


## Many-Body Delocalization as a Quantum Avalanche

Thimothée Thiery,<sup>1,\*</sup> François Huveneers,<sup>2,†</sup> Markus Müller,<sup>3,‡</sup> and Wojciech De Roeck<sup>1,§</sup>

<sup>1</sup>*Instituut voor Theoretische Fysica, KU Leuven, 3001 Leuven, Belgium*

<sup>2</sup>*Université Paris-Dauphine, PSL Research University, CNRS, CEREMADE, 75016 Paris, France*

<sup>3</sup>*Paul Scherrer Institute, PSI, 5232 Villigen, Switzerland*

 (Received 13 October 2017; revised manuscript received 31 August 2018; published 2 October 2018)

We propose a multiscale diagonalization scheme to study disordered one-dimensional chains, in particular, the transition between many-body localization (MBL) and the ergodic phase, expected to be governed by resonant spots. Our scheme focuses on the dichotomy of MBL versus validity of the eigenstate thermalization hypothesis. We show that a few natural assumptions imply that the system is localized with probability one at criticality. On the ergodic side, delocalization is induced by a quantum avalanche seeded by large ergodic spots, whose size diverges at the transition. On the MBL side, the typical localization length tends to the inverse of the maximal entropy density at the transition, but there is a divergent length scale related to the response to an inclusion of large ergodic spots. A mean-field approximation analytically illustrates these results and predicts a power-law distribution for thermal inclusions at criticality.

DOI: [10.1103/PhysRevLett.121.140601](https://doi.org/10.1103/PhysRevLett.121.140601)

**Introduction.**—The phenomenology and theory of many-body localization (MBL), i.e., the absence of thermalization in interacting quantum systems [1–16], challenges our understanding of statistical mechanics. In  $d = 1$ , the main outstanding issue is the nature of the transition [17–29] that separates the MBL from the ergodic (thermalizing) phase. To describe it, several phenomenological renormalization schemes have been introduced [19,20,26,27], with partially conflicting predictions.

In the present Letter, we develop a theory which is rooted in two microscopic principles. The first principle, governing nonresonant couplings, is the spectral perturbation theory. The second principle is the use of the random matrix theory for resonant couplings [30–36], which strikingly predicts an “avalanche” instability: An infinite localized system can be thermalized by a finite ergodic seed if the typical localization length  $\zeta$  exceeds a critical  $\zeta_c$  [36]. We implement these principles in the form of a multistep diagonalization procedure [37–42], described compactly below and in more detail in the companion paper [43].

Analyzing first the general consequences of this scheme, we find that the critical point must be localized with probability one. This conclusion, which rests on a few basic facts and does not involve any detailed analysis, contrasts with predominantly numerical RG studies [19,27], which reported the half-chain entanglement entropy at the critical point to follow a (subthermal) volume law. Our result instead implies that the bipartite entanglement entropy of typical cuts is discontinuous at the transition, as in Ref. [21], and that the typical localization  $\zeta$  does not diverge. The latter is a direct consequence of the explicit upper bound  $\zeta \leq \zeta_c$  [36]. On the other hand, we do identify a length scale  $\ell^*$  that does diverge as  $(\zeta - \zeta_c)^{-1}$  as

one approaches the transition from the MBL side; see Fig. 1. This is caused by the divergent susceptibility of the sample to the insertion of large ergodic spots. In our scheme, such spots trigger delocalization by an avalanche instability, a central aspect that distinguishes our work from previous approaches [19,20,26,27]. The validity of the avalanche scenario and the associated bound on  $\zeta$  were recently verified through high-precision numerics [44], and we show here that it leads to a consistent and physical picture of the MBL transition. On the thermal side, instead, we find no divergent correlation length, but only a diverging crossover length  $L_+$ , beyond which typical samples appear thermal.  $L_+$  is associated with a typical timescale of local thermalization,  $t_+$ , that diverges quasi-exponentially at the transition.

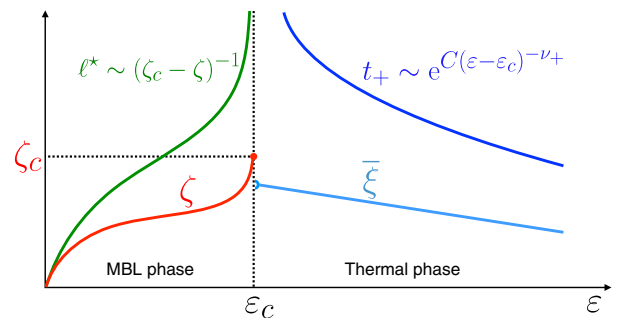


FIG. 1. Phase diagram where  $1/\epsilon$  quantifies the disorder strength.  $\zeta$  is the typical localization length. It is bounded by the critical localization length  $\zeta_c$ , which equals the inverse of the entropy density.  $\ell^*$  quantifies the susceptibility to the insertion of large ergodic spots.  $\bar{\zeta}$  is the average eigenstate correlation length and  $t_+$  a typical local thermalization timescale in the thermal phase.

We illustrate these aspects with a mean-field approximation of our scheme. While it introduces some oversimplifications, as discussed below, it offers a concrete implementation of the main ideas developed in this Letter and yields several conclusions that have been confirmed by a full numerical analysis of the scheme [43]. This includes the fact that the upper bound  $\zeta_c$  is indeed saturated at the transition, as well as a power-law distribution of the sizes of thermal inclusions at criticality.

*Multistep diagonalization scheme.*—We consider a chain of spins  $S_i$ , with  $d_s$  states per site  $i$ , with a generic local Hamiltonian, which we write in the form

$$H = \sum_I D_I + \sum_I V_I. \quad (1)$$

Here,  $I$  denotes a stretch of consecutive sites that the operators  $D_I$  and  $V_I$  act on. The operators  $D_I$  act diagonally in a preferred basis that, for concreteness, we take to be the  $S_i^z$  basis. The  $V_I$  are not diagonal, and we refer to them as “couplings”. We express lengths in units of the lattice spacing  $a$ . A special role is played by the entropy density  $s = \log(d_s)/a$ . We now diagonalize the system by iteratively eliminating couplings; see Fig. 2.

*Perturbative couplings:* The distinction between perturbative and resonant couplings is at the heart of our procedure. We declare a coupling  $V_I$  “perturbative” if typical eigenstates of  $D_I + V_I$  are small perturbations of the eigenstates of  $D_I$  and can hence be obtained by the perturbation theory. Following Ref. [38], we prefer to think in terms of a unitary transformation  $U_I$  that eliminates the coupling  $V_I$  to lowest order by acting on

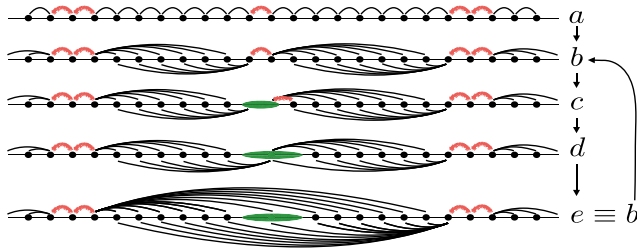


FIG. 2. Evolution of the spin chain during the diagonalization procedure. Black dots are spins, and the green ellipse is a group of nonperturbatively diagonalized spins. Arcs symbolize couplings that act on all spins they embrace. Red ones are resonant, and black ones are perturbative. (a) Initial nearest-neighbor couplings. (b) All perturbative couplings that do not touch resonant couplings have been eliminated, which generates weak couplings to the adjacent resonant spots. (c) Some resonant sites have been fused into a green spot, requiring a reevaluation of the existing couplings. Some arcs to sites close to the green spot have thereby become resonant (red). (d) The new resonant coupling is fused into the green spot. No new red couplings emerged. They are eliminated in step (e). A next resonant spot can be fused now, as in (b), until no resonances are left.

$H$  as  $H \rightarrow U_I H U_I^\dagger$ . This is achieved by choosing  $U_I = e^{A_I}$  with  $\langle \eta' | A_I | \eta \rangle := (\langle \eta' | V_I | \eta \rangle / E_I(\eta') - E_I(\eta))$ . Here,  $\eta, E_I(\eta)$  are eigenstates and eigenvalues of  $D_I$ . This procedure is meaningful if

$$\mathcal{G} \equiv \max_{\eta'} |\langle \eta' | A_I | \eta \rangle| < 1 \text{ for typical } \eta; \quad (2)$$

see Supplemental Material [45]. If  $V_I$  is not perturbative, i.e., if  $\mathcal{G} \geq 1$ , then we call the coupling  $V_I$  resonant and we do *not* eliminate it. By eliminating a perturbative coupling, we generate new, but usually smaller, couplings. Indeed, whenever  $I \cap J \neq \emptyset$ , we will create a new coupling  $V_{I'} = U_I V_J U_I^{-1} - V_J \approx [A_I, V_J]$  (at first order) with  $I' \equiv I \cup J$ .

*Resonant couplings:* We first eliminate all perturbative couplings that do not touch resonant regions; see Fig. 2. We then assume that the remaining resonant couplings induce full ergodicity locally, i.e., within their range. We thus diagonalize them by a unitary that we consider as an effective random matrix. Such a strong dichotomy has been theoretically predicted [36], and these predictions are in remarkable agreement with numerics [44]. The random matrix ansatz remains consistent throughout the scheme if the perturbation theory is used as much as possible to “isolate” resonances from their environment [36,38]. For this reason, we diagonalize resonances in order of increasing size and leave untouched perturbative couplings that link to ergodic spots [43,46]. The diagonalization of a resonant region alters the perturbative couplings attached to it and can turn them resonant (and thereby potentially start an avalanche). If they remain perturbative, they can be eliminated in the next step as the scheme is iterated; see Fig. 2. After a number of iterations which scales logarithmically with the system size, all couplings will be eliminated.

*End of the procedure:* The final result of our procedure is encoded in the diagonalizing unitary  $U$ . It is obtained as a product of unitaries  $U_{I_n}$ , each acting on a single stretch  $I_n$ , as described above. A region  $Y$  is ergodic if  $Y = \cup_{\gamma} I_{\gamma}$  for a collection of intersecting stretches  $\{I_{\gamma}\}$  and such that all  $U_{I_{\gamma}}$  are nonperturbative. The full system is thermal if the whole sample is ergodic; otherwise, it is by definition MBL. From  $U$ , one obtains local integrals of motion [38,47–49] by inverse conjugation:

$$\tau_i = U^\dagger S_i^z U. \quad (3)$$

We decompose  $\tau_i = \sum_I \tau_{i,I}$  in spatial components, where  $\tau_{i,I}$  acts on the stretch  $I$ . The decay of  $\tau_{i,I}$  with increasing  $|I|$  defines the typical localization length:

$$\zeta^{-1} = - \lim_{|I| \rightarrow \infty} \frac{\langle \log \|\tau_{i,I}\| \rangle}{|I|}, \quad (4)$$

where  $\|\cdot\|$  is the operator norm [50], whose crucial role will be explained below, and where  $\langle \cdot \rangle$  denotes the disorder

average, assuming that some randomness enters in the Hamiltonian (1). The locality properties of the unitary  $U$  also reflect in the spatial decay of couplings  $V_I$  (created and eliminated in the course of our scheme), as they arise from a transformation inverse to (3).  $\zeta$  can thus also be defined by replacing  $\tau_{i,l}$  in (4) by  $V_I$ .

*General analysis of the scheme.*—For any disorder realization, at a finite length  $L$ , let  $\rho_{\text{th}}(L)$  be the density of sites in an ergodic region, as determined by the scheme. Let us now assume that our scheme has the two following properties. (i) Denoting the inverse disorder strength by  $\varepsilon$  (see below), we assume that  $\langle \rho(L) \rangle_\varepsilon$  is continuous and nondecreasing as  $\varepsilon$  increases. (ii) For any given disorder realization, all sites declared thermal remain so if more sites are added to the chain at one or both ends. One should probably expect (ii) to hold only up to rare exceptions (e.g., coupling a site strongly to a very disordered region might increase the effective disorder on that site; see also [43]). Yet, it seems reasonable that, despite the neglect of such rare cases, our assumptions capture correctly the phenomenology of the scheme.

We first deduce that  $\langle \rho(L) \rangle_\varepsilon$  reaches a limit  $\rho^*(\varepsilon)$  as  $L \rightarrow \infty$ , for any  $\varepsilon$ . Indeed, (ii) directly implies the superadditivity property  $\rho(L+L') \geq (L/L+L')\rho(L) + (L'/L+L')\rho(L')$ , and then the limit exists by Fekete's superadditivity lemma. Second, we show a concentration property: For any disorder strength  $\varepsilon$  and for all but a vanishing fraction of the samples, the thermal density approaches  $\rho^*(\varepsilon)$ , i.e.,  $P[|\rho(L) - \rho^*(\varepsilon)| > \delta] \rightarrow 0$  for any  $\delta > 0$  as  $L \rightarrow \infty$ . Since this property is valid, in particular, at the critical point  $\varepsilon = \varepsilon_c$ , it implies that it must be either localized or thermal with probability 1. To see the concentration of  $\rho$ , let us fix  $\delta > 0$  and let  $L_0$  be large enough so that  $|\langle \rho(L_0) \rangle_\varepsilon - \rho^*(\varepsilon)| < \delta/2$ . Let us then consider a “product” system made of blocks of size  $L_0$  that are decoupled from each other. (ii) implies that  $P[\rho(L) > a] \geq P_{\text{prod}}[\rho(L) > a]$ . The concentration property that we seek holds definitely true for the product system. We thus conclude that  $P[\rho(L) - \rho^*(\varepsilon) < -\delta]$  goes to 0 as  $L \rightarrow \infty$ , and, since the average value of  $\rho$  converges to  $\rho^*(\varepsilon)$ , we also conclude that  $P[\rho - \rho^*(\varepsilon) > \delta]$  must vanish as  $L \rightarrow \infty$ .

It remains to decide whether the critical point is localized or thermal. Since  $\langle \rho(L) \rangle_\varepsilon$  is nondecreasing in both  $L$  and  $\varepsilon$  and continuous in  $\varepsilon$  by (i) and (ii), we conclude that  $\rho^*(\varepsilon)$  is nondecreasing and left-continuous. Hence, either  $\rho^*(\varepsilon)$  is actually continuous at the transition and the critical point is thermal, or it has a jump at the transition and the critical point is localized. It is clear that the localization length  $\zeta$  should diverge as  $\rho^*$  approaches 1 (cf. the “rule of halted decay: below). Hence, the bound  $\zeta \leq \zeta_c$  [36,44] implies that  $\rho_c < 1$  and, thus, that the critical point is localized.

*Understanding the transition.*—We now develop a general picture for the transition. By making a simple mean-field assumption, we obtain explicit analytical results that

substantiate this picture. Detailed numerics in Ref. [43] yields further support for our theory.

*Resonances and scales:* The simplest resonances are associated to couplings  $V_I$  on single bonds  $I = \{i, i+1\}$ . Let  $\varepsilon$  be the probability that such a  $V_I$  is resonant, and let us use this as a measure for the inverse disorder strength. We call a *bare spot of the order of  $k$*  a set of  $k$  adjacent resonant bonds. The density of such bare spots is  $\varepsilon^k/a$ , and their distribution is the only randomness taken into account: The localized material between these spots is homogeneous with bare localization length  $\zeta_1$ . In other words, we consider a bimodal distribution of nearest-neighbor couplings. We parametrize [43]  $\zeta_1(\varepsilon) = -1/\log(\varepsilon/K)$  with a nonuniversal constant  $K$ . In our scheme, we treat the smallest resonant spots first, and thus it is natural to think of the order  $k$  as an effective scale. We introduce the *running* localization length  $\zeta_k$  as above, but using a unitary  $U$  that eliminates only the spots of the order of  $k' < k$  (alternatively, replacing  $|I| \rightarrow \infty$  by  $|I| \sim ae^{-k}$ ).  $\zeta_k$  is thus the effective localization length seen by spots of the order of  $k$ . It increases with  $k$ , since increasingly more effects of resonant spots are included. Indeed, a calculation yields the important *rule of halted decay*: If a fraction  $\rho_I$  of a stretch  $I$  is thermal, then one finds  $\|V_I\| \sim e^{-(1-\rho_I)|I|/\zeta_1}$  for couplings  $V_I$  that are relevant to our scheme; see [45].

A spot of the order of  $k$  melts (or thermalizes) a region of length  $\ell_k$  on each side of the bare spot. The fraction of space occupied by such thermal regions is  $\rho_k \equiv \varepsilon^k(k + 2\ell_k/a)$ . Hence,  $\rho_k$  is the density of additional thermalized regions that has to be accounted for when passing from  $\zeta_k$  to  $\zeta_{k+1}$ . How precisely  $\zeta$  is assumed to increase does not affect the resulting key features. Using the rule of halted decay, a simple possibility is the mean-field approximation (see [45])

$$\zeta_{k+1}^{-1} = (1 - \rho_k)\zeta_k^{-1}. \quad (5)$$

*Large resonant spots:* We now derive an expression for the length of the melted region  $\ell_k$ , in agreement with the theory developed in Refs. [36,44]. The couplings linking a bare spot of the order of  $k \gg 1$  to its close surroundings are typically resonant in the early iterations of the scheme; see Fig. S2(d) in [45]. After thermalizing  $\ell$  spins on each side, the couplings  $V_{E-l}, V_{E-r}$  from the spot to the spins  $l, r$  just outside it originate microscopically from the (by now rotated) couplings  $\tilde{V}_{E-l}, \tilde{V}_{E-r}$  between the spins  $l, r$  and the peripheral spins of the bare spot. Those scaled as  $\|\tilde{V}_{E-l}\|, \|\tilde{V}_{E-r}\| \sim g_0 e^{-\ell/\zeta_k}$ . Since we have diagonalized the spot by a random unitary, any structure distinguishing the coupling operators from random matrices [i.e., eigenstate thermalization hypothesis (ETH) behavior] has been erased, but the norm of the operators is preserved. Hence, we know that  $\|V_{E-l,r}\| = \|\tilde{V}_{E-l,r}\|$ . As this coupling is now indeed a random matrix acting on a space of dimension  $d_\ell \equiv e^{s(ak+2\ell)}$  (since the spot has grown on two sides), its matrix

elements scale as  $g_0 e^{-\ell/\zeta_k} d_\ell^{-1/2}$ . These become perturbative only once they fall below the level spacing  $\sim d_\ell^{-1}$ , i.e., if  $e^{s(ak+2\ell)/2} e^{-\ell/\zeta_k} \leq 1$ . Thus, spins are thermalized up to distance  $\ell = \ell_k$  with

$$\ell_k = k \frac{sa}{2} \left( \frac{1}{\zeta_k} - s \right)^{-1}. \quad (6)$$

Since  $\zeta_k \rightarrow \zeta$  as  $k \rightarrow \infty$ , but  $\ell_k < \infty$  in the MBL phase, we derive a bound on the typical localization length

$$\zeta \leq \zeta_c \equiv s^{-1}. \quad (7)$$

If this bound were violated, a finite spot would trigger an avalanche and delocalize an arbitrarily large system. As we have seen above, a system appears less localized at larger scales. The picture that emerges is thus that at criticality  $\lim_{k \rightarrow \infty} \zeta_k = s^{-1}$ , with the transition being driven by infinitely large spots.

*Discussion:* From the relation (6) and assuming the recursion equation (5), we can render the flow of  $\zeta_k$  near the transition explicit; see [45]. We find a transition at  $\varepsilon = \varepsilon_c \in ]0, 1[$  defining three regimes: (i) *Localized regime*  $\varepsilon < \varepsilon_c$ .—At a large scale,  $\zeta_k \rightarrow \zeta < \zeta_c$ , and, from Eq. (6),

$$\ell_k/k \rightarrow \ell^* \sim (\zeta_c - \zeta)^{-1} \quad (8)$$

as  $k \rightarrow \infty$ , where  $\ell_k/k$  represents the susceptibility of the material to the insertion of a bare thermal spot of size  $k$ . We expect  $\zeta_c - \zeta \sim (\varepsilon_c - \varepsilon)^{\nu_-}$ , leading to  $\ell^* \sim (\varepsilon_c - \varepsilon)^{-\nu_-}$ . Numerics yields  $\nu_- \in [1/2, 1]$  to be nonuniversal: It depends on the parameter  $K$ . Moreover,  $\ell^*$  is the scale at which the power law distribution of thermal spot sizes is cut off by an exponential tail; see [45].

(ii) *Critical regime*  $\varepsilon = \varepsilon_c$ .—At a large scale,  $\zeta_k \rightarrow \zeta_c$ , and the bound (7) is saturated. From Eq. (6), the susceptibility  $\ell_k/k$  diverges thus as  $k \rightarrow \infty$  ( $\ell_k \sim \varepsilon_c^{-k/2}$ ). Yet, the system is localized, and the thermal density  $\rho$  is strictly smaller than 1. The probability of having a thermal region of size  $\ell$  centered on a given site scales as  $p(\ell) \sim \ell^{-\tau}$ , with  $\tau = 3$ . While the *typical* half-chain entanglement entropy  $S$  is hence bounded, its *average* (over samples) diverges as  $\bar{S} \sim \int^L \ell^2 p(\ell) d\ell \sim \log(L)$  with system size  $L$ .

(iii) *Thermal regime*  $\varepsilon > \varepsilon_c$ .—At small scales, the system appears localized ( $\zeta_k < \zeta_c$ ), but at a finite  $k_*$ ,  $\zeta_{k_*} \geq \zeta_c$  implying  $\ell_{k_*} = \infty$  and triggering an avalanche. The critical core size  $k_*$  diverges logarithmically  $k_* \sim \nu_+ [\log(\varepsilon - \varepsilon_c) / \log(\varepsilon_c)]$  with  $\nu_+ > 0$ .

*Finite-size scaling.*—Let us evaluate the probability  $p(\varepsilon, L)$  of a chain of length  $L$  to be thermal. For large  $L$ , to exponential accuracy, we find  $p(\varepsilon, L) = \exp(-L/L_-)$  with  $L_- \sim |\varepsilon - \varepsilon_c|^{-\nu_-}$  in the MBL phase (by requiring a thermal spot to cover the whole system). On the thermal side, the system is ergodic unless it contains no explosive spots. This yields  $p(\varepsilon, L) = 1 - (1 - \varepsilon^{k_*})^L \approx 1 - \exp(-L/L_+)$  with  $L_+ = \varepsilon^{-k_*} \sim |\varepsilon - \varepsilon_c|^{-\nu_+}$ , with a

nonuniversal  $\nu_+ \in [1, 2]$ . In the critical fan, one finds  $p(\varepsilon, L) \sim L^{-\beta}$  with  $\beta = \tau - 2 = 1$ , as follows from the estimate  $p(\varepsilon, L) \sim L \int_{\ell > L} p(\ell) d\ell$ . In particular, closely above the transition point,  $p(\varepsilon, L)$  has a nonmonotonic behavior as a function of  $L$ : It first decreases polynomially and then grows to 1 exponentially fast.

*Correlation lengths.*—The MBL transition does not manifest itself in thermodynamic correlation functions but only in dynamic properties such as eigenstate correlation functions like

$$\text{Cor}(O_0, O_\ell) \equiv |\langle \Psi, O_0 O_\ell \Psi \rangle - \langle \bar{\Psi}, O_0 \Psi \rangle \langle \Psi, O_\ell \Psi \rangle|, \quad (9)$$

for local operators  $O_i$  acting around site  $i$ , where the average is over both  $\Psi$  and disorder. Numerical studies [42,51] had suggested that

$$\bar{\xi}^{-1} = \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \log \text{Cor}(O_0, O_\ell) \quad (10)$$

diverges at the transition while being finite in both bulk phases. Let us thus analyze what our scheme predicts for  $\bar{\xi}$ .

*MBL phase:* A detailed analysis [45] yields the following picture: If  $O_0, O_\ell$  are located on either side of, but just outside, a thermal region (bare spot + full melted region), then  $\text{Cor}(O_0, O_\ell) \sim 1$ . This is easiest understood by realizing that the couplings connecting the spot to sites 0 or  $\ell$  are barely nonresonant, because those sites are just outside the melted regions. This self-organized criticality affects the diagonalizing unitaries, yielding  $\text{Cor}(O_0, O_\ell) \sim 1$ . Taking into account the probability of thermal regions, we conclude that  $\bar{\xi} \sim \ell^*$ .

*Thermal phase:* It is often suggested that  $\bar{\xi}$  as defined in (10) also diverges from the thermal side. However, within our scheme, an infinite chain on the thermal side is thermal with probability 1, so that ETH applies and the correlator (9) thus decays exponentially in  $\ell$ . Yet, if  $\bar{\xi}$  is defined [51] via Eq. (10) with local operators acting on the two ends of the chain, then, with probability  $e^{-L/L_+}$ , the correlator probes a localized chain, and it can hence fail to decay. By exhibiting [45] a mechanism for long-range correlations, we put the lower bound  $(\log L_+)^{-1} \sqrt{L_+}$  on the associated divergent correlation length.

*Critical slowing down.*—While there is no natural diverging correlation length on the thermal side, there is definitely a diverging timescale  $t_+$ : the inverse of local thermalization rates for typical spins. Indeed, if a spin is eventually thermalized by a thermal region emanating from a bare spot at distance  $\ell$ , the Fermi golden rule roughly yields a flipping rate  $e^{-2\ell/\zeta} \Gamma_0$  with a microscopic rate  $\Gamma_0$ . In the thermal phase, most spins will be thermalized by a bare spot of the order of  $k_* = k_*(\varepsilon)$ , located at a typical distance  $\varepsilon^{-k_*} \sim L_+ \sim (\varepsilon - \varepsilon_c)^{-\nu_+}$ . Therefore,  $t_+ \sim e^{C(\varepsilon - \varepsilon_c)^{-\nu_+}}$  diverges (quasi)exponentially. While this result concerns typical sites in the chain, transport over long distances  $L \gg L_+$  is dominated by rare localized Griffiths regions

[14,52–56], realized here as regions containing no explosive spots with  $k \geq k_*$ . This leads [45] to the subdiffusive scaling  $R(L) \sim L^{C(\varepsilon-\varepsilon_c)^{-\nu_+}}$  of resistance  $R = R(L)$  close to the transition.

*Beyond mean field.*—We emphasize that the mechanism driving the transition and its consequences [such as the divergence of the susceptibility ( $\ell^*$ ) and of the correlation length  $\bar{\xi}$ , the saturation of the bound (7), etc.] follow simply from the consistency of a scale-by-scale analysis. They do not rely on the specific flow equation (5) we used to analyze the scheme more quantitatively. Thus, as expected, they are confirmed by the numerical analysis [43] of our full, nonsimplified scheme. The most important new feature that emerges there is the fact that ergodic spots of low order  $k$  arranged close to each other in a fractal pattern have the same delocalizing power as a spot of much higher order. This means that the concept of a bare ergodic spot should be refined. As a consequence, the tail of the subcritical probability  $p(\ell)$  of thermal spots is a stretched exponential rather than an exponential. Hence,  $\bar{\xi}$  is *strictu sensu* always infinite in the localized phase. However, replacing  $1/\ell$  by  $1/\ell^b$  in the definition (10), with an appropriate choice of  $b$ , yields a finite  $\bar{\xi}$ , which diverges at the transition. Furthermore, in the full scheme the exponents  $\tau$ ,  $\beta$ , and  $\nu_{\pm}$  are modified [their (non)universality being hard to assess], e.g.,  $3 > \tau > 2$ , implying  $\bar{S} \sim L^{3-\tau} \ll L$  at the transition. In particular, the rigorous Harris bounds [57] on  $\nu_{\pm}$  (violated by the mean field) are satisfied, as it has to be. As  $\nu_-$  seems unrelated to avalanches, we expect  $\nu_+ \neq \nu_-$  also beyond the mean field, but we have not been able to settle this numerically in the full scheme.

*Conclusion.*—By studying the role of ergodic spots, we have argued that the MBL-to-ETH critical point is localized. The transition occurs as a quantum avalanche kicked off by the largest ergodic spots. The associated divergent length scales all derive from a divergent susceptibility to large ergodic spots.

We are grateful to A. Scardicchio, D. Huse, and V. Khemani for helpful insights and discussions. F. H. benefited from the support of the projects EDNHS ANR-14-CE25-0011 and LSD ANR-15-CE40-0020-01 of the French National Research Agency (ANR). W. D. R. acknowledges the support of the Flemish Research Fund FWO under Grant No. G076216N. T. T. and W. D. R. have been supported by the InterUniversity Attraction Pole phase VII/18 dynamics, geometry and statistical physics of the Belgian Science Policy. T. T. acknowledges support as a Research Foundation (FWO, Flanders) fellow.

\* thimothee.thiery@kuleuven.be

† huveneers@ceremade.dauphine.fr

‡ markus.mueller@psi.ch

§ wojciech.deroeck@kuleuven.be

- [1] P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
- [2] L. Fleishman and P. W. Anderson, *Phys. Rev. B* **21**, 2366 (1980).
- [3] D. Basko, I. Aleiner, and B. Altshuler, *Ann. Phys. (Amsterdam)* **321**, 1126 (2006).
- [4] I. Gornyi, A. Mirlin, and D. Polyakov, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [5] M. Žnidarič, T. Prosen, and P. Prelovšek, *Phys. Rev. B* **77**, 064426 (2008).
- [6] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [7] A. Pal and D. A. Huse, *Phys. Rev. B* **82**, 174411 (2010).
- [8] V. Ros, M. Müller, and A. Scardicchio, *Nucl. Phys.* **B891**, 420 (2015).
- [9] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, *Phys. Rev. Lett.* **113**, 107204 (2014).
- [10] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **91**, 081103 (2015).
- [11] R. Nandkishore and D. A. Huse, *Annu. Rev. Condens. Matter Phys.* **6**, 15 (2015).
- [12] E. Altman and R. Vosk, *Annu. Rev. Condens. Matter Phys.* **6**, 383 (2015).
- [13] D. A. Abanin and Z. Papić, *Ann. Phys. (Berlin)* **529**, 1700169 (2017).
- [14] D. J. Luitz and Y. B. Lev, *Ann. Phys. (Berlin)* **529**, 1600350 (2017).
- [15] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D. A. Huse, and M. Knap, *Ann. Phys. (Berlin)* **529**, 1600326 (2017).
- [16] J. Z. Imbrie, V. Ros, and A. Scardicchio, *Ann. Phys. (Berlin)* **529**, 1600278 (2017).
- [17] T. Grover, arXiv:1405.1471.
- [18] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. X* **5**, 041047 (2015).
- [19] A. C. Potter, R. Vasseur, and S. Parameswaran, *Phys. Rev. X* **5**, 031033 (2015).
- [20] R. Vosk, D. A. Huse, and E. Altman, *Phys. Rev. X* **5**, 031032 (2015).
- [21] V. Khemani, S.-P. Lim, D. N. Sheng, and D. A. Huse, *Phys. Rev. X* **7**, 021013 (2017).
- [22] V. Khemani, D. N. Sheng, and D. A. Huse, *Phys. Rev. Lett.* **119**, 075702 (2017).
- [23] F. Setiawan, D.-L. Deng, and J. H. Pixley, *Phys. Rev. B* **96**, 104205 (2017).
- [24] A. K. Kulshreshtha, A. Pal, T. B. Wahl, and S. H. Simon, arXiv:1707.05362.
- [25] S. A. Parameswaran, A. C. Potter, and R. Vasseur, *Ann. Phys. (Berlin)* **529**, 1600302 (2017).
- [26] L. Zhang, B. Zhao, T. Devakul, and D. A. Huse, *Phys. Rev. B* **93**, 224201 (2016).
- [27] P. T. Dumitrescu, R. Vasseur, and A. C. Potter, *Phys. Rev. Lett.* **119**, 110604 (2017).
- [28] T. Devakul and R. R. P. Singh, *Phys. Rev. Lett.* **115**, 187201 (2015).
- [29] S. P. Lim and D. N. Sheng, *Phys. Rev. B* **94**, 045111 (2016).
- [30] J. Deutsch, *Phys. Rev. A* **43**, 2046 (1991).
- [31] M. Srednicki, *Phys. Rev. E* **50**, 888 (1994).
- [32] M. Rigol, V. Dunjko, and M. Olshanii, *Nature (London)* **452**, 854 (2008).

- [33] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.* **65**, 239 (2016).
- [34] A. Chandran, A. Pal, C. Laumann, and A. Scardicchio, *Phys. Rev. B* **94**, 144203 (2016).
- [35] P. Ponte, C. Laumann, D. A. Huse, and A. Chandran, *Phil. Trans. R. Soc. A* **375**, 20160428 (2017).
- [36] W. De Roeck and F. Huveneers, *Phys. Rev. B* **95**, 155129 (2017).
- [37] F. Wegner, *Ann. Phys. (Berlin)* **506**, 77 (1994).
- [38] J. Z. Imbrie, *J. Stat. Phys.* **163**, 998 (2016).
- [39] L. Rademaker and M. Ortuño, *Phys. Rev. Lett.* **116**, 010404 (2016).
- [40] D. Pekker, B. K. Clark, V. Oganesyan, and G. Refael, *Phys. Rev. Lett.* **119**, 075701 (2017).
- [41] C. Monthus, *J. Phys. A* **49**, 305002 (2016).
- [42] L. Rademaker, M. Ortuño, and A. M. Somoza, *Ann. Phys. (Berlin)* **529**, 1600322 (2017).
- [43] T. Thiery, M. Müller, and W. De Roeck, [arXiv:1711.09880](https://arxiv.org/abs/1711.09880).
- [44] D. J. Luitz, F. Huveneers, and W. De Roeck, *Phys. Rev. Lett.* **119**, 150602 (2017).
- [45] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.121.140601> for details about (A) the resonance condition, (B) the rule of halted decay, (C) the mean-field model, (D) the calculation of correlation lengths, (E) the scheme in physical time, (F) the comparison with previous RG schemes.
- [46] It is easy to see that one gets nonsense by eliminating couplings in the wrong order. For example, assume that we would always eliminate first the couplings that touch a resonant spot. Then such couplings would almost always be resonant, and a single resonant spot would almost always thermalize the full chain.
- [47] M. Serbyn, Z. Papić, and D. A. Abanin, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [48] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [49] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phys. Rev. B* **90**, 174202 (2014).
- [50] The operator norm is defined as  $\|O\| = \sup_{\psi: \|\psi\|_2=1} \|O\psi\|_2$ , where  $\psi$  runs over vectors in Hilbert space and  $\|\cdot\|_2$  is the usual  $L_2$  norm.
- [51] L. Zhang, V. Khemani, and D. A. Huse, *Phys. Rev. B* **94**, 224202 (2016).
- [52] Y. B. Lev, G. Cohen, and D. R. Reichman, *Phys. Rev. Lett.* **114**, 100601 (2015).
- [53] D. J. Luitz, N. Laflorencie, and F. Alet, *Phys. Rev. B* **93**, 060201 (2016).
- [54] S. Gopalakrishnan, K. Agarwal, E. A. Demler, D. A. Huse, and M. Knap, *Phys. Rev. B* **93**, 134206 (2016).
- [55] M. Žnidarič, A. Scardicchio, and V. K. Varma, *Phys. Rev. Lett.* **117**, 040601 (2016).
- [56] A. Nahum, J. Ruhman, and D. A. Huse, *Phys. Rev. B* **98**, 035118 (2018).
- [57] A. Chandran, C. R. Laumann, and V. Oganesyan, [arXiv:1509.04285](https://arxiv.org/abs/1509.04285).