Many-body (de)localization in large quantum chains

Elmer V. H. Doggen,[1, *] Frank Schindler,[2] Konstantin S. Tikhonov,[1, 3]
Alexander D. Mirlin,[1, 3, 4, 5] Titus Neupert,[2] Dmitry G. Polyakov,[1] and Igor V. Gornyi[1, 3, 4, 6]

[1]*Institut für Nanotechnologie, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
[2]*Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[3]*L. D. Landau Institute for Theoretical Physics RAS, 119334 Moscow, Russia*
[4]*Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*
[5]*Petersburg Nuclear Physics Institute, 188300 St. Petersburg, Russia*
[6]*A. F. Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia*

We theoretically study the quench dynamics for an isolated Heisenberg spin chain with a random on-site magnetic field, which is one of the paradigmatic models of a many-body localization transition. We use the time-dependent variational principle as applied to matrix product states, which allows us to controllably study chains of a length up to $L = 100$ spins, i.e., much larger than $L \simeq 20$ that can be treated via exact diagonalization. For the analysis of the data, three complementary approaches are used: (i) determination of the exponent β which characterizes the power-law decay of the antiferromagnetic imbalance with time; (ii) similar determination of the exponent β_Λ which characterizes the decay of a Schmidt gap in the entanglement spectrum, (iii) machine learning with the use, as an input, of the time dependence of the spin densities in the whole chain. We find that the consideration of the larger system sizes substantially increases the estimate for the critical disorder W_c that separates the ergodic and many-body localized regimes, compared to the values of W_c in the literature. On the ergodic side of the transition, there is a broad interval of the strength of disorder with slow subdiffusive transport. In this regime, the exponents β and β_Λ increase, with increasing L, for relatively small L but saturate for $L \simeq 50$, indicating that these slow power laws survive in the thermodynamic limit. From a technical perspective, we develop an adaptation of the "learning by confusion" machine learning approach that can determine W_c.

I. INTRODUCTION

Many-body localization (MBL) refers to the localization of particles by disorder in the presence of interactions at non-zero energy density (for recent reviews, see [1–3]), as opposed to the conventional Anderson localization [4] which describes non-interacting particles in the presence of disorder. MBL is of fundamental interest for understanding metal-insulator transitions and disordered superconductors. From a numerical perspective, MBL is a notoriously difficult problem to describe because of its many-body nature, sensitivity to finite-size effects and the requirement for ensemble averaging over many realizations of disorder. On a conceptual level, the study of MBL ties to the thermalization of quantum systems and the bridge between microscopic dynamics and quantum statistics [5].

The study of MBL did not begin in earnest until the development of landmark theories [6, 7] predicting a temperature-driven transition to a localized phase, now known as an MBL phase. At the same time, advances in the increase of computational power and the development of new algorithms based on tensor networks like matrix product states (MPS) have dramatically accelerated the numerical study of disordered interacting systems, albeit thus far restricted to ground states, short times, boundary-driven systems or very strong disorder

[8–12]. Experimentally, MBL has been reported in ultra-cold atoms [13, 14], trapped ions [15], and dipolar spins in diamond [16]. Analytically, the existence of an MBL region in the phase diagram of a disordered quantum spin chain has been proven based on an assumption about energy level repulsion [17].

In this work, we investigate a Heisenberg chain with isotropic interactions and a random magnetic field, which has become one of the paradigmatic models for the investigation of the MBL-related physics. In particular, Ref. [18] considered system sizes L from $L = 8$ to $L = 16$ and found crossing points in level statistics and relaxation of spin modulations plotted as a function of disorder W (in notations of the present work) for various system sizes. By analogy with the scaling analysis of the Anderson transition in non-interacting systems, these crossing points serve as an indication of the MBL transition. However, a rather strong drift of the crossing points was found, from $W \simeq 2$ to $W \simeq 3$, as L increased from 8 to 16, making it difficult to locate the transition. In a later work [19], systems of larger sizes, up to $L = 22$, were investigated via exact diagonalization (for the current state-of-the-art, see Ref. [20]), and the estimate for the transition point $W_c \simeq 3.7$ was obtained, consistent with the upper bound $W_c \lesssim 4$ found in Ref. [21]. On the other hand, Ref. [22] used a numerical linked-cluster expansion for the entanglement entropy in the thermodynamic limit and obtained a lower bound $W_c \gtrsim 4.5$ which appears to be in conflict with the exact-diagonalization results quoted above.

---

* Corresponding author: elmer.doggen@kit.edu

All in all, the current status of the numerically obtained results is still controversial; simulations of small systems can easily miss physics relevant at longer length scales even if ameliorated by finite-size scaling approaches. Most importantly, the strong dependence of the apparent transition point on the system size calls for a detailed numerical study of physical observables characterizing the MBL transition in systems of size L that is well beyond the reach of exact diagonalization. To achieve this goal, we consider the MPS-based approach, which allows us to controllably study the long-time quench dynamics around the MBL transition in spin chains of length up to $L = 100$. We also apply machine-learning techniques to further analyze the data obtained from MPS.

Clearly, not only the position of the MBL transition is of interest but also a more detailed insight into the physics of phases around it. Previous works on one-dimensional (1D) systems indicated that a part of the delocalized phase adjacent to the transition is characterized by a slow, subdiffusive dynamics [23–27]. However, it was also found [28], within exact-diagonalization studies of systems with size up to $L = 24$, that an apparent exponent characterizing this subdiffusive phase changes substantially with the system size. This poses a question of whether the subdiffusive behavior is a genuine property of the system in the $L \to \infty$ limit, or merely a transient feature which characterizes the relatively small systems. To answer this question, it is important to numerically study substantially larger systems. This is another motivation for the present investigation of the MBL physics in a Heisenberg chain within the MPS-based approach.

II. MODEL AND METHOD

A. Disordered Heisenberg XXZ chain

We consider the Heisenberg XXZ chain with an on-site random field on a lattice of length L with open boundary conditions, as described by the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{L} \left[\frac{J}{2} \left(S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+ \right) + \Delta S_i^z S_i^{z+1} + h_i S_i^z \right]. \tag{1}$$

Here S_i^+, S_i^-, and S_i^z are the standard spin-1/2 Pauli operators corresponding to the site i, and the on-site field h_i takes random values according to a uniform distribution $h_i \in [-W, W]$. In the following, we set $J, \hbar \equiv 1$ as a choice of units, and put $\Delta = 1$, i.e., we consider the isotropic Heisenberg chain, unless stated otherwise. Using a Jordan-Wigner transformation, this model can be mapped to the model of nearest-neighbor interacting hard-core bosons with an on-site disordered potential. We consider the zero spin sector $\sum_i \langle S_i^z \rangle = 0$, which corresponds to half-filling in the particle picture. The Anderson model of non-interacting particles in a disordered potential is retrieved for $\Delta = 0$. The model (1) was used

to discuss the MBL transition and the subdiffusive behavior in the delocalized phase, see the references above.

B. Time-dependent variational principle

To study the model (1), we employ a recently developed numerical method for describing the time evolution of 1D lattice systems, which is based on the Dirac-Frenkel time-dependent variational principle (TDVP) [29] as applied to MPS [30, 31]. Contrary to traditional time-evolution algorithms, such as the time-dependent density matrix renormalization group (t-DMRG) [32, 33] or time-evolving block decimation (TEBD) [34], the TDVP does not rely on a Suzuki-Trotter decomposition of local Hamiltonian terms. Instead, the time-dependent wave function $|\psi\rangle(t)$ is given by

$$\frac{d|\psi\rangle}{dt} = -i\mathcal{P}_{\text{MPS}}\mathcal{H}|\psi\rangle, \tag{2}$$

where $\mathcal{P}_{\text{MPS}}$ projects the time-evolved wave function back onto the variational MPS manifold, with a dimension typically much smaller than the dimension of the complete Hilbert space 2^L.

From the perspective of accuracy of MPS simulations, the worst-case scenario for a local Hamiltonian is realized when the von Neumann entanglement entropy for a bipartition of the chain into two parts grows linearly [35], which is known as *volume-law entanglement* and should be contrasted to the *area-law entanglement* characteristic of localized systems [36]. It is well known that in such a worst-case scenario, the required bond dimension χ (which controls the dimension of the variational manifold) grows exponentially in time if it is required that the truncation error is kept below some finite value. This provides a fundamental limit to the maximum time reached using MPS simulations under the condition that $\langle \psi_{\text{MPS}} | \psi_{\text{exact}} \rangle \approx 1$ [37]. Within the traditional DMRG framework, recent developments have extended the maximum possible times significantly [38], although the fundamental limit remains. On the other hand, under certain conditions it is expected [39–41] that the TDVP can provide reasonable estimates for the transport properties despite a potentially large truncation error. A key advantage of the method is that unlike t-DMRG and TEBD, it conserves the total energy by construction.

In the model (1), an MBL regime characterized by area-law entanglement has been identified for strong disorder (large W). On the other hand, for weak disorder (small W) an ergodic regime has been reported, described by the eigenstate thermalization hypothesis (ETH) [42, 43] and exhibiting volume-law entanglement. In the present work, we focus on the range of sufficiently strong disorder $W \geq 2$, which includes the whole MBL phase as well as a part of the delocalized phase. As we have verified (see detailed explanations below), the MPS approach works reliably up to long times considered in our study not only in the MBL phase but also in the
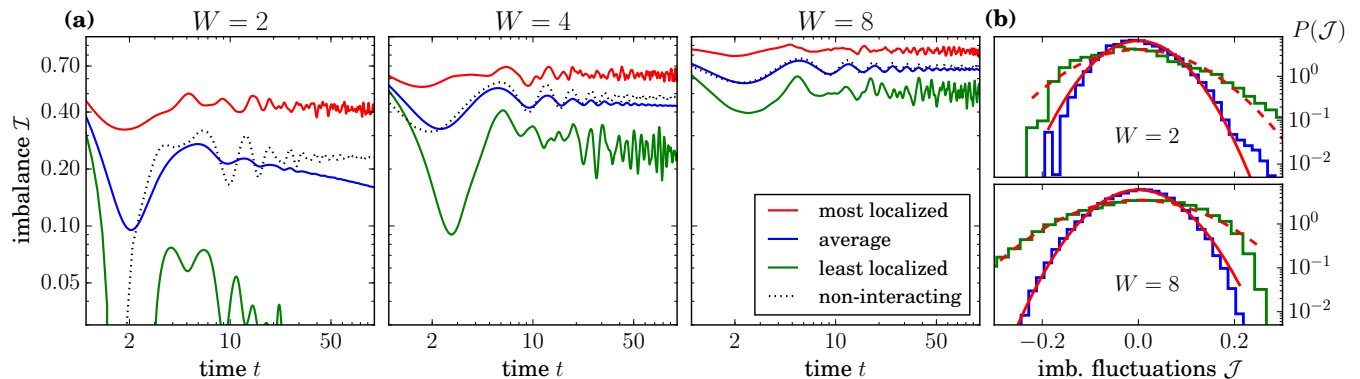
FIG. 1. **(a)** Spin imbalance (3) as a function of time t for a spin chain of length $L = 100$ and various disorder strengths $W = 2, 4, 8$. The blue line shows the disorder average over $R = \mathcal{O}(10^3)$ realizations and the dotted black line shows the non-interacting case $\Delta = 0$ using 10000 independent disorder realizations. The red (green) line shows the most (least) localized single realization. **(b)** Probability distribution function of the fluctuations of the imbalance around the average $\mathcal{J}(t) \equiv \mathcal{I}(t) - \overline{\mathcal{I}(t)}$ for $W = 2$ (top), $W = 8$ (bottom), $L = 50$ (green) and $L = 100$ (blue) over the time interval $t \in [50, 100]$. The solid (dashed) red line shows a Gaussian fit for $L = 100$ ($L = 50$).

delocalized phase with $W \geq 2$ since the corresponding dynamics is very slow.

III. NUMERICAL RESULTS (TDVP)

In this section, we will detail results obtained using the direct analysis of the TDVP data for the quench dynamics in disordered Heisenberg chains (1) of lengths $L = 16$, 50 and 100 with disorder strength ranging from $W = 2$ to $W = 8$.

A. Imbalance

We follow the time dynamics of an initial unentangled (product) Néel state $|\psi\rangle = \{\uparrow, \downarrow, \ldots\}$, where we consider many different realizations $R \gg 1$ of disorder. Whether the system is in the delocalized or localized phase can be quantified using the *imbalance*:

$$\mathcal{I}(t) = \frac{1}{L} \sum_{i=1}^{L} (-1)^i \langle S_i^z(t) \rangle, \qquad (3)$$

a quantity that measures how much of the initial anti-ferromagnetic order remains at time t [44]. By definition, $\mathcal{I}(0) = 1$, and for a fully ergodic, thermalized state, $\mathcal{I}(t) \to 0$ at long times on average, which means that the system completely loses memory of the initial anti-ferromagnetic order. The disorder average is denoted as $\overline{\mathcal{I}(t)} \equiv (1/R) \sum_r \mathcal{I}_r(t)$, where $r = \{1, \ldots, R\}$ labels the different realizations.

The imbalance is appealing from several perspectives. It is a global characteristic of the system which signifies its degree of localization, but at the same time it is computed from purely local quantities (average spins). In

view of this locality, the imbalance can be readily measured in an experimental setting using cold atoms [13]. We can also compare directly the localization properties of an interacting system to those of non-interacting Anderson insulators. The non-interacting value $\mathcal{I}_\mathrm{A}(t)$ is obtained by computing the time evolution of the Néel state through the exact diagonalization of (1) with $\Delta = 0$ (cf. Ref. [21]). In Fig. 1, we show the dynamics of imbalance for both the interacting and non-interacting cases using the parameters $L = 100$, $W = \{2, 4, 8\}$.
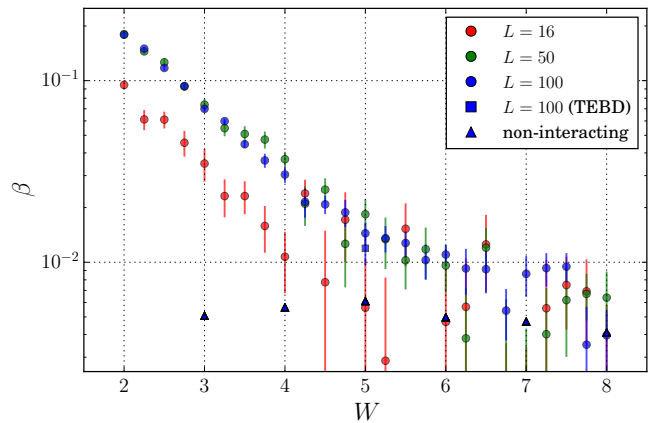

FIG. 2. Power-law exponent β corresponding to the decay of the imbalance $\mathcal{I} \propto t^{-\beta}$ over the window $t \in [50, 100]$, for various system sizes L. For the cases $L = 50$ and 100, the 10 sites near each edge of the chain were not considered. The square symbol shows the result for an independent implementation of the time-evolving block decimation (TEBD) algorithm [45]. Triangles indicate a weak finite-time decay of $\mathcal{I}_\mathrm{A}$ for the non-interacting case $\Delta = 0$, $L = 100$. Error bars are 1σ-intervals based on a bootstrapping procedure.

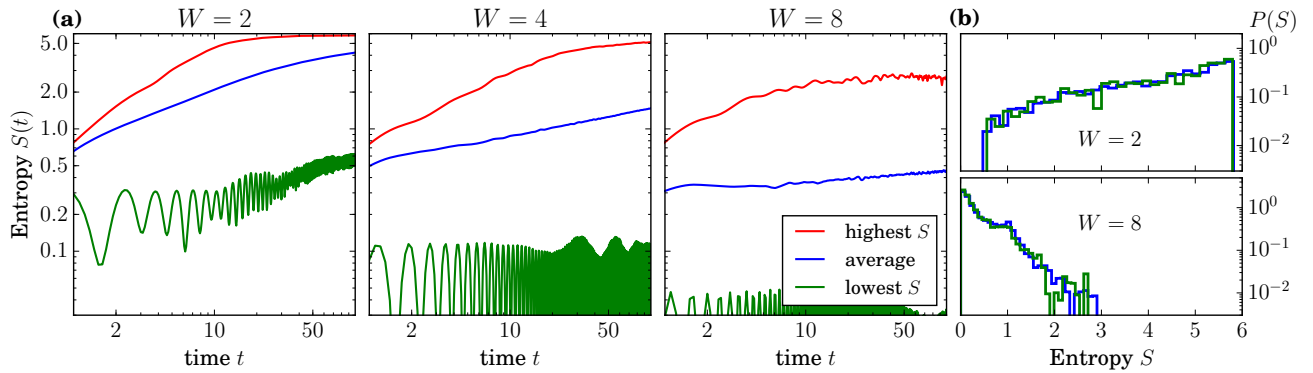For a thermalized system, a power-law decay $\mathcal{I}(t) \propto$

FIG. 3. **(a)** The von Neumann entropy of entanglement S as a function of time for the same parameters as in Fig. 1. **(b)** Distributions of S over the interval $t \in [95, 100]$. The blue (green) lines show the result for $L = 100$ ($L = 50$).

$t^{-\beta}$ is expected [44]. In fact, within the Boltzmann equation, one would get an exponential decay with time, since the imbalance corresponds to a mode with a large wave vector q (namely, $q = \pi$) describing, in contrast to the total spin, a non-conserved quantity. However, taking into account the coupling of this mode to the low-q diffusive (or subdiffusive) mode associated with the spin density spreading suggests a power-law decay (by analogy with the long-time tails related to the return probability that are found for other observables). Indeed, previous numerical results indicated that the exponent β of the imbalance decay is qualitatively similar to the subdiffusion exponent found from the return probability, the mean-square displacement of a spin excitation, and the low-frequency dependence of the conductivity (see Ref. [44] and references therein).

In Fig. 2, we show the exponent β computed numerically using a least-squares fitting algorithm for various values of W and L from the imbalance at sufficiently long times $t \in [50, 100]$. We choose $W = 2$ as the weakest disorder to be considered. This is based on the requirement that the value $\chi = 64$ of the bond dimension used in our computations is sufficient to ensure that the value of β is insensitive to a further increase of χ (see Appendix A). As a further check, we compute β using a completely independent implementation of the time-evolving block decimation (TEBD) algorithm for $W = 5$, $L = 100$, $\chi = 64$, finding agreement within error bars. In addition, for a few values of $W \geq 4$ and for the system size $L = 100$ we follow the time evolution until $t = 300$ using $\chi = 32$ and find a good agreement with β obtained in the window $t \in [50, 100]$ (not shown), as well as checking that for bond dimension up to $\chi = 96$, β is insensitive to increasing the time window up to $t = 200$ for $L = 50$ and $W = 5$ (see Appendix A). The number of different disorder realizations for each choice of W and L is typically of the order of $500 - 1000$.

For small systems that can be treated by exact diagonalization (such as $L = 16$), we find a good agreement with previous results [27]. Specifically, the value of β, which continuously decreases with increasing W, van-

ishes within error bars at $W \approx 4$. However, an increase of the system size from $L = 16$ to $L = 50$ leads to a substantial increase of β, consistent with a trend observed for relatively small systems in Ref. [28]. As a result, the value of disorder for which β vanishes also increases [46]. On the other hand, the results for β obtained for $L = 50$ and $L = 100$ agree within error bars. We have further diminished finite-size effects by excluding the 10 sites on each far end of the chain for the cases $L = 50, 100$ for the computation of $\mathcal{I}$, so that we only consider the bulk of the system.

We thus conclude that, after the initial increase for $L \lesssim 50$, the exponent β saturates, i.e., essentially reaches its thermodynamic-limit ($L \to \infty$) value. This implies two important conclusions. First, the slow, subdiffusive transport appears to be a genuine property of a long chain on the ergodic side of the MBL transition and not just a finite-size effect. Second, the estimated value of the critical disorder is $W_c \gtrsim 5$, i.e., considerably larger than the values suggested by the exact-diagonalization analysis of small systems. This estimate is a lower bound determined by considering the lower end of our error estimate with two standard deviations, plus a small systematic error shown as triangle symbols in Fig. 2 [47]. For weaker disorder, we can confidently conclude the system is ergodic.

In order to understand how representative the average value of the imbalance is, we have also studied its fluctuations, $\mathcal{J}(t) \equiv \mathcal{I}(t) - \overline{\mathcal{I}(t)}$. The root-mean-square amplitude of the fluctuations of the imbalance, $\sqrt{\overline{\mathcal{J}^2(t)}}$, has been found to be very slowly varying with time for long times. This allows us to consider the probability density function $P[\mathcal{J}(t)]$, where $t \in [50, 100]$. In this window, $\mathcal{J}(t)$ is, to an excellent approximation, Gaussian-distributed (see Fig. 1b), with a standard deviation σ roughly proportional to $1/\sqrt{L}$. This implies self-averaging of the imbalance. A similar behavior is found for a non-interacting, Anderson-localized system. The wide distributions for smaller system sizes complicate the accurate determination of β around the transi-
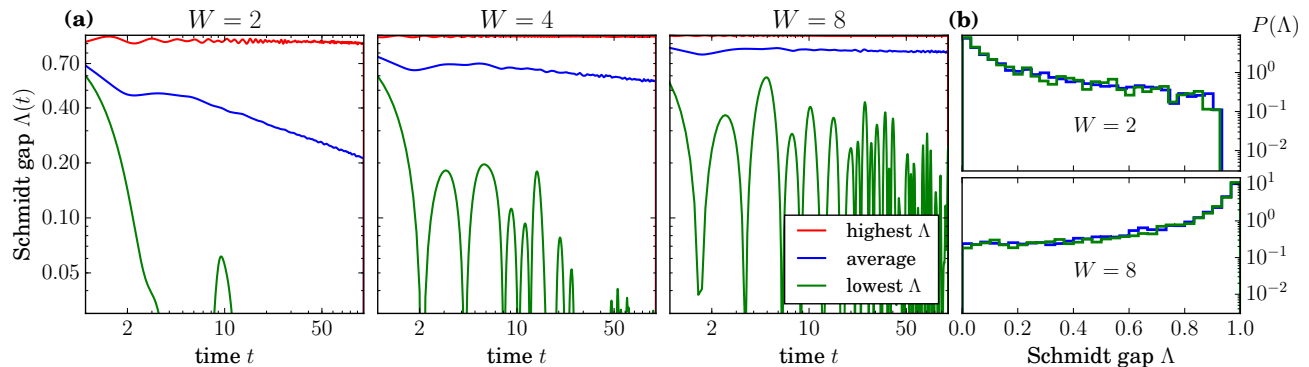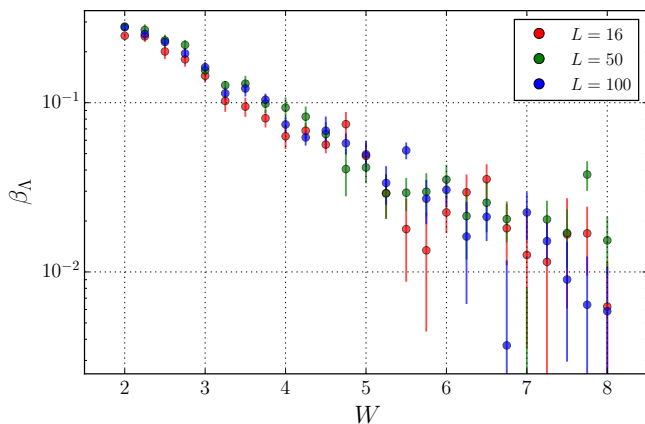
FIG. 4. **(a)** The Schmidt gap Λ as a function of time for the same parameters as in Fig. 1. **(b)** Distributions of Λ over the interval $t \in [95, 100]$. The blue (green) lines show the result for $L = 100$ ($L = 50$).

FIG. 5. Power-law exponent β_Λ corresponding to the decay of the disorder-averaged Schmidt gap $\Lambda \propto t^{-\beta_\Lambda}$ over the window $t \in [50, 100]$, for various system sizes L. Error bars are 1σ-intervals based on a bootstrapping procedure.

tion, so that considering larger system sizes is very beneficial also from this point of view.

B. Entanglement and Schmidt gap

The entropy of entanglement characterizes the spread of correlations in the system. On the ergodic side, power laws have been predicted for the growth of entanglement close to the transition [27]: $S(t) \propto t^{\beta_S}$. Another quantity of interest is the so-called Schmidt gap Λ, which is defined as the difference between the two largest values of the entanglement spectrum (see Ref. [36] and references therein) for a bipartition of the system into subsystems A and B: $\Lambda = \lambda_1 - \lambda_2$. Here the entanglement eigenvalues $\lambda_i \geq 0$ [48] are the eigenvalues of the effective entanglement Hamiltonian H_e defined through the reduced density matrix of the system after tracing out the degrees of freedom of the subsystem B: $H_e = -\ln \text{Tr}_B \rho$ [49].

To explain the connection between the Schmidt gap

and the more frequently used von Neumann entropy of entanglement, we recall that the latter can be expressed in terms of the entanglement eigenvalues λ_i as follows :

$$S = -\sum_i \lambda_i \log_2 \lambda_i \simeq -\sum_{i=1}^\chi \lambda_i \log_2 \lambda_i. \qquad (4)$$

Here the eigenvalues λ_i are ordered in a descending way by convention. The approximate equality holds as long as the entanglement remains relatively low. For our problem, we find that the effect of the cutoff by the bond dimension χ is negligible for $W \gtrsim 4$ from the perspective of S (see Fig. 3).

Contrary to the entanglement entropy, the evaluation of the Schmidt gap requires knowledge of only the first two entanglement eigenvalues and hence is less sensitive to the value of the bond dimension. A further appealing property of this quantity is that, for a thermalized system, $\Lambda(t) \to 0$ for $t \to \infty$, while for a localized system it is expected that the Schmidt gap remains finite in the long-time limit, $\Lambda(t) \to \mathcal{O}(1)$, or at most decays logarithmically. In this sense, the $t \to \infty$ behavior of the Schmidt gap allows one to distinguish between the localized and delocalized phases in the same way as for the imbalance. The entanglement entropy and Schmidt gap for $L = 100$ and representative choices of W are depicted in Figs. 3 and 4, respectively, which also show the distributions of these quantities close to the final time $t = 100$ considered here. The shape of the distributions found for the entropy of entanglement is in qualitative agreement with a detailed study of such distributions for eigenstates of small systems using exact diagonalization [50]. That these distributions are converged with respect to system size reflects the slow spread of entanglement close to the transition (even on the ergodic side).

We have determined the Schmidt gap by using a bipartition in the middle of the chain. Interestingly, we find that the averaged Schmidt gap shows in the delocalized phase a power-law decay similar to that as $\mathcal{I}(t)$. The corresponding power-law exponent β_Λ is shown in Fig. 5. The results demonstrate a striking qualitative similarity
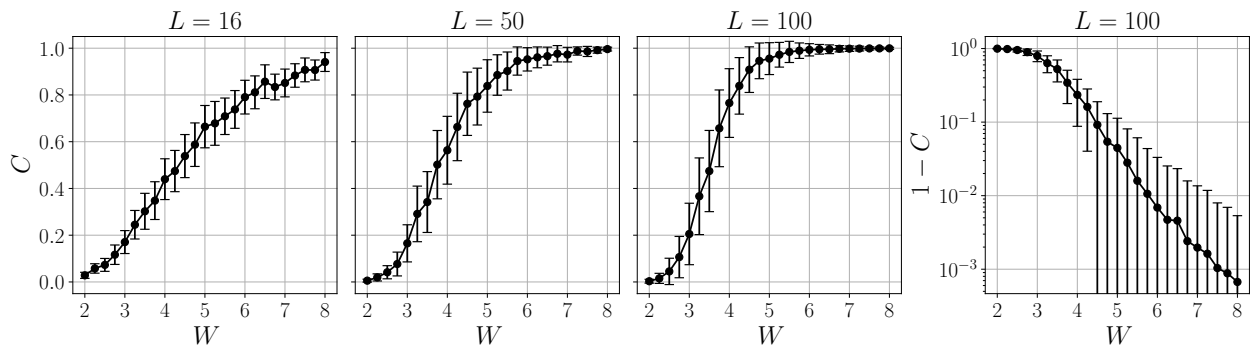
FIG. 6. Results of the supervised machine learning algorithm. Shown is the average confidence C with which time-series (spin densities $\langle S^z \rangle$ as a function of time) taken from TDVP for a given disorder magnitude are classified as belonging to the localized phase, where we labeled $W = 8$ as localized and $W = 2$ as delocalized. A plateau emerges at large W for $L = 100$, indicating a transition to the MBL regime. The rightmost panel shows $1 - C$ on a semi-log scale. Error bars are 1σ-intervals based on 100 independent training sessions (these should be understood as being cut off at $C = 0$ and $C = 1$).

to those for the imbalance exponent β, Fig. 2. However, the found large-L values of β_Λ are somewhat above those for β. The corresponding estimate for the critical disorder of the MBL transition obtained as a point where the Schmidt gap exponent β_Λ vanishes within error bars is $W_c \simeq 6$, i.e., even larger than $W_c \simeq 5$ found from results for the imbalance. For the entropy of entanglement S, however, we clearly see the effect of the cutoff at $S = 6$ imposed by the bond dimension $\chi = 64$ (see Fig. 4a, left panel), where the power-law behavior is disturbed around $t \approx 30$. Considering that the transport properties have nonetheless converged with χ, this indicates that the TDVP can indeed provide reasonable results even if there are significant cutoff effects in terms of the entropy. Moreover, the Schmidt gap, which still shows a clean power-law behavior at $W = 2$, appears to be less sensitive to this numerical cutoff.

IV. MACHINE LEARNING

To further corroborate our analysis, we apply methods from machine learning [51], which has emerged recently as a powerful tool to analyze localization phenomena [52–57], to our data obtained using the TDVP. We use two algorithms: a partially supervised approach that has previously been employed in Ref. [52], and a fully unsupervised method based on the "learning by confusion" scheme introduced in Ref. [58]. The combination of traditional numerical analysis and machine learning is mutually reinforcing: understanding localization though machine learning amounts to learning machine learning through localization.

A. Supervised classification algorithm

For the supervised learning approach, we train a feedforward neural network to distinguish data at two ex-

tremes of our data set: $W = 2$ (delocalized) and $W = 8$ (presumed to be localized). We choose a single hidden layer network with a ReLU activation function for the hidden layer (of size ~ 10) and a Softmax activation function for the output layer (of size 2, corresponding to the two classes we are distinguishing). As input data, we use time-series $\langle S_i^z(t) \rangle$, $i = 1 \ldots L$, evaluated at equidistant time-steps $t \in [50, 51, \ldots, 100]$ taken from the full TDVP time evolution at $W = 2$ and $W = 8$. We choose a simple cross entropy error function with ℓ_2 regularization, cf. Ref. [52].

After convergence of the training set error, we apply the trained network on a test set containing data at $W = 2$ and $W = 8$, finding a classification accuracy of larger than 99%. We subsequently apply the trained network to TDVP time-series at intermediate disorder strengths $W = 2.25, 2.5, \ldots, 7.75$, to determine the average confidence C with which data corresponding to a given disorder strength is classified as either belonging to the delocalized ($W = 2$) or localized ($W = 8$) class of time-series we trained with. Successful training implies that $C = 0$ for $W = 2$ and $C = 1$ for $W = 8$, while $C \in [0, 1]$ at intermediate disorder strengths quantifies how similar the time evolution is to the extreme values $W = 2$ and $W = 8$. The results are shown in Fig. 6.

This approach clearly indicates the formation of a plateau for $L = 100$, suggesting the presence of a many-body localized phase over a range of disorder strengths, but its thermal counterpart gives way to an extended crossover region. In contrast, no plateau or a less pronounced one is present in the data for system sizes $L = 16$ and $L = 50$, respectively, showing the importance of considering larger systems.

These results can be interpreted in analogy with the analysis of the power-law exponents β and β_Λ: over a range of disorder strengths, the imbalance does not decay and the spin densities "look similar" to the case $W = 8$, down to the critical disorder. However, the algorithm can still pick up differences between a large value of the power law and a small one, leading to a non-zero C even
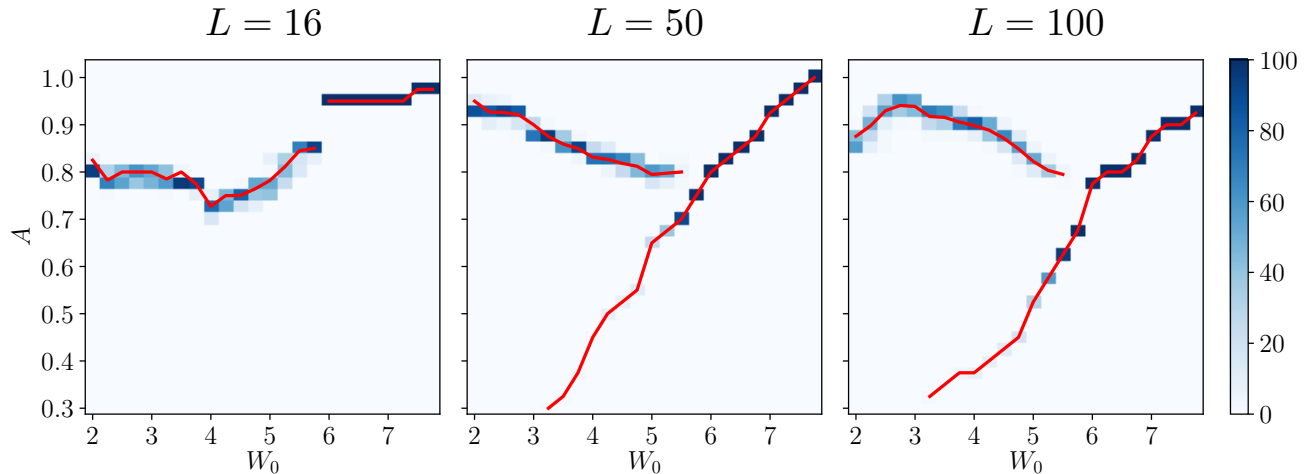
FIG. 7. Results of the unsupervised machine learning algorithm, based on "learning by confusion" for input data with $L = 16$, $L = 50$, and $L = 100$. Shown is the distribution of the accuracy $\langle A(W_0) \rangle$ over 100 training instances with different initial conditions for each W_0. For system sizes $L = 50$ and $L = 100$, the distribution is bimodal in the interval $3.25 < W_0 < 5.5$ with the two maxima tracing out the red lines as a function of W_0.

in the delocalized regime, so that the behavior of β and $1 - C$ is qualitatively similar (see Fig. 6, rightmost panel). The plateau, within error bars, starts at $W \gtrsim 4.5$ [59]. Hence, one should not expect the shape of the curves to approach a step-like function in the limit $L \to \infty$ (the sharpening with system size can be associated with the narrowing distributions $\mathcal{J}$, see Fig. 1b), in contrast to level statistics [19]. A precise determination of the critical disorder strength, relying on knowledge of the extent of the "plateau" in the limit $L \to \infty$, is therefore difficult to obtain using this approach, and placing the transition for instance at the midpoint $C = 1/2$ leads to an underestimate for the critical disorder [52]. Nonetheless, in the region where β vanishes within error bars, the supervised learning approach is consistent with the analysis of the decay of the imbalance in the sense that in this region (belonging to the MBL plateau, as defined above) $C \gtrsim 0.99$ for $L = 100$. In the next section, we will detail an unsupervised method that is able to determine W_c more precisely.

B. Unsupervised confusion algorithm

In Ref. [58] a scheme called "learning by confusion" for the unsupervised detection of phase transitions for data ordered along a one-dimensional parameter space (here: the strength of disorder $W \in [W_{\min}, W_{\max}]$) has been proposed. For this, an arbitrary parameter W_0 is fixed and a feed-forward neural network is trained (in a supervised fashion) to distinguish data with $W < W_0$ from data with $W > W_0$. The accuracy $\langle A(W_0) \rangle$ of classification that the trained network achieves when applied to a test set is evaluated. The process is repeated for different choices of W_0. The resulting function $\langle A(W_0) \rangle$ has two global maxima at $W_0 = W_{\min}$ and $W_0 = W_{\max}$, where

the accuracy is trivially 1 because all the data can be classified as belonging to one phase. In Ref. [58] it was observed that an additional local maximum of $\langle A(W_0) \rangle$ occurs when W_0 equals the location of a phase transition W_c, as it would be easiest for the network to classify the data into two sets for this choice of separation. Thus, in presence of a single phase transition as a function of W_0, the curve $\langle A(W_0) \rangle$ is expected to take a W-shape.

We apply this algorithm, using the same neural network architecture and the same type of input data as in the supervised case, while dividing the training data by $W_0 = 2.25, 2.5, \ldots, 8$. Each time we determine the average accuracy $\langle A(W_0) \rangle$ as the percentage of correctly classified sets $\langle S_i^z(t) \rangle$ (with respect to the division W_0) in a test set (which does not coincide with the training data). We observe that the results depend strongly on the initial conditions for the training, i.e., on the random initial choice of weights and biases for the network. Even for input data from the largest system sizes $L = 100$, no consistent W-shape emerges in $\langle A(W_0) \rangle$. This is due to substantial fluctuations of $\langle A(W_0) \rangle$ for fixed W_0 between training runs, in particular around the putative transition region.

The fluctuations of $\langle A(W_0) \rangle$ resulting from the initial conditions of the training are, however, not random. We claim that they carry information that can be used to locate the transition. We observe that over a large region of W_0, they follow a bimodal distribution as shown in Fig. 7. We interpret the two branches as instances of trained networks that identify localized and ergodic features, respectively [60]. The transition should then be identified as the position where both types of networks occur with the same probability, which happens around $W_0 \approx 5$. Viewing the ETH-MBL transition as a crossover (at least from the perspective of the finite-size and finite-time data that serves as input of the neural network),

this analysis puts $W_0 \approx 5.5 - 6$ as an upper bound to the crossover region. Above this value the distribution of $\langle A(W_0) \rangle$ becomes unimodal.

V. SUMMARY AND OUTLOOK

The quench dynamics of large disordered spin chains in the Heisenberg model has been investigated by means of the time-dependent variational principle for matrix product states. We have studied the long-time behavior of the imbalance and found a regime, occurring in a broad range of parameters of the system, with slow, yet finite transport. In terms of averaged transport, we find that the average, typical and median imbalance all lead to the same power-law exponent β. Our results imply that the ergodic regime extends (at least) up to disorder $W_c \simeq 5$. We observe a substantial shift of the exponent β (and, as a result, of an estimate for the MBL transition point W_c) when we go from relatively small systems (that can be exactly diagonalized) to large systems with $L = 50$ and 100. On the other hand, we do not see any significant difference between the results for $L = 50$ and $L = 100$, which indicates saturation of the exponent β. This favors the conclusion that the subdiffusive behavior is a true long-time asymptotic behavior. These findings have been supported by the results for the dynamics of the Schmidt gap in the entanglement spectrum, which shows a very similar behavior.

We have complemented a conventional analysis of the data (with fitting the average values to power laws) by a machine-learning analysis of the whole time dependences of individual spins in various realizations of disorder. We have chosen two approaches, a supervised and a novel unsupervised algorithm. The latter is based on learning by confusion, but also exploits the stochastic nature of the learning process. This approach provides a more accurate way to determine the position of the transition, and could be useful for wider applications in determining phase transitions. The results from both the supervised and unsupervised machine-learning methods support our conclusions concerning the bound for the critical disorder strength $W_c \gtrsim 5$. Thus, we can reinforce machine learning by combining it with a "traditional" analysis of numerical data, verifying that machine learning tools can be reliably applied.

Our results, showing a very slow transport on the ergodic side of the MBL transition ($W < W_c$) support the expectation that the system looks essentially localized at the transition point ($W = W_c$) [18]. From this point of view, the MBL transition has much in common with the Anderson transition on random regular graphs (RRG) [61]. The latter problem is viewed as a toy model of the MBL transition, even though this connection is less precise for short-range interaction models than for those with power-law interactions [62]. A slow dynamics in the RRG model was recently studied numerically in Ref. [63]. From the perspective of the machine-learning approaches

applied in this work, the notion that the transition point is localized in character manifests itself in that it is located near the edge of the "plateau" characterizing the MBL phase, rather than at the midpoint between the two regimes. In terms of the confusion-based algorithm, the transition is located at the crossing between "delocalized" and "localized" branches.

A slow dynamics near the transition and a localized character of the critical point are also qualitatively consistent with the avalanche mechanism of the transition developed recently in Refs. [64–66]. This scenario predicts an extended weakly delocalized regime above the "nominal" threshold $W_c^{(0)}$ and implies that previous numerical results might be tainted by finite-size effects. Specifically, within this scenario, static "ergodic spots" (spatial regions with anomalously weak disorder) embedded in the nominally localized phase [67] thermalize the rest of the system below the true MBL transition at $W_c > W_c^{(0)}$. In small systems, such spots are typically not found and the chains appear to be localized for $W > W_c^{(0)}$; however, with increasing L the probability of finding ergodic spots increases and the true transition at W_c becomes well resolved.

Indeed, we find that considering larger systems leads to a substantial increase in delocalization, in agreement with the avalanche scenario. Moreover, the unsupervised machine-learning analysis, showing the coexistence of localizing and delocalizing realizations of disorder in the range $3.5 \lesssim W \lesssim 5.5$ (see Fig. 7), may also be interpreted in terms of the avalanche-induced delocalization. Further, consistent with the prediction of Thiery *et al.* [65, 66], the system "looks localized" near the transition, as evidenced by the strong similarity between the spin dynamics for the non-interacting case $\Delta = 0$ and the interacting case close to our lower bound for the MBL transition. However, at the present stage, we cannot explicitly confirm (or falsify) this mechanism, as this would require a specific analysis of observables that would have a distinct behavior within this mechanism. Another possibility consistent with our results is the existence of a glass-like crossover regime [63].

Future work inspired by our results could focus on different energy densities in addition to just the center of the band that has been considered here, in order to map the transition line in the energy-disorder plane on the basis of data for large systems. Further, it is also interesting and instructive to analyze the phase diagram of long XXZ chains in the interaction-disorder plane. Clearly, a better analytical understanding of the slow dynamics near the MBL transition would be very important. In this context, it is interesting to note that a similar slow transport near the transition (or, more accurately, apparent transition) is also found for quasiperiodic systems [13] as well as in two-dimensional disordered systems [68], where the influence of rare bottlenecks is expected to be negligible. It remains to be seen whether this slow transport has a common origin in all these situations.

A so-called "dreaming" protocol within the neural-

network framework, which generates the configurations that are representative for given phases, could be very useful for developing the analytical theories of the subdiffusive phase. Furthermore, this type of machine-learning approach can be envisaged to simulate the dynamics at much longer times (and, perhaps, in much larger systems), which is currently inaccessible by the most advanced numerical tools. Finally, a hotly debated question is the very existence of the true ($L \to \infty$) MBL transition in various models. We hope that our work will pave the way for numerical investigations of the physics associated with the MBL transition on large systems, which is of obvious importance for shedding light on this issue.
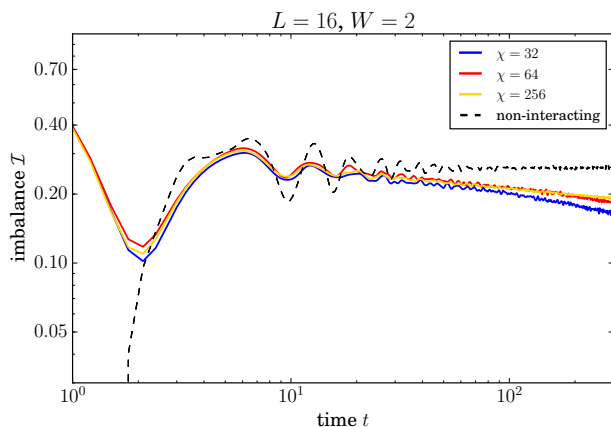
FIG. 8. Time evolution of the disorder-averaged imbalance (3) as computed by the TDVP with a truncated bond dimension $\chi = 32, 64$ and by exact numerics $\chi = 256$ for a small system $L = 16$, up to $t = 300$. We consider several hundred independent realizations of disorder. For the sake of comparison, also the non-interacting case $\Delta = 0$ is shown.

Appendix A: Convergence of the TDVP

Since the TDVP truncates the entanglement spectrum, it is of interest to investigate the convergence with respect to bond dimension [40]. We perform two benchmarks: first, we compare the time evolution of the TDVP to numerically exact numerics for a small system $L = 16$. Here, we consider the disorder average of independent disorder samples, since we are interested in convergence at the level of this average, rather than at the level of individual realizations. The exact numerics are obtained simply by performing TDVP time evolution with an unrestricted bond dimension. For a system of size $L = 16$ an MPS with maximum bond dimension $\chi = 256$ captures the time evolution exactly (for a system of size L, non-truncated MPS have a maximum bond dimension of $2^{L/2}$ in the center of the chain). The result is shown in Fig. 8, which shows the time evolution of the initial state up to times $t = 300$. Recall that in the main body of the paper, we considered times only up to $t = 100$. Secondly, we consider a larger system $L = 50$ and consider whether the result converges with bond dimension, where we consider times up to $t = 200$ and bond dimensions up to $\chi = 96$, higher than in the main body of the paper. The result is shown in Fig. 9.

From both of these approaches, we conclude that for the times we consider (up to $t = 100$), a bond dimension of $\chi = 64$ provides an excellent approximation for not too weak disorder $W = 2$; an approximation that only improves with increasing disorder. This conclusion is further corroborated by comparison with a completely independent implementation of the time-evolving block decimation method, for which we compute the power law decay coefficient β for $W = 5$, finding excellent agreement with the TDVP (see Fig. 2). Therefore, our numerical results are certainly accurate in the disorder and time window we have considered in this work. Moreover, in the moderately strongly disordered case $W = 5$ we find good agreement between power laws obtained in the window $t \in [50, 100]$ and those in the window $t \in [100, 200]$, providing evidence for the survival of these power laws in the long-time limit.

[1] Rahul Nandkishore and David A. Huse, "Many-body localization and thermalization in quantum statistical mechanics," Ann. Rev. Cond. Mat. Phys. **6**, 15–38 (2015).
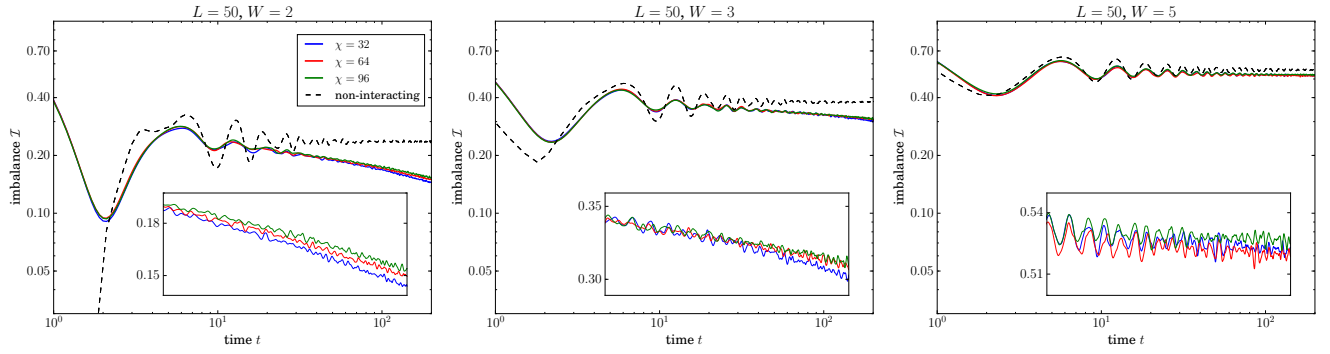
FIG. 9. Time evolution of the disorder-averaged imbalance (3) as computed by the TDVP with bond dimension $\chi = 32, 64, 96$ for a moderately large system $L = 50$ up to $t = 200$ and choices of disorder strength $W = 2, 3, 5$. For comparison, also the non-interacting case $\Delta = 0$ is shown. The fitted coefficients for the power law decay of the imbalance in the window $t \in [100, 200]$ for $W = 5$ are $\beta(\chi = 32) = 0.025 \pm 0.005$, $\beta(\chi = 64) = 0.020 \pm 0.005$, $\beta(\chi = 96) = 0.020 \pm 0.004$, which is in good agreement (within error bars) with the value found for $t \in [50, 100]$ as presented in the main text. Insets show a zoomed region in the time window $t \in [50, 200]$ on a log-log scale.

[2] Ehud Altman and Ronen Vosk, "Universal dynamics and renormalization in many-body-localized systems," Ann. Rev. Cond. Mat. Phys. **6**, 383–409 (2015).

[3] Dmitry A. Abanin and Zlatko Papić, "Recent progress in many-body localization," Ann. Phys. (Berl.) **529**, 1700169 (2017).

[4] Ferdinand Evers and Alexander D. Mirlin, "Anderson transitions," Rev. Mod. Phys. **80**, 1355 (2008).

[5] M. Rigol, V. Dunjko, and M. Olshanii, "Thermalization and its mechanism for generic isolated quantum systems," Nature **452**, 854 (2008).

[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.M. Basko, I.L. Aleiner, and B.L. Altshuler, "Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states," Ann. Phys. (N. Y.) **321**, 1126 – 1205 (2006).

[8] Marko Žnidarič, Antonello Scardicchio, and Vipin Kerala Varma, "Diffusive and subdiffusive spin transport in the ergodic phase of a many-body localizable system," Phys. Rev. Lett. **117**, 040601 (2016).

[9] Vedika Khemani, Frank 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).

[10] Xiongjie Yu, David Pekker, and Bryan K. Clark, "Finding matrix product state representations of highly excited eigenstates of many-body localized hamiltonians," Phys. Rev. Lett. **118**, 017201 (2017).

[11] Thorsten B. Wahl, Arijeet Pal, and Steven H. Simon, "Efficient representation of fully many-body localized systems using tensor networks," Phys. Rev. X **7**, 021018 (2017).

[12] Elmer V. H. Doggen, Gabriel Lemarié, Sylvain Capponi, and Nicolas Laflorencie, "Weak- versus strong-disorder superfluid—bose glass transition in one dimension," Phys. Rev. B **96**, 180202 (2017).

[13] Michael Schreiber, Sean S. Hodgman, Pranjal Bordia, Henrik P. Lüschen, Mark H. Fischer, Ronen Vosk, Ehud Altman, Ulrich Schneider, and Immanuel Bloch, "Observation of many-body localization of interacting fermions in a quasirandom optical lattice," Science **349**, 842–845 (2015).

[14] Jae-yoon Choi, Sebastian Hild, Johannes Zeiher, Peter Schauß, Antonio Rubio-Abadal, Tarik Yefsah, Vedika Khemani, David A. Huse, Immanuel Bloch, and Christian Gross, "Exploring the many-body localization transition in two dimensions," Science **352**, 1547–1552 (2016).

[15] 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).

[16] G. Kucsko, S. Choi, J. Choi, P. C. Maurer, H. Zhou, R. Landig, H. Sumiya, S. Onoda, J. Isoya, F. Jelezko, E. Demler, N. Y. Yao, and M. D. Lukin, arXiv:1609.08216.

[17] John Z. Imbrie, "On many-body localization for quantum spin chains," J. Stat. Phys. **163**, 998–1048 (2016).

[18] Arijeet Pal and David A. Huse, "Many-body localization phase transition," Phys. Rev. B **82**, 174411 (2010).

[19] David J. Luitz, Nicolas Laflorencie, and Fabien Alet, "Many-body localization edge in the random-field heisenberg chain," Phys. Rev. B **91**, 081103 (2015).

[20] F. Pietracaprina, N. Macé, D. J. Luitz, and F. Alet, arXiv:1803.05395 (2018).

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

[22] Trithep Devakul and Rajiv R. P. Singh, "Early breakdown of area-law entanglement at the many-body delocalization transition," Phys. Rev. Lett. **115**, 187201 (2015).

[23] Maksym Serbyn, Z. Papić, and Dmitry A. Abanin, "Universal slow growth of entanglement in interacting strongly disordered systems," Phys. Rev. Lett. **110**, 260601 (2013).

[24] Yevgeny Bar Lev, Guy Cohen, and David R. Reichman, "Absence of diffusion in an interacting system of spinless

fermions on a one-dimensional disordered lattice," Phys. Rev. Lett. **114**, 100601 (2015).

[25] Kartiek Agarwal, Sarang Gopalakrishnan, Michael Knap, Markus Müller, and Eugene Demler, "Anomalous diffusion and griffiths effects near the many-body localization transition," Phys. Rev. Lett. **114**, 160401 (2015).

[26] Kartiek Agarwal, Ehud Altman, Eugene Demler, Sarang Gopalakrishnan, David A. Huse, and Michael Knap, "Rare-region effects and dynamics near the many-body localization transition," Ann. Phys. (Berl.) **529**, 1600326 (2016).

[27] David J. Luitz, Nicolas Laflorencie, and Fabien Alet, "Extended slow dynamical regime close to the many-body localization transition," Phys. Rev. B **93**, 060201 (2016).

[28] Soumya Bera, Giuseppe De Tomasi, Felix Weiner, and Ferdinand Evers, "Density propagator for many-body localization: Finite-size effects, transient subdiffusion, and exponential decay," Phys. Rev. Lett. **118**, 196801 (2017).

[29] P. A. M. Dirac, "Note on exchange phenomena in the thomas atom," Proc. Cambridge Philos. Soc. **26**, 376–385 (1930).

[30] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete, "Time-dependent variational principle for quantum lattices," Phys. Rev. Lett. **107**, 070601 (2011).

[31] Jutho Haegeman, Christian Lubich, Ivan Oseledets, Bart Vandereycken, and Frank Verstraete, "Unifying time evolution and optimization with matrix product states," Phys. Rev. B **94**, 165116 (2016).

[32] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, "Time-dependent density-matrix renormalization-group using adaptive effective hilbert spaces," J. Stat. Mech.: Th. Exp. **2004**, P04005 (2004).

[33] Steven R. White and Adrian E. Feiguin, "Real-time evolution using the density matrix renormalization group," Phys. Rev. Lett. **93**, 076401 (2004).

[34] G. Vidal, "Efficient classical simulation of slightly entangled quantum computations," Phys. Rev. Lett. **91**, 147902 (2003).

[35] Pasquale Calabrese and John Cardy, "Evolution of entanglement entropy in one-dimensional systems," J. Stat. Mech.: Th. Exp. **2005**, P04010 (2005).

[36] Nicolas Laflorencie, Phys. Rep. **646**, 1 – 59 (2016).

[37] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," Ann. Phys. (N. Y.) **326**, 96 – 192 (2011).

[38] C. Karrasch, J. H. Bardarson, and J. E. Moore, "Reducing the numerical effort of finite-temperature density matrix renormalization group calculations," New J. Phys. **15**, 083031 (2013).

[39] E. Leviatan, F. Pollmann, J. H. Bardarson, D. A. Huse, and E. Altman, arXiv:1702.08894 (2017).

[40] Benedikt Kloss, Yevgeny Bar Lev, and David Reichman, "Time-dependent variational principle in matrix-product state manifolds: Pitfalls and potential," Phys. Rev. B **97**, 024307 (2018).

[41] Mario Berta, Fernando G. S. L. Brandão, Jutho Haegeman, Volkher B. Scholz, and Frank Verstraete, arXiv:1709.07423 (2017).

[42] J. M. Deutsch, "Quantum statistical mechanics in a closed system," Phys. Rev. A **43**, 2046–2049 (1991).

[43] M. Srednicki, "Chaos and quantum thermalization," Phys. Rev. E **50**, 888–901 (1994).

[44] David J. Luitz and Yevgeny Bar Lev, "The ergodic side of the many-body localization transition," Ann. Phys. (Berl.) **529**, 1600350–n/a (2017).

[45] Daniel Jaschke, Michael L. Wall, and Lincoln D. Carr, "Open source matrix product states: Opening ways to simulate entangled many-body quantum systems in one dimension," Comput. Phys. Commun. **225**, 59 – 91 (2018).

[46] It should be noted that for small $\beta = \mathcal{O}(10^{-2})$, we can no longer distinguish a power-law decay over our limited time window from, say, a logarithmic correction. Nevertheless, this procedure yields a sensitive probe for the decay of $\mathcal{I}(t)$ in the sense that *any* statistically significant monotonous decay will result in a non-zero β.

[47] The imbalance of a non-interacting system $\mathcal{I}_\text{A}(t)$, while asymptotically approaching a positive constant, shows a weak decay with time. We have checked that fitting of this decay by a power law in the considered time interval yields fictitious values of β in the range $\beta \approx 0.004 - 0.006$. This implies a lower bound on the true exponent β of a many-body delocalized system that we can distinguish from zero.

[48] In numerical language, the entanglement eigenvalues are also the squares of the Schmidt numbers, obtained easily in the MPS representation through singular value decompositions.

[49] Hui Li and F. D. M. Haldane, "Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-abelian fractional quantum hall effect states," Phys. Rev. Lett. **101**, 010504 (2008).

[50] David J. Luitz, "Long tail distributions near the many-body localization transition," Phys. Rev. B **93**, 134201 (2016).

[51] Christopher Bishop, *Pattern Recognition and Machine Learning* (Springer-Verlag, New York, 2006).

[52] Frank Schindler, Nicolas Regnault, and Titus Neupert, "Probing many-body localization with neural networks," Phys. Rev. B **95**, 245134 (2017).

[53] Tomohiro Mano and Tomi Ohtsuki, "Phase diagrams of three-dimensional anderson and quantum percolation models using deep three-dimensional convolutional neural network," J. Phys. Soc. Jpn. **86**, 113704 (2017).

[54] Jordan Venderley, Vedika Khemani, and Eun-Ah Kim, "Machine learning out-of-equilibrium phases of matter," Phys. Rev. Lett. **120**, 257204 (2018).

[55] E. P. L. van Nieuwenburg, E. Bairey, and G. Refael, arXiv:1712.00450 (2017).

[56] Y.-T. Hsu, X. Li, D.-L. Deng, and S. Das Sarma, arXiv:1805.12138 (2018).

[57] Wei Zhang, Lei Wang, and Ziqiang Wang, arXiv:1807.02954 (2018).

[58] Evert P. L. van Nieuwenburg, Ye-Hua Liu, and Sebastian D. Huber, "Learning phase transitions by confusion," Nat. Phys. **13**, 435 (2017).

[59] The start of the plateau can be associated with the value of W where the lower bound of the error bar in the right panel of Fig. 6 touches the upper bound of the error bar for the training at $W = 8$.

[60] It is interesting to note that the "MBL branch" for large chains ($L = 50$ and 100) starts at $W \simeq 3 - 4$, which is close to the estimate for the critical disorder that was previously obtained from the exact diagonalization in relatively small chains.

[61] K. S. Tikhonov, A. D. Mirlin, and M. A. Skvortsov, "Anderson localization and ergodicity on random regular graphs," Phys. Rev. B **94**, 220203 (2016).

[62] K. S. Tikhonov and A. D. Mirlin, "Many-body localization transition with power-law interactions: Statistics of eigenstates," Phys. Rev. B **97**, 214205 (2018).

[63] G. Biroli and M. Tarzia, "Delocalized glassy dynamics and many-body localization," Phys. Rev. B **96**, 201114 (2017).

[64] Wojciech De Roeck, Francois Huveneers, Markus Müller, and Mauro Schiulaz, "Absence of many-body mobility edges," Phys. Rev. B **93**, 014203 (2016).

[65] T. Thiery, F. Huveneers, M. Müller, and W. De Roeck, arXiv:1706.09338 (2017).

[66] T. Thiery, M. Müller, and W. De Roeck, arXiv:1711.09880 (2017).

[67] Sarang Gopalakrishnan, Markus Müller, Vedika Khemani, Michael Knap, Eugene Demler, and David A. Huse, "Low-frequency conductivity in many-body localized systems," Phys. Rev. B **92**, 104202 (2015).

[68] Pranjal Bordia, Henrik Lüschen, Sebastian Scherg, Sarang Gopalakrishnan, Michael Knap, Ulrich Schneider, and Immanuel Bloch, "Probing slow relaxation and many-body localization in two-dimensional quasiperiodic systems," Phys. Rev. X **7**, 041047 (2017).

[69] Ashley Milsted, Jutho Haegeman, Tobias J. Osborne, and Frank Verstraete, "Variational matrix product ansatz for nonuniform dynamics in the thermodynamic limit," Phys. Rev. B **88**, 155116 (2013).

[70] M. L. Wall and Lincoln D. Carr, "Out-of-equilibrium dynamics with matrix product states," New J. Phys. **14**, 125015 (2012).

[71] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).

[72] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015), software available from tensorflow.org.

[73] J. D. Hunter, "Matplotlib: A 2d graphics environment," Comput. Sci. Eng. **9**, 90–95 (2007).

[74] O. Tange, "Gnu parallel - the command-line power tool," ;login: The USENIX Magazine **36**, 42–47 (2011).