Strong ergodicity breaking due to local constraints in a quantum system

Sthitadhi Roy$^{1,2,*}$ and Achilleas Lazarides$^3,†$

$^1$Rudolf Peierls Centre for Theoretical Physics, Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, United Kingdom
$^2$Physical and Theoretical Chemistry, Oxford University, South Parks Road, Oxford OX1 3QZ, United Kingdom
$^3$Interdisciplinary Centre for Mathematical Modelling and Department of Mathematical Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, United Kingdom

Quantum systems that violate the eigenstate thermalisation hypothesis thereby falling outside the paradigm of conventional statistical mechanics are of both intellectual and practical interest. We show that such a breaking of ergodicity may arise purely due to local constraints on random many-body Hamiltonians. As an example, we study an ergodic quantum spin-1/2 model which acquires a localised phase upon addition of East-type constraints. We establish its phenomenology using spectral and dynamical properties obtained by exact diagonalisation. Mapping the Hamiltonian to a disordered hopping problem on the Fock space graph we find that potentially non-resonant bottlenecks in the Fock-space dynamics, caused by spatially local segments of frozen spins, lie at the root of localisation. We support this picture by introducing and solving numerically a class of random matrix models that retain the bottlenecks. Finally, we obtain analytical insight into the origins of localisation using the forward-scattering approximation. A numerical treatment of the forward-scattering approximation yields critical points which agree quantitatively with the exact diagonalisation results.

I. INTRODUCTION

The fundamental question of how closed quantum systems admit a thermodynamic description has motivated the study of their excited eigenstates and out-of-equilibrium dynamics [1, 2]. A central concept is the Eigenstate Thermalisation Hypothesis (ETH), satisfied by generic ergodic systems [3–6]. Positing that eigenstate expectation values of local observables are smooth functions of the eigenenergies, ETH amounts to a statement that the energy, an integral of motion, is a state variable so that local observables in the long-time dynamical state are fully determined by its value. Generic systems satisfy this by default, and any violation of the ETH is therefore interesting. In a recent development it was realised that one way ETH can be violated is the presence of many-body localisation (MBL) [7–11] (see Refs. [12–14] for reviews and further references therein). Specifying macroscopic properties of MBL systems requires an extensive set of emergent quasi-local integrals of motion [15–19].

Many of the universal properties of ETH systems are well described within the framework of random matrix theory [3–6, 20], where one studies random matrices incorporating the relevant symmetries of the physical system instead of actual Hamiltonians. The physical intuition here is that a random matrix is the “least structured” object which still captures the characteristic properties of the Hamiltonians of interest. Inspired by this approach, and focusing on Hamiltonians with finite local Hilbert space dimensions (spin-1/2 systems constituting possibly the simplest example), we ask what properties a random many-body Hamiltonian matrix must satisfy so as to capture the defining properties of MBL systems, namely, violation of ETH and absence of spatial transport.

As the question pertains to the many-body Hamiltonian, it is natural to treat the problem directly as one of localisation in Fock space [21–30]. Any many-body Hamiltonian of our type can be interpreted as a disordered hopping problem on the Fock space of the system.

* sthitadhi.roy@chem.ox.ac.uk
† A.Lazarides@lboro.ac.uk
1 Assuming no other conserved quantities, for example, total spin, momentum, particle number etc.

(a)
↓↓↓↓↑
↓↓↓↑↑
↓↓↑↑↑
↓↑↑↑↑
↓↑↓↑↑
↓↑↓↓↑
(b)

FIG. 1. The Fock-space graph for a system of $N(=5)$ spins-$1/2$ where each vertex is a $σ^z$-product state as indicated and the links correspond to spin-flip terms in the Hamiltonian. (a) The links in red are present for both the QREM as well as for the EastREM whereas the those in grey are present only for the QREM; the constraints in the EastREM switches off the latter. (b) Illustration of how the constraints generally increase the shortest distance between two vertices. For the pair of vertices chosen, the path with the QREM is only of length one (green) whereas for the EastREM the shortest path shown in black is longer.
Considering the Fock basis states as vertices and the hop-
pings as links between them, one can view the Hamilt-
onian matrix as a graph (see Fig. 1). The question then
translates to what properties the graph must have for the
system to exhibit non-ergodic behaviour.

In this paper we provide one answer to the question: con-
strained hopping on the Fock space due to local kinetic
constraints on the real-space degrees of freedom can lead
to localisation. Such constrained dynamics in translation
invariant systems have been to show exhibit slow dy-
namics and metastable behaviour [31–33], but here we
focus on the possibility of a localised phase and accom-
panying localisation transitions in the eigenstates. Using
exact numerical and approximate analytical techniques
we demonstrate that constraining Fock-space connectiv-
ity leads to localisation, and that this is not due to dis-
order as the unconstrained version of our model is disor-
dered but never localised. The Fock space of our model
is not fragmented but is rather fully connected, thus the
physics here is qualitatively different from cases where
the constraints fragment Fock space [34–36]. Further-
more our model breaks ergodicity strongly, as signified
by the presence of a phase where all eigenstates are lo-
calised, thus also differing from weak ergodicity break-
ing as in the case of quantum many-body scars [37–39].
In our case, ergodicity breaking is due to the states in
Fock space naturally grouping into clusters, with dense
intra-cluster but sparse inter-cluster connections. This
leads to potentially non-resonant bottlenecks in the Fock
space, which is the root of localisation. This establishes
the central result of this work – how constrained quantum
dynamics can lead to localisation without fragmenting the
Fock space.

Structure of the paper

We start with an overview of the paper in Sec. II which
sets up the Fock space we work with and states the main
results of this paper. In Sec. III we introduce a kinetically
constrained quantum spin-1/2 model to put the ideas on
a concrete footing. The constrained quantum model is
based on the quantum random energy model (QREM) which acts as our reference unconstrained model. The
QREM has been shown to be completely delocalised ex-
cept for a vanishing fraction of eigenstates at the spec-
tral edges [40–43]. We then impose East model-like con-
straints [44–46], calling the resulting model the EastREM. Section IV is dedicated to the phenomenology of
the model: we map out its phase diagram using spec-
tral and eigenstate properties in Sec. IV A, finding that
a fully localised phase emerges, unlike for the QREM.
Dynamical properties further support this as shown in
Sec. IV B where we find that an initial state retains its
memory locally in space, reflecting the locality of the con-
straints. In Sec. V we discuss how the constraints impose
a particular structure in the Hamiltonian and construct
Hamiltonians random apart from having this structure,
showing that they still display the dynamics of interest.
Finally, in Sec. VI we use the forward scattering approx-
imation (FSA) to obtain a (semi)-analytical understand-
ing of localisation in this model: Secs. VI A and VI B
present an analytical treatment of the FSA for the spec-
tral bulk and edges respectively, while Sec. VI C presents
a numerical treatment of the FSA, finding agreement
with the numerical results of Sec. IV. The FSA explicitly
demonstrates the role of constraints and reveals clearly
the distinction between the unconstrained (QREM) and
constrained (EastREM) versions of the model, explicitly
demonstrating the role of the constraints.

II. OVERVIEW

Fock space offers a natural viewpoint from which to
approach the problem as any many-body Hamiltonian
can be interpreted as a hopping problem on the Fock-
space graph,

$$H = \sum_{\alpha=1}^{N_H} \mathcal{E}_\alpha |\alpha\rangle \langle \alpha| + \sum_{\alpha \neq \beta} \Gamma_{\alpha\beta} |\alpha\rangle \langle \beta|.$$  (1)

Here the set of basis states \{\langle \alpha|\} are the sites on the \(N_H\)-
dimensional Fock-space graph, of which there are expo-
entially many (in system size), \(N_H \sim e^N\). The diagonal
elements of the Hamiltonian \(\mathcal{E}_\alpha\) are the on-site energies
in this Fock space. The off-diagonal elements \(\Gamma_{\alpha\beta}\) then
represent hopping amplitudes. The off-diagonal part of
the Hamiltonian, \(H_{\text{offdiag}}\) also allows us to define a dis-

cance on the Fock space graph – the distance between
two states \(|\alpha\rangle\) and \(|\beta\rangle\), denoted as \(r_{\alpha\beta}\), is defined as
the length of the shortest path between them following the

links generated by \(H_{\text{offdiag}}\).

Hamiltonian matrices as in Eq. (1) which are associ-
ated with a many-body system (short-ranged with lo-
cal degrees of freedom) in general have matrix elements
which satisfy two generic features. First, the Fock-space
site energies scale as \(\sqrt{N}\), such that one can define an
effective on-site disorder strength on the Fock space as
\(W_{\text{FS}}^2 := N_H^{-1} \sum_\alpha (\mathcal{E}_\alpha^2) = W^2 N\) and \(W\) an \(O(1)\)
number. This simply reflects that for generic short-ranged
systems, each Fock-space site energy is an extensive sum
of random numbers. Secondly, the off-diagonal matrix

lement numbers are of magnitude \(O(1)\) and, crucially,
the average connectivity of the Fock-space sites is exten-
sive: \(N_H^{-1} \sum_{\alpha \neq \beta} (\Gamma_{\alpha\beta}^2) \sim N\). This is a result of the fact
that for short-ranged systems, the Hamiltonian connects
a state with an extensive number of different states each
differing from the initial one only locally.

\(^2\) The adjacency matrix of this graph is closely related to the
Hamiltonian expressed in the Fock basis.
If all the Fock-space site energies are independent of each other, Eq. (1) can be interpreted as an Anderson localisation problem on a graph with connectivity $N$, hopping amplitude $\Gamma$, and disorder strength $W_{FS}$. Applying the localisation criterion for Bethe lattices \cite{47}, which we expect to work well for cases with diverging connectivity, one finds the critical disorder strength $W_{c}$ satisfies $2\sqrt{\ln N} \ln \left(\frac{W_{c} \sqrt{N}}{2\Gamma}\right) = 1$, so that $W_{c}$ diverges in the thermodynamic limit.\footnote{The expression is obtained from the so-called ‘upper limit approximation’, however the exact solution was shown to differ from it by a factor $\approx e/2$ \cite{47, 48}.} A localised phase therefore does not exist, at least in the bulk of the spectrum. We therefore ask what additional ingredients are minimally required to stabilise a many-body localised phase without altering the generic features mentioned above.

Elsewhere \cite{30}, one answer to this question was provided: strong correlations in the $E_{\alpha}$, which render the problem fundamentally different from an Anderson localisation problem on a high-dimensional graph.\footnote{Here, strong correlations means that two basis states $|\alpha\rangle$ and $|\beta\rangle$ finitely distant from each other have Fock-space site energies finitely different from each other in the thermodynamic limit, $|E_{\alpha} - E_{\beta}| \sim O(1)$.} In fact, this is precisely the scenario for local Hamiltonians where the presence of a localised phase has been argued for on analytical as well as numerical grounds \cite{9, 11, 23, 30, 49–51}.

In this work, we take the complementary perspective and show that, depending on the pattern and distribution of connectivities, a fully localised phase may occur even for completely uncorrelated Fock-space disorder, non-fractured (i.e., fully-connected) Fock space, and typically extensive connectivity for each site. We demonstrate this for the case of spatially local kinetic constraints, which create bottlenecks in the Fock space but leave it fully connected (every site is accessible from every other).

Although our Fock space is not fragmented, it can be reorganised into sparsely connected clusters. The picture that emerges is one of sites densely interconnected within each cluster, but sparse interconnections between clusters. In other words, the constraints suppress links between sites belonging to different clusters. We show that this is the fundamental mechanism which leads to a fully many-body localised phase in both real and Fock spaces, despite the Fock-space site energies being uncorrelated and the Fock space not being fragmented — this constitutes the central result of this work.

As a concrete setting we consider a system of $N$ quantum spins-1/2 (denoted by the set of Pauli matrices, $\{\sigma^{i}\}$) where the Fock-space basis states are simply the classical configurations – product states in the $\sigma^{z}$-basis. Assigning independent random energies to the $2^{N}$ configurations leads to the random energy model (REM) \cite{52} which, upon addition of spin-flip terms $\sigma^{x}$ to the Hamiltonian becomes the QuantumREM (QREM). This will be our reference unconstrained model and has no localised phase in the bulk of the spectrum. Imposing East-like constraints in the spin-flip terms, that is, allowing a particular spin flip only if the spin to its right is pointing up, results in a constrained model which we call the EastREM. The construction of the model and a discussion of implications of the constraints for the structure of the connectivity of Fock space constitutes Sec. III.

The phenomenology of the model is established in Sec. IV. We present results for the statistics of level spacing ratios and participation entropies of the eigenstates on the Fock space which reveal a phase diagram with a fully localised phase. Dynamical autocorrelations from time evolving an initial product state also show non-ergodic behaviour in the form of retention of memory of initial configuration. In fact, the real-space profile of the dynamical autocorrelation directly reflects the effect of the corresponding local kinetic constraints. Finally, we identify the clusters made up of densely connected states and then construct a Hamiltonian matrix where the clusters are described by GOE Hamiltonians but the matrix elements connecting different clusters are as for the EastREM. This random matrix analogue to the EastREM, which we call GOEastREM, displays the relevant features of the EastREM, demonstrating that the clustering is the crucial ingredient.

Analytical insights into the origin of the localisation on the Fock space graph are obtained from the FSA, discussed in Sec. VI. The FSA is an approximation for the non-local propagator on the Fock space which takes into account the contribution only from the shortest paths between two Fock-space sites. As the constraints essentially have the effect of modifying the statistics of shortest paths on the Fock space, the FSA is ideally suited for analysing the EastREM and exposing its differences from the QREM. As elaborated in Sec. VI, two aspects of the statistics of shortest paths are crucial, (i) the scaling of the number of Fock space sites separated by distance $r$ with both system size and $r$, and (ii) the scaling of number of paths between such Fock space sites separated by $r$. These features of the Fock space are inputs to the FSA, and the results predict an appearance of localised states in the spectral bulk of the EastREM contrary to the QREM. We also corroborate the theoretical predictions from the FSA with a numerical treatment of the FSA by enumerating the directed paths on the Fock space, and we find that the critical point so obtained is concomitant with that obtained from exact diagonalisation studies of Sec. IV.

III. CONSTRAINED QUANTUM MODEL

Our prototypical model for a kinetically constrained quantum system is one made of $N$ spins-1/2, derived from the QREM by imposing constraints. The $\sigma^{z}$-product states constitute the basis states of our Fock space $|\alpha\rangle \equiv \{|\sigma^{z}_{i}\rangle\alpha\}$ and to each of them is associated
an independent random energy $\mathcal{E}_\alpha$ drawn from a normal distribution with zero mean and variance $N$. The diagonal (first) part of the Hamiltonian of Eq. (1) is given by

$$H_{\text{REM}} = \sum_{\alpha=1}^{2^N} \mathcal{E}_\alpha \langle \{\sigma_z^i\}_\alpha \rangle \langle \{\sigma_z^i\}_\alpha \rangle,$$

(2)

with $\mathcal{E}_\alpha \sim \mathcal{N}(0,N)$. Henceforth, we will use the terms spin-configuration $\langle \{\sigma_z^i\}_\alpha \rangle$ and Fock-space site $|\alpha\rangle$ interchangeably.

The QREM is obtained by adding to $H_{\text{REM}}$ unconstrained single flips generated by the Hamiltonian

$$H_X = \Gamma \sum_{i=1}^{N} \sigma_i^x,$$

(3)

which corresponds to the second (hopping) term of Eq. (1), such that the total Hamiltonian is

$$H_{\text{QREM}} = H_{\text{REM}} + H_X.$$

(4)

In terms of Fock space sites, the QREM Hamiltonian is precisely a $N$-dimensional hypercube with $N_H = 2^N$ vertices each of which has a connectivity of exactly $N$. Each of the $N$ links on any vertex corresponds to a flip of a particular spin as the single spin-flips induced by $H_X$ are unconstrained. Another direct implication of this is that for any Fock-space site, the number of Fock-space sites at a distance $r$ is $^{(N)}_r$.

Localisation or lack thereof in the QREM was studied in Ref. [42] where it was found that the model is ergodic in the spectral bulk for infinitesimally small $\Gamma$ while the spectral edges can have localised eigenstates, so that there are mobility edges at finite energy densities $\epsilon = E/N \sim \Gamma$. However, as the width of the density of states $\sim \sqrt{N}$, in the thermodynamic limit the localised eigenstates occupy only a vanishing fraction of the spectrum. Generic quantum dynamics therefore exhibit ergodic behaviour and we consider the QREM, our reference unconstrained model to be ergodic at all $\Gamma \neq 0$.

The EastREM, our model for a constrained quantum system, is obtained from the QREM by imposing local constraints of the East type such that the Hamiltonian is given by

$$H_{\text{EastREM}} = H_{\text{REM}} + H_{\text{East}},$$

(5)

where

$$H_{\text{East}} = \frac{\Gamma}{2} \sum_{i=1}^{N} \sigma_i^x (1 + \sigma_{i+1}^z)$$

(6)

where we impose periodic boundary conditions, resulting in a Fock space that is not fragmented.$^6$ The constraint modifies the hopping on the Fock space (the second term in Eq. (1)) so that it allows a spin at real-space site $i$ to be flipped if and only if the spin at site $i + 1$ is pointing up. Hence, in terms of hopping in Fock space, it has the effect of switching off all the hopping amplitudes of the QREM Hamiltonian that corresponded to a flip of a spin with the spin to its right pointing up. A visual demonstration is shown in Fig. 1(a) where the red links are present for both the QREM and the EastREM while the blue links are present only on the QREM. This has a number of consequences.

Firstly, the constraints lead to a suppression of the average connectivity, although it still scales as $N$. Secondly, the distribution of connectivities, which is a delta-function at $Z = N$, for the former it is a binomial distribution and has support on lower values of $Z$. (b) The distribution of the shortest distances from the domain wall state for $H_{\text{East}}$ (solid lines). The dashed lines show the same for the $H_X$, where it is simply the binomial distribution.

\[\text{FIG. 2. (a) The distribution of the connectivities } Z \text{ on the Fock space generated by } H_{\text{East}} \text{ (6) for different system sizes } N. \text{ The red dashed line shows the same for } H_X \text{ (3). While for the latter it is a delta-function at } Z = N, \text{ for the former it is a binomial distribution and has support on lower values of } Z. \text{ (b) The distribution of the shortest distances from the domain wall state for } H_{\text{East}} \text{ (solid lines). The dashed lines show the same for the } H_X, \text{ where it is simply the binomial distribution.} \]

$^5$ Note that for the offdiagonal part of the Hamiltonian given by $H_X$ of Eq. (3), the distance between two Fock-space sites is the same as the usual Hamming distance – the number spins different between the two configurations.

$^6$ In the case of open boundary conditions, Fock space is fragmented as if the rightmost $w > 0$ spins are all down, they remain frozen for all time.
Fig. 8), the importance of which will become apparent in Sec. VI. All of the above suggest a general tendency of the constraints to localise a state on the Fock-space. While qualitative now, these pictures will be important later when we formalise the above ideas using the FSA on the Fock-space.

In real space, a qualitative picture of the origins of localisation due to the constraints is as follows. Due to the East-like constraints, any contiguous block of down spins is slow to thermalise as it can only do so in a sequential fashion starting from the right edge of the block. Spins deep inside such blocks, say at a distance $r$ away from the right edge of the block can flip only at $r^{th}$ order in perturbation theory. By contrast, for the QREM any spin is free to flip and they can do so in any order. Furthermore, even the “liquid” regions of the chain, which are regions initially without such frozen blocks are affected by the constraints dynamically. Thermalising the “liquid” regions involves flipping the up spins to down creating new constrained regions, which eventually arrest the dynamics.

IV. PHENOMENOLOGY

A. Spectral properties and MBL phase diagram

To establish the phenomenology of the EastREM in terms of the spectral properties and obtain an MBL phase diagram we use two commonly studied numerical diagnostics: statistics of level spacing ratios and participation entropies of the eigenstates on the Fock space.

The level spacing ratio, $s_n$, is defined as $[9, 11, 53]$

$$s_n = \min(\Delta_n, \Delta_{n+1})/\max(\Delta_n, \Delta_{n+1})$$

with $\Delta_n = E_n - E_{n-1}$, where the $E_n$s denote the consecutive eigenenergies. For an ergodic system, $s_n$ has a Wigner-Dyson distribution, reflecting the presence of level repulsions, so that $\langle s \rangle \approx 0.53$. A localised system on the other hand has uncorrelated eigenvalues resulting in $s_n$ having a Poisson distribution and $\langle s \rangle \approx 0.386$.

The eigenstates on the Fock space also carry signatures of ergodicity breaking $[49, 54, 55]$. The $q^{th}$ participation entropy of an eigenstate $|\psi\rangle$ defined via $S_q^p(|\psi\rangle) = \frac{1}{q-1} \ln \left[ \sum_\alpha |\langle \alpha |^q \right]$ scales as $[49]$

$$S_q^p(|\psi\rangle) = a_q \ln N_H + b_q \ln \ln N_H. \quad (7)$$

In the ergodic phase $a_q \approx 1$ as a consequence of the eigenstate being spread over the entire Fock space whereas in the MBL phase $a_q < 1$ indicating that the support of the eigenstate is a vanishing fraction of the Fock space dimension in the thermodynamic limit.

Numerically analysing the two diagnostics using exact diagonalisation we obtain the MBL phase diagram in the $\epsilon$-$\Gamma$ plane shown in Fig. 3. We emphasise that the density of states is a Gaussian with a width proportional to $\sqrt{N}$. Hence, any finite energy density corresponds to the edges of the spectrum where only a vanishing fraction of the eigenstates live in the thermodynamic limit (see Fig. 3a). It is the middle of the spectrum, $\epsilon = 0$, defined via $\text{Tr}[H]$, which determines the generic dynamical behaviour of the system.

The critical $\Gamma$ can be obtained from the mean level spacing ratio by collapsing the data for various $N$ onto a common function of $g(\Gamma - \Gamma_c)N^{1/\nu}$. Such an exercise leads to the set of critical $\Gamma_c$ at different energy densities shown by the black circles in Fig. 3(b). Representative plots of the raw data of the mean level spacing ratios in the spectral bulk and edges are shown in panels (c) and (d) respectively.

The critical line in the $\Gamma$-$\epsilon$ plane so obtained shows a good agreement with that of the deviation of $a_1$ from 1, the second diagnostic for the MBL transition. For the EastREM, a clear MBL phase emerges at $\epsilon = 0$ with a transition to the ergodic phase at $\Gamma_c \approx 0.17$. This is qualitatively different from the QREM where at $\epsilon = 0$, the model is ergodic at all finite values of $\Gamma$. Additionally, in the spectral edges (finite $\epsilon$), the transition from the MBL to ergodic phase occurs at a larger value of $\Gamma$ in the EastREM compared to the QREM; this indicates a parametric increase of the robustness of localised phase in the presence of the constraints.

B. Non-ergodic dynamics

As a dynamical signature of ergodicity breaking, we study the autocorrelation function

$$A(t) = \frac{1}{N} \sum_{i=1}^{N} \langle \psi_0 | \sigma_i^z(t) \sigma_i^z(0) | \psi_0 \rangle, \quad (8)$$

where the initial state is chosen to be the domain wall (DW) state, $| \psi_0 \rangle = \downarrow \downarrow \cdots \downarrow \uparrow \uparrow \cdots$. The DW state has an extensive connectivity of $N/2$ on the Fock-space graph, so that arrested dynamics starting from this initial state, if present, cannot be due to a subextensive connectivity of the initial state. At the same time, it contains an extensively large blockaded segment of down spins thus proving to be a convenient choice for clearly demonstrating the effect of the constraints. We stress that our choice of the initial state is not special; the phase diagram in Fig. 3(a) shows that there exists a phase where all the eigenstates are localised.

We employ the kernel polynomial method $[56]$ using Chebyshev polynomials which allows us to evolve systems with $N = 20$ up to very long times, $t \sim 10^4$. The results for $A(t)$ are shown in Fig. 4. For $\Gamma < \Gamma_c$, $A(t)$ saturates to a finite values at long times. The saturated value does not depend on system size, suggesting that the system retains memory of its initial condition in the thermodynamic limit at infinite times. This clearly signifies a strong breaking of ergodicity. In contrast at larger values of $\Gamma$, $A(t)$ slowly decays with both $t$ and $N$. The autocorrelation saturates to a finite value for finite $N$, but this
saturation value decays with $N$ such that in the thermodynamic limit the autocorrelation decays to zero at long times. This is the hallmark of an ergodic system. While it is difficult to precisely determine the critical value of $\Gamma$ separating the two dynamical phases, which we estimate to be in the vicinity of $\Gamma_c \approx 0.17$ (consistently with the exact diagonalisation results of Sec. IV A), the existence of one is clear.

We now turn to the study of the spatially resolved dynamical autocorrelation. We define the autocorrelation functions separately for the left and right halves of the chain (which for the DW initial state correspond to the blockaded and non-blockaded regions),

$$A_{L(R)}(t) = \frac{2}{N} \sum_{i=1}^{N/2(N)} \langle \psi_0 | \sigma^z_i(t) \sigma^z_i(0) | \psi_0 \rangle.$$  \hspace{1cm} (9)

![FIG. 3. Localisation phase diagram of the EastREM in the $\Gamma$-$\epsilon$ plane. (a) The total density of states $\rho(E)$ is a Gaussian with a variance $\sim N$ ($N$ being the system size) such that in terms of energy densities $\epsilon = E/N$, the fraction of eigenstates at all finite $\epsilon$ is vanishingly small in the thermodynamic limit. Note that $\epsilon = 0$ corresponds to the middle of the spectrum. (b) The ergodic region (blue) is characterised by the first participation ratio’s volume law coefficient in Eq. (7), $a_1 \approx 1$ as shown by the colour-map, whereas $a_1 < 1$ in the MBL phase (light region). The black dots show the critical $\Gamma$ extracted from the level spacing ratios for the EastREM whereas the red squares denote the critical $\Gamma$ line for the QREM [43]. The black dashed line denotes the result obtained from a numerical treatment of the FSA (Sec. VI C). (c)-(d) Representative plots of the mean level spacing ratio, $\langle s \rangle$, versus $\Gamma$ for different system sizes $N$ for the bulk and edges of the spectrum respectively. All the data was averaged over 1000 disorder realisations and the statistical errorbars estimated using 500 bootstrap resamplings.](image)

![FIG. 4. Autocorrelation in the EastREM as a function of time, $A(t)$, starting from the domain wall state for various $\Gamma$ values and system sizes $N$. For small values of $\Gamma$, the $A(t)$ saturates to a finite value with no perceptible finite-size effects whereas for larger values of $\Gamma$, it decays with both $t$ and $N$ like an ergodic system. The critical $\Gamma$ can be estimated to lie in the vicinity of $\Gamma_c \approx 0.17$. The data was obtained by averaging over 1000 disorder realisations for all systems sizes and statistical errors estimated using standard bootstrap resampling.](image)

![FIG. 5. Panels (a) and (d) show the spin-expectation value on the chain (of length $N = 20$) at different times (denoted by the colourbar) for $\Gamma = 0.1$ in the localised phase and 0.3 in the delocalised phase. While in the former, both the halves of the chain fail to thermalise, in the latter, both of them do thermalise although the initially blockaded left half is slower. The other panels show this via the behaviour of $A_L(t)$ and $A_R(t)$, defined in Eq. (9), with $t$ and $N$: In the localised phase (top panels) neither half thermalises, while in the delocalised phase (bottom panels) both $L$ and $R$ eventually thermalise, with the initially blockaded $L$ region thermalising slower.](image)
As anticipated in Sec. III, the breaking of ergodicity manifested in localised behaviour can be attributed to two effects.

(i) Because of the East-like constraint of Eq. (6), any block of contiguous down spins is slow to “melt” since the only spin in that block that can change dynamically is the one on the rightmost edge. The entire block can therefore melt only sequentially starting from the right. In other words, for spin- configurations with such “solid” blocks of frozen spins, a large number of channels out of these configurations, which involve flipping of spins deep in the frozen block are simply unavailable. Moreover, this also has the effect of supressing the total number of pathways on the Fock space from one configuration to another. For example, there is a single shortest path on the Fock-space graph that connects the DW state to the all-up state. Contrarily for the QREM, the corresponding number of shortest paths is \( \left( \frac{N}{2} \right)! \sim e^N \).

(ii) In the localised phase, the apparent liquid regions made up of segments of up spins also don’t thermalise, see Fig. 5(a) and (c). The mechanism underlying this is the creation of new blockades dynamically. Once a single spin (say at site \( i \)) is flipped from up to down, the one at \( i-1 \) is frozen until the \( i^{th} \) spin is flipped back up. However, this flipping is unavoidable; thermalising the region requires, by definition, that the quantum state explore all other spin configurations in the Fock space, and these naturally posses segments of down spins creating new constrained regions which eventually may lead to localisation.

V. MINIMALLY STRUCTURED CONSTRAINED MODEL

To demonstrate the two effects mentioned at the end of Sec. IV B, we now construct a new model in which the second effect is removed by hand while the first left in.\(^7\) To do so, we recognise that the first effect above, namely the slowness of the melting of blockaded regions, is due to the relatively small number of matrix elements leading out of clusters of states all of which include the same blockaded island, while the second relates to dynamics inside each cluster.

The model we construct consists of GOE matrices describing each of these clusters, with each of these matrix blocks connected to the others by matrix elements which are identical to the same matrix elements as in the EastREM. It is hence a hybrid of a random matrix with the EastREM, and arguably the least structured model that still displays one of the features of the EastREM, namely, the difficulty of melting the blockaded islands. Unlike the EastREM, liquid regions will remain liquid under the dynamics of the new model, being fully chaotic as their dynamics is described by the random matrix.

To construct our new model, first we group together spin configurations so that all states with a given length of blockaded down spins starting from the leftmost spin are in the same group:

\[
\text{Cluster# 0 : } \begin{array}{c}
\downarrow \cdots \downarrow \\
N/2-1 \\
\uparrow \\
N/2
\end{array}, \\
\text{Cluster# 1 : } \begin{array}{c}
\downarrow \cdots \downarrow \\
N/2-1 \\
\uparrow \\
N/2
\end{array}, \\
\vdots
\text{Cluster# } i : \begin{array}{c}
\downarrow \cdots \downarrow \\
N/2-i \\
\uparrow \\
N/2+i-1
\end{array}, \\
\text{(10)}
\]

where the \( \circ \) denotes sites which could either up or down spins. We note two features of this separation of Fock space into clusters:

(i) Firstly, hoppings in the EastREM between different clusters correspond to progressively melting the solid block. This is because EastREM only allows either the rightmost spin of a blockaded island or the first spin after the island to flip, and either of

\(^7\) We focus on constructing a model appropriate to an initial domain wall state, but this is not a special choice and any spin-configuration could have been used.
these flips results in a state in cluster \(i \pm 1\) so that transitions are only allowed between clusters \(i\) and \(i \pm 1\) by the EastREM rules.

(ii) Secondly, flipping spins in the liquid regions corresponds to Fock-space hoppings within a cluster. These lead to formation of new constraints as discussed in Sec. IV B and stop the apparently liquid regions from thermalising in the MBL phase.

In the bottom two panels of Fig. 6 we show a representation of the Hamiltonian matrix of the EastREM (left) and GOEastREM (right) in the basis of the Fock states, arranged so that states in the same block are next to each other. The black lines correspond to the boundaries between blocks, so that the square blocks along the diagonal of the matrices correspond to transitions inside each cluster while the off-diagonal blocks to transitions between the clusters.

To allow spins to flip freely in the liquid regions without the formation of new blockades, we randomise all matrix elements between states in the same cluster while keeping the matrix elements between clusters as in the EastREM model; in other words, we make the blocks on the diagonal in Fig. 6 GOE matrices while keeping everything outside them identical to the EastREM. This has the effect of allowing all intra-cluster transitions (that is, dynamics in the liquid region) with no constraints while keeping the inter-cluster transitions (corresponding to island melting) as in the EastREM model.

Fig. 6 also makes it evident that decreasing the size of an island by more than 1 still cannot be done by a single application of the Hamiltonian (there are still no matrix elements connecting clusters that are not nearest neighbours). Melting an island is thus slow, involving a time \(\mathcal{O}(\Gamma w)\) for an island of length \(w\), like in the EastREM. On the other hand, the GOE structure of the intra-cluster Hamiltonians means that the effect of constraints within the cluster is no longer there as such new constraints cannot be created in the liquid regions.

Hence out of the two effects identified earlier, namely, slow dynamics/localisation in the already frozen region and formation of new frozen regions, the latter has been eliminated in the GOEastREM. This is confirmed in the dynamical autocorrelations in the GOEastREM starting from the domain-wall state as shown in Fig. 7. The results are for \(\Gamma = 0.1\) which corresponds to the MBL phase for the EastREM. The left half of the system which corresponds to the solid region fails to thermalise as in the EastREM as indicated by the saturation of \(A_L(t)\) with both \(t\) and \(N\). On the other hand, the right half rapidly thermalises, resulting in the systematic decay of the saturation values of \(A_R(t)\) with \(N\), in stark contrast to the EastREM. This demonstrates that, as anticipated, the non-ergodic behaviour shown by the segment of up spins in the EastREM was indeed caused by the formation of new blockades, as the GOEastREM removes that mechanism. At the same time, as the GOEastREM preserves the constraints which lead to non-thermalisation of segments of down spins, similar to the EastREM, indicating that the same mechanism is at play in both the models.

VI. FORWARD SCATTERING APPROXIMATION

To provide analytical insight we now turn to the forward scattering approximation (FSA), which is an approximation to the non-local (in Fock space) Green’s function to lowest order in \(\Gamma\) and amounts to a stability analysis of the trivially localised phase at vanishing hopping \(\Gamma = 0\) [Eqs. (3)-(5)].

Considering an arbitrary initial state which we label by \(\alpha = 0\) and which is an eigenstate of the unperturbed \(\Gamma = 0\) Hamiltonian (that is, a \(\sigma^z\)-product state), the weight of the perturbed eigenstate on an arbitrary spin-configuration \(\{\sigma_i^z\}_\alpha\), denoted as \(\psi(\{\sigma_i^z\}_\alpha)\), is

\[
\psi(\{\sigma_i^z\}_\alpha) = \sum_{p \in \text{paths}^\ast(0, \alpha)} \prod_{\beta \in p} \frac{\Gamma}{\mathcal{E}_\alpha - \mathcal{E}_\beta},
\]

where \(\text{paths}^\ast(0, \alpha)\) is the set of all shortest paths from the unperturbed \(\alpha = 0\) state to \(|\alpha\rangle\). The \(\mathcal{E}_\alpha\), as before, are the random Fock-space site energies defined in Eq. (2) and are normally distributed, \(\mathcal{E}_\alpha \sim \mathcal{N}(0, \mathcal{N})\). In this setting, the breakdown of localisation is signalled by the probability of resonance at arbitrarily large distances \(r\) on the Fock space from the site \(\alpha = 0\) approaching unity such that under the state spreads to Fock-space sites such distances at finite \(\Gamma\). The delocalisation criterion can be formally expressed as

\[
\lim_{r \to \infty} \mathbb{P} \left( \frac{\ln |\psi_r|^2}{2r} > -\xi^{-1} \right) \to 1,
\]

---

8 For a given disorder realisation, the diagonal elements, \(\{\mathcal{E}_\alpha\}\), are also the same as that of the EastREM.
that the distributions decay faster than a power-law and hence state. On a log-log scale, the curves bend downwards showing paths, space site distant by where \( N \) scaled with that of the QREM \((N!\) decays exponentially with \( N \), showing a strong suppression of the number of paths due to the constraints in the EastREM.

where \( \psi_r \) denotes the wavefunction amplitude on a Fock-space site distant by \( r \) from the initial state and \( \xi \), an analogue of the localisation length on the Fock space. Note that, the delocalisation criterion of Eq. (12) gives a conservative estimate in that it provides a lower bound on the critical \( \Gamma \) as it is enough for the maximum of \( \psi_r \) over all states in the middle of the spectrum \((\xi)\) to be resonant, \( p_r^{\text{res}} \), can be computed as the net support of the distribution \( P_F \) on \( F_r \geq -r \ln \Gamma_r \),

\[
P_r^{\text{res}} = \int_{-r \ln \Gamma}^{\infty} dF_r \; P_F(F_r) \approx \frac{1}{(r-1)!} \left[ r \ln \frac{\sqrt{N}}{\Gamma_r} \right]^{r-1} \exp \left[ -r \ln \frac{\sqrt{N}}{\Gamma_r} \right]. \tag{16}
\]

As each of the \( n_r^{(s)} \) sites at distance \( r \) are independent, the probability that there is no resonance at distance \( r \) is given by \( Q_r = (1 - p_r^{\text{res}})^{n_r^{(s)}} \approx e^{-n_r^{(s)}p_r^{\text{res}}} \). The ratio \( \lambda_r = n_r^{(s)}p_r^{\text{res}}/n_r^{(s)}p_r^{\text{res}} \) is a monotonically decreasing function of \( r \). Hence, if for some \( r \) we have \( \lambda_r < 1 \) then \( n_r^{(s)}p_r^{\text{res}} \to 0 \) as \( r \to \infty \) and consequently \( Q_r \to 1 \); this signals the stability of localisation as \( Q_r \) is the probability of no resonances at distance \( r \).

Using Eq. (16), the localisation criterion \( \lambda_r < 1 \) can be rewritten as

\[
\Gamma \leq \frac{n_r^{(p)}}{n_{r+1}^{(p)}} \frac{n_r^{(s)}}{n_{r+1}^{(s)}} \frac{\sqrt{N}}{\ln N} \left( 1 + \frac{1}{r} \right)^{-r}, \tag{17}
\]

for a finite \( r \) in the limit of \( N \to \infty \). Hence, for localisation to persist until a finite value of \( \Gamma \), we require the dependence of \( n_r^{(s)} \) and \( n_r^{(p)} \) on \( r \) and \( N \) to be such that \( K(r,N) \) does not scale with \( N \).

For the QREM, \( n_r^{(s)} = \binom{N}{r} \) and \( n_r^{(p)} = r! \), and the ratio \( K(r,N) \sim N^{-1/2}/r \). Thus the RHS of Eq. (17) scales as \( N^{-1/2} \) and vanishes in the thermodynamic limit rendering localisation impossible. On the other hand, for
the EastREM, the ratio \( K(r, N) \) computed numerically does show an absence of dependence on \( N \) for finite \( r \) (see Fig. 9); the range of \( r \) over which this holds grows with \( N \), suggesting that a localisation-delocalisation transition is indeed possible at a finite \( \Gamma \) for \( \epsilon = 0 \) in the thermodynamic limit.

Note that the qualitative difference between the QREM and EastREM with regard to the ratio \( n_r^{(s)}/n_{r+1}^{(s)} \) arises purely from the constraints. In the QREM, after one flips \( r \ll N \) spins, one is free to flip any of the \( N - r \approx N \) spins in the next step, which leads to the ratio \( n_r^{(s)}/n_{r+1}^{(s)} \) scaling as \( N \). On the other hand, in the EastREM, flipping some spins from up to down creates new blockades and so the number states available on successive steps don’t scale as fast as in the QREM. This argument in conjunction with the FSA presents an analytical picture of how the constraints affect the distribution of distances on the Fock space which in turn lead to a constraint-induced localised phase in the EastREM, unlike the QREM.

**B. Enhancement of localisation at spectral edges**

Let us now consider the situation at a finite energy density \( \epsilon_0 = \mathcal{E}_0/N \) which, as the spectral width \( \propto \sqrt{N} \), corresponds to the edges of the spectrum. Even though the density of states is exponentially small in this energy region, it is nevertheless important for the dynamics.

The distribution of \( \Gamma/(\mathcal{E}_0 - \mathcal{E}_\alpha) \) is fat-tailed so that single sites can become resonant. As most of the \( \mathcal{E}_\alpha \)s are \( \sim \sqrt{N} \) these resonances are rare, so we focus on paths with a single resonance; for a Fock-space site \( |\alpha\rangle \) at distance \( r \) to be resonant it is sufficient for a single site to be resonant. For a resonance to occur at distance \( r \) (and not before), we require that \( \mathcal{E}_\beta \sim \sqrt{N} \) for all but the last \( \beta \) on the shortest path but \( |\mathcal{E}_0 - \mathcal{E}_\alpha| \ll 1 \). In this scenario, the amplitude on the Fock-space site \( \alpha \) at distance \( r \) an

\[
\psi_r = n_r^{(p)} \left( \frac{\gamma}{\epsilon_0} \right)^{r-1} \frac{\Gamma}{\mathcal{E}_0 - \mathcal{E}_\alpha}, \tag{18}
\]

where as before \( n_r^{(p)} \) is the average number of paths to sites at distance \( r \) and we have implicitly assumed that all paths are independent. As the distribution of the number of shortest paths, \( n_r^{(p)} \) is not fat-tailed (see Fig. 8(a)), using the average is justified.

From Eq. (18), a resonance at the last site requires that \( |\psi_r| > 1 \) or equivalently \( |\mathcal{E}_0 - \mathcal{E}_\alpha| < n_r^{(p)} \Gamma \left( \frac{\gamma}{\epsilon_0} \right)^{r-1} \). Thus the probability of the state being resonant is

\[
p_{r}^{\text{res}} = \int_{\mathcal{E}_0 - n_r^{(p)} \Gamma \left( \frac{\gamma}{\epsilon_0} \right)^{r-1}}^{\mathcal{E}_0 + n_r^{(p)} \Gamma \left( \frac{\gamma}{\epsilon_0} \right)^{r-1}} d\mathcal{E}_\alpha \frac{1}{\sqrt{2\pi N}} \exp \left[ -\frac{\mathcal{E}_\alpha^2}{2N} \right] \tag{19}
\]

\[
\approx \sqrt{\frac{2}{\pi N}} \exp \left[ -\frac{\mathcal{E}_\alpha^2}{2N} \right] n_r^{(p)} \Gamma \left( \frac{\gamma}{\epsilon_0} \right)^{r-1}. \tag{20}
\]

Since the \( \mathcal{E}_\alpha \)s are i.i.d. random variables, the expression above holds for any state at a distance \( r \). The probability that none of the \( n_r^{(s)} \) sites at distance \( r \) is resonant then is simply given as before by \( Q_r = (1 - p_{r}^{\text{res}}n_r^{(s)}) \approx e^{-n_r^{(s)}p_{r}^{\text{res}}} \). Localisation persists if \( Q_r \to 1 \) as \( N \to \infty \) whenever \( r \) is a finite-fraction of \( N \); we thus define \( x = r/N \) which will be useful later.

Using the expression in Eq. (20), \( Q_r \) can be written as

\[
Q_r = \exp \left[ -ke^{Nf(r, N, \epsilon_0)} \right], \tag{21}
\]
where $f$ is

$$f(r, \epsilon_0) = -\frac{\epsilon_0^2}{2} + x \ln \frac{\Gamma}{\epsilon_0} + \frac{1}{N} \ln (n_r^{(s)} n_r^{(p)}) - x \ln N,$$  \label{eq:22}

where we have only kept terms that survive in the thermodynamic limit.

Crucially, $Y(r, N)$ in Eq. (22) is only a function of $x$, $Y(r, N) = Y(x)$, which in turn means that $f(r, N, \epsilon_0) = f(x, \epsilon_0)$. This can be trivially shown for the QREM using $n_r^{(p)} = r! \approx (r/e)^r$ and $n_r^{(s)} = \binom{N}{r}$ that $Y(x) = -(1-x) \ln(1-x) - x$. For the EastREM such analytic expressions $n_r^{(p)}$ and $n_r^{(s)}$ are not available, but the numerically obtained form in Fig. 10(a) shows that $Y(r, N)$ indeed is simply a function of $x = r/N$. The localisation condition $Q_r \rightarrow 1$ as $N \rightarrow \infty$ then requires that

$$\max_x f(x, \epsilon_0) < 0.$$  \label{eq:23}

The critical $\Gamma$ can be obtained by solving the equation

$$-\frac{\epsilon_0^2}{2} + x_s \ln \frac{\Gamma_c}{\epsilon_0} + Y(x_s) = 0$$  \label{eq:24}

where $f(x, \epsilon_0)$ is maximised at $x_s$.

We solve Eq. (24), for both the QREM and the EastREM, showing the results in Fig. 10. We find that in the EastREM, localisation persists to a larger value of $\Gamma$.

C. Numerical treatment of FSA

We now locate the transition numerically exactly within the FSA for small system sizes. To do this, we rewrite the delocalisation criterion of Eq. (12) as

$$\lim_{r \rightarrow \infty} P(\Lambda_r > \ln\left(\frac{1}{\Gamma_c}\right)) \rightarrow 0,$$  \label{eq:25}

where $\Lambda_r = \ln |\psi_r|^2/2r - \ln \Gamma$ and for $\Gamma < \Gamma_c$. Our strategy is to directly calculate the amplitudes $\psi_r$ within the FSA (by obtaining the shortest paths numerically) and from those obtain the distribution of Eq. (25). Eq. (25) then says that the upper limit of its support then determines the critical $\Gamma_c$. Without loss of generality, for this calculation we shall take the DW state as the initial state as before.

We note here that while for the QREM to each state there corresponds only one state at hamming distance $N$, for the EastREM there are $\sim e^N$ such states. Indeed, in the EastREM, the number of configurations at distance $r$ is peaked at $r = N$. Hence, one can argue that studying the statistics of $\Lambda_N$ can overestimate $(1/\Gamma_c)$, thus underestimating $\Gamma_c$ as the likelihood of having a resonance at $r = N$ is quite high simply due to a large fraction of the configurations having $r = N$. As our main result is that $\Gamma_c > 0$ and in general larger than for the QREM, underestimating it is not a problem.

To calculate $P(\Lambda_N)$ starting from the domain-wall state we construct a matrix

$$\mathcal{T} = \Gamma \sum_{\beta, \gamma} A_{\beta, \gamma} \langle \beta | \gamma \rangle$$

with $A_{\beta, \gamma} = 1$ if $r_{(0, \gamma)} < r_{(0, \beta)}$ and $\langle \beta | H_{\text{East}} | \gamma \rangle \neq 0$, where $r_{(0, \beta)}$ is the hamming distance between $|\beta\rangle$ and the domain-wall state. That is, $A_{\beta, \gamma} = 1$ if the transition between the two states $|\beta\rangle$ and $|\gamma\rangle$ is allowed by the Hamiltonian and it increases the distance from the domain-wall state. The amplitude on the configuration $|\alpha\rangle$, at a hamming distance $r$ is then given by

$$\psi(|\sigma_i^r\rangle_\alpha) = \langle \alpha | T^r | 0 \rangle,$$

from which $|\psi_r| = \max_{r_{(0, \alpha)}=r} \langle \psi(|\sigma_i^r\rangle_\alpha) \rangle$.

The distribution of $\Lambda_N$ so obtained is shown in Fig. 11(a) for various $N$. It clearly has finite support, which has an upper bound that becomes sharper with increasing $N$. This is consistent with the conclusion that as $N \rightarrow \infty$, there exists a sharp value $1/\Gamma_c$ above which the distribution has no weight, as required by Eq. (25).

Since the distribution tends to get sharper with $N$, one can argue that the critical $1/\Gamma_c$ can be estimated as

$$\lim_{N \rightarrow \infty} N \langle \Lambda_N \rangle = \ln(1/\Gamma_c).$$  \label{eq:26}

In order to estimate this limiting value, we fit $N \langle \Lambda_N \rangle$ as a function of $N$ to a form

$$N \langle \Lambda_N \rangle = a + N \ln(1/\Gamma_c) + bN\gamma,$$  \label{eq:27}

where the last term takes into account the slowly decreasing fluctuations in $\Lambda_N$ with increasing $N$. The fit is shown in Fig. 11(b) with the best fit parameters yielding $\ln(1/\Gamma_c) = 1.77 \pm 0.02$ which implies $\Gamma_c = 0.17 \pm 0.01$.

Note that the transition criterion in Eq. (25) can also be equivalently stated as

$$\lim_{N \rightarrow \infty} C(\Lambda_N) = \ln(1/\Gamma_c),$$  \label{eq:28}

where $C(\Lambda_N)$ is the cumulative distribution corresponding to $P(\Lambda_N)$. For finite-sized systems, we plot the $1 - C(\Lambda_N)$ in Fig. 11(c) and observe a clear crossing of the data for various system sizes. The value at which the crossing occurs and which we identify as the critical point matches remarkably well with that obtained from the finite-size scaling analysis of $\langle \Lambda_N \rangle$, as shown by the grey shaded region in Fig. 11(c). More importantly, the critical value so obtained, $\Gamma_{c,\text{FSA}}$, is in excellent agreement with the infinite temperature $\Gamma_c$ obtained from exact diagonalisation, see Fig. 3.
FIG. 11. Numerical treatment of FSA for the EastREM. (a) The distributions \(P(\Lambda_N)\) for different \(N\). (b) The circles show the data for \(N(\Lambda_N)\) and the red dashed line shows a fit to the form in Eq. (27). The best fit parameters are \(\ln(1/\Gamma) = 1.77 \pm 0.02\) and \(\gamma = 0.09\), the low value of the latter highlighting the slow decay of the fluctuations in \(\Lambda_N\) with \(N\). (c) The survival function, \(1 - C(\Lambda_N)\) corresponding to \(P(\Lambda_N)\) shows a clear crossing of the data for various \(N\) suggesting a critical point well in agreement with that extracted from (b) as shown by the grey shaded region.

VII. DISCUSSION

In conclusion, we have shown that local constraints can induce strong ergodicity breaking manifested as localisation in quantum many-body systems, crucially without shattering the Fock space. The locality of the constraints allows us to identify the long-lived local spatial configurations responsible for the dynamical arrest, which in Fock space turn out to correspond to dynamical bottlenecks caused by sparse connectivity between clusters of states. These results are exemplified by a quantum random energy model with East-like constraints, which we introduce and call the EastREM. We provide further support for the picture of isolated regions of Fock space by constructing and studying a random matrix model with GOE blocks for each cluster but the same intercluster connections as the EastREM. This model, which we name the GOEastREM, is the minimally structured model possessing the localisation mechanism we have identified. We finally obtain analytical insight by applying the FSA in Fock space. The constraints modify the distance (in Fock space) dependence of the number of accessible sites and paths to them, and the FSA shows how this leads to localisation. In addition to providing this insight, the FSA is in excellent agreement with the ED results (see Fig. 3).

At this juncture, a number of potential directions for future work present themselves. An immediate direction of interest is a systematic study of the statistical mechanics of the paths on the Fock space by treating them as directed polymers on a correlated but random landscape. The replica trick [57] is ideally suited to obtain further analytical insight into the problem as the non-local propagator on the Fock space is expected to be dominated only by a few paths which pass through the resonant bottlenecks. In fact, in the context of many-body localisation in traditionally studied short-ranged disordered spin chains, a classical percolation proxy on the Fock space was recently introduced [27, 28]. The effect of constraints on such a percolation picture and the potential connections to the directed polymer picture could shed light on the nature of the transition.

A different question is whether an approach based on random unitary circuits, recently used to study universal properties of ergodic systems [58–60], can be generalised to include local constraints such that ergodicity is broken. In fact, there have been works on including conservation laws [61–63] as well as ergodicity breaking [64] in unitary circuits. The question then is to modifying the structure of the unitary gates in the circuit such that the scrambling is constrained locally, analogously to having a conserved degree of freedom locally. The physics in this scenario remains fundamentally different from unitary circuits with conservation laws which shatter the Hilbert space [34, 35].

Looking further afield, periodically driven (Floquet) systems have emerged as one of the more active areas of research in quantum dynamics. The main difficulty in seeing interesting physics with them is that ergodic systems inevitably heat up under driving [65–67]. Two routes to arresting this heating have been integrability [68] and Floquet-MBL [69, 70], both of which rely on breaking ergodicity to prevent heating. It is then natural to ask whether the present method of breaking ergodicity with local constraints can also prevent the heating up of driven quantum systems, without explicitly fragmenting the Fock space (thus rendering the physics distinct from that of scars in Floquet systems [71, 72]).

Finally one might ask whether many body localisation originating from correlations in Fock space [30], can be identified as being caused by emergent constraints due to the correlated Fock-space disorder.

ACKNOWLEDGMENTS

We would like to thank D. E. Logan for illuminating discussions about the FSA. This work was in part supported by EPSRC Grants No. EP/N01930X/1 and EP/S020527/1.

[1] M. Srednicki, The approach to thermal equilibrium in quantized chaotic systems, J. Phys. A: Math. Gen. 32, 1163 (1999).
[44] F. Ritort and P. Sollich, Glassy dynamics of kinetically constrained models, Advances in Physics 52, 219 (2003).
[45] J. P. Garrahan, P. Sollich, and C. Toninelli, Kinetically constrained models, in Dynamical heterogeneities in glasses, colloids, and granular media, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos (Oxford University Press, Oxford, 2011).
[46] J. P. Garrahan, Aspects of non-equilibrium in classical and quantum systems: Slow relaxation and glasses, dynamical large deviations, quantum non-ergodicity, and open quantum dynamics, Physica A: Statistical Mechanics and its Applications 504, 130 (2018).

[47] R. Abou-Chacra, D. J. Thouless, and P. W. Anderson, A self-consistent theory of localization, Journal of Physics C: Solid State Physics 6, 1734 (1973).
[48] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. 109, 415 (1958).
[49] D. J. Thouless, A. Matveev, and P. W. Anderson, A model of the localization transition, Phys. Rev. B 24, 2252 (1981).
[50] Y. Bar Lev, G. Cohen, and D. R. Reichman, Absence of diffusion in an interacting system of spinless fermions on a one-dimensional disordered lattice, Phys. Rev. Lett. 114, 100601 (2015).
[51] J. Z. Imbrie, On many-body localization for quantum spin chains, J. Stat. Phys. 163, 998 (2016).
[52] C. W. von Keyserlingk, T. Rakovszky, and F. Pollmann, Many-body localization in non-equilibrium systems, Phys. Rev. X 4, 021052 (2014).
[53] A. De Luca and A. Scardicchio, Spectral statistics in spatially extended chaotic quantum many-body systems, Phys. Rev. Lett. 123, 210603 (2019).
[54] A. Chan, A. De Luca, and J. T. Chalker, Spectral statistics in spatially extended chaotic quantum many-body systems, Phys. Rev. Lett. 121, 060601 (2018).
[55] A. Lazarides, A. Das, and R. Moessner, Equilibrium states of generic quantum systems subject to periodic driving, Phys. Rev. E 90, 012110 (2014).
[56] L. D’Alessio and M. Rigol, Long-time behavior of isolated periodically driven interacting lattice systems, Phys. Rev. X 4, 041048 (2014).
[57] P. Ponte, A. Chandran, Z. Papić, and D. A. Abanin, Periodically driven ergodic and many-body localized quantum systems, Annals of Physics 353, 196 (2015).
[58] A. Lazarides, A. Das, and R. Moessner, Periodic thermodynamics of isolated quantum systems, Phys. Rev. Lett. 112, 150401 (2014).
[59] A. Lazarides, A. Das, and R. Moessner, Fate of many-body localization under periodic driving, Phys. Rev. Lett. 115, 030402 (2015).
[60] P. Ponte, Z. Papić, F. Huveneers, and D. A. Abanin, Many-body localization in periodically driven systems, Phys. Rev. Lett. 114, 140401 (2015).
[61] B. Mukherjee, S. Nandy, A. Sen, and K. Sengupta, Collapse and revival of quantum many-body scars via floquet engineering, arXiv preprint arXiv:1907.08212 (2019).
[62] A. Haldar, D. Sen, R. Moessner, and A. Das, Scars in strongly driven floquet matter: resonance vs emergent conservation laws, arXiv preprint arXiv:1909.04064 (2019).