Many-body localization transition in a two-dimensional disordered quantum dimer model

Hugo Théveniaut, Zhihao Lan, and Fabien Alet

1Laboratoire de Physique Théorique, IRSAMC, Université de Toulouse, CNRS, UPS, 31062 Toulouse, France
2Centre for the Mathematics and Theoretical Physics of Quantum Non-equilibrium Systems and School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom

(Dated: February 13, 2019)

Many-body localization is a unique physical phenomenon driven by interactions and disorder for which a quantum system can evade thermalization. While the existence of a many-body localized phase is now well-established in one-dimensional systems, its fate in higher dimension is an open question. We present evidence for the occurrence of a many-body localization transition in a two-dimensional quantum dimer model with interactions and disorder. This conclusion is reached on the basis of large-scale simulations for static and dynamical properties of a consequent number of observables. Our results pave the way for a generic understanding of occurrence of a many-body localization transition in dimension larger than one, and highlight the unusual quantum dynamics that can be present in constrained systems.

I. INTRODUCTION

The emergence of thermalization in isolated quantum systems, which are not in contact with a bath and solely follow unitary Hamiltonian dynamics, has long been a central issue in statistical mechanics. This question has been revived thanks to an extraordinary increase in the degree of control and isolation reached in cold-atom experiments. The central concept of Eigenstate Thermalization Hypothesis (ETH), a conjecture on expectation values of local operators in the eigenstates of the isolated system, allows to provide a physical understanding and justification of this thermalization. The ETH is found to be valid for the generic majority of many-body systems. Recent years have witnessed a vast body of research concentrating on a situation where the ETH fails, namely the phenomenon of many-body localization (MBL), which occurs most noticeably in the presence of strong disorder.

In a many-body localized system, eigenstates at the same energy density do not form a statistical ensemble, and local expectation values strongly differ from one eigenstate to the other. MBL states sustain a low level of quantum entanglement in the form of an area law, e.g., in contrast with ETH states which obey a volume law for entanglement entropy, and have been found to be able to sustain long-range order in cases where equilibrium states are forbidden to. Dynamics in MBL systems markedly differ from those of thermalizing systems, with e.g., long-time memory of the initial state, absence of transport for conserved quantities (localization), and a slow propagation of quantum information with logarithmic growth of entanglement. The latter example allows to distinguish MBL from Anderson localization (where entanglement is bounded) which occurs in non-interacting systems.

The existence of many-body localization in one dimension is now established and reasonably well-understood in one-dimensional (1d) quantum lattice systems with short-range interactions, thanks to a concerted effort of approaches, including numerical simulations, phenomenological renormalization approaches, rigorous results, and cold-atom experiments. Crucial to this understanding is the provable existence of local integral of motions, denoted as MBL states (for localized bits): these emergent localized operators, which diagonalize the Hamiltonian of the system at strong disorder, explain most of the salient features of the MBL phase. In the typical setting of lattice models, the MBL phase indeed emerges in the presence of strong enough disorder, and is separated from an ETH phase at low disorder by a many-body localization transition involving an exponential number of eigenstates. The reviews provide an introduction and recent insights on this alive field of research.

A. Many-body localization in higher dimensions

An outstanding question which remains open is the fate of many-body localization in dimension higher than one. A basic observation in this case is that localization is more difficult to achieve as there are more transport channels than in one dimension. The Anderson localization survives in two dimensions (2d) for all disorder strengths, but with a much higher localization length, and needs an increasingly stronger critical disorder in higher dimensions.

In the many-body case, analytical arguments were presented that suggest the intrinsic instability of a MBL phase in dimension higher than one. They are based on the inevitable existence of thermal ‘bubbles’ (created by e.g. local regions of space with low disorder), which growth with time is energetically favorable and which eventually thermalize any MBL sample. The thermalization conveyed in this scenario is asymptotic, meaning that it may require an asymptotically large time before signs of localization (for instance memory of the initial state) disappear. In these works, the MBL region is represented in terms of the l-bits, the exponentially localized integral of motions which form a complete set of operators spanning the Hilbert space.

In spite of these arguments, three recent remarkable numerical simulations present signatures of a MBL phase in two-dimensional models of spins or bosons, in either static or dynamic quantities. These results are very intriguing, but the methods used in these works are not unbiased: an approximate unitary diagonalization transformation based on the lowly-entangled nature or localized nature of MBL...
eigenstates is used, which misses by construction the ETH phase (we note nevertheless that Ref. 46 presents evidence for the transition in 2d, but without a finite-size or finite-entanglement scaling analysis). Other exact diagonalization studies in 2d are limited to small systems (lattices with at most 16 sites), such that no finite-size scaling is possible. Finally, a quantum Monte Carlo treatment of a Hamiltonian constructed from $\ell$-bits also argues for the existence of MBL in a 2d spin system\textsuperscript{21}. On the experimental side, recent advances in cold-atom experiments present clear dynamical signs of localization (in terms of memory of an initial imbalance of atoms) on the longest accessible time scale in two dimensional systems, either for bosonic atoms in random potential or fermionic gases in quasi-periodic potentials\textsuperscript{53}.

In this work, we present a comprehensive study of a two-dimensional strongly interacting disordered model, based on extensive numerical simulations on fairly large systems (square lattices with up to 64 sites). These simulations are exact, unbiased, and allow to study both ETH and MBL regimes and the transition between them. We study an important range of clusters of different sizes and shapes, allowing to follow the evolution of many ETH or MBL indicators with size. This is made possible thanks to two main specificities of our work: we push large-scale parallel numerical methods to their limits (see Sec. II B), and we work on a constrained model (described in Sec. II A) that presents specific features appropriate for localization and numerical studies (Sec. I B).

### II. MODEL AND NUMERICAL METHODS

#### A. Model

This work focuses on a quantum dimer model on a square lattice with random potential terms. Quantum dimer models have a long history in condensed matter as low-energy effective models. They have been first introduced by Rokhsar and Kivelson\textsuperscript{72} in the context of the Resonating Valence Bond theory of high-temperature superconductivity, an interest which has been revived with the use of QDMs to describe experiments on underdoped cuprates\textsuperscript{73}. QDMs are also prominent representations of highly frustrated magnets where dimers represent nearest-neighbor spin singlets\textsuperscript{74}. Dimer models are indeed often more tractable than frustrated spin models from which they can be systematically derived\textsuperscript{75}, while capturing the essential physics. For instance, the existence of 2d quantum spin liquids exhibiting topological order was first demonstrated in a QDM on a non-bipartite lattice\textsuperscript{76}.

Quantum dimer models are defined in a configuration space where only configurations which fulfill the dimer constraint are allowed (dimer coverings, see Fig. 1). Their Hamiltonian contains a potential term that attributes an energy to every interactions encoded in a Hamiltonian), which would rather favor delocalization and thermalization. This competition provides strong motivation to investigate the possibility of MBL in quantum constrained models. The interplay between constraints, interactions and disorder has been explored in recent works on 1d model\textsuperscript{19,71} (for Ref. 70, equivalent to a dimer model on a ladder), which constraints are relevant for the physics of cold-atomic chains in the regime of Rydberg blockade resulting in rich phase diagrams, with \textit{e.g.} different mechanisms to produce MBL\textsuperscript{30}.

Another important aspect (crucial for the numerical approach taken in this work) worth mentioning is the fact that the constraints overall \textit{reduce} the size of the phase space for allowed configurations. For instance, the size of the Hilbert space scales as \( \sim 1.34^N \) for dimer coverings of square lattices with \( N \) sites, instead of \( 2^N \) for spin 1/2 models (or \( 4^N \) if no constraints were present). This greatly helps numerics based on exact approaches, as this enlarges the set of available finite samples in 2d. Given that exact diagonalization simulations have been instrumental in the study and understanding of MBL, this is also an important positive feature in the search for MBL in quantum constrained models.

The plan of the paper is as follows. We first describe the model and the 2d lattices used in this work in Sec. II A, along with a brief description of the exact numerical methods employed (Sec. II B). We then describe in detail physical properties of the eigenstates of this model in Sec. III, including spectral and eigenstate statistics, expectation values of local observables as well as half-system entanglement entropy. Sec. IV is devoted to a machine-learning analysis of entanglement spectra. Sec. V then presents an analysis of dynamical properties after a quench. Finally, we summarize our results and present perspectives in Sec. VI.
Dimer configuration

Potential term

Quantum dimer model with random potential

Winding numbers

FIG. 1. Quantum dimer models. Top left panel: a configuration of dimers is valid when each lattice site is touched by one and only one dimer. Right panel: illustration of the quantum dimer model on the square lattice with a disordered potential Eq. 1. Plaquettes $p$ of the lattice where a pair of parallel dimers is present (highlighted in green) contribute a potential $V_p$ (different for each plaquette) to the potential energy of this configuration. The kinetic term flips dimers on such plaquettes such that the orientation of the two dimers on this plaquette is changed. In the bottom right panel, the flip of the plaquette $p = 20$ (surrounded by red lines) is represented, resulting in a new configuration with a different potential energy. Bottom left panel: on the square (bipartite) lattice with periodic boundary conditions, all dimer configurations can be classified by a pair of winding numbers $(W_x, W_y)$ defined as follows. Consider a horizontal (vertical) line—such as represented in the figure—perpendicular to lattice bonds. $W_x$ ($W_y$) is then equal to the algebraic number (i.e. ±1 where the sign is defined by the sublattice on a fixed side of the line) of dimers in this configuration crossing this line. The sign ±1 is represented by arrows on lattice bonds. Note that the winding numbers do not depend on the position of the line, and they cannot be changed by a local reconfiguration of dimers. The represented configurations all belong to the $(W_x, W_y) = (0, 0)$ sector in which all calculations are performed in this work. Periodic boundary conditions are represented with the help of duplicated grey sites on the boundaries.

dimer covering, and a kinetic term that usually flips dimers along a loop, allowing to pass from one dimer configuration to another. The specific QDM Hamiltonian that we consider in this work is based on the square lattice and contains a potential term with a random component, taking the following form

$$H = -t \sum_{\Box} \left( \langle \Box | \langle \Box | + \langle \Box | \langle \Box | \right) + \sum_{\Box_p} V_p \left( \langle \Box | \langle \Box | + \langle \Box | \langle \Box | \right).$$

(1)

The sums run over all plaquettes $\Box_p$ of the square lattice, $V_p$ is a random potential different for each plaquette which is drawn uniformly from a box distribution $V_p \in [-V, V]$. The kinetic energy scale is set to $t = 1$. The various terms are illustrated in Fig. 1. At $V = 0$ (no disorder), Ref. 77 showed that eigenstates in the middle of the spectrum obey the eigenstate thermalization hypothesis. Quite importantly, there is no limit where Eq. 1 can be mapped to a model of free particles moving in a random potential (no Anderson localization limit).

Quantum dimer models based on coverings of bipartite lattices (such as the square lattice) have an important property: they display conserved topological numbers, known as winding numbers, when considering systems with periodic boundary conditions. The definition of the two conserved winding numbers (denoted $W_x$ and $W_y$) is given in Fig. 1. It is easy to see that they are conserved by the Hamiltonian Eq. 1, and by any local moves (i.e. that do not go through periodic boundary conditions), hence their topological nature. $W_x$ and $W_y$ both range from $-L/2$ to $L/2$ on a square lattice. The Hamiltonian is block-diagonal, with sectors identified by $(W_x, W_y)$. We note that the resulting abelian $U(1)$ symmetry of the model Eq. 1 is compatible with many-body localization. Without loss of generality, we will restrict to the largest sector $(W_x = 0, W_y = 0)$ in what follows.

B. Lattices and numerical methods

This work uses two-dimensional lattices (represented in Fig. 13 in the Appendix) with $N$ sites and $2N$ bonds of the following form: regular square lattices with $N = L^2$ sites (with $L = 4, 6, 8$), rectangular lattices with $N = L_x \times L_y$ sites (with $L_{x,y} = 4, 6, 8, 10$) as well as tilted square lattices with $N = 32, 40$ and 52 sites (see Appendix for details). Periodic boundary conditions are taken. All the lattices used in this work are presented in Tab. I, together with the corresponding Hilbert space sizes in the $(W_x = 0, W_y = 0)$ sector.
We use different numerical methods to treat exactly the Hamiltonian Eq. 1. In Sec. III, we first use methods which allow to obtain exact eigenstates of $H$. As usually done in the MBL context, we concentrate on eigenstates with eigenvalue $E$ located right in the middle of the many-body spectrum, that is at $\epsilon = 0.5$ with $\epsilon = (E - E_{\text{min}})/(E_{\text{max}} - E_{\text{min}})$ where $E_{\text{min}}$/$E_{\text{max}}$ are the extremal energies of $H$. Full numerical diagonalization is possible up to $N = 36$ sites. We also use massively parallel shift-invert diagonalization\textsuperscript{79} that allows to target eigenstates at $\epsilon = 0.5$ on larger samples up to $N = 52$. We typically obtain about 100 eigenstates near $\epsilon = 0.5$. Results are averaged over more than 500 (up to 10,000) realizations of disorder. To show the interest of working with constrained models, the Hilbert space of the largest cluster on which we can obtain exact eigenstates ($N = 52$) is almost twice as large as the one for $N = 24$ spins 1/2 (in the zero magnetization sector), which constitutes the state of the art for the analysis of MBL in spin chains\textsuperscript{79–81}. Obtaining 10 eigenstates in the middle of the spectrum for one disorder sample at $N = 52$ requires about 20 minutes on 10,000 CPU cores.

In Sec. V, we also consider dynamical properties of various observables after a quench. The knowledge of all eigenstates allows to consider properties at arbitrary time up to $N = 36$. We also use Krylov expansion time evolution techniques\textsuperscript{82}, which allow simulations up to time of the order of thousand plaquette flips (larger than typical experimental time scales) on large clusters up to $N = 64$. Note that simulations on the largest sample $N = 64$ are extremely demanding (see sizes of Hilbert space in Tab. I). In the study of dynamics, we average over between 30 and 1000 realizations of disorder for each simulation set.

### III. PROPERTIES OF EIGENSTATES

#### A. Spectral statistics

We first consider the simplest property of a many-body system, namely its energy spectrum. The nature of the distribution of energy levels can already discriminate between different regimes: ergodic systems display level repulsion while localized systems display a random distribution of level spacing characterized by Poisson statistics. In the many-body context, and in order to counter the fact that the density of states is not uniform, it is useful to consider the distribution of consecutive ratios as introduced in Ref. 83. We define gaps in the many-body spectrum as $s_n = E_n - E_{n-1}$ and consider the consecutive reduced gap ratio $r_n = \frac{\min(s_n, s_{n+1})}{\max(s_n, s_{n+1})}$, for which $0 \leq r \leq 1$. The distribution $P(r)$ and average value of $\langle r \rangle = \int_0^1 rP(r)dr$, averaged over eigenstates and disorder realizations, have been computed for random matrix ensembles (corresponding to ergodic systems) and for independent energy levels (Poisson statistics)\textsuperscript{84}. For real Hamiltonians such as Eq. 1, the random matrix ensemble to be considered is the Gaussian Orthogonal Ensemble (GOE). Fig. 2 displays both the distribution $P(r)$ and its average $\langle r \rangle$ for various values of disorder and system sizes. We compare them to predictions for the GOE ensemble\textsuperscript{84} $P_{\text{GOE}}(r) = \frac{27}{16}\frac{r}{(1+r^2)^2}$, $\langle r_{\text{GOE}} \rangle \simeq 0.5307$ and Poisson statistics $P_{\text{Poisson}}(r) = \frac{\pi}{4\sqrt{r}}$, $\langle r_{\text{Poisson}} \rangle \simeq 0.38629$. Considering first the distribution (Fig. 2b), we find that the agreement with $P_{\text{GOE}}(r)$ for low values of disorder ($V = 3$) and $P_{\text{Poisson}}(r)$ for high values ($V = 25, 30$) is excellent, with no adjustable parameter. Intermediate values of disorder $V = 12, 17$ present distributions which display a crossover.
behavior between the two limit distributions for the system size considered $N = 40$. The average gap ratio $\langle r \rangle$ also displays an interesting crossover between the two limiting cases, with different system sizes showing an apparent crossing point (see inset of top panel of Fig. 2 for square samples $N = \{32, 36, 40\}$ around $V \approx 15 - 20$. We note that rectangular samples also show the same crossover between the GOE and Poisson limits, but do not cross exactly at the same value which can be attributed to their different aspect ratios. The critical value of the gap ratio $\langle r \rangle^* \approx 0.392$ is smaller than for the 1d standard MBL model, indicating that the putative transition point looks even closer to the Poissonian localized limit. Note that for the largest disorder $V = 30$ considered here, we also checked on the largest system available to full diagonalization ($N = 36$) that no band-like structure is present in the energy spectrum, a feature which was found to mimic MBL on too small system sizes.

### B. Eigenstate statistics

We next consider the similarity between eigenstates at the same energy density. It is well-known that eigenstates in an ergodic phase look ‘similar’ (e.g. they present similar expectation values for local observables), while eigenstates in a MBL phase are very different (e.g. local observables can differ strongly from one eigenstate to the following one in the many-body spectrum). To quantify the degree of similarity, it is useful to consider the Kullback-Leibler divergence between two eigenstates $|n\rangle$ and $|n'\rangle$ located nearby in the spectrum (here at the same energy density $\epsilon = 0.5$). It is defined as $KL = \sum_p p_i \log(p_i/p'_i)$, where $p_i$ and $p'_i$ are the participation coefficients, i.e. the square amplitudes $p_i = |\langle n|i \rangle|^2$, $p'_i = |\langle n'|i \rangle|^2$ of eigenstates in a given basis (here the computational basis where dimer occupations are diagonal). The sum runs over the full Hilbert space, and the eigenstates are normalized such that $\sum_i p_i = \sum_i p'_i = 1$. Eigenstates at ‘infinite temperature’ ($\beta = 0.5$) in an ergodic phase should be very close to eigenstates of GOE random matrices, for which one has $KL_{GOE} = 2$. On the other hand, for eigenstates which are many-body localized, we expect $KL$ to diverge with system size as the eigenstates are increasingly different.

Fig. 3 precisely shows this behavior, with expectation values $\langle KL \rangle$ converging to their limiting GOE value at low disorder, and diverging for large disorder. We also observe a crossing point at $V \approx 15$, albeit slightly drifting, for different system sizes. This drifting behavior is also observed for 1d MBL in the random-field Heisenberg model. The value of $KL$ at the transition point is larger than for the 1d spin model, pointing again to a critical point closer to the localized limit.

### C. Dimer occupation and columnar imbalance

**Dimer occupation** — A qualitative understanding of the behavior of eigenstates in the different regimes is provided by considering local observables, and in particular by lattice bond occupation by dimers. For each lattice bond $b$, we compute the dimer occupation $\langle n_b | n \rangle$ in the eigenstate $|n\rangle$ with $n_b |i\rangle = 1$ if the bond $b$ is occupied by a dimer in the basis configuration $|i\rangle$, 0 otherwise.

Fig. 4 represents graphically the dimer occupation $\langle n_b | n \rangle$ in three different eigenstates, consecutive in the many-body spectrum, for three values of disorder strength. In the low-disorder phase ($V = 1$), all dimer occupations are equally likely: $\langle n_b \rangle \approx 1/4$ for all bonds and all eigenstates. The dimer occupations at large disorder ($V = 30$) mainly overlap with a single dimer configuration, which differs totally from one eigenstate to the next one. Resonances between configu-
computing $I$ for all dimer configurations in the Hilbert space) is represented in the $V = 30$ panel for comparison (there are $1/N$ values possible for the imbalance $I$). The location of the peaks of $P(I)$ matches precisely those of $P_{\text{conf}}(I)$, with however different amplitudes. This indicates that at large $V$, eigenstates are close (but not exactly equal) to pure dimer configurations, in agreement with the qualitative view provided by the snapshots of Fig. 4.

D. Participation entropies

To understand the localization properties of the eigenfunctions in the Hilbert (configuration) space, we next consider the participation entropy

$$S_p(|n\rangle) = -\sum_i p_i \log(p_i)$$

(3)

where $p_i = |\langle n|i\rangle|^2$ and the sum runs over all basis states $|i\rangle$ of the computational basis. At low disorder, we expect eigenfunctions to be fully delocalized with a participation entropy scaling as $S_p = a \log(\mathcal{H}) - \ldots$ (with $a = 1$ for a fully-delocalized state and ... denoting finite size corrections). In the large disorder regime on the other hand, based on data for the MBL phase in the 1d random field Heisenberg\cite{1}, we do not expect a true localization in configuration space but instead a volumic, multifractal, behavior where $S_p = a \log(\mathcal{H})$ with $a$ strictly less than 1 (and increasingly small as disorder is increased). Fig. 6 presents the numerical data for the average participation entropy divided by the log of the size of the Hilbert space $\langle S_p \rangle / \log(\mathcal{H})$. Two regimes can be clearly distinguished: one at low $V$ where data tend to their maximal value ($a = 1$) with system size, and another at larger disorder where curves tend to group together towards sensibly smaller values of the volume law coefficient $a$. Even though there is no unique consistent crossing point in the data (note that there is no generic argument predicting such a crossing point as $S_p$ exhibits a volume law scaling in both regimes), we find that

![Diagram](image-url)

FIG. 5. Probability distribution of columnar imbalance $I$ for different values of disorder. On the bottom right panel is also represented (circles) the distribution for pure dimer $P_{\text{conf}}(I)$ (normalized such that $P_{\text{conf}}(I = 0) = P_{V=30}(I = 0)$). Data for the tilted square lattice with $N = 40$ sites, with more than $10,000$ eigenstates for each value of disorder. Inset (top right): definition of the phase $\phi_p$ used in the definition of the columnar imbalance, on top of the columnar configuration for which $I = 1$.

rations which differ by one or two plaquette flips can be evidenced by almost equal occupation of parallel dimers around a plaquette in some eigenstates (e.g. left-middle plaquette for the left-most eigenstate at $V = 30$). At intermediate disorder ($V = 17$), while for the two first left eigenstates dimer occupations look similar (at a superficial level), the right eigenstate configuration is markedly different.

Columnar imbalance — We now consider expectation values for another local observable, which will be useful later when discussing dynamics in Sec. V. We define the imbalance with respect to some specific state, the columnar configuration represented in the inset of Fig. 5, as:

$$I = \frac{1}{N_p} \sum_{\square_p} e^{i\phi_p} \left( 2 \langle \square_p | \hat{J} | \square_p \rangle + \langle \square_p | \hat{J} | \square_p \rangle - \hat{J} \right).$$

(2)

The phase $\phi_p$ carried out by each plaquette is simply chosen to be $0$ or $\pi$ on a columnar arrangement (see inset of Fig. 5) such that the imbalance of this perfect columnar state is maximal and equal to $I = 1$. At thermal equilibrium and infinite temperature, the imbalance vanishes on average.

We display in Fig. 5 the probability distribution $P(I)$ of $\langle n|I|n\rangle$ taken over eigenstates $|n\rangle$ located at $\epsilon = 0.5$ for different values of disorder and computed on the tilted square lattice with $N = 40$ sites. For small values of disorder (left column), the probability distribution is gaussian around the mean value 0, as expected for an ETH phase. The typical width of the distribution (which is very small for the smallest disorder $V = 1$) broadens with disorder. Close to $V \approx 15$, peaks for specific values of the imbalance start to appear in the distribution which become clearly visible for all strong values of disorder $V \geq 20$ (right panels). The distribution of imbalance for pure dimer configurations $P_{\text{conf}}(I)$ (obtained by

![Diagram](image-url)

FIG. 6. Average participation entropy divided by the log of the size of Hilbert space as a function of disorder for different system sizes. Square samples are color highlighted.
data for $S_p/\log \mathcal{H}$ for different sizes tend to regroup themselves around $V \simeq 15-20$, consistent with previous estimates of a transition point.

E. Entanglement entropy

Many-body eigenstates in ergodic or localized phases have strikingly different entanglement properties. Consider the von Neumann entanglement entropy of an eigenstate $|n\rangle$ for a system divided in two parts $A$ and $B$:

$$S(|n\rangle) = -\text{Tr}_A \rho_A \log(\rho_A)$$

with $\rho_A$ the reduced density matrix obtained by tracing out bond degrees of freedom in $B$ $\rho_A = \text{Tr}_B |n\rangle \langle n|$. For our samples, we choose equal parts $A$ and $B$ containing identical number of lattice bonds and sites (see Appendix for a graphical representation of the entanglement cut).

The entanglement entropy $S$, averaged over eigenstates and realizations of disorder, is represented in Fig. 7. In the ETH phase, a volume law of the entanglement entropy is expected with $S$ scaling with the number of degrees of freedom in $A$. This is what is found in the low-$V$ regime of Fig. 7, with the following caveat: due to the constrained nature of the Hilbert space, not all configurations of $A$ are compatible with those of $B$ (see a discussion in Ref. 87 and in the Appendix) and the number of these ‘constraint sectors’ in $\rho_A$ depends in particular on the area between $A$ and $B$. This results in different samples with the same volume having different constraints on $\rho_A$ when bipartitioned. This is clearly seen when computing entanglement entropy on random states in the different samples denoted as $S_{\text{random}}$ and represented by the star symbol in

![Fig. 7. Average entanglement entropy $S$ of subsystem $A$ as a function of disorder for different system sizes. The star symbol on the $V = 0$ axis is the entanglement entropy of a random state $S_{\text{random}}$. Inset: Area law scaling of the average entanglement entropy divided by the size of the surface $A$ between $A$ and $B$ as a function of disorder (see Appendix for definition of $A$ for tilted clusters). Square samples are color highlighted.](image)

![Fig. 8. Standard deviation of entanglement entropy std$(S)$ of subsystem $A$ as a function of disorder for different system sizes, scaled by the entanglement entropy of a random state $S_{\text{random}}$. Square samples are color highlighted.](image)

![Fig. 9. Entanglement entropy $S$ of subsystem $A$ as a function of disorder for different system sizes. The star symbol on the $V = 0$ axis is the entanglement entropy of a random state $S_{\text{random}}$. Inset: Area law scaling of the average entanglement entropy divided by the size of the surface $A$ between $A$ and $B$ as a function of disorder (see Appendix for definition of $A$ for tilted clusters). Square samples are color highlighted.](image)

IV. MACHINE LEARNING ANALYSIS OF ENTANGLEMENT SPECTRA

As a complementary approach, we study the quantum dimer model Eq. 1 using machine learning techniques and follow the supervised scheme introduced in Ref. 89. As inputs representative of the two phases, we provide entanglement
spectra obtained deep in the ETH and the MBL phases and we train a neural network to classify them accordingly. The supervised learning has been used in multiple occasions in the context of delocalization-localization transitions\textsuperscript{90–92} producing results in good qualitative agreement with more conventional analyses. In particular, Refs. 90, 92, and 93 showed the interest of using labeled entanglement spectra.

For this scheme, we consider the $N = 32, 36, 40$ square samples and the largest rectangular sample $N = 6 \times 8$. For each of them, we provide 10,000 entanglement spectra (2000 for $N = 40, 6 \times 8$) including between 100 and 200 disorder realizations per disorder strength at $V = 1$ for the ETH-labelled phase and $V = 30$ for the MBL-labelled phase. The spectra being rather large (1972 values for $N = 32$, up to 21286 for $N = 6 \times 8$), we found that sorting them allowed for both perfect training and test accuracies for all system sizes. We used a fully-connected neural network consisting of one hidden layer of 32 neurons followed by two softmax output neurons.

We follow a cross-validation procedure where we randomly selected half of the dataset to form the training dataset, the rest being assigned to the test set. This process is repeated multiple times, generating new training and test partitions each time. This allowed us to track whether the neural networks were overfitting depending on the training conditions. Namely, we checked that data from $V = 1$ and $V = 30$ give perfect training and test accuracies for each system size.

Fig. 9 displays features that are consistent with the previous analysis. The left panel displays the average output of the neural network for the different samples as a function of disorder strength. At low $V$, the machine learning analysis validates a fully-ETH phase (i.e. where all samples are classified ETH) that extends up to a value $V = V_1$, and at large $V$ a fully-MBL phase (with more than 99% accuracy) at large $V \geq V_2$ ($V_1 \approx 6$ and $V_2 \geq 20$ for the largest $N = 6 \times 8$ sample). Notice how these bounding values (in particular $V_1$) shift to higher values of disorder with system size. This reflects that ETH is easier to disrupt on a too small sample, in perfect agreement with the trend in all other observables discussed in Sec. III. We find no crossing point with system size in the current data, different from what is observed with the 2d Ising model\textsuperscript{99}.

The right panel of Fig. 9 shows the standard deviation of the neural network output as a function of disorder. The standard deviation is low in both limits where the phases are well distinguished (at low and large $V$), and peaks at an intermediate value of disorder. The location of the peak (which shifts with system size) is the point where the neural network has most difficulties to classify phases. It can be interpreted as the finite-size estimate of the transition point. Notice the similarity between the standard deviation of the neural network output and the standard deviation of the entanglement entropy (Fig. 8), in particular the positions of the peaks are almost the same for both quantities for the different sample sizes.

In conclusion of this section, we find that a neural network only fed with entanglement spectra is able to learn how to correctly distinguish the ETH and MBL phases for the quantum dimer model with random potential Eq. 1 as well as to provide finite-size estimates of the transition point between the two. This automated method gives results in very good agreement with the analysis based on more standard, feature-enginered, estimators of the phases presented in the previous Sec. III. Our results also indicate that the machine learning analysis is also subject to finite-size effects, indicating that the use of numerical data coming from a single size may not be able to provide quantitative results for the study of phase transitions. One noticeable interest of the neural network analysis (already pinpointed earlier\textsuperscript{90,92}) is that the required amount of data and overall computational cost is considerably lower than with more traditional observables to obtain approximately similar quality of prediction: for instance good statistics on the gap ratio (Fig. 2) requires $\approx 100$ times more realizations of disorder than with the machine learning analysis (Fig. 9).

V. DYNAMICAL PROPERTIES AFTER A QUENCH

In this section, we consider dynamical properties after a quench of the model Eq. 1, in order to probe for the arrested transport dynamics or slow propagation of information expected in a MBL phase. We consider quenches from an initial product state (with no entanglement) and take here simply the columnar configuration depicted in the inset of Fig. 5 (a product state). This columnar configuration is the most flippable state, and is thus well connected to other dimer configurations by the action of the Hamiltonian. We thus intuitively expect this state to rapidly thermalize if it does. The choice of a columnar configuration echoes the experimental protocols for which the system is often intialized in a charge density state\textsuperscript{30,53}, or numerical simulations which choose a N\'eel state in spin chains, both states sharing the same properties as the columnar configuration.
We start with the dynamics of the columnar imbalance defined in Eq. 2. By definition, $I(t = 0) = 1$, whereas $I = 0$ when averaged over all eigenstates (see Fig. 5). In the long-time limit, $I(t)$ is thus expected to vanish in a thermal phase, whereas a non-vanishing long-time value would signal a localized phase. Note that on finite systems, $I(t \to \infty)$ never strictly vanishes. Fig. 10 displays the columnar imbalance as a function of time, with panels corresponding to different disorder strengths $V$. In the low-disorder limit, the imbalance quickly decays to a vanishing value on short time-scales (this is most readily seen at $V = 1$). When disorder is increased, decays become slower and for intermediate disorder values $V = 10, 15, 20$, the saturation value is not reached for the largest clusters within the time scale studied (for $N = 32, 36$ data not shown at larger time show that the saturation is reached within less than a percent for $t = 1000$ for all disorders considered). We can nevertheless observe that the imbalance clearly decreases with system size for the largest time available, in a manner compatible with a vanishing value in the thermodynamic limit for $V \lesssim 15$. For larger values of disorder, one clearly observes that the saturation value is much larger and, within available time and cluster sizes, the imbalance does not tend to vanish: this is most speaking for $V = 30$ where more than 80% of the initial imbalance is retained even on the largest clusters. Pinpointing exactly a transition point with dynamics is difficult due to the time scale limitations for the largest clusters, but the present data appear to be compatible with a transition for $V$ in-between 15 and 20, similar to what is found with the analysis of eigenstate data in Sec. III.

The nature of the growth of entanglement entropy $S(t)$ after a quench is an important feature to distinguish MBL in one-dimension, where it takes a logarithmic form with time\textsuperscript{18,19}, from an ETH regime (with ballistic growth $S(t)$\textsuperscript{96}) or an Anderson localized regime (where it saturates to a constant $O(1)$). In both MBL and ETH regimes, the saturation value at long time is expected to fulfill a \textit{volume} law, with a prefactor equal to (smaller than) the corresponding thermal entropy for the ETH (MBL) case\textsuperscript{97,98}.

We represent in Fig. 11 the dynamics after a quench from the columnar configuration of the average entanglement entropy $\bar{S}(t)$ normalized by the entanglement entropy of a random state $S^{\text{random}}$ (which is the thermal entropy up to $O(1)$ corrections) for different disorder strengths. The ETH regime with a rapid growth of entanglement is readily identified at low disorder $V = 1$, where $\bar{S}(t)$ reaches its saturation value (close to the maximum value reached by a random state) in less than 10 plaquette flips. In the top panels where the same scale is used for the vertical axis, the entanglement entropy is observed to grow slower towards its saturation value, with visible finite-size effects. Again the time and size limitations do not strictly allow to draw definitive conclusions, but it appears very likely that for large enough clusters, the entanglement entropy will reach its thermal value at long time for $V \lesssim 15$. On the bottom panels (disorders $V = 20, 25, 30$) on the other hand, the entanglement entropy develops at an even much slower pace. The logarithmic horizontal scale chosen for the bottom panels highlights the striking feature that the growth of entanglement entropy appears to behave as $\log(t)$ for the largest clusters which are not hampered by finite-size satura-
FIG. 11. Average entanglement entropy $S(t)$ after a quench from the columnar configuration, scaled by the (extensive) entanglement entropy of a random state $S_{\text{random}}$, for different values of disorder $V$ and different system sizes $N$ of square clusters. The top four panels are represented using the same vertical scale, and a linear scale for the time axis. The bottom three panels use a logarithmic scale for the time axis, highlighting the slow growth of entanglement for the largest values of disorder.

FIG. 12. Growth of the average participation entropy $S_p(t)$, normalized by its maximal value $\log(H)$ after a quench from the columnar configuration for different values of disorder $V$ and different system sizes $N$. As in Fig. 11, the top four panels use the same vertical scale, and a linear scale for the time axis whereas the bottom three panels use a log scale for the time axis.

We note that if we naively export the argument (based on the exponentially-decaying interactions between local integral of motions) that explains the logarithmic growth of entanglement in 1d MBL, we would expect a $\log^2(t)$ growth of entanglement in 2d MBL, which is not compatible with the data at hand within our numerical limitations.
We finally present data for the evolution after a quench of the participation entropy $S_p(t)$ in Fig. 12. The dynamics of the participation entropy is almost identical to the one of the entanglement entropy discussed above: at small disorder, rapid saturation to the maximum attainable value $\log(\mathcal{H})$, slower growth at intermediate values and logarithmic growth for the largest samples for strong disorder (note again the logarithmic scale on the bottom panels of Fig. 12). The only noticeable difference in the dynamics of both entropies comes from different finite-size effects at short time, with participation data showing an improved nesting of curves as system size is increased which is not visible for entanglement. From participation entropy, we conclude again on slow dynamics starting from $V \gtrsim 15 - 20$ consistent with the existence of a localized phase.

VI. DISCUSSION AND PERSPECtIVES

We presented a comprehensive and extensive large-scale exact numerical study of the localization properties of a two-dimensional constrained quantum many-body system with disorder. We interpret all the data, both for eigenstate or dynamical properties, as consistent with the existence of two distinct phases: an ETH phase at low disorder and a many-body localized phase at strong disorder, which are separated by a transition located around disorder strength $V \sim 15 - 20$.

Our evidence for a MBL transition in the 2d quantum dimer model comes from numerical simulations on finite lattices. Of course, finite-size simulations can always be argued to artificially detect a MBL phase even when only a ETH phase occurs in the thermodynamic limit. We would like to emphasize that the level of numerical evidence for a MBL transition in the 2d quantum dimer model is in our opinion and experience similar to the one obtained for the standard model of MBL in one dimension$^{21,22}$, with similar or larger Hilbert space sizes and time scales probed in the current work. Also even if the behavior at large values of disorder (say $V = 30$) is ultimately ergodic, the time scales and/or system sizes needed to probe ergodicity would be extremely large. An experimental realization of Eq. 1 or of a similar constrained model in 2d would see localization for all practical purposes on the time scales available in the lab. We deliberately avoided to attempt to perform a data collapse analysis to extract e.g. critical exponents due to the two following limiting points: the moderate linear system sizes that we can reach as well as the non-uniform aspect ratio between available samples would likely provide biased estimates for critical exponents. For one-dimensional MBL, the larger linear length scales that can be reached have been argued (see e.g. Ref. 99) not to be large enough to provide correct estimates of asymptotic critical behavior. The situation is likely to be the same here. With these numerical limitations in mind, we can nevertheless observe that critical values of the gap ratio or of the Kullback-Leibler divergence of eigenstates are closer to their Poisson than their ETH limits, indicating that the transition point is even less ergodic than for the one-dimensional MBL transition in the random-field Heisenberg spin chain$^{22}$. Going back to the $l$-bits arguments which predict the asymptotic unstaibility of MBL in two dimensions$^{43-45}$, we remark that we do not know how to simply construct $l$-bits for the quantum dimer model. The crucial point is that there is no tensor product structure for the Hilbert space of quantum dimer models, thus constructing a basis of commuting operators with local support appears ill-defined. As the bubble argument is explicitly based on $l$-bits, it cannot strictly speaking apply here.

Regarding experiments, the relative strong disorder needed to induce the MBL phase does not provide immediate relevance of our results in frustrated magnets, where QDMs naturally emerge. However, the artificial realizations of quantum dimer models and other related 2d constrained models have been extensively studied in cold-atomic and mesoscopic systems. Precise experimental protocols have been proposed using Ryberg atoms$^{100}$, trapped ions$^{101}$, large-spin ultracold atoms$^{102}$, Josephson Junction arrays$^{103}$ or superconducting quantum circuits$^{104}$. Further explorations are needed to understand which experimental platforms are the most suited to implement random potential or interactions. Constrained models in 1d have already been implemented (with no randomness) exploiting Rydberg blockade in atomic chains$^{60}$ already exhibiting rich unusual quantum dynamics.

There are several perspectives opened up by our work. First, the roadmap to two-dimensional MBL can be exploited using the QDM on other two-dimensional lattices, allowing to test for universality and to search for other features. Two interesting cases are worth mentioning: (i) the honeycomb lattice, which has an effective smaller local Hilbert space allowing in principle to reach larger samples, (ii) the kagome lattice, which possesses conserved $Z_2$ quantum numbers, allowing the exciting possibility of 2d topological order in MBL states$^{76}$. The possibility of MBL in other quantum constrained models such as quantum ice or loop models also provide an interesting follow-up.

The large-scale numerical techniques that we use in the present work would be useful to investigate other aspects of the model Eq. 1 not addressed here, in particular to make connection with 1d models where MBL is best understood. The current investigation focused on the ‘infinite temperature limit’, that is excited states in the middle of the spectrum $(\epsilon = 0.5)$. It would be interesting to construct a full energy-disorder phase diagram to see if a many-body mobility edge exists as in 1d spin chain models$^{22}$. Also the detailed dynamics occurring in the ETH phase remains to be investigated. The relatively slow dynamics exhibited around $V \sim 10, 15$ in Figures 10, 11, 12 could be signs of a precursor subdiffusive regime, similar to what is observed in 1d systems$^{105-107}$. Finally, it would be intriguing to see if the machine learning techniques used in Sec. IV would be able to distinguish 1d from 2d MBL, and if not, to use neural networks trained on 1d spin chain models to probe 2d MBL (transfer learning).

On the analytical side, the absence of an obvious scheme for constructing $l$-bits points towards the interest of building an alternative effective description for the many-body localized phase in such 2d constrained models. Recall as well that there is no free-particle limit to start with in the model.
Eq. 1. The height description\textsuperscript{108} of dimer coverings (and of constrained models in general) appears as a good candidate to develop such a framework. Finally from a purely classical perspective, we note that the strong disorder limit $V \to \infty$ constitutes an interesting classical statistical mechanics problem (with for instance non-trivial degeneracies), unsolved to the best of our knowledge and distinct from other interacting dimer models\textsuperscript{109}.

ACKNOWLEDGMENTS

We gratefully thank Gabriel Meyer, Juan P. Garrahan and Stephen Powell for participation and useful discussions at an initial stage of this project. We also thank Nicolas Laflorencie for useful comments. H.T. is supported by a grant from the Fondation CFM pour la Recherche. Z.L. acknowledges support from EPSRC Grant No. EP/M019691/1. This work benefited from the support of the project THERMOLOC ANR-16-CE30-0023-02 of the French National Research Agency (ANR) and by the French Programme Investissements d’Avenir under the program ANR-11-IDEX-0002-02, reference ANR-10-LABX-0037-NEXT. We acknowledge PRACE for awarding access to HLRS’s Hazel Hen computer based in Stuttgart, Germany under grant number 2016153659, as well as the use of HPC resources from CALMIP (grants 2017-P0677 and 2018-P0677) and GENCI (grant x2018050225). The shift-invert\textsuperscript{72} and Krylov time evolution numerical calculations are based on the linear algebra libraries PETSc\textsuperscript{110,111}, SLEPc\textsuperscript{112}, MUMPS\textsuperscript{113,114} and STRUMPACK\textsuperscript{115,116}. The neural network analysis used the library TensorFlow\textsuperscript{117}.

Appendix: Lattices

1. Lattice geometries

We use regular and tilted square lattices as well as rectangular lattice with $N$ sites and $2N$ bonds. Regular square lattices have $N = L^2$ sites, and we used samples with $L = 4, 6, 8$. Rectangular lattices, which have $N = L_x \times L_y$ sites (with $L_{x, y} = 4, 6, 8, 10$), have a non-unity aspect ratio, which can affect the analysis of finite-size effects. Finally, tilted square lattices host $N = p^2 + q^2$ sites, and we use the combination $p = 4, q = 4$ (with $N = 32$ sites), $p = 6, q = 2$ ($N = 40$), and $p = 6, q = 4$ ($N = 52$). Tilted square lattices have a unity aspect ratio, but they do not have all symmetries of the full regular square: this is however irrelevant as the Hamiltonian Eq. 1 does not allow any lattice symmetries due to disorder. All these lattices, which are represented in Fig. 13, have in common that they host a $(W_x, W_y) = (0, 0)$ winding sector: this is true when $L, L_x, L_y, p$ and $q$ are even. One can also easily define columnar state on these samples (see example for $N = 52$ in Fig. 13).

2. Entanglement cut

When computing bipartite entanglement properties, there is some freedom in choosing where is located the entanglement cut, \textit{i.e.} which degrees of freedom belong to $A$ or $B$ in Eq. 4. We choose the cuts displayed in Fig. 13, where all lattices are divided in two parts with equal number of bonds. With this cut, the area $A$ between $A$ and $B$ is expected to scale with $L$ for regular square lattices, $L_y$ for rectangular lattices, and $\sqrt{p^2 + q^2}$ for tilted square lattices.

This choice of cuts for the regular square, rectangular and $N = 32$ tilted square lattice also induces the following property: one of the two windings numbers (the one associated to a line parallel to the cut, e.g. $W_y$ for the rectangular lattices) is conserved by the cut, but not the other one. The reduced density matrix on $A$ is block-diagonal with blocks labelled by winding numbers in $A ($W_A^A$ which is forced to be 0, and $W_A^B$), which we take advantage of in numerical computations. The existence of blocks in the reduced density matrices and its effect on entanglement has been discussed in other models with conservation laws\textsuperscript{87,118,120}.

We note also a supplementary specificity of constrained models for entanglement properties (see recent discussion in Ref. 87). Due to the constraint and even within the conserved-number blocks \textit{e.g.} configurations of $A$ with $W_A^A$ and configurations of $B$ with $W_B^B$ such that $W_A^A + W_B^B = W_A^B$, not all configurations of $A$ are compatible with those of $B$. This property has an important consequence, for instance when computing volume-law entanglement, as different samples with the same volume can have different constraints on configurations in $A$ and $B$ when bipartitioned. This is illustrated in the main text with the example of the $8 \times 4$ and $4 \times 8$ rectangular samples. This property can also cause difficulties in comparing scaling of entanglement with system size for different samples.

![FIG. 13. Lattices used in this work: regular square lattices with $N = L^2$ sites (top left panel with $L = 6$), rectangular lattices with $N = L_x \times L_y$ (top middle panel with $L_x = 6$ and $L_y = 8$), and tilted square lattices (bottom row with $N = 32, 40, 52$ sites). The red/blue bond decomposition is used for the entanglement partition. Top right: columnar state for the tilted $N = 52$ square lattice.](image-url)
