilbert space is split into (approximately) disconnected sectors, corresponding to differ-

ent families of scarred eigenstates. In the last work, we introduce a quantum machine learning algorithm based on quantum variational autoencoders, that can detect different families of scars. We identify previously overlooked families of scars in the PXP model and demonstrate a successful application of our algorithm to the spin ladder model with a fragmented Hilbert space.

STRESZCZENIE

Hamiltonian typowego, wielociałowego układu kwantowego posiada termalne stany własne, jak przewiduje hipoteza o termalizacji stanów własnych (ETH). Powoduje to, że w układzie rozpoczynającym ewolucję ze stanu o skończonej gęstości energii, długoczasowe wartości oczekiwane obserwabli są opisywane przez zespół mikrokanoniczny i zależą jedynie od wartości energii, zaś pamięć o szczegółowych własnościach stanu początkowego szybko zanika. Istnieją natomiast układy, które nie spełniają ETH i nie ulegają termalizacji. Wśród przykładów znajdują się modele całkowalne, zlokalizowane wielociałowo (MBL), modele z wielociałowymi bliznami kwantowymi (QMBS), a także z pofragmentowaną przestrzenią Hilberta. W niniejszej pracy przedstawiamy nowe narzędzia numeryczne służące do badań nad procesem łamania ergodyczności w układach kwantowych.

Przyjmuje się, że za przejście pomiędzy fazą MBL oraz fazą ergodyczną jest odpowiedzialny mechanizm tzw. lawiny kwantowej. Początkowy bąbel ergodyczny, wytworzony przez losową fluktuację własności układu, może oddziaływać z otoczeniem, rosnąć, a w konsekwencji termalizować resztę układu. Szczegółowo badamy mechanizm lawiny kwantowej w modelu XXZ ze spinem $1/2$ oraz nieporządkiem. Proponujemy algorytm uczenia maszynowego oparty na sieciach neuronowych o architekturze autoenkodera, który pozwala nam na wykrywanie bąbli ergodycznych w czasie i przestrzeni z użyciem jedynie dwuciałowych funkcji korelacji mierzalnych eksperymentalnie. Używając powyższej metody, badamy rozkłady wielkości bąbli w różnych reżimach modelu, a także identyfikujemy rzadkie zdarzenia, w których następuje termalizacja. W kolejnej pracy pochylamy się nad wpływem bąbli ergodycznych na fazę MBL, w sposób kontrolowany umieszczając bąbel ergodyczny po jednej stronie układu. Wykorzystując dwuciałowe funkcje korelacji, pokazujemy, że aby zaobserwować zatrzymanie się lawin kwantowych należy precyzyjnie przygotować stan początkowy, nawet w silnie zlokalizowanym reżimie modelu. Potwierdzamy także, że bąbel ergodyczny modelowany przez Hamiltonian będący ortogonalną macierzą losową z rozkładu Gaussa, jaki zwykle zakłada się w teoriach opisujących lawiny kwantowe, ma bardzo podobny wpływ na zachowanie układu, jak realny bąbel ergodyczny. Trzecia praca koncentruje się na propozycji nowego sposobu pomiaru propagacji informacji kwantowej. Porównujemy ewolucje czasowe funkcji autokorelacji pojedynczego spinu, wyliczając ich różnicę między układem z L spinami oraz jego ograniczeniem do $\ell < L$ spinów. Zachowanie się proponowanej wielkości analizujemy w reżimie ergodycznym, z lokalizacją Andersona oraz MBL modelu

XXZ z nieporządkiem, oraz porównujemy je do wyników uzyskanych za pomocą innych narzędzi.

W układach z wielociałowymi bliznami kwantowymi prawie wszystkie stany własne są termalne, ale na przestrzeni całego spektrum istnieje też niewielka liczba specjalnych stanów własnych o niskim splątaniu. Przestrzeń Hilberta jest podzielona na (prawie) niepołączone ze sobą sektory, odpowiadające kolejnym rodzinom blizn kwantowych. W niniejszej pracy wprowadzamy algorytm kwantowego uczenia maszynowego oparty o kwantowy autoenkoder wariacyjny, dzięki któremu możliwe staje się wykrywanie rodzin blizn kwantowych. Identyfikujemy pominięte wcześniej rodziny blizn w modelu PXP i pokazujemy skuteczne działanie naszego algorytmu na przykładzie modelu drabiny spinowej z pofragmentowaną przestrzenią Hilberta.

ACKNOWLEDGMENTS

First and foremost, I want to warmly thank my supervisor, Jakub Zakrzewski, who agreed to guide me through my research journey over the past four years. I truly enjoyed learning from him and working together in our quest to understand complex quantum many-body phenomena. His exceptional passion, enthusiasm, and sense of humor greatly inspired and influenced me, both as an early-stage researcher and as a person.

I extend my sincere gratitude to my long-time collaborators from the Institute of Photonic Sciences in Castelldefels (Barcelona) – Maciej Lewenstein and Piotr Sierant – with whom all the results comprising this thesis were obtained. Maciej’s broad expertise, experience, and ability to provide constructive feedback were invaluable during times of struggle. I am especially grateful to Piotr Sierant for sharing his expert-level insights on many different topics in computational physics. He was always ready to assist with any technical problems I encountered and contributed with his deep understanding of non-equilibrium physics during our discussions of the results. I am also thankful to Korbinian Kottmann, who contributed to one of the projects in this thesis with his profound knowledge and experience in machine learning. Additionally, I would like to thank the collaborators with whom I had the pleasure of working on a project on machine learning in femtosecond physics, which is not included in this thesis: Philip H. Bucksbaum, Marcelo F. Ciappina, Andrew S. Maxwell, and Nicholas Werby.

I greatly appreciate the technical support provided by Roman Marcinek. Finally, I would like to thank my colleagues at the Jagiellonian University in Kraków – Adith, Ali, Arindam, Konrad, Krzysiek, Kuba, Mateusz, Mykhailo, Pedro, Piotr, Sasha, Staszek, Titas, and Weronika – whose constant openness and friendliness made the work environment very pleasant.

Na końcu chciałbym podziękować moim rodzicom, Marcie i Piotrowi, za ogromne wsparcie, które dostawałem od nich przez całe życie, i bez którego napisanie tej pracy nie byłoby możliwe. Słowa wdzięczności kieruję także do mojej kochanej żony, Magdy, która w najtrudniejszych chwilach była dla mnie nieocenionym oparciem.

Contents

1	Introduction	15
1.1	Quantum chaos and the random matrix theory	15
1.2	Thermalization, ergodicity and the Eigenstate Thermalization Hypothesis	18
1.3	Breakdown of ergodicity	19
2	Many-body localization	23
2.1	Introduction	23
2.1.1	Definition of many-body localization	24
2.1.2	The disordered XXZ model	25
2.1.3	Status of MBL: major problems	25
2.2	Article 1	30
2.3	Article 2	31
2.4	Article 3	33
3	Quantum many-body scars	35
3.1	Introduction	35
3.2	The PXP model	36
3.3	Article 4	37
4	Conclusions	41
A	Exact time evolution based on the Chebyshev polynomial expansion	45
	Articles	57

LIST OF ACRONYMS

RMT	Random matrix theory
GOE	Gaussian orthogonal ensemble
GUE	Gaussian unitary ensemble
EE	Entanglement entropy
ETH	Eigenstate thermalization hypothesis
AL	Anderson localization
MBL	Many-body localization
LIOM	Local integral of motion
QMBS	Quantum many-body scars
ED	Exact diagonalization
TL	Thermodynamic limit
CFD	Correlation function difference
QVAE	Quantum variational autoencoder

PREFACE

A generic isolated many-body quantum system reaches thermal equilibrium, acting as its own thermal bath, and being accurately described by a microcanonical ensemble. However, some systems violate this rule and exhibit non-ergodic behavior. There are examples of interacting quantum systems that are integrable and do not thermalize, but only in the fine-tuning scenario and small perturbations destroy integrability. In the presence of disorder, some finite interacting systems are in the many-body localized regime, characterized by a complete set of local integrals of motion. Another class of non-ergodic systems are those with fragmented Hilbert spaces, in which disconnected parts of the Hilbert space emerge and the system initialized in one of such sectors cannot explore the other sectors. Some systems thermalize or not, depending on the choice of the initial state, such as those hosting quantum many-body scars.

This Introduction aims to serve as a guide for the reader of the four papers on ergodicity breaking in quantum many-body systems, which constitute the PhD Thesis and are attached at the end:

- [1] Tomasz Szȩdra, Piotr Sierant, Korbinian Kottmann, Maciej Lewenstein, and Jakub Zakrzewski, *"Detecting ergodic bubbles at the crossover to many-body localization using neural networks"*, Phys. Rev. B **104**, L140202 (2021)
- [2] Tomasz Szȩdra, Piotr Sierant, Maciej Lewenstein, and Jakub Zakrzewski, *"Catching thermal avalanches in the disordered XXZ model"*, Phys. Rev. B **109**, 134202 (2024)
- [3] Tomasz Szȩdra, Piotr Sierant, Maciej Lewenstein, and Jakub Zakrzewski, *"Tracking locality in the time evolution of disordered systems"*, Phys. Rev. B **107**, 054204 (2023)
- [4] Tomasz Szȩdra, Piotr Sierant, Maciej Lewenstein, and Jakub Zakrzewski, *"Unsupervised detection of decoupled subspaces: Many-body scars and beyond"*, Phys. Rev. B **105**, 224205 (2022)

The Introduction covers elementary concepts in the field of thermalization and ergodicity breaking, omitted in the publications due to the lack of space and the typical article format. The physical models considered in the publications are also briefly outlined. The results obtained in each of the papers are summarized. Additional commentaries aim to put these works into a broader literature context and propose new, promising research directions.

This Introduction is organized as follows. In Ch. 1 we start with the basic concepts of thermalization, ergodicity, and ergodicity breaking.

We begin Ch. 2 with a brief description of many-body localization and the disordered XXZ model and proceed to Articles [1, 2], which make a step forward to answer the question of the robustness of the MBL phase to thermalization in the thermodynamic limit against the putative quantum avalanche scenario. We also cover the topic of the propagation of quantum information in quantum many-body systems and present Article [3], proposing a novel probe of information spreading based on the differences of single-body correlation functions in the system and its enlarged copy. Chapter 3 introduces quantum many-body scars and the PXP model. We summarize Article [4], in which we design a quantum machine learning algorithm for detecting clusters of similar, scarred eigenstates in an otherwise thermalizing spectrum. Several comments about related works are made. We conclude in Ch. 4.

Throughout this work we assume natural units with $\hbar = 1$.

INTRODUCTION

The *ergodic hypothesis* by Boltzmann [5], dating back more than 120 years, states that when the classical statistical system reaches equilibrium, it visits the whole phase space during in the time evolution. The time spent in each part of the phase space is proportional to its volume, defining a statistical ensemble. When the long-time observables become equivalent to the ensemble averages, the system is said to have *thermalized*.

Quantum thermalization, ergodicity, and, in particular, ergodicity breaking, are the key ideas in Articles [1–4] which constitute the PhD Thesis. Nevertheless, a typical research article is written for specialists, assuming prior knowledge of the basic ideas in the field. This Chapter aims to fill this gap and introduce the reader into the concepts of thermalization and ergodicity in quantum mechanics. We start with the effective description of quantum chaos through the random matrix theory in Sec. 1.1. In Sec. 1.2 we introduce the notion of quantum thermalization and the eigenstate thermalization hypothesis. Ways to break ergodicity in quantum systems are briefly discussed in Sec. 1.3.

1.1 QUANTUM CHAOS AND THE RANDOM MATRIX THEORY

Works by Wigner (and later of Dyson, Mehta and others) in the 1950's and 1960's [6–9] tackled the problem of explaining the properties of the energy spectra of complex heavy nuclei. Wigner quickly realized that it is quite hopeless to understand these systems from a microscopic point of view by calculating the exact energy levels and the Hamiltonian eigenstates. Instead, he concentrated on the statistical properties of the spectrum, drawing the first important connection between the quantum systems and the random matrix theory. Wigner discovered that if one considers a small part of the spectrum with a constant density of states, in a generic non-fine-tuned basis the Hamiltonian looks like a random matrix.

Random matrix theory (RMT) defines a random ensemble $P(H)$ of square matrices H . Let us restrict to the case when the $N \times N$ matrices are real and symmetric (called Gaussian orthogonal ensemble (GOE), denoted by $\beta = 1$ and corresponding to Hamiltonians with time-reversal symmetry) or hermitian (Gaussian unitary ensemble (GUE),

$\beta = 2$, without time-reversal symmetry). For the sake of simplicity, let us consider only the GOE which is uniquely defined by the following requirements [10]:

1. The ensemble is invariant under every *orthogonal* transformation

$$H' = W^T H W, \quad P(H') dH' = P(H) dH, \quad (1.1)$$

where W is any real orthogonal matrix and the measure is $dH = 2^{N(N-1)/4} \prod_{1 \leq i \leq N} dH_{ii} \prod_{1 \leq i < j \leq N} dH_{ij}$.

2. The matrix elements H_{ij} (with $i \leq j$) are independent random variables.

Point 1. assures that no specific basis is favorable, hence all states behave in the same way when acted upon by the matrix H . Point 2. gives a possibility for an analytic solution of many problems in the RMT. Similarly, one can define a GUE ensemble by requiring invariance under *unitary* transformation. The conditions 1., 2. uniquely determine the probability distribution:

$$P(H) = \frac{1}{Z} \exp \left(-\frac{\beta}{2a^2} \text{Tr}(H^2) \right) \quad (1.2)$$

where a sets the overall energy scale and Z is the normalization constant. Indeed, due to the presence of the trace, the probability is invariant under any orthogonal/unitary rotations of the matrix H . Within the RMT, one can calculate statistical properties of the eigenvalues and the eigenvectors of the Hamiltonian.

EIGENVALUES The adjacent energy level gaps $s_n = E_{n+1} - E_n$ come from the Wigner-Dyson distribution

$$P(s) = A_\beta s^\beta \exp(-B_\beta s^2), \quad (1.3)$$

with A_β, B_β found by normalizing the probability distribution and fixing the mean level spacing to unity [7]. Eq. (1.3) was derived exactly for 2×2 matrices and approximates the gaps distribution very well also for large matrices [11]. In 1984, Bohigas, Giannoni and Schmit (BGS) [12] conjectured that the level statistics of *generic* quantum systems that have a chaotic classical limit are described by the RMT and the Wigner-Dyson distribution. This was indeed confirmed in heavy nuclei [13], quantum chaotic billiards [14], hydrogen atom in a strong magnetic field [15], spin lattices [16] and many other models.

Calculating the level spacing statistics in practice requires a procedure of *level unfolding* (renormalizing the mean energy density to unity) which can be quite challenging for real spectra where the lo-

cal energy density fluctuates [16]. To deal with this problem, one computes the gap ratio [17] defined as

$$r_n = \frac{\min(s_n, s_{n+1})}{\max(s_n, s_{n+1})}, \quad (1.4)$$

which is not sensitive to the local density of states. For the GOE, the mean gap ratio reads $\bar{r}_{\text{GOE}} \approx 0.5307$ [18]. For an integrable system, the energy gaps are described by the Poisson distribution $P(s) \propto e^{-Cs}$ and the mean gap ratio is $\bar{r}_{\text{PS}} \approx 0.386294$ [19].

EIGENSTATES Let us discuss the structure of the eigenvectors of H . For brevity, we concentrate on the GOE random ensemble. Since any two Hamiltonians that are related by an orthogonal transformation belong to the same ensemble, there can be no structure in the eigenvectors of H . The eigenvectors of H are just random unit vectors $(v_1, v_2, \dots, v_{\mathcal{D}})$ in the Hilbert space of dimension $\mathcal{D}$, satisfying the normalization condition

$$P(v_1, v_2, \dots, v_{\mathcal{D}}) \propto \delta(\sum_i v_i^2 - 1). \quad (1.5)$$

One of the central concepts in quantum many-body physics is the entanglement, defined for a given bipartition $\mathcal{A}, \mathcal{B}$ of the whole system $\mathcal{AB}$. It can be quantified by the von Neumann bipartite entanglement entropy (EE)

$$\mathcal{S}_{\text{vN}} = -\text{Tr}(\rho_{\mathcal{A}} \ln \rho_{\mathcal{A}}), \quad (1.6)$$

where $\rho_{\mathcal{A}}$ is the reduced density matrix for subsystem $\mathcal{A}$,

$$\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}} |\psi\rangle \langle \psi|, \quad (1.7)$$

and $|\psi\rangle$ is the quantum state. Let us concentrate on one-dimensional systems. The RMT predicts that eigenstates entanglement entropy is proportional to the volume of the subsystem size $L_{\mathcal{A}}$ as $\mathcal{S}_{\text{vN}} = s(E)L_{\mathcal{A}} + C$, which is called *volume law*. Global conservation laws such as total magnetization affect the constant C [20]. On the other hand, for example for the groundstates of generic gapped Hamiltonians, the EE saturates at a constant value, $\mathcal{S}_{\text{vN}} \approx \text{const}$, which is known as the *area law*¹. Specifically, for an equal bipartition, $L_{\mathcal{A}} = L/2$, the average entropy reaches the maximal Page value [21], $\mathcal{S}_{\text{vN, RMT}} = \ln \mathcal{D}_{\mathcal{A}} - \mathcal{D}_{\mathcal{A}}^2 / (2\mathcal{D}) = (L \ln 2 - 1)/2$, where $\mathcal{D}$ and $\mathcal{D}_{\mathcal{A}}$ denote the dimension of the full and subsystem Hilbert space, respectively.

¹ We only discussed spatial dimension $d = 1$. In a general case, volume of the system is proportional to L^d and its area to L^{d-1} . Hence, $\mathcal{S}_{\text{vN}} \propto L^d$ is the volume law, while $\mathcal{S}_{\text{vN}} \propto L^{d-1}$ is the area law.

1.2 THERMALIZATION, ERGODICITY AND THE EIGENSTATE THERMALIZATION HYPOTHESIS

Let us define thermalization in the quantum case. We choose an initial state with a mean energy² E . The system is said to thermalize if a joint probability distribution of physical observables $O(t)$ reaches a microcanonical prediction and remains close to it for almost all future time. Another definition says that the system is thermal or ergodic if the reduced density matrix of any small subsystem $\mathcal{A}$ of the whole system $\mathcal{AB}$, $\rho_{\mathcal{A}}(t) = \text{Tr}_{\mathcal{B}} |\psi(t)\rangle \langle \psi(t)|$, evolves to a Gibbs density matrix

$$\lim_{t \rightarrow \infty} \rho_{\mathcal{A}}(t) = \text{Tr}_{\mathcal{B}} \rho^{\text{eq}} \approx \rho_{\mathcal{A}}^{\text{eq}}, \quad \rho^{\text{eq}} = \frac{e^{-\beta H}}{Z}, \quad (1.8)$$

where H is the Hamiltonian, β is the inverse temperature defined by the expectation value of the energy of the initial state (there is no external bath), and Z is a partition function. Equation (1.8) implies that subsystem $\mathcal{B}$ acts as a thermal bath to subsystem $\mathcal{A}$. Thus, the long-time expectation values of operators $O(t)$ are equal to their thermal expectation values.

We can write the time evolution of the initial state $|\psi_0\rangle$ in the eigenbasis of the Hamiltonian:

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t} |n\rangle, \quad (1.9)$$

where $c_n = \langle n | \psi_0 \rangle$. A time-dependent observable $O(t)$ reads

$$O(t) = \langle \psi(t) | O | \psi(t) \rangle = \sum_n |c_n|^2 O_{nn} + \sum_{n,m \neq n} c_n^* c_m e^{i(E_m - E_n)t} O_{nm}, \quad (1.10)$$

where $O_{nm} = \langle n | O | m \rangle$. The first term in the sum corresponds to a long-time average, while the second term encodes time-dependent fluctuations and defines the timescale of equilibration. In turn, the matrix elements O_{mn} determine the thermalization properties of the system. In the 1990's, Deutsch [23] and Srednicki [24, 25] observed that the RMT, although correctly capturing the energy level statistics, does not accurately predict matrix elements O_{mn} that model thermalization in real physical systems, and proposed an extension to the RMT – the Eigenstate Thermalization Hypothesis (ETH). The ETH is a surmise for the operator matrix elements in the Hamiltonian eigenbasis,

$$O_{mn} = O(\bar{E}) \delta_{mn} + e^{-S(\bar{E})/2} f_O(\bar{E}, \omega) R_{mn}, \quad (1.11)$$

² For a typical initial product state and system of volume V , the energy width of the initial state is $\Delta \propto \sqrt{V}$, which is small compared to the energy spectrum full width scaling proportionally to V [22].

where $\bar{E} = (E_m + E_n)/2$, $\omega = E_n - E_m$, $S(E)$ is the thermodynamic entropy at energy E (i.e., $S(E) = \ln \rho(E)$ where $\rho(E)$ is the density of states), and, importantly, $O(\bar{E})$ and $f_O(\bar{E}, \omega)$ are both smooth functions of their arguments and do not scale with the system size. R_{mn} is a random variable with zero mean and unit variance. It is generally expected that all physical observables (i.e., observables for which statistical mechanics applies) satisfy the ETH Ansatz (1.11). There is no restriction on the locality of operators but operators such as projectors on the system eigenstates $P_n = |n\rangle \langle n|$ are excluded [14]. The long-time expectation value of the average of the operator $O(t)$, Eq. (1.10), is the microcanonical expectation value $O(E)$, i.e., the ETH predicts thermalization. For a review on thermalization and the ETH, see Ref. [14].

1.3 BREAKDOWN OF ERGODICITY

Although the ETH seems to be a universal description of interacting quantum many-body systems, questions about the existence of systems that violate it immediately arise. Without the need to satisfy the constraints of statistical mechanics, non-thermalizing quantum systems can potentially uncover new, unconventional, non-equilibrium phases of matter, and provide memories to store quantum information.

We will briefly discuss the classes of ergodicity breaking in quantum systems. If the system does not go into equilibrium, it necessarily violates the diagonal part of the ETH. The system may behave differently depending on the initial state. Systems for which thermalization fails for all initial states are said to exhibit strong ETH violation or strong ergodicity breaking. For some systems, while most generic initial states thermalize, there exist a few special initial states which do not – we will refer to such a phenomenon as a weak ergodicity breaking. These thermalization properties must be reflected in the structure of the Hamiltonian eigenstates. In case of strong ergodicity breaking, all³ eigenstates are non-thermal (in the sense of not satisfying Eq. (1.8)), while in the case of weak ergodicity breaking almost all of them are thermal except a small number. Origin of the non-ergodicity in isolated many-body quantum systems can come from one of the following phenomena:

INTEGRABILITY An integrable system possesses a complete set of exponentially many integrals of motion, i.e., operators that commute with the Hamiltonian, and falls into the category of strong ergodicity breaking. In many cases, operators corresponding to the integrals of motion can be constructed analytically and the model can be solved exactly. Profound examples of exactly solvable models are XXZ chain without external disorder [26] and the Hubbard [27] model, both in

³ More precisely, all eigenstates except those near the edges of the spectrum.

one dimension. Although in classical mechanics the Kolmogorov-Arnold-Moser (KAM) theorem [28–30] guarantees the stability of the periodic trajectories in the phase space (also known as KAM tori) under small perturbations, there is no rigorous KAM theorem for quantum systems and they are generically expected to be unstable against perturbations. Thus, from an experimental perspective, all integrable interacting models are fine-tuned. Energy gaps in the spectra of integrable systems follow Poisson statistics. The eigenstates are generically volume-law entangled, so it is hard to distinguish them from thermal eigenstates using this criterion.

MANY-BODY LOCALIZATION The interplay of interactions and strong disorder leads to the appearance of a complete set of so-called local integrals of motion (LIOMs) which are almost conserved during the time evolution. This constitutes an interesting example of emergent integrability – the many-body localization (MBL), which is an analogue of Anderson localization for interacting systems. Numerical results obtained for finite systems confirm that the spectral properties of MBL Hamiltonians are characteristic to integrable systems (Poisson statistics), and that all eigenstates are non-thermal and show the area-law of entanglement.

QUANTUM MANY-BODY SCARS There are systems for which most but a few eigenstates are thermal. Special eigenstates with features not fulfilling the ETH are called quantum many-body scars (QMBS) and can appear anywhere in the spectrum. In a typical system, the number of these special eigenstates is exponentially smaller than the number of thermal eigenstates, i.e. their number grows polynomially with the system size, while the number of all eigenstates grows exponentially. Hence, these systems exhibit weak ergodicity breaking. Presence of special eigenstates does not affect the level statistics which are captured by the Wigner-Dyson distribution. Quantum scars are typically area-law entangled. In some models exhibiting scars, like the PXP model [31] described in Ch. 3, there are families of QMBS spaced almost equally in energy, giving rise to long-time revivals from specific simple initial states.

HILBERT SPACE FRAGMENTATION Presence of some dynamical constraints or hidden symmetries can lead to the emergence of dynamically disconnected subspaces of the Hilbert space, which is called Hilbert space fragmentation [32]. The presence of subspaces is not caused by any simple symmetries of the Hamiltonian, such as translation, reflection, etc. If the wave function is initialized in one of the subspaces, it stays there during the time evolution and does not spread to other subspaces. The Hamiltonian has a block-diagonal form, and the properties of the eigenstates in each block can be different: some

blocks may scale their size with the whole system and be integrable, some may fulfill the ETH, while some may have a fixed small size independent of the total Hilbert space size. Fragmented Hilbert space can be spanned by both area-law and volume-law eigenstates. The number of eigenstates which violates the ETH in the thermodynamic limit determines whether we are dealing with strong Hilbert space fragmentation (finite fraction of the Hilbert space size), or weak Hilbert space fragmentation (measure zero set). Due to the presence of non-trivial symmetries and overlapping symmetry sectors, level statistics of such systems can be different from the Wigner-Dyson and Poisson statistics [19, 33]. Sometimes the weak Hilbert space fragmentation is considered as equivalent to quantum scarring [32, 34].

MANY-BODY LOCALIZATION

2.1 INTRODUCTION

In the 1958 paper [35], P. W. Anderson showed that when non-interacting quantum particles are placed in a static disordered medium¹, the particle transport is fully inhibited. This phenomenon is now called Anderson localization (AL). The single-particle eigenstates of such a system are exponentially localized in real space, $|\psi(x)| \propto e^{-|x-x_0|/\xi_{\text{AL}}}$, where ξ_{AL} is the localization length and x_0 determines the center of mass. If a wavepacket is initially localized around a certain point, it will not spread during the time evolution on the lengthscales larger than the localization length ξ_{AL} . AL is a general wave phenomenon that, loosely speaking, happens due to destructive interference of the waves scattering on the potential barriers due to disorder. For a review, see Ref. [36].

A natural way to extend the Anderson model is to take interactions into account. In the limit of weak interactions, the localization is destroyed and the thermal eigenstates describe the system, see e. g. Ref. [37]. It was quickly realized that to see a lack of thermalization, one needs to set the strength of interactions to values comparable to the disorder strength [38]. Isolated, interacting, and disordered systems that do not thermalize on a sufficiently long timescale are called many-body localized.

This Chapter aims to present and summarize Articles [1–3], providing new insights about the MBL in finite systems. Before we begin with the Articles, we provide a definition of MBL in Sec. 2.1.1, and in Sec. 2.1.2 introduce the *drosophila* model in the studies of MBL, the disordered XXZ spin-1/2 chain. The status of the field of MBL, with the main problems outlined in Sec. 2.1.3, have shaped the motivation for the studies made in the Articles. Each Article is then commented on in Secs. 2.2, 2.3, 2.4.

¹ Precisely, Anderson considered an effective non-interacting description of strongly interacting electrons moving in a disordered medium.

2.1.1 Definition of many-body localization

The many-body localized *phase* is defined in the thermodynamic limit (TL) through the long-time average of the expectation value of some few-body physical operator O by requiring it to disagree with the microcanonical expectation value,

$$\lim_{L \rightarrow \infty} \lim_{t_0 \rightarrow \infty} \frac{1}{t_0} \int_0^{t_0} dt O(t) \neq \text{Tr}\{\rho_{\text{MC}} O\}. \quad (2.1)$$

The order of size L and time t limits matters, and usually one first takes a limit with time and then with size going to infinity. If one studies the properties of the eigenstates for finite L , the limit of $t \rightarrow \infty$ is taken implicitly. In the time dynamics from a quench scenario, one usually calculates the time evolution of some observable up to a long time, then tries to fit a long-time behavior and performs a scaling to an infinite system size. Finite systems, if they do not thermalize at the longest relevant Heisenberg timescale given by the inverse energy level spacing, are in the MBL *regime* rather than *phase*.

Alternatively, the MBL can be defined as a phase/regime with a complete set of emergent local integrals of motion (LIOMs) [39, 40]. The eigenstates of any MBL Hamiltonian, due to low entanglement, can be smoothly transformed into simple product states through the action of a quasilocal unitary operator U . In the case of spin-1/2 Hamiltonians expressed in terms of Pauli operators $\sigma_i^{x,y,z}$, denoting the aforementioned product state as an eigenstate of another set of spin-1/2 operators τ_i^z , one can write

$$\tau_i^z = U \sigma_i^z U^\dagger, \quad (2.2)$$

where τ_i^z are the LIOMs. Unitary U is quasilocal, i.e., it can be factored into a sequence of maximally n -site operators with decreasing norms, $U = \Pi_i \dots U_{i,i+1,i+2}^{(3)} U_{i,i+1}^{(2)}$ with the Frobenius operator norm satisfying $\|1 - U_{i,i+1,\dots,i+n}^{(n)}\|_F^2 < e^{-n/\xi}$, and ξ is interpreted as the LIOM localization length. The full Hamiltonian transforms into

$$H = \sum_i J_i \tau_i^z + \sum_{i < j} J_{ij} \tau_i^z \tau_j^z + \sum_{i < j < k} J_{ijk} \tau_i^z \tau_j^z \tau_k^z + \dots, \quad (2.3)$$

with the couplings

$$J_{ij} \propto e^{-|i-j|/\kappa}, \quad J_{ijk} \propto e^{-|i-k|/\kappa}, \dots, \quad (2.4)$$

where κ is a constant. Indeed, $[\tau_i^z, H] = 0$, and the system described by LIOMs is integrable. Although a unitary U also exists in the ergodic regime, it is not quasilocal.

2.1.2 The disordered XXZ model

A paradigmatic model exhibiting the MBL behavior, studied by us, is the spin-1/2 XXZ model defined by the following Hamiltonian:

$$H_{\text{XXZ}} = J \sum_{i=1}^{L-1} (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z) + \sum_{i=1}^L W_i \sigma_i^z, \quad (2.5)$$

where $J = 1$ sets the energy scale, $\Delta = 1$ is the interaction amplitude, L is the system size, and W_i is a random variable uniformly distributed in the interval $[-W, W]$. The Hamiltonian conserves the total magnetization $\sigma_{\text{tot}}^z = \sum_i \sigma_i^z$. In the following, we will always concentrate on the largest $\sigma_{\text{tot}}^z = 0$ symmetry sector of dimension $\binom{L}{L/2}$. Through the Jordan-Wigner transformation [41], the model can be transformed into an equivalent model of interacting spinless fermions hopping on a lattice with random chemical potential on each site. For $W = 0$ the model is integrable through the Bethe Ansatz [26]. Strong ergodicity breaking in this model happens (at least for a finite system size) at a sufficiently large disorder amplitude W , currently estimated at around $W_c \approx 5.4$ [42].

2.1.3 Status of MBL: major problems

For an illustration of the major difficulties in analyzing the ETH-MBL transition in disordered XXZ and other models, we highlight a few approaches in the following paragraphs.

ENERGY SPECTRA Statistics of the mean energy level gap ratio, Eq. (1.4), can reveal if the system exhibits chaotic or integrable behavior. Figure 2.1a) shows $\bar{r}$ in the middle of the spectrum as a function of disorder strength W for different system sizes. The curve transitions from the GOE value for the weak disorder, $\bar{r}_{\text{GOE}} \approx 0.53$, to Poisson statistics $\bar{r}_{\text{PS}} \approx 0.39$ in the large disorder limit. However, upon increasing the system size, the transition point steadily drifts towards larger values of W . The departure of $\bar{r}$ from $\bar{r}_{\text{GOE}}$ starts at $W_T(L) \propto L$ [44], and the crossing point for consecutive system sizes drifts as $W^*(L) \approx W_c + C/L$, where C is a constant [17]. Extrapolation of the latter trend to $L \rightarrow \infty$ gives the critical disorder strength $W_c \approx 5.4$ [42], but a question on which, if any, of the two trends prevails in the TL is not yet resolved [45].

EIGENSTATES Eigenstates of the Hamiltonian contain more information than the energy spectrum. An insightful property used to characterize the eigenstates is the von Neumann EE, see Eq. (1.6). Figure 2.1b) shows the rescaled EE $s_E = \mathcal{S}_{\text{vN}}/\mathcal{S}_{\text{vN, RMT}}$ (data extracted from Ref. [42]) and a crossover between the volume-law entangle-

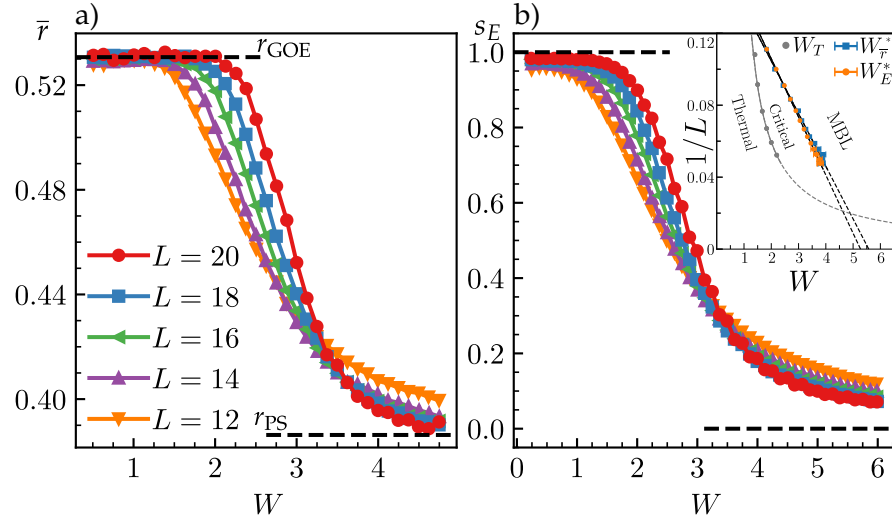


Figure 2.1: a) Values of the middle-spectrum average gap ratios $\bar{r}$ of the XXZ model as a function of the disorder strength W . The crossing point of the lines, as well as the disorder at which departure from the GOE prediction is observed, drifts significantly with the system size. b) Rescaled entanglement entropy $s_E = \mathcal{S}_{\text{vN}}/\mathcal{S}_{\text{RMT}}$ as a function of disorder strength: change from volume law for weak disorder to area law for strong disorder. Inset: Transition disorder strengths as a function of system size L , marking the departure from the GOE prediction at $W^T(L)$ or the crossing point at $W^*(L)$ (crossing point is shown also for the rescaled entanglement entropy s_E). Dashed line marks extrapolation to $L \rightarrow \infty$. Numerical data for panels a), and b) was extracted from [43]. The inset is adapted from [42] with permission.

ment for the weak disorder ($s_E \rightarrow 1$), and the area-law entanglement ($s_E \rightarrow 0$) for the large disorder. Significant drifts of the transition crossing point disorder strength with system size make the extrapolation to $L \rightarrow \infty$ similarly ambiguous as for $\bar{r}$.

Over the past several years, many other works have tried to characterize the eigenstates, see reviews [45–47]. These attempts include the details of the eigenstate entanglement structure along the crossover [48], trace distance between the eigenstates [49] as well as the reaction to a small perturbation of the Hamiltonian measured through the fidelity susceptibility [50]. Typically, there are similar problems with an unambiguous extrapolation of the crossover disorder strength to the critical disorder strength in the thermodynamic limit.

DYNAMIC QUANTITIES The MBL phase is defined as the lack of thermalization of observables to microcanonical values during the time evolution. Thus, studying the time dependence of observables in disordered many-body systems is a natural setup for distinguishing quantum ergodicity from MBL. Since performing time evolution for a chosen initial state with methods such as Chebyshev kernel polynomial expansion [51, 52] (see Appendix A for more details), Krylov space time evolution [53, 54] or tensor networks algorithms [55–58] is not as complex as solving a (more general) exact diagonalization (ED) problem, larger system sizes than in ED are within reach, which is particularly important as studies of MBL are concerned about the limit of large system sizes.

A typical quantity employed in the studies of MBL, dating back to the early experiments with ultracold fermions in an optical lattice [59], is the correlation function

$$C(t) = \frac{1}{c_0} \sum_{i=l_0}^{L-l_0} \langle \sigma_i^z(t) \sigma_i^z(0) \rangle, \quad (2.6)$$

where $l_0 \ll L$ is set to eliminate the boundary effects, and c_0 is a constant set by requiring $C(0) = 1$. The behavior of $C(t)$ quantifies the (auto)correlations between the initial projection of the z spin components and their time-evolved values. For the Néel initial state $|\psi(0)\rangle = |\downarrow\uparrow\downarrow\uparrow\dots\rangle$, the correlation function $C(t)$ is also called the *imbalance*². If the imbalance stays close to 1, the system retains the memory of the initial state and is regarded as many-body localized. Extensive studies of the imbalance in the XXZ model [60] in systems of size up to $L = 200$ and $W = 8$ have revealed a very slow power-law decay as $C(t) \propto t^{-\beta}$, where $\beta \approx 10^{-3}$, at least up to times $t = 1500$.

² In the interacting spinless fermion picture of the XXZ model, the Néel state corresponds to a density wave with every even site occupied by a fermion and every odd site empty. In this case, the imbalance measures if the time evolution retains the memory of the *imbalanced* initial state or leads to the redistribution of fermions equally on even and odd sites.

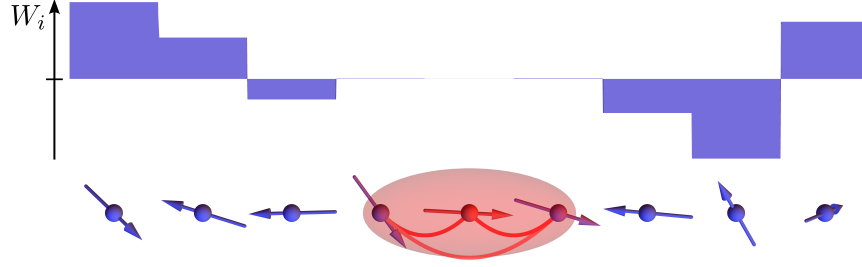


Figure 2.2: Due to a locally weak random disorder, an ergodic bubble (in red) of size $\ell \ll L$ (here $\ell = 3$) can emerge in the system, which happens with probability $p_\ell \rightarrow 1$ in the limit of large system $L \rightarrow \infty$. The ergodic bubble can thermalize its surroundings through interactions. In the quantum avalanche scenario of the ever-expanding single bubble, the whole system is eventually thermalized. Red lines mark strong two-body correlations within the bubble.

Whether the slow decay continues until the largest physically relevant Heisenberg timescale, exponential in the system size, remains unknown. At larger disorder strength, the dynamics slow down even more, with $\beta \propto e^{-\gamma W}$ [45], making the determination of whether the system eventually thermalizes extremely challenging, if not impossible.

A detailed description of all other studies of time-dependent wave function properties, including experimental works, goes beyond the scope of this short note to the subject. Notable examples include the investigation of the entanglement entropy [42, 61–66], inverse participation ratio [67], spin transport properties [68–70], out-of-time-ordered correlators [71, 72], survival probability [73], and operator spreading [74]. Nevertheless, all of the studies are limited to either small systems or finite times, denying a definite answer about the TL.

MODELING: QUANTUM AVALANCHES One can take an alternative approach to the studies of eigenstates and dynamical properties resulting from a given Hamiltonian and try *modeling* certain microscopic physical mechanisms, assuming that they drive the transition. Predictions of these models can be then compared with the numerical simulations for small systems and followed by a reasonable assumption that the mechanism works universally across system sizes and translates to the thermodynamic limit. So far, a few models that are believed to describe the essential aspects of the ETH-MBL transition have been proposed, see Ref. [45] for a recent overview. Articles [1, 2] concentrate on one of them, the so-called *quantum avalanche scenario* introduced by Ref. [75]. In this setting, thermalization is caused by rare thermal regions, also known as *ergodic bubbles* or *Griffiths regions* [76–79], that emerge due to accidentally low local disorder, see Fig. 2.2. The ergodic bubble thermalizes its surroundings, leading to a larger ergodic bubble with more potential for further thermalization. In a

single realization of random disorder from the interval $[-W, W]$, the probability that at ℓ consecutive sites the absolute disorder value is smaller than W_{thr} is $p_\ell = (W_{\text{thr}}/W)^\ell$. Thus, the weak disorder regions are exponentially rare with their size. On the other hand, in the thermodynamic limit of $L \rightarrow \infty$, having at least one accidental ergodic bubble of size at least ℓ (where $\ell \ll L$ is arbitrary) is a random event that happens with probability arbitrarily close to 1. Given that, studying the fate of the MBL region in the presence of ergodic bubbles across the finite-size ETH-MBL crossover is very important for understanding whether localization persists in the thermodynamic limit. The avalanches have been studied numerically in small systems [80–86]. The problem with the avalanche mechanism is that it remains to be verified how well its idealized model corresponds to behaviors in “real” systems – this is the topic partially covered by Articles [1, 2]. A formal description of the avalanche mechanism is presented in Article [2].

We should mention that several alternative mechanisms driving the MBL transition have been proposed. These include the study of the many-body resonance mechanism [77, 87, 88] (with potential links to the quantum avalanches [86]), and the analysis of the fate of strong impurities from Ref. [89].

SUMMARY Clear signs of emergent integrability have been detected in finite systems, but the verification of whether the MBL survives to the thermodynamic limit remains elusive. One possibility is to look at the features of the energy levels and the eigenstates. This approach is limited to small system sizes due to an exponential growth of the Hilbert space dimension with the system size L , and extrapolation to the thermodynamic limit is not supported by solid theoretical arguments, yielding inconclusive results. Another possibility, allowing one to reach larger system sizes, is to compute the time evolution of the system in a quench scenario. Here, the problem with the exponential slowdown of the dynamics with disorder amplitude makes the determination of even the existence of the ETH-MBL transition beyond the reach of state-of-the-art numerics. Finally, one can model the MBL transition by an idealized microscopic mechanism, the relevance of which for real systems is not entirely understood.

2.2 ARTICLE 1: DETECTING ERGODIC BUBBLES AT THE CROSSOVER TO MANY-BODY LOCALIZATION USING NEURAL NETWORKS

SUMMARY Although the construction of the perturbative model of quantum avalanches [75] can lead to interesting insights about the MBL transition in an idealized setting, one still needs to verify whether it accurately describes the dynamics of “real” many-body systems such as the disordered XXZ model. In this Article we propose a neural-network-based algorithm for the detection of ergodic bubbles based on the time evolution of the two-body correlation functions.

Using Chebyshev’s polynomial kernel time evolution method, described in Appendix A, we construct a dataset with the correlation functions as a function of time and space for the initial Néel state and the XXZ Hamiltonian with random and quasiperiodic disorder. We train a neural network with an autoencoder architecture, to encode and decode the time series with two-body correlations $\langle \sigma_i^z(t) \sigma_j^z(t) \rangle$ in the ergodic regime of the model. In this way, the network learns the universal properties of correlation function evolutions in the thermalizing system. Later, we use the neural network for the data across the ETH-MBL crossover. By inspecting the error of the input reconstruction, we can identify spatiotemporal regions corresponding to ergodic bubbles.

Our results confirm the relevance of rare Griffiths regions for the quantum avalanche process. We study the distribution of ergodic bubble sizes which changes from an exponential decay in the MBL regime of the model to a power law in the critical regime, which is compatible with the avalanche scenario of thermalization.

Additionally, we compare our results with a method from earlier Ref. [48] for detecting *entanglement clusters* in quantum wave functions, previously applied to eigenstates. Here, the algorithm is applied to time-evolved states and yields result in qualitative agreement with the ergodic bubble detection by neural networks, as outlined in the Supplemental Material of [1]. Recently, Ref. [90] has studied a related concept of *correlation clusters*, i. e. clusters of strongly correlated spins on a graph with an adjacency matrix defined by pairwise connected correlations. A comparison of both methods would be an interesting subject of future work.

AUTHOR’S CONTRIBUTIONS T. S. performed all numerical calculations for system sizes $L \leq 16$, proposed, implemented, and tested the neural networks for the ergodic bubble detection, reporting the results and discussing them in the group. Numerical calculations for system size $L = 24$ were supplied by P. S.. All figures (except Fig. 1 in [1], prepared by K. K.) were produced by T. S., who also drafted the initial version of the manuscript, and served as the corresponding author.

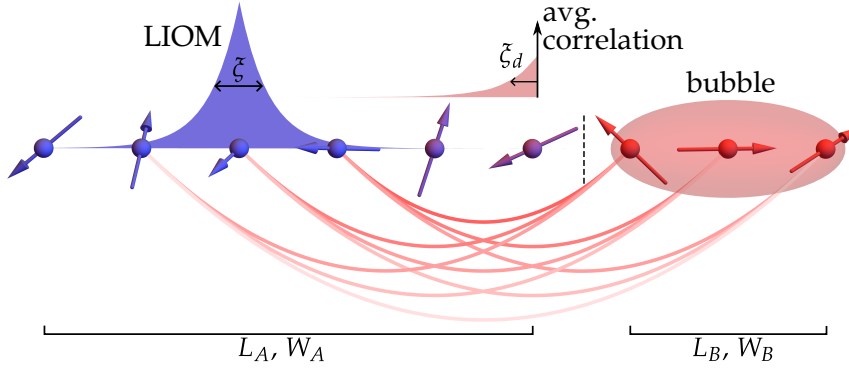


Figure 2.3: Model of an engineered ergodic bubble studied in Article [2]. The strongly disordered subsystem A of length L_A is modeled by LIOMs interacting with sites from a weakly disordered subsystem B of size L_B . The approximately exponential decay of two-body connected correlations of each spin with the bath (averaged over the bath), defines the decay length $\xi_d(t)$. Lines denote the strength of interactions between the spins in the LIOM and the bath, falling off exponentially with the distance, with a red (white) color corresponding to stronger (weaker) interactions. These interactions can potentially lead to a decay of the LIOM with time.

2.3 ARTICLE 2: CATCHING THERMAL AVALANCHES IN THE DISORDERED XXZ MODEL

SUMMARY As mentioned in Sec. 2.1.3, although the ergodic bubbles almost surely occur in disordered systems in the thermodynamic limit, they are exponentially rare in finite systems. This severely complicates the study of their influence on MBL, even using special ergodic bubble detection methods such as that based on neural networks [1]. This motivation has led us to the consideration of a disordered Heisenberg spin chain model in which an ergodic bubble of size L_B is *engineered* as a region with weak disorder $W_B = 0.5$ (deeply in the ergodic regime) and placed at one end of the chain of length L_A and disorder W_A , see Fig. 2.3. In this setting, the seed for an avalanche is present in *each* disorder realization. One should note that the first propositions of such investigations date back to Ref. [81] which looks at the system's eigenstates. Time evolutions were studied in Ref. [91] experimentally implementing the Bose-Hubbard model with quasiperiodic onsite disorder potential, and in Ref. [92] considering the XXZ chain with periodically enabled and disabled coupling to an ergodic bubble.

We use the Chebyshev polynomial expansion of the time evolution operator, Appendix A, to compute the time evolution of the initial quantum state with the disordered XXZ Hamiltonian. During the simulation, we measure the imbalance $C(t)$, Eq. (2.6), and the connected two-body correlation function

$$\langle \sigma_i^z(t) \sigma_j^z(t) \rangle_c = \langle \sigma_i^z(t) \sigma_j^z(t) \rangle - \langle \sigma_i^z(t) \rangle \langle \sigma_j^z(t) \rangle \quad (2.7)$$

and for each spin i we calculate the average correlation with the sites in the ergodic bubble. In the MBL regime, at a fixed time t , the average correlation decays approximately exponentially with the distance from the interface, and a fitted decay length $\xi_d(t)$ quantifies the distance thermalized by the ergodic bubble.

Our analysis of $\xi_d(t)$ identifies conditions under which the system does not thermalize even in the presence of the ergodic bubble. We provide analytical arguments that without an avalanche, the $\xi_d(t) \propto \ln t$ growth will saturate at a time exponential in the ergodic bubble size. The logarithmic in-time growth of $\xi_d(t)$ beyond this timescale, up to the Heisenberg time, as well as faster-than-logarithmic growth, are compatible with the scenario of quantum avalanches but do not exclude other microscopic mechanisms responsible for thermalization. Our interpretation corrects that from an earlier Ref. [91] which erroneously attributes $\xi_d(t) \propto \ln t$ to the termination of avalanches.

The numerical results confirm the expectations based on the simplified model of quantum avalanches but highlight the importance of the choice of the initial state. Even deep in the MBL regime, the growth $\xi_d(t) \propto \ln t$ continues until the largest timescales if the evolution starts in the Néel state. If the evolution starts from a fine-tuned state: product of eigenstate in the considered spin chain and Néel state in the ergodic bubble, we see the termination of avalanches at expected timescales. However, extraction of the critical disorder strength from the analysis of $\xi_d(t)$ in finite-size systems poses a challenge because it is not clear how one should go to the TL with both lengthscales L_A, L_B . We further show that the LIOM localization length ξ can be extracted from the dynamics of $\xi_d(t)$, and analyze the influence of the bath properties on the thermalization process, demonstrating that the ergodic bubble governed by a GOE Hamiltonian faithfully reproduces results of a weak disorder ergodic bubble. We also investigate the behavior of correlations within the bubble when it is in contact with an MBL system. Finally, we show that the kicked XXZ model of Ref. [92] has a positive bias towards thermalization not observed in our model.

AUTHOR'S CONTRIBUTIONS T. S. performed all numerical calculations and data analysis, reporting the results and discussing them in the group. All figures were produced by T. S., who also drafted the initial version of the manuscript and served as the corresponding author in the journal submission.

Derivation of the independence of the decay rate, Γ_j in the Article [2], on the bubble size (hence, correction of the error from an earlier Ref. [91]) is a contribution from P. S..

Using the CFD, we characterize three different regimes of the disordered XXZ model: ergodic, MBL, and AL. We uncover and eliminate technical problems in the interpretation of the CFD that arise due to the overall decrease of the correlation function oscillation amplitude in an interacting system. In the ergodic regime, we find that the CFD quickly grows due to the non-locality of the dynamics and identify a power law relation between the time and space for a fixed value of the CFD, with an exponent suggesting a superdiffusive spread of quantum information captured by the CFD. In the AL regime, we observe an information lightcone, with the spread of the CFD slowing down exponentially with the distance from the removed site. In the MBL regime, a similar trend is observed. A slow growth in time of the characteristic length scale that describes the exponential decay of the CFD is identified in the MBL regime.

AUTHOR'S CONTRIBUTIONS T. S. performed all numerical calculations, except the tensor network time evolution, and the data analysis, reporting the results and discussing them in the group. All figures were produced by T. S., who also drafted the initial version of the manuscript.

reduced density matrices of subsystem composed just of a single site. This would add spatial resolution to the Loschmidt echo.

QUANTUM MANY-BODY SCARS

3.1 INTRODUCTION

In 2017, an experiment with a quantum simulator based on 51 Rydberg atoms in a one-dimensional lattice revealed an unexpected dynamical behavior [103]. For specific, simple initial states the system showed oscillations of the wave function return probability at long times, while for generic initial states the system quickly relaxed to a thermal ensemble, which is a signature of ergodicity. This was intriguing as the effective Hamiltonian describing the system did not possess any straightforward conservation laws other than energy, eliminating integrability as a possible explanation of this effect. Localization phenomena were also ruled out due to dependence on the initial state, as well as the lack of any external disorder.

Soon later, the lattice with Rydberg atoms was modeled by an effective, idealized “PXP” model [31], which we describe in Sec. 3.2. Ergodicity breaking was traced back to the presence of a small number of special, ETH-violating eigenstates in the highly excited part of the spectrum, submerged in an otherwise thermal sea of eigenstates [31, 104]. It was found that the system violates the ETH *weakly*, i.e., the non-thermal eigenstates comprise a vanishing fraction of the whole Hilbert space. Oscillatory revivals of the system originate from the fact that these eigenstates have low entanglement entropy, an increased fidelity with the simple initial state, and are approximately equally spaced in energy.

The special eigenstates are called *quantum many-body scars*. It is now understood that scarring is connected with the decomposition of the Hilbert space into two sectors

$$\mathcal{H} \approx \mathcal{H}_{\text{scar}} \oplus \mathcal{H}_{\text{thermal}}, \quad (3.1)$$

where $\mathcal{H}_{\text{scar}}$ is the scarred subspace, and $\mathcal{H}_{\text{thermal}}$ is the subspace spanned by thermal eigenvectors of the Hamiltonian. These subspaces are approximately, or sometimes exactly, decoupled from each other in the sense that when the time evolution of the system starts in one subspace, the probability density stays there for a long time.

The scarring phenomenon has been discovered in plenty of other interacting quantum models and several practical uses in quantum technologies have been found. Early Ref. [105] uses scars to prepare Schrödinger cat states in the Rydberg atom array. The special entanglement properties of the QMBS can be utilized for entanglement-enhanced metrology [106–108]. Robustness of scars against disorder can be potentially used in near-term quantum devices as a resource for quantum hardware calibration [109]. Refs. [110, 111] use scars embedded in the spectrum of the Floquet time evolution operator of a driven PXP model to construct period-doubling discrete time crystals [112]. For an extensive summary on the subject of QMBS, the reader is referred to recent reviews in Refs. [113–115].

This Chapter aims to present and summarize the Article [4]. In Sec. 3.2 we shortly discuss the PXP model and the problem we want to solve. Section 3.3 is a commentary on Article [4].

3.2 THE PXP MODEL

Let us come back to the experiment with an array of Rydberg atoms [103]. A Rydberg atom can be effectively viewed as a two-level system, i.e., spin-1/2, where one projection of the spin, in this context usually denoted as $|\circ\rangle$, corresponds to a groundstate of the Rydberg atom, while the other projection $|\bullet\rangle$ to the highly excited (principal quantum number of order $n \sim O(50)$) Rydberg state. In presence of the microwave field, the atom performs Rabi oscillations between these two levels. Let us notice that a highly excited atom has a large dipole moment because the electron spends most of time far from the nucleus. When two Rydberg atoms are put together, they interact via the van der Waals dipole-dipole interactions, strongly dependent on the distance r between the atoms as $\propto 1/r^6$. If the atoms are close enough, the energy penalty of the van der Waals interaction prohibits the situation in which both atoms are excited $|\dots\bullet\bullet\dots\rangle$. This phenomenon is called a Rydberg blockade [116–118] and constitutes an effective constraint on the dynamics of the array of Rydberg atoms in a lattice with period smaller than the Rydberg blockade radius: an atom cannot transition to an excited state if any of its neighbors is in the excited state.

The effective model of a perfect Rydberg blockade is called the PXP model and is defined by the following Hamiltonian:

$$H_{\text{PXP}} = \sum_{i=1}^L P_{i-1} \sigma_i^x P_{i+1}, \quad (3.2)$$

where we assume periodic boundary conditions ($P_{L+1} \equiv P_1$, $P_0 \equiv P_L$). Projectors $P_i = (1 - \sigma_i^z)/2 = |\circ\rangle_i \langle\circ|_i$ ensure that no two neighboring spins can be both in the $|\bullet\rangle$ state at the same time: states like

$|\dots\bullet\bullet\dots\rangle$ are annihilated by the Hamiltonian and thus excluded from the Hilbert space. Without the projectors, the Hamiltonian given by Eq. (3.2) would correspond to simple Rabi oscillations of individual spins with the operator $\sigma_i^x = |\circ\rangle_i\langle\bullet|_i + |\bullet\rangle_i\langle\circ|_i$, but the Rydberg blockade makes the model interacting.

Simulations of the PXP model have confirmed the experimental results: when the system is initialized in a simple initial product state $|\mathbb{Z}_2\rangle = |\circ\bullet\circ\bullet\dots\rangle$, i.e. a charge density wave with period 2, it performs long-time revivals of the wave function and local observables [31, 104]. Other initial states such as period-3 density wave $|\mathbb{Z}_3\rangle = |\circ\circ\bullet\circ\circ\bullet\dots\rangle$ show similar revivals with different frequency, whereas other “simple” states such as period-4 density wave $|\mathbb{Z}_4\rangle$ or no excitations $|\circ\circ\circ\circ\dots\rangle$ thermalize. Inspection of the eigenstates of the PXP Hamiltonian in Fig. 3.1 reveals that there is a number of special eigenstates, approximately equally spaced in energy, with an increased overlap with the initial state $|\mathbb{Z}_2\rangle$ (similarly for $|\mathbb{Z}_3\rangle$ - plot not shown), as well as the low entanglement entropy. These eigenstates form *families* living in the corresponding subspaces of the Hilbert space and manifestly violate the ETH prediction. However, one can see that the entanglement entropy for some other states is also low and a certain regular structure is visible, but no families are identified. In Article [4] we aim to characterize this part of the spectrum.

3.3 ARTICLE 4: “UNSUPERVISED DETECTION OF DECOUPLED SUBSPACES: MANY-BODY SCARS AND BEYOND”

SUMMARY The von Neumann entanglement entropy of the PXP Hamiltonian (3.2) shows a rich structure in the low-entanglement part. Although the two simplest scar families can be traced back to have a large overlap with the $|\mathbb{Z}_2\rangle$ and $|\mathbb{Z}_3\rangle$ “parent” states, the question whether other eigenstates also form families (clusters) of scars with a common “parent” state, remained open. The problem is complicated by the fact that the eigenvectors are mutually orthogonal to each other so standard similarity measures from linear algebra, such as the inner product, surely fail. Article [4] introduces a new paradigm in the automatic detection of clusters of eigenvectors from the same subspace of the Hilbert space based on an unsupervised quantum-inspired machine learning algorithm.

The main idea relies on using the quantum variational autoencoder (QVAE) circuit, Fig. 3.2, as a measure of similarity between the eigenvectors of a given Hamiltonian presumably hosting families of quantum scars or the fragmented Hilbert space. The purpose of the QVAE is to “compress” the input quantum state on N qubits onto a smaller number of $N - N_{\text{trash}}$ qubits. The remaining N_{trash} “trash” qubits are supposed to be always in the same product state and all

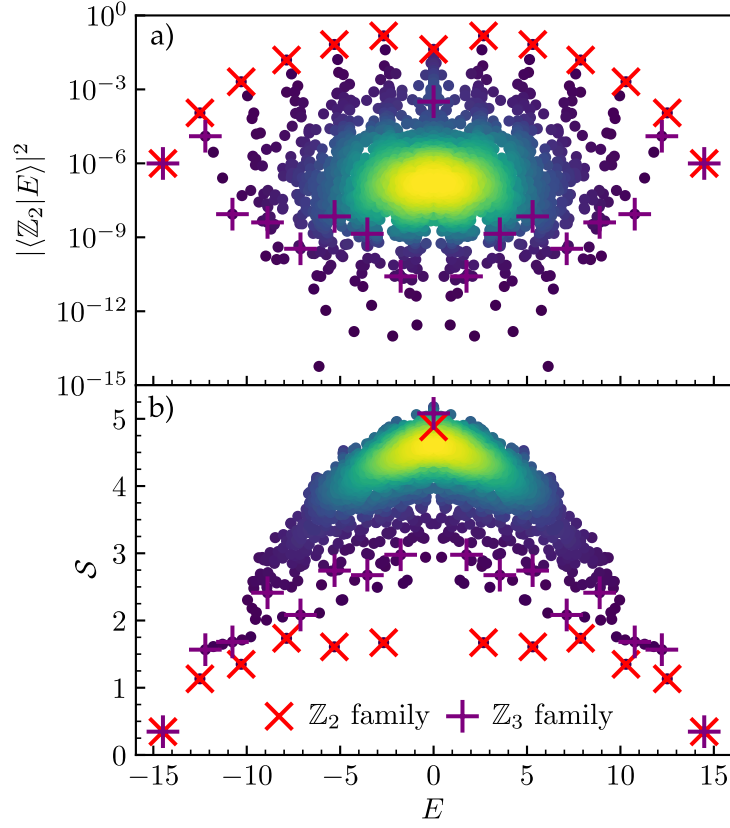


Figure 3.1: a) Fidelity of the $|\mathbb{Z}_2\rangle$ state with the eigenstates of the PXP model for system size $L = 24$, restricted to the zero momentum and inversion-symmetric symmetry sector. There is a small number of special eigenstates with a large overlap, approximately equally spaced in energy. Colors denote the density of states, with yellow (purple) corresponding to large (small) density. In a similar way, eigenstates corresponding to the $\mathbb{Z}_3$ family can be identified. b) Von Neumann bipartite entanglement entropy calculated for the same eigenstates. Scars corresponding to the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ families show anomalously low entropy. However, there are more low entropy states, characterization of which is our goal.

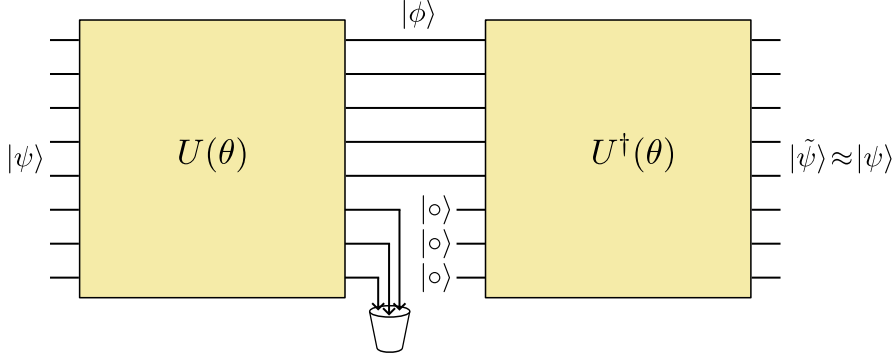


Figure 3.2: Schematic representation of a quantum variational autoencoder. Horizontal lines correspond to qubits. The input state $|\psi\rangle$ is transformed by a parametrized unitary operator $U(\theta)$. The transformation is variationally optimized so that the resulting state is $|\phi\rangle \otimes |\circ\rangle^{\otimes N_{\text{trash}}}$, where N_{trash} is a number of so-called “trash qubits”, which do not carry any information about the input state, and latent representation $|\phi\rangle$ is defined on a smaller number of qubits than $|\psi\rangle$. In contrast to the neural network autoencoder, the encoder $U(\theta)$ automatically provides a decoder $U^\dagger(\theta)$. During decoding, the state of the trash qubits on the encoder output is disregarded and a new product state is prepared.

information about the input is supposed to be transferred to the other qubits. Such a situation is achieved after variational optimization of the circuit parameters on a training dataset to minimize the cost function quantifying the independence of the trash qubits on the input. In our case, the trash qubits are supposed to always be in the product state $|\circ\rangle^{\otimes N_{\text{trash}}}$, so the cost function is the number of “•” measured on trash qubits. Successful training of the QVAE relies on the assumption that the training dataset consists of states with low entanglement [119].

We have found suitable circuit hyperparameters for the PXP model by testing the scar family detection on the known $\mathbb{Z}_2$ and $\mathbb{Z}_3$ scar families. In order to reach relatively large system sizes of $N = 24$, we have restricted the unitary gates of the circuit to be compatible with the constraints of the PXP model Hilbert space. We then proceeded to analyze the rest of the low-entanglement part of the spectrum. As a result, we have identified five new families of scars and their characteristic features.

In order to show that our algorithm is not limited only to a specific PXP model, we also considered the spin ladder model with exactly decoupled subspaces from Ref. [120]. The QVAE algorithm allowed us to successfully identify hidden symmetries in the Hilbert space.

RELATED WORKS There were other attempts at detecting scars with the use of machine learning and other algorithms. Refs. [121, 122] concentrate on identifying single-particle scars with ML on classical, image-like data. Working with many-body models, a recent work [123]

proposes multidimensional scaling (of which the principal component analysis is a special case) to detect whether an initial state corresponds to scarring dynamics or not based on spin configurations stochastically probed during time evolution. In a series of works, Moudgalya et al. [124–127] introduce the concept of commutant algebras to find invariant subspaces in the Hilbert space with numerical and analytical tools. Ref. [49] discusses an alternative similarity measure for eigenstates based on trace distances between the reduced density matrices, potentially also useful for detecting hidden symmetries. Ref. [128] proposes an algorithm for finding *local* conservation laws in models with Hilbert space fragmentation and MBL.

The algorithm proposed by us is limited by the need of the exact diagonalization (ED), required to propose candidates for the scarred eigenstates, which are then processed by the QVAE. For Hilbert spaces of dimension beyond $\sim 10^5$, ED becomes infeasible. After the publication of our paper, there were at least two attempts to obtain scarred eigenstates in the spectra of large systems beyond ED. Both are based on low entanglement of the scarred eigenstates. Ref. [129] uses the density matrix renormalization group algorithm for matrix product states with a small bond dimension to find scar eigenstates in the highly excited part of the spectrum. Ref. [119] does essentially the same for shallow quantum circuits. Both works show that our algorithm could potentially be used in the future for larger systems than attainable with the ED.

AUTHOR’S CONTRIBUTIONS T. S. has initiated the project, applied for funding, performed all numerical calculations, analyzed the data and presented it at the group meetings, prepared the first draft of the paper, including all figures, and the supplemental material. He was the corresponding author.

The idea to showcase the use of QVAE for the spin ladder model is a contribution from P. S..

CONCLUSIONS

In a series of Articles [1–4] we make a step forward in understanding ergodicity breaking phenomena in quantum many-body systems and provide novel numerical tools for their analysis.

Two of our studies of MBL concentrate on modeling the breakdown of localization through quantum avalanches. To check the relevance of this mechanism in the disordered XXZ model, in Article [1] we propose an algorithm based on neural networks that can detect ergodic bubbles in the time evolutions of two-body correlation functions in the disordered XXZ spin chain. We find qualitative and quantitative differences between the average bubble size growth in time and the bubble size distribution in the MBL, critical, and ergodic regimes. Overall, these results are consistent with previous claims that quantum avalanches can be responsible for destroying localization [75]. It is also worth noting that results obtained with neural networks are in qualitative agreement with an algorithm based on quantum mutual information for identifying entanglement clusters [48] but because of the use of simple correlation functions require much less computational and experimental resources. This constitutes an excellent example of how to avoid extensive quantum calculations and delegate a difficult quantum data analysis problem to a deep learning model. It should be pointed out that it would be very interesting to compare the ergodic bubbles detected by the neural network with the correlation clusters from later Ref. [90], constructed on similar data but using graph theory to find correlated “communities” of lattice sites.

In the follow-up Article [2], instead of looking for accidental ergodic bubbles, we *plant* an ergodic bubble in the system for every disorder realization at one end of the chain. We employ connected two-body correlations of spins with the bubble to investigate how far the bubble spreads from the initial setting depending on the disorder strength. Our central result is the demonstration that avalanche suppression in the MBL regime is not as easy to observe as it was previously claimed in Ref. [91], i.e., an unbounded growth of the volume thermalized by the bubble proportional to $\ln t$ shall be interpreted as a highly non-perturbative avalanche spreading process. We also discover that the initial state plays a significant role for the fate of an avalanche in the considered model. We highlight the technical difficulties in data

analysis for non-self-averaging quantities, such as the investigated correlation decay rate, which is a warning for all future numerical works on MBL in the disordered XXZ model. Furthermore, our results are compared with those of Ref. [92], studying the XXZ model with a weakly disordered bath connected periodically. We find a fundamental difference in the system's behavior – the time-dependent model has a natural tendency towards thermalization. Thus, Ref. [92] may be misinterpreting thermalization coming from the driving with that coming from the avalanches due to the bath. It would be very interesting to check if the same applies to other Floquet models with the ETH-MBL transition, such as the disordered kicked Ising model from Ref. [130]. The setup with a planted ergodic bubble considered by us is quite general and not restricted only to the disordered XXZ model. Our work, by providing the tested numerical analysis tools and their exhaustive characterization, opens new pathways in the study of thermalization caused by ergodic inclusions for various other Hamiltonians.

The Article [3] introduces a novel method of measuring the propagation of quantum information in many-body systems using single-body autocorrelation functions. The autocorrelation is computed at a certain moment in time in a system of size L and its smaller copy of size $\ell < L$. The correlation function difference (CFD) informs whether the presence of additional sites affects the local measurement, which is a proxy for measuring the transfer of quantum information. Benchmarking this probe in the different regimes of the disordered XXZ model, we unravel and solve technical difficulties for potential future users. By working only with simple, experimentally accessible observables, the CFD becomes a feasible tool for investigations of quantum many-body dynamics. In particular, our method does not require the reversal of time necessary for OTOC, and provides the spatial resolution unlike the Loschmidt echo. As a future work, it would be interesting to find further links between the CFD and some other quantum information propagation probes, and apply the CFD to other quantum models.

In the Article [4], we propose a quantum many-body scar clusterization algorithm based on an unsupervised quantum machine learning platform. As a key result, we show that a quantum variational autoencoder can be trained to serve as a similarity measure between the PXP Hamiltonian eigenstates and identify known and new families of scars. The proposed algorithm can be used for other quantum models as we highlight by applying the QVAE to the disordered spin-1/2 ladder. For other models, the suspicion that hidden symmetries are embedded in the Hilbert space may come from unconventional level statistics [19]. Nevertheless, the requirement for a practical trainability of the chosen circuit architecture is the weak, area-law entanglement, which eliminates the algorithm's applicability to a large class of models with volume-law entangled eigenstates. Concerning future work, it would

be interesting to find links between the mathematical framework of commutant algebras [124–127] for detecting hidden symmetries, and the QVAE clustering. One could as well try applying the trace distance between the eigenstates [49] to investigate families of scars in the PXP model. Newly developed methods for scarred excited state calculations with tensor networks [129] and quantum circuits [119] may allow for applying our algorithm to much larger systems in the future.

EXACT TIME EVOLUTION BASED ON THE CHEBYSHEV POLYNOMIAL EXPANSION

Let us consider the problem of numerically solving the Schrödinger equation (SE)

$$i \frac{d}{dt} |\psi\rangle = H |\psi\rangle, \quad (\text{A.1})$$

where the Hamiltonian H does not depend on time. Formally, the solution is given by

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle, \quad U(t) = e^{-iHt}. \quad (\text{A.2})$$

The matrix exponent in Eq. (A.2) can be easily computed for any time t if the Hamiltonian is in the diagonal form after exact diagonalization. The ED has a $O(\mathcal{D}^3)$ computational cost and $O(\mathcal{D}^2)$ memory cost for keeping the resulting eigenstates. Storing only the time-evolved wavefunction $|\psi(t)\rangle$ at a time instant t would bear a significantly smaller memory cost of $O(\mathcal{D})$. Thus, it is memory-efficient to propagate the state $|\psi(0)\rangle$ in many timesteps Δt up to time t , each time applying the unitary operator, $|\psi(\tau + \Delta t)\rangle = U(\Delta t) |\psi(\tau)\rangle$.

The simplest scheme to propagate Eq. (A.1) by a timestep Δt is to expand the operator $U(\Delta t)$ into a Taylor series

$$U(\Delta t) \approx 1 - iH\Delta t + \dots \quad (\text{A.3})$$

In a typical many-body problem, excluding all-to-all interacting models, the Hamiltonian matrix is sparse in the computational basis, and this can be utilized to efficiently compute the action of the Hamiltonian matrix on the state vector with the numerical cost of $O(d\mathcal{D})$ where d is the average number of non-zero elements of the matrix in a row. However, a numerical time evolution based on this expansion is unstable because it does not preserve the time-reversal symmetry of the SE. Although Eq. (A.3) can be symmetrized to restore the time-reversal symmetry and preserve the norm and energy of the state, the timestep Δt must be very small in order to limit the error of the Taylor expansion, making the scheme impractical for long time evolutions [51]. The problem boils down to finding the polynomial approximation of $U(t)$ in terms of H . It has been shown [51] that the

lowest approximation error is achieved using the expansion based on the Chebyshev polynomials¹

$$U(\Delta t) \approx e^{-ib\Delta t} \left(J_0(a\Delta t) + 2 \sum_{k=1}^N (-i)^k J_k(a\Delta t) T_k(\mathcal{H}) \right), \quad (\text{A.4})$$

where $a = (E_{\max} - E_{\min})/2$, $b = (E_{\max} + E_{\min})/2$, and $E_{\min}(E_{\max})$ is the minimal (maximal) eigenvalue of the Hamiltonian H . The rescaled Hamiltonian is $\mathcal{H} = (H - b)/a$, and $J_k(t)$ is the Bessel function of the first kind of order k ,

$$J_k(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m+k+1)} \left(\frac{x}{2} \right)^{2m+k}, \quad (\text{A.5})$$

where $\Gamma(z)$ is the gamma function. The Chebyshev polynomials of the first kind $T_n(x)$ are defined on the interval $[-1, 1]$ by

$$T_n(x) = \cos(n \arccos x), \quad (\text{A.6})$$

satisfy the orthogonality relation with respect to the inner product

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) \frac{dx}{\sqrt{1-x^2}}, \quad (\text{A.7})$$

and can be computed using the recurrence relation

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x). \end{aligned} \quad (\text{A.8})$$

The order of expansion N is determined numerically so that the resulting approximate operator preserves the norm of the state vector it acts upon up to a desired accuracy. The extremal values of the energy $E_{\min}$, $E_{\max}$ are computed using the Lanczos algorithm [131].

¹ We adopt the notation from Ref. [45].

BIBLIOGRAPHY

- [1] T. Szóldra, P. Sierant, K. Kottmann, M. Lewenstein, and J. Zakrzewski, “Detecting ergodic bubbles at the crossover to many-body localization using neural networks”, *Phys. Rev. B* **104**, L140202 (2021).
- [2] T. Szóldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, “Catching thermal avalanches in the disordered XXZ model”, *Phys. Rev. B* **109**, 134202 (2024).
- [3] T. Szóldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, “Tracking locality in the time evolution of disordered systems”, *Phys. Rev. B* **107**, 054204 (2023).
- [4] T. Szóldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, “Unsupervised detection of decoupled subspaces: many-body scars and beyond”, *Phys. Rev. B* **105**, 224205 (2022).
- [5] L. Boltzmann, *Vorlesungen über Gastheorie*, Vol. 98 (J.A. Barth, Leipzig, 1896).
- [6] E. P. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions”, *Annals of Mathematics* **62**, 548–564 (1955).
- [7] E. P. Wigner, in Gatlinburg conference on neutron physics by time-of-flight, 2309 (Oak Ridge National Laboratory, 1957), p. 59.
- [8] E. P. Wigner, “On the distribution of the roots of certain symmetric matrices”, *Annals of Mathematics* **67**, 325–327 (1958).
- [9] M. Mehta, “On the statistical properties of the level-spacings in nuclear spectra”, *Nuclear Physics* **18**, 395–419 (1960).
- [10] O. Bohigas, “Random matrix theories and chaotic dynamics”, (1991).
- [11] F. Haake, *Quantum signatures of chaos* (Springer, Berlin, 2010).
- [12] O. Bohigas, M. J. Giannoni, and C. Schmit, “Characterization of chaotic quantum spectra and universality of level fluctuation laws”, *Phys. Rev. Lett.* **52**, 1–4 (1984).
- [13] O. Bohigas, R. U. Haq, and A. Pandey, *Fluctuation properties of nuclear energy levels and widths: comparison of theory with experiment* (D Reidel, Netherlands, 1983).
- [14] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, “From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics”, *Advances in Physics* **65**, 239–362 (2016).

- [15] D. Wintgen and H. Friedrich, “Classical and quantum-mechanical transition between regularity and irregularity in a Hamiltonian system”, *Phys. Rev. A* **35**, 1464–1466 (1987).
- [16] L. F. Santos and M. Rigol, “Onset of quantum chaos in one-dimensional bosonic and fermionic systems and its relation to thermalization”, *Phys. Rev. E* **81**, 036206 (2010).
- [17] V. Oganesyan and D. A. Huse, “Localization of interacting fermions at high temperature”, *Phys. Rev. B* **75**, 155111 (2007).
- [18] 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).
- [19] O. Giraud, N. Macé, E. Vernier, and F. Alet, “Probing symmetries of quantum many-body systems through gap ratio statistics”, *Phys. Rev. X* **12**, 011006 (2022).
- [20] L. Vidmar and M. Rigol, “Entanglement entropy of eigenstates of quantum chaotic Hamiltonians”, *Phys. Rev. Lett.* **119**, 220603 (2017).
- [21] D. N. Page, “Average entropy of a subsystem”, *Phys. Rev. Lett.* **71**, 1291–1294 (1993).
- [22] M. Rigol, V. Dunjko, and M. Olshanii, “Thermalization and its mechanism for generic isolated quantum systems”, *Nature* **452**, 854 EP – (2008).
- [23] J. M. Deutsch, “Quantum statistical mechanics in a closed system”, *Phys. Rev. A* **43**, 2046–2049 (1991).
- [24] M. Srednicki, “Chaos and quantum thermalization”, *Phys. Rev. E* **50**, 888–901 (1994).
- [25] M. Srednicki, “The approach to thermal equilibrium in quantized chaotic systems”, *Journal of Physics A: Mathematical and General* **32**, 1163–1175 (1999).
- [26] R. Orbach, “Linear antiferromagnetic chain with anisotropic coupling”, *Phys. Rev.* **112**, 309–316 (1958).
- [27] E. H. Lieb and F. Y. Wu, “Absence of Mott transition in an exact solution of the short-range, one-band model in one dimension”, *Phys. Rev. Lett.* **20**, 1445–1448 (1968).
- [28] A. N. Kolmogorov, “On conservation of conditionally periodic motions for a small change in Hamilton’s function”, *Dokl. Akad. Nauk SSSR* **98**, 527–530 (1954).
- [29] V. I. Arnol’d, “Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian”, *Uspehi Mat. Nauk* **18**, 13–40 (1963).

- [30] J. Moser, “On invariant curves of area-preserving mappings of an annulus”, *Nachr. Akad. Wiss. Göttingen Math.-Phys.* **II**, 1–20 (1962).
- [31] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, “Weak ergodicity breaking from quantum many-body scars”, *Nature Physics* **14**, 745–749 (2018).
- [32] P. Sala, T. Rakovszky, R. Verresen, M. Knap, and F. Pollmann, “Ergodicity breaking arising from Hilbert space fragmentation in dipole-conserving Hamiltonians”, *Physical Review X* **10**, 011047 (2020).
- [33] L. Herviou, J. H. Bardarson, and N. Regnault, “Many-body localization in a fragmented Hilbert space”, *Phys. Rev. B* **103**, 134207 (2021).
- [34] V. Khemani, M. Hermele, and R. Nandkishore, “Localization from Hilbert space shattering: From theory to physical realizations”, *Phys. Rev. B* **101**, 174204 (2020).
- [35] P. W. Anderson, “Absence of diffusion in certain random lattices”, *Phys. Rev.* **109**, 1492–1505 (1958).
- [36] A. Lagendijk, B. v. Tiggelen, and D. S. Wiersma, “Fifty years of Anderson localization”, *Physics Today* **62**, 24–29 (2009).
- [37] J. Colbois, F. Alet, and N. Laflorencie, “Interaction-driven instabilities in the random-field XXZ chain”, (2024), [arXiv:2403.09608 \[cond-mat.dis-nn\]](#).
- [38] 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).
- [39] 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).
- [40] D. A. Huse, R. Nandkishore, and V. Oganesyan, “Phenomenology of fully many-body-localized systems”, *Phys. Rev. B* **90**, 174202 (2014).
- [41] P. Jordan and E. Wigner, “Über das Paulische Äquivalenzverbot”, *Zeitschrift für Physik* **47**, 631–651 (1928).
- [42] P. Sierant, M. Lewenstein, and J. Zakrzewski, “Polynomially filtered exact diagonalization approach to many-body localization”, *Phys. Rev. Lett.* **125**, 156601 (2020).
- [43] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, “Ergodicity breaking transition in finite disordered spin chains”, *Phys. Rev. B* **102**, 064207 (2020).
- [44] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, “Quantum chaos challenges many-body localization”, *Phys. Rev. E* **102**, 062144 (2020).

- [45] P. Sierant, M. Lewenstein, A. Scardicchio, L. Vidmar, and J. Zakrzewski, “Many-body localization in the age of classical computing”, (2024), [arXiv:2403.07111 \[cond-mat.dis-nn\]](#).
- [46] D. J. Luitz and Y. B. Lev, “The ergodic side of the many-body localization transition”, *Annalen der Physik* **529**, 1600350 (2017).
- [47] F. Alet and N. Laflorencie, “Many-body localization: an introduction and selected topics”, *Comptes Rendus Physique* **19**, 498–525 (2018).
- [48] L. Herviou, S. Bera, and J. H. Bardarson, “Multiscale entanglement clusters at the many-body localization phase transition”, *Phys. Rev. B* **99**, 134205 (2019).
- [49] R. Khasseh, J. Zhang, M. Heyl, and M. A. Rajabpour, “Identifying quantum many-body integrability and chaos using eigenstate trace distances”, *Phys. Rev. Lett.* **131**, 216701 (2023).
- [50] D. Sels and A. Polkovnikov, “Dynamical obstruction to localization in a disordered spin chain”, *Phys. Rev. E* **104**, 054105 (2021).
- [51] H. Tal-Ezer and R. Kosloff, “An accurate and efficient scheme for propagating the time dependent Schrödinger equation”, *The Journal of Chemical Physics* **81**, 3967–3971 (1984).
- [52] H. Fehske and R. Schneider, *Computational many-particle physics* (Springer, Germany, 2008).
- [53] A. Nauts and R. E. Wyatt, “New approach to many-state quantum dynamics: the recursive-residue-generation method”, *Phys. Rev. Lett.* **51**, 2238–2241 (1983).
- [54] D. J. Luitz and Y. B. Lev, “The ergodic side of the many-body localization transition”, *Annalen der Physik* **529**, 1600350 (2017).
- [55] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, “Time-dependent variational principle for quantum lattices”, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [56] T. Koffel, M. Lewenstein, and L. Tagliacozzo, “Entanglement entropy for the long-range Ising chain in a transverse field”, *Phys. Rev. Lett.* **109**, 267203 (2012).
- [57] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, “Unifying time evolution and optimization with matrix product states”, *Phys. Rev. B* **94**, 165116 (2016).
- [58] T. Chanda, P. Sierant, and J. Zakrzewski, “Time dynamics with matrix product states: many-body localization transition of large systems revisited”, *Phys. Rev. B* **101**, 035148 (2020).
- [59] 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**, 842–845 (2015).

- [60] P. Sierant and J. Zakrzewski, “Challenges to observation of many-body localization”, *Phys. Rev. B* **105**, 224203 (2022).
- [61] 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).
- [62] M. Serbyn, Z. Papić, and D. A. Abanin, “Universal slow growth of entanglement in interacting strongly disordered systems”, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [63] D. J. Luitz and Y. Bar Lev, “Anomalous thermalization in ergodic systems”, *Phys. Rev. Lett.* **117**, 170404 (2016).
- [64] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, “Probing entanglement in a many-body-localized system”, *Science* **364**, 256–260 (2019).
- [65] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, “Evidence for unbounded growth of the number entropy in many-body localized phases”, *Phys. Rev. Lett.* **124**, 243601 (2020).
- [66] 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).
- [67] M. Schiulaz, E. J. Torres-Herrera, F. Pérez-Bernal, and L. F. Santos, “Self-averaging in many-body quantum systems out of equilibrium: chaotic systems”, *Phys. Rev. B* **101**, 174312 (2020).
- [68] Y. Bar Lev, G. Cohen, and D. R. Reichman, “Absence of diffusion in an interacting system of spinless fermions on a one-dimensional disordered lattice”, *Phys. Rev. Lett.* **114**, 100601 (2015).
- [69] S. Bera, G. De Tomasi, F. Weiner, and F. Evers, “Density propagator for many-body localization: finite-size effects, transient subdiffusion, and exponential decay”, *Phys. Rev. Lett.* **118**, 196801 (2017).
- [70] F. Weiner, F. Evers, and S. Bera, “Slow dynamics and strong finite-size effects in many-body localization with random and quasiperiodic potentials”, *Phys. Rev. B* **100**, 104204 (2019).
- [71] X. Chen, T. Zhou, D. A. Huse, and E. Fradkin, “Out-of-time-order correlations in many-body localized and thermal phases”, *Annalen der Physik* **529**, 1600332 (2017).
- [72] X. Mi et al., “Information scrambling in quantum circuits”, *Science* **374**, 1479–1483 (2021).

- [73] E. J. Torres-Herrera, A. M. García-García, and L. F. Santos, “Generic dynamical features of quenched interacting quantum systems: survival probability, density imbalance, and out-of-time-ordered correlator”, *Phys. Rev. B* **97**, 060303 (2018).
- [74] A. Weisse, R. Gerstner, and J. Sirker, “Operator spreading and the absence of many-body localization”, (2024), [arXiv:2401.08031 \[cond-mat.dis-nn\]](#).
- [75] W. De Roeck and F. Huveneers, “Stability and instability towards delocalization in many-body localization systems”, *Phys. Rev. B* **95**, 155129 (2017).
- [76] T. Vojta, “Quantum Griffiths effects and smeared phase transitions in metals: theory and experiment”, *J. Low Temp. Phys.* **161**, 299–323 (2010).
- [77] S. Gopalakrishnan, M. Müller, V. Khemani, M. Knap, E. Demler, and D. A. Huse, “Low-frequency conductivity in many-body localized systems”, *Phys. Rev. B* **92**, 104202 (2015).
- [78] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, “Anomalous diffusion and Griffiths effects near the many-body localization transition”, *Phys. Rev. Lett.* **114**, 160401 (2015).
- [79] N. Pancotti, M. Knap, D. A. Huse, J. I. Cirac, and M. C. Bañuls, “Almost conserved operators in nearly many-body localized systems”, *Phys. Rev. B* **97**, 094206 (2018).
- [80] D. J. Luitz, F. m. c. Huveneers, and W. De Roeck, “How a small quantum bath can thermalize long localized chains”, *Phys. Rev. Lett.* **119**, 150602 (2017).
- [81] M. Goihl, J. Eisert, and C. Krumnow, “Exploration of the stability of many-body localized systems in the presence of a small bath”, *Phys. Rev. B* **99**, 195145 (2019).
- [82] P. J. D. Crowley and A. Chandran, “Avalanche induced coexisting localized and thermal regions in disordered chains”, *Phys. Rev. Res.* **2**, 033262 (2020).
- [83] A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, “Avalanches and many-body resonances in many-body localized systems”, *Phys. Rev. B* **105**, 174205 (2022).
- [84] D. Sels, “Bath-induced delocalization in interacting disordered spin chains”, *Phys. Rev. B* **106**, L020202 (2022).
- [85] Y.-T. Tu, D. Vu, and S. Das Sarma, “Avalanche stability transition in interacting quasiperiodic systems”, *Phys. Rev. B* **107**, 014203 (2023).
- [86] H. Ha, A. Morningstar, and D. A. Huse, “Many-body resonances in the avalanche instability of many-body localization”, *Phys. Rev. Lett.* **130**, 250405 (2023).

- [87] B. Villalonga and B. K. Clark, “Eigenstates hybridize on all length scales at the many-body localization transition”, (2020), [arXiv:2005.13558 \[cond-mat.dis-nn\]](#).
- [88] P. J. D. Crowley and A. Chandran, “A constructive theory of the numerically accessible many-body localized to thermal crossover”, *SciPost Phys.* **12**, 201 (2022).
- [89] D. Sels and A. Polkovnikov, “Thermalization of dilute impurities in one-dimensional spin chains”, *Phys. Rev. X* **13**, 011041 (2023).
- [90] K. Hémero, F. Pollmann, and A. Smith, “Identifying correlation clusters in many-body localized systems”, *Phys. Rev. B* **105**, 064202 (2022).
- [91] J. Léonard, S. Kim, M. Rispoli, A. Lukin, R. Schittko, J. Kwan, E. Demler, D. Sels, and M. Greiner, “Probing the onset of quantum avalanches in a many-body localized system”, *Nature Physics* **19**, 481–485 (2023).
- [92] J. C. Peacock and D. Sels, “Many-body delocalization from embedded thermal inclusion”, *Phys. Rev. B* **108**, L020201 (2023).
- [93] A. Peres, “Stability of quantum motion in chaotic and regular systems”, *Phys. Rev. A* **30**, 1610–1615 (1984).
- [94] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, “Dynamics of Loschmidt echoes and fidelity decay”, *Physics Reports* **435**, 33–156 (2006).
- [95] A. I. Larkin and Y. N. Ovchinnikov, “Quasiclassical Method in the Theory of Superconductivity”, *Soviet Journal of Experimental and Theoretical Physics* **28**, 1200 (1969).
- [96] Y. Huang, Y.-L. Zhang, and X. Chen, “Out-of-time-ordered correlators in many-body localized systems”, *Annalen der Physik* **529**, 1600318 (2017).
- [97] R.-Q. He and Z.-Y. Lu, “Characterizing many-body localization by out-of-time-ordered correlation”, *Phys. Rev. B* **95**, 054201 (2017).
- [98] Y. Chen, “Universal logarithmic scrambling in many body localization”, (2016), [arXiv:1608.02765 \[cond-mat.dis-nn\]](#).
- [99] M. Serbyn, M. Knap, S. Gopalakrishnan, Z. Papić, N. Y. Yao, C. R. Laumann, D. A. Abanin, M. D. Lukin, and E. A. Demler, “Interferometric probes of many-body localization”, *Phys. Rev. Lett.* **113**, 147204 (2014).
- [100] G. De Tomasi, S. Bera, J. H. Bardarson, and F. Pollmann, “Quantum mutual information as a probe for many-body localization”, *Phys. Rev. Lett.* **118**, 016804 (2017).

- [101] A. S. Holevo, “Bounds for the quantity of information transmitted by a quantum communication channel”, *Problems Inform. Transmission* **9**, 177–183 (1973).
- [102] D. Yuan, S.-Y. Zhang, Y. Wang, L.-M. Duan, and D.-L. Deng, “Quantum information scrambling in quantum many-body scarred systems”, *Phys. Rev. Res.* **4**, 023095 (2022).
- [103] H. Bernien et al., “Probing many-body dynamics on a 51-atom quantum simulator”, *Nature* **551**, 579–584 (2017).
- [104] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, “Quantum scarred eigenstates in a Rydberg atom chain: entanglement, breakdown of thermalization, and stability to perturbations”, *Phys. Rev. B* **98**, 155134 (2018).
- [105] A. Omran et al., “Generation and manipulation of Schrödinger cat states in Rydberg atom arrays”, *Science* **365**, 570–574 (2019).
- [106] S. Dooley, “Robust quantum sensing in strongly interacting systems with many-body scars”, *PRX Quantum* **2**, 155134 (2021).
- [107] J.-Y. Desaulles, F. Pietracaprina, J. Goold, and S. Pappalardi, “Extensive multipartite entanglement from $su(2)$ quantum many-body scars”, *Phys. Rev. Lett.* **129**, 020601 (2022).
- [108] S. Dooley, S. Pappalardi, and J. Goold, “Entanglement enhanced metrology with quantum many-body scars”, *Phys. Rev. B* **107**, 035123 (2023).
- [109] I. Mondragon-Shem, M. G. Vavilov, and I. Martin, “Fate of quantum many-body scars in the presence of disorder”, *PRX Quantum* **2**, 030349 (2021).
- [110] W. Deng and Z.-C. Yang, “Using models with static quantum many-body scars to generate time-crystalline behavior under periodic driving”, *Phys. Rev. B* **108**, 205129 (2023).
- [111] K. Huang and X. Li, “Engineering subharmonic responses beyond prethermalization via Floquet scar states”, *Phys. Rev. B* **109**, 020601 (2024).
- [112] K. Sacha, “Discrete time crystals and related phenomena”, in *Time crystals*, Vol. 129 (Springer International Publishing, Cham, 2020), pp. 39–172.
- [113] M. Serbyn, D. A. Abanin, and Z. Papić, “Quantum many-body scars and weak breaking of ergodicity”, *Nature Physics* **17**, 675–685 (2021).
- [114] S. Moudgalya, B. A. Bernevig, and N. Regnault, “Quantum many-body scars and Hilbert space fragmentation: a review of exact results”, *Reports on Progress in Physics* **85**, 086501 (2022).
- [115] A. Chandran, T. Iadecola, V. Khemani, and R. Moessner, “Quantum many-body scars: a quasiparticle perspective”, *Annual Review of Condensed Matter Physics* **14**, 443–469 (2023).

- [116] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, “Fast quantum gates for neutral atoms”, *Phys. Rev. Lett.* **85**, 2208–2211 (2000).
- [117] M. D. Lukin, M. Fleischhauer, R. Cote, L. M. Duan, D. Jaksch, J. I. Cirac, and P. Zoller, “Dipole blockade and quantum information processing in mesoscopic atomic ensembles”, *Phys. Rev. Lett.* **87**, 037901 (2001).
- [118] E. Urban, T. A. Johnson, T. Henage, L. Isenhower, D. D. Yavuz, T. G. Walker, and M. Saffman, “Observation of Rydberg blockade between two atoms”, *Nature Physics* **5**, 110–114 (2009).
- [119] G. Cenedese, M. Bondani, A. Andreanov, M. Carrega, G. Benenti, and D. Rosa, *Shallow quantum circuits are robust hunters for quantum many-body scars*, 2024, [arXiv:2401.09279 \[quant-ph\]](#).
- [120] T. Iadecola, “Exact localized and ballistic eigenstates in disordered chaotic spin ladders and the Fermi-Hubbard model”, *Phys. Rev. Lett.* **123**, 036403 (2019).
- [121] Y. A. Kharkov, V. E. Sotskov, A. A. Karazeev, E. O. Kiktenko, and A. K. Fedorov, “Revealing quantum chaos with machine learning”, *Phys. Rev. B* **101**, 064406 (2020).
- [122] C.-D. Han, C.-Z. Wang, and Y.-C. Lai, “Meta-machine-learning-based quantum scar detector”, *Phys. Rev. Appl.* **19**, 064042 (2023).
- [123] H. Cao, D. G. Angelakis, and D. Leykam, *Unsupervised learning of quantum many-body scars using intrinsic dimension*, 2024, [arXiv:2401.07795 \[quant-ph\]](#).
- [124] S. Moudgalya and O. I. Motrunich, “Hilbert space fragmentation and commutant algebras”, *Phys. Rev. X* **12**, 011050 (2022).
- [125] S. Moudgalya and O. I. Motrunich, “Numerical methods for detecting symmetries and commutant algebras”, *Phys. Rev. B* **107**, 224312 (2023).
- [126] S. Moudgalya and O. I. Motrunich, “From symmetries to commutant algebras in standard Hamiltonians”, *Annals of Physics* **455**, 169384 (2023).
- [127] S. Moudgalya and O. I. Motrunich, *Exhaustive characterization of quantum many-body scars using commutant algebras*, 2023, [arXiv:2209.03377 \[cond-mat.str-el\]](#).
- [128] P. Łydźba, P. Prelovšek, and M. Mierzejewski, “Local integrals of motion in dipole-conserving models with Hilbert space fragmentation”, *Phys. Rev. Lett.* **132**, 220405 (2024).
- [129] S.-Y. Zhang, D. Yuan, T. Iadecola, S. Xu, and D.-L. Deng, “Extracting quantum many-body scarred eigenstates with matrix product states”, *Phys. Rev. Lett.* **131**, 020402 (2023).

- [130] P. Sierant, M. Lewenstein, A. Scardicchio, and J. Zakrzewski, “Stability of many-body localization in Floquet systems”, *Phys. Rev. B* **107**, 115132 (2023).
- [131] C. Lanczos, “An iteration method for the solution of the eigenvalue problem of linear differential and integral operators”, *Journal of Research of the National Bureau of Standards* **45**, 255282 (1950).

ARTICLES

Detecting ergodic bubbles at the crossover to many-body localization using neural networksTomasz Szoldra^{1,*}, Piotr Sierant^{2,1}, Korbinian Kottmann³, Maciej Lewenstein^{3,4} and Jakub Zakrzewski^{1,5}¹*Instytut Fizyki Teoretycznej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*²*The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, I-34151 Trieste, Italy*³*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology,**Avenida Carl Friedrich Gauss 3, E-08860 Castelldefels (Barcelona), Spain*⁴*ICREA, Passeig Lluís Companys 23, E-08010 Barcelona, Spain*⁵*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*

(Received 9 June 2021; revised 14 October 2021; accepted 15 October 2021; published 29 October 2021)

The transition between ergodic and many-body localized (MBL) phases is expected to occur via an avalanche mechanism, in which *ergodic bubbles* that arise due to local fluctuations in system properties thermalize their surroundings leading to delocalization of the system, unless the disorder is sufficiently strong to stop this process. We propose an algorithm based on neural networks that allows us to detect the ergodic bubbles using experimentally measurable two-site correlation functions. Investigating the time evolution of the system, we observe a logarithmic in time growth of the ergodic bubbles in the MBL regime. The distribution of the size of ergodic bubbles converges during time evolution to an exponentially decaying distribution in the MBL regime, and a power-law distribution with a thermal peak in the critical regime, supporting thus the scenario of delocalization through the avalanche mechanism. Our algorithm permits us to pinpoint the quantitative differences in the time evolution of systems with random and quasiperiodic potentials, as well as to identify rare (Griffiths) events. Our results open different pathways in studies of the mechanisms of thermalization of disordered many-body systems and beyond.

DOI: [10.1103/PhysRevB.104.L140202](https://doi.org/10.1103/PhysRevB.104.L140202)

Introduction. Many-body localization (MBL) [1,2] is a phenomenon that prevents strongly disordered quantum many-body systems from reaching thermal equilibrium [3–5]. In this dynamical phase, the transport is suppressed [6] due to the presence of a complete set of local integrals of motion [7–10] that account also for the logarithmic growth of entanglement entropy in time evolution [11–13], or the area-law entanglement of the eigenstates (for recent reviews, see Refs. [14,15]). While the properties of the MBL phase are considerably well understood, its stability in the thermodynamic limit has recently been vividly debated [16–21], and it may depend on various system specific properties [22].

The delocalization of MBL was proposed to occur via an “avalanche” mechanism [23,24]. In this approach, small ergodic regions (“ergodic bubbles”) immersed among insulating blocks delocalize their surroundings and grow at the expense of the localized regions. If the disorder is not sufficiently strong to stop this process, the system becomes delocalized. The avalanche mechanism was soon incorporated into the real space renormalization group approaches [25–28] that suggest the transition to MBL is of the Kosterlitz-Thouless type. This conclusion is supported by a structure of entanglement clusters in the eigenstates of a system at the MBL transition identified in Ref. [29] with the help of quantum mutual information [30]. A related concept of the “entanglement length” was introduced in Ref. [31].

The signatures of MBL have been observed experimentally in setups of ultracold atoms [32–35], ions [36], or superconducting qubits [37–39]. Most of the experiments investigated the time evolution of the initial product states probing correlations between the initial and time-evolved occupations of lattice sites. Strong correlations persisting in the long time limit of time evolution indicate the memory of the initial state and the onset of MBL [40]. Recent experiments [41–43] were able to directly measure the density correlations between various lattice sites, hence opening up new directions in investigations of quantum dynamics.

The aim of this Letter is to propose a scheme to detect and study the dynamics of the ergodic bubbles at the crossover to MBL relying on the two-site correlation functions that are directly measurable in experiments. In this way our study parallels recent attempts to understand the delocalization of the dynamics via the avalanche mechanism and many-body resonances scenario [44–46]. To achieve our goal we employ recurrent neural networks (RNNs) that are tailored to time-dependent data and have achieved unprecedented success in natural language processing tasks in recent years [47–49]. Our scheme allows us to identify the ergodic bubbles, i.e., the regions of the system, where the dynamics is ergodic, and to study their time evolution. We pinpoint the mechanism of delocalization of the MBL phase by probing the distribution of the ergodic bubbles. Comparing various types of disorder, we detect rare Griffiths regions [50–54] and discuss their impact on delocalization of the MBL phase.

*tomasz.szoldra@doctoral.uj.edu.pl

The models. From now on we concentrate on the random-field Heisenberg model with the Hamiltonian

$$H = J \sum_{i=1}^L \vec{S}_i \cdot \vec{S}_{i+1} + \sum_{i=1}^L W_i S_i^z, \quad (1)$$

where $\vec{S}_i$ are spin-1/2 operators, $J = 1$ is fixed as the energy unit, and periodic boundary conditions are assumed. Setting W_i to be independent random variables uniformly distributed in the interval $[-W, W]$, we obtain the model with random disorder, widely studied in the context of MBL [55–65]. Estimates of the critical disorder strength vary between $W_c^{\text{RD}} \approx 4$ [66,67] and $W_c^{\text{RD}} \gtrsim 5$ [31,68,69]. We consider also the quasiperiodic potential $W_i = W \cos(2\pi\zeta i + \phi)$, where $\phi \in [0, 2\pi]$ is a random phase, W characterizes the strength of the potential, and the golden ratio $\zeta = (\sqrt{5} - 1)/2$ is used. The transition to the MBL phase for that ζ value occurs at $W_c^{\text{QP}} \approx 2.5$ [70–73]. The random potential features rare regions in which disorder may be anomalously weak or strong. There are no such fluctuations in the quasiperiodic case. Hence, the MBL transition may be qualitatively different in the two cases [74].

In what follows we use RNNs to analyze the approach to thermal equilibrium in the course of the time evolution of system (1). The neural networks have been applied to studies of properties of eigenstates of disordered quantum many-body systems [75–78] or in investigations of gross features of their time dynamics [79–81].

Ergodic bubble detection. We define a set of neighboring spins to be an *ergodic bubble* (or *cluster*) if the local observables associated with them have properties characteristic to an ergodic/thermal phase. In particular, the observables undergo fluctuations in the course of time evolution and this feature may be used for their direct detection. Nevertheless, providing a well-defined criterion for distinguishing the ergodic from the nonergodic evolutions remains a nontrivial task. In this section we propose an algorithm employing neural networks that learns how the ergodic evolutions “look like,” and later assigns the time evolutions of experimentally accessible observables to an ergodic or nonergodic class.

At the core of our framework is the anomaly detection scheme [82] (similar techniques were recently used for mapping out phase diagrams of the quantum many-body systems [83–85]). Precisely speaking, we consider a neural network of an autoencoder architecture, whose goal is to spot the characteristic features of the input data, “compress” it into a latent representation (the *encoder* network), and, from there, to accurately reproduce the original input (the *decoder* network)—compare Fig. 1. The autoencoder is trained on the *normal* data until it efficiently reproduces them, and then evaluated on a data set consisting of *normal* and *anomalous* data to identify the *anomalous* data based on the high reconstruction loss. In our case, the *normal* data correspond to the time series of observables in the ergodic regime at weak disorder, as opposed to the nonergodic evolutions for large disorder, regarded as the *anomalous* data.

Specifically, as input data we use the two-site correlation functions $C_{i,d}(t) = \langle S_i^z(t) S_{i+d}^z(t) \rangle$. Let us define a collection of these correlations $x_i(t)$ that is supposed to characterize

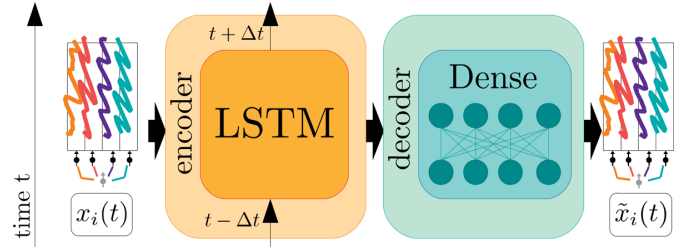


FIG. 1. Neural network architecture: A snapshot $x_i(t)$ of two-point correlation functions is input to the multichannel time series autoencoder. The object $x_i(t)$ is processed by two consecutive LSTM layers in the encoder. The LSTMs additionally receive (pass) a hidden state from (to) the previous (next) time step, thus creating a time order. The decoder is simply composed of three dense (fully connected) layers. The network outputs the time series $\tilde{x}_i(t)$ that reconstructs the input signal $x_i(t)$.

site i at time t . It consists of correlation functions $C_{i,d}(t)$ corresponding to d_0 neighbors of site i , sampled across n_t discrete points in time,

$$x_i(t) = \{C_{i,d}(t + n\Delta t)\}_{d=-d_0, \dots, -1, 1, \dots, d_0, n=0, \dots, n_t-1}, \quad (2)$$

where, in our implementation, we take $\Delta t = 0.5$, $d_0 = 2$, and $n_t = 32$. The four-channel time series of length 32 starting at time t , $x_i(t)$, constitutes a single input to the neural network. Having a trained autoencoder network we can perform an anomaly detection by measuring the reconstruction loss $l_i(t) = \|\tilde{x}_i(t) - x_i(t)\|$, where $\|\cdot\|$ is the Euclidean norm (see Fig. 2). In order to classify a site i at time t , one needs to set a loss cutoff parameter l_{cut} . In our algorithm, the thermal bubble is defined as a set of adjacent sites which satisfy $l_i \leq l_{\text{cut}}$, that is, for which the network succeeds in reconstructing the temporal evolution of observables with loss no larger than l_{cut} . Thus, we have a quantitative tool to detect the ergodic clusters

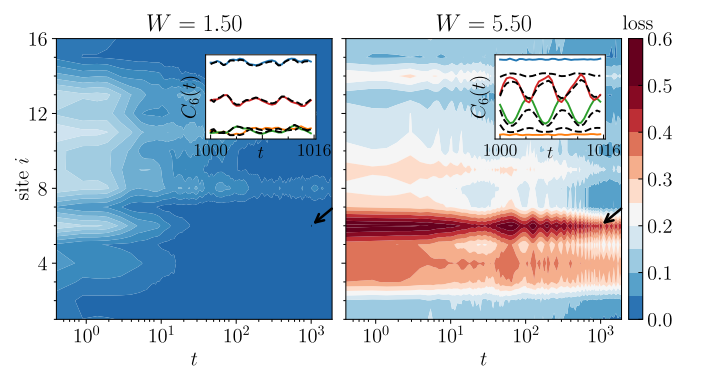


FIG. 2. Neural network loss—a measure of nonergodicity in the system. In the ergodic regime $W = 1.5$, the nonergodic (high loss) regions quickly disappear; they persist in the MBL phase at $W = 5.5$. Both evolutions were calculated for the same uniform disorder realization but with a different amplitude. Insets show the two-site correlations (input of the NN, solid lines), Eq. (2), and the NN output (dashed lines). The NN learned to reconstruct ergodic evolutions perfectly (left inset). It performs significantly worse on the nonergodic data (right inset) it has not encountered during training, as expected in the anomaly detection scheme.

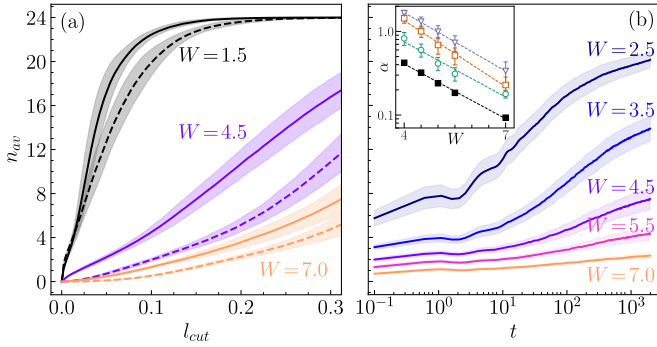


FIG. 3. (a) The average size n_{av} of the ergodic bubbles as a function of the loss cutoff l_{cut} at time $t = 20$ (dashed lines) and $t = 1980$ (solid lines) for random disorder. The ergodic ($W = 1.5$), critical ($W = 4.5$), and MBL ($W = 7.0$) regimes differ qualitatively by the dependence of n_{av} on l_{cut} and by the amount the bubbles grow in time. (b) Logarithmic in time growth of the average bubble size for $l_{cut} = 0.15$ and several disorder strengths W . Inset: See text.

and to differentiate them from sites at which the anomaly is detected and the dynamics is not ergodic.

Among many neural network architectures, particularly well suited for multichannel time series processing, are RNNs that naturally utilize the temporal order by recursively including the data from previously processed time steps. In our case, the *encoder* part of the network consists of a special type of RNN with two layers of long short-term memory (LSTM) architecture [86] with 64 and 32 units (see Fig. 1). As a decoder we use three dense time-distributed layers of dimension 128 and 64 and 32. The training data set consists of time evolutions from $t = 100$ to $t = 2000$ corresponding to 100 realizations of uniformly distributed disorder of strength $W \in \{0.1, 0.2, \dots, 0.5\}$ for system size $L = 16$. For a detailed comparison with the networks trained on larger values of disorder, as well as on the quasiperiodic distribution of disorder, see the Supplemental Material [87]. We stop the training after a certain value 1.7×10^{-3} of the validation loss is achieved.

Dynamics of ergodic bubbles. We denote the average in a given state as $\langle \cdot \rangle_\psi$ and average over 2400 disorder realizations as $\langle \cdot \rangle_W$. To control the risk that a single neural network model accidentally learns a random feature of the data which another model does not, we independently train 30 models on the same data and denote the average over them as $\langle \cdot \rangle_N$. All error bars in the following plots correspond to one standard deviation in this averaging. We will consider the distributions of the number of spins n in an ergodic bubble, as well as the average size of ergodic bubbles $n_{av} \equiv \langle \langle n \rangle_\psi \rangle_W \rangle_N$. If there are no ergodic bubbles in a time-evolved state $|\psi\rangle$, we manually define $\langle n \rangle_\psi = 0$. Initially, the spins are fully uncorrelated and the initial state is chosen to be a Néel state $|\psi\rangle = |\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\cdots\rangle$. The time evolution with Hamiltonian (1) is calculated numerically by means of a kernel polynomial method [88–90] for a system consisting of $L = 24$ spins.

We start our studies of the MBL transition by checking how the average size of ergodic bubbles depends on the cutoff l_{cut} [see Fig. 3(a)]. In the ergodic regime the increase of n_{av} with l_{cut} is much sharper than for the critical and MBL regimes which signifies that our algorithm indeed detects the prevalent

ergodicity in the system. Moreover, comparing the results for $t = 20$ and $t = 1980$, we observe that the number of ergodic regions increase in the course of the evolution, but the growth is much smaller in the MBL than in the ergodic and critical regimes. We may now choose a “reasonable” value of the cutoff $l_{cut} = 0.15$, for which, at $W = 1.5$, the neural network treats almost the entire chain as a large thermal bubble and in the critical regime it finds a mixture of ergodic bubbles and nonergodic regions.

In Fig. 3(b) we present the bubble dynamics for the chosen value of cutoff $l_{cut} = 0.15$ and different disorder strengths. Due to a limited temporal resolution of our classification scheme [the time window in which the snapshot $x_i(t)$ of two-site correlation functions has a length of $n_t \Delta t = 16$], in the initial stage of the evolution we detect ergodic clusters of nonvanishing size. In the course of time evolution the ergodic bubbles thermalize their surroundings and, consequently, the average size of the ergodic bubbles grows with time. Our results indicate that the growth of the bubbles is logarithmic in time.

The inset of Fig. 3(b) shows the logarithmic growth rates α , fitted from $n_{av}(t) = \alpha \log t + C$ for different loss cutoffs $l_{cut} = 0.1, 0.15, 0.2$ (from top to bottom, open markers), considered as a function of W . They are surprisingly well compatible with the similar logarithmic growth rates of the bipartite von Neumann entanglement entropy $S(\mathcal{A}) = -\text{Tr}_{\mathcal{A}}(\rho_{\mathcal{A}} \ln \rho_{\mathcal{A}})$, where $\rho_{\mathcal{A}}$ is the reduced density matrix of subsystem $\mathcal{A}$ consisting of spins indexed by $i = 1 \dots L/2$ (bottom, solid markers). That is, in both cases, $\alpha(W) \approx C \exp(-\beta W)$, where $\beta = (0.55 \pm 0.05)$ while C is different for entropy and clusters for dimensional reasons. Importantly, this conclusion holds for arbitrary cutoff $l_{cut} \in [0.1, 0.2]$, indicating that our algorithm serves also as an indirect qualitative measure of the entanglement entropy. The decrease of the growth rate of ergodic bubbles with disorder strength W is consistent with the exponential slowdown of dynamics [18,91] with increasing W at the crossover to MBL regime.

The qualitative correspondence between the growth of ergodic bubbles and spreading of the entanglement in the system can be explained by a simple argument. Let us assume that the chain contains an ergodic cluster of length n . If the cut that separates the subsystem $\mathcal{A}$ from its surroundings does not split the cluster, the contribution from this realization to the average entropy is relatively small. However, if the cut splits the cluster, the contribution will be much larger because $\mathcal{A}$ contains spins strongly entangled with the rest of the ergodic bubble. According to the volume law of entanglement, this contribution will be proportional to the number of cluster sites included in $\mathcal{A}$ which is proportional to n_{av} . Therefore, we may conclude that S is proportional to n_{av} . Moreover, we are convinced that the network detects strongly entangled parts of the system without the need to directly quantify the entanglement.

Distribution of ergodic bubbles. Further insights into MBL transition are obtained from an analysis of the distributions of the ergodic bubble size n . In Fig. 4(a) we present such distributions for random disorder, in the initial stage $t = 20$, and after a long evolution time $t = 1980$. In both cases, in the MBL regime ($W = 7$), the distribution decays exponentially as $\exp(-n\gamma)$, consistently with an intuitive expectation

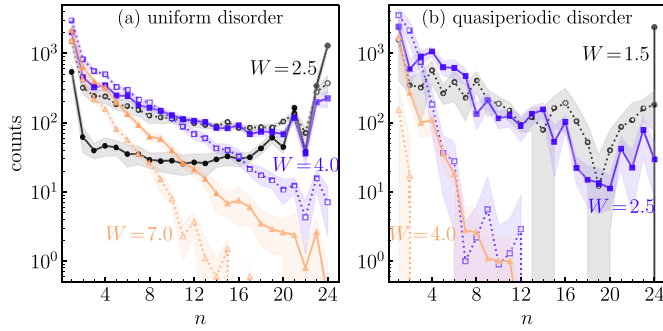


FIG. 4. Histograms of the bubble sizes n at time $t = 20$ (open markers) and $t = 1980$ (solid markers), at a cutoff level $l_{\text{cut}} = 0.15$ and (a) uniform and (b) quasiperiodic distributions of disorder, 2400 realizations.

that ergodic bubbles correspond to rare events in which the disorder is anomalously weak on n neighboring lattices sites. Indeed, if $p < 1$ is the probability to have the weak disorder on a certain site, the probability to find n such sites is $p^n = e^{n \log p}$. In the vicinity of the critical regime ($W = 4$), the ergodic bubbles are distributed, at $t = 1980$, according to a heavy-tailed distribution, the determination of which is a hard problem given the data change by less than one order of magnitude. There emerges a thermal peak at $n \approx L$. Indications of a heavy-tailed distribution of the power-law type $n^{-\beta}$ and the presence of a thermal peak in the critical regime were observed in renormalization group schemes [92] as well as in exact diagonalization studies [93]. Exactly at the transition one expects the exponent β governing the decay of ergodic clusters to be equal to 2 [27]. The smaller value of β for $W = 4$ suggests that system becomes ergodic in the thermodynamic limit at this disorder strength. Here, $\beta = 2$ is obtained around $W = 5.3$. While pinpointing exactly the transition point would require much larger system sizes and evolution times, our results support the picture of the MBL transition occurring via rare avalanches [23]. Finally, on the ergodic side of the crossover ($W = 2.5$) we observe the abundance of dominant thermal clusters of size $n \approx L$. During the time evolution from $t = 20$ to $t = 1980$, the number of large ergodic bubbles grow in time at the expense of the smaller ones. This process eventually leads to delocalization of the whole system. The transition from an exponential to a heavy-tailed distribution in Fig. 4, with the values of β for $W = 3.5$ – 7.0 , is presented in the Supplemental Material [87].

The situation is qualitatively similar for the system with a quasiperiodic potential as shown in Fig. 4(b). In the MBL regime ($W = 4$) the probability to detect an ergodic bubble of size n is exponentially small in n . This suggests a presence of a similar rare event mechanism as in the case of random disorder. Such a mechanism, in the absence of fluctuations in the quasiperiodic potential, might be associated with rare configurations of the state of the system [94]. In the critical regime ($W = 2.5$), we find that the bubble size is distributed according to a power law $n^{-2.2}$, although there are no traces of the thermal peak at $n \approx L$. This suggests that delocalization in a quasiperiodic system occurs, similarly to the random case, via an avalanche mechanism [95]. Finally, in the ergodic

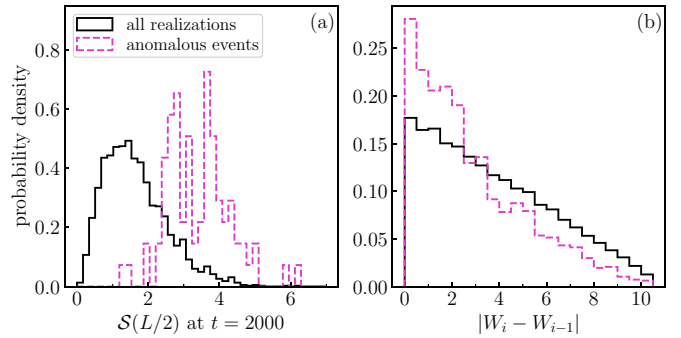


FIG. 5. Properties of 2400 realizations of uniform disorder $W = 5.5$ (solid black line), including those 74 for which very large clusters $n \geq L - 1 = 23$ were detected in the course of evolution from $t = 0$ to $t = 1980$ by a single neural network model (dashed pink line).

regime ($W = 1.5$), the broad distribution of ergodic bubbles at small time $t = 20$ quickly evolves into a single thermal peak at $n \approx L$. The small variance of the size of ergodic bubbles both in ergodic and thermal regimes suggests that the MBL transition in quasiperiodic systems is more stable than in the random case [96].

Griffiths regions. For the random disorder, even at $W = 5.5$, our algorithm detects ergodic bubbles of size $n \geq L - 1$ (see the Supplemental Material [87] for a corresponding plot similar to Fig. 2). Figure 5 shows a comparison of the properties of all disorder realizations with the properties of those anomalous realizations. Exceptionally large bipartite entanglement entropy $S(\mathcal{A})$ and strongly resonant values of disorder distinguish these anomalous realizations from the background. This shows that our algorithm detects rare Griffiths events that arise due to fluctuations in the random disorder. For a quasiperiodic potential the atypically large bubbles were not detected at all if the cutoffs l_{cut} were set at the same level. Even if we tune the cutoffs so that the number of detected anomalous realizations is comparable to the uniform disorder case, their features are not distinctive (for details, see the Supplemental Material [87]).

Summary. We proposed an algorithm based on neural networks, which allows us to detect and study the dynamics of ergodic bubbles in disordered many-body systems—the subject intensively discussed recently [44–46]. The algorithm learns itself the features of ergodic time evolution and employs the anomaly detection scheme to identify lattice sites, at which the evolution is nonergodic. Detected ergodic bubbles grow logarithmically in time in the MBL regime, which indicates that our algorithm captures the features of entanglement in the system without quantifying it directly. The distributions of the size of ergodic bubble, exponential in the MBL regime and heavy tailed with a thermal peak in the critical regime, support the avalanche scenario of delocalization of the MBL phase. Those results are in qualitative agreement with the results obtained when the algorithm of Ref. [29] is used to identify entanglement clusters in the state of the system *during time evolution*, as we show in the Supplemental Material [87]. The approach of Ref. [29] relies on the quantum mutual information of numerous subsystems that is very hard to measure in practice. In contrast, our approach employs only two-site correlation functions readily accessible

in present-day experiments. Hence, our algorithm allows for a quantitative investigation of the mechanisms of thermalization at the MBL transition, not only in the numerics, but also in the experimental setup.

Acknowledgments. We would like to thank Fabien Alet for helpful comments and a critical reading of the manuscript. T.S. and J.Z. acknowledge support by PL-Grid Infrastructure and by National Science Centre (Poland) under OPUS-18 Project No. 2019/35/B/ST2/00034. M.L. acknowledges support from ERC AdG NOQIA, Agencia Estatal de Investigación (“Severo Ochoa” Center of Excellence CEX2019-000910-S, Plan National FIDEUA PID2019-106901GB-I00/10.13039/501100011033, FPI), Fundació Privada Cellex, Fundació Mir-Puig, and from Generalitat de Catalunya (AGAUR Grant No. 2017 SGR 1341, CERCA program, QuantumCAT _U16-011424, cofunded by ERDF

Operational Program of Catalonia 2014-2020), MINECO-EU QUANTERA MAQS [funded by State Research Agency (AEI) PCI2019-111828-2/10.13039/501100011033], EU Horizon 2020 FET-OPEN OPTOLogic (Grant No. 899794), and the National Science Centre, Poland-Symfonia Grant No. 2016/20/W/ST4/00314, Marie Skłodowska-Curie Grant STRETCH No. 101029393, “La Caixa” Junior Leaders fellowships (ID100010434), and EU Horizon 2020 under Marie Skłodowska-Curie grant agreement No. 847648 (LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012). This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 713729 (K.K.). P.S. acknowledges the support of Foundation for Polish Science (FNP) through scholarship START.

-
- [1] D. Basko, I. Aleiner, and B. Altshuler, *Ann. Phys. (NY)* **321**, 1126 (2006).
 - [2] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
 - [3] J. M. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
 - [4] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
 - [5] L. D’Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.* **65**, 239 (2016).
 - [6] R. Nandkishore and D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
 - [7] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
 - [8] V. Ros, M. Mueller, and A. Scardicchio, *Nucl. Phys. B* **891**, 420 (2015).
 - [9] J. Z. Imbrie, *Phys. Rev. Lett.* **117**, 027201 (2016).
 - [10] M. Mierzejewski, M. Kozarzewski, and P. Prelovšek, *Phys. Rev. B* **97**, 064204 (2018).
 - [11] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
 - [12] J. H. Bardarson, F. Pollmann, and J. E. Moore, *Phys. Rev. Lett.* **109**, 017202 (2012).
 - [13] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
 - [14] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, *Rev. Mod. Phys.* **91**, 021001 (2019).
 - [15] F. Alet and N. Laflorencie, *C. R. Phys.* **19**, 498 (2018).
 - [16] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, *Phys. Rev. E* **102**, 062144 (2020).
 - [17] R. K. Panda, A. Scardicchio, M. Schulz, S. R. Taylor, and M. Znidaric, *Europhys. Lett.* **128**, 67003 (2020).
 - [18] P. Sierant, D. Delande, and J. Zakrzewski, *Phys. Rev. Lett.* **124**, 186601 (2020).
 - [19] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, *Phys. Rev. Lett.* **124**, 243601 (2020).
 - [20] D. Abanin, J. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. Parameswaran, F. Pollmann, A. Potter, M. Serbyn, and R. Vasseur, *Ann. Phys.* **427**, 168415 (2021).
 - [21] D. Sels and A. Polkovnikov, *arXiv:2009.04501*.
 - [22] P. Sierant, E. Gonzalez Lazo, M. Dalmonte, A. Scardicchio, and J. Zakrzewski, *Phys. Rev. Lett.* **127**, 126603 (2021).
 - [23] W. De Roeck and F. Huveneers, *Phys. Rev. B* **95**, 155129 (2017).
 - [24] D. J. Luitz, F. Huveneers, and W. De Roeck, *Phys. Rev. Lett.* **119**, 150602 (2017).
 - [25] T. Thiery, F. Huveneers, M. Müller, and W. De Roeck, *Phys. Rev. Lett.* **121**, 140601 (2018).
 - [26] A. Goremykina, R. Vasseur, and M. Serbyn, *Phys. Rev. Lett.* **122**, 040601 (2019).
 - [27] P. T. Dumitrescu, A. Goremykina, S. A. Parameswaran, M. Serbyn, and R. Vasseur, *Phys. Rev. B* **99**, 094205 (2019).
 - [28] A. Morningstar and D. A. Huse, *Phys. Rev. B* **99**, 224205 (2019).
 - [29] L. Herviou, S. Bera, and J. H. Bardarson, *Phys. Rev. B* **99**, 134205 (2019).
 - [30] G. De Tomasi, S. Bera, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **118**, 016804 (2017).
 - [31] J. Gray, S. Bose, and A. Bayat, *Phys. Rev. B* **97**, 201105(R) (2018).
 - [32] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**, 842 (2015).
 - [33] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, *Science* **352**, 1547 (2016).
 - [34] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, *Phys. Rev. Lett.* **119**, 260401 (2017).
 - [35] H. P. Lüschen, S. Scherg, T. Kohlert, M. Schreiber, P. Bordia, X. Li, S. Das Sarma, and I. Bloch, *Phys. Rev. Lett.* **120**, 160404 (2018).
 - [36] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, *Nat. Phys.* **12**, 907 (2016).
 - [37] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, B. Foxen, M. Giustina, E. Jeffrey, J. Kelly, E. Lucero, J. Mutus, M. Neeley, C. Quintana, and D. Sank, *Science* **358**, 1175 (2017).
 - [38] Q. Guo, C. Cheng, Z.-H. Sun, Z. Song, H. Li, Z. Wang, W. Ren, H. Dong, D. Zheng, Y.-R. Zhang, R. Mondaini, H. Fan, and H. Wang, *Nat. Phys.* **17**, 234 (2021).

- [39] M. Gong, G. D. de Moraes Neto, C. Zha, Y. Wu, H. Rong, Y. Ye, S. Li, Q. Zhu, S. Wang, Y. Zhao, F. Liang, J. Lin, Y. Xu, C.-Z. Peng, H. Deng, A. Bayat, X. Zhu, and J.-W. Pan, *Phys. Rev. Research* **3**, 033043 (2021).
- [40] A. Nico-Katz, A. Bayat, and S. Bose, [arXiv:2009.04470](https://arxiv.org/abs/2009.04470).
- [41] M. Rispoli, A. Lukin, R. Schittko, S. Kim, M. E. Tai, J. Léonard, and M. Greiner, *Nature (London)* **573**, 385 (2019).
- [42] Q. Guo, C. Cheng, H. Li, S. Xu, P. Zhang, Z. Wang, C. Song, W. Liu, W. Ren, H. Dong, R. Mondaini, and H. Wang, [arXiv:2011.13895](https://arxiv.org/abs/2011.13895).
- [43] W. Morong, F. Liu, P. Becker, K. S. Collins, L. Feng, A. Kyprianidis, G. Pagano, T. You, A. V. Gorshkov, and C. Monroe, [arXiv:2102.07250](https://arxiv.org/abs/2102.07250).
- [44] P. J. D. Crowley and A. Chandran, [arXiv:2012.14393](https://arxiv.org/abs/2012.14393).
- [45] A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, [arXiv:2107.05642](https://arxiv.org/abs/2107.05642).
- [46] D. Sels, [arXiv:2108.10796](https://arxiv.org/abs/2108.10796).
- [47] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, Stroudsburg, PA, 2014), pp. 1724–1734.
- [48] I. Sutskever, O. Vinyals, and Q. V. Le, *Proc. Adv. Neural Inf. Process. Syst.* **4**, 3104 (2014).
- [49] D. Bahdanau, K. H. Cho, and Y. Bengio, in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (International Conference on Learning Representations, ICLR, 2015).
- [50] T. Vojta, *J. Low Temp. Phys.* **161**, 299 (2010).
- [51] S. Gopalakrishnan, M. Müller, V. Khemani, M. Knap, E. Demler, and D. A. Huse, *Phys. Rev. B* **92**, 104202 (2015).
- [52] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, *Phys. Rev. Lett.* **114**, 160401 (2015).
- [53] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, *Ann. Phys.* **529**, 1600326 (2017).
- [54] N. Pancotti, M. Knap, D. A. Huse, J. I. Cirac, and M. C. Bañuls, *Phys. Rev. B* **97**, 094206 (2018).
- [55] L. F. Santos, G. Rigolin, and C. O. Escobar, *Phys. Rev. A* **69**, 042304 (2004).
- [56] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [57] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [58] T. C. Berkelbach and D. R. Reichman, *Phys. Rev. B* **81**, 224429 (2010).
- [59] S. Bera, H. Schomerus, F. Heidrich-Meisner, and J. H. Bardarson, *Phys. Rev. Lett.* **115**, 046603 (2015).
- [60] T. Enss, F. Andraschko, and J. Sirker, *Phys. Rev. B* **95**, 045121 (2017).
- [61] P. Sierant and J. Zakrzewski, *Phys. Rev. B* **101**, 104201 (2020).
- [62] L. Colmenarez, P. A. McClarty, M. Haque, and D. J. Luitz, *SciPost Phys.* **7**, 64 (2019).
- [63] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, *Phys. Rev. B* **102**, 064207 (2020).
- [64] N. Laflorencie, G. Lemarié, and N. Macé, *Phys. Rev. Research* **2**, 042033(R) (2020).
- [65] L. Vidmar, B. Krajewski, J. Bonca, and M. Mierzejewski, [arXiv:2105.09336](https://arxiv.org/abs/2105.09336).
- [66] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103(R) (2015).
- [67] N. Macé, F. Alet, and N. Laflorencie, *Phys. Rev. Lett.* **123**, 180601 (2019).
- [68] T. Devakul and R. R. P. Singh, *Phys. Rev. Lett.* **115**, 187201 (2015).
- [69] P. Sierant, M. Lewenstein, and J. Zakrzewski, *Phys. Rev. Lett.* **125**, 156601 (2020).
- [70] S. Iyer, V. Oganesyan, G. Refael, and D. A. Huse, *Phys. Rev. B* **87**, 134202 (2013).
- [71] M. Lee, T. R. Look, S. P. Lim, and D. N. Sheng, *Phys. Rev. B* **96**, 075146 (2017).
- [72] S. Bera, T. Martynec, H. Schomerus, F. Heidrich-Meisner, and J. H. Bardarson, *Ann. Phys.* **529**, 1600356 (2017).
- [73] E. V. H. Doggen and A. D. Mirlin, *Phys. Rev. B* **100**, 104203 (2019).
- [74] V. Khemani, D. N. Sheng, and D. A. Huse, *Phys. Rev. Lett.* **119**, 075702 (2017).
- [75] F. Schindler, N. Regnault, and T. Neupert, *Phys. Rev. B* **95**, 245134 (2017).
- [76] Y.-T. Hsu, X. Li, D.-L. Deng, and S. Das Sarma, *Phys. Rev. Lett.* **121**, 245701 (2018).
- [77] P. Huembeli, A. Dauphin, P. Wittek, and C. Gogolin, *Phys. Rev. B* **99**, 104106 (2019).
- [78] H. Théveniaut and F. Alet, *Phys. Rev. B* **100**, 224202 (2019).
- [79] E. van Nieuwenburg, E. Bairey, and G. Refael, *Phys. Rev. B* **98**, 060301(R) (2018).
- [80] E. V. H. Doggen, F. Schindler, K. S. Tikhonov, A. D. Mirlin, T. Neupert, D. G. Polyakov, and I. V. Gornyi, *Phys. Rev. B* **98**, 174202 (2018).
- [81] A. Bohrdt, S. Kim, A. Lukin, M. Rispoli, R. Schittko, M. Knap, M. Greiner, and J. Léonard, *Phys. Rev. Lett.* **127**, 150504 (2021).
- [82] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, *ACM Comput. Surv.* **54**, 1 (2021).
- [83] K. Kottmann, P. Huembeli, M. Lewenstein, and A. Acín, *Phys. Rev. Lett.* **125**, 170603 (2020).
- [84] N. Käming, A. Dawid, K. Kottmann, M. Lewenstein, K. Sengstock, A. Dauphin, and C. Weitenberg, *Mach. Learn.: Sci. Technol.* **2**, 035037 (2021).
- [85] K. Kottmann, P. Corboz, M. Lewenstein, and A. Acín, *SciPost Phys.* **11**, 025 (2021).
- [86] S. Hochreiter and J. Schmidhuber, *Neural Comput.* **9**, 1735 (1997).
- [87] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.104.L140202> for further analysis of the bubble size histograms, comparison with an entanglement-based clustering method, and details of the neural network training.
- [88] H. Tal-Ezer and R. Kosloff, *J. Chem. Phys.* **81**, 3967 (1984).
- [89] C. Leforestier, R. Bisseling, C. Cerjan, M. Feit, R. Friesner, A. Gulberg, A. Hammerich, G. Jolicard, W. Karlein, H.-D. Meyer, N. Lipkin, O. Roncero, and R. Kosloff, *J. Comput. Phys.* **94**, 59 (1991).
- [90] H. Fehske and R. Schneider, *Computational Many-Particle Physics* (Springer, Berlin, 2008).
- [91] T. Chanda, P. Sierant, and J. Zakrzewski, *Phys. Rev. Research* **2**, 032045(R) (2020).

- [92] P. T. Dumitrescu, R. Vasseur, and A. C. Potter, [Phys. Rev. Lett. **119**, 110604 \(2017\)](#).
- [93] X. Yu, D. J. Luitz, and B. K. Clark, [Phys. Rev. B **94**, 184202 \(2016\)](#).
- [94] S. Gopalakrishnan and S. Parameswaran, [Phys. Rep. **862**, 1 \(2020\)](#).
- [95] S.-X. Zhang and H. Yao, [Phys. Rev. Lett. **121**, 206601 \(2018\)](#).
- [96] S.-X. Zhang and H. Yao, [arXiv:1906.00971](#).

Supplemental material for “Detecting ergodic bubbles at the crossover to many-body localization using neural networks”

Tomasz Szoldra,¹ Piotr Sierant,^{2,1} Korbinian Kottmann,³ Maciej Lewenstein,^{3,4} and Jakub Zakrzewski^{1,5}

¹*Instytut Fizyki Teoretycznej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*

²*The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, 34151, Trieste, Italy*

³*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*

⁴*ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain*

⁵*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, Kraków, Poland*

(Dated: October 14, 2021)

ERGODIC BUBBLE SIZE HISTOGRAMS

In FIG. S1 we show how the ergodic bubble size distribution changes from a heavy-tailed decay (here we assume a power-law $Cn^{-\beta}$) with a thermal peak in the critical regime $W = 4$, into an exponential decay $C \exp(-n\gamma)$ when $W = 7$ in the many-body localized phase. All fits are performed on $n = 3 \dots 22$ to minimize the influence of the thermal peak on the fit, and averaged over 10 independently trained neural networks with the corresponding error bars representing the standard deviation of this averaging. We extract the fitted parameters and present them in FIGs. S1(d),(e). We notice that on the MBL side γ increases with an increasing disorder strength. Moreover, it is expected [1] that at the critical point $\beta = 2$ and our results range from $\beta = 0.62(14)$ at $W = 3.5$ to $\beta = 3.20(26)$ at $W = 7.0$, with $\beta = 2$ around $W = 5.3$.

In FIG. S2 we provide an example of an anomalous re-alization/Griffiths event, detected by the neural network. An initially strongly non-ergodic regions thermalizes after some time due to the interaction with its ergodic surroundings.

ALTERNATIVE CLUSTERING ALGORITHM BASED ON QUANTUM MUTUAL INFORMATION

In this part of the Supplemental material we compare our neural network-based algorithm with a scheme introduced in [2] that detects so-called *entanglement clusters*. We start with a short revision of this algorithm, point out important differences between entanglement- and ergodic clusters, and then apply this method to the time evolving quantum states, obtaining results in qualitative agreement with our neural network approach.

The input to the entanglement clustering algorithm is the many-body wave function $|\psi\rangle$. After the clusterization procedure, entanglement clusters should contain indices of spins strongly entangled with each other and, at the same time, weakly entangled between separate clusters. The clusters are constructed starting from the whole system and subdividing it recursively into two parts $\mathcal{A}$, $\mathcal{B}$

containing $n_{\mathcal{A}}$, $n_{\mathcal{B}}$ sites each, at every step minimizing the normalized quantum mutual information (QMI) of the splitting,

$$0 \leq i(\mathcal{A}, \mathcal{B}) = \frac{\mathcal{S}(\mathcal{A}) + \mathcal{S}(\mathcal{B}) - \mathcal{S}(\mathcal{A} \cup \mathcal{B})}{\min(n_{\mathcal{A}}, n_{\mathcal{B}})} \leq 2 \ln 2, \quad (\text{S1})$$

where the von Neumann entropy of subsystem $\mathcal{A}$ reads

$$\mathcal{S}(\mathcal{A}) = -\text{Tr}_{\mathcal{A}}(\rho_{\mathcal{A}} \ln \rho_{\mathcal{A}}), \quad (\text{S2})$$

$\rho_{\mathcal{A}}$ is its reduced density matrix, and periodic boundary conditions are assumed. The number of all such possible bipartitions to choose from grows exponentially with the system size. To make the problem numerically tractable, only a certain class, the *continuous* bipartitions, are left as candidates for the splitting. (“Continuous” means that the splitting does not introduce holes in a single cluster. For example, a cluster consisting of sites $\{1, 2, 3, 4, 5, 6\}$ can be divided into $\{1, 2, 5, 6\}$ and $\{3, 4\}$ but not into $\{1, 3, 5, 6\}$ and $\{2, 4\}$. However, holes in the clusters may occur due to periodic boundary conditions assumed at each step, e.g. splitting $\{1, 2, 5, 6\}$ can result in $\{1, 6\}$ and $\{2, 5\}$.) Splitting is performed recursively until the cluster size is equal to 1 and the final representation of the wave function forms a binary tree. Each node of the tree contains indices of the lattice sites, as well as the QMI of the minimal bipartition i . Then, the user chooses a certain cutoff of the quantum mutual information i_{cut} and traverses the tree starting at the root. If the current node contains at least two sites and $i < i_{\text{cut}}$, one goes down to the node’s two children. In the opposite case, $i \geq i_{\text{cut}}$, the list of sites at the current node is considered an *entanglement cluster*. A list of all such clusters is the final output of the clustering algorithm at the QMI cutoff level i_{cut} .

There is an important difference in the sense of entanglement and ergodic clusters: every site must be a member of one and only one entanglement cluster whereas it doesn’t necessarily have to be assigned to an ergodic cluster; the observables associated with single spins from an entanglement cluster may show either ergodic- or non-ergodic kind of dynamics. This means that the QMI clustering algorithm detects a somewhat different family of clusters than the neural network. On the other hand, it is

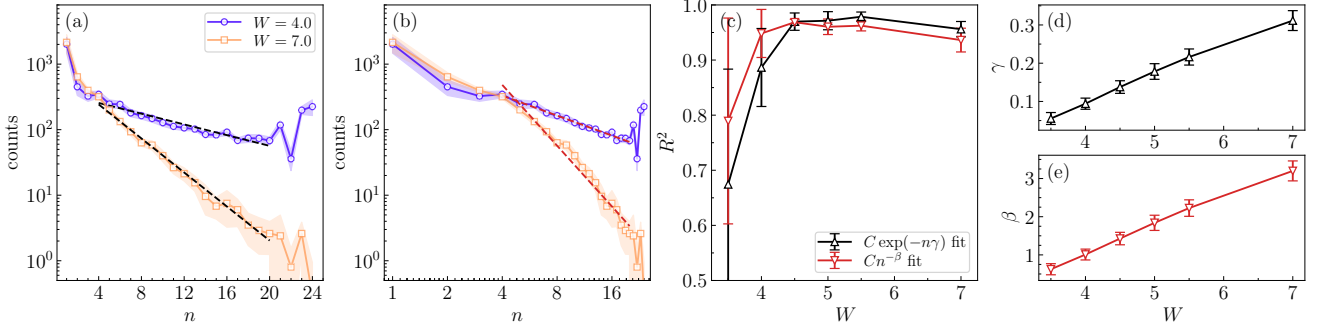


FIG. S1. Transition from an exponential to a power-law decay of ergodic bubble size n distributions for loss cutoff $l_{\text{cut}} = 0.15$ and system size $L = 24$. (a): Bubble size histograms at large time $t = 1980$. Exponential fit $C \exp(-n\gamma)$ (dashed line) deviates from the $W = 4.0$ datapoints and agrees very well with those for $W = 7.0$. (b): same as (a) but with power-law fits $C n^{-\beta}$ (notice the logarithmic horizontal scale). Here, the fit describes datapoints at $W = 4.0$ better than for $W = 7.0$. All fits were performed between $n = 3$ and $n = 20$ only. (c): Measure of the fit quality $R^2 = 1 - \text{RSS}/\text{TSS}$, where RSS, TSS - residual and total sum of squares, for exponential and power-law fits like in (a) and (b) but for more disorder strengths. Power-law fits the data better in the critical regime $W \lesssim 4.5$. On the MBL side, $W \gtrsim 4.5$, the exponential decay has a lower relative error. (d): Fitted values of γ . (e): Fitted values of β .

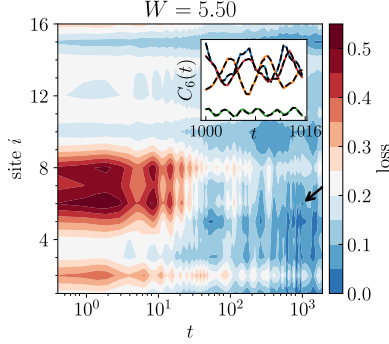


FIG. S2. Neural network loss for an anomalous $W = 5.5$ disorder realization, detected at the cutoff level $l_{\text{cut}} = 0.15$, system size $L = 16$. We notice the disappearance of the large loss (non-ergodic) region due to thermalization by its environment.

known that if at least a few spins are strongly entangled, their observables will converge towards thermal values and thermal evolutions, thus large entanglement clusters become equivalent to ergodic clusters. It is, however, not meaningful to compare very small ergodic clusters ($n = 1$ or 2) with similarly small entanglement clusters. Additionally, in our approach, a time series of spin correlations is needed to clusterize the system, whereas in the QMI case the clusters can be defined at time t based on the entanglement properties of the instantaneous wave function $|\psi(t)\rangle$. Furthermore, entanglement clusters can have holes (see the discussion in the preceding paragraph), and NN clusters cannot, but, as argued in ref. [2], QMI clusters with holes amount to less than 5-10% of all clusters in a typical scenario. For a summary of this paragraph see TABLE I.

So far, the QMI clustering procedure has been applied

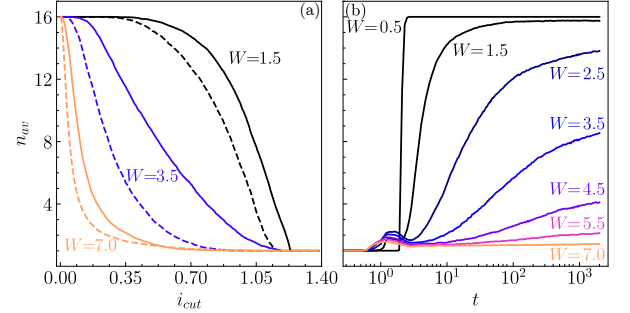


FIG. S3. a) Average size of entanglement clusters n_{av} at time instants $t = 20$ (dashed lines) and $t = 1900$ (solid lines) for different QMI cutoffs i_{cut} and random disorder strengths W , system size $L = 16$. b) Average cluster size growth in time for cutoff $i_{\text{cut}} = 0.5$.

Entanglement clusters	Ergodic clusters
$n \geq 1$	$n \geq 0$
$\sum_i n_i = L$	$\sum_i n_i \leq L$
defined for $ \psi(t)\rangle$	defined for $\{\psi(t), \dots, \psi(t + n\Delta t)\}$

TABLE I. Comparison of some properties of entanglement clusters and ergodic clusters.

only to the middle-spectrum eigenstates of the disordered Heisenberg Hamiltonian [2]. Here, we extend those studies to characterize the cluster growth in time and compare the results with the neural network clustering. The system size is limited to $L = 16$ due computationally expensive entanglement entropy calculations. As in the Letter, we denote the averaging of the cluster size over a single state as $\langle \cdot \rangle_\psi$ and over disorder realizations as $\langle \cdot \rangle_W$. In FIG. S3a) we present the average entanglement cluster size $n_{\text{av}} = \langle \langle n \rangle_\psi \rangle_W$ for different cutoff values i_{cut} at

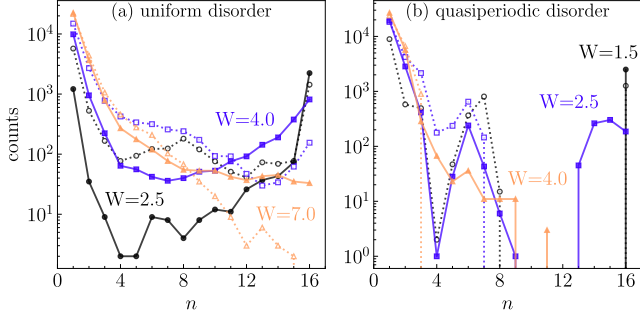


FIG. S4. Histograms of the entanglement cluster sizes n detected at the QMI cutoff level $i_{\text{cut}} = 0.4$ by the QMI clustering algorithm close to the start of the evolution $t = 20$ (empty markers) and after a long time $t = 1900$ (filled markers) for 2500 disorder realizations with system size $L = 16$.

time $t = 20$ and $t = 1900$. We confirm that, similarly to the NN case, the transitions with cutoff are qualitatively different in the ergodic, critical and MBL regimes and that the clusters grow regardless of the cutoff. Next, we select an intermediate value of the cutoff $i_{\text{cut}} = 0.5$, in FIG. S3b) we analyze the average cluster size evolution in time and, like for neural networks, observe a logarithmic growth $n_{\text{av}}(t) = \alpha \log t + C$ in the MBL phase. We find no relation between the growth rates α of entanglement entropy and entanglement clusters because the latter strongly depend on the chosen threshold l_{cut} .

In FIG. S4 we plot the histograms of entanglement cluster sizes in the ergodic, critical and MBL regime for uniform and quasiperiodic disorder. In the ergodic case, we notice the disappearance of initially large number of small clusters and the appearance of the power-law distribution $p(n) \sim n^{-X}$ with a peak for large clusters dominating the distribution after a long time evolution. In the critical regime, after the evolution, we detect a mixture of all cluster sizes. This is in agreement with the main result from the study of eigenstates [2], namely that at the critical point small entanglement clusters are entangled together to form larger clusters – choosing a constant cutoff we are able to observe clusters of nearly all sizes. In contrast, in the localized phase, the clusters become small and independent of each other, with the cluster distribution described by $p(n) \sim \exp(-n\gamma)$ without the thermal peak. All qualitative conclusions about the properties of the cluster distributions are thus in agreement with the NN results presented in the Letter. A one order of magnitude difference in the number of the smallest clusters is connected with the different meanings of ergodic and entanglement clusters, as described in the previous paragraph – between the two methods one can only compare the properties of the largest clusters.

Griffiths regions We can verify whether the anomalously large clusters, at large disorders, detected by the QMI clustering algorithm, have the same properties as the Griffiths events [3–7] found by the neural network.

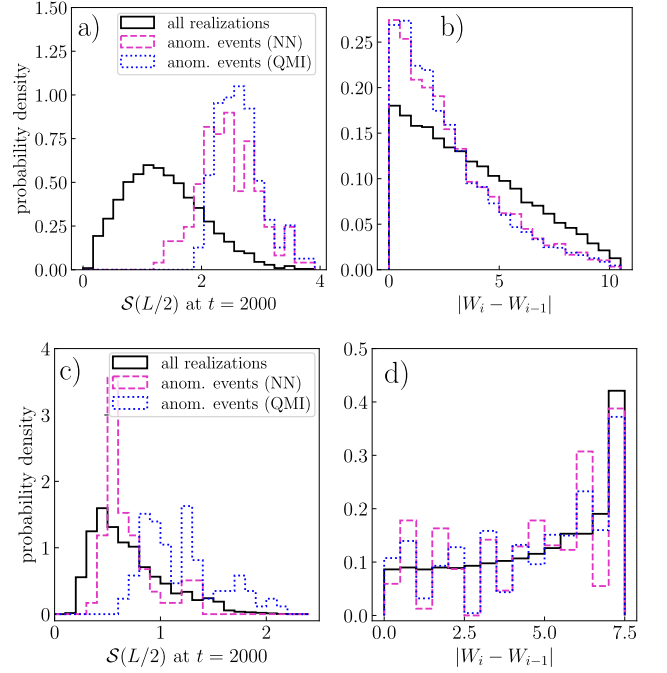


FIG. S5. Anomalous realizations for uniform (first row) and quasiperiodic (second row) disorder for system size $L = 16$. At the same values of cutoffs as in the uniform disorder case, no anomalous realizations are detected at all. If the cutoffs are tuned to $i_{\text{cut}} = 0.33$, $l_{\text{cut}} = 0.09$ (presented here) to obtain a similar number of anomalous realizations as in FIG. 5 from the Letter, these realizations are not as clearly distinguishable from the background as in the uniform disorder case (maybe except for a fraction of the high-entropy states detected by QMI).

In FIG. S5 we plot histograms similar to FIG. 5 from the Letter, except the system size is now $L = 16$ (this does not change NN results qualitatively). FIGs. S5(a),(b) clearly show, that for the uniform disorder $W = 5.5$, the anomalous events detected by both methods at the cutoff levels $l_{\text{cut}} = 0.15$, $i_{\text{cut}} = 0.50$ have nearly the same distributions of entanglement entropies and resonant disorders. Applying both methods to the quasiperiodic disorder with amplitude $W = 4.0$ (FIG. S5), we do not detect any anomalous realizations at the same cutoff levels. If we tune the cutoffs so that finding a large cluster is more probable ($l_{\text{cut}} = 0.25$, $i_{\text{cut}} = 0.20$), a comparable number of anomalous disorder realizations is detected, but their features are not distinct from the background, except slightly larger entropies found by the entanglement clustering method. This difference is not surprising because entropy properties are the object based on which the entanglement clusters are built.

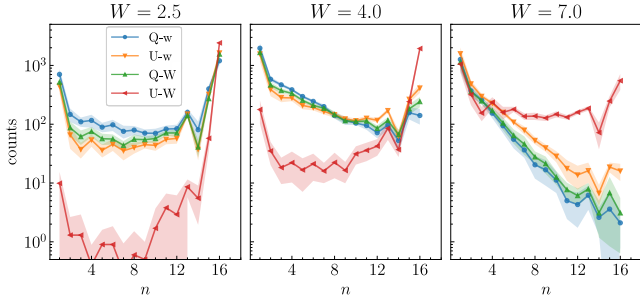


FIG. S6. Comparison of ergodic bubble size distributions obtained using neural networks trained on different input data (see TABLE II) at loss cutoff $l_{\text{cut}} = 0.15$ in the $L = 16$ system.

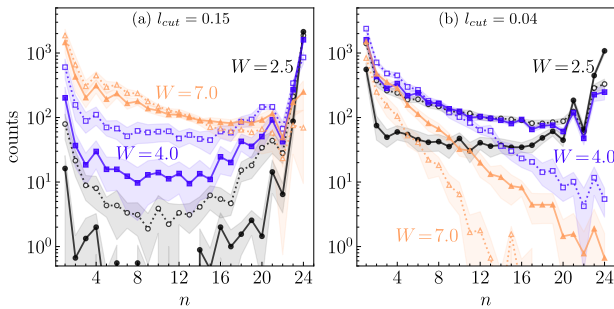


FIG. S7. Bubble size histograms for the model trained on $W = 0.1 - 1.0$. For the same value of the loss cutoff $l_{\text{cut}} = 0.15$ as for the model trained on $W = 0.1 - 0.5$ presented in the Letter we obtain a histogram with more large bubbles, even for $W = 7.0$. (b) By tuning the cutoff down to 0.04 we can recover the same qualitative histogram as in FIG. 4a in the Letter. Here, $L = 24$.

NEURAL NETWORK TRAINING

In the Letter we present results obtained using a network trained on evolutions corresponding to uniform disorder in the range $W = 0.1 - 0.5$ and time $t = 100 - 2000$, which we assume to be all ergodic. To perform a consistency check and verify how the input data affects the model performance, we also train models on larger values of disorder up to $W = 1.0$, as well as on quasiperiodic disorder of the same amplitudes.

Specifically, the training dataset consists of 100 realizations of disorder from the set called “w”: $W \in \{0.1, 0.2, \dots, 0.5\}$ or “W”: $W \in \{0.1, 0.2, \dots, 1.0\}$. We train separately on disorders coming from the uniform (models called “U-w”, “U-W”) and quasiperiodic (“Q-w”, “Q-W”) distributions, as shown in TABLE II. We stop the training after a certain value $1.7 \cdot 10^{-3}$ of the validation loss is achieved. Validation data corresponds to 40 disorder realizations that the network does not use

to tune the weights but only to terminate the training when the desired level of loss is achieved.

model name	dist. of disorder	values of disorder
“U-w”	uniform	$\{0.1, 0.2, \dots, 0.5\}$
“U-W”	uniform	$\{0.1, 0.2, \dots, 1.0\}$
“Q-w”	quasiperiodic	$\{0.1, 0.2, \dots, 0.5\}$
“Q-W”	quasiperiodic	$\{0.1, 0.2, \dots, 1.0\}$

TABLE II. Distributions and values of disorder used for training four classes of neural network models. At each value of disorder, time series of its 100 random realizations were used as input. From each model class, 10 neural networks were trained independently on the same data.

In FIG. S6 we present cluster size histograms similar to FIG. 4 from the Letter but for all considered models and only at the end of the evolution $t = 1900$ and $L = 16$. We observe that all networks except “U-W” give consistent results. We suspect that this is due to the fact that in the “U-W” training scenario the timescale of the full thermalization of the system is longer than $t = 100$.

In FIG. S7 we further evaluate the “U-W” networks for $L = 24$. They show a similar behavior as for $L = 16$ - under the same loss cutoff, they detect large bubbles even at $W = 7.0$ (FIG. S7(a)). However, we find out that upon decreasing the threshold, the histograms from FIG. 4(a) in the Letter can be qualitatively recovered. This means that if one allows for a change in the threshold, the training procedure is robust with respect to the input training data.

It is important to note that the neural networks are trained on the $L = 16$ data only. They can later be applied to any system size L which is a strong advantage of the method. It is possible because the two-site correlations the network uses as an input are local observables characterizing a single site and not the whole system of length L .

-
- [S1] P. T. Dumitrescu, A. Goremykina, S. A. Parameswaran, M. Serbyn, and R. Vasseur, *Phys. Rev. B* **99**, 094205 (2019).
 - [S2] L. Herviou, S. Bera, and J. H. Bardarson, *Phys. Rev. B* **99**, 134205 (2019).
 - [S3] T. Vojta, *J. Low Temp. Phys.* **161**, 299 (2010).
 - [S4] S. Gopalakrishnan, M. Müller, V. Khemani, M. Knap, E. Demler, and D. A. Huse, *Phys. Rev. B* **92**, 104202 (2015).
 - [S5] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, *Phys. Rev. Lett.* **114**, 160401 (2015).
 - [S6] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, *Annalen der Physik* **529**, 1600326 (2017).
 - [S7] N. Pancotti, M. Knap, D. A. Huse, J. I. Cirac, and M. C. Bañuls, *Phys. Rev. B* **97**, 094206 (2018).

Catching thermal avalanches in the disordered XXZ model

Tomasz Szoldra^{1,2,*}, Piotr Sierant³, Maciej Lewenstein^{3,4} and Jakub Zakrzewski^{2,5}¹*Doctoral School of Exact and Natural Sciences, Jagiellonian University, Łojasiewicza 11, PL-30-348 Kraków, Poland*²*Instytut Fizyki Teoretycznej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*³*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*⁴*ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain*⁵*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, PL-30-348 Kraków, Poland*

(Received 9 February 2024; accepted 14 March 2024; published 8 April 2024)

We study the XXZ model with a random magnetic field in contact with a weakly disordered spin chain, acting as a finite thermal bath. We revise Fermi's golden rule description of the interaction between the thermal bath and the XXZ spin chain, contrasting it with a nonperturbative quantum avalanche scenario for the thermalization of the system. We employ two-point correlation functions to define the extent ξ_d of the thermalized region next to the bath. Unbounded growth of ξ_d proportional to the logarithm of time or faster is a signature of an avalanche. Such a behavior signifies the thermalization of the system, as we confirm numerically for a generic initial state in the ergodic and critical regimes of the XXZ spin chain. In the many-body localized regime, a clear termination of avalanches is observed for specifically prepared initial states and, surprisingly, is not visible for generic initial product states. Additionally, we extract the localization length of the local integrals of motion and show that a bath made out of a weakly disordered XXZ chain has a similar effect on the system as a bath modeled by a Hamiltonian from a Gaussian orthogonal ensemble of random matrices. We also comment on the result of the earlier study [Phys. Rev. B **108**, L020201 (2023)], arguing that the observed thermalization is due to external driving of the system and does not occur in the autonomous model. Our work reveals experimentally accessible signatures of quantum avalanches and identifies conditions under which termination of the avalanches may be observed.

DOI: [10.1103/PhysRevB.109.134202](https://doi.org/10.1103/PhysRevB.109.134202)

I. INTRODUCTION

The ergodic hypothesis of Boltzmann [1] stipulates that all microstates of a many-body system are equiprobable over a sufficiently long period of time. Such an equilibrium state can be described within the framework of statistical mechanics with several macroscopic variables. The eigenstate thermalization hypothesis (ETH) [2–5] describes the equilibrium of isolated quantum many-body systems, warranting that, after the thermalization process is complete, local observables, including their higher order correlation functions, are described by appropriate ensembles of statistical mechanics [6–10].

While the applicability of the ETH has been confirmed for a wide range of quantum many-body systems, see, e.g., Refs. [11–16], a phenomenon of many-body localization (MBL) [17–21] has been proposed as a general mechanism that prevents many-body systems from reaching the thermal equilibrium in the presence of a strong disorder. Numerical results for spin-1/2 XXZ chains [22,23], bosonic models [24–27], or systems of spinful fermions [28–31] indicate a rapid slow-down of dynamics with increasing disorder strength [32–35], as confirmed experimentally in setups of ultracold atoms [36–41], ions [42], or superconducting qubits [43–46]. Early studies interpreted this behavior as a stable

MBL phase [47–49]. In such a dynamical phase of matter, the information about details of the initial state of the system is kept indefinitely in the values of local integrals of motion (LIOMs) [50–56]. At the same time, the transport is suppressed [57], and the entanglement spreads slowly [58–60]. However, it has recently become apparent that the numerical evidence about the dramatic slow down of the dynamics with increasing disorder strength [61–64], and the related properties of highly excited eigenstates [65–72], do not necessarily imply an existence of a stable MBL phase and could be rather interpreted in terms of an MBL regime [73–75], in which the thermalization occurs only at enormous time scales and only for sufficiently large system size [76,77].

The question of the existence of a stable MBL phase is actively pursued, and the answer may be model-dependent as suggested by recent numerical studies [76,78,79]. One approach to tackle this question, fundamental for our understanding of nonequilibrium many-body physics, is to investigate a mechanism that underlies the slow thermalization of a strongly disordered many-body system. Spatially uncorrelated disorder, assumed in the standard models of MBL, necessarily leads to the emergence of Griffiths regions, where the disorder is locally weaker. The Griffiths regions play an essential role in the physics of quantum phase transitions [80] and affect the slow dynamics of disordered many-body systems [81–84]. The significance of Griffiths regions was further highlighted when an *avalanche* mechanism

*tomasz.szoldra@doctoral.uj.edu.pl

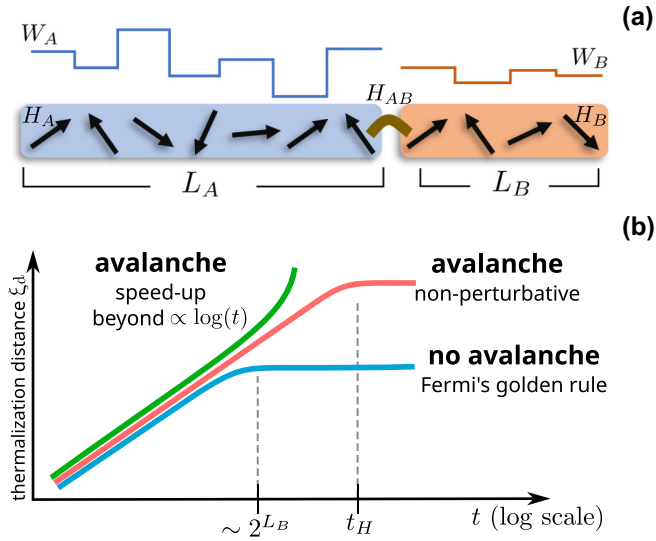


FIG. 1. (a) Scheme of our setup: two spin-1/2 XXZ chains of sizes L_A , L_B , governed by Hamiltonians H_A , H_B , with different amplitudes of random disorder, connected at the interface by H_{AB} . (b) Sketch of the signatures of quantum avalanches obtained in Sec. II for the model from (a). First-order perturbative scheme predicts the growth of thermalized region as $\propto \ln(t)$ up to $t \sim 2^{L_B}$ when the system starts to feel the discreteness of the bath spectrum (blue curve). In the avalanche scenario, the bath continually increases its size and the logarithmic growth continues until the Heisenberg time t_H (red curve) or speeds up (green curve).

for MBL transition was proposed [85]. The avalanche mechanism assumes that the Griffiths regions of anomalously small disorder become *ergodic bubbles*, and describes, under certain simplifying assumptions, how the ergodic bubbles grow by thermalizing their surroundings. The simplified model of avalanche spreading is itself an interesting model of ergodicity breaking [77,86–89], and constitutes the basis of various phenomenological approaches to MBL [90–93]. Nevertheless, understanding the relevance of the avalanche mechanism for the dynamics of disordered many-body systems [94–99] and how to model the spread of avalanches [75,100–102] remain the major challenges in the field of MBL.

In this work, to study the fate of such ergodic bubbles in a controlled setting, we follow the approach of [103–105] and divide the lattice into subsystems A and B with different disorder strengths W_A and $W_B < W_A$. The subsystem A plays the role of an MBL system, while the subsystem B is the ergodic bubble, see Fig. 1(a). By monitoring the growth of correlations between the bath B and the MBL system A , we investigate how the bath influences the dynamics of the subsystem A , and whether it induces its thermalization. Quantitative analysis of the correlations between the subsystems allows us to pinpoint the onset of the propagation of avalanches that destabilize the MBL system. We distinguish three different types of behavior of the system, presented in Fig. 1(b). If the evolution starts from a generic initial state, the system thermalizes in the ergodic and critical regime and the quantum avalanches spread faster than $\propto \ln(t)$. For the same initial state, but in the MBL regime, we still observe a nontrivial increase of the thermalization distance $\xi_d(t) \propto \ln(t)$. The latter behavior is an artifact of the initial state. Considering a product state of

a highly excited eigenstate of subsystem A and arbitrary state in subsystem B as the initial state, we demonstrate that in the MBL regime the growth of $\xi_d(t)$ terminates at a time set by the bath size signifying the halt of spreading of avalanches.

The rest of this paper is structured as follows. In Sec. II, we outline a general picture of the possible scenarios for interaction of an ergodic bath with an MBL system. Subsequently, in Sec. III, we present the disordered spin-1/2 XXZ model and describe the correlation functions that allow us to investigate the propagation of avalanches, followed by numerically extracted signatures of avalanche propagation in our setup. In Sec. IV, we discuss a situation when the Hamiltonian of the subsystem B is replaced by a GOE random matrix, study the sensitivity to the disorder strength in the bath, and discuss the influence of the subsystem A on the dynamics of the bath B . Finally, in Sec. V, we show that periodic turning on and off of the coupling between subsystems A and B leads to much stronger tendency towards delocalization in comparison to the case of the time independent coupling.

II. ERGODIC BUBBLES IN MBL SYSTEM

Typically, MBL is studied in systems with uncorrelated on-site disorder, with values drawn from a uniform distribution on an interval $[-W, W]$. The probability that a stretch of ℓ consecutive sites, denoted as subsystem B , experiences a weaker disorder $W_B < W$ is exponentially small in ℓ and equal to $p_\ell = (W_B/W)^\ell$. However, the probability of finding an ergodic bubble, i.e., a thermal subsystem with weak disorder strength W_B of any finite size ℓ , tends to unity with increasing system size L . Such a subsystem B plays a role of a *thermal bath* for the neighboring spins. The fate of this bath, and its influence on the dynamics of the surrounding spins are of vital importance for the stability of MBL in the system.

We consider a 1D disordered spin chain described by spin-1/2 operators $\vec{S}_i = (\hat{X}_i, \hat{Y}_i, \hat{Z}_i)$ on sites $i = 1, \dots, L$, and model the situation of interest by the following Hamiltonian:

$$H = H_A + H_B + H_{AB}, \quad (1)$$

where H_A describes a generic MBL subsystem A , H_B is the Hamiltonian of the ergodic bubble, identified here as the subsystem B and H_{AB} is the coupling between the subsystems A and B .

The density of states of the thermal subsystem is, at the leading order, given by $\rho_B \sim 2^{L_B}$. Since the subsystem B is ergodic, the matrix elements of a local operator O_i localized around site i in B , in the eigenbasis of H_B , $H_B|\psi_n\rangle = E_n|\psi_n\rangle$, are given by the ETH ansatz [2,3,5,106]

$$\langle \psi_n | O_i | \psi_m \rangle = O(\bar{E})\delta_{nm} + \rho_B^{-\frac{1}{2}} f_O(\bar{E}, \omega) R_{mn}, \quad (2)$$

where $\bar{E} = (E_m + E_n)/2$, $\omega = E_n - E_m$ is the energy difference, R_{mn} is a random variable of zero mean and unit variance, and the factor $\rho_B^{-\frac{1}{2}}$ encodes the thermodynamic entropy for the eigenstates in the middle of the spectrum of H_B [107]. Importantly, $O(\bar{E})$ and $f_O(\bar{E}, \omega)$ are both smooth functions of their arguments.

The MBL Hamiltonian H_A can be expressed in terms of the set of LIOM operators, $\{\tau_i^z\}$, $i = 1, \dots, L_A$ [49]. Each of the LIOMs can be expanded in the microscopic spin operators

$(\hat{X}_j, \hat{Y}_j, \hat{Z}_j)$ as

$$\hat{\tau}_j^z = \zeta \hat{Z}_j + \sum_{n=1}^{\infty} a_j^{(n)} \hat{A}_j^{(n)}, \quad (3)$$

where ζ is the overlap of $\hat{\tau}_j^z$ with Z_j , and $\hat{A}_j^{(n)}$ is an operator acting nontrivially only at sites $j-n, \dots, j, \dots, j+n$. The coefficients of longer-range operators decay as $a_j^{(n)} \sim e^{-n/\xi}$ [51]. This decay defines the characteristic lengthscale ξ and is a crucial property that determines the phenomenology of the MBL systems by ensuring the quasilocality of LIOMs.

The bath-MBL coupling H_{AB} involves microscopic spin operators supported at the interface between the subsystems A and B . Due to the exponential decay of LIOMs (3), this coupling can be rewritten as $H_{AB} = \sum_{j=1}^{L_A} H_{AB}^{(j)}$ with

$$H_{AB}^{(j)} = (V e^{-r/\xi} O_i \tau_j^- + \text{H.c.}), \quad (4)$$

where O_i acts on the degrees of freedom of the ergodic subsystem B , and $r = |i - j|$ is the distance between the LIOM and the bath. To write (4), we have neglected all subleading terms involving higher order products of LIOM operators and we have kept only the local operator O_i . These assumptions are sufficient for the qualitative analysis of the leading terms dictating the influence of the bath on the MBL system.

Let us assume that at time $t = 0$, we prepare the system in the state

$$|\phi_{\text{in}}\rangle = |\tau_1 \dots \tau_j \dots \tau_{L_A}\rangle \otimes |\psi_n\rangle \quad (5)$$

with a fixed configuration of LIOMs, $\tau_j = \uparrow, \downarrow$, and with the ergodic bath initialized in the eigenstate $|\psi_n\rangle$ with energy E_n .¹

The initial state of the MBL subsystem A may slowly thermalize due to the coupling H_{AB} to the ergodic bath B . In the following, we analyze two extreme scenarios for time evolution of the system. First, we assume that the ergodic bath remains unaltered during the time evolution, and employ the first order of perturbation theory to estimate the decay rate of the LIOM configuration [108]. We contrast this perturbative scenario with the nonperturbative avalanche mechanism [85] which assumes that the thermalized spins become members of the ergodic bubble, increasing the potential of the bath to thermalize the subsequent degrees of freedom.

1. Perturbative coupling to a static bath

Let us estimate the time scale for thermalization of the LIOM $\hat{\tau}_j$ at site j . To that end, from the bath-MBL interaction we keep only the coupling term $H_{AB}^{(j)}$, and note that, in the first order perturbation theory, the initial state $|\phi_{\text{in}}\rangle$ (5)

is coupled to

$$|\phi_{\text{fin}}\rangle = |\tau_1 \dots \bar{\tau}_j \dots \tau_{L_A}\rangle \otimes |\psi_m\rangle, \quad (6)$$

where $\bar{\tau}_j$ is the opposite spin to τ_j and $|\psi_m\rangle$ is an eigenstate of the bath with $m \neq n$. The probability that the LIOM $\hat{\tau}_j$ remains in its initial eigenstate τ_j falls down exponentially in time with a decay rate given by the Fermi golden rule [108],

$$\Gamma_j \sim |\mathcal{T}_j|^2 \rho_B, \quad (7)$$

where the matrix element $\mathcal{T}_j = \langle \phi_{\text{fin}} | H_{AB}^{(j)} | \phi_{\text{in}} \rangle$ can be rewritten with the help of the ETH ansatz (2), giving rise to

$$\Gamma_j \sim V^2 e^{-2r/\xi}, \quad (8)$$

where we have neglected the energy dependence of the spectral function f_O and of the density of states. We note that the rate Γ_j is independent of the bath size L_B since the density of states ρ_B in (7) is canceled by the square of $\rho_B^{-1/2}$ from (2). Moreover, (8) is valid only on timescales shorter than ρ_B^{-1} . At larger times, the discreteness of the spectrum of H_B is resolved, and transitions from the state $|\phi_{\text{in}}\rangle$ occur only in the rare cases of matching between the energies of $|\phi_{\text{in}}\rangle$, $|\phi_{\text{fin}}\rangle$. Therefore LIOM at site j is thermalized at timescale $t_j \sim 1/\Gamma_j \propto e^{2r/\xi}$ increasing exponentially with the distance r from the bath, provided that $t_j < \rho_B^{-1}$. In other words, the distance ξ_d of penetration of the ergodic bath B into the MBL subsystem A will grow in time as $\xi_d \propto \ln t$, and then will saturate at a time proportional to 2^{L_B} . Therefore the MBL in the subsystem A remains stable in this scenario. Indeed, the difference of energies of the states $|\phi_{\text{in}}\rangle$, $|\phi_{\text{fin}}\rangle$ is, at minimum (in a typical case), proportional to ρ_B^{-1} . Hence, the first order perturbation theory correction to the state $|\phi_{\text{in}}\rangle$ is of the order of

$$h_r^{\text{stat}} \sim \frac{\mathcal{T}_{L_A-r}}{\rho_B^{-1}} \sim e^{(-1/\xi)r}, \quad (9)$$

i.e., it is vanishing exponentially with the increasing distance r from the bath.

2. Avalanche mechanism

The above description assumes that the events of thermalization of subsequent LIOMs are taking place independently, and that the thermalized LIOMs do not influence the dynamics of the system at the later stages of the process. In the following, we analyze the consequences of an opposite assumption.

Let us now consider thermalization of the LIOM $\hat{\tau}_{L_A}$ closest to the bath. Equation (8) implies that the probability remaining in the eigenstate τ_{L_A} of the LIOM is significantly below unity after time $t_{L_A} \propto e^{2/\xi}$. According to the avalanche mechanism [85], once the state of the LIOM $\hat{\tau}_{L_A}$ has relaxed, the spin at site $j = L_A$ becomes a member of the ergodic bath, and, consequently the density of states of the new bath B' is increased by a factor of 2, $\rho_{B'} = 2\rho_B$. After r iterations of such a process, the density of the states of the enlarged bath is $\rho_{B'} = 2^r \rho_B$. We consider a state $|\tau_1 \dots \tau_{L_A-r}\rangle \otimes |\phi_n\rangle$, where $|\phi_n\rangle$ is an eigenstate of the enlarged bath comprised of $L_B + r$ spins, and calculate the first order perturbation theory correction to this state due to the coupling H_{AB} , which leads us to

$$h_r^{\text{aval}} \sim \frac{\mathcal{T}_{L_A-r}}{(2^r \rho_B)^{-1}} \sim e^{((\ln 2)/2 - 1/\xi)r}. \quad (10)$$

¹The eigenstate of the bath $|\psi_n\rangle$ is chosen as the initial state for simplicity. Due to the ergodicity of the bath, product states in real space become superpositions of eigenstates $|\psi_n\rangle$ with structureless coefficients, leading to qualitatively the same behavior. In contrast, the assumption of an initial state being an eigenstate in the subsystem A is of crucial importance, see Sec. III C.

In contrast to the perturbative expression for the static bath B , Eq. (9), the hybridization ratio h_r^{aval} is exponentially enhanced by a factor $e^{r(\ln 2)/2}$, associated with the growth of $\rho_{B'}$ upon incorporation of the neighboring spins. As a consequence, thermalization of the MBL subsystem may be a self-sustaining process if the growth of the bath density of states overcomes the decrease of the interaction strength with the distance r , i.e., when $(\ln 2)/2 - 1/\xi > 0$. When this condition is met, an avalanche propagates throughout the system, and the system thermalizes. According to the Fermi golden rule rate (8), the avalanche propagation may still be exponentially slow, i.e., the distance ξ_d of penetration of the MBL system may grow in time $\xi_d \propto \ln t$. However, the density of states of the enlarged bath increases after each act of thermalizing and incorporation of the neighboring spin, hence, (8) is valid up to the time scale ρ_B^{-1} , which continually increases when the avalanche propagates. The discussed scenario predicts an unbounded logarithmic growth of the thermalization length $\xi_d \propto \ln t$. Higher order perturbative corrections, as well as non-perturbative mechanisms associated with the growth of the thermal inclusion, beyond the simplified processes discussed above, may speed up the spreading of avalanche when $(\ln 2)/2 - 1/\xi > 0$. Thus, we *define* that avalanches spread in the investigated disordered many-body system whenever the presence of a thermal inclusion induces logarithmic, or any faster, growth of ξ_d in time. This definition does not agree with that of Ref. [103] in which $\xi_d \propto \log t$ is classified as a lack of an avalanche. Disagreement comes from the fact that the expression for the decay rate Γ_j , Eq. (8) is valid only at times $t < \rho_B^{-1}$, which, in the absence of the avalanche spreading, is bounded by the constant 2^{L_B} .

III. DISORDERED XXZ SPIN CHAIN WITH ENGINEERED ERGODIC BUBBLE

The simple scenarios for the time evolution of an MBL system in contact with an ergodic bath described in Sec. II rely on a number of assumptions which are not straightforwardly testable. For that reason, the direct relevance of the avalanche mechanism for the dynamics of disordered systems remains not clear. The first assumption regards the form of LIOM operators, see Eq. (3). While there are numerical approaches that allow to construct LIOMs for strongly disordered systems [55,109–119], probing of the properties of LIOMs at large time scales, and especially in the vicinity of the crossover to ETH phase remains a challenge, and one has to resort to perturbative schemes which are not rigorously controlled [51]. Moreover, the avalanche mechanism assumes that each thermalized spin joins the ergodic bath which still can be described by the ETH ansatz (2). A strict validity of this assumption is unclear [94]. Finally, ergodic bubbles in strongly disordered systems form due to rare disorder configurations, and hence, their dynamics can be probed only indirectly, with specific tools [120].

The above reasons motivate us to analyze numerically the process of avalanche propagation in a system in which we assure the presence of the ergodic bubble by dividing the lattice into two subsystems as shown in Fig. 1(a).

A. The system

We consider the XXZ-spin model with random magnetic field, widely studied in the context of MBL [23,33,61,62,67,73,82,95,121–137], coupled to a bath modelled also by XXZ-spin system. The Hamiltonian reads

$$\begin{aligned} H &= H_A + H_B + H_{AB}, \\ H_A &= J \sum_{i=1}^{L_A-1} X_i X_{i+1} + Y_i Y_{i+1} + \Delta_A Z_i Z_{i+1} \\ &\quad + \sum_{i=1}^{L_A} W_i Z_i, \\ H_B &= J \sum_{i=L_A+1}^{L-1} X_i X_{i+1} + Y_i Y_{i+1} + \Delta_B Z_i Z_{i+1} \\ &\quad + \sum_{i=L_A+1}^L W_i Z_i, \\ H_{AB} &= J(X_i X_{i+1} + Y_i Y_{i+1} + \Delta_{AB} Z_i Z_{i+1}), \end{aligned} \quad (11)$$

where $\vec{S}_i = (X_i, Y_i, Z_i)$ are spin-1/2 operators at lattice site i , $J = 1$ fixes the energy unit (and J^{-1} is the characteristic time scale which we call a tunneling time having in mind the equivalent spinless fermions expression of the model which arises upon the Jordan-Wigner transformation). Open boundary conditions are assumed. Unless stated otherwise, we set the interaction strengths $\Delta_A = \Delta_B = \Delta_{AB} = 1$. The chain consists of two subsystems of sizes L_A and L_B , with the total system size $L = L_A + L_B$. The on-site magnetic fields, W_i , are independent random variables drawn from the uniform distribution in the interval that differs between the two subsystems:

$$W_i \in \begin{cases} [-W_A, W_A] & \text{if } 1 \leq i \leq L_A, \\ [-W_B, W_B] & \text{otherwise,} \end{cases} \quad (12)$$

see Fig. 1(a). The subsystem B serves as a thermal bath that seeds the avalanche and initializes the process of thermalization of the subsystem A , and, consequently, of the whole chain. For that reason, we fix the disorder in the subsystem B to be weak. If not stated otherwise, we put $W_B = 0.5$. At this disorder strength the Heisenberg spin chain thermalizes quickly, since $W_B = 0.5$ is well below the various estimates for critical disorder strength for MBL transition [23,67,137–139]. We work in the $Z = 0$ total magnetization sector. Time evolution is computed exactly using the Chebyshev polynomial expansion of the time evolution operator [140,141].

Observables describing the system (except for imbalance from Fig. 7) are calculated as a median over 1000 random disorder realizations unless specified explicitly. Using a robust estimator, such as median, instead of a usual arithmetic average, ensures convergence of the measured quantities in the number of disorder realizations. Median captures a typical behavior of the system and is not sensitive to few accidental disorder realizations, particularly affecting the mean value at the largest disorder strengths, deep in the MBL regime. For a comparison between results obtained using the mean and the median, see Appendix A.

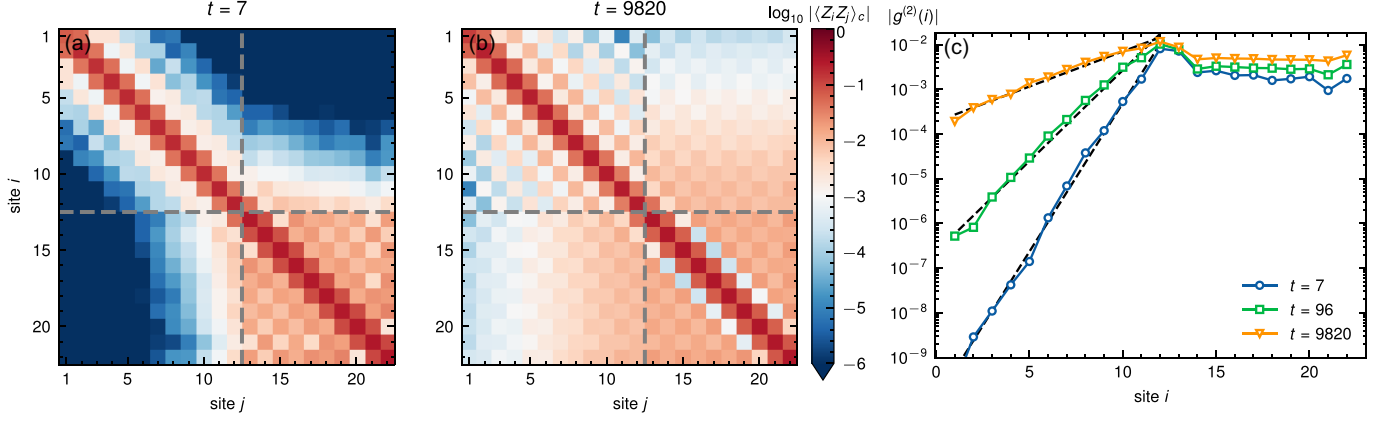


FIG. 2. Model parameters are $L_A = 12$, $L_B = 10$, $W_A = 4$, and $W_B = 0.5$. (a) Initially, correlations within subsystem A (first 12 sites) are short-ranged, while correlations within the bath (sites 13–22) spread over the whole bath. (b) After a sufficiently long time, correlations between the bath and the MBL subsystem are developed and the system thermalizes. Checkerboard pattern appears due to a specific initial Néel state. (c) Average (over bath) of two-body correlations for each lattice site. Correlation decay length ξ_d is obtained by fitting an exponential decay with distance, Eq. (14) (dashed lines).

B. The avalanche spreading

In order to describe the process of the avalanche spreading and of the ensuing thermalization of the subsystem A quantitatively, we fix the initial state as the Néel state and calculate the two-point connected correlation function $\langle Z_i Z_j \rangle_c = \langle Z_i Z_j \rangle - \langle Z_i \rangle \langle Z_j \rangle$ for sites i and j .

Typical behavior of the $\langle Z_i Z_j \rangle_c$ correlation function for $W_A = 4$ and $W_B = 0.5$ is presented in Fig. 2. Initially, at time scale of one tunneling time only the short-range correlations are present throughout the system among the neighboring sites $|i - j| \approx 1$ (plot not shown). After several tunneling times, $t \approx 7$, the sites within the bath B develop mutual correlations, while the correlations in the subsystem A remain short-ranged, see Fig. 2(a). Subsystems A and B begin to develop nontrivial correlations due to their connection via a bond between the sites L_A , $L_A + 1$. During the course of time evolution of the system, the correlations gradually increase, and, at a large time $t = 9820$, we observe that the whole subsystem A develops significant correlations with the bath B . The correlation between a given site i_A from A is, to a good approximation, the same with all of the sites i_B from the bath B . This occurs because the spins in the bath are strongly correlated with each other and as such can no longer be treated individually. In contrast, the strength of the correlation between the spins from subsystem A with spins of the bath B decreases with the distance to the edge of the bath.

Our aim is to find out if the correlations between the bath B and the subsystem A are a nonperturbative effect of the propagation of a quantum avalanche, or are only a perturbative result of the bath proximity, as described in Sec. II. To quantify how the correlations with the bath penetrate through the subsystem A, following Ref. [103], we define a bath-averaged correlation function for each site,

$$g^{(2)}(i) = \overline{\langle Z_i Z_j \rangle_c} |_{j \in L_B}. \quad (13)$$

An example of the $g^{(2)}(i)$ function is shown in Fig. 2(c). Clearly, the correlations with the bath in the subsystem A fall off exponentially with the distance from the interface between A and B. This leads us to the definition of the correlation decay

length ξ_d which is obtained by fitting

$$|g^{(2)}(i)| = c \exp\left(-\frac{L_A - i}{\xi_d}\right) \quad (14)$$

at a given instant of time. We note that the employed definition allows us to accurately determine the decay length ξ_d without introducing unwanted artifacts of the approach employed in Ref. [103]. In particular, this method is less sensitive to finite-size effects and the discreteness of the lattice, see Appendix B for more details. In the following, we will probe the thermalization of the subsystem A mainly focusing on the time dependence of the decay length ξ_d .

C. Signatures of avalanches due to the bath

Exact diagonalization studies of the disordered XXZ model [67] suggest the presence of three distinct regimes of the system characteristics, depending on the disorder strength W . For the lowest disorder values $W < W^T(L)$ the system follows ETH, its energy level statistics follow GOE predictions and the system thermalizes in a generic way. In contrast, for disorder larger than a certain value $W > W^*(L)$, the spectrum shows signs of emergent integrability (MBL). If an MBL transition occurs in the system, then $\lim_{L \rightarrow \infty} W^T(L) = \lim_{L \rightarrow \infty} W^*(L) = W_c$, where W_c is the critical disorder strength. In that case, for finite systems, there is a finite transient critical regime $W^T(L) < W < W^*(L)$ with multifractal entanglement structure in the eigenstates [95]. This might suggest that, in these three regimes, the system will react differently to the attached finite thermal bath. Our expectation is that in the ergodic and critical regimes, any bath will lead to a growth of $\xi_d(t)$ faster than $\propto \ln(t)$, consistent with the avalanche scenario. In the MBL case, the growth $\propto \ln(t)$ should saturate around time $t \sim 2^{L_B}$ when the subsystem A starts to resolve the discreteness of the bath's energy spectrum.

The results of our calculations of the correlation decay length $\xi_d(t)$ in all three regimes with a bath of size $L_B = 6$ and the Néel initial state are presented in Figs. 3(a)–3(c). Figure 3(a) shows the system behavior in the ergodic regime with $W_A = 2.5$. The avalanche is clearly spreading through

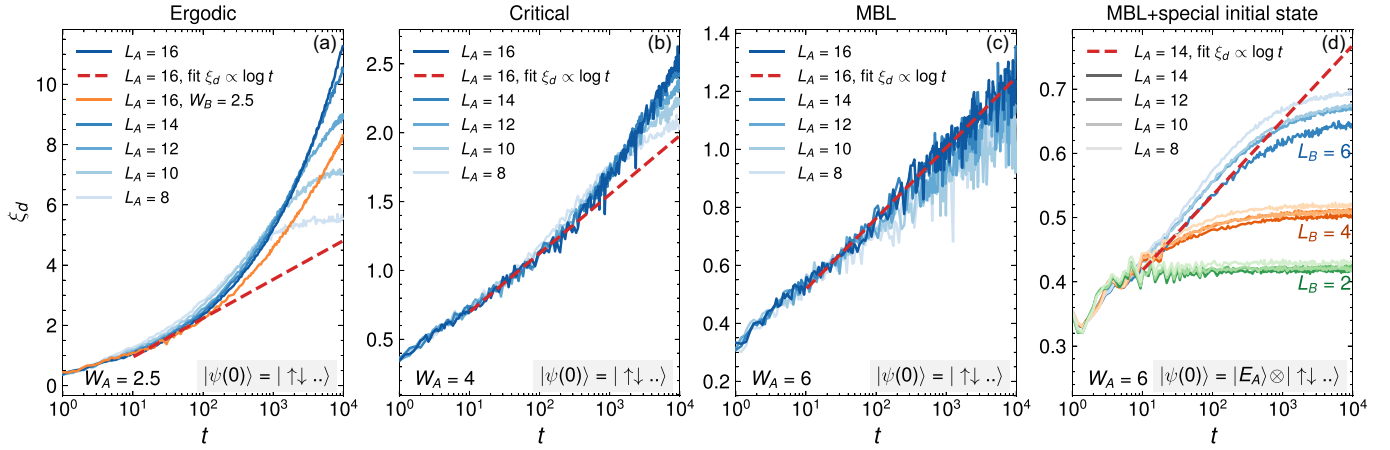


FIG. 3. Growth of the correlation decay length $\xi_d(t)$ in the three regimes of disorder W_A , different subsystem A sizes L_A , bath size $L_B = 6$ and initial Néel state [(a)–(c)] or $L_B = 2, 4, 6$ and a special initial state (d). Thermal avalanches are observed in the ergodic and critical regimes [(a) and (b)] as a growth faster than $\xi_d \propto \ln(t)$. Even though there are no quantum avalanches in the MBL regime (c), the growth of ξ_d continues as $\propto \ln(t)$ much beyond the timescale $t \sim 2^{L_B}$. To see termination of avalanches in the MBL regime, one can start with a fine-tuned initial state, a product state of the subsystem A eigenstate energy E_A (we choose E_A closest to 0), and a Néel state in subsystem B (d). For reference, linear fits $\xi_d \propto \ln(t)$ are performed between times $t = 10$ – 100 (abd) or $t = 10$ – $10\,000$ (c) and denoted by dashed red lines. In (d), 5×10^4 disorder realizations were used. For (d), we employ a special fitting method described in Appendix B.

the system, as suggested by a positive curvature of the decay length $\xi_d(t)$, cf. linear fit between times $t = 10$ – 100 , implying faster-than-logarithmic growth with time. It suggests that some assumptions of the perturbation theory introduced in Sec. II are not valid. In this case, the subsystem A cannot be described by a set of LIOMs. For the smallest system sizes, the positive curvature is also visible, but the growth of $\xi_d(t)$ saturates before reaching $t = 10^4$. It is a finite-size effect occurring when ξ_d becomes of the order of L_A . Increasing L_A systematically increases accessible ξ_d before saturation. We also notice that, in fact, no thermal bath is needed to see features compatible with avalanches – the curve for $W_B = W_A = 2.5$ and $L_A = 16$ behaves qualitatively the same as the case with a bath $W_B = 0.5$. This means that the system can serve as a bath for itself, in agreement with the usual understanding of the thermalization in quantum many-body systems.

In Fig. 3(b), we switch to the critical regime $W_A = 4$. The growth of $\xi_d(t)$ is slower than in the ergodic case, yet it is still faster than $\xi_d \sim \ln(t)$, suggesting the spread of avalanche. With increasing subsystem size L_A , the correlation decay length ξ_d slightly increases, i.e., thermalization is enhanced. This is compatible with exact diagonalization studies [67] which show that increase of the system size increases the disorder strength at which the ETH-MBL crossover is observed.

Figure 3(c) shows the apparent instability of the strongly disordered MBL system ($W_A = 6$) in contact with a thermalizing bath when starting from an initial Néel state. We observe that ξ_d grows as $\propto \ln(t)$ for the whole simulation time $t = 10^4$. Recall that according to the perturbative arguments presented in Sec. II, if the logarithmic growth of ξ_d continues beyond $t \approx 2^{L_B}$, one deals with a continually increasing timescale of thermalization by the bath [see Fermi’s golden rule in Eq. (8)]. This suggests spreading of an avalanche which thermalizes the system and may be surprising as other signatures of MBL, such as energy level statistics, entanglement entropy, Thouless time scaling [67] suggest a deeply localized regime for this value of disorder in the absence of the bath.

On the other hand, a nontrivial long-time dynamics is also observed for the bipartite von Neumann entanglement entropy which grows as $\propto \ln t$ [74] even in MBL regime. Other works [75] predict a much larger value of disorder $W > 18$ needed for the MBL phase. Looking for the explanation of this counterintuitive behavior of the MBL system, we have checked that increasing the disorder strength to $W_A = 20$ does not remove the logarithmic growth until the latest accessible times $t = 10^4$. Removing the interactions in subsystem A and going to Anderson localized phase by setting $\Delta_A = 0$ in subsystem A also leads to a similar long-time logarithmic growth, see details in Appendix C.

It turns out that, in order to observe the termination of a thermal avalanche, as predicted by the perturbation theory for sufficiently small LIOM localization lengths, one has to necessarily fulfill one of its assumption: the initial state must be an eigenstate of the subsystem A . Similar setup was considered in Ref. [105] with the purpose of analyzing energy-resolved features of the spectrum and in Ref. [142] for perturbatively coupled chaotic systems. One can still start with the Néel state in subsystem B as its eigenstates are featureless, and so is their superposition. Results for this initial condition are presented in Fig. 3(d). Clearly, the penetration of the thermal bath into the MBL subsystem is logarithmic with time and then stops around time on the order of 2^{L_B} , in full agreement with the perturbative argument. Behavior of $\xi_d(t)$ is qualitatively the same for many subsystem sizes L_A and fixed bath L_B .

Summarizing, the long-time growth of $\xi_d(t)$ for $W_A = 6$ in Fig. 3(c) seems to be a feature of the MBL regime rather than a signature of an avalanche spreading and thermalization of the system. When the initial state is an eigenstate of the subsystem A , we do not observe any signatures of avalanche spreading for the same system parameters, see Fig. 3(d). The influence of the bath is fully captured with the first order perturbation theory and the highly nonperturbative effects of the avalanche mechanism are not observed at $W_A = 6$. This dependence of

the ξ_d behavior on the initial condition is not occurring at smaller disorder strengths, when the system thermalization genuinely begins, for instance at $W_A = 4$ (data not shown, see Appendix C). Therefore our analysis indicates that the access to the eigenstates of H_A is required for observation of the stop of logarithmic growth of ξ_d in time. Preparation of eigenstate of H_A is experimentally a nontrivial task, increasing the difficulty of an experiment that could observe the halt of quantum avalanche spreading. Finally, we emphasize that we are able to observe the stopping of avalanches only at the considered (experimentally relevant) time scales and system sizes. We cannot exclude that the avalanche starts to spread and the system thermalizes at $W_A = 6$ for larger L_A or longer times.

D. Extraction of the LIOM localization length from time evolution

Linear fits from Fig. 3 allow us to extract the localization length of the LIOMs ξ , see Eq. (3), and compare it with a predicted threshold for avalanche, Eq. (10), $\xi_{\text{aval}} = 2/\ln(2) \approx 2.9$. In both perturbative and nonperturbative pictures, the LIOM at distance r decays around time $t_j \propto \exp(2r/\xi)$. In numerical data from Fig. 3, distance r is measured by the correlation decay rate ξ_d . Thus one can perform a linear fit

$$\xi_d(t) = \frac{\xi}{2} \ln t + \text{const} \quad (15)$$

and extract the LIOM localization length ξ . For $W_A = 2.5$, we get $\xi = 1.1 < \xi_{\text{aval}}$ between times $t = 10 - 100$ which grows to $\xi = 5.4 > \xi_{\text{aval}}$ between times $t = 1000 - 10000$ and is compatible with the presence of an avalanche in Fig. 3(a). In the critical regime $W_A = 4$, the obtained value $\xi = 0.4$ is smaller than the threshold ξ_{aval} . In the localized regime $W_A = 6$, the LIOM localization length extracted from the initial Néel state scenario is $\xi = 0.21$ and is also well below the threshold for avalanche. A lower $\xi = 0.10$ is obtained for the special initial state, signaling some additional effects happening in the Néel state case which enhance the nontriviality of the dynamics in that case.

IV. BATH FEATURES

This section is devoted to the features of the bath that might be relevant for the thermalization process. In Sec. IV A, we show that the bath can be approximated well by all-to-all interacting system modeled by a GOE matrix. In Sec. IV B, we find that the system dynamics does not change qualitatively upon changing disorder amplitude W_B in the bath (provided W_B is small enough so the bath is ergodic by itself). Finally, in Sec. IV C, we study the evolution of correlations within the bath when connected to subsystem A in its three different regimes, capturing the moment when the bath starts to grow. Additional analyses, such as change of the decay length for a fixed bath and changing subsystem A size L_A , or changing disorder strength W_A , are shown in Appendix E.

A. Modeling the bath as a GOE matrix

Many approaches to thermalization in MBL through avalanche scenario assumed that the bath can be modeled by a random matrix from the Gaussian orthogonal ensemble

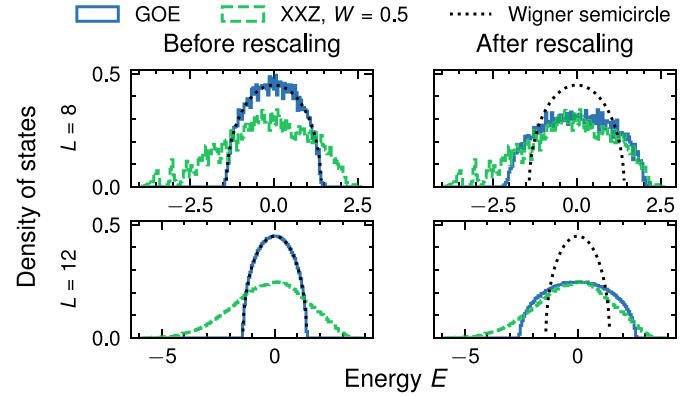


FIG. 4. Comparison of the density of states for the XXZ model with uniform $W = 0.5$ disorder (green dashed line) and the GOE density of states (blue solid line) for the same system size L . Before rescaling, the GOE density of states reproduces the Wigner semicircle law (black dotted line). After rescaling the GOE density by a constant multiplicative factor, as described in the text, the two densities start overlapping in the major part of the spectrum. Histograms are averaged over 100 disorder and GOE realizations.

(GOE) which should behave identically to a region thermalized due to accidentally low disorder values/differences [85,86,143,144]. The full Hamiltonian for such a setting reads

$$H' = J \sum_{i=1}^{L_A} \vec{S}_i \cdot \vec{S}_{i+1} + \sum_{i=1}^{L_A} W_i Z_i + \alpha H_{B,\text{GOE}}, \quad (16)$$

where $H_{B,\text{GOE}}$ is a GOE matrix defined for L_B sites in subsystem B, and α is an energy rescaling factor we will specify later. Having access to the measures directly probing the avalanches, we verify this assumption numerically by comparing the two scenarios. Algorithm for numerical construction of the $H_{B,\text{GOE}}$ in the total S_z magnetization conserving basis is given in Appendix D. Before, we proceed to present the results, we explain a caveat that needs to be taken into account when comparing the two models.

The density of states of $H_{B,\text{GOE}}$ is given by the Wigner semicircle law [145,146],

$$\rho(E) = \frac{1}{\pi} \sqrt{2 - E^2}. \quad (17)$$

Its spectrum spans an interval of energies $-\sqrt{2} < E < \sqrt{2}$. In order to reproduce bandwidth of the bath in the XXZ model, the GOE Hamiltonian is multiplied by a factor $\alpha = (p_{95} - p_5)/(2\sqrt{2})$, where p_{95} and p_5 are the 95th and 5th percentile of the disorder-averaged XXZ density of states at $W = 0.5$. The result of rescaling energies is presented in Fig. 4. The rescaling gives a real symmetric random matrix whose middle spectrum energy gaps reproduce the middle spectrum gaps of the XXZ spin chain with disorder $W = 0.5$ in the zero magnetization sector. Numerical values of the rescaling factors for different lengths of the chain $L = 2-12$ and other “replicated” disorder strengths $W = 0, 0.5$, and 1 are listed in Appendix D.

Naively, one could expect that replacing the nearest-neighbor interacting low-disorder spin bath with an all-to-all interacting GOE bath will speed up the thermalization in the

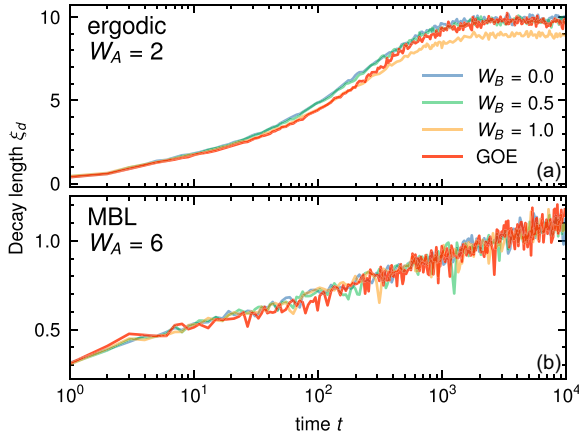


FIG. 5. Decay length growth with time. $L_A = 8$ and $L_B = 6$. Precise value of the bath disorder strength W_B does not qualitatively change the behavior, irrespective if it is in the ergodic (a) or the MBL side of the crossover (b). Thus our analysis is restricted to $W_B = 0.5$ for the rest of the paper. We compare it with the bath modelled as a GOE matrix coupled to a single spin at the end of the chain. The GOE matrix is rescaled to replicate the density of states of the bath with $W_B = 0.5$, see text.

system. The GOE Hamiltonian scrambles quantum information very fast, on the timescale $\mathcal{O}(1)$ [147,148]. In contrast, the XXZ model has Thouless time that scales as $\mathcal{O}(L^2)$ at best [66,73]. In spite of this, we do not observe any significant differences between these two cases when inspecting the correlation decay length $\xi_d(t)$ for subsystem A in both ergodic ($W_A = 2$) and MBL regime ($W_A = 6$), see Fig. 5. To our understanding, it is not a surprise that the long-time $t \geq \mathcal{O}(L_B^2)$ behaviors are the same for the GOE bath and low-disorder XXZ model bath, as in both cases the baths have already thermalized and are in a similar featureless quantum state with random probability amplitudes dependent only on the initial energy. However, an unprecedented agreement between these scenarios also at initial stages of time evolution remains a puzzle. One possible explanation is that if the subsystem A is connected with the bath only at its end by a local term in the Hamiltonian, thermalization of the distant site from the interface is mostly governed by the rate of propagation of correlations through subsystem A up to the interface, which does not depend on the bath properties.

B. Disorder strength W_B

As mentioned in Sec. IV A, the low-disorder bath can be well modelled by the GOE Hamiltonian. What is more, $\xi_d(t)$ calculated in Fig. 5 does not change with the disorder strength of the bath W_B as long as W_B stays far on the ergodic side of the crossover, $W_B = 0, 0.5$, and 1 . This is particularly important as we are trying to use the “engineered” bath to model “accidental” baths naturally present in the system due to accidentally small disorder values. Interestingly, even for the case of $W_B = 0$ the thermalizing behavior is consistently the same. Integrability of the bath is broken in this case by the interaction with subsystem A .

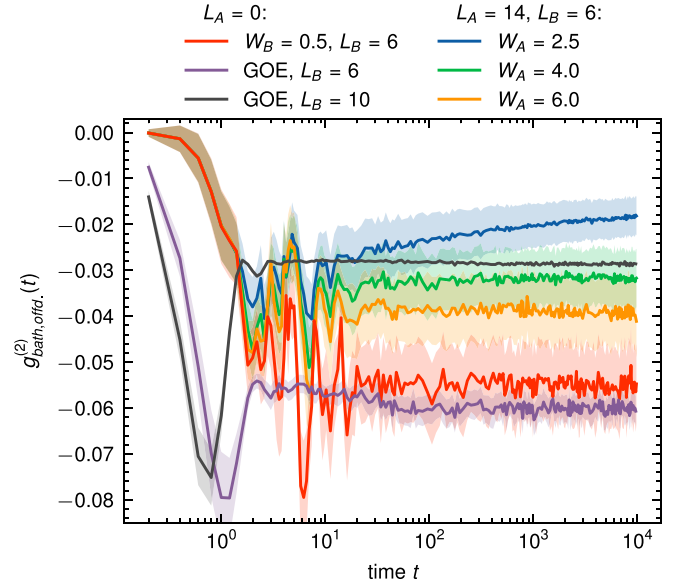


FIG. 6. Mean off-diagonal correlations within the bath as a function of time. Shades denote standard deviations of the median. Results for the disconnected bath ($L_A = 0$) with weak disorder $W_B = 0.5$ (red line) and the GOE bath of different sizes (purple and gray line) are compared with the $L_A = 14$, $L_B = 6$ case. Quantum avalanche is visible in the $W_A = 2.5$ case since growth of offdiagonal correlations indicates an increase of the effective size of the bath.

C. Correlations within the bath

We now concentrate on the reaction of the bath when it is interacting with the disordered XXZ chain. As it was suggested in earlier studies [94,149], the ergodic bath may be destroyed when in contact with an MBL system by the so-called “MBL proximity” effect. In an opposite scenario, a quantum avalanche may extend the bath into the MBL subsystem. Here we study the off-diagonal correlations of spins within the bath to characterize the bath during the time evolution.

Figure 6 shows the average (over the bath) of the offdiagonal correlations within the bath,

$$g_{\text{bath, offd.}}^{(2)} = \frac{1}{L_B(L_B - 1)} \sum_{\substack{i,j=L_A+1 \\ i \neq j}}^L \langle Z_i Z_j \rangle_c. \quad (18)$$

As a limiting case, we check a bath that is modeled by the GOE Hamiltonian from Sec. IV A with $L_A = 0$ (no subsystem A) and $L_B = 6, 10$. Clearly, the correlations saturate fast at a negative value and remain stationary - all pairs of spins within the bath develop correlations between each other until no more correlations can be built further. The larger the system, the smaller the absolute value of the bath averaged correlations, since there are more possibilities to transition from a single site. A similar thermalizing behavior is observed for a weakly disordered spin bath ($W_B = 0.5$), also for $L_A = 0$, but reaching the equilibrium takes longer compared to the GOE case for the same size $L_B = 6$, see discussion in Sec. IV A.

Let us check what happens when the subsystem A is attached in Fig. 6. We examine different disorder strengths $W_A = 2.5, 4$, and 6 for fixed $L_A = 14$ and $L_B = 6$. At $W_A = 6$,

we see clear signatures of saturation of bath correlations around time $t \approx 10$. Similar behavior is observed in the critical regime $W_A = 4$ but the saturation occurs at a larger value. Since the correlations saturate at a steady value, we may conclude that the bath does not change its size at least to the accuracy offered by this probe. Comparison to the baseline values for the free and GOE baths suggests an increase of the effective bath size for $W_A = 4$ in comparison to the $W_A = 6$ case. Comparing with GOE bath for $L_B = 10$ leads us to the conclusion that the size of the effective bath for $W_A = 4$ case cannot be larger than 10 sites. Interestingly, a continuous dynamic drift in the offdiagonal correlations is observed for $W_A = 2.5$. This is another sign of a quantum avalanche spreading, in addition to the correlation decay length presented in Fig. 3. Additional sites from subsystem A join the bath by developing correlations with its constituents. The increase in the effective size of the bath weakens the correlations therein.

All in all, the dynamics of the correlation functions within the bath allows to probe the thermalization of the full system from a different perspective, providing indirect information about the increase of the effective bath size.

V. KICKED XXZ MODEL

In their recent work, Peacock and Sels [104] study a model similar to ours: a strongly disordered Heisenberg spin chain A with a bath B consisting of weakly disordered spins ($W_B = 0.5$). However, in their model, the Hamiltonian is periodic in time. The unitary evolution operator over one period of time reads

$$U = e^{-iH_{AB}\tau} e^{-i(H_A+H_B)\tau}, \quad (19)$$

where $H_{A,B,AB}$ are defined in Eq. (12). The subsystems A and B evolve separately for half of the period $T = 2\tau = 20$ and for the second half only the A - B coupling yields the dynamics with H_A and H_B being put to zero. We dub this system “kicked” as the form of the unitary, (19), allows one to interpret the dynamics as evolving with $H_A + H_B$ for period τ and then undergoing a δ kick of strength $H_{AB}\tau$ resembling thus famous kicked rotator or kicked top models [150]. The time dynamics is studied by looking at the imbalance

$$\mathcal{I}(t) = \frac{4}{L_A - 4} \sum_{i=3}^{L_A-2} \langle Z_i(0)Z_i(t) \rangle, \quad (20)$$

for system A of length L_A and B of length L_B – note that two spins on each side of the system A are not included to minimize edge effects.

We compare the imbalance obtained for the kicked model (19) with that obtained for autonomous system governed by Eq. (12) for the same parameter values. Reproducing the results of [104] we note that in the presence of the thermal inclusion the system A shows a much faster decay than in the case $W_B = W_A$ without a thermal inclusion. This difference is noticeable even for strongly localized chains at $W_A = 12$. At sufficiently long times, the imbalance decays, suggesting thermalization of the system in the wide range of disorder values $W_A = 4$ –12. The onset of the decay of the imbalance scales with disorder as $t \rightarrow te^{-aW_A}$, with constant a found to be $a = 2.1$ for $L_A = 12$, $L_B = 10$, $W_B = 0.5$. The latter

behavior bears some resemblance to the Thouless time scaling $t_{Th} = L^2 e^{W/\Omega}$ (where Ω is a constant) proposed for XXZ spin chain in Ref. [73].

Intuitively, one could expect that since the bath B is connected with the subsystem A at all times in our case, contrary to kicked type of dynamics generated by (19), the autonomous system should thermalize faster. On the other hand, the energy is absorbed from the driving in the kicked case, favoring the thermalization of the kicked system, unless the driving is very fast [151].

The comparison of both models is presented in Fig. 7. The first column of Figs. 7(a) and 7(e) shows the imbalance as a function of time in both models. It is clear that although the signs of thermalization in the kicked model start to appear at longer times ($t \approx 200$ –300) than in the time-independent model, the acceleration of the imbalance decay at largest disorder values is much more evident in the kicked case. This is further demonstrated in the middle column of Figs. 7(b) and 7(f) which shows how the difference of imbalance in the case of no bath ($W_A = W_B$) and with the bath ($W_B = 0.5$) accelerates with time for the kicked model even for $W_A = 12$. In contrast, the growth of the imbalance difference does not accelerate for the time-independent model. On the contrary, it rather saturates, at sufficiently large W_A , to a constant (which is possibly dependent on L_A). This result is consistent with the MBL regime and is in a stark difference with the kicked case.

Finally, in the rightmost column of Figs. 7(c), 7(d) and 7(g), we check whether the rescaling of time $t \rightarrow te^{-aW_A}$, which captures the dynamics of the kicked model in a broad regime of W_A , applies also to the energy-conserving model. We find the optimal a describing the data by fitting a fourth degree polynomial to the imbalance and minimizing the fit error by tuning a . To check if a is universal and does not depend on disorder W_A , our fits are performed separately for small disorders ($W_A \in [2, 5]$) on the ergodic side of the crossover [Fig. 7(c)] and for large disorders ($W_A \in [10, 13]$) on the localized side. Fitting errors are measured only in the corresponding fitting range. The resulting optimal a are $a = 6$ and 11, respectively. In the latter case, the data can be in fact almost equally well described by any $a \in [8, 15]$, as the fit error does not change in this interval by more than a few percent. Notably, due to the presence of the exponent, the difference between $a = 6$ and $a = 11$ cases is huge, the rescaled times differ by 15(!) orders of magnitude in the middle of the interval of disorder strengths considered. The lack of agreement with a found for low disorder values suggests that the time-independent system does not have a universal timescale dictated by the disorder value. On the other hand, for the kicked model, the same procedure gives a ranging from $a = 1.6$ for the largest disorders $W_A \in [10, 12]$ to $a = 2.3$ for the smallest disorders $W_A \in [4, 6]$, and the optimal a is well-defined in both cases, staying in agreement with $a = 2.1$ from Ref. [104]. We note that the dependence of a on the interval of disorder strengths used to determine its value, is opposite in the two models. This highlights further the difference in the behavior of the two models at large W_A . Details on how we determine values of the constant a are described in Appendix G.

Summarizing, the results of Ref. [104] are reproduced for the kicked case (note that we use numerically exact

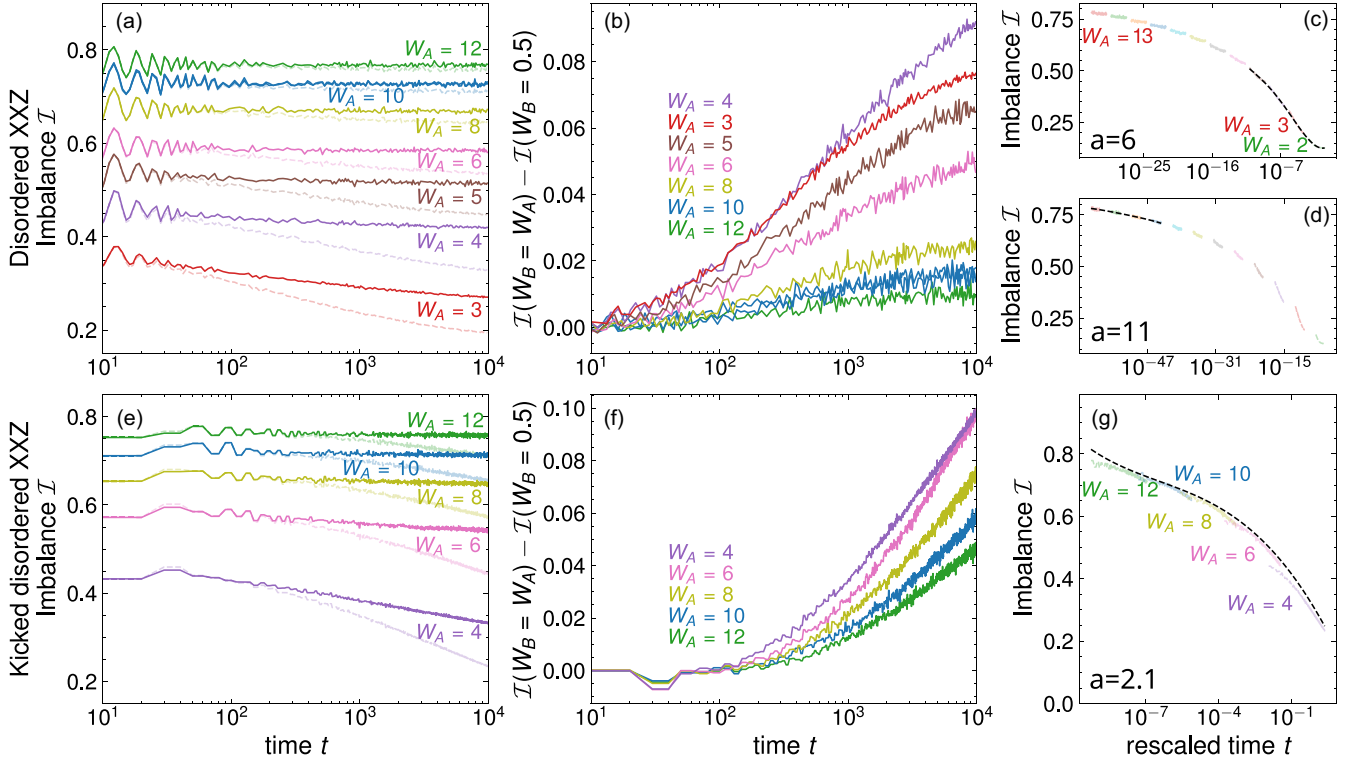


FIG. 7. Comparison of imbalance on sites 3–10 between kicked XXZ model by Peacock and Sels and the energy-conserving model for $L_A = 12$ and $L_B = 10$. Solid lines: $W_A = W_B$, bright lines: $W_B = 0.5$ [(a), (c)–(e), and (g)]. The speedup of imbalance decay happens due to the lack of energy conservation (e) because it is not observed in the time-independent case for the same parameters (a). [(b) and (f)] Differences of imbalance with $W_B = W_A$ and $W_B = 0.5$. Speedup of imbalance decay is visible for the kicked model, while saturation is observed for the time-independent model. Right column: Imbalance with rescaled time axis $t \rightarrow te^{-aW_A}$ for the case $W_B = 0.5$. Optimal constant a is obtained by minimizing the error of a fourth degree polynomial fit (dashed black line) and a chosen range of disorder. While in the kicked case (g), there is a single universal constant $a \approx 2.1$ describing the curve for all $W_A \in [2, 12]$, in the disordered XXZ model case the optimal constant a drifts from $a \approx 6$ for $W_A \in [2, 5]$ to $a \in [8, 15]$ for $W_A \in [10, 13]$. See text and Appendix G for more details on the fitting procedure.

Chebyshev time propagation rather than the tensor network time-evolved block decimation method (TEBD) that is approximate). The results for the kicked and time independent case differ qualitatively for the considered system size. The tendency towards the thermalization of the setup of [104], even at the strong disorder $W_A = 12$, does not translate to the energy-conserving setup. While the rescaling $t \rightarrow te^{-aW_A}$ suggests uniformity of the thermalizing behavior in the kicked case across the broad interval of disorder strengths, the same rescaling does not allow to collapse the results for our autonomous case with a reasonable accuracy. Importantly, it is the autonomous, time independent case which is relevant for our understanding of an avalanche spreading in isolated quantum spin chains. Indeed, if disorder fluctuations produce a region of anomalously weak disorder, its interaction with the surrounding spins is captured by a time independent Hamiltonian.

VI. CONCLUSIONS AND DISCUSSION

We have analyzed the effects of rare thermal bubbles, which arise in disordered spin chains due to disorder fluctuations on thermalization of the system. While the behavior of the ergodic bubbles determines the stability of MBL phase,

their direct investigation is not straightforward since the probability of appearance of an ergodic bubble is exponentially small in its size. For that reason, we consider a scenario of a finite thermal bath (XXZ spin chain with low disorder) attached to an XXZ spin chain with uniformly distributed random disorder which plays the role of an MBL subsystem A. Based on the time evolution of the system, we study the decay length $\xi_d(t)$ of connected two-body correlations $\langle Z_i Z_j \rangle_c$ of the MBL subsystem with the bath. Our perturbative analysis of the system-bath coupling identifies the conditions under which the MBL system is not thermalized by the bath. In this case, $\xi_d(t) \propto \ln(t)$ for a time scaling exponentially in the bath size, followed by a saturation. In contrast, the quantum avalanche mechanism includes nonperturbative effects, leading to thermalization, manifested as a growth $\xi_d(t) \propto \ln(t)$ or faster until time scale which scales exponentially with the size of the total system.

We perform an extensive numerical study, obtaining a superlogarithmic growth of $\xi_d(t)$ in the critical and ergodic regimes. Such behavior is consistent with the quantum avalanche mechanism, although we cannot exclude that a different microscopic mechanism plays a role in the observed thermalization process. Surprisingly, the $\xi_d(t) \propto \ln(t)$ growth

until the Heisenberg time is also observed in the deep MBL regime if the initial state is set to be a simple product state. However, the above result is not a sign of thermalization but rather a long-timescale feature of the dynamics, not captured by the theory. The fact that the system does not thermalize becomes apparent only if the time evolution starts from an initial state being a product state of an energy eigenstate in the MBL subsystem A and Néel state in the bath. In this case, $\xi_d(t)$ saturates at a constant after a time fixed by the bath size. The onset of thermalization is visible in the same setting in the ergodic and critical regimes. Thus we conclude that the observation of $\xi_d(t) \propto \ln(t)$ long-time growth from an initial product state does not allow us to unambiguously determine whether the avalanches propagate or not.

As we show in Sec. III D the rate of growth of $\xi_d(t)$ allows us to extract the LIOM localization length ξ . Extracted LIOM localization lengths in all three regimes (ergodic, critical or MBL) of subsystem A are consistent with the predicted ETH-MBL crossover at $\xi = 2/\ln(2)$. We further demonstrate that the thermalization process does not depend on the details of the bath, as long as the bath is an ergodic system on its own. In particular, changing the bath disorder strength between $W_B = 0-1$ does not qualitatively change the decay length growth with time. Moreover, instead of the XXZ spin chain with low disorder W_B serving as a bath, one can attach a bath whose Hamiltonian is a properly rescaled GOE matrix and the thermalization characteristics manifested by the behavior of $\xi_d(t)$, remain the same to a good approximation.

We also inspect the offdiagonal correlations within the bath during the time evolution. The correlations in the bath disconnected from the MBL subsystem A quickly saturate to time independent values. Similarly, when the bath is connected to XXZ spin chain in the MBL regime, the correlations saturate at a similar time scale, reflecting the lack of non-trivial dynamics between the MBL system and the bath. In contrast, when the XXZ spin chain in subsystem A is in an ergodic regime, we observe a steady relaxation of the off-diagonal correlations within the bath, which is a direct consequence of a quantum avalanche increasing the effective bath size.

Finally, we contrast our calculations of the imbalance with the results from the kicked XXZ model [104]. Unlike in the kicked model, the time rescaling $t \rightarrow te^{-aW_A}$ does not allow us to collapse the results for time evolution of the imbalance for a wide range of disorder strength, W_A , in our energy conserving setup. Also, in our model, the change in the behavior of imbalance with time, with and without a bath, is not as dramatic as for the kicked case. This suggests that thermalization in the kicked model is enhanced by the driving and one should not generically expect the same behavior for the energy-conserving Hamiltonian model, relevant for the question of MBL in autonomous systems.

Our results shed new light on the avalanche mechanism present in the disordered XXZ model. We have provided numerical evidence consistent with the avalanche spreading from an engineered thermal inclusion and have shown that their fate depends on the disorder strength W_A . At the same time, we have demonstrated that an observation of the termination of spreading of quantum avalanches is more challenging as the system has to be prepared in an eigenstate of the MBL subsystem A . Notably, preparation of area-law entangle-

ment initial states of the MBL subsystem A , is experimentally feasible [152].

Our work demonstrates that the disorder strengths in the XXZ spin chain at which the avalanches propagate and are terminated correspond to regimes of disorder strengths identified respectively as ergodic and MBL in the exact diagonalization studies of the model, for example see Ref. [67]. The parametric shift of the disorder strengths required for termination of avalanches observed in Refs. [75,100] with respect to the exact diagonalization studies may, therefore, be attributed to the assumptions about the nature of the ergodic bubble made in Refs. [75,100]. The results presented in our work are valid for ergodic bubbles placed on several lattice sites and do not support those assumptions. Nevertheless, we cannot exclude the possibility that ergodic bubbles of much larger sizes, which can be found in thermodynamically large systems, seed avalanches that propagate in presence of much stronger disorder. The length and time scales associated with such rare events are, most likely, not relevant for numerical or experimental investigations of disordered many-body systems.

The numerical data presented in this work are freely available from Ref. [153] or from the authors upon a reasonable request.

ACKNOWLEDGMENTS

P.S. acknowledges interesting discussions with M. Hopjan, R. Świątek, and L. Vidmar. This research has been supported by the National Science Centre (Poland) under project 2019/35/B/ST2/00034 (T.S.) and under the OPUS call within the WEAVE program 2021/43/I/ST3/01142 (J.Z.). The research was also supported by a grant from the Priority Research Area (Visibility and Mobility) and Research Area DigiWorld under the Strategic Programme Excellence Initiative at Jagiellonian University (T.S.). We gratefully acknowledge Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational Grant No. PLG/2023/016370.

P.S. and M.L. acknowledge support from: European Research Council AdG NOQIA; MCIN/AEI (PGC2018-0910.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-I00, Plan National STAMEENA PID2022-139099NB, I00, project funded by MCIN/AEI/10.13039/501100011033 and by the “European Union NextGenerationEU/PRTR” (PRTR-C17.I1), FPI); QUANTERA MAQS PCI2019-111828-2); QUANTERA DYNAMITE PCI2022-132919, QuantERA II Programme co-funded by European Union’s Horizon 2020 program under Grant Agreement No 101017733); Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM ENIA project call - Quantum Spain project, and by the European Union through the Recovery, Transformation and Resilience Plan - NextGenerationEU within the framework of the Digital Spain 2026 Agenda; Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program, AGAUR Grant No. 2021 SGR

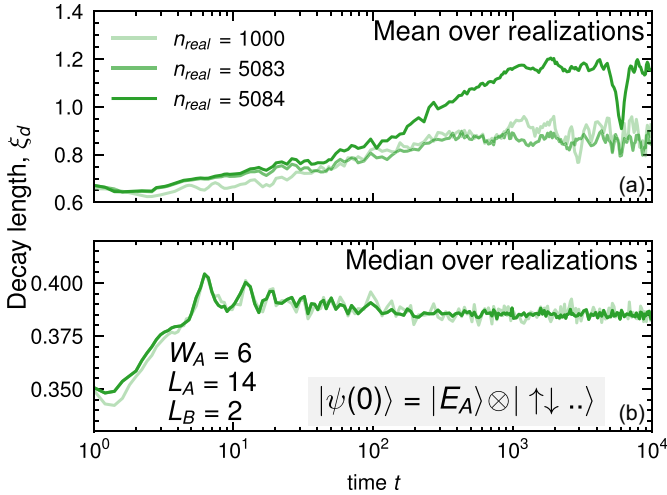


FIG. 8. Growth of the decay length $\xi_d(t)$ calculated for noninteracting Anderson model ($\Delta_A = 0$) with an interacting bath ($\Delta_B = 1 = \Delta_{AB}$) for different numbers of random realizations n_{real} . Including a single extra anomalous realization of disorder (low disorder close to the interface) affects the mean over realizations significantly (panel a), while the median stays unchanged (panel b).

01452, QuantumCATU16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020); Barcelona Supercomputing Center MareNostrum (FI-2023-3-0024); Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, European Commission, European Climate, Infrastructure and Environment Executive Agency (CINEA), or any other granting authority. Neither the European Union nor any granting authority can be held responsible for them (EU Quantum Flagship PASQuanS2.1, 101113690, EU Horizon 2020 FET-OPEN OPTologic, Grant No 899794), EU Horizon Europe Program (This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101080086 NeQSTGrant Agreement 101080086 NeQST); ICFO Internal "QuantumGaudi" project; European Unions Horizon 2020 program under the Marie Skłodowska-Curie grant agreement No 847648; "La Caixa" Junior Leaders fellowships, La Caixa" Foundation (ID 100010434); CF/BQ/PR23/11980043.

APPENDIX A: USING MEDIAN OVER DISORDER REALIZATIONS

Throughout the analysis of the data supporting the paper we have encountered issues with the convergence of the mean over realizations due to rare thermalization events. We have observed the largest difficulties with the convergence for the case of an Anderson insulator in subsystem A, $\Delta_A = 0$, connected to an interacting bath, $\Delta_B = 1 = \Delta_{AB}$, see Eq. (11) in the main text. Figure 8 shows the decay length $\xi_d(t)$ obtained by fitting an exponential decay to the $g^{(2)}(i)$ function calculated as a mean [Fig. 8(a)] or median [Fig. 8(b)] over random disorder realizations. We have found that even with more than 5000 realizations, including a single anomalous realization in

the calculation leads to a drastic change in the mean value of $g^{(2)}(i)$, while not affecting the median $g^{(2)}(i)$. We have checked that this originates from a huge increase of the mean $g^{(2)}(i)$ function due to thermalization caused by accidentally small values of disorder in part of subsystem A close to the bath, i.e., an anomalous rare event. Thus, to stay consistent with the rest of the results presented in the main text, we use the median for all system configurations. It is important to note that for larger values of $\xi_d \gtrsim 1$ we have observed no difference between the mean and the variance.

APPENDIX B: COMMENT ON THE METHOD OF THE DECAY LENGTH FITTING

We define the correlation decay length ξ_d according to Eq. (14) in the main text. To calculate ξ_d , we fit Eq. (14) to numerical data for $|g^{(2)}(i)| > 10^{-8}$ in order not to include numerical noise.

Another method, chosen in Ref. [103] for bosonic system, is to calculate the first moment,

$$\xi_{d,\text{FM}} = - \sum_{i=1}^{L_A} (L_A - i) g^{(2)}(i). \quad (\text{B1})$$

We believe this definition lacks an easily interpretable meaning in the spin system as it has a unit of magnetization multiplied by distance. Secondly, comparison between different system sizes is complicated as for a larger system size L_A there are more terms in the sum above and the change in value of $\xi_{d,\text{FM}}$ could be related solely to the size change. Thus, for our purpose of measuring the distance of bath penetration into the MBL system, we propose an alternative to decay length obtained from exponential fit, i.e., the "center-of-mass" correlation length,

$$\xi_{d,\text{CoM}} = \frac{\sum_{i=1}^{L_A} (L_A - i) |g^{(2)}(i)|}{\sum_{i=1}^{L_A} |g^{(2)}(i)|}. \quad (\text{B2})$$

Although the two definitions of ξ_d and $\xi_{d,\text{CoM}}$ seem more or less equivalent, practical calculation of $\xi_{d,\text{CoM}}$ is more sensitive to (i) discreteness of the lattice and (ii) finite size effects. To demonstrate that, we perform a simple numerical experiment in which we assume correlations with perfect artificial exponential decay with

$$\xi_d(t) = \log_{10} t \quad (\text{B3})$$

and use two methods to recover this value for finite system with $L_A = 14$. Result is shown in Fig. 9. The method of fitting an exponential decay gives a correct $\xi_d(t)$, and there are no finite size distortions. On the other hand, the center-of-mass method underestimates ξ_d for the smallest values of $\xi_d \lesssim 0.5$, while having an incorrect positive curvature instead of being a straight line. For systems deeply in the MBL phase, with a low ξ_d , this could be incorrectly interpreted as a speedup due to avalanche spreading, whereas the true decay length grows as in Eq. (B3). Going further up to $\xi_d \gtrsim L_A/5$, the linear trend on the logarithmic scale is recovered and the center-of-mass result differs from the true value by an additive factor of -0.5 . This is not problematic, as a constant shift does not change the conclusions about avalanches. However, at larger

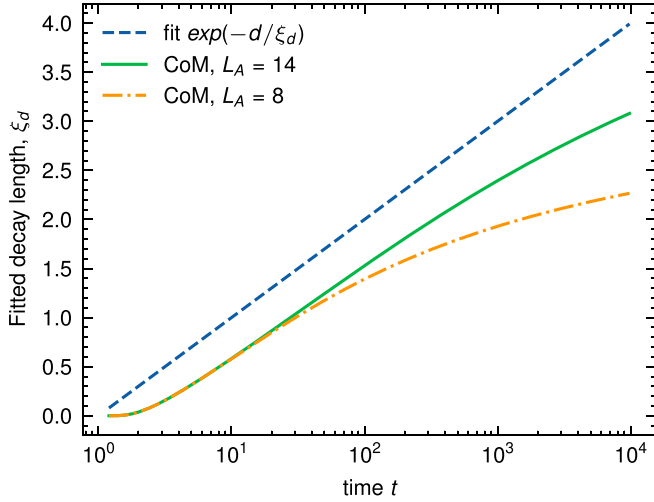


FIG. 9. Decay lengths obtained by fitting an exponential decay $\exp(-d/\xi_d)$ (solid line) (with d —distance to interface) and calculating the center of mass, Eq. (B2) (dashed and dotted-dashed lines) assuming an artificial model of correlations with $\xi_d(t) = \log_{10} t$. Center-of-mass calculations, Eq. (B2), would falsely suggest that the long-time growth of $\xi_d(t)$ is slower than $\xi_d \sim \log_{10}(t)$, and are visibly affected by finite size effects even for $\xi_d/L_A \sim 5$. Due to discreteness of the lattice, for $\xi_d \lesssim 0.5$, the faster-than-logarithmic growth is erroneously implied by the CoM result.

values, $\xi_d \gtrsim L_A/5$, a clear negative curvature starts to appear due to finite size effects. This underestimation of the decay length could be interpreted as a lack of avalanches, in case the real decay length actually speeds up. As a consequence, in the main text we use exponential decay fits instead of the center-of-mass method. Even though the assumption of an exponential decay is needed here, the method does not falsely under- or overestimate the result, and change the conclusion about avalanches.

Due to small values of the decay length ξ_d in Fig. 3(d) in the main text, using the exponential fit revealed “jumps” of ξ_d in time once new sites reached the threshold of $|g^{(2)}(i)| =$

10^{-8} and started to be included in the fit. In order to avoid this artifact, we calculated $\xi_{d,\text{CoM}}$, and translated the result back to ξ_d by utilizing the empirical relation between ξ_d and $\xi_{d,\text{CoM}}$ found in Fig. 9. In this way, we are able to compare the result with the other results obtained solely by fitting an exponential decay.

APPENDIX C: CONDITIONS FOR OBSERVING TERMINATION OF AVALANCHES

In Fig. 3(c) in the main paper, we have observed a rather unintuitive behavior: for $W_A = 6$, $\xi_d(t)$ continues its growth until the latest times as $\xi_d(t) \propto \ln(t)$, in contrary to other measures from literature that suggest strong MBL in this regime of disorder strength and sizes. Thus we speculated that this behavior is caused by either an initial state not being an eigenstate of H_A assumed in theory from Sec. II or by interactions. Results of these investigations are presented in Fig. 10.

Figure 10(a) with the Anderson insulator ($\Delta_A = 0$) starting from an initial Néel state shows that the long-time dynamics does not disappear even in the case of no interactions. This is surprising as Anderson localization suppresses transport through the system. This suggests the explanation of the long-time growth of $\xi_d(t)$ for the MBL system in Fig. 3(c) of the main text—it is a feature of the dynamics and not an indication of the quantum avalanche. Indeed, setting the initial state in the Anderson insulator case to the product of an Anderson orbital (eigenstate with energy E_A which we choose to be closest to 0) and the Néel state in the bath, Fig. 10(b), we restore the saturation of $\xi_d(t)$ growth at large times.

Finally, we check whether the special choice of the initial state affects the behavior of the model in the ergodic and critical regime. In Fig. 10(c), we demonstrate that for $W_A = 2.5$ in the ergodic regime there is a speed up of $\xi_d(t)$ beyond $\xi_d(t) \propto \ln(t)$, a signature of a quantum avalanche. The same happens in the critical regime $W_A = 4$ in Fig. 10 but the speed up appears later.

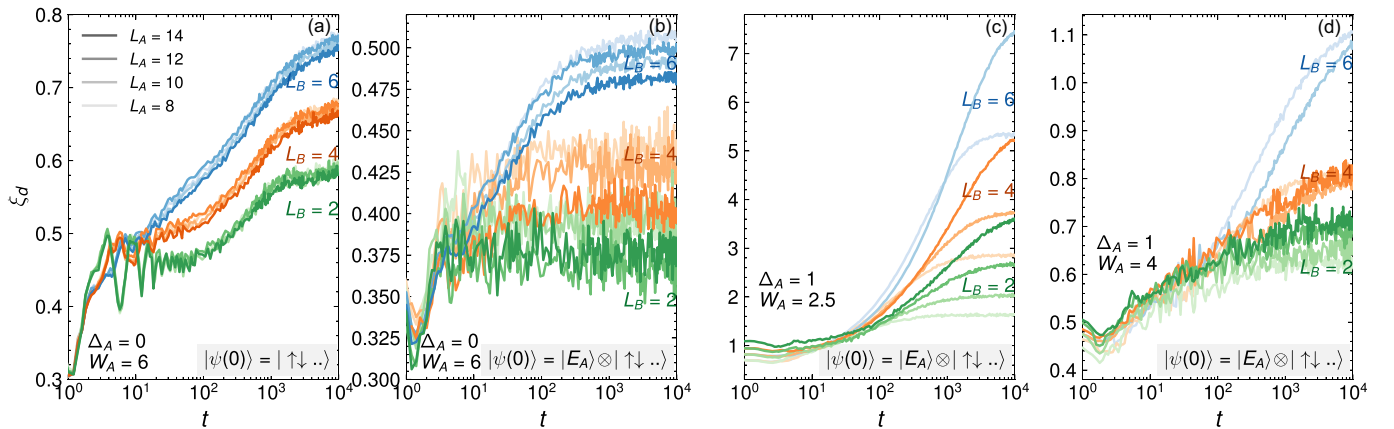


FIG. 10. Correlation decay length $\xi_d(t)$ calculations in different regimes of the XXZ model, supplementing Fig. 3 in the main text. (a) Even the localized Anderson insulator ($\Delta_A = 0$) shows a steady increase of the decay length with time for an initial Néel state. (b) Anderson insulator and an initial state being a product of an Anderson orbital in subsystem A (we choose energy E_A closest to 0) and Néel state in the bath allows us to observe saturation in $\xi_d(t)$. (c) Starting from a special initial state still shows thermalization in the interacting system in its ETH regime, $W_A = 2.5$. The same is observed in the critical regime with $W_A = 4$ in (d). In all cases, $\Delta_B = 1 = \Delta_{AB}$.

ALGORITHM 1. Embedding the GOE Hamiltonian in a Z-conserving basis.

```

basis ← Z-conserving basis of total dimension  $N$ , given by integers which contain spin configurations in binary format.
 $H_{XXZ}$  ← Standard XXZ Hamiltonian matrix of dimension  $N \times N$  with hopping, interaction and disorder for the first
 $L_A$  spins, expressed in basis. It also includes hopping and interaction for the pair of spins  $L_A$  (last in subsystem A)
and  $L_A + 1$  (first in subsystem B).
 $\alpha$  ← GOE spectrum rescaling factor from Table A1.
 $A$  ← Matrix of  $2^{L_B} \times 2^{L_B}$  random Gaussian variables of mean 0 and variance 1.
 $G \leftarrow 0.5 * (A + A^T)$  // GOE matrix of (excess) dimension  $2^{L_B} \times 2^{L_B}$ .
maskB ←  $(1 \ll L_B) - 1$  // Uses binary left shift  $\ll$ . Binary number that has 1 on the  $B$  part of the spin
configuration and 0 on A, eg. 000011 for  $L_A = 4$ ,  $L_B = 2$ .
 $X$  ← Empty list of dimension  $N$ .
for  $i$  in range( $N$ ) do
    // To find normalization factor of the GOE matrix (number of non-zero elements) so that the final spectrum is
    // the Wigner semicircle  $-\sqrt{2} < E < \sqrt{2}$ .
     $X[i] \leftarrow \text{basis}[i] \& \text{maskB}$  // Binary AND. Gives configuration on subsystem B.
end
 $n \leftarrow$  Number of unique binary numbers in  $X$ .
for  $i$  in range( $N$ ) do
     $\text{configAi} \leftarrow \text{basis}[i] \gg L_B$ 
     $\text{configBi} \leftarrow \text{basis}[i] \& \text{maskB}$ 
    for  $j$  in range( $N$ ) do
         $\text{configAj} \leftarrow \text{basis}[j] \gg L_B$ 
         $\text{configBj} \leftarrow \text{basis}[j] \& \text{maskB}$ 
        if  $\text{configAi} == \text{configAj}$  then
             $H_{XXZ}[i, j] \leftarrow H_{XXZ}[i, j] + 0.5 * \alpha * G[\text{configBi}, \text{configBj}] / \sqrt{n}$  // Matrix elements of  $G$  selected based on
            // integers uniquely defining configuration in the  $B$  part of the chain.
             $H_{XXZ}[j, i] \leftarrow H_{XXZ}[i, j]$ 
        end
    end
end
return  $H_{XXZ}$ 

```

APPENDIX D: GOE HAMILTONIAN—MAGNETIZATION CONSERVATION AND RESCALING

In order to reproduce the finite spin chain bath by the GOE, the $H_{B, \text{GOE}}$ matrix is supposed to conserve the Z component of the total magnetization in the whole system. The prescription we use for the GOE matrix construction is described as algorithm1.

Furthermore, we rescale the energy spectrum of a GOE matrix from $-\sqrt{2} \leq E \leq \sqrt{2}$ to replicate the middle part of energy spectrum of the weakly disordered XXZ chain. We calculate the spectrum of the XXZ chain and find $\alpha = (p_{95} - p_5) / (2\sqrt{2})$ where p_x is the x -th percentile of the energy distribution. Values of α for different system sizes and disorder values are listed in Table I. Result of the spectrum rescaling is presented in the main paper as Fig. 4. We have also verified in Fig. 11 that the total spectrum of the combined XXZ and GOE Hamiltonians, given by Eq. (16), reproduces the original spectrum of the XXZ model with two regions of different disorder amplitudes, Eq. (12).

APPENDIX E: FURTHER NUMERICAL EXPERIMENTS

In this section of Appendix, we perform further numerical studies of the interplay of the system and bath parameters and their effect on thermalization.

1. Fixed bath, changing subsystem A size

In Fig. 12, we check how the long-time value of $\xi_d(t)$ changes with L_A for two bath sizes $L_B = 2$ and 6. In particular, we observe that in the ETH phase ($W_A = 2.5$) ξ_d grows with

TABLE I. Factors used to rescale the GOE Hamiltonian. In the main paper, values for $W = 0.5$ were used.

W	L	scale α
0	2	0.354
	4	0.837
	6	1.113
	8	1.353
	10	1.574
	12	1.720
0.5	2	0.428
	4	0.927
	6	1.232
	8	1.459
	10	1.666
	12	1.823
1	2	0.600
	4	1.115
	5	1.468
	8	1.684
	10	1.915
	12	2.091

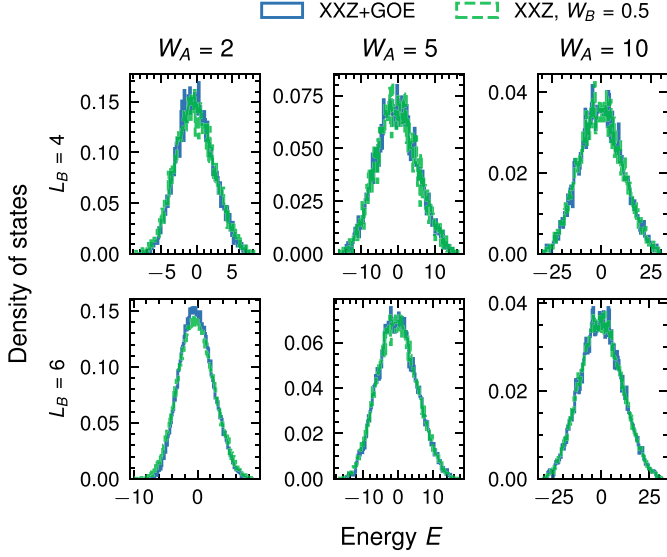


FIG. 11. When bath of size L_B is modeled by a GOE matrix rescaled by a proper factor α , the spectrum of the model remains approximately unchanged. Here, $L_A = 10$. Single disorder realization.

L_A faster than linearly. This means that for a larger system, a larger fraction of the system is correlated with the bath. Extrapolating this to the large system size limit, $L_A \gg 1$, almost the entire system is thermalized by a finite bath before reaching the Heisenberg time t_H . On the other hand, disorder values $W_A = 4, 6$, and 10 give a linear dependence of ξ_d on L_A and a constant fraction of the system stays localized for each setting, not being affected by the presence of the bath.

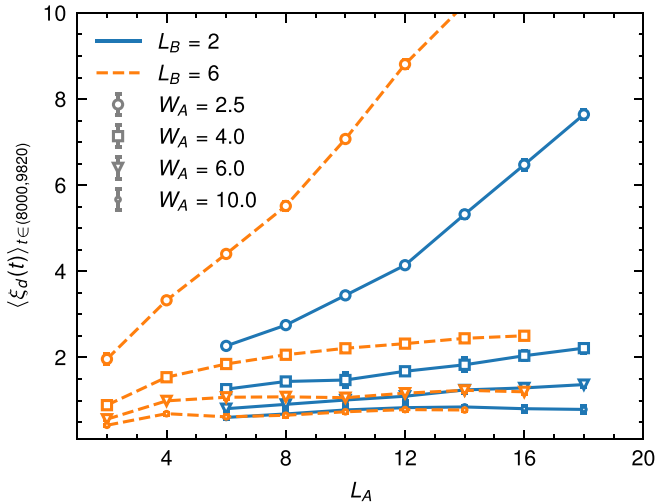


FIG. 12. Fitted bath correlation decay length ξ_d as a function of A system size L_A , averaged over time window $t \in (8000, 9820)$. Increasing the bath size from $L_B = 2$ to $L_B = 6$ enhances thermalization. However, if the growth is not faster than linear with L_A , the thermalized part of chain A will occupy at most a finite fraction of the chain A , and some localized sites far from the interface will not be affected by the bath. They may still be fully thermalized but at timescales exponential in system size L_A .

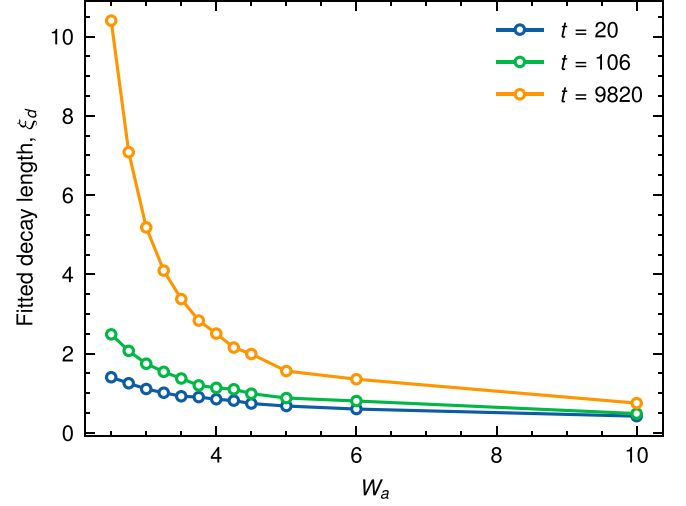


FIG. 13. Fitted bath correlation decay length ξ_d as a function of disorder strength W_A . $L_A = 14$, $L_B = 6$.

2. Fixed bath and L_A , changing W_A

Restricting ourselves to a smaller system size of $L_A = 14$, $L_B = 6$, we check how the late-time decay length depends on the disorder strength W_A . The result is presented in Fig. 13. This plot shows that the decay length decreases very fast between $W_A = 2.5$ to $W_A \approx 5$, and then stays approximately constant around value $\xi_d \approx 0.5$ for higher W_A . It suggests that MBL prevents the spread of avalanches at strong disorder, yet no sharp transition can be identified due to finite sizes of the systems of interest.

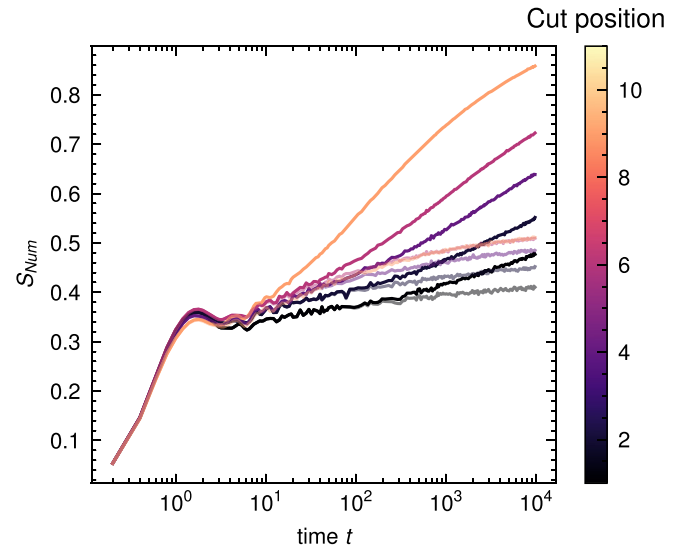


FIG. 14. Number entropy S_{Num} for $L_A = 12$, $L_B = 10$, $W_A = 4$, and $W_B = 0.5$ (solid lines) or $W_B = 4$ (transparent lines). Depending on the position of the cut, a speedup beyond $\propto \ln(t)$ is observed in the case $W_B = 0.5$ sooner or later, whereas the growth is slower than logarithmic when $W_B = 4$. This suggests that due to the presence of the bath, a thermal avalanche is traveling through the system. Thus conclusions from the analysis of the number entropy and correlation decay length are fully compatible.

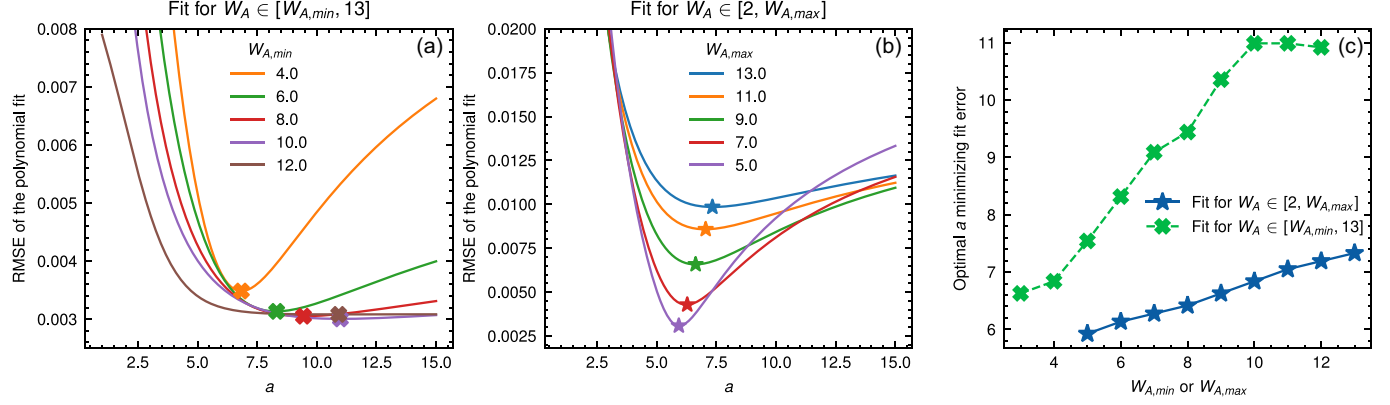


FIG. 15. Disordered XXZ model. [(a) and (b)] Root-mean-squared error (RMSE) of the 4-th degree polynomial fits to imbalance from Fig. 7 with rescaled time axis according to $t \rightarrow te^{-aW_A}$. Lowest values of the error, marked by crosses or stars, correspond to datapoints forming a universal curve due to rescaling time. Panel (a) shows fit errors when fitting the curve for the largest imbalances - interval of disorder amplitudes used for fitting is marked by the color, see legend. Minima are very shallow when fitting at large disorder values, signaling ambiguity of optimal a in that case. Panel (b) shows the RMSE of the fit limited to the data for smallest disorder values, down to $W_A = 2.0$. (c) Optimal a that describes universality shifts from around 6.5 to 11 when fitting to largest disorders. It stays around 6–7 when fitting to smallest disorder values, signaling the lack of a globally universal behavior for all disorder values. Fit was performed between times $t = 90$ to $t = 9820$ to not include initial oscillations of the imbalance, and for disorder values $W_A = 2, 3, 4, \dots, 13$. In all cases, $W_B = 0.5$.

APPENDIX F: COMPARISON WITH OTHER MEASURES—ENTANGLEMENT ENTROPY

In order to get a view on thermalization process from another point, we calculate the von Neuman entanglement entropy of the system. In particular, we are interested in the number entropy part, which in the spinless fermions description of the XXZ chain measures the uncertainty in the number of particles in the subsystem. The formula reads

$$\mathcal{S}_{num}(t) = - \sum_n p(n) \ln p(n), \quad (F1)$$

where $p(n)$ is the probability of total magnetization $n = \sum_{i \in \text{cut}} Z_i$ of a chosen subsystem of the whole system, and $\rho(n)$ is the block of the reduced density matrix in magnetization n sector of the Hilbert space.

In Fig. 14, we notice that in the case of no bath ($W_B = W_A = 4, L_A = 12, L_B = 10$), the number entropy grows with

time slower than $\ln t$. Bath with $W_B = 0.5$ causes a speedup in the number entropy growth. While the first speedup for cut between sites 11 and 12 can be observed around $t = 10$, the further the cut position is, the longer it takes for the speedup to be visible. Nevertheless, entanglement entropy for the furthest cut from the interface obtains a speedup beyond logarithmic growth when the bath is present. This can be interpreted as an avalanche reaching the boundary of the system. This conclusion stays in agreement with the result obtained for the same system by measuring simpler two-body correlation functions and their decay length, shown in Fig. 3(b) in the main paper.

APPENDIX G: UNIVERSAL SCALING OF TIME WITH DISORDER

Here we present details on the determination of optimal constant a in a universal time rescaling law $t \rightarrow te^{-aW_A}$ in the discussion of Sec. V.

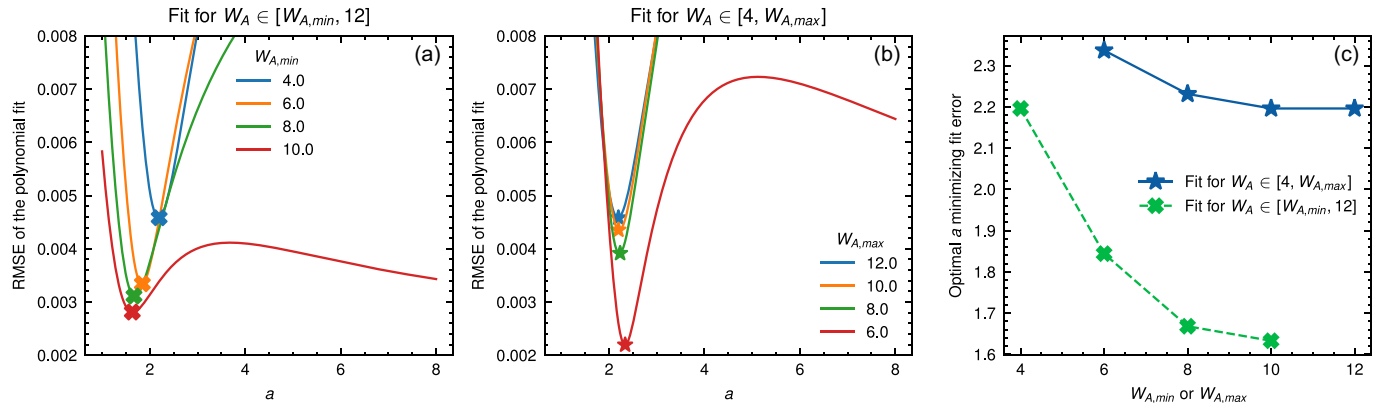


FIG. 16. Same as in Fig. 15 but for the kicked XXZ model. Minima in (a) and (b) are well defined and do not change significantly depending on the fitting interval of disorders W_A . (c) shows that an optimal a minimizing the RMSE has an inverse trend when compared to the time-independent case in Fig. 15. Nevertheless, its values stay consistently around 2, while in the time-independent case they drift from 6.5 to 11 with the change of fitting disorder interval. Here, $W_A = 4, 6, 8, 10, 12$, and $W_B = 0.5$.

To find a numerically, we fix the value of a , rescale time $t \rightarrow te^{-aW_A}$, and perform a fourth degree polynomial fit to the imbalance from Fig. 7 on the logarithmic time scale. The fit is performed either for a small interval of disorder values W_A , or for all data. Root-mean-squared error (RMSE) of this fit measures the quality of a as a potential constant explaining the trends in the data that can be captured by a continuous curve. We repeat the fitting for a range of different values of a , and the optimal a is the one with the minimal RMSE of the fit.

Results for fitting a to the data from our disordered XXZ model with a bath, described by Eq. (12) in the main text, are presented in Fig. 15. When the fit is performed for the strongest disorder amplitudes up to $W_A = 13$ [Fig. 15(a)], the optimal a has values around $a \approx 6.5$ –11, the RMSE minima shift significantly when changing the range of disorders used for fitting, and the minima are very shallow i.e. a is not properly defined. Fitting to the data corresponding to the weakest disorder amplitudes, down to $W_A = 2$, gives a different result: optimal a stays around 6–7.5 and the minima are deeper [see

Fig. 15(b)]. Optimal a for all considered ranges of disorder are summarized in Fig. 15(c): constant a drifts between different ranges of disorder amplitudes used for fitting, and the two regimes of large and small disorder values give inconsistent a . This means that the model lacks a universal scaling of time $t \rightarrow te^{-aW_A}$.

The same analysis, repeated for the kicked XXZ model of Peacock and Sels [104], confirms the presence of a single universal constant a that drifts much less with the disorder interval considered. Figures 16(a) and 16(b) show RMSE of the fourth-degree polynomial fit to the data which has large (small) disorder values. The minima are well defined in both cases, and the optimal a stays around 2 with an absolute error of around 0.3. Thus there exists a single a that is universal irrespective of the disorder strength, in full agreement with the statements of Ref. [104]. It is interesting, though, that all trends observed in this case in Fig. 16(c), are reverse to trends for the time-independent model, Fig. 16(c). At present, it is not clear to us why such a change in trend is present.

-
- [1] G. Gallavotti, Ergodicity, ensembles, irreversibility in boltzmann and beyond, *J. Stat. Phys.* **78**, 1571 (1995).
 - [2] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
 - [3] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
 - [4] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, *Nature (London)* **452**, 854 (2008).
 - [5] 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).
 - [6] L. Foini and J. Kurchan, Eigenstate thermalization hypothesis and out of time order correlators, *Phys. Rev. E* **99**, 042139 (2019).
 - [7] S. Pappalardi, L. Foini, and J. Kurchan, Eigenstate thermalization hypothesis and free probability, *Phys. Rev. Lett.* **129**, 170603 (2022).
 - [8] S. Pappalardi, F. Fritzsche, and T. Prosen, General eigenstate thermalization via free cumulants in quantum lattice systems, *arXiv:2303.00713*.
 - [9] S. Pappalardi, L. Foini, and J. Kurchan, Microcanonical windows on quantum operators, *Quantum* **8**, 1227 (2024).
 - [10] M. Fava, J. Kurchan, and S. Pappalardi, Designs via free probability, *arXiv:2308.06200*.
 - [11] M. Rigol, Breakdown of thermalization in finite one-dimensional systems, *Phys. Rev. Lett.* **103**, 100403 (2009).
 - [12] L. F. Santos and M. Rigol, Onset of quantum chaos in one-dimensional bosonic and fermionic systems and its relation to thermalization, *Phys. Rev. E* **81**, 036206 (2010).
 - [13] R. Steinigeweg, J. Herbrych, and P. Prelovšek, Eigenstate thermalization within isolated spin-chain systems, *Phys. Rev. E* **87**, 012118 (2013).
 - [14] E. Khatami, G. Pupillo, M. Srednicki, and M. Rigol, Fluctuation-dissipation theorem in an isolated system of quantum dipolar bosons after a quench, *Phys. Rev. Lett.* **111**, 050403 (2013).
 - [15] W. Beugeling, R. Moessner, and M. Haque, Finite-size scaling of eigenstate thermalization, *Phys. Rev. E* **89**, 042112 (2014).
 - [16] C. Schönle, D. Jansen, F. Heidrich-Meisner, and L. Vidmar, Eigenstate thermalization hypothesis through the lens of auto-correlation functions, *Phys. Rev. B* **103**, 235137 (2021).
 - [17] 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).
 - [18] 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).
 - [19] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, *Phys. Rev. B* **75**, 155111 (2007).
 - [20] M. Žnidarič, T. Prosen, and P. Prelovšek, Many-body localization in the Heisenberg XXZ magnet in a random field, *Phys. Rev. B* **77**, 064426 (2008).
 - [21] A. Pal and D. A. Huse, Many-body localization phase transition, *Phys. Rev. B* **82**, 174411 (2010).
 - [22] A. De Luca and A. Scardicchio, Ergodicity breaking in a model showing many-body localization, *Europhys. Lett.* **101**, 37003 (2013).
 - [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] P. Sierant and J. Zakrzewski, Many-body localization of bosons in optical lattices, *New J. Phys.* **20**, 043032 (2018).
 - [26] T. Orell, A. A. Michailidis, M. Serbyn, and M. Silveri, Probing the many-body localization phase transition with superconducting circuits, *Phys. Rev. B* **100**, 134504 (2019).

- [27] M. Hopjan and F. Heidrich-Meisner, Many-body localization from a one-particle perspective in the disordered one-dimensional Bose-Hubbard model, *Phys. Rev. A* **101**, 063617 (2020).
- [28] R. Mondaini and M. Rigol, Many-body localization and thermalization in disordered Hubbard chains, *Phys. Rev. A* **92**, 041601(R) (2015).
- [29] P. Prelovšek, O. S. Barišić, and M. Žnidarič, Absence of full many-body localization in the disordered Hubbard chain, *Phys. Rev. B* **94**, 241104(R) (2016).
- [30] J. Zakrzewski and D. Delande, Spin-charge separation and many-body localization, *Phys. Rev. B* **98**, 014203 (2018).
- [31] M. Kozarzewski, P. Prelovšek, and M. Mierzejewski, Spin subdiffusion in the disordered Hubbard chain, *Phys. Rev. Lett.* **120**, 246602 (2018).
- [32] D. J. Luitz, N. Laflorencie, and F. Alet, Extended slow dynamical regime close to the many-body localization transition, *Phys. Rev. B* **93**, 060201(R) (2016).
- [33] S. Bera, G. De Tomasi, F. Weiner, and F. Evers, Density propagator for many-body localization: Finite-size effects, transient subdiffusion, and exponential decay, *Phys. Rev. Lett.* **118**, 196801 (2017).
- [34] F. Weiner, F. Evers, and S. Bera, Slow dynamics and strong finite-size effects in many-body localization with random and quasiperiodic potentials, *Phys. Rev. B* **100**, 104204 (2019).
- [35] D. Sels and A. Polkovnikov, Thermalization of dilute impurities in one-dimensional spin chains, *Phys. Rev. X* **13**, 011041 (2023).
- [36] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Observation of many-body localization of interacting fermions in a quasirandom optical lattice, *Science* **349**, 842 (2015).
- [37] 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).
- [38] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, Observation of slow dynamics near the many-body localization transition in one-dimensional quasiperiodic systems, *Phys. Rev. Lett.* **119**, 260401 (2017).
- [39] H. P. Lüschen, S. Scherg, T. Kohlert, M. Schreiber, P. Bordia, X. Li, S. Das Sarma, and I. Bloch, Single-particle mobility edge in a one-dimensional quasiperiodic optical lattice, *Phys. Rev. Lett.* **120**, 160404 (2018).
- [40] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Probing entanglement in a many-body-localized system, *Science* **364**, 256 (2019).
- [41] M. Rispoli, A. Lukin, R. Schittko, S. Kim, M. E. Tai, J. Léonard, and M. Greiner, Quantum critical behaviour at the many-body localization transition, *Nature (London)* **573**, 385 (2019).
- [42] 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).
- [43] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, B. Foxen, M. Giustina, E. Jeffrey, J. Kelly, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank *et al.*, Spectroscopic signatures of localization with interacting photons in superconducting qubits, *Science* **358**, 1175 (2017).
- [44] A. Y. Guo, M. C. Tran, A. M. Childs, A. V. Gorshkov, and Z.-X. Gong, Signaling and scrambling with strongly long-range interactions, *Phys. Rev. A* **102**, 010401(R) (2020).
- [45] Q. Guo, C. Cheng, H. Li, S. Xu, P. Zhang, Z. Wang, C. Song, W. Liu, W. Ren, H. Dong, R. Mondaini, and H. Wang, Stark many-body localization on a superconducting quantum processor, *Phys. Rev. Lett.* **127**, 240502 (2021).
- [46] W. Morong, F. Liu, P. Becker, K. S. Collins, L. Feng, A. Kyprianidis, G. Pagano, T. You, A. V. Gorshkov, and C. Monroe, Observation of Stark many-body localization without disorder, *Nature (London)* **599**, 393 (2021).
- [47] R. Nandkishore and D. A. Huse, Many-body-localization and thermalization in quantum statistical mechanics, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
- [48] F. Alet and N. Laflorencie, Many-body localization: An introduction and selected topics, *C. R. Phys.* **19**, 498 (2018).
- [49] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [50] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phenomenology of fully many-body-localized systems, *Phys. Rev. B* **90**, 174202 (2014).
- [51] V. Ros, M. Mueller, and A. Scardicchio, Integrals of motion in the many-body localized phase, *Nucl. Phys. B* **891**, 420 (2015).
- [52] 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).
- [53] J. Z. Imbrie, Diagonalization and many-body localization for a disordered quantum spin chain, *Phys. Rev. Lett.* **117**, 027201 (2016).
- [54] T. B. Wahl, A. Pal, and S. H. Simon, Efficient representation of fully many-body localized systems using tensor networks, *Phys. Rev. X* **7**, 021018 (2017).
- [55] M. Mierzejewski, M. Kozarzewski, and P. Prelovšek, Counting local integrals of motion in disordered spinless-fermion and Hubbard chains, *Phys. Rev. B* **97**, 064204 (2018).
- [56] S. J. Thomson and M. Schiró, Time evolution of many-body localized systems with the flow equation approach, *Phys. Rev. B* **97**, 060201(R) (2018).
- [57] M. Žnidarič, A. Scardicchio, and V. K. Varma, Diffusive and subdiffusive spin transport in the ergodic phase of a many-body localizable system, *Phys. Rev. Lett.* **117**, 040601 (2016).
- [58] 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).
- [59] M. Serbyn, Z. Papić, and D. A. Abanin, Universal slow growth of entanglement in interacting strongly disordered systems, *Phys. Rev. Lett.* **110**, 260601 (2013).
- [60] F. Iemini, A. Russomanno, D. Rossini, A. Scardicchio, and R. Fazio, Signatures of many-body localization in the dynamics of two-site entanglement, *Phys. Rev. B* **94**, 214206 (2016).

- [61] E. V. H. Doggen, F. Schindler, K. S. Tikhonov, A. D. Mirlin, T. Neupert, D. G. Polyakov, and I. V. Gornyi, Many-body localization and delocalization in large quantum chains, *Phys. Rev. B* **98**, 174202 (2018).
- [62] T. Chanda, P. Sierant, and J. Zakrzewski, Time dynamics with matrix product states: Many-body localization transition of large systems revisited, *Phys. Rev. B* **101**, 035148 (2020).
- [63] E. V. H. Doggen and A. D. Mirlin, Many-body delocalization dynamics in long aubry-andré quasiperiodic chains, *Phys. Rev. B* **100**, 104203 (2019).
- [64] T. Chanda, P. Sierant, and J. Zakrzewski, Many-body localization transition in large quantum spin chains: The mobility edge, *Phys. Rev. Res.* **2**, 032045(R) (2020).
- [65] R. K. Panda, A. Scardicchio, M. Schulz, S. R. Taylor, and M. Žnidarič, Can we study the many-body localisation transition?, *Europhys. Lett.* **128**, 67003 (2020).
- [66] P. Sierant, D. Delande, and J. Zakrzewski, Thouless time analysis of anderson and many-body localization transitions, *Phys. Rev. Lett.* **124**, 186601 (2020).
- [67] P. Sierant, M. Lewenstein, and J. Zakrzewski, Polynomially filtered exact diagonalization approach to many-body localization, *Phys. Rev. Lett.* **125**, 156601 (2020).
- [68] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Evidence for unbounded growth of the number entropy in many-body localized phases, *Phys. Rev. Lett.* **124**, 243601 (2020).
- [69] 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).
- [70] 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.* **427**, 168415 (2021).
- [71] D. Sels and A. Polkovnikov, Dynamical obstruction to localization in a disordered spin chain, *Phys. Rev. E* **104**, 054105 (2021).
- [72] B. Krajewski, L. Vidmar, J. Bonča, and M. Mierzejewski, Restoring ergodicity in a strongly disordered interacting chain, *Phys. Rev. Lett.* **129**, 260601 (2022).
- [73] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, *Phys. Rev. E* **102**, 062144 (2020).
- [74] P. Sierant and J. Zakrzewski, Challenges to observation of many-body localization, *Phys. Rev. B* **105**, 224203 (2022).
- [75] A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, Avalanches and many-body resonances in many-body localized systems, *Phys. Rev. B* **105**, 174205 (2022).
- [76] P. Sierant, E. G. Lazo, M. Dalmonte, A. Scardicchio, and J. Zakrzewski, Constraint-induced delocalization, *Phys. Rev. Lett.* **127**, 126603 (2021).
- [77] W. De Roeck, F. Huveneers, B. Meeus, and O. A. Prośniak, Rigorous and simple results on very slow thermalization, or quasi-localization, of the disordered quantum chain, *Physica A* **631**, 129245 (2023).
- [78] P. Sierant, M. Lewenstein, A. Scardicchio, and J. Zakrzewski, Stability of many-body localization in floquet systems, *Phys. Rev. B* **107**, 115132 (2023).
- [79] B. Krajewski, L. Vidmar, J. Bonča, and M. Mierzejewski, Strongly disordered anderson insulator chains with generic two-body interaction, *Phys. Rev. B* **108**, 064203 (2023).
- [80] T. Vojta, Quantum Griffiths effects and smeared phase transitions in metals: Theory and experiment, *J. Low Temp. Phys.* **161**, 299 (2010).
- [81] S. Gopalakrishnan, K. Agarwal, E. A. Demler, D. A. Huse, and M. Knap, Griffiths effects and slow dynamics in nearly many-body localized systems, *Phys. Rev. B* **93**, 134206 (2016).
- [82] K. Agarwal, S. Gopalakrishnan, M. Knap, M. Müller, and E. Demler, Anomalous diffusion and griffiths effects near the many-body localization transition, *Phys. Rev. Lett.* **114**, 160401 (2015).
- [83] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, Rare-region effects and dynamics near the many-body localization transition, *Ann. Phys.* **529**, 1600326 (2017).
- [84] N. Pancotti, M. Knap, D. A. Huse, J. I. Cirac, and M. C. Bañuls, Almost conserved operators in nearly many-body localized systems, *Phys. Rev. B* **97**, 094206 (2018).
- [85] W. De Roeck and F. Huveneers, Stability and instability towards delocalization in many-body localization systems, *Phys. Rev. B* **95**, 155129 (2017).
- [86] D. J. Luitz, F. Huveneers, and W. De Roeck, How a small quantum bath can thermalize long localized chains, *Phys. Rev. Lett.* **119**, 150602 (2017).
- [87] J. Šuntajs and L. Vidmar, Ergodicity breaking transition in zero dimensions, *Phys. Rev. Lett.* **129**, 060602 (2022).
- [88] J. Šuntajs, M. Hopjan, W. D. Roeck, and L. Vidmar, Similarity between a many-body quantum avalanche model and the ultrametric random matrix model, [arXiv:2308.07431](https://arxiv.org/abs/2308.07431).
- [89] K. Pawlik, P. Sierant, L. Vidmar, and J. Zakrzewski, Many-body mobility edge in quantum spin models, [arXiv:2308.01073](https://arxiv.org/abs/2308.01073).
- [90] T. Thiery, F. m. c. Huveneers, M. Müller, and W. De Roeck, Many-body delocalization as a quantum avalanche, *Phys. Rev. Lett.* **121**, 140601 (2018).
- [91] A. Goremykina, R. Vasseur, and M. Serbyn, Analytically solvable renormalization group for the many-body localization transition, *Phys. Rev. Lett.* **122**, 040601 (2019).
- [92] P. T. Dumitrescu, A. Goremykina, S. A. Parameswaran, M. Serbyn, and R. Vasseur, Kosterlitz-thouless scaling at many-body localization phase transitions, *Phys. Rev. B* **99**, 094205 (2019).
- [93] A. Morningstar and D. A. Huse, Renormalization-group study of the many-body localization transition in one dimension, *Phys. Rev. B* **99**, 224205 (2019).
- [94] I.-D. Potirniche, S. Banerjee, and E. Altman, Exploration of the stability of many-body localization in $d > 1$, *Phys. Rev. B* **99**, 205149 (2019).
- [95] L. Herviou, S. Bera, and J. H. Bardarson, Multiscale entanglement clusters at the many-body localization phase transition, *Phys. Rev. B* **99**, 134205 (2019).
- [96] T. Szödra, P. Sierant, K. Kottmann, M. Lewenstein, and J. Zakrzewski, Detecting ergodic bubbles at the crossover to many-body localization using neural networks, *Phys. Rev. B* **104**, L140202 (2021).

- [97] K. Hémerly, F. Pollmann, and A. Smith, Identifying correlation clusters in many-body localized systems, *Phys. Rev. B* **105**, 064202 (2022).
- [98] D. C. W. Foo, N. Swain, P. Sengupta, G. Lemarié, and S. Adam, Stabilization mechanism for many-body localization in two dimensions, *Phys. Rev. Res.* **5**, L032011 (2023).
- [99] L. Colmenarez, D. J. Luitz, and W. D. Roeck, Ergodic inclusions in many-body localized systems, *Phys. Rev. B* **109**, L081117 (2024).
- [100] D. Sels, Bath-induced delocalization in interacting disordered spin chains, *Phys. Rev. B* **106**, L020202 (2022).
- [101] H. Ha, A. Morningstar, and D. A. Huse, Many-body resonances in the avalanche instability of many-body localization, *Phys. Rev. Lett.* **130**, 250405 (2023).
- [102] Y.-T. Tu, D. Vu, and S. Das Sarma, Avalanche stability transition in interacting quasiperiodic systems, *Phys. Rev. B* **107**, 014203 (2023).
- [103] J. Léonard, S. Kim, M. Rispoli, A. Lukin, R. Schittko, J. Kwan, E. Demler, D. Sels, and M. Greiner, Probing the onset of quantum avalanches in a many-body localized system, *Nat. Phys.* **19**, 481 (2023).
- [104] J. C. Peacock and D. Sels, Many-body delocalization from embedded thermal inclusion, *Phys. Rev. B* **108**, L020201 (2023).
- [105] Y.-T. Tu, D. Vu, and S. Das Sarma, Localization spectrum of a bath-coupled generalized aubry-andré model in the presence of interactions, *Phys. Rev. B* **108**, 064313 (2023).
- [106] M. Srednicki, The approach to thermal equilibrium in quantized chaotic systems, *J. Phys. A: Math. Gen.* **32**, 1163 (1999).
- [107] P. C. Burke and M. Haque, Entropy and temperature in finite isolated quantum systems, *Phys. Rev. E* **107**, 034125 (2023).
- [108] C. C. Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics: Vol. Two* (John Wiley and Sons, 2002).
- [109] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, Constructing local integrals of motion in the many-body localized phase, *Phys. Rev. B* **91**, 085425 (2015).
- [110] L. Rademaker and M. Ortuño, Explicit local integrals of motion for the many-body localized state, *Phys. Rev. Lett.* **116**, 010404 (2016).
- [111] M. Ortuño, A. M. Somoza, and L. Rademaker, Construction of many-body eigenstates with displacement transformations in disordered systems, *Phys. Rev. B* **100**, 085115 (2019).
- [112] D. Pekker, B. K. Clark, V. Oganesyan, and G. Refael, Fixed points of wegner-wilson flows and many-body localization, *Phys. Rev. Lett.* **119**, 075701 (2017).
- [113] S. P. Kelly, R. Nandkishore, and J. Marino, Exploring many-body localization in quantum systems coupled to an environment via wegner-wilson flows, *Nucl. Phys. B* **951**, 114886 (2020).
- [114] S. J. Thomson, D. Magano, and M. Schirò, Flow equations for disordered Floquet systems, *SciPost Phys.* **11**, 028 (2021).
- [115] S. J. Thomson and M. Schirò, Local integrals of motion in quasiperiodic many-body localized systems, *SciPost Phys.* **14**, 125 (2023).
- [116] T. E. O'Brien, D. A. Abanin, G. Vidal, and Z. Papić, Explicit construction of local conserved operators in disordered many-body systems, *Phys. Rev. B* **94**, 144208 (2016).
- [117] P. Peng, Z. Li, H. Yan, K. X. Wei, and P. Cappellaro, Comparing many-body localization lengths via nonperturbative construction of local integrals of motion, *Phys. Rev. B* **100**, 214203 (2019).
- [118] B. Leipner-Johns and R. Wortis, Charge- and spin-specific local integrals of motion in a disordered hubbard model, *Phys. Rev. B* **100**, 125132 (2019).
- [119] S. Adami, M. Amini, and M. Soltani, Structural properties of local integrals of motion across the many-body localization transition via a fast and efficient method for their construction, *Phys. Rev. B* **106**, 054202 (2022).
- [120] T. Szóldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, Unsupervised detection of decoupled subspaces: Many-body scars and beyond, *Phys. Rev. B* **105**, 224205 (2022).
- [121] T. C. Berkelbach and D. R. Reichman, Conductivity of disordered quantum lattice models at infinite temperature: Many-body localization, *Phys. Rev. B* **81**, 224429 (2010).
- [122] S. Bera, H. Schomerus, F. Heidrich-Meisner, and J. H. Bardarson, Many-body localization characterized from a one-particle perspective, *Phys. Rev. Lett.* **115**, 046603 (2015).
- [123] T. Enss, F. Andraschko, and J. Sirker, Many-body localization in infinite chains, *Phys. Rev. B* **95**, 045121 (2017).
- [124] M. Serbyn and J. E. Moore, Spectral statistics across the many-body localization transition, *Phys. Rev. B* **93**, 041424(R) (2016).
- [125] C. L. Bertrand and A. M. García-García, Anomalous thouless energy and critical statistics on the metallic side of the many-body localization transition, *Phys. Rev. B* **94**, 144201 (2016).
- [126] L. F. Santos, G. Rigolin, and C. O. Escobar, Entanglement versus chaos in disordered spin chains, *Phys. Rev. A* **69**, 042304 (2004).
- [127] L. F. Santos, Integrability of a disordered heisenberg spin-1/2 chain, *J. Phys. A: Math. Gen.* **37**, 4723 (2004).
- [128] L. Colmenarez, P. A. McClarty, M. Haque, and D. J. Luitz, Statistics of correlation functions in the random Heisenberg chain, *SciPost Phys.* **7**, 064 (2019).
- [129] P. Sierant and J. Zakrzewski, Level statistics across the many-body localization transition, *Phys. Rev. B* **99**, 104205 (2019).
- [130] P. Sierant and J. Zakrzewski, Model of level statistics for disordered interacting quantum many-body systems, *Phys. Rev. B* **101**, 104201 (2020).
- [131] M. Schiulaz, E. J. Torres-Herrera, F. Pérez-Bernal, and L. F. Santos, Self-averaging in many-body quantum systems out of equilibrium: Chaotic systems, *Phys. Rev. B* **101**, 174312 (2020).
- [132] E. J. Torres-Herrera, G. De Tomasi, M. Schiulaz, F. Pérez-Bernal, and L. F. Santos, Self-averaging in many-body quantum systems out of equilibrium: Approach to the localized phase, *Phys. Rev. B* **102**, 094310 (2020).
- [133] T. Szóldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, Tracking locality in the time evolution of disordered systems, *Phys. Rev. B* **107**, 054204 (2023).
- [134] N. Macé, F. Alet, and N. Laflorencie, Multifractal scalings across the many-body localization transition, *Phys. Rev. Lett.* **123**, 180601 (2019).
- [135] N. Laflorencie, G. Lemarié, and N. Macé, Chain breaking and kosterlitz-thouless scaling at the many-body localization transition in the random-field heisenberg spin chain, *Phys. Rev. Res.* **2**, 042033(R) (2020).

- [136] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Ergodicity breaking transition in finite disordered spin chains, *Phys. Rev. B* **102**, 064207 (2020).
- [137] J. Gray, S. Bose, and A. Bayat, Many-body localization transition: Schmidt gap, entanglement length, and scaling, *Phys. Rev. B* **97**, 201105(R) (2018).
- [138] T. Devakul and R. R. P. Singh, Early breakdown of area-law entanglement at the many-body delocalization transition, *Phys. Rev. Lett.* **115**, 187201 (2015).
- [139] N. Macé, N. Laflorencie, and F. Alet, Many-body localization in a quasiperiodic Fibonacci chain, *SciPost Phys.* **6**, 50 (2019).
- [140] H. Tal-Ezer and R. Kosloff, An accurate and efficient scheme for propagating the time dependent schrödinger equation, *J. Chem. Phys.* **81**, 3967 (1984).
- [141] H. Fehske and R. Schneider, *Computational Many-Particle Physics* (Springer, Germany, 2008).
- [142] J. J. Pulikkottil, A. Lakshminarayan, S. C. L. Srivastava, M. F. I. Kieler, A. Bäcker, and S. Tomsovic, Quantum coherence controls the nature of equilibration and thermalization in coupled chaotic systems, *Phys. Rev. E* **107**, 024124 (2023).
- [143] P. J. D. Crowley and A. Chandran, Avalanche induced coexisting localized and thermal regions in disordered chains, *Phys. Rev. Res.* **2**, 033262 (2020).
- [144] P. J. D. Crowley and A. Chandran, A constructive theory of the numerically accessible many-body localized to thermal crossover, *SciPost Phys.* **12**, 201 (2022).
- [145] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. Math.* **62**, 548 (1955).
- [146] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. Math.* **67**, 325 (1958).
- [147] Y. Sekino and L. Susskind, Fast scramblers, *J. High Energy Phys.* **10** (2008) 065.
- [148] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, *J. High Energy Phys.* **11** (2017) 048.
- [149] R. Nandkishore, Many-body localization proximity effect, *Phys. Rev. B* **92**, 245141 (2015).
- [150] F. Haake, *Quantum Signatures of Chaos* (Springer, Berlin, 2010).
- [151] D. Abanin, W. De Roeck, W. W. Ho, and F. Huveneers, A rigorous theory of many-body prethermalization for periodically driven and closed quantum systems, *Commun. Math. Phys.* **354**, 809 (2017).
- [152] S. Liu, S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao, Probing many-body localization by excited-state variational quantum eigensolver, *Phys. Rev. B* **107**, 024204 (2023).
- [153] <https://chaos.if.uj.edu.pl/ZOA/opendata/>.

Tracking locality in the time evolution of disordered systems

Tomasz Szołdra^{1,2}, Piotr Sierant³, Maciej Lewenstein^{3,4} and Jakub Zakrzewski^{2,5}¹*Szkoła Doktorska Nauk Ścisłych i Przyrodniczych, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*²*Instytut Fizyki Teoretycznej, Wydział Fizyki, Astronomii i Informatyki Stosowanej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*³*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*⁴*ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain*⁵*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, PL-30-348 Kraków, Poland*

(Received 28 November 2022; accepted 9 February 2023; published 21 February 2023)

Using local density correlation functions for a one-dimensional spin system, we introduce a correlation function difference (CFD) which compares correlations on a given site between a full system of size L and its restriction to $\ell < L$ sites. We show that CFD provides useful information on transfer of information in quantum many-body systems by considering the examples of ergodic, Anderson, and many-body localized regimes in a disordered XXZ spin chain. In the ergodic phase, we find that the propagation of CFD is asymptotically faster than the spin transport but slower than the ballistic propagation implied by the Lieb-Robinson bound. In contrast, in the localized cases, we unravel an exponentially slow relaxation of CFD. Connections between CFD and other observables detecting nonlocal correlations in the system are discussed.

DOI: [10.1103/PhysRevB.107.054204](https://doi.org/10.1103/PhysRevB.107.054204)

I. INTRODUCTION

The understanding of dynamics of interacting quantum many-body systems has been a subject of extensive research over recent years. The eigenstate thermalization hypothesis [1–4] anticipates that such systems are spectrally faithful to random matrix theory predictions, and when prepared in out-of-equilibrium states, they evolve in time to an equilibrium state dependent only on a few global constants of motion but independent of the details of the initial state. One exception to this ergodic paradigm is provided by a class of integrable models [5]. Similarly, ergodicity may be broken in the presence of a sufficiently strong disorder due to the phenomenon of many-body localization (MBL) [6–8]. MBL is characterized by a simultaneous occurrence of various features such as the emergent integrability of the system (with the existence of a complete set of local integrals of motion [9,10]) resulting in Poissonian spectral statistics [11–17], area-law entangled eigenstates [18,19], logarithmic growth of the entanglement entropy in time dynamics of initially low-entangled states [20–22], and a long memory of the initial state in the dynamics [23–26] (for recent reviews see [27,28]). Interestingly, the last years have brought new developments to our understanding of nonequilibrium quantum many-body physics. On one hand, nonergodic dynamics has been observed in disorder-free systems such as tilted lattices [29–38], in models with global constraints [39–45], or, notably, in implementations of lattice gauge theories [46–52]. On the other hand, following a seminal work [53], an intensive debate developed concerning the very existence of MBL in the thermodynamic limit [54–60]. While it seems that there exist models showing a genuine transition to MBL phase that survives the thermodynamic limit [61,62], the paradigmatic model for MBL, i.e.,

the disordered Heisenberg chain, still defies a definite answer. Some recent works [58,63] suggest that a stable MBL phase occurs in this model only at a very strong disorder shifting the previously considered border between ergodic and MBL phases [57,64] by an almost one order of magnitude in disorder amplitude. Other studies negate altogether the existence of MBL highlighting a presence of a persistent (but extremely slow) growth of the classical part of the entanglement entropy, the so-called number entropy [65,66] (see, however, [67] for opposite interpretations of those results). While often sophisticated measures (like an entanglement entropy or a spectral form factor) are considered, it seems worthwhile to focus also on observables more easily accessible experimentally, such as the imbalance addressed in the first experiments showing signatures of MBL [68,69] or the corresponding two-point density correlation functions. Still, even for those simpler observables, it can be shown that no conclusive answer about the stability of the MBL phase can be obtained on timescales offered by present day experiments as well as state-of-the-art simulations using tensor networks [70].

The aim of this work is to introduce an observable that allows one to probe the locality of dynamics in quantum many-body systems that would be experimentally more accessible than the out-of-time-order correlator (OTOC) function, conventionally employed in studies of thermalization and quantum information scrambling [71–76] (see [77–81] for studies of disordered systems) or spin echo and double electron-electron resonances (DEER) employed in [82]. To that end, we consider a 1D lattice system comprising L sites with open boundary conditions (OBCs), and we examine two settings: the first, in which we calculate the two times density correlation function $C_d(t)$ in a given site (later called simply the density correlation function for the full system of L sites);

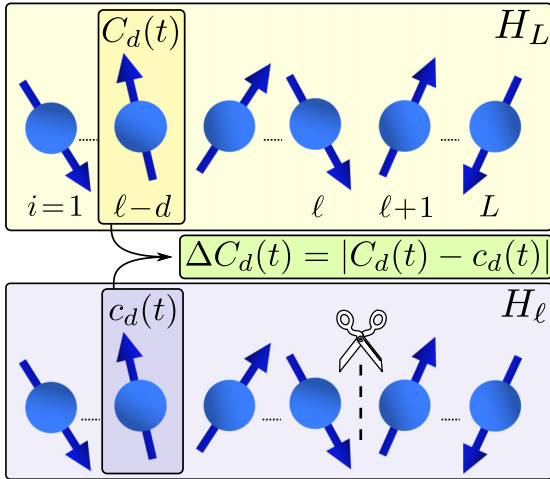


FIG. 1. Scheme for tracking locality of quantum dynamics in a spin lattice. We compare the time evolution of the single-site correlation function $C_d(t)$ at site $i = l - d$ in a system comprising L sites (upper panel) with $c_d(t)$ at the same site but in a system of $\ell < L$ sites (bottom panel).

the second in which we compute the similar density correlation function $c_d(t)$ for a system restricted to $\ell < L$ sites. Our setup is illustrated in Fig. 1. The observable of our interest is a correlation function difference (CFD), $\Delta C_d(t)$, calculated as

$$\Delta C_d(t) = |C_d(t) - c_d(t)|; \quad (1)$$

see Sec. II for a precise definition of $C_d(t)$, $c_d(t)$. The CFD reflects changes in the time evolution of density correlation functions that occur when the system size is increased from ℓ to L . The realization of random disorder on the first ℓ sites stays the same. This quantifies the locality of quantum dynamics: if the dynamics at site d is not affected by degrees of freedom residing on sites separated from d by a certain length scale ξ , we expect a vanishing CFD $\Delta C_d(t) \approx 0$ as long as $l - d > \xi$. In contrast, a nonzero CFD indicates a nontrivial impact of the degrees of freedom at sites $i > \ell$ on time evolution at site d .

We note that CFD is particularly relevant for experiments in which $C_d(t)$ is directly measurable [83,84] as it requires a density measurement with a single-site resolution. Moreover, the change of the system size from ℓ to L in Fig. 1 can be alternatively interpreted as an introduction of a local tunneling term that couples sites $i < \ell$ and $i \geq \ell$ to the Hamiltonian for L sites. This provides a conceptual link between CFD and earlier studies of deformation of many-body eigenstates due to local perturbations [85–88]. A similar scenario was considered also in the context of local quench of a 1D quantum spin chain in which two parts of the chain are joined [89–91] or split [92]. In that case, the dynamics and the relevant time scales are dictated by the physics of low-energy excitations (in particular, by the underlying conformal field theory for critical spin chains). In contrast, in this work we probe dynamics of highly out-of-equilibrium initial states.

Section II contains the definition of disordered Heisenberg model on which we benchmark CFD. In Sec. III we study the dynamics of the model for three distinct regimes: ergodic phase with delocalized eigenstates, Anderson localized

(noninteracting) model and MBL regime. We show that CFD behaves differently in each of the three regimes, reflecting the various degrees of locality of dynamics in the system. We also relate our observable to the number entropy which measures the fluctuations of the number of particles in different parts of the system. Finally, we conclude in Sec. V.

II. THE MODEL

To study characteristic properties of CFD in different dynamical phases of matter, we consider the disordered XXZ spin chain, with the Hamiltonian

$$H_L = \sum_{i=1}^{L-1} J(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y) + \Delta S_i^z S_{i+1}^z + \sum_{i=1}^L h_i S_i^z, \quad (2)$$

where S_i^k , $k = x, y, z$ are spin-1/2 operators at site i and the on-site random magnetic field h_i is uniformly distributed in the $[-W, W]$ interval where W is the disorder amplitude. Further on, we take $J = 1$ as an energy unit. The model (2) maps to spinless interacting fermions via a Jordan-Wigner transformation where J corresponds to tunneling amplitude between sites while Δ gives the strength of nearest-neighbor interactions.

We shall employ CFD to investigate the process of quantum information transfer in this system and compare the three regimes: ergodic phase ($\Delta = 1$, $W < W_c$), Anderson-localized phase ($\Delta = 0$, $W > 0$), and the MBL regime ($\Delta = 1$, $W > W_c$), where W_c is the critical disorder strength at which, at relatively small system sizes, the crossover between delocalized and MBL regimes occurs in an interacting system. At the system sizes considered, $L \approx 20$, the crossover is located around $W_c \approx 3.5$ [57,64]. We study time dynamics from an initial Néel state, a separable state with every second spin up and every second down (we limit the analysis to a conserved total magnetization sector $S_z = 0$ for even L) corresponding to every second site occupied and every second site empty in the fermionic language. While the early experiments [68] considered a global observable, i.e., a staggered magnetization (in other words the imbalance between odd and even sites) over a whole lattice, we consider the single-site correlation functions defined as $C_i(t) = \overline{C_i^s(t)}$, where the overline denotes the disorder average and $C_i^s(t) = \langle S_i^z(t) S_i^z(0) \rangle$ is the quantum mechanical expectation value for the i th spin at time t . The standard imbalance is proportional to a sum of all $C_i(t)$. Here we use the single-site correlation functions to calculate the CFD according to Eq. (1). To evaluate $c_i(t)$, we modify the Hamiltonian (2) by removing the link between sites $i = \ell$, $\ell + 1$, i.e., by removing the term $J(S_\ell^x S_{\ell+1}^x + S_\ell^y S_{\ell+1}^y) + \Delta S_\ell^z S_{\ell+1}^z$. Such a cut may be realized (probably with some experimental difficulty) using, for example, the technique of subwavelength optical barriers [93]. Alternatively, projection of site-resolved optical potentials [94] in conjunction with optical box potentials [95] can be used to separately investigate dynamics of the system comprising ℓ and L sites. Similarly, CFD may be studied experimentally in other quantum simulators such as ion chains or Rydberg atom arrays; for a review see [96].

III. CFD IN VARIOUS DYNAMICAL REGIMES

A. Ergodic phase

We start our investigation of CFD by considering the ergodic phase of the disordered XXZ spin chain, setting $W = 1$ and $\Delta = 1$ in (2). We note that, in the fermionic interpretation of the model, the particle transport in the ergodic phase of the XXZ spin chain is, depending on the value of disorder strength W , diffusive for $0 < W \lesssim 0.55$ (for $\Delta = 1$) or subdiffusive for larger disorder strengths $W < W_c$ [97–99]. Moreover, for a short-range Hamiltonian we study, one expects that the velocity of information spreading in the system is bounded by the Lieb-Robinson velocity [100,101], which implies that local excitations propagate within a causal light cone.

With those two points as guiding principles, we compute the CFD, ΔC_d , for the XXZ spin chain at $W = 1$ and $\Delta = 1$ by calculating the single-site correlation functions $c_i(t)$ and $C_i(t)$ for system sizes $l = 26$ and $L = 28$, respectively, using the Chebyshev expansion of the time evolution operator [102]. The resulting CFD, shown in Fig. 2(a), becomes nonzero at times that increase with the distance d from the cut. This is in accordance with the hypothesis that CFD probes correlations in the system that become increasingly nonlocal during the course of the time evolution. Importantly, the initial increase of ΔC_d is followed by a damping of CFD at later times. The latter behavior does not reflect changes in the nonlocal content of the dynamics of the system, but rather is caused by the decrease of temporal fluctuations of $C_d(t)$ and $c_d(t)$ due to presence of interactions [103,104].

In order to isolate this phenomenon, we quantify the temporal fluctuations of $C_d(t)$ [$c_d(t)$] by calculating a rolling standard deviation

$$\sigma(C_d) = \overline{[C_d^s(t) - \langle C_d^s(t) \rangle_{\Delta t}]^2}_{\Delta t}, \quad (3)$$

where the overline denotes the disorder average, and $\langle \cdot \rangle_{\Delta t}$ is an average over time window of some small size Δt around time t (we take $\Delta t = 5$; as checked the results are similar for $\Delta t = 1, 10$). The inset in Fig. 2(a) shows that the decay of $\sigma(C_d)$ is approximately described by a power law in the interval of times considered. This decay of the amplitude of temporal oscillations of the single-site correlation functions is the main factor responsible for the late-time decrease of CFD. We shall discuss in details the mechanism of this effect in the following section. To compensate for this effect we inspect the dimensionless quantity $\Delta C_d / \sigma(C_d(t))$, which we refer to as the rescaled CFD.

Figure 2(b) shows the spatiotemporal dependence of the rescaled CFD in the ergodic phase of the disordered XXZ spin chain. In particular, by considering the curves at which $\Delta C_d / \sigma(C_d(t)) = \text{const}$ we find that a given value of the rescaled CFD is reached after a time scaling as a power law, $t \propto d^\alpha$, with the distance from the cut and $1 < \alpha < 2$. The power α (dependent on the isovalue curve as denoted in the figure) is consistently greater than unity for small and medium d indicating that the information about the cut revealed by the rescaled CFD spreads superdiffusively. Thus, CFD is not related to a particle (spin) transport, which is subdiffusive for the parameters used [97–99]. This is quite interesting as CFD contains density correlations which are manifestly linked with the transport. Therefore, our results suggest that

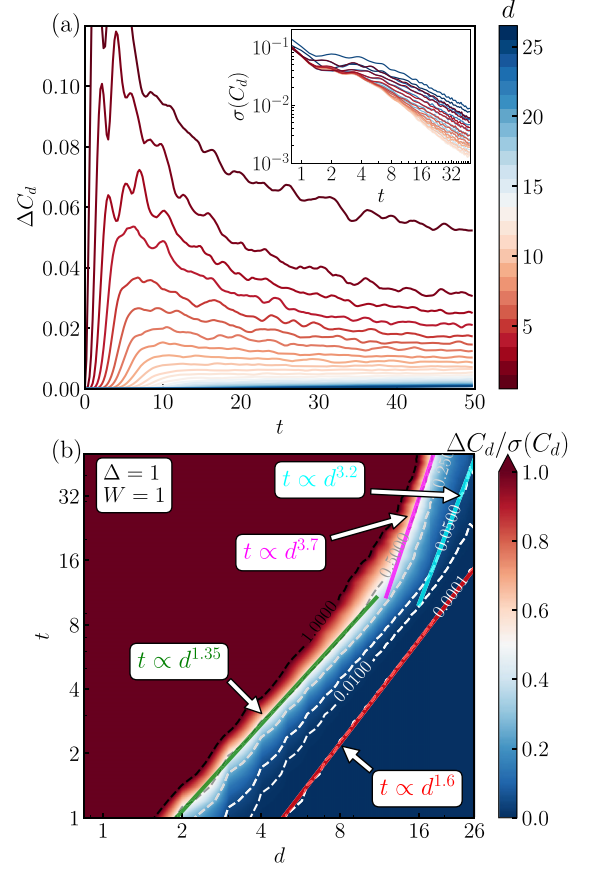


FIG. 2. The CFD, ΔC_d , in the ergodic phase of the XXZ chain with disorder strength $W = 1$, length $L = 28$, and a cut at $\ell = 26$. (a) Time dependence of ΔC_d for various distances d from the cut. The inset presents the rolling standard deviation $\sigma(C_d)$, Eq. (3), that quantifies temporal fluctuations of the single-site correlation function $C_d(t)$. (b) The rescaled CFD, $\Delta C_d / \sigma(C_d)$, as function of the distance from the cut and time t in the log-log scale; the white dashed lines denote isovalues of $\Delta C_d / \sigma(C_d)$, numbers indicate the values for different lines. Those almost straight lines in the log-log plot are thus well fitted piecewise by a power law $t \propto d^\alpha$ with values of the exponent indicated. Averages are taken over 1000 disorder realizations.

CFD captures effects of quantum information spreading in the chain that occurs faster in time than the particle transport but still slower than the ballistic dependence associated with the Lieb-Robinson bound. This implies that changes of $C_d(t)$ with respect to $c_d(t)$ that contribute to CFD are not due to a direct transfer of particle density n_d ¹ from the cut but rather are due to changes in details of dynamics of particles at distance d from the cut. Such local changes of dynamics are associated with accumulated effect of interactions of many particles between the cut and the considered site. Note also that the power of t vs d dependence suddenly grows for large d . This is an artificial finite size effect resulting from the boundary conditions.

¹Particle number n_i is related to z component of the spin via Jordan-Wigner transformation $n_i = S_i^z + 1/2$ in a fully local fashion.

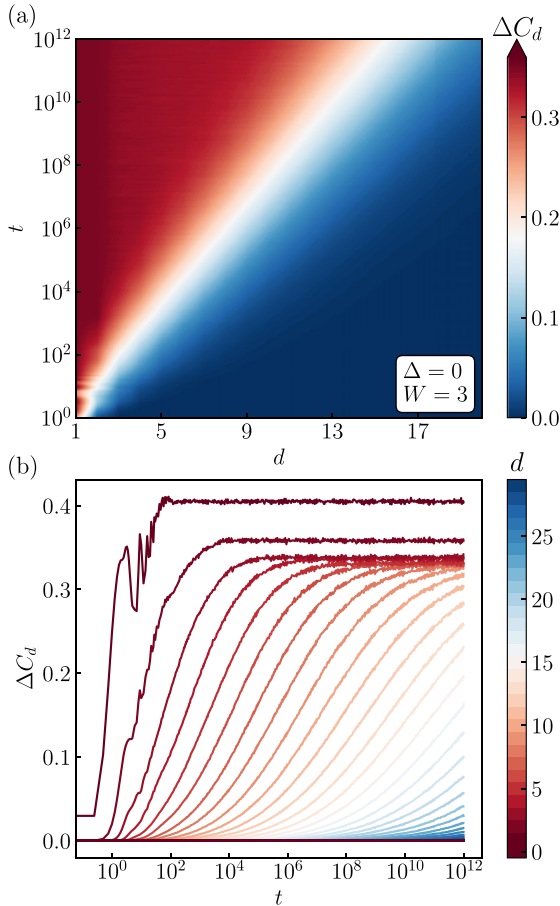


FIG. 3. (a) “Information log-lightcone” in the Anderson localized chain with disorder strength $W = 3$. Notice the logarithmic scale for time. (b) Time evolution of ΔC at different distances from the perturbed site, $L = 100$, $\ell = 50$. For a single site, a logarithmic growth in time can be observed until saturation.

B. Anderson localization

We now switch off the interactions by setting $\Delta = 0$ in (2). The model is Anderson localized for any $W > 0$, with localization length given by the approximate formula [105]

$$\xi_{\text{loc}} = \frac{3}{\log(1 + W^2/4)}, \quad (4)$$

which reproduces the well-known formulas for the weak-disorder limit $\xi_{\text{loc}} = 12/W^2$ as well as for large disorder $\xi_{\text{loc}} = 3/\log(W^2/4)$.

Given the exponential localization of single-particle eigenstates of the system, we expect that the spreading of quantum correlations from the center of the perturbation (i.e., the bond where the system is cut) should be exponentially suppressed compared to the previously analyzed case. Indeed, this is directly reflected in the behavior of CFD, as shown in Fig. 3(a) where the “information lightcone,” defined by a condition $\Delta C_d(t) = \text{const}$, is visible but on the logarithmic timescale. To perform the time evolution, we map the system to a free fermion model [70,106] which, upon exact diagonalization of the resulting Hamiltonian, allows us to access the correlation function at arbitrary time t . Disorder averages are taken over 50 000 realizations throughout this section. Similarly, in

Fig. 3(b) we observe that the growth of $\Delta C_d(t)$ begins at the time $\log t \propto d$. At later times, CFD increases logarithmically in time, $\Delta C_d(t) \sim \alpha \log t$, with α slightly decreasing with d , and then saturates to a certain long-time value. For $d = 0, 1$ this value is slightly larger than for $d \geq 2$ due to the boundary conditions (OBCs) of the chain, which strongly influence the value of $c_d(t)$ for $d \approx 1$ and do not affect $C_d(t)$. Importantly, for the Anderson insulator, in contrast to the interacting case $\Delta = 1$ discussed in the preceding section, we do not observe any decay of CFD at late times. This confirms the prediction that the latter behavior occurs only in the presence of interactions in the system.

The phenomenology of the CFD behavior in the Anderson localized system can be explained by the following perturbative arguments, which rely on the fact of exponential localization of single-particle eigenstates with localization length ξ_{loc} (4).

To simplify the following considerations, let us switch to the fermionic language in which $t_{i,i+1} \equiv J(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y)$ in (2) becomes a tunneling term and the on-site disorder term is $h_i S_i^z = h_i(n_i - 1/2)$, where n_i is a fermionic number operator at site i . Furthermore, let us consider the strong disorder limit $W \gg 1$, so that the localization length is $\xi_{\text{loc}} \ll 1$, according to (4). In such a situation, a fermion initially localized at site j at distance $d \gg \xi_{\text{loc}}$ from the cut in the system (recall Fig. 1) will be oscillating with frequency $\omega_0 = J^2/(\epsilon_j - \epsilon_{j\pm 1})$ between the site j and a neighboring unoccupied site $j \pm 1$, where $\epsilon_j \approx h_j$ in the strong disorder limit. Hence, the single-site density correlation function, $c_d(t) \sim \langle n_i(t) \rangle$, is oscillating with frequency ω_0 . When we perturb the system by introducing the link between the sites ℓ and $\ell + 1$, the frequency of oscillations of the function $C_d(t)$ changes to $\omega'_0 = J^2/(\epsilon'_j - \epsilon'_{j\pm 1})$ since the energies of the states localized at sites j and $j \pm 1$ are changed to ϵ'_j and $\epsilon'_{j\pm 1}$. In the strong disorder limit, the energy shifts $\Delta\epsilon_j = \epsilon'_j - \epsilon_j$ are given by the second-order perturbative expressions

$$\Delta\epsilon_j = \sum_{k \neq j} \frac{|\langle \phi_j | t_{\ell, \ell+1} | \phi_k \rangle|^2}{\epsilon_j - \epsilon_k}, \quad (5)$$

where $|\phi_i\rangle$ denotes an eigenstate of H_ℓ localized at site i . The sum over k extends over the whole lattice. However, due to the exponential localization of wave functions, the matrix element in the numerator is proportional to $e^{-(|\ell-j|+|\ell-k|)/\xi_{\text{loc}}}$. Hence, for a fixed distance from the cut $d = |\ell - j|$, the energy shift is exponentially small in d , $\Delta\epsilon_j \sim e^{-d/\xi_{\text{loc}}}$ (a shift of similar order is experienced at the site $j \pm 1$). Moreover, $\Delta\epsilon_j$'s fluctuate between disorder realizations, and in some instances the sum in (5) may be dominated by a single resonance term for which $\epsilon_j - \epsilon_k \ll 1$. The difference in the frequencies ω_0 and ω'_0 leads to a dephasing of the oscillations of $C_d(t)$ and $c_d(t)$ which takes place at time $t_{\text{deph}} \approx |\omega'_0 - \omega_0|^{-1} \approx e^{d/\xi_{\text{loc}}}$.

A comparison of $C_d^s(t)$ and $c_d^s(t)$ for a single disorder realization presented in Fig. 4(a) immediately confirms our predictions. We observe a gradual dephasing of the oscillations of $C_d^s(t)$ with respect to $c_d^s(t)$ which at time $t_{\text{deph}} \approx 10^3$ leads to a saturation of CFD at a value proportional to the rolling standard deviation $\sigma(C_d)$, which, in contrast to the interacting case, saturates quickly in time, as shown in Fig. 4(b). The fact that the saturation value of CFD becomes

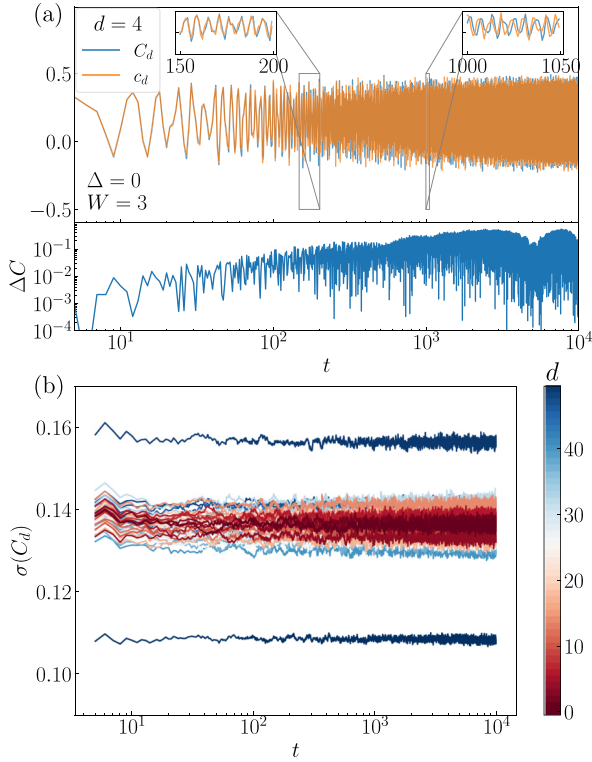


FIG. 4. Single realization (a) of time dynamics in the Anderson model for disorder amplitude $W = 3$, $L = 100$, $\ell = 50$. Upper panel shows both C_d and c_d for $d = 4$. The left inset presents early time interval $t \in [150, 200]$, in which C_d and c_d are in phase. In the right inset, for $t \in [1000, 1050]$, we observe that the dephasing between the oscillations of $C_d(t)$ and $c_d(t)$ has already occurred, yielding the major contribution to CFD, $\Delta C_d(t)$. (b) The disorder averaged rolling standard deviation $\sigma(C_d)$ of C_d . It clearly saturates fast at a d -dependent value showing that the oscillations persists in time with a similar amplitude.

independent of d for $d \gtrsim \xi_{\text{loc}}$ can be easily understood. When, at a given distance d from the cut, the time t exceeds the dephasing time t_{deph} , the value of CFD, ΔC_d , is determined by the properties of local oscillations of fermions and properties of eigenstates in the vicinity of the site ℓ . Both contributions become independent of d upon the disorder average. Finally, the fact that $\Delta C_d(t) \sim \alpha(d) \log t$, with α decreasing with d , can be expected from the fact that the distribution of t_{deph} is broad and resonances $\epsilon_j - \epsilon_k \ll 1$ in (5) yields, for certain disorder realizations, growth of $\Delta C_d(t)$ even at times $t \ll e^{d/\xi_{\text{loc}}}$.

Our predictions are further confirmed by the histograms in Fig. 5, where we investigate various contributions to the CFD. First, we consider a difference of the rolling mean values of $C_d^s(t)$ and $c_d^s(t)$, take its absolute value, and consider its distribution over disorder realizations and intervals of time between $t = 5 \times 10^3 - 10^4$. We observe that the histogram is strongly peaked at $|\langle C_d \rangle_t - \langle c_d \rangle_t| \lesssim 0.03$, confirming that the difference in the average values of $C_d^s(t)$ and $c_d^s(t)$ has a minor contribution to the CFD $\Delta C_d(t)$. In contrast, the rolling mean value $\langle \Delta C \rangle_t$ of CFD has a much broader distribution that extends over to ≈ 0.4 . The extent of this distribution is comparable to the extent of the distribution of $\sigma(C_d)$, i.e., the

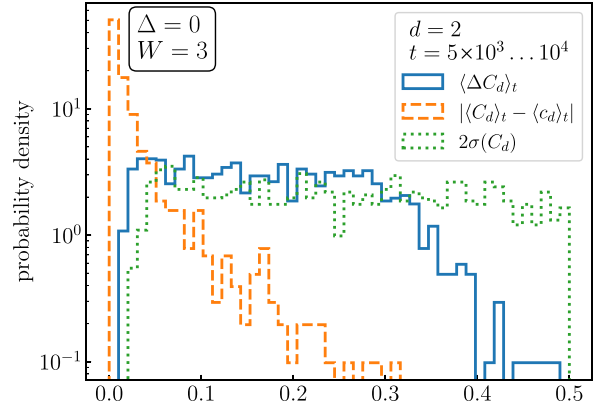


FIG. 5. Histograms showing that at late times ΔC_d comes from dephasing of $C_d(t)$ vs $c_d(t)$. The system is noninteracting ($\Delta = 0$), the disorder strength is set to $W = 3$, and $L = 16$, $\ell = 14$. We compare (i) the distribution (over disorder realizations) of the absolute value difference $|\langle C_d \rangle_t - \langle c_d \rangle_t|$ between the rolling mean values of $C_d^s(t)$ and $c_d^s(t)$; (ii) the rolling mean value $\langle \Delta C_d \rangle_t$ of CFD; and (iii) the rolling standard deviation $\sigma(C_d)$ of $C_d(t)$. The rolling window corresponds to $\Delta t = 10$ tunneling times.

rolling standard deviation of the oscillations of $C_d(t)$ around the mean. This shows that the main contribution to the CFD, $\Delta C_d(t)$, is due to the dephasing mechanism rather than due to the much smaller difference of the mean values.

Moreover, we observe that the localization length ξ_{loc} is a universal length scale that describes the system for different values of W . In Fig. 6 we show that the CFD, $\Delta C_d(t)$, at fixed time t , decreases exponentially as a function of d/ξ_{loc} , with the agreement between the curves becoming more evident at large times. Figure 7 shows the time t_{thr} of reaching a given, fixed value of the CFD, $\Delta C_d(t)$. This time scales exponentially with d/ξ_{loc} for many values of W , again confirming the presence of “information log-lightcone” as well as the fact that ξ_{loc} provides the relevant length scale. The slope in Fig. 7

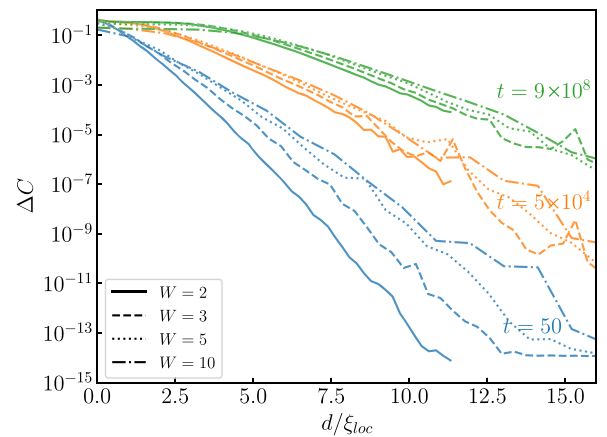


FIG. 6. Exponential decrease of the CFD $\Delta C_d(t)$ with distance d from the cut in the system with $L = 100$, $\ell = 50$, measured in units of characteristic length scale for the problem, the localization length ξ_{loc} . With increasing time, universal behavior becomes more prominent. Saturation at large values of d/ξ_{loc} is due to accumulation of numerical errors in the time propagation.

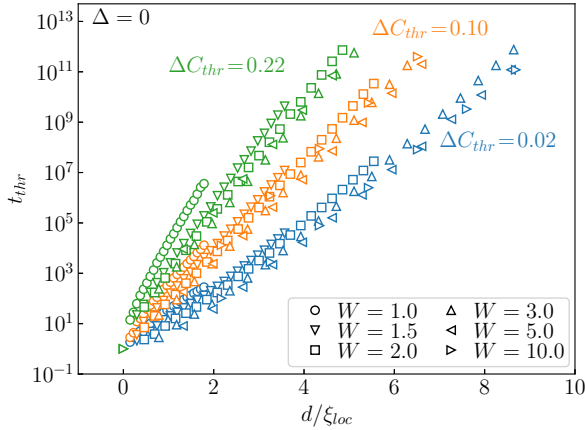


FIG. 7. Time t_{thr} of reaching a given value of the CFD, $\Delta C_d(t_{\text{thr}}) = \Delta C_{\text{thr}}$, in the Anderson model [Eq. (2)] with $\Delta = 0$ for different disorder strengths W , $L = 100$, $\ell = 50$.

depends on the saturation level in ΔC_d , which confirms that the distribution of dephasing time t_{deph} has nonvanishing weight even at $t_{\text{deph}} \ll e^{d/\xi_{\text{loc}}}$. This trend, according to expectations, is more pronounced at smaller disorder strengths, e.g., $W = 1$.

Analysis performed for the Anderson model shows that in the noninteracting localized system the CFD, $\Delta C_d(t)$, changes on timescales that grow exponentially with the distance, with a universal length scale set by the localization length ξ_{loc} . Due to the simplicity of the noninteracting case, we were able to pinpoint the dephasing mechanism that is responsible for the main contribution to CFD. In the next section, we perform an analogous study but in presence of interactions for the MBL phase.

C. MBL phase

Let us now consider the interacting case, setting $\Delta = 1$ in (2). We start by considering $W = 5$, the disorder amplitude which is well within the MBL regime for small system sizes analyzed here (recall that $W_c \approx 3.5$ is the estimation [64] of the position of ergodic-MBL crossover at $L \approx 20$). To calculate the time evolution we use the Chebyshev expansion of the time evolution operator [102]. Similarly to the noninteracting case, we observe a logarithmic lightcone for the CFD $\Delta C_d(t)$ for $\Delta = 1$; cf. Fig. 8(a). Another feature shared by the interacting and the Anderson localized models is the linear growth of $\Delta C_d(t)$ with $\log t$ presented in Fig. 8(b). There is, however, a notable difference in the long-time behavior of $\Delta C_d(t)$ between the two cases considered. While in the AL case, the values of $\Delta C_d(t)$ always grow and saturate at a constant value given by the $C_d(t)$ long-time oscillation amplitude (except for $d = 0, 1$, which saturate at slightly larger values due to OBC around site ℓ ; see discussion in Sec. III B), in the MBL case, at small values of d , $\Delta C_d(t)$ first grows, then saturates, and finally at still longer times decreases towards an equilibrium value. The phenomenology of the initial, logarithmic in time, growth of the CFD at all distances d can be explained in the same manner as for the AL case. Here the local integrals of motion (LIOM) [9,10,18,107] play the role of projectors on localized eigenstates, and the main contribution to CFD comes from the dephasing mechanism.

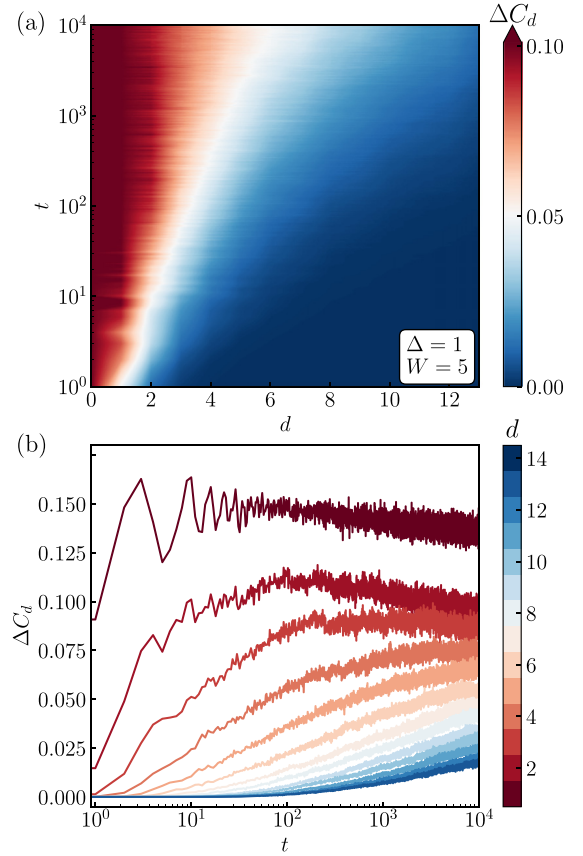


FIG. 8. (a) “Information log-lightcone” in the many-body localized chain with disorder strength $W = 5$, $L = 16$, $\ell = 14$. Notice the logarithmic scale for time. (b) Time evolution of ΔC_d at different distances from the cut at ℓ . Change of monotonicity is observed due to decrease of the $C_d(t)$ oscillation amplitude with time; for details see Fig. 9.

The change in monotonicity of CFD at later times is analogous to the previously analyzed ergodic case; cf. Sec. III A. It can be technically explained by an inspection of the oscillation amplitudes of $C_d(t)$ and $c_d(t)$ calculated as the rolling standard deviations of the signal $\sigma(C_d(t))$ for the rolling window of $\Delta t = 10$, according to Eq. (3). In contrast to the Anderson localized phase, where $\sigma(C_d(t))$ strictly approaches a constant as a function of time t (see the bottom panel in Fig. 4 for this effect), temporal fluctuations of $C_d(t)$ decrease logarithmically in time in the interacting model. This decay is significant, as $\sigma(C_d(t))$ decreases by a factor of even 30% between short times of order 10 and longer times of order 10^4 ; see Fig. 9(a). The decay of $\sigma(C_d(t))$ reflects the fact that oscillations present in $C_d(t)$ [and $c_d(t)$] actually decay in amplitude with time, as observed earlier for interacting systems in [103,104]. To compensate for this effect we may inspect the dimensionless rescaled CFD $\Delta C_d/\sigma(C_d(t))$ shown in Fig. 9(b). It grows monotonically in time and does not saturate before time $t = 10^4$. This suggests that slow but nontrivial dynamics of the density in the system persist to the longest times considered.

Finally, we consider time evolution of larger systems ($L = 100, 200$) calculated with the TDVP algorithm [108–111]. We choose a relatively strong disorder $W = 10$, which allows us

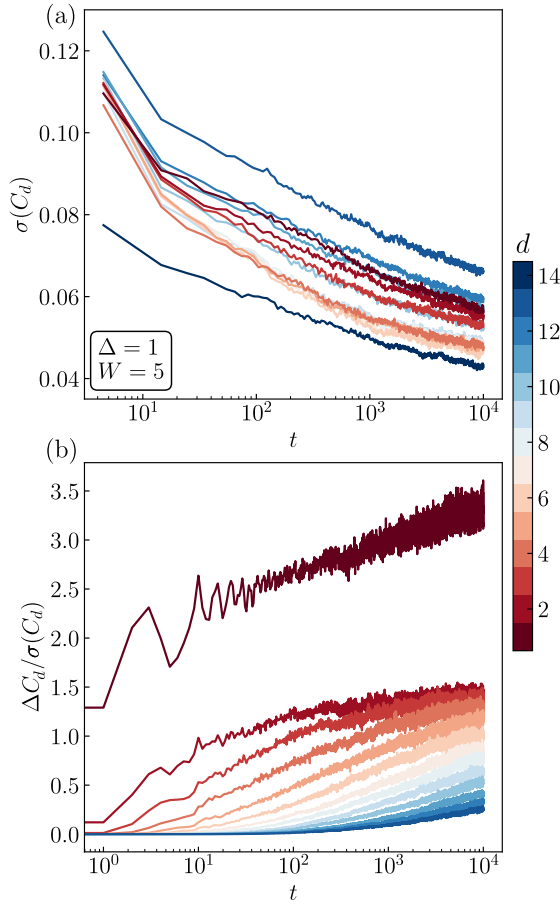


FIG. 9. (a) Rolling standard deviation of $C_d(t)$ with time window of size $\Delta t = 10$ and system size $L = 16$, cut at $\ell = 14$. Deviation from the mean decreases with time, explaining the decrease of ΔC_d at long times observed in Fig. 8. (b) $\Delta C_d(t)$ in units of the rolling standard deviation. The decreasing trend is inverted.

to calculate the time evolution of the system to times of the order of $10^3 J^{-1}$ with modest bond dimensions $\chi < 200$. In Fig. 10 we plot the CFD as function of the distance d from the cut in the system, observing an exponential decay $\Delta C_d \sim e^{-d/\xi_{\text{MBL}}(t)}$. The characteristic lengthscale $\xi_{\text{MBL}}(t)$ slowly increases with time t , consistently with nontrivial dynamics of imbalance at those timescales reported in [70]. Thus the analysis of CFD in the MBL regime does not bring a definitive answer about the locality of the dynamics that would be signified by $\xi_{\text{MBL}}(t) \xrightarrow{t \rightarrow \infty} \text{const.}$ Nevertheless, our results emphasize the analogies in the behavior of CFD between the MBL regime and Anderson insulator, showing similar degrees of locality in dynamics of the systems at finite times.

IV. RELATION TO OTHER LOCALITY MEASURES

The CFD, $\Delta C_d(t)$ is a local probe of dynamics and can be compared with other measures of locality, for instance, with the entanglement entropy. We note, however, that the latter quantity is much more difficult to access experimentally.

In particular, the von Neuman bipartite entanglement entropy $S_d(t)$ of subsystem A consisting of sites $i \in [1, \dots, (\ell - d)]$ can be written as a sum of two terms (see, e.g.,

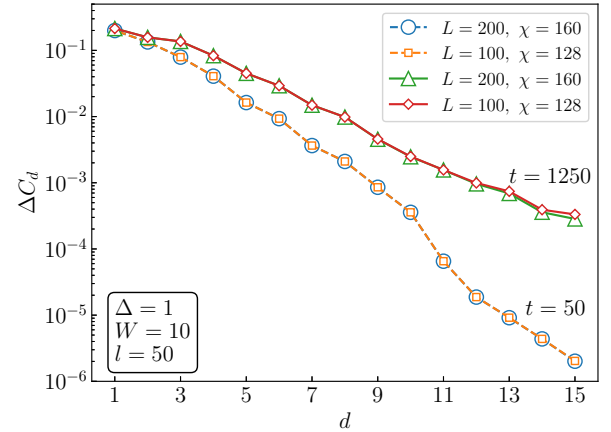


FIG. 10. Exponential decrease of the CFD ΔC_d with the distance from the cut d for XXZ chain with $L = 200, 100$, with the cut at $\ell = 50$. The disorder strength is fixed as $W = 10$, and CFD is calculated at times $t = 50$ and $t = 1250$. Time evolution of the system was found using tensor network TDVP algorithm. Average is calculated over 1008 disorder realizations.

[37,70,84,112–116] for other applications):

$$S_d(t) = S_{\text{num},d}(t) + S_{\text{conf},d}(t), \quad (6)$$

with the configurational entropy given by

$$S_{\text{conf},d}(t) = - \sum_n p(n) \text{Tr}[\rho(n) \ln \rho(n)], \quad (7)$$

where $p(n)$ denotes the probability that $n = \sum_{i=1}^{\ell-d} S_i^z$ and $\rho(n)$ is the block of the reduced density matrix in sector with magnetization n , while the number entropy reads

$$S_{\text{num},d}(t) = - \sum_n p(n) \log p(n). \quad (8)$$

In the spinless fermion description of the model, the number entropy contains information about the distribution of the number of particles in one part of the chain, with a total number of particles conserved, while the configurational entropy measures the number of possible particle arrangements within the subsystem.

We investigate the absolute differences of the two entropies $\Delta S_{\text{conf},d}(t)$, $\Delta S_{\text{num},d}(t)$ after performing the cut between sites ℓ and $\ell + 1$, in the same manner as we investigate $\Delta C_d(t)$. In Fig. 11 we find that the time dependence of $\Delta C_d(t)$ is compatible with the difference of the number entropy $\Delta S_{\text{num},d}(t)$. This is expected $\Delta S_{\text{num},d}(t)$: both $C_d(t)$ and $S_{\text{num},d}(t)$ provide information about a flow of particles between two subsystems. For small d , i.e., when the cut is close to the site studied, $\Delta C_d(t)$ and $\Delta S_{\text{num},d}(t)$ dynamics strongly differ; apparently the correlation between both quantities is small and affected by local dynamics in the vicinity of the cut. On the other hand, for sufficiently large d , say, $d > 6$, $\Delta C_d(t)$ time dependence mimics that of $\Delta S_{\text{num},d}(t)$. The latter seems to grow slowly in time with the growth being similar to $\ln \ln t$ time dependence resembling the recent findings of such a growth in the number entropy itself [65,66]. Since the matter seems to be controversial [67], especially due to the difficulty of treating really large system sizes, the similarity between $\Delta C_d(t)$ and

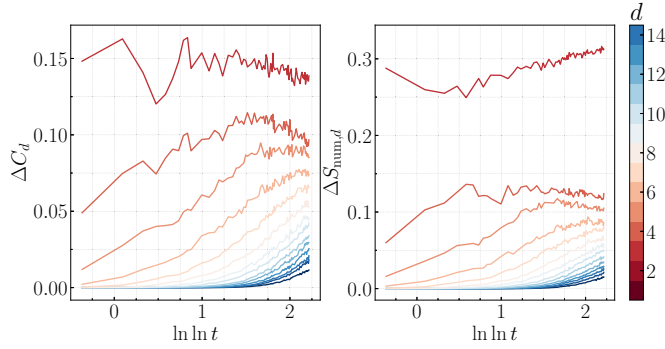


FIG. 11. Side-by-side comparison of $\Delta C_d(t)$ and the difference in the number entropy $\Delta S_{\text{num},d}(t)$ for system size $L = 16$, cut at $\ell = 14$.

$\Delta S_{\text{num},d}(t)$ time dynamics for larger d suggests than one could also use the former in experimental studies to possibly resolve the issues related to number entropy.

Additionally, we do not expect that CFD may bring some useful information on the configurational entropy that provides, let us recall, information about possible particle rearrangements in one of the subsystems. This inherently quantum information may be recovered from spin-spin correlations at different sites [117,118], but the analysis of the behavior of this quantity is beyond the scope of the present work.

V. CONCLUSIONS

We have introduced CFD, a local quantity that allows us to quantify the degree of locality of dynamics of quantum many-body systems. The CFD, $\Delta C_d(t)$, compares time evolution of single-site density correlation functions in two situations: when the 1D system comprises of L sites and when the system is cut at site $\ell < L$. A vanishing CFD at a distance d from the cut indicates that the degrees of freedom residing at sites beyond the cut do not affect the dynamics at the considered site. Conversely, a nonvanishing value of $\Delta C_d(t)$ for a certain d quantifies the degree of influence of sites beyond the cut on the dynamics at the considered distance from the cut. The determination of CFD requires a single-site resolution for measurement of density correlation functions and a single-site control over the Hamiltonian of the system. Both conditions are satisfied by present day experiments with ultracold systems.

In order to benchmark the behavior of CFD, $\Delta C_d(t)$, in various dynamical phases of matter, we have studied the time evolution of disordered XXZ spin chain. Varying the parameters of the system, we tuned the system between an ergodic phase, Anderson insulator phase, and MBL regime. We have shown that the dynamics of the considered dynamical regimes are characterized by distinct behaviors of CFD. In accordance with the expectation that the dynamics quickly gets nonlocal in the ergodic phase, we have observed a quick increase of CFD, finding that the condition $\Delta C_d(t) = \text{const}$ implies a power law relation between the distance from the cut, d , and time, t , at which CFD reaches a given value. For the Anderson localized system, we have found that CFD increases logarithmically slowly in time. Consequently, the condition $\Delta C_d(t) = \text{const}$ yields the relation $t \sim e^{d/\xi_{\text{loc}}}$ be-

tween the time and space, where ξ_{loc} is the localization length of single-particle eigenstates that becomes the relevant length scale for the problem. Finally, we have demonstrated a similar dependence for MBL regime. In the latter case, we observed a slow increase in time of the characteristic length scale that governs the exponential decay of CFD, in analogy with earlier studies of dynamics in MBL regime.

Our study demonstrates the utility of CFD in quantifying the degree of locality of quantum dynamics of many-body systems. Observing the diffusive behavior in the ergodic phase and exponentially slow dephasing in the localized case, we have shown that CFD is sensitive to dynamics of the density of particles, but remains oblivious to more complex quantum correlations in the system. Natural possible directions for future work include extending CFD to systems of bosons, understanding its behavior in the absence of symmetries such as the conservation of the total number of particles, and studying analogous differences of other, possibly nonlocal, observables such as two-point correlation functions. Another avenue for future research is to investigate the behavior of CFD in disordered many-body systems with real space renormalization group approaches [119–121].

The numerical data presented in this work are freely available from [122] or from the authors upon a reasonable request.

ACKNOWLEDGMENTS

We thank F. Evers for the illuminating discussions. The numerical computations have been possible thanks to PL-Grid Infrastructure. We acknowledge collaboration within QUANTERA DYNAMITE PCI2022-132919. T.S. and J.Z. received support from National Science Centre (Poland) under Projects No. 2019/35/B/ST2/00034, No. 2021/03/Y/ST2/00186, No. 2021/43/I/ST3/01142 as well as from the grant of Priority Research Area DigiWorld under the Strategic Programme Excellence Initiative at Jagiellonian University. P.S. and M.L. acknowledge support from ERC AdG NOQIA; Ministerio de Ciencia y Innovación Agencia Estatal de Investigaciones (PGC2018-097027-B-I00/10.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-I00, FPI, QUANTERA MAQS PCI2019-111828-2, Proyectos de I+D+I “Retos Colaboración” QUSPIN RTC2019-007196-7); MICIIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat de Catalunya; Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program, AGAUR Grant No. 2021 SGR 01452, QuantumCAT \ U16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020); Barcelona Supercomputing Center MareNostrum (FI-2022-1-0042); EU Horizon 2020 FET-OPEN OPTologic (Grant No. 899794); EU Horizon Europe Program (Grant Agreement 101080086 - NeQST), National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314); ICFO Internal “QuantumGaudi” project; European Union’s Horizon 2020 research and innovation program under the Marie-Sklódowska-Curie Grant Agreement No. 101029393 (STREDCH) and No. 847648 (“La Caixa” Junior Leaders fellowships ID100010434

LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013). Views and opinions expressed in this work are, however, those of the author(s) only and do not necessarily reflect those of

the European Union, European Climate, Infrastructure and Environment Executive Agency (CINEA), nor any other granting authority. Neither the European Union nor any granting authority can be held responsible for them.

-
- [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] L. Vidmar and M. Rigol, Generalized gibbs ensemble in integrable lattice models, *J. Stat. Mech.: Theory Exp.* (2016) 064007.
 - [6] 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).
 - [7] D. Basko, I. Aleiner, and B. Altshuler, Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states, *Ann. Phys.* **321**, 1126 (2006).
 - [8] A. Pal and D. A. Huse, Many-body localization phase transition, *Phys. Rev. B* **82**, 174411 (2010).
 - [9] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phenomenology of fully many-body-localized systems, *Phys. Rev. B* **90**, 174202 (2014).
 - [10] V. Ros, M. Mueller, and A. Scardicchio, Integrals of motion in the many-body localized phase, *Nucl. Phys. B* **891**, 420 (2015).
 - [11] M. Serbyn and J. E. Moore, Spectral statistics across the many-body localization transition, *Phys. Rev. B* **93**, 041424(R) (2016).
 - [12] W. Buijsman, V. Cheianov, and V. Gritsev, Random Matrix Ensemble for the Level Statistics of Many-Body Localization, *Phys. Rev. Lett.* **122**, 180601 (2019).
 - [13] P. Sierant and J. Zakrzewski, Level statistics across the many-body localization transition, *Phys. Rev. B* **99**, 104205 (2019).
 - [14] P. Sierant and J. Zakrzewski, Model of level statistics for disordered interacting quantum many-body systems, *Phys. Rev. B* **101**, 104201 (2020).
 - [15] W.-J. Rao, Critical level statistics at the many-body localization transition region, *J. Phys. A: Math. Theor.* **54**, 105001 (2021).
 - [16] B. De, P. Sierant, and J. Zakrzewski, On intermediate statistics across many-body localization transition, *J. Phys. A: Math. Theor.* **55**, 014001 (2022).
 - [17] W.-J. Rao, Random matrix model for eigenvalue statistics in random spin systems, *Physica A* **590**, 126689 (2022).
 - [18] 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).
 - [19] 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.
 - [20] M. Žnidarič, T. Prosen, and P. Prelovšek, Many-body localization in the Heisenberg XXZ magnet in a random field, *Phys. Rev. B* **77**, 064426 (2008).
 - [21] 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).
 - [22] M. Serbyn, Z. Papić, and D. A. Abanin, Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems, *Phys. Rev. Lett.* **110**, 260601 (2013).
 - [23] R. Nandkishore and D. A. Huse, Many-body-localization and thermalization in quantum statistical mechanics, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
 - [24] S. Bera, G. De Tomasi, F. Weiner, and F. Evers, Density Propagator for Many-Body Localization: Finite-Size Effects, Transient Subdiffusion, and Exponential Decay, *Phys. Rev. Lett.* **118**, 196801 (2017).
 - [25] P. Sierant and J. Zakrzewski, Many-body localization of bosons in optical lattices, *New J. Phys.* **20**, 043032 (2018).
 - [26] J. Herbrych, M. Mierzejewski, and P. Prelovšek, Relaxation at different length scales in models of many-body localization, *Phys. Rev. B* **105**, L081105 (2022).
 - [27] F. Alet and N. Laflorencie, Many-body localization: An introduction and selected topics, *C. R. Phys.* **19**, 498 (2018).
 - [28] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
 - [29] M. Schulz, C. A. Hooley, R. Moessner, and F. Pollmann, Stark Many-Body Localization, *Phys. Rev. Lett.* **122**, 040606 (2019).
 - [30] 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).
 - [31] S. R. Taylor, M. Schulz, F. Pollmann, and R. Moessner, Experimental probes of Stark many-body localization, *Phys. Rev. B* **102**, 054206 (2020).
 - [32] T. Chanda, R. Yao, and J. Zakrzewski, Coexistence of localized and extended phases: Many-body localization in a harmonic trap, *Phys. Rev. Res.* **2**, 032039(R) (2020).
 - [33] R. Yao and J. Zakrzewski, Many-body localization of bosons in an optical lattice: Dynamics in disorder-free potentials, *Phys. Rev. B* **102**, 104203 (2020).
 - [34] S. Scherg, T. Kohlert, P. Sala, F. Pollmann, B. Hebbe Madhusudhana, I. Bloch, and M. Aidelsburger, Observing non-ergodicity due to kinetic constraints in tilted Fermi-Hubbard chains, *Nat. Commun.* **12**, 4490 (2021).
 - [35] Q. Guo, C. Cheng, H. Li, S. Xu, P. Zhang, Z. Wang, C. Song, W. Liu, W. Ren, H. Dong, R. Mondaini, and H. Wang, Stark Many-Body Localization on a Superconducting Quantum Processor, *Phys. Rev. Lett.* **127**, 240502 (2021).
 - [36] W. Morong, F. Liu, P. Becker, K. S. Collins, L. Feng, A. Kyprianidis, G. Pagano, T. You, A. V. Gorshkov, and C. Monroe, Observation of Stark many-body localization without disorder, *Nature (London)* **599**, 393 (2021).

- [37] R. Yao, T. Chanda, and J. Zakrzewski, Many-body localization in tilted and harmonic potentials, *Phys. Rev. B* **104**, 014201 (2021).
- [38] R. Yao, T. Chanda, and J. Zakrzewski, Nonergodic dynamics in disorder-free potentials, *Ann. Phys.* **435**, 168540 (2021).
- [39] 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).
- [40] J. Feldmeier, F. Pollmann, and M. Knap, Emergent Glassy Dynamics in a Quantum Dimer Model, *Phys. Rev. Lett.* **123**, 040601 (2019).
- [41] R. M. Nandkishore and M. Hermele, Fractons, *Annu. Rev. Condens. Matter Phys.* **10**, 295 (2019).
- [42] P. Sala, T. Rakovszky, R. Verresen, M. Knap, and F. Pollmann, Ergodicity Breaking Arising from Hilbert Space Fragmentation in Dipole-Conserving Hamiltonians, *Phys. Rev. X* **10**, 011047 (2020).
- [43] V. Khemani, M. Hermele, and R. Nandkishore, Localization from Hilbert space shattering: From theory to physical realizations, *Phys. Rev. B* **101**, 174204 (2020).
- [44] P. Sierant, E. G. Lazo, M. Dalmonte, A. Scardicchio, and J. Zakrzewski, Constraint-Induced Delocalization, *Phys. Rev. Lett.* **127**, 126603 (2021).
- [45] T. Szoldra, P. Sierant, M. Lewenstein, and J. Zakrzewski, Unsupervised detection of decoupled subspaces: Many-body scars and beyond, *Phys. Rev. B* **105**, 224205 (2022).
- [46] N. J. Robinson, A. J. A. James, and R. M. Konik, Signatures of rare states and thermalization in a theory with confinement, *Phys. Rev. B* **99**, 195108 (2019).
- [47] A. J. A. James, R. M. Konik, and N. J. Robinson, Nonthermal States Arising from Confinement in One and Two Dimensions, *Phys. Rev. Lett.* **122**, 130603 (2019).
- [48] F. M. Surace, P. P. Mazza, G. Giudici, A. Leroze, A. Gambassi, and M. Dalmonte, Lattice Gauge Theories and String Dynamics in Rydberg Atom Quantum Simulators, *Phys. Rev. X* **10**, 021041 (2020).
- [49] T. Chanda, J. Zakrzewski, M. Lewenstein, and L. Tagliacozzo, Confinement and Lack of Thermalization after Quenches in the Bosonic Schwinger Model, *Phys. Rev. Lett.* **124**, 180602 (2020).
- [50] M. C. Bañuls, R. Blatt, J. Catani, A. Celi, J. I. Cirac, M. Dalmonte, L. Fallani, K. Jansen, M. Lewenstein, S. Montangero *et al.*, Simulating lattice gauge theories within quantum technologies, *Eur. Phys. J. D* **74**, 165 (2020).
- [51] M. Aidelsburger, L. Barbiero, A. Bermudez, T. Chanda, A. Dauphin, D. González-Cuadra, P. R. Grzybowski, S. Hands, F. Jendrzejewski, J. Jünemann *et al.*, Cold atoms meet lattice gauge theory, *Philos. Trans. R. Soc. A* **380**, 20210064 (2022).
- [52] A. S. Aramthottil, U. Bhattacharya, D. González-Cuadra, M. Lewenstein, L. Barbiero, and J. Zakrzewski, Scar states in deconfined $\mathbb{Z}_2$ lattice gauge theories, *Phys. Rev. B* **106**, L041101 (2022).
- [53] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, *Phys. Rev. E* **102**, 062144 (2020).
- [54] R. K. Panda, A. Scardicchio, M. Schulz, S. R. Taylor, and M. Žnidarič, Can we study the many-body localisation transition? *Europhys. Lett.* **128**, 67003 (2020).
- [55] 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.* **427**, 168415 (2021).
- [56] P. Sierant, D. Delande, and J. Zakrzewski, Thouless Time Analysis of Anderson and Many-Body Localization Transitions, *Phys. Rev. Lett.* **124**, 186601 (2020).
- [57] P. Sierant, M. Lewenstein, and J. Zakrzewski, Polynomially Filtered Exact Diagonalization Approach to Many-Body Localization, *Phys. Rev. Lett.* **125**, 156601 (2020).
- [58] A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, Avalanches and many-body resonances in many-body localized systems, *Phys. Rev. B* **105**, 174205 (2022).
- [59] P. J. D. Crowley and A. Chandran, A constructive theory of the numerically accessible many-body localized to thermal crossover, *SciPost Phys.* **12**, 201 (2022).
- [60] D. M. Long, P. J. D. Crowley, V. Khemani, and A. Chandran, Phenomenology of the prethermal many-body localized regime, [arXiv:2207.05761](https://arxiv.org/abs/2207.05761).
- [61] P. Sierant, M. Lewenstein, A. Scardicchio, and J. Zakrzewski, Stability of many-body localization in kicked Ising model, [arXiv:2203.15697](https://arxiv.org/abs/2203.15697).
- [62] J. Šuntajs and L. Vidmar, Ergodicity Breaking Transition in Zero Dimensions, *Phys. Rev. Lett.* **129**, 060602 (2022).
- [63] D. Sels, Bath-induced delocalization in interacting disordered spin chains, *Phys. Rev. B* **106**, L020202 (2022).
- [64] 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).
- [65] M. Kiefer-Emmanouilidis, R. Unanyan, M. Fleischhauer, and J. Sirker, Evidence for Unbounded Growth of the Number Entropy in Many-Body Localized Phases, *Phys. Rev. Lett.* **124**, 243601 (2020).
- [66] 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).
- [67] R. Ghosh and M. Žnidarič, Resonance-induced growth of number entropy in strongly disordered systems, *Phys. Rev. B* **105**, 144203 (2022).
- [68] 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).
- [69] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, Observation of Slow Dynamics near the Many-Body Localization Transition in One-Dimensional Quasiperiodic Systems, *Phys. Rev. Lett.* **119**, 260401 (2017).
- [70] P. Sierant and J. Zakrzewski, Challenges to observation of many-body localization, *Phys. Rev. B* **105**, 224203 (2022).
- [71] Y. Sekino and L. Susskind, Fast scramblers, *J. High Energy Phys.* **10** (2008) 065.
- [72] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, *J. High Energy Phys.* **08** (2016) 106.
- [73] J. Maldacena and D. Stanford, Remarks on the Sachdev-Ye-Kitaev model, *Phys. Rev. D* **94**, 106002 (2016).
- [74] D. J. Luitz and Y. Bar Lev, Information propagation in isolated quantum systems, *Phys. Rev. B* **96**, 020406(R) (2017).

- [75] A. Bohrdt, C. B. Mendl, M. Endres, and M. Knap, Scrambling and thermalization in a diffusive quantum many-body system, *New J. Phys.* **19**, 063001 (2017).
- [76] D. J. Luitz and Y. Bar Lev, Emergent locality in systems with power-law interactions, *Phys. Rev. A* **99**, 010105(R) (2019).
- [77] Y. Huang, Y.-L. Zhang, and X. Chen, Out-of-time-ordered correlators in many-body localized systems, *Ann. Phys.* **529**, 1600318 (2017).
- [78] R. Fan, P. Zhang, H. Shen, and H. Zhai, Out-of-time-order correlation for many-body localization, *Sci. Bull.* **62**, 707 (2017).
- [79] B. Swingle and D. Chowdhury, Slow scrambling in disordered quantum systems, *Phys. Rev. B* **95**, 060201(R) (2017).
- [80] R.-Q. He and Z.-Y. Lu, Characterizing many-body localization by out-of-time-ordered correlation, *Phys. Rev. B* **95**, 054201 (2017).
- [81] A. Smith, J. Knolle, R. Moessner, and D. L. Kovrizhin, Logarithmic Spreading of Out-of-Time-Ordered Correlators without Many-Body Localization, *Phys. Rev. Lett.* **123**, 086602 (2019).
- [82] M. Serbyn, M. Knap, S. Gopalakrishnan, Z. Papić, N. Y. Yao, C. R. Laumann, D. A. Abanin, M. D. Lukin, and E. A. Demler, Interferometric Probes of Many-Body Localization, *Phys. Rev. Lett.* **113**, 147204 (2014).
- [83] M. Rispoli, A. Lukin, R. Schittko, S. Kim, M. E. Tai, J. Léonard, and M. Greiner, Quantum critical behaviour at the many-body localization transition, *Nature (London)* **573**, 385 (2019).
- [84] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Probing entanglement in a many-body-localized system, *Science* **364**, 256 (2019).
- [85] P. Sierant, A. Maksymov, M. Kuś, and J. Zakrzewski, Fidelity susceptibility in Gaussian random ensembles, *Phys. Rev. E* **99**, 050102(R) (2019).
- [86] M. Pandey, P. W. Claeys, D. K. Campbell, A. Polkovnikov, and D. Sels, Adiabatic Eigenstate Deformations as a Sensitive Probe for Quantum Chaos, *Phys. Rev. X* **10**, 041017 (2020).
- [87] A. Maksymov, P. Sierant, and J. Zakrzewski, Energy level dynamics across the many-body localization transition, *Phys. Rev. B* **99**, 224202 (2019).
- [88] S. J. Garratt and S. Roy, Resonant energy scales and local observables in the many-body localized phase, *Phys. Rev. B* **106**, 054309 (2022).
- [89] P. Calabrese and J. Cardy, Time Dependence of Correlation Functions Following a Quantum Quench, *Phys. Rev. Lett.* **96**, 136801 (2006).
- [90] V. Eisler and I. Peschel, Evolution of entanglement after a local quench, *J. Stat. Mech.: Theory Exp.* (2007) P06005.
- [91] J.-M. Stéphan and J. Dubail, Local quantum quenches in critical one-dimensional systems: Entanglement, the Loschmidt echo, and light-cone effects, *J. Stat. Mech.: Theory Exp.* (2011) P08019.
- [92] A. Zamora, J. Rodríguez-Laguna, M. Lewenstein, and L. Tagliacozzo, Splitting a critical spin chain, *J. Stat. Mech.: Theory Exp.* (2014) P09035.
- [93] F. Jendrzejewski, S. Eckel, T. G. Tiecke, G. Juzeliūnas, G. K. Campbell, L. Jiang, and A. V. Gorshkov, Subwavelength-width optical tunnel junctions for ultracold atoms, *Phys. Rev. A* **94**, 063422 (2016).
- [94] W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice, *Nature (London)* **462**, 74 (2009).
- [95] N. Navon, R. P. Smith, and Z. Hadzibabic, Quantum gases in optical boxes, *Nat. Phys.* **17**, 1334 (2021).
- [96] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. A. Eriksson, K.-M. C. Fu *et al.*, Quantum simulators: Architectures and opportunities, *PRX Quantum* **2**, 017003 (2021).
- [97] Y. Bar Lev, G. Cohen, and D. R. Reichman, Absence of Diffusion in an Interacting System of Spinless Fermions on a One-Dimensional Disordered Lattice, *Phys. Rev. Lett.* **114**, 100601 (2015).
- [98] M. Žnidarič, A. Scardicchio, and V. K. Varma, Diffusive and Subdiffusive Spin Transport in the Ergodic Phase of a Many-Body Localizable System, *Phys. Rev. Lett.* **117**, 040601 (2016).
- [99] B. Bertini, F. Heidrich-Meisner, C. Karrasch, T. Prosen, R. Steinigeweg, and M. Žnidarič, Finite-temperature transport in one-dimensional quantum lattice models, *Rev. Mod. Phys.* **93**, 025003 (2021).
- [100] E. H. Lieb and D. W. Robinson, The finite group velocity of quantum spin systems, *Commun. Math. Phys.* **28**, 251 (1972).
- [101] C. L. Baldwin, A. Ehrenberg, A. Y. Guo, and A. V. Gorshkov, Disordered Lieb-Robinson bounds in one dimension, *arXiv:2208.05509*.
- [102] H. Fehske and R. Schneider, *Computational Many-Particle Physics* (Springer, Berlin, 2008).
- [103] M. Serbyn, Z. Papić, and D. A. Abanin, Quantum quenches in the many-body localized phase, *Phys. Rev. B* **90**, 174302 (2014).
- [104] S. Nandy, F. Evers, and S. Bera, Dephasing in strongly disordered interacting quantum wires, *Phys. Rev. B* **103**, 085105 (2021).
- [105] A. C. Potter, R. Vasseur, and S. A. Parameswaran, Universal Properties of Many-Body Delocalization Transitions, *Phys. Rev. X* **5**, 031033 (2015).
- [106] I. Peschel and V. Eisler, Reduced density matrices and entanglement entropy in free lattice models, *J. Phys. A: Math. Theor.* **42**, 504003 (2009).
- [107] A. Chandran, I. H. Kim, G. Vidal, and D. A. Abanin, Constructing local integrals of motion in the many-body localized phase, *Phys. Rev. B* **91**, 085425 (2015).
- [108] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-Dependent Variational Principle for Quantum Lattices, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [109] T. Koffel, M. Lewenstein, and L. Tagliacozzo, Entanglement Entropy for the Long-Range Ising Chain in a Transverse Field, *Phys. Rev. Lett.* **109**, 267203 (2012).
- [110] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Unifying time evolution and optimization with matrix product states, *Phys. Rev. B* **94**, 165116 (2016).
- [111] T. Chanda, P. Sierant, and J. Zakrzewski, Time dynamics with matrix product states: Many-body localization transition of large systems revisited, *Phys. Rev. B* **101**, 035148 (2020).
- [112] N. Schuch, F. Verstraete, and J. I. Cirac, Nonlocal Resources in the Presence of Superselection Rules, *Phys. Rev. Lett.* **92**, 087904 (2004).

- [113] N. Schuch, F. Verstraete, and J. I. Cirac, Quantum entanglement theory in the presence of superselection rules, *Phys. Rev. A* **70**, 042310 (2004).
- [114] W. Donnelly, Decomposition of entanglement entropy in lattice gauge theory, *Phys. Rev. D* **85**, 085004 (2012).
- [115] X. Turkeshi, P. Ruggiero, V. Alba, and P. Calabrese, Entanglement equipartition in critical random spin chains, *Phys. Rev. B* **102**, 014455 (2020).
- [116] P. Sierant, K. Biedroń, G. Morigi, and J. Zakrzewski, Many-body localization in presence of cavity mediated long-range interactions, *SciPost Phys.* **7**, 008 (2019).
- [117] G. D. Chiara and A. Sanpera, Genuine quantum correlations in quantum many-body systems: A review of recent progress, *Rep. Prog. Phys.* **81**, 074002 (2018).
- [118] T. Schweigler, V. Kasper, S. Erne, I. Mazets, B. Rauer, F. Cataldini, T. Langen, T. Gasenzer, J. Berges, and J. Schmiedmayer, Experimental characterization of a quantum many-body system via higher-order correlations, *Nature (London)* **545**, 323 (2017).
- [119] R. Vosk and E. Altman, Many-Body Localization in One Dimension as a Dynamical Renormalization Group Fixed Point, *Phys. Rev. Lett.* **110**, 067204 (2013).
- [120] C. Monthus, Strong disorder renormalization for the dynamics of many-body-localized systems: Iterative elimination of the fastest degree of freedom via the Floquet expansion, *J. Phys. A: Math. Theor.* **51**, 275302 (2018).
- [121] P. Ruggiero and X. Turkeshi, Quantum information spreading in random spin chains, *Phys. Rev. B* **106**, 134205 (2022).
- [122] <https://chaos.if.uj.edu.pl/ZOA/opendata/>.

Unsupervised detection of decoupled subspaces: Many-body scars and beyond

Tomasz Szoldra^{1,*}, Piotr Sierant², Maciej Lewenstein^{2,3} and Jakub Zakrzewski^{1,4}

¹*Instytut Fizyki Teoretycznej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*

²*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*

³*ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain*

⁴*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, PL-30-348 Kraków, Poland*



(Received 27 January 2022; revised 17 May 2022; accepted 25 May 2022; published 16 June 2022)

Highly excited eigenstates of quantum many-body systems are typically featureless thermal states. Some systems, however, possess a small number of special, low-entanglement eigenstates known as quantum scars. We introduce a quantum-inspired machine-learning platform based on a quantum variational autoencoder (QVAE) that detects families of scar states in spectra of many-body systems. Unlike a classical autoencoder, QVAE performs a parametrized unitary operation, allowing us to compress a single eigenstate into a smaller number of qubits. We demonstrate that the autoencoder trained on a scar state is able to detect the whole family of scar states sharing common features with the input state. We identify families of quantum many-body scars in the PXP model beyond the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ families and find dynamically decoupled subspaces in the Hilbert space of disordered, interacting spin-ladder model. The possibility of an automatic detection of subspaces of scar states opens new pathways in studies of models with a weak breakdown of ergodicity and fragmented Hilbert spaces.

DOI: [10.1103/PhysRevB.105.224205](https://doi.org/10.1103/PhysRevB.105.224205)

I. INTRODUCTION

Recent progress in noisy, intermediate-scale quantum (NISQ) computers [1–3] lead to a fast development of algorithms suited for use on these machines [4] with the purpose of achieving a quantum advantage in various areas: physics, machine learning, quantum chemistry, and combinatorial optimization. Of particular importance are variational quantum algorithms [5], in which quantum circuits are applied to quantum states, whose parameters are optimized with classical feedback loops. Physical applications include variational quantum eigensolvers [6–8], algorithms for ground-state preparation [9], time-evolution simulations [10–12], or quantum variational autoencoders (QVAE) [13–15]. The autoencoders encode the input data into a reduced representation and then use it to reconstruct the data with the optimal fidelity. As such, autoencoders are basic tools for data compression in machine learning. In turn, the task of QVAE is to realize the unitary transformation that transfers the input entangled n -qubits state into a product state of k , relevant, entangled qubits and $n - k$ “trash” separable qubits [see Eq. (1) below]. QVAE have been realized experimentally in a photonic device [16] and recently employed in investigation of quantum phase transitions [17]. In this work we demonstrate the applicability of QVAE in an analysis of properties of highly excited eigenstates in quantum many-body systems.

According to eigenstate thermalization hypothesis (ETH) [18–20], a small subsystem of an isolated, interacting quantum many-body system is described by a thermal density

matrix after a long-time evolution, irrespectively of the initial nonequilibrium state. However, some systems violate this paradigm of quantum ergodicity and exhibit a long-time behavior dependent on the initial state. Examples of such nonergodic systems include integrable systems [21] and many-body localized phases in the presence of quenched disorder [22–26]. Another mechanism of ergodicity breaking in a form of persistent oscillations for particularly chosen initial states has been discovered in the experiment with ultracold Rydberg atoms [27]. This behavior arises due to the presence of few atypical, almost equally spaced eigenstates with low entanglement entropy, the so-called quantum many-body scars (QMBS) [28,29] that are embedded in the otherwise thermal spectrum of a quantum many-body system. For initial states with high overlap with a few QMBS, one observes long-lived oscillations of observables, whereas for generic initial conditions the system quickly approaches the thermal equilibrium state. Several theories explaining the emergence of QMBS were proposed starting long time ago with the notion of “scars of symmetries” [30] (see also [31]): a spectrum generating algebra [32,33], Krylov restricted thermalization [34], projector embedding [35], and the presence of symmetric, coupled subspaces [36]. The QMBS occur in PXP model [37], describing Rydberg atoms chain, but also in Affleck-Kennedy-Lieb-Tasaki (AKLT) model [38,39], quantum local random networks [40], frustrated magnetic lattices [41], lattice gauge theories [42–44], optical lattices [45], or spin systems [46,47].

The aim of this work is to provide a scheme to detect families of QMBS based on QVAE. A *family* of QMBS is formed by eigenstates which (a) have an increased overlap with some “parent” state, (b) are characterized by a sub-volume-law entanglement entropy. The property (a) enables our algorithm

*tomasz.szoldra@doctoral.uj.edu.pl

to implicitly extract the features of the parent state from a single eigenstate in the training and encode them in QVAE. The probability that training succeeds is enhanced by the property (b). The performance of the trained autoencoder on other eigenstates serves as a measure of similarity between the eigenstates. The other representatives of the family of QMBS are found as eigenstates for which the performance of QVAE is significantly better than the typical behavior. In the following, we first describe details of the scheme. Then we apply it to detect the $\mathbb{Z}_2$ family and to discover different families of QMBS in the PXP model [28,37] and to identify subspaces of decoupled eigenstates in the spin ladder of [46].

II. QUANTUM VARIATIONAL AUTOENCODER

The QVAE aims to compress the n -qubit input state $|\psi\rangle$ into a k -qubit state $|\phi\rangle$ (where $k < n$), i.e., to perform a unitary transformation $U(\theta)$ parametrized by the circuit parameters θ ,

$$|\psi\rangle \rightarrow U(\theta)|\psi\rangle = |\phi\rangle \otimes |0\rangle^{\otimes(n-k)}, \quad (1)$$

where the last $n - k \equiv N_{\text{trash}}$ qubits are called “trash” qubits. The QVAE circuit starts with randomly initialized parameters θ which are variationally adjusted (here, using simultaneous perturbation stochastic approximation optimizer [48–51]) to “optimal” value $\theta = \theta^*$ satisfying Eq. (1) for a given set of input states $\{|\psi_i\rangle\}$. At each optimization step, the unitary $U(\theta)$ is applied on $\{|\psi_i\rangle\}$ and the trash qubits are measured giving either “0” or “1.” The total number of 1 defines the cost function that the optimizer tries to minimize in the next parameter update. (Cost function is the Hamming distance between the measured bit string of 1 and 0 on all trash qubits and the desired $|0\rangle^{\otimes(n-k)}$ trash qubit configuration.) Unlike for classical autoencoders, training a unitary encoder $U(\theta)$ automatically provides the decoder $U^\dagger(\theta)$ that can be applied to the compressed state $|\phi\rangle \otimes |0\rangle^{\otimes(n-k)}$ to reconstruct the input state $|\psi\rangle$. The cost function fulfills the requirement of locality on the trash qubits which is critical for circumventing the “barren plateaus” of the cost landscape and trainability of the model (see [52,53] for further details).

The architecture of the quantum circuit has to be expressive (i.e., able to encode a large class of quantum states with a few trainable parameters θ) and to possess a large entangling capability to transfer the entanglement of the whole system out of the trash qubits [54]. Building on the previous results [52,54] we choose alternating layered ansatz consisting of layers with single-qubit rotations around the y axis by an angle $\theta \in [0, 2\pi]$, $R_y(\theta) = \exp(-i\sigma_y\theta/2)$, and two-qubit controlled-Z (CZ) gates that apply a σ_z operator on one of the qubits if the other one is in the state $|1\rangle$ and act as an identity if the other qubit is in the state $|0\rangle$. Each of L layers of QVAE consists of $R_y(\theta)$ rotations of all qubits and CZ entangling operations between the neighboring qubits, with the pairs of entangled qubits alternating from layer to layer following a checkerboard pattern (see [55]).

In our scheme, as an input state $|\psi\rangle$ to train QVAE we take a single scar state that belongs to a given family of QMBS in a considered many-body system. To identify the other scars from the same family, we evaluate the performance of QVAE on eigenstates from the spectrum of the model. The numerical complexity of the procedure is thus $O(\mathcal{D})$ times

the number of iterations for QVAE training and $O(\mathcal{D}^2)$ for the comparison of eigenstates where $\mathcal{D}$ is the dimension of the Hilbert space. This is lower than the exact diagonalization cost $O(\mathcal{D}^3)$ required for generation of the input data.

III. SCARS IN THE PXP MODEL

The PXP Hamiltonian reads

$$\hat{H} = \sum_i \hat{P}_{i-1} \sigma_i^x \hat{P}_{i+1} \quad (2)$$

with periodic boundary conditions, where the projectors $\hat{P}_i = (1 - \sigma_i^z)/2$ ensure that neighboring spin-up states are separated at least by one lattice site, hence implementing the Rydberg blockade phenomenon [56] as a constraint on the Hilbert space. Certain specific initial states like $\mathbb{Z}_2 = |0101\dots\rangle$, $\mathbb{Z}_3 = |001001\dots\rangle$ and product states that contain domain walls between $\mathbb{Z}_2$ and $\mathbb{Z}_3$ configurations give rise to persistent long-time oscillations of the local observables and the revivals of the wave function, while other states like $|\mathbb{Z}_0\rangle = |0000\dots\rangle$ and $|\mathbb{Z}_4\rangle = |00010001\dots\rangle$ show fast relaxation without revivals. The presence of families of $\mathbb{Z}_2$ and $\mathbb{Z}_3$ quantum scars gives rise to this behavior [57–65]. Some of the scarred states in the PXP model were constructed exactly as matrix product state (MPS) with a finite bond dimension [66], from which the family of $\mathbb{Z}_2$ scars was generated as quasiparticle excitations.

Input data to QVAE correspond to the eigenstates of the PXP model obtained through exact diagonalization for the system size $N = 24$ in the inversion-symmetric and zero-momentum sector with the Hilbert space dimension $\mathcal{D} = 2359$ ($\mathcal{D} \ll 2^L$ due to constraints and periodic boundary conditions). Local constraints of the PXP model allow to reduce the computational cost of the procedure by considering only the projection of QVAE onto the constrained subspace of Hilbert space. To that end, it suffices to substitute $R_y(\theta) \rightarrow \tilde{R}_y(\theta)$, $CZ_i \rightarrow E_i$ in the circuit, with $\tilde{R}_y(\theta)$ rotating qubit i only if qubits $i - 1, i + 1$ are in the $|0\rangle$ state (identity otherwise), and E_i acting on four qubits $i - 1, \dots, i + 2$, performing the entangling operation of qubits i and $i + 1$ if qubits $i - 1, i + 2$ are in the state $|0\rangle$ (identity otherwise). Exact matrix forms of these operators are given in [55]. This version of the QVAE will be referred to as the CQVAE. We should note that the translational and inversion symmetries of the original Hamiltonian are not exploited in the CQVAE because these symmetries are manifestly broken by the considered few-qubit gates. Thus, for $N = 24$, the CQVAE still acts on 2359 eigenstates but each of them is expressed in a 103 682-dimensional Hilbert space.

We select a $\mathbb{Z}_2$ scar at energy $E \approx -2.67$ as the training input state $|\chi_0\rangle$. This state has a significant overlap with the $\mathbb{Z}_2$ configuration, i.e., $|\langle\chi_0|\mathbb{Z}_2\rangle|^2 = 0.15$ (cf. with a value expected in the high-temperature thermal ensemble $1/\mathcal{D} \approx 0.0004$). Figure 1(a) shows performance of a trained autoencoder on eigenstates from the PXP Hamiltonian spectrum. Indeed, we see that the $\mathbb{Z}_2$ scars are characterized by a significant drop in the CQVAE cost. Plots of CQVAE reconstruction fidelity (not shown) also reveal high-fidelity peaks on $\mathbb{Z}_2$ scars. In this way, the family of $\mathbb{Z}_2$ scars can be identified in an automatic way. Interestingly, the $\mathbb{Z}_3$ family has the largest

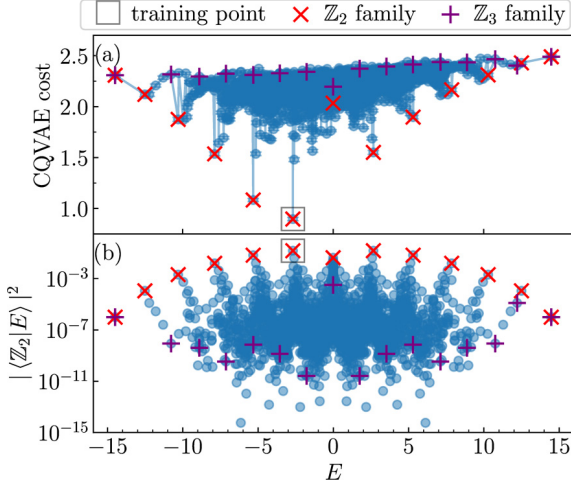


FIG. 1. (a) Performance of the CQVAE with $N_{\text{trash}} = 8$ and $L = 7$ trained on the strongest $\mathbb{Z}_2$ scar of the PXP model with $N = 24$ close to the middle of the spectrum, applied to all eigenstates. Best performance is observed in the eigenstates from the $\mathbb{Z}_2$ scars family that have an increased overlap with the $|\mathbb{Z}_2\rangle$ state, as presented in (b). Lines serve as a guide to the eye. Error bars come from averaging over 32 independent trainings.

cost even though it has low entanglement entropy showing that CQVAE learned to distinguish the real-space patterns $\mathbb{Z}_3$ scar states from the dominant configurations of $\mathbb{Z}_2$ family [see Fig. 1(b)]. Interestingly, at smaller system sizes, e.g., $N = 12$, we can recover the “parent” state $|\chi\rangle \approx |\mathbb{Z}_2\rangle$ for the family by optimizing the *input* of the trained CQVAE to minimize cost (see [55] for details).

The next step is to find other scarred families. We select eigenstates with a low entanglement entropy $S = -\text{Tr} \rho_A \ln \rho_A < 2.7$, where ρ_A is the reduced density matrix of the half of the spin chain, train the CQVAE on each of them, and calculate the cost on other eigenstates. Pairs of eigenstates that have a low cost when training on both of them are regarded as belonging to the same scar family. This property is transitive, i.e., if eigenstates $|E_1\rangle, |E_2\rangle$ have a small cost and $|E_2\rangle, |E_3\rangle$ as well, then a set $|E_1\rangle, |E_2\rangle, |E_3\rangle$ is regarded as one family. Figure 2 shows five families discovered by the CQVAE in the PXP model for $N = 24$, with an example of the training results on one representative of the family No. 1 [Fig. 2(a)]. The eigenstates from the family No. 1 are characterized by increased overlaps with several Fock states, the example of $|i_1\rangle = |101010010010100100100100\rangle$ state is shown in Fig. 2(b). Other Fock states with high overlaps with the family No. 1 contain a mixture of the same number of rearranged three $\mathbb{Z}_2$ and six $\mathbb{Z}_3$ patterns. The same holds true for families No. 2–No.5 of scar states found for $N = 24$. Patterns with a larger period can be found, e.g., family No. 2 has increased overlaps with four $\mathbb{Z}_3$ and three $\mathbb{Z}_4$ configurations. In that way, our QVAE-based scheme allows to explain the presence and identify relations between eigenstates with low entanglement entropy of the PXP model [see Fig. 2(c)]. Since the families of scar states do not have a single simple representative Fock state, their classification, especially in system-size independent manner is more involved; in [55] we show the details of the families for $N = 24$.

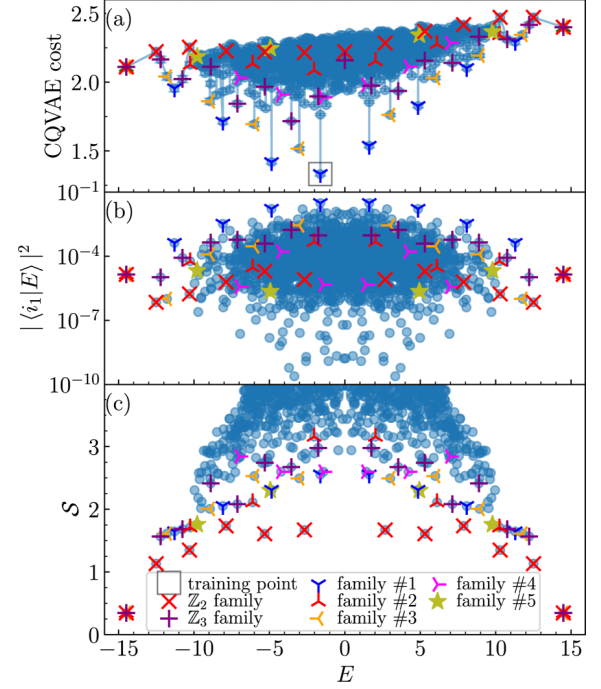


FIG. 2. (a) CQVAE cost of eigenstates of the PXP model, $N = 24$, reveals a scar “family No. 1” upon training on one of its representatives. Similar plots for other scarred families found by the algorithm are given in [55]. (b) Overlaps of the eigenstates with state $|i_1\rangle$. (c) Entanglement entropy vs energy.

IV. DISORDERED SPIN LADDER

Consider a spin ladder with Hamiltonian

$$H = H^{\parallel} + H^{\perp} = \frac{1}{4} \sum_{k=1}^{L-1} h_{k,k+1}^{\parallel} + \frac{1}{4} \sum_{k=1}^L h_k^{\perp}, \quad (3a)$$

where

$$h_{k,k+1}^{\parallel} = \sigma_k^x \sigma_{k+1}^x + \sigma_k^y \sigma_{k+1}^y + \tau_k^x \tau_{k+1}^x + \tau_k^y \tau_{k+1}^y, \\ h_k^{\perp} = J(\sigma_k^x \tau_k^x + \sigma_k^y \tau_k^y) + \Delta_k \sigma_k^z \tau_k^z + h_k(\sigma_k^z + \tau_k^z). \quad (3b)$$

$k = 1, \dots, N$ labels the rungs of the ladder, and spins on the left and right legs of the ladder are represented by Pauli matrices σ_k^{α} and τ_k^{α} ($\alpha = x, y, z$), respectively. Values of h_k are drawn from a uniform distribution in the interval $[-h, h]$, and we set $J = 1$, $h = 0.1$, $\Delta_k = 1$. The model has a $U(1)$ symmetry associated with the total magnetization $Z = \sum_{k=1}^N (\sigma_k^z + \tau_k^z)$ and a $\mathbb{Z}_2$ symmetry associated with the exchange of the ladder legs $\sigma_k^{\alpha} \leftrightarrow \tau_k^{\alpha}$. Even though this model has signatures of quantum ergodicity (e.g., energy-level spacings follow the Wigner-Dyson distribution), one can analytically construct exact invariant subspaces of the Hamiltonian (3) that are not related to any local conserved quantity as shown in [46]. It is first noticed that the eigenstates of $h_k^{\perp}$ on a single rung are $|S\rangle = (|1_1^0\rangle - |1_0^1\rangle)/\sqrt{2}$ (“singlet”), $|T\rangle = (|1_1^0\rangle + |1_0^1\rangle)/\sqrt{2}$ (“triplet”), $|D\rangle = |1_0^0\rangle$ (“doublon”), $|H\rangle = |1_1^1\rangle$ (“holon”), where the first (second) row of the vector corresponds to the left (right) leg of the ladder. Product of such states is an eigenstate of the total leg Hamiltonian $H^{\perp}$. By examining the action of the remaining $H^{\parallel}$ on the two-rung

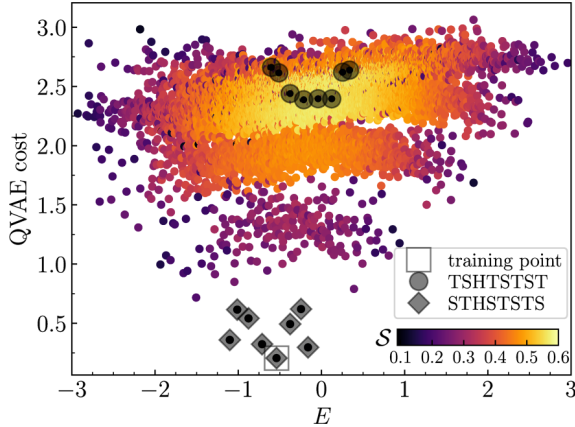


FIG. 3. Cost function of QVAE trained on an eigenstate from the one-holon subspace of the spin-ladder model with $N = 8$ (16 spins in total), evaluated on the eigenstates from $Z = 1$ symmetry sector. Color bar denotes the bipartite entanglement entropy for 4 rungs. Results averaged over 56 independent trainings of the QVAE with $N_{\text{trash}} = 5$ and $L = 10$.

states $\{|ST, SH, TH, HH \dots\rangle\}$ one shows that $H^\parallel$ annihilates configurations ST, TS, HH, DD , and moves H (or D) around if S or T are its neighbors. It follows that configurations $|STSTST \dots\rangle$ and $|TSTSTS \dots\rangle$ are annihilated by $H^\parallel$ (they are a “vacuum background”). Upon inserting a given number of only holons (or doublons) between them, e.g., $|STSH_jTSH_kTS \dots\rangle$, one constructs an invariant subspace with a given number of the four letters that are conserved under the action of the total Hamiltonian H . Dimension of such a subspace after r insertions of doublons (or holons) is $\binom{N}{r}$.

Construction of invariant subspaces in this model required an involved theoretical insight [46]. Here, by employing the QVAE we can detect their presence in an automatic way. Let us restrict only to the $Z = 1$ symmetry sector of the Hamiltonian (3) with $N = 8$ rungs and a single disorder realization. We encode the ladder state onto the spatially one-dimensional quantum circuit, by mapping the left (right) leg of the ladder to odd (even) sites of the circuit. In this manner, the neighboring spins of the ladder are mapped to sites of the circuit that are close to each other. Training of the QVAE on a generic eigenstate and application to all others gives a featureless QVAE cost. However, if we train on an eigenstate from the invariant subspace represented by state $|STSTHTSTS\rangle$ with 1 holon H , 4 singlets S , and 3 triplets T , we get a significantly lower error on all 8 eigenstates that span this subspace (see Fig. 3). The QVAE cost on the other invariant subspace $|TSTSHSTST\rangle$ with a different number of singlets and triplets [but an identical entanglement entropy (notice the color scale in Fig. 3)] is comparable to the QVAE cost on generic, highly entangled, eigenstates. Similarly, a QVAE trained on an eigenstate from the subspace $|TSTSHSTST\rangle$ yields small costs on eigenstates from this subspace whereas the cost on eigenstates from $|STSTHTSTS\rangle$ is substantial (plot not shown). This is in a full analogy with results for PXP model and demonstrates how QVAE identifies and distinguishes families of scar states.

V. CONCLUSION

We proposed a scheme based on QVAE that allows to identify families of nonergodic eigenstates of quantum many-body systems. We validated our scheme on the $\mathbb{Z}_2$ family of scar states of the PXP model. Then, our scheme was employed to demonstrate presence of families of scar states in spectrum of PXP model beyond the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ families. To confirm the generality of our approach, we used it to identify the family of scar states in the disordered spin ladder (3) as well as in a PXP-like model with a three-body blockade [55]. The use of QVAE is crucial in our scheme. QVAE learns from a single, high-dimensional measurement point (a single eigenstate), in contrast to classical autoencoders that require a larger set of training data. The flow of the entanglement entropy through the layers of the autoencoder has a physical meaning and results in a compression of the quantum state. By respecting the laws of quantum mechanics, QVAE becomes a versatile tool in studying eigenstates of many-body systems allowing, for instance, for a direct implementation of the local constraints of PXP model on the QVAE. Finally, we also point out that having the eigenstates one can try to identify scarred states by calculating the inverse participation ratio (IPR) with a negligible numerical cost instead of costly entanglement entropy. Moreover, instead of exact diagonalization used here, one may use an approximate algorithm, e.g., DMRG-X [67], to obtain a subset of eigenstates serving as an input to QVAE.

While all calculations performed here used classical machine, hardware implementation of QVAE on a physical quantum computer seems straightforward [17]. Although preparation of ground states of selected Hamiltonians is possible by the variational quantum eigensolvers, algorithms that provide excited states are more involved [4]. Hence, the preparation of the input states is the most challenging step of our scheme that is feasible only for limited system sizes. However, to navigate through the exponentially large Hilbert space one can use a prior knowledge about the scar states (e.g., their energy) which may extend the interval of system sizes accessible on current quantum hardware.

ACKNOWLEDGMENTS

We are grateful to K. Kottmann for valuable suggestions. The numerical computations have been possible thanks to PL-Grid Infrastructure. This research was partially funded by the Priority Research Area Digiworld under the program Excellence Initiative—Research University at the Jagiellonian University in Krakow. The works of T.S. and J.Z. have been realized within the Opus Grant No. 2019/35/B/ST2/00034, financed by National Science Centre (Poland). P.S. and M.L. acknowledge support from ERC AdG NOQIA; Agencia Estatal de Investigación (R&D Project No. CEX2019-000910-S, funded by MCIN/AEI/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-I00, FPI, QUANTERA MAQS PCI2019-111828-2, Proyectos de I + D + I “Retos Colaboración” QUSPIN RTC2019-007196-7); Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya through the European Social Fund FEDER and CERCA program (AGAUR Grant No. 2017 SGR 134, Quantum-

CAT U16-011424, co-funded by ERDF Operational Program of Catalonia 2014–2020); EU Horizon 2020 FET-OPEN OPTologic (Grant No. 899794); National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314); European Union's Horizon 2020 research and innova-

tion programme under the Marie-Skłodowska-Curie Grant Agreement No. 101029393 (STREDCH) and No. 847648 ("La Caixa" Junior Leaders fellowships ID100010434: LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013).

-
- [1] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 - [2] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li *et al.*, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, *Phys. Rev. Lett.* **127**, 180501 (2021).
 - [3] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, *et al.*, Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light, *Phys. Rev. Lett.* **127**, 180502 (2021).
 - [4] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
 - [5] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
 - [6] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
 - [7] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
 - [8] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Variational quantum algorithms for discovering Hamiltonian spectra, *Phys. Rev. A* **99**, 062304 (2019).
 - [9] W. W. Ho and T. H. Hsieh, Efficient variational simulation of non-trivial quantum states, *SciPost Phys.* **6**, 029 (2019).
 - [10] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).
 - [11] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, *Quantum* **3**, 191 (2019).
 - [12] S. Endo, J. Sun, Y. Li, S. C. Benjamin, and X. Yuan, Variational Quantum Simulation of General Processes, *Phys. Rev. Lett.* **125**, 010501 (2020).
 - [13] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum autoencoders for efficient compression of quantum data, *Quantum Sci. Technol.* **2**, 045001 (2017).
 - [14] C. Bravo-Prieto, Quantum autoencoders with enhanced data encoding, *Machine Learning: Sci. Technol.* **2**, 035028 (2021).
 - [15] C. Cao and X. Wang, Noise-Assisted Quantum Autoencoder, *Phys. Rev. Applied* **15**, 054012 (2021).
 - [16] A. Pepper, N. Tischler, and G. J. Pryde, Experimental Realization of a Quantum Autoencoder: The Compression of Qutrits via Machine Learning, *Phys. Rev. Lett.* **122**, 060501 (2019).
 - [17] K. Kottmann, F. Metz, J. Fraxanet, and N. Baldelli, Variational quantum anomaly detection: Unsupervised mapping of phase diagrams on a physical quantum computer, *Phys. Rev. Research* **3**, 043184 (2021).
 - [18] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
 - [19] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
 - [20] 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).
 - [21] B. Sutherland, *Beautiful Models: 70 Years of Exactly Solved Quantum Many-body Problems* (World Scientific, Singapore, 2004).
 - [22] 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).
 - [23] D. Basko, I. Aleiner, and B. Altshuler, Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states, *Ann. Phys.* **321**, 1126 (2006).
 - [24] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
 - [25] F. Alet and N. Laflorencie, Many-body localization: An introduction and selected topics, *C. R. Phys.* **19**, 498 (2018).
 - [26] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
 - [27] 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).
 - [28] 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).
 - [29] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Quantum scarred eigenstates in a rydberg atom chain: Entanglement, breakdown of thermalization, and stability to perturbations, *Phys. Rev. B* **98**, 155134 (2018).
 - [30] D. Delande and J. C. Gay, Scars of Symmetries in Quantum Chaos, *Phys. Rev. Lett.* **59**, 1809 (1987).
 - [31] M. Serbyn, D. A. Abanin, and Z. Papić, Quantum many-body scars and weak breaking of ergodicity, *Nat. Phys.* **17**, 675 (2021).
 - [32] 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).
- [33] T. Iadecola and M. Schecter, Quantum many-body scar states with emergent kinetic constraints and finite-entanglement revivals, *Phys. Rev. B* **101**, 024306 (2020).
- [34] S. Moudgalya, A. Prem, R. Nandkishore, N. Regnault, and B. A. Bernevig, Thermalization and its absence within Krylov subspaces of a constrained Hamiltonian, *Memorial Volume for Shoucheng Zhang* (World Scientific, 2021), Chap. 7, pp. 147–209.
- [35] N. Shiraishi and T. Mori, Systematic Construction of Counterexamples to the Eigenstate Thermalization Hypothesis, *Phys. Rev. Lett.* **119**, 030601 (2017).
- [36] C. J. Turner, J.-Y. Desaulles, K. Bull, and Z. Papić, Correspondence Principle for Many-Body Scars in Ultracold Rydberg Atoms, *Phys. Rev. X* **11**, 021021 (2021).
- [37] I. Lesanovsky and H. Katsura, Interacting fibonacci anyons in a rydberg gas, *Phys. Rev. A* **86**, 041601(R) (2012).
- [38] S. Moudgalya, S. Rachel, B. A. Bernevig, and N. Regnault, Exact excited states of nonintegrable models, *Phys. Rev. B* **98**, 235155 (2018).
- [39] D. K. Mark, C.-J. Lin, and O. I. Motrunich, Unified structure for exact towers of scar states in the Affleck-Kennedy-Lieb-Tasaki and other models, *Phys. Rev. B* **101**, 195131 (2020).
- [40] F. M. Surace, M. Dalmonte, and A. Silva, Quantum local random networks and the statistical robustness of quantum scars, [arXiv:2107.00884](https://arxiv.org/abs/2107.00884).
- [41] 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).
- [42] F. M. Surace, P. P. Mazza, G. Giudici, A. Lerose, A. Gambassi, and M. Dalmonte, Lattice Gauge Theories and String Dynamics in Rydberg Atom Quantum Simulators, *Phys. Rev. X* **10**, 021041 (2020).
- [43] G. Magnifico, M. Dalmonte, P. Facchi, S. Pascazio, F. V. Pepe, and E. Ercolessi, Real time dynamics and confinement in the Z_n Schwinger-Weyl lattice model for 1+1 QED, *Quantum* **4**, 281 (2020).
- [44] A. S. Aramthottil, U. Bhattacharya, D. González-Cuadra, M. Lewenstein, L. Barbiero, and J. Zakrzewski, Scar states in deconfined Z_2 lattice gauge theories, [arXiv:2201.10260](https://arxiv.org/abs/2201.10260).
- [45] H. Zhao, J. Vovrosh, F. Mintert, and J. Knolle, Quantum Many-Body Scars in Optical Lattices, *Phys. Rev. Lett.* **124**, 160604 (2020).
- [46] T. Iadecola and M. Žnidarič, Exact Localized and Ballistic Eigenstates in Disordered Chaotic Spin Ladders and the Fermi-Hubbard Model, *Phys. Rev. Lett.* **123**, 036403 (2019).
- [47] M. Schecter and T. Iadecola, Weak Ergodicity Breaking and Quantum Many-Body Scars in Spin-1 XY Magnets, *Phys. Rev. Lett.* **123**, 147201 (2019).
- [48] J. Spall, Accelerated second-order stochastic optimization using only function measurements, in *Proceedings of the 36th IEEE Conference on Decision and Control* (IEEE, Piscataway, NJ, 1997), Vol. 2, p. 1417.
- [49] J. C. Spall, An overview of the simultaneous perturbation method for efficient optimization, *Johns Hopkins APL Tech. Dig.* **19**, 482 (1998).
- [50] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [51] M. S. ANIS, Abby-Mitchell, H. Abraham, AduOffei, R. Agarwal, G. Agliardi, M. Aharoni, I. Y. Akhalwaya, G. Aleksandrowicz *et al.*, QISKIT: An open-source framework for quantum computing.
- [52] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, *Nat. Commun.* **12**, 1791 (2021).
- [53] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [54] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms, *Adv. Quantum Technol.* **2**, 1900070 (2019).
- [55] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.105.224205> for discussion of the QVAE/CQVAE implementation details, scar family identification procedure, full list of detected scar families, optimizing the QVAE input state, and short analysis of the QVAE applied to a PXP-like model with three-body constraints. It includes Refs. [48–52, 68–77].
- [56] A. Browaeys and T. Lahaye, Many-body physics with individually controlled rydberg atoms, *Nat. Phys.* **16**, 132 (2020).
- [57] V. Khemani, C. R. Laumann, and A. Chandran, Signatures of integrability in the dynamics of Rydberg-blockaded chains, *Phys. Rev. B* **99**, 161101(R) (2019).
- [58] S. Choi, C. J. Turner, H. Pichler, W. W. Ho, A. A. Michailidis, Z. Papić, M. Serbyn, M. D. Lukin, and D. A. Abanin, Emergent SU(2) Dynamics and Perfect Quantum Many-Body Scars, *Phys. Rev. Lett.* **122**, 220603 (2019).
- [59] W. W. Ho, S. Choi, H. Pichler, and M. D. Lukin, Periodic Orbits, Entanglement, and Quantum Many-Body Scars in Constrained Models: Matrix Product State Approach, *Phys. Rev. Lett.* **122**, 040603 (2019).
- [60] T. Iadecola, M. Schecter, and S. Xu, Quantum many-body scars from magnon condensation, *Phys. Rev. B* **100**, 184312 (2019).
- [61] A. A. Michailidis, C. J. Turner, Z. Papić, D. A. Abanin, and M. Serbyn, Stabilizing two-dimensional quantum scars by deformation and synchronization, *Phys. Rev. Research* **2**, 022065(R) (2020).
- [62] 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).
- [63] 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. Research* **2**, 043305 (2020).
- [64] K. Bull, J.-Y. Desaulles, and Z. Papić, Quantum scars as embeddings of weakly broken lie algebra representations, *Phys. Rev. B* **101**, 165139 (2020).
- [65] F. M. Surace, M. Votto, E. G. Lazo, A. Silva, M. Dalmonte, and G. Giudici, Exact many-body scars and their stability in constrained quantum chains, *Phys. Rev. B* **103**, 104302 (2021).
- [66] C.-J. Lin and O. I. Motrunich, Exact Quantum Many-Body Scar States in the Rydberg-Blockaded Atom Chain, *Phys. Rev. Lett.* **122**, 173401 (2019).

- [67] V. Khemani, F. Pollmann, and S. L. Sondhi, Obtaining Highly Excited Eigenstates of Many-Body Localized Hamiltonians by the Density Matrix Renormalization Group Approach, *Phys. Rev. Lett.* **116**, 247204 (2016).
- [68] K. Nakaji and N. Yamamoto, Expressibility of the alternating layered ansatz for quantum computation, *Quantum* **5**, 434 (2021).
- [69] P. Weinberg and M. Bukov, QuSpin: A python package for dynamics and exact diagonalisation of quantum many body systems part I: spin chains, *SciPost Phys.* **2**, 003 (2017).
- [70] P. Weinberg and M. Bukov, QuSpin: A python package for dynamics and exact diagonalisation of quantum many body systems. Part II: bosons, fermions and higher spins, *SciPost Phys.* **7**, 020 (2019).
- [71] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016), <http://www.deeplearningbook.org>.
- [72] S. Dooley and G. Kells, Extreme many-body scarring in a quantum spin chain via weak dynamical constraints, *Phys. Rev. B* **105**, 155127 (2022).
- [73] J.-Y. Desaulles, A. Hudomal, C. J. Turner, and Z. Papić, Proposal for Realizing Quantum Scars in the Tilted 1D Fermi-Hubbard Model, *Phys. Rev. Lett.* **126**, 210601 (2021).
- [74] D. Erhan, Y. Bengio, A. Courville, and P. Vincent, Visualizing higher-layer features of a deep network, Technical Report, Université de Montréal (2009), https://www.researchgate.net/publication/265022827_Visualizing_Higher-Layer_Features_of_a_Deep_Network/.
- [75] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, Understanding neural networks through deep visualization, *arXiv:1506.06579*.
- [76] C. Olah, A. Mordvintsev, and L. Schubert, Feature visualization, *Distill* (2017).
- [77] J.-Y. Desaulles, K. Bull, A. Daniel, and Z. Papić, Hypergrid subgraphs and the origin of scarred quantum walks in the many-body hilbert space, *arXiv:2112.06885*.

Supplemental material for "Unsupervised detection of decoupled subspaces: many-body scars and beyond"

Tomasz Szoldra,¹ Piotr Sierant,² Maciej Lewenstein,^{2,3} and Jakub Zakrzewski^{1,4}

¹*Instytut Fizyki Teoretycznej, Uniwersytet Jagielloński, Łojasiewicza 11, PL-30-348 Kraków, Poland*

²*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*

³*ICREA, Passeig Lluis Companys 23, 08010 Barcelona, Spain*

⁴*Mark Kac Complex Systems Research Center, Uniwersytet Jagielloński, Kraków, Poland*

QVAE ARCHITECTURE

The architecture of QVAE is an Alternating Layered Ansatz [1, 2] shown in Fig. S1. Optimal angles θ that minimize cost defined in the main text are found by the SPSA optimizer [3–6] with a random initial guess and automatic determination of the learning and perturbation rate implemented in Qiskit [6]. In order to utilize the constraints imposed on the Hilbert space in the PXP model, we implement the unitary circuit $U(\theta)$ from scratch using the QuSpin Python library [7, 8]. The code is available upon request.

Hyperparameters of the circuits: number of trash qubits N_{trash} , layers L , measurement shots in the determination of cost N_{shots} and training iterations N_{iter} were found heuristically by performing a grid search for smaller PXP system with $N = 18$ and choosing hyperparameters maximizing drops of cost on the $\mathbb{Z}_2$ scars if trained on another $\mathbb{Z}_2$ scar. Optimal parameters read $N_{\text{trash}} = 6$, $L = 5$, $N_{\text{shots}} = 300$, $N_{\text{iter}} = 20000$, yet other sets with those quantities changed by factors of up to 25% still yielded statistically significant drops of $\mathbb{Z}_2$ scars costs. These results were the starting point for more computationally expensive $N = 24$ case - they were scaled linearly with the system size and adjusted heuristically to optimal values $N_{\text{trash}} = 8$, $L = 7$, $N_{\text{shots}} = 600$, $N_{\text{iter}} = 50000$ used to produce Figs. 1, 2 in the main text.

Inspection of the training outcomes can give intuition about the interplay of hyperparameters. It was observed that overfitting, i.e., a trivial learning of a perfect representation of the training eigenstate, which is easily detectable by a large drop in cost of the training point compared to all other eigenstates and no detection of other $\mathbb{Z}_2$ scars in the benchmarking process, is caused by too many parameters (too large L) or too weak compression ratio (too small N_{trash}). The number of shots N_{shots} controls the variance of cost between iterations which should remain at around 1-10% to overcome local minima but not jump too far in the cost landscape, in full analogy to the sizes of batches in the stochastic gradient descent algorithm [9]. We also observed that optimal QVAE circuits had, unsurprisingly, more layers than CQVAE circuits for the same number of qubits, because of the need to represent an exponentially larger part of the Hilbert space with more parameters. We found that CQVAE was

untrainable for the same number of layers as the QVAE. We expect this is due to the presence of four-qubit operations in CQVAE which makes the optimization problem more difficult. We also noticed that the number of layers L should be large enough so that trash qubits are contained within "light-cone" of the first qubit.

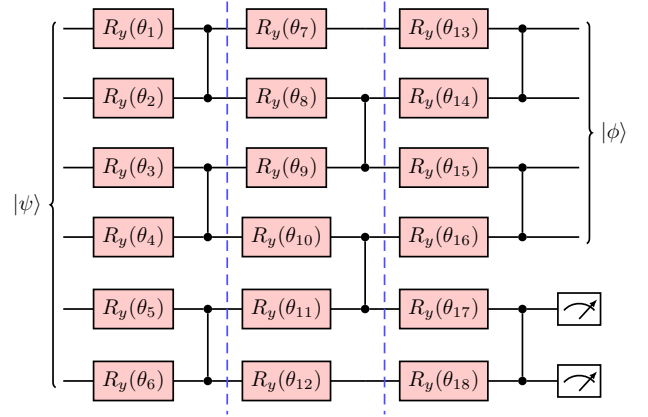


FIG. S1. Quantum Variational Autoencoder (QVAE) composed of the single-qubit rotations around the y axis parametrized by angles $\theta = (\theta_1, \dots, \theta_{18})$ and two-qubit controlled-Z gates that entangle the neighboring qubits. Separate layers are denoted by dashed vertical lines.

CQVAE OPERATORS

Below we list exact matrix forms operators used in the QVAE variant restricted to constrained Hilbert space of the PXP model. Rotation of one qubit around the y axis is performed only if its neighbours are in the state $|0\rangle$ according to

$$\tilde{R}_y(\theta) = \begin{pmatrix} |000\rangle & |010\rangle & |001\rangle & |100\rangle & |101\rangle \\ \cos(\theta/2) & \sin(\theta/2) & & & \\ -\sin(\theta/2) & \cos(\theta/2) & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \quad (\text{S1})$$

where states denoted above the matrix enumerate the 3-qubit computational basis with constraints. Similarly, two neighboring qubits are entangled only if their neighbors are in state $|0\rangle$ by the following operator:

$$E_i = \begin{pmatrix} |0000\rangle & |0010\rangle & |0100\rangle & |0001\rangle & |0101\rangle & |1000\rangle & |1010\rangle \\ 1 & & & & & & \\ & 1/\sqrt{2} & 1/\sqrt{2} & & & & \\ & 1/\sqrt{2} & -1/\sqrt{2} & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{pmatrix} \quad (\text{S2})$$

In this case, the entangling operator is different than in the CZ gate (which acts not trivially only on 2-qubit configuration $|11\rangle$ excluded by the PXP model constraints).

DETECTION OF FAMILIES IN THE PXP MODEL

We provide technical details concerning the scars identification in the PXP model. In Fig. S2 we show output of CQVAE used to determine members of family #1. We see that there is an overall growth of cost with distance in energy. Nevertheless, family members can be detected one by one using transitivity (see main text). Similar plots have been used to detect other families. Clusters of mutually related eigenstates (incl. transitivity) were identified by graph toolbox. Namely, we regard all eigenstates as graph nodes. For each training eigenstate, its corresponding node gets a directed link to nodes on which there are drops in QVAE cost. Then a community detection tool **CommunityGraphPlot** in Wolfram Mathematica is used to list all isolated families with more than one node and bidirectional links.

CHARACTERIZATION OF FAMILIES IN THE PXP MODEL

In this paragraph we further characterize QMBS beyond the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ families detected by the QVAE in the PXP model. For families #1 – #5 we notice an increased overlap with a set of Fock states listed in Table S1. For readability, Fock states are represented by $\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_4, \mathbb{Z}_5$ patterns that appear in them. For example, state denoted as 2 – 2 – 3 – 2 – 3 – 3 – 3 – 3 – 3 corresponds to Fock state $|1010100100100100100\rangle$. To illustrate the increased overlaps with Fock states from Table S1, we calculate probabilities of finding the eigenstates in any of those Fock states,

$$P(i) = \sum_{j=1}^{n_i} |\langle E | f_{ij} \rangle|^2, \quad (\text{S3})$$

name	Fock state pattern	$(\#\mathbb{Z}_2, \#\mathbb{Z}_3, \#\mathbb{Z}_4, \#\mathbb{Z}_5)$
family #1	2-2-3-2-3-3-3-3-3	(3,6,0,0)
	2-3-2-3-2-3-3-3-3	(3,6,0,0)
	2-2-3-3-2-3-3-3-3	(3,6,0,0)
	2-2-3-3-3-2-3-3-3	(3,6,0,0)
family #2	3-3-3-4-3-4-4	(0,4,3,0)
	3-3-4-3-3-4-4	(0,4,3,0)
	3-3-4-3-4-3-4	(0,4,3,0)
family #3	2-2-2-3-2-3-2-2-3-3	(6,4,0,0)
family #4	2-2-2-2-2-2-2-2-3-3	(9,2,0,0)
	2-2-2-2-2-2-2-2-3-3	(9,2,0,0)
	2-2-2-2-2-2-3-2-2-3	(9,2,0,0)
	2-2-2-2-2-2-3-2-2-3	(9,2,0,0)
	2-2-2-2-2-3-2-2-2-3	(9,2,0,0)
family #5	3-4-4-4-4-5	(0,1,4,1)
	3-4-4-4-5-4	(0,1,4,1)
	3-4-4-5-4-5	(0,1,3,2)
	3-4-4-5-4-4	(0,1,4,1)
	3-4-5-3-4-5	(0,2,2,2)
	3-4-5-3-5-4	(0,2,2,2)
	3-4-5-4-9	(0,1,2,1)
	3-4-5-4-3-5	(0,2,2,2)
	3-4-5-4-8	(0,1,2,1)
	3-4-9-3-5	(0,2,1,1)

TABLE S1. List of Fock states with increased overlaps with corresponding QMBS families. Last column contains the number of occurrences of corresponding patterns in the Fock states.

where # i labels the family and $\{|f_{ij}\rangle\}_j$ are Fock states listed next to family # i in Table S1, and plot them in Fig. S4.

We find that for families #1 – #4 the numbers of patterns are conserved, and family #5 gives more than one number of occurrences of each pattern. Family #4 is composed of a maximal number of $\mathbb{Z}_2$ patterns that can be supplemented by $\mathbb{Z}_3$ to fill the whole system with $N = 24$. A similar family was observed for other system sizes: $N = 18$ (6,2,0,0) and $N = 30$ (12,2,0,0) at shifted energies, see Fig. S5. It is at present not clear what is the mechanism behind the conservation of the "domain wall" number within a single family.

INTERPRETING THE QVAE CIRCUIT: INPUT STATE OPTIMIZATION

In the main text we demonstrate that QVAE allows one to find multiple scar families among eigenstates of the PXP model. In the paragraph above we listed physical features specific to the found families, indirectly indicating "what order parameter the QVAE learns". Here, we employ an alternative method to interpret the trained QVAE circuits. Variationally optimizing the input, starting from a random state, we simultaneously minimize the

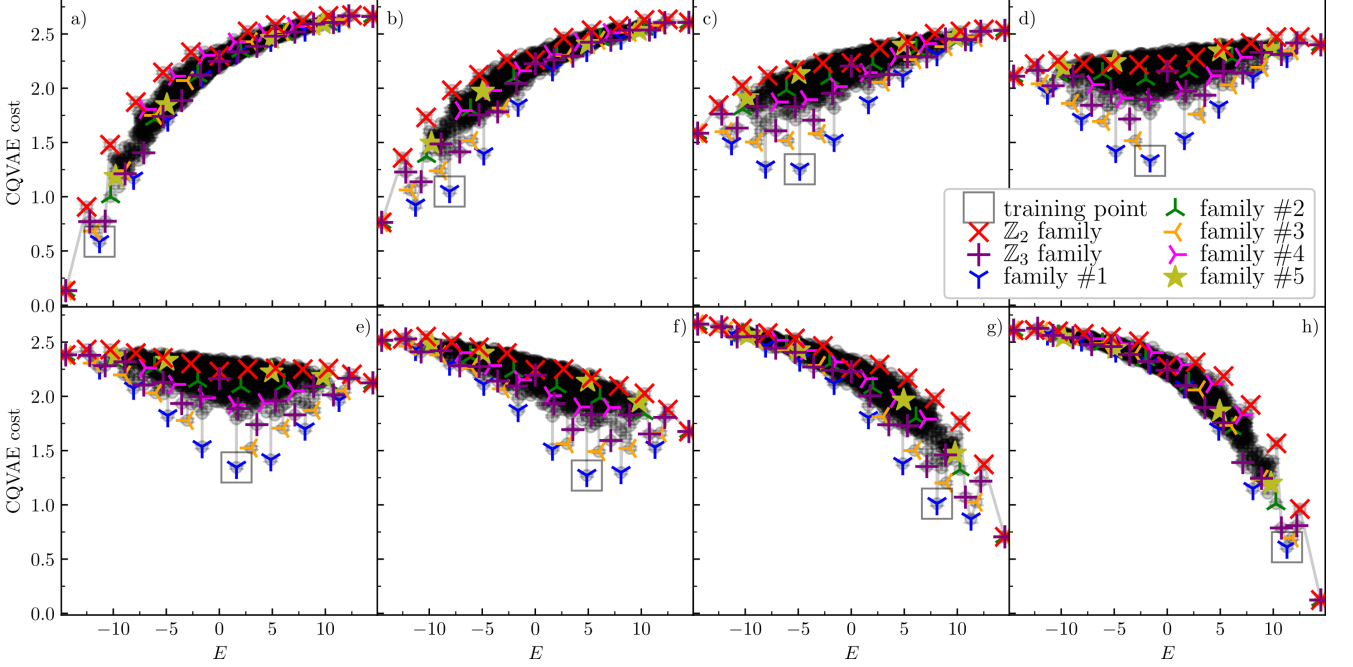


FIG. S2. CQVAE cost vs energy for family #1 for all eigenstates within this family. Above data were used to identify members of family #1.

cost on trash qubits for an ensemble of independently trained QVAEs with constant circuit parameters. This approach is in stark contrast to standard training where the input is given and the circuit parameters are variationally adjusted. We optimize over an ensemble rather than a single QVAE for robustness of the result. Optimization of the input is a popular method of understanding layers in classical deep neural networks working with real-world pictures [10, 11], with a lot of optimization results especially entertaining to look at for humans [12]. Due to the generative nature of the process starting with random noise, input optimization is sometimes called "dreaming" and we adopt this naming here.

In FIG. S3a) we show local magnetization $\langle \sigma_i^z \rangle$ in an example "dreamt" state $|\chi_{dream}\rangle$ found for ensemble of 100 CQVAEs trained on a $\mathbb{Z}_2$ scar eigenstate in the middle of the spectrum. A $\mathbb{Z}_2$ -like magnetization pattern in $|\chi_{dream}\rangle$ can be noticed. It is also visible in the structure factor $S(q) = \sum_j \exp(iqj) \langle \sigma_j^z \rangle$ in FIG. S3b) which has a peak at $q = \pi$. Therefore, even if we assume no prior knowledge that the $|\mathbb{Z}_2\rangle$ state represents the $\mathbb{Z}_2$ scar family, we are able to recover $|\chi_{dream}\rangle \approx |\mathbb{Z}_2\rangle$, serving as a single representative of this family. Unfortunately, since the optimization is done in the input space, due to an exponential increase of computational time with system size, this method is not feasible for large systems. In FIG. S3 we used system size $N = 12$, 50000 input optimization steps, each with 100 shots for cost estimation. Respective CQVAEs consisted of $L = 2$ layers, $N_{trash} = 4$ trash qubits, and were trained with $N_{shots} = 100$ shots

for cost estimation and $N_{iter} = 1000$ parameter optimization steps.

PXP-LIKE MODEL WITH A THREE-BODY BLOCKADE

To demonstrate the universality of our method, we also apply it to a modified PXP model with another type of blockade [13, 14]. The Hamiltonian reads

$$\hat{H} = \sum_i \hat{P}_{i-1,i,i+1} \sigma_i^x, \quad (S4)$$

where $\hat{P}_{i-1,i,i+1} = 1 - |1_{i-1}1_i1_{i+1}\rangle \langle 1_{i-1}1_i1_{i+1}|$. This is a weaker constraint than in the PXP model. QVAE, working in the full Hilbert space, $N_{trash} = 5$, $L = 10$, applied to this model with $N = 16$ detects scars of the $\mathbb{Z}_2$ type, see Fig. S6. No other families have been found by the algorithm.

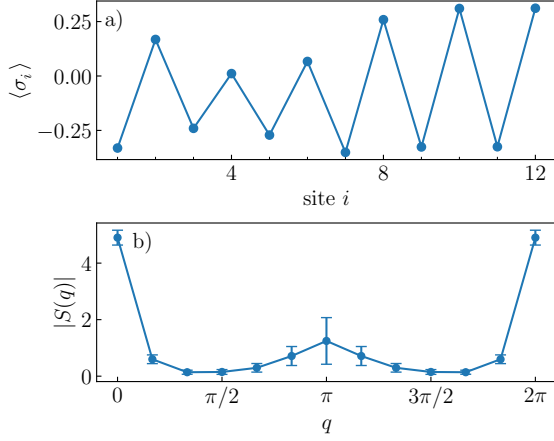


FIG. S3. "Dreaming": optimizing the input state to minimize the QVAE cost (see text) simultaneously for 100 QVAEs trained on a single $\mathbb{Z}_2$ scarred eigenstate, $N = 12$. We get a "representative" state $|\chi_{dream}\rangle \approx |\mathbb{Z}_2\rangle$ for the scar family. a) Local magnetization of the "dream" state $|\chi_{dream}\rangle$ shows a $\mathbb{Z}_2$ -like pattern. b) Structure factor of the calculated magnetization (averaged over 15 independent optimizations) has a peak at $q = \pi$, signaling a period-2 ordering in the position space.

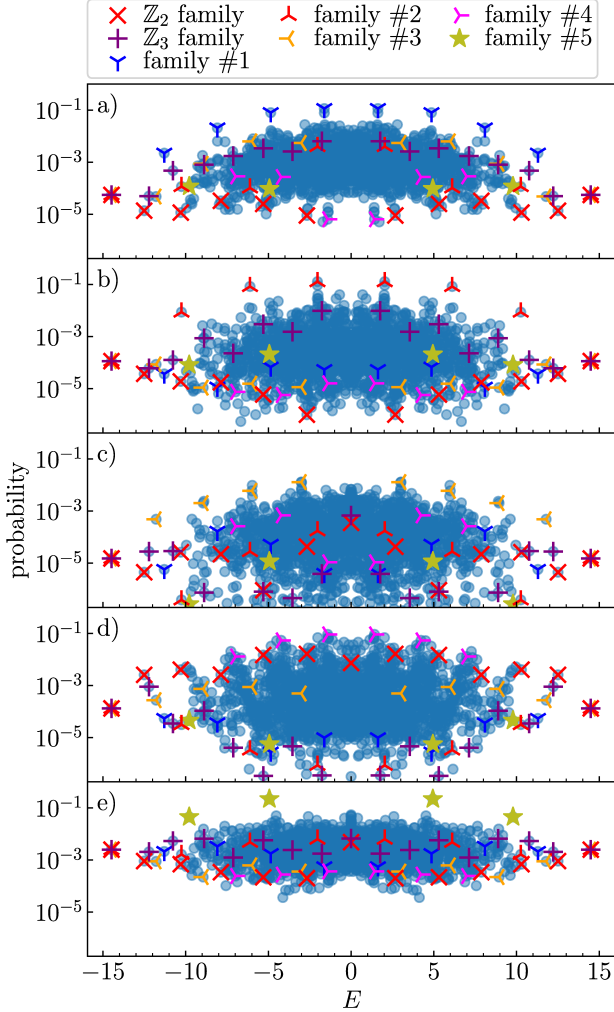


FIG. S4. Probability (S3) of finding eigenstate in any of the Fock states the scarred family #1 – #5 (a-e, respectively) have a large overlap with, see Table S1.

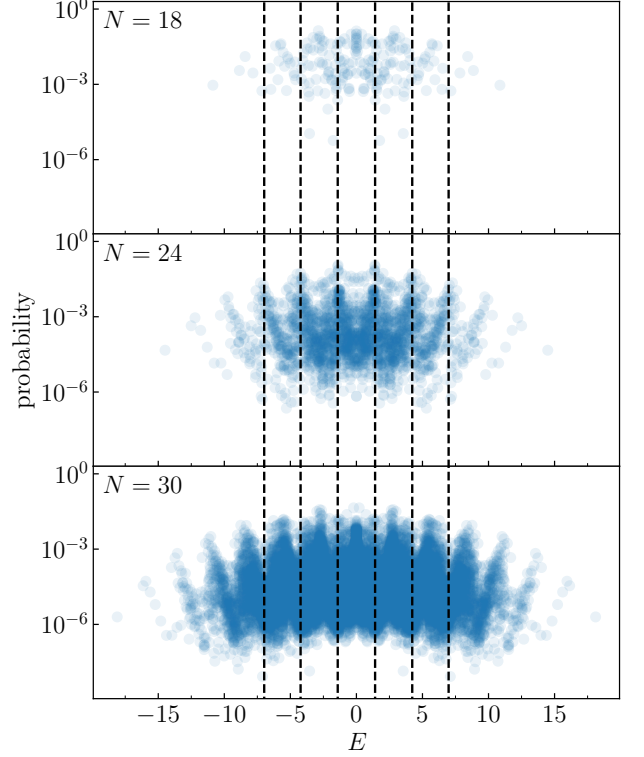


FIG. S5. Probability (S3) of finding eigenstate in Fock states with (6,2), (9,2), (12,2) $\mathbb{Z}_2$ and $\mathbb{Z}_3$ patterns for $N = 18, 24, 30$, respectively. Energies of family #4 found for $N = 24$ by CQVAE are denoted by dashed vertical lines.

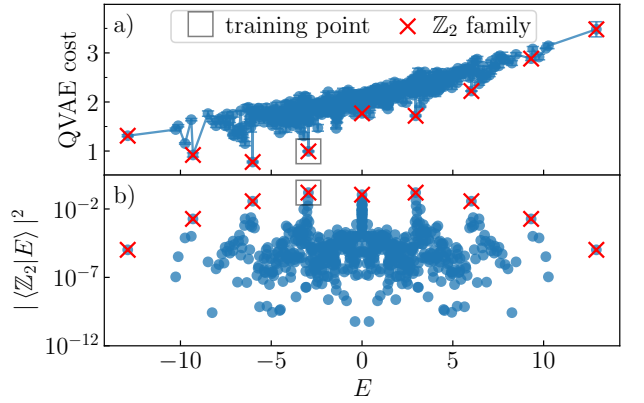


FIG. S6. $\mathbb{Z}_2$ scars detected by the QVAE in the variant of the PXP model with a three-body constraint.

-
- [S1] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, *Nature Communications* **12**, 1791 (2021).
- [S2] K. Nakaji and N. Yamamoto, Expressibility of the alternating layered ansatz for quantum computation, *Quantum* **5**, 434 (2021).
- [S3] J. Spall, Accelerated second-order stochastic optimization using only function measurements, *Proceedings of the 36th IEEE Conference on Decision and Control* **2**, 1417 (1997).
- [S4] J. C. Spall, An overview of the simultaneous perturbation method for efficient optimization, *Johns Hopkins Apl Technical Digest* **19**, 482 (1998).
- [S5] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature* **549**, 242–246 (2017).
- [S6] M. S. ANIS, Abby-Mitchell, H. Abraham, AduOffei, R. Agarwal, G. Agliardi, M. Aharoni, I. Y. Akhalwaya, G. Aleksandrowicz, and T. A. et al., Qiskit: An open-source framework for quantum computing (2021).
- [S7] P. Weinberg and M. Bukov, QuSpin: a Python Package for Dynamics and Exact Diagonalisation of Quantum Many Body Systems part I: spin chains, *SciPost Phys.* **2**, 003 (2017).
- [S8] P. Weinberg and M. Bukov, QuSpin: a Python Package for Dynamics and Exact Diagonalisation of Quantum Many Body Systems. Part II: bosons, fermions and higher spins, *SciPost Phys.* **7**, 20 (2019).
- [S9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016) <http://www.deeplearningbook.org>.
- [S10] D. Erhan, Y. Bengio, A. Courville, and P. Vincent, Visualizing higher-layer features of a deep network, *Technical Report, Univeristé de Montréal* (2009).
- [S11] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, Understanding neural networks through deep visualization (2015).
- [S12] C. Olah, A. Mordvintsev, and L. Schubert, Feature visualization, *Distill* 10.23915/distill.00007 (2017), <https://distill.pub/2017/feature-visualization>.
- [S13] S. Dooley and G. Kells, Extreme many-body scarring in a quantum spin chain via weak dynamical constraints, *Phys. Rev. B* **105**, 155127 (2022).
- [S14] J.-Y. Desaulles, K. Bull, A. Daniel, and Z. Papić, Hypergrid subgraphs and the origin of scarred quantum walks in the many-body hilbert space (2021), arXiv:2112.06885 [quant-ph].