LOCOBATIC THEOREM FOR DISORDERED MEDIA AND VALIDITY OF LINEAR RESPONSE

PRELIMINARY DRAFT

WOJCIECH DE ROECK, ALEXANDER ELGART, AND MARTIN FRAAS

Abstract. The adiabatic theorem of quantum mechnaics cannot hold or systems in a localization regime, because the evolution of eigenvectors as a parameter is varied (also called "spectral flow") is generically non-local. However, there is a remnant of the adiabatic theorem, which we call the "locobatic theorem". It refers to the fact that adiabatic evolution corresponds, with high probability, to the spectral flow of a local restriction of the Hamiltonian. We make the above statements precise for a class of Hamiltonians describing a particle in a disordered background. An important application is a justification of the linear response formula for the Hall conductivity in a 2D system with the Fermi energy lying in a mobility gap.

1. Introduction

1.1. Adiabatic theorems.

In quantum mechanics, a central problem is to solve and understand the linear initial value problem (IVP):

$$i\dot{\psi}(t) = H(t)\psi(t), \quad \psi(0) = \psi_o, \tag{1.1}$$

where $H(t)$ is a self-adjoint family of operators on some Hilbert space $\mathcal{H}$ (the so-called Hamiltonian), and ψ_o is a normalized vector on $\mathcal{H}$ (the initial wave packet of the system). The solution of the IVP becomes trivial in the case of time-independent operators $H(t) = H_o$ and the initial state ψ_o being an eigenvector for H_o. In this case, the evolution $\psi(t)$ coincides with ψ_o up to an acquired phase.

A more interesting and physically realistic situation arises when the dependence on time in $H(t)$ is present, but is slow (adiabatic). In this case, the evolution $\psi(t)$ is expected to follow the spectral evolution of the Hamiltonian $H(t)$ (the assertion known as the *adiabatic theorem of quantum mechanics*). Of course, slow is a relative concept, and we need to quantify the reference time scale for these purposes. In the standard adiabatic theorem, such parameter is given by the spectral gap in $H(t)$ (note that energy has units of time^{-1} in (1.1)). To make the statement more quantitative, it is convenient to consider the family $H(\epsilon t)$, where ϵ is a small (adiabatic) parameter, and the physical time t runs over the long interval $[0, 1/\epsilon]$. After a change of variables $s = \epsilon t$ where s is a rescaled time, the relevant IVP becomes

$$i\epsilon\dot{\psi}_\epsilon(s) = H(s)\psi_\epsilon(s), \quad \psi_\epsilon(0) = \psi_o, \quad s \in [0, 1]. \tag{1.2}$$

We denote by $U_\epsilon(s)$ the corresponding propagator, i.e. the unitary operator that solves the IVP

$$i\epsilon\partial_s U_\epsilon(s) = H(s)U_\epsilon(s), \quad U_\epsilon(0) = \mathbb{I}. \tag{1.3}$$

Let us assume that the spectrum $\sigma(H(s))$ of the operator $H(s)$ contains a set $\mathcal{S}(s)$ which is isolated from the rest of the spectrum by a uniform distance g (the spectral gap). Denoting by $P(s)$ the spectral projection of $H(s)$ onto $\mathcal{S}(s)$, and assuming that $P(0)\psi_o = \psi_o$, the (qualitative) adiabatic theorem states that

$$\lim_{\epsilon\to 0} \|\psi_\epsilon(s) - P(s)\psi_\epsilon(s)\| = 0, \tag{1.4}$$

provided $H(s)$ is smooth. In fact a stronger statement holds true, namely

$$\lim_{\epsilon\to 0} \|U_\epsilon(s)P(0)U_\epsilon^*(s) - P(s)\| = 0, \tag{1.5}$$

Date: August 10, 2021.

and one can make the error estimate for the norm above explicit, in terms of its ϵ and g dependencies, see e.g., Lemma 4.5 below.

The adiabatic theorem and its derivatives play an important role in the various branches of quantum and statistical mechanics (e.g., our second result below concerns the derivation of the Green–Kubo transport formula using this theory). The first results on adiabatic behavior go back to the dawn of quantum mechanics and are due to Born and Fock (1928). The modern adiabatic theory was initiated by Kato in 1950 [Ka] and has since been since studied intensively in mathematical physics literature. The adiabatic theorem has been extended to a situation where the family $P(s)$ is smooth, but no gap is present, [B, AE, AHS]. This situation usually occurs for a ground state in the threshold of continuous spectrum. More recently, the adiabatic theorem was established for certain systems characterized by a spectral gap but non-smooth $P(s)$, [BDF]. This situation arises in the context of the thermodynamic limit for many-body systems. This paper considers a case where *both* conditions fail to hold, in general. Such a situation occurs in the localized regime of a disordered H, as will be described now

1.2. **Localized systems and resonant hybridization.** We will be interested in the family

$$H(s) = H + \beta W(s) \tag{1.6}$$

on $\ell^2(\mathbb{Z}^d)$, where H is a disordered Hamiltonian, β is a small parameter and $W(\cdot)$ is a uniformly bounded family of smooth operators (the driving). We say that an open interval $J_{loc} \subset \sigma(H)$ is a *mobility gap* or a region of exponential localization if the spectrum of H in J_{loc} is of pure point type and there exist constants $0 < C, m, \mu < \infty$, such that for each eigenpair $(E_i, \psi_i), E_i \in J_{loc}$ one can find $x_i \in \mathbb{Z}^d$, called a *localization center* for ψ_i, satisfying

$$|\psi_i(x)| \leq C\,|x|^m\, e^{-\mu|x-x_i|}. \tag{1.7}$$

The prototypical example of such H is the Anderson model $H = -\Delta + V_\omega$ on $l^2(\mathbb{Z}^d)$, where Δ denotes the discrete Laplacian and V_ω is a multiplication operator, $V_\omega\psi(x) = \omega(x)\psi(x)$ for $\psi \in \ell^2(\mathbb{Z}^d)$, where $\omega(x)$ is drawn as an i.i.d. random variable with some joint probability distribution μ. The Anderson Hamiltonian is known to display exponential localization in the vicinity of spectral edges, at large values of disorder (condition on the distribution μ) and in dimension $d = 1$, for almost all configurations ω. We will not attempt to cite the extensive literature for reviews, results and open problems concerning this model and its variants but rather refer the interested reader to the recent monograph [AW] on the subject.

For the purposes of the present paper, we stress that one should not expect much uniformity of the localization properties as a function of s or β. More concretely, even if the property (1.7) holds for all $s \in [0, 1]$, we do not expect the constant C to be bounded in s on any open subset of $[0, 1]$, provided that W is sufficiently nontrivial. The destruction of such uniform localization properties should intuitively proceed via a mechanism known as *resonant hybridization*, see e.g. [AW, Chapter 15]. However, as far as we are aware, there are no prior results that work for $\mathbb{Z}^d$ systems, for any d illustrated adequately by the two-level system with a Hamiltonian $H(s)$ of the form

$$H(s) = \begin{pmatrix} g & s \\ s & -g \end{pmatrix}, \quad s \in (-1, 1), \quad g \ll 1.$$

When $s = 0$, an eigenbasis is $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. These remains approximate eigenvectors for $H(s)$ provided $|s| \ll g$. However, the picture is different for the case when the relation between the energy gap $2g$ and the tunneling amplitude $|s|$ is reversed: When $g \ll |s|$ an approximate eigenbasis is given by $\{e_1 \pm e_2\}$. I.e., the eigenfunctions are no longer localized in the basis $\{e_i\}$ and instead are given by hybridized functions which are combinations of these vectors.

In a generic disordered system such two-level description emerges when one wants to single out the interaction behavior of a pair of spatially separated eigenstates and as such is present at all scales. This leads us to the above-mentioned intuition that C cannot be not locally bounded. Let us also mention that, as we see in this example, the hybridization phenomenon is usually tied to an avoided level crossing. If we consider the spectral flow of eigenvectors as a function of s, then we see that this flow will transition between e_1 and e_2 in a time span of approximate length g. If this phenomenon occurs as well in our extended disordered system, then this means that the spectral flow is very nonlocal, as $e_{1,2}$ can be localized arbitrarily far away from each other. More precisely, if we consider a finite volume restriction of H, say to a box with side length $\mathcal{L}$, we can then label the eigenstates $\psi_{i,s}$ so that for each i, $t \mapsto \psi_{i,s}$ is continuous. However, we expect do the modulus of continuity to diverge badly as $\mathcal{L} \to \infty$, see Theorem 1.1.

We are not aware of any prior results making this heuristics precise and therefore, in Appendix A we show the emergence of the hybridization rigorously for a one dimensional system. Specifically, we prove Theorem A.18, which qualitatively can be formulated as

Theorem 1.1. *Let H_o be the standard Anderson model in 1d. Then, under some additional regularity assumptions on the random potential and mild assumptions on W, the eigenfunction hybridization occurs on all scales with the scale-independent probability. The corresponding eigenvalues exhibit avoided level crossings.*

1.3. **The locobatic behaviour.** Theorem 1.1 above leads to an interesting question: The folk adiabatic theorem suggests that dynamics should follow the spectral data, i.e., the spectral flow $s \mapsto \psi_{i,s}$, when the adiabatic parameter ϵ is small enough. However, as we saw above, this spectral flow is extremely nonlocal, whereas the physical evolution can not be arbitrarily nonlocal. The way that this dilemma is resolved, we believe, is that the physical evolution of an initial eigenvector is for most values of s stays close to one of the global eigenvectors $\psi_{i,s}$, but the index s varies wildly with s. A simpler take on this is that the evolution of the initial eigenvector stays, for all times s, close to an instantaneous eigenvector of the restriction of H to a local box around the support of the initial eigenvector. We will refer to this statement as a *locobatic theorem*, and state it quantitatively as Theorem 1.5 below.

In order to state this assertion properly, we need to introduce the necessary framework first.

An operator K acting on $\ell^2\left(\mathbb{Z}^d\right)$ is range-r for some $r \in \mathbb{N}$ if

$$K(x,y) := \langle \delta_x, K\delta_y \rangle = 0 \text{ provided } |x-y| > r, \quad x,y \in \mathbb{Z}^d,$$

where $|x-y|$ stands for the ℓ^∞ distance in $\mathbb{Z}^d$.

Assumption 1.2. The operators $H(s)$ are uniformly bounded, smooth, range-r, self-adjoint operators, acting on $\ell^2\left(\mathbb{Z}^d\right)$, of the form $H(s) = H + \beta W(s)$. In addition,

$$\|H(s)\| \le C, \quad \|W^{(k)}(t)\| \le C_k, \quad W^{(k)}(0) = W^{(k+1)}(1) = 0,$$

for some constants C, C_k and $k \in \mathbb{N}_0$.

For any $\Theta \subset \mathbb{Z}^d$, we denote by H^Θ the canonical restriction $\chi_\Theta H \chi_\Theta$ of H to $\ell^2(\Theta)$.

Assumption 1.3 (Finite range of disorder correlations)**.** For any pair of subsets Θ, Φ of $\mathbb{Z}^d$ that satisfy $\operatorname{dist}(\Theta, \Phi) > r$ the operators H^Θ and H^Φ are statistically independent.

For any region $\Theta \subset \mathbb{Z}^d$ and $x, y \in \Theta$, we define

$$|x-y|_\Theta = \min\left(|x-y|, \left(\operatorname{dist}(x, \partial_1\Theta) + \operatorname{dist}(y, \partial_1\Theta)\right)\right), \tag{1.8}$$

with the interior boundary $\partial_1\Theta = \{x \in \Theta, \operatorname{dist}(x, \Theta^c) = 1\}$. This distance function regards $\partial_1\Theta$ as a single point. It permits us to state that there is exponential decay in the bulk without ruling out absence of decay along the boundary due to delocalized edge modes. With this preparation, our assumption of Anderson localization in an interval J_{loc} for H reads

Assumption 1.4 (Fractional moment condition on J_{loc})**.** There exist $q \in (0,1)$ and $C_q, \mu > 0$ such that for any subset Θ of $\mathbb{Z}^d$ we have

$$\sup_{E \in J_{loc}} \mathbb{E}\left(\left|(H^\Theta - E - i0)^{-1}(x,y)\right|^q\right) \le C_q e^{-\mu|x-y|_\Theta} \text{ for all } x, y \in \Theta, \tag{1.9}$$

where $\mathbb{E}(\cdot)$ stands for expectations with respect to ω.

Since the next result is easier stated in finite volume, we introduce a periodized restriction of $H(s)$ to a discrete torus $\mathbb{T} = \mathbb{T}^d_M$ which we identify with the square $[1, M]^d$ with opposite faces identified. This is defined as

$$H^{\mathbb{T}}(x,y) = \sum_{m,n \in M\mathbb{Z}^d} H(x+m, y+n), \quad x, y \in \mathbb{T}. \tag{1.10}$$

Our two main parameters are the adiabaticity parameter ϵ and the driving strength β, introduced earlier in (1.2) and (1.6), respectively. In our results we will use three exponents,

$$\xi = \frac{d}{q}, \quad p_1 > 2d + 1/2 + d/q, \quad p_2 > \max\left(d + \frac{1}{2} + \xi, 2\xi\right),$$

with fixed p_1, p_2 satisfying the inequalities. We allow for the system size M to be arbitrarily large and all our estimates will be uniform in M.

The following is then the 'locobatic theorem'. It is based on the emergence of a local structure for the spectral data associated with a torus, once we partition it into smaller boxes of the linear size ℓ.

Theorem 1.5 (Locabatic theorem on a torus). *Assume assumption 1.2 and assumptions 1.3–1.4 for $H(0)$. We introduce a scale parameter $\ell \in \mathbb{N}$ satisfying*

$$\ell^{p_2} \le \epsilon^{-1} \le Ce^{c\sqrt{\ell}}, \qquad \beta^{1/p_1} \le 1/\ell \tag{1.11}$$

Let J'_{loc} be any closed interval contained in J_{loc}. With probability at least $1 - e^{-c\sqrt{\ell}}$, the following holds true for a fraction of at least $1 - e^{-c\sqrt{\ell}}$ of eigenstates ψ of $H^{\mathbb{T}}$ with eigenvalue $E \in J'_{loc}$: There is region $R \subset \mathbb{T}$ with $\mathrm{diam}(R) \le c\ell^{3/2}$ *and an isolated spectral patch $S(0) \subset \sigma(H^R(0))$ such that*

(i) For all s, the spectral patch remains isolated from the rest of the spectrum $\sigma(H^R(s))$. We denote the associated spectral projector by $P(s)$.

(ii) The solution $\psi_\epsilon(s)$ of the IVP started from ψ satisfies

$$\max_{s \in [0,1]} \|(1 - P(s))\psi_\epsilon(s)\| \le C\left(\epsilon \ell^{d+1/2+\xi} + e^{-c\sqrt{\ell}}\right). \tag{1.12}$$

The bound can be improved for $s = 1$; For any $N \in \mathbb{N}$,

$$\|(1 - P(1))\psi_\epsilon(1)\| \le C_N \left(\epsilon^N \left(\ell^{N(d+1/2+\xi)} + \ell^{(2N+1)\xi}\right) + e^{-c\sqrt{\ell}}\right). \tag{1.13}$$

If the spectrum of $H^{\mathcal{R}}$ is *level-spaced*, i.e. if the probability of a spacing significantly smaller than $|\mathcal{R}|$ is small (as one can prove, e.g., for the standard Anderson model [KM] and at the bottom of the spectrum for more general random models, [DE]), then with large probability the spectral patch $S(s)$ consists of a simple eigenvalue and hence $P(s)$ is a rank-one projector. Moreover, with large probability, one can argue that for a large fraction of times s, the range of $P(s)$ stays close to an eigenprojection of the global Hamiltonian $H^{\mathbb{T}}(s)$. We don't expect this property to hold for all times s on a basis of the hybridization result, Theorem 1.1, which shows that the physical evolution cannot follow the non-local spectral flow.

We will use generic constants C, c, whose values can change from line to line, but they will be uniform in M, ϵ, β and in the scale parameter ℓ introduced below. They will however depend, in general, on the other parameters and constants introduced above (such as the range r and the probability distribution μ, as well as constants C_q, C_k, etc.).

1.4. **Justification of linear response.** One of the main applications of the locobatic theorem is a proof of validity of the linear response relation for Hall conductivity.

We describe the setup. Take $d = 2$ and let $\mathbb{Z}^2 \ni x = (x_1, x_2)$. We will denote by Λ_n the characteristic function of the set $\{x \in \mathbb{Z}^2 : x_n \ge 0\}$, $n = 1, 2$. We consider a Hamiltonian of a form

$$H(s) = H_0 + \beta g(s)\Lambda_2,$$

corresponding to an electric potential $\beta g(s)$ applied across the 2-direction. The function g

(i) $g \in C^\infty[-1,1]$
(ii) $g(s) = 0$ for $s \leq s_0$ for some $s_0 > -1$.
(iii) $g(s) = 1$ for $s \geq 0$

In the previous sections, we considered the adiabatic evolution from $s = 0$ to $s = 1$, but now it is more natural to consider the time interval $[-1, 1]$. From time -1 to 0 we adiabatically switch on the perturbation $\beta g(s)\Lambda_2$, an electric field pointing in x_2-direction, localised along the $x_2 = 0$-axis. The total charge passing through a fiducial line, $x_1 = 0$, from time $t = 0$ up to a time $t = T$ is given by

$$Q = \int_0^T j(t)dt = \int_0^T \mathrm{tr}(P_\epsilon(t) - P)J)dt,$$

where j is the current, $P = P_{<E_F}(H_0)$ is the Fermi projection of the unperturbed Hamiltonian, $P_\epsilon(t)$ is the solution of the driven Schrödinger equation with $P_\epsilon(t) = P$, and $J = i[H, \Lambda_1]$ is the current observable (the subtraction of P inside the trace corresponds to the removal of the so-called persistent current). As we show in the proof, the product $(P_\epsilon(t) - P)J)$ is indeed a trace-class operator, even if the neither of the two factors separately is trace-class. Upon rescaling the total time as $T = \epsilon^{-1}$ and introducing the scaled time $s = \epsilon t$, we get

$$Q = \frac{1}{\epsilon}\int_0^1 \mathrm{tr}(P_\epsilon(s) - P)J)ds,$$

where $P_\epsilon(s)$ solves the adiabatic Schrödinger equation

$$i\epsilon\partial_s P_\epsilon(s) = [H(s), P_\epsilon(s)], \quad P_\epsilon(-1) = P.$$

The Hall conductance is defined as a proportionality constant between the applied potential difference (spatial integral of the electric field) and the current flowing in the perpendicular direction, i.e. the measured conductance σ_m is defined by a relation

$$Q = \sigma_m \frac{\beta}{\epsilon}\int_0^1 g(s)ds,$$

which gives

$$\sigma_m = \frac{1}{\beta}\int_0^1 \mathrm{tr}(P_\epsilon(s) - P)J)ds.$$

We show the validity of linear response in this system by establishing that the limit

$$\lim_{\beta\to 0}\frac{1}{\beta}\int_0^1 \mathrm{tr}(P_\epsilon(s) - P)J)ds$$

exists for a protocol such that $\epsilon \ll \beta$ and is equal to the conductance σ obtained from the Kubo formula, see e.g., [AG],

$$\sigma := \mathrm{tr}(P[[P, \Lambda_1], [P, \Lambda_2]]). \tag{1.14}$$

The condition $\epsilon \ll \beta$ ensures that a macroscopic amount of charge is transported during the process. The shape of g determines the state preparation protocol. In our case, it corresponds to a common choice in which the state is adiabatically prepared before the measurement of the current takes place (see e.g. [AG]). The most intuitive setup would be to take $\epsilon \to 0$ first and only then $\beta \to 0$. This is certainly beyond the reach of our theorem and we do not know whether in that setup the expression still equals σ[1].

Our result reads

Theorem 1.6. *Suppose that H satisfies Assumptions 1.3–1.4 with E_F lying in the interior of J_{loc}. Assume moreover that $e^{-\beta^{-p_1/2}} < \epsilon < \beta^{p_2 p_1}$. Then*

$$|\sigma - \sigma_m| \leq C\frac{\epsilon}{\beta^p} + \mathcal{O}(\beta^\infty + \epsilon^\infty),$$

[1] It is very likely that the evolution at very small ϵ is delocalized, but we do not quite grasp the implications of that for our problem.

holds with probability $1 - e^{-\beta^{-p'}}$ *for some integers* p, p'.

1.5. **Adiabatic theorem in the strong operator topology.** An interesting application of our results for a finite geometry is the following assertion that holds for $\mathbb{Z}^d$.

Theorem 1.7. *Let* $E_F \in J_{loc}$ *and suppose that Assumption 1.2–1.4 hold. Then (cf.* (1.5)*)*

$$\text{s-lim}_{\epsilon,\beta\to 0} U_\epsilon(s) P_{E_F}(0) U_\epsilon^*(s) = P_{E_F}(s), \quad s \in [0,1]$$

almost surely, provided we take a limit maintaining the relation

$$e^{-1/\beta^p} < \epsilon < \beta^{p_2 p_1}, \tag{1.15}$$

with $p = p_1/2$.

Remark 1.8. For $\epsilon = \beta$ this statement was proven earlier in [ES]. Let us also mention that the choice of the topology (the strong operator topology) here is essential - in the absence of the spectral gap the result is not expected to hold in the norm operator topology, see [ES] for a counterexample. A major difference between the prior work and our result is that the (formal) total variation of the perturbation, $\epsilon^{-1}\beta \|W\|$ blows up as $\epsilon, \beta \to 0$, thanks to the constraints on ϵ and β.

Additional notation. By $\Lambda_R(y) \subset \mathbb{Z}^d$ we will denote a cube $\Lambda_R = \Lambda_R(y) := \big([-R,R]^d + y\big) \cap \mathbb{Z}^d$ for $y \in \mathbb{Z}^d$, with the side length $2R$. For a subset $\Phi \subset \mathbb{Z}^d$, we will denote by $\partial_\ell \Phi$ it's ℓ-extended boundary, i.e.,

$$\partial_\ell \Phi = \{x \in \Phi : \ \text{dist}\,(x, \Phi^c) \le \ell\}. \tag{1.16}$$

By Φ_ℓ we will denote

$$\Phi_\ell = \Phi \setminus \partial_\ell \Phi. \tag{1.17}$$

For a Hermitian operator H, we denote by $P_J(H)$ the spectral projection of H on the set $J \subset \mathbb{R}$. For an operator X we denote $\bar{X} := 1 - X$. For $\mathcal{A} \subset \mathbb{T}$, $c \in \mathbb{R}_+$, and $\ell \in \mathbb{N}$, let $\rho_{\mathcal{A}}^\ell$ be a (scaled) distance function

$$\rho_{\mathcal{A}} := \rho_{\mathcal{A}}^\ell(x) = \frac{\text{dist}\,(\mathcal{A}, \{x\})}{\sqrt{\ell}}. \tag{1.18}$$

Let K_ℓ^c be a dilation of an operator K with respect to $\rho_{\mathcal{A}}^\ell$ of the form

$$K_\ell^c = \mathrm{e}^{-c\,\rho_{\mathcal{A}}^\ell} K \,\mathrm{e}^{c\,\rho_{\mathcal{A}}^\ell}. \tag{1.19}$$

We will denote

$$\|K\|_{c,\ell} = \|K_\ell^c\| \tag{1.20}$$

in the sequel. This norm is multiplicative, i.e.,

$$\|AB\|_{c,\ell} \le \|A\|_{c,\ell} \|B\|_{c,\ell} \tag{1.21}$$

for a pair of operators A, B.

1.6. **Outline of the proofs.** We will now comment on the arguments pertaining to the proof of our core assertions, namely Theorems 2.1–2.2 below. We will not comment on the derivations of the remaining results, as they follow via the more standard strategy.

We first introduce the concepts of local and ultra-local structures. In order to describe our constructions with the least possible number of parameters we will use the scale variable $\ell \in \mathbb{N}$ introduced in Theorem 1.5. It will be convenient to formulate them on a torus $\mathbb{T}$ whose linear dimension is $\mathcal{L} = e^{c\sqrt{\ell}}$ but this condition can be relaxed.

Let $J \subset J_{loc}$ and let $\{(E_n, \psi_n)\}$ be a collection of eigenpairs for $H^{\mathbb{T}}(0)$ with energies J. We will say that $H^{\mathbb{T}}(0)$ possesses an *ultralocal structure* in J if there exists a disjoint collection $\{\mathcal{T}_\gamma\}$ of subsets of $\mathbb{T}$ such that $\text{diam}\,(\mathcal{T}_\gamma) \le C\ell^{3/2}$ for each γ with the following property: For each ψ_n, there exists γ such that

$$\big\|\bar{P}_{J_\gamma}\big(H^{\mathcal{T}_\gamma}(0)\big)\,\psi_n\big)\big\| \le e^{-c\sqrt{\ell}}. \tag{1.22}$$

Let us note that random Schrödinger operators $H(0)$ satisfying Assumption 1.4 possess the ultra-local property with probability $\ge 1 - e^{-c\sqrt{\ell}}$ provided the length of the interval J is of the order $\ell^{-\xi}$, in fact a stronger statement holds true, see Theorem 5.4 below. Unfortunately, localization in the usual sense (or in an ultra-local sense to this matter) breaks down under perturbations due to the hybridization phenomenon. So the first step is to identify a weaker notion than ultra-locality that however remains stable under small perturbations.

We will say that $H^{\mathbb{T}}(s)$ possesses a *local structure* in $J \subset J_{loc}$ if there exists a disjoint collection $\{\mathcal{T}_\gamma\}$ of subsets of $\mathbb{T}$ such that $\operatorname{diam}(\mathcal{T}_\gamma) \leq \ell^{3/2}$ for each γ with the following properties:

(i) (Local Gap) There exist intervals $J_\gamma = [E_\gamma^-, E_\gamma^+]$ comparable in length to J such that

$$J_\gamma \subset J \text{ and } \operatorname{dist}\big(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma}(s))\big) \geq \Delta; \tag{1.23}$$

(ii) (Support of spectral projections) Let $\mathcal{T} := \cup_\gamma \mathcal{T}_\gamma$. Then

$$\big\|P_J(s)\chi_{\mathbb{T}\setminus\mathcal{T}_{8\ell}}\big\| \leq e^{-c\sqrt{\ell}}, \tag{1.24}$$

and

$$\|P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \chi_{\partial_\ell\mathcal{T}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\partial_\ell\mathcal{T}} - \chi_{\mathcal{T}_{8\ell}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\mathcal{T}_{8\ell}}\| \leq e^{-c\sqrt{\ell}}. \tag{1.25}$$

The unperturbed Hamiltonian possesses a local structure for small but not too small Δ. As we shall see in the proof of Theorem 2.1, the local structure is stable under the perturbation, i.e., if the Hamiltonian possesses a local structure for $s = 0$ on J, it possesses it for *all* s on a slightly smaller interval J', provided β is sufficiently small. The reason for this stability is related to the fact that under small local perturbations an eigenstate with energy E is well described in terms of the thin spectral projection about E of the unperturbed operator. Since the latter is supported in the localized patches $\mathcal{T}_\gamma$, so is the eigenstate. The locality property is fully compatible with the hybridization effect: Even if initially the state is ultra-local (concentrated in a single patch $\mathcal{T}_{\gamma_o}$), it can (and in fact will) hybridize to a number of different patches $\mathcal{T}_\gamma$ as s increases.

The scaling of various objects with ℓ depends on q, d and our choice of sub-exponential error $\exp(-c\sqrt{\ell})$. The correct scaling of Δ and β to ensure existence of local structure is given in Theorem 2.1.

Once the local structure for the family $H(s)$ is established, one can use an (enhanced) version of the standard, gapped adiabatic theorem (Lemma 4.5) to control the behavior of the individual spectral patches $P_{J_\gamma}\big(H^{\mathcal{T}_\gamma}(s)\big)$. Here we explicitly invoke part (1.8.(i)) of the definition above. This, in turn, allows us to control a physical evolution of spectral data $Q(s)$ for $H^{\mathbb{T}}(s)$ near the energy E, see Section 4.5 for details. Finally, we show that it translates to the adiabatic theorem for the (distorted) Fermi projection, Theorem 2.2. The principal idea here is that the removal of the spectral data $Q(s)$ on one hand creates a spectral gap for H (making the standard adiabatic theorem applicable) and on the other does not distort the adiabatic behavior of the system too much since $Q(s)$ itself evolves adiabatically, the feature we verified in the previous step.

2. Adiabatic theorem on a torus

We shorthand $P_J(s) := P_J(H^{\mathbb{T}}(s))$ and $P_J := P_J(0)$ in this section.

We will show Section 5 that Anderson type models posses a local structure. In fact, a stronger statement holds true:

Theorem 2.1 (Local structure of $H^{\mathbb{T}}(s)$)**.** *Suppose that H satisfies Assumptions 1.3–1.4 and the family $H(s)$ satisfies Assumption 1.2. Let*

$$\mathcal{L} = e^{c_1\sqrt{\ell}}, \quad V_\ell = \ell^{d+1/2}, \quad \delta = c_2\ell^{-\xi}, \quad \Delta = c_3 V_\ell^{-1}\ell^{-\xi}, \tag{2.1}$$

and suppose that $\beta \leq \ell^{-p_1}$. Then there exists constants $c, c_1, c_2, c_3, c_4, c_5, c_6$ such that for ℓ large enough $H^{\mathbb{T}}(s)$ possesses a local structure for the energy interval $J = (E - 6\delta, E + 6\delta)$: One can find a disjoint collection $\{\mathcal{T}_\gamma\}$ of subsets of Λ such that $|\mathcal{T}_\gamma| \leq c_4 V_\ell$, $\operatorname{diam}(\mathcal{T}_\gamma) \leq c_5\ell^{3/2}$ for each γ and that the following conditions are met:

(i) (Local Gap) There exists intervals $J_\gamma = [E_\gamma^-, E_\gamma^+]$ *such that*

$$(E - 3\delta, E + 3\delta) \subset J_\gamma \subset J \text{ and } \operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma}(s))\right) \geq \Delta; \tag{2.2}$$

(ii) (Support of spectral projections) Let $\mathcal{T} := \cup_\gamma \mathcal{T}_\gamma$*. Then*

$$\left\| P_J(s) \chi_{\Lambda \setminus \mathcal{T}_{8\ell}} \right\| \leq e^{-c\sqrt{\ell}}, \tag{2.3}$$

and

$$\| P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \chi_{\partial_\ell \mathcal{T}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\partial_\ell \mathcal{T}} - \chi_{\mathcal{T}_{8\ell}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\mathcal{T}_{8\ell}} \| \leq e^{-c\sqrt{\ell}}. \tag{2.4}$$

(iii) (Exponential Decay of Correlations) Let $\mathcal{A}_o = \partial_\ell \mathcal{T}_\gamma \cup (\mathcal{T}_\gamma)_{8\ell}$*, then (with* $\mathcal{A} = \mathcal{A}_o$ *in* (1.18)–(1.20)*) we have*

$$\left\| \left(H^{\mathcal{T}_\gamma}(s)) - z \right)^{-1} \right\|_{c,\ell} \leq \frac{\ell^{3d}}{\Delta} \frac{1}{\langle Im\, z \rangle}, \tag{2.5}$$

for $z \in \mathbb{C}$ *with* $\operatorname{Re}(z) = E_\gamma^\pm$*.*

holds true with probability $> 1 - e^{-c_6\sqrt{\ell}}$*.*

The dependence on β *here is deterministic, i.e. there exist a subset of configurations of probability* $> 1 - e^{-c_6\sqrt{\ell}}$ *such that the conclusions hold for all* $\beta < \ell^{-p_1}$*.*

An additional statement we will establish is

Theorem 2.2 (Locabatic theorem for distorted Fermi projection)**.** *In the setting of Theorem 2.2, assume in addition that*

$$e^{-c\sqrt{\ell}} \leq \epsilon \leq \ell^{-p_2}, \tag{2.6}$$

and fix $N \in \mathbb{N}$*. Then for* ℓ *large enough, there exists a smooth family of orthogonal projections* $\mathcal{Q}(s)$ *with the following properties:*

(i) $\left\| [\mathcal{Q}(s), H^{\mathbb{T}}(s)] \right\| \leq C_N \left(\epsilon \Delta^{-1} + e^{-c\sqrt{\ell}} \right)$*;*

(ii) $\left\| P_{<E-6\delta}(H^{\mathbb{T}}(s)) \bar{\mathcal{Q}}(s) \right\| + \left\| \mathcal{Q}(s) P_{>E+6\delta}(H^{\mathbb{T}}(s)) \right\| \leq C_N \left(\epsilon \Delta^{-1} + e^{-c\sqrt{\ell}} \right)$*;*

(iii) If we denote by $\mathcal{Q}_\epsilon(s)$ *the solution of IVP* $i\epsilon \dot{\mathcal{Q}}_\epsilon(s) = [\mathcal{Q}_\epsilon(s), H^{\mathbb{T}}(s)]$*,* $\mathcal{Q}_\epsilon(0) = \mathcal{Q}(0)$*, we have*

$$\| \mathcal{Q}_\epsilon(s) - \mathcal{Q}(s) \| \leq C_N \left(\epsilon^N \left(\frac{1}{\Delta^N} + \frac{1}{\delta^{2N+1}} \right) + e^{-c\sqrt{\ell}} \right). \tag{2.7}$$

Furthermore, for $s = 0$ *and* $s = 1$*, the inequalities in (i) and (ii) hold without the terms proportional to* ϵ*.*

3. Consequences of Theorem 2.2

We will use the notation of Theorems 2.2 and 2.1. The purpose of this section is the

Proof of Theorem 1.7. Set $\ell = \lfloor \beta^{-p_1} \rfloor$. By Theorem 2.1, Proposition 5.5 and Lemma 5.6, the probability of the event that conditions (2.2)–(5.30), $\|(P_E - P_E^{\mathbb{T}})\| < e^{-c\mathcal{L}}$, and $|P_E(H^{\mathbb{T}})(x,y)| < e^{-c|x-y|}$ for all x, y hold on the scale ℓ is $\geq 1 - e^{-c\sqrt{\ell}}$. Thus, by Borel-Cantelli's lemma, for almost all random configurations $\omega \in \Omega$, there exists β_0 such that the event holds for all $\beta < \beta_o$. Furthermore, for a.e. ω, E is not an eigenvalue of H thanks to Lemma 5.1. From now on we will fix the configuration ω for which all this conditions are satisfied and will assume that β is below the corresponding threshold value β_o.

A condition

$$e^{-1/\beta^{\frac{p_1}{2}}} < \epsilon < \beta^{p_1 p_2}$$

implies that $e^{-c\sqrt{\ell}} < \epsilon < \ell^{-p_2}$. Hence for β sufficiently small, the conclusions Theorem 2.2.(i)–2.2.(iii) of Theorem 2.2 are satisfied.

We note that using Proposition B.3, $\|(P_E - P_E^{\mathbb{T}})\| < e^{-c\mathcal{L}}$, and $|P_E(H^{\mathbb{T}})(x,y)| < e^{-c|x-y|}$ we have

$$\left\| \left(U_\epsilon(s,0) P_E U_\epsilon(0,s) - U_\epsilon^{\mathbb{T}}(s,0) P_E^{\mathbb{T}} U_\epsilon^{\mathbb{T}}(0,s) \right) \chi_{\Lambda_{\mathcal{L}/3}} \right\| \leq C e^{-c\mathcal{L}}. \tag{3.1}$$

where in these steps we have used Proposition B.3, (5.8), (5.9) (that now hold determenistically by the virtue of our assumptions on ω and β), and Proposition B.3 one more time, respectively.

We will also need the following assertion, which is of the independent interest.

Lemma 3.1. *Let $E_\pm \in J_{loc}$ be such that $|E_+ - E_-| \le \delta$. Let $P_\pm(s) := P^{\mathbb{T}}_{E_\pm}(s)$. Then*

$$\left\|\bar{P}_+(0))U^{\mathbb{T}}_\epsilon(0,s)P_-(s))\chi_{\Lambda_L}\right\| \le D_n\left(\epsilon^n + e^{-1/\beta^p}\right), \tag{3.2}$$

where D_n is some n-dependent coefficient depending only on C, C_k in Assumption 1.2.

Proof of Lemma 3.1. Using Theorem 2.2, we have

$$\|\bar{P}_+(0)(U^{\mathbb{T}})^*_\epsilon(s)P_-(s) - \bar{P}_+(0)\bar{\mathcal{Q}}(0)(U^{\mathbb{T}})^*_\epsilon(s)\mathcal{Q}(s)P_-(s)\| \le Ce^{-1/\beta^p},$$

and by the same proposition

$$\left\|\bar{P}_+(0)\bar{\mathcal{Q}}(0)(U^{\mathbb{T}})^*_\epsilon(s)\mathcal{Q}(s)P_-(s)\right\| \le C_N\epsilon^N + e^{-1/\beta^p}.$$

□

We are now ready to complete the proof. We first observe that it suffices to establish that

$$\lim_{\epsilon,\beta\to 0}\|(U_\epsilon(s,0)P_E(0)U_\epsilon(0,s) - P_E(s))\,\chi_{\Lambda_L}\| = 0. \tag{3.3}$$

Using (3.1), we deduce that it suffices to show that

$$\lim_{\epsilon,\beta\to 0}\left\|\left(U^{\mathbb{T}}_\epsilon(s,0)P_E(0)U^{\mathbb{T}}_\epsilon(0,s) - P^{\mathbb{T}}_E(s)\right)\chi_{\Lambda_L}\right\| = 0. \tag{3.4}$$

We next decompose

$$\begin{aligned} U^{\mathbb{T}}_\epsilon(s,0)P_E(0)U^{\mathbb{T}}_\epsilon(0,s) - P^{\mathbb{T}}_E(s) &= U^{\mathbb{T}}_\epsilon(s,0)P_E(0)U^{\mathbb{T}}_\epsilon(0,s)\bar{P}^{\mathbb{T}}_{E+\delta}(s) \\ &\quad - U^{\mathbb{T}}_\epsilon(s,0)\bar{P}_E(0)U^{\mathbb{T}}_\epsilon(0,s)P^{\mathbb{T}}_{E-\delta}(s) \\ &\quad + \left(U^{\mathbb{T}}_\epsilon(s,0)P_E(0)U^{\mathbb{T}}_\epsilon(0,s) - P^{\mathbb{T}}_E(s)\right)P^{\mathbb{T}}_{(E-\delta,E+\delta)}(s), \end{aligned}$$

and bound the first term using Lemma 3.1. To bound the last one, we note that $\left\{H^{\mathbb{T}}(s)\right\}$ (extended as an operator on $\ell^2(\mathbb{Z}^d)$) converges to $H(0)$ in the strong resolvent sense as $\epsilon \to 0$. Hence for any interval (a,b), we have

$$\text{s-lim}_{\epsilon,\beta\to 0}P^{\mathbb{T}}_{(a,b)}(s) = P_{(a,b)}(0),$$

[RS, Theorem VIII.24]. Hence $P^{\mathbb{T}}_{(E-\delta,E+\delta)}(s) \overset{SOT}{\to} P_{\{E\}}(0) = 0$, and (3.4) follows.

□

Remark 3.2. For $\epsilon \ge \beta$ Theorem 1.7 follows from the result in [ES] (where it si proven for $\epsilon = \beta$, but the argument is still valid for $\epsilon \ge \beta$). Here we are focused on the regime $\epsilon \ll \beta$. We did not attempt to extend the theorem to the remaining interval $\beta \ge \epsilon \ge \beta^{p_1p_2}$.

4. Adiabatic theory for localized spectral patches

Throughout this section we will work on the torus, in the setting of Theorem 2.1. To simplify the notation, we will shorthand $H(s) := H^{\mathbb{T}}(s)$ in this section. In addition, we will use the assumption of Theorem 2.2), namely that

$$e^{-c\sqrt{\ell}} \le \epsilon \le \ell^{-p_2}, \quad p_2 > \max\{d + \frac{1}{2} + \frac{d}{q}, 2\frac{d}{q}\}.$$

This, in particular, implies that for ℓ large enough $\epsilon^{-1}e^{-c\sqrt{\ell}} \le e^{-c\sqrt{\ell}}$ and that $\epsilon/\Delta \ll 1$. We will use this repeatedly. We will also assume that $1 \ge \Delta \ge \beta > 0$ (in fact, the above conditions imply $\Delta \gg \beta$ for ℓ large, but it will only matter later on).

4.1. **Kato's operator.** Let $1 \geq \Delta \geq \beta > 0$ and let $H(s)$ be a smooth family of self-adjoint operators on $[0,1]$ such that

Assumption 4.1. *(a)* $\|H(s)\| \leq C$ and $\left\|H^{(k)}(s)\right\| \leq \beta C_k$ for $k \in \mathbb{N}$, where $H^{(k)}(s)$ stands for the k-th derivative of $H(s)$ with respect to the s variable;

(b) There exist $E_{1,2} \in \mathbb{R}$ and $\Delta > 0$ such that $\min_{s\in[0,1]} \operatorname{dist}(\sigma(H(s)), \{E_1, E_2\}) \geq 2\Delta$;

(c) $H^{(k)}(s) = 0$ for $s = \{0,1\}$ and $k \in \mathbb{N}$.

Throughout this section, we will denote by $P(s)$ the spectral projection of $H(s)$ onto the interval $[E_1, E_2]$ and will use the shorthand $R_z(s)$ for $(H(s) - z)^{-1}$. For an operator A (which can be s-dependent) we define the operator $X_A(s)$ by

$$X_A(s) = \frac{1}{2\pi} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} R_{ix+E_j}(s)\, A\, R_{ix+E_j}(s)\, dx. \tag{4.1}$$

It has been introduced by Kato in his work on adiabatic theorem, and henceforth we will refer to it as Kato's operator.

We note that, for $H(s)$ satisfying Assumption 4.1,

$$\max_{j=1,2} \left\|R_{ix+E_j}(s)\right\| \leq \left(x^2 + \Delta^2\right)^{-1/2} \tag{4.2}$$

and consequently

$$\|X_A(s)\| \leq \frac{\|A\|}{\pi} \int_{-\infty}^{\infty} \left(x^2 + \Delta^2\right)^{-1} dx \leq \Delta^{-1}\|A\|. \tag{4.3}$$

Using the Leibniz rule and (B.2), it is straightforward to see that, more generally,

$$\left\|X_A^{(k)}(s)\right\| \leq C_k \|A\|_k, \quad k \in \mathbb{Z}_+, \tag{4.4}$$

where $\|\cdot\|_k$ denotes the Sobolev-type norm

$$\|A\|_k = \sum_{j=0}^{k} \left\|A^{(j)}(s)\right\|. \tag{4.5}$$

The importance of the Kato's operator is related to the fact that it solves the commutator equation

$$[H(s), X_A(s)] = [P(s), A], \tag{4.6}$$

which plays a role in a construction of adiabatic theory for gapped Hamiltonians, in particular in the Nenciu's expansion presented below.

To handle the adiabatic behavior of localized spectral patches we will also need to understand the locality properties of the Kato's operator.

Lemma 4.2. *Let $A(s)$ be a smooth family of operators on $[0,1]$. Suppose that in addition to Assumption 4.1, there exists some set $\mathcal{A}$ and $M, c > 0$ such that*

$$\left\|R_{ix+E_j}(s)\right\|_{c,\ell} \leq M\langle x\rangle^{-1}, \quad j = 1, 2. \tag{4.7}$$

Then

$$\left\|e^{c\rho^\ell_{\mathcal{A}}} X^{(1)}_{A(s)}(s)\right\| \leq C\left(\beta M^2 |\ln\Delta| + \beta M \Delta^{-1}\right) \left\|e^{c\rho^\ell_{\mathcal{A}}} A(s)\right\| + CM|\ln\Delta| \left\|e^{c\rho^\ell_{\mathcal{A}}} A^{(1)}(s)\right\|. \tag{4.8}$$

Proof. We will be suppress the s-dependence in the proof. Using (B.2) and (1.21), we can bound

$$\begin{aligned}\left\|e^{c\rho^\ell_{\mathcal{A}}} X^{(1)}_A\right\| \leq \sum_{j=1}^{2} \Bigg(&\frac{C\beta}{\pi} \int_{-\infty}^{\infty} \left\|R_{ix+E_j}\right\|^2_{c,\ell} \left\|e^{c\rho^\ell_{\mathcal{A}}} A\right\| \left\|R_{ix+E_j}\right\| dx \\ &+ \frac{1}{\pi} \int_{-\infty}^{\infty} \left\|R_{ix+E_j}\right\|_{c,\ell} \left\|e^{c\rho^\ell_{\mathcal{A}}} A^{(1)}\right\| \left\|R_{ix+E_j}\right\| dx \\ &+ \frac{C\beta}{\pi} \int_{-\infty}^{\infty} \left\|R_{ix+E_j}\right\|_{c,\ell} \left\|e^{c\rho^\ell_{\mathcal{A}}} A\right\| \left\|R_{ix+E_j}\right\|^2 dx \Bigg).\end{aligned}$$

Using (4.7) and Assumption 4.1.(b), we get (4.8). □

4.2. **Nenciu's expansion.** An elegant approach for the analysis of the adiabatic behavior of gapped systems was discovered by Nenciu [N]. We will use it as a starting point in our construction.

Lemma 4.3 (Nenciu's expansion)**.** *Let $H(s)$ be a smooth family of self-adjoint operators on $[0,1]$ that satisfies Assumption 4.1. Let $B_n(s)$ be a smooth family defined recursively as follows: $B_0(s)=P(s)$ and, for $n\in\mathbb{N}$,*

$$B_n(s)=\Big(\bar{P}(s)X_{\dot{B}_{n-1}(s)}(s)P(s)+h.c.\Big)+S_n(s)-2P(s)S_n(s)P(s), \tag{4.9}$$

where

$$S_n(s)=\sum_{j=1}^{n-1}B_j(s)B_{n-j}(s). \tag{4.10}$$

Then we have

(i)

$$\dot{B}_n(s)=-i\left[H(s),B_{n+1}(s)\right] \tag{4.11}$$

for all $n\in\mathbb{Z}_+$;

(ii) $B_n(s)=0$ for $s=\{0,1\}$ and $n\in\mathbb{N}$;

(iii) We have

$$\sup_s\left\|B_n^{(k)}(s)\right\|\le C_{n,k}\Delta^{-n},\quad k,n\in\mathbb{Z}_+. \tag{4.12}$$

Proof. Property 4.3.(i) is due to Nenciu, [N]. Property 4.3.(ii) follows directly from the recursive definition of B's. We establish 4.3.(iii) by induction:

Induction base: For $n=0$ and arbitrary k, the bound $\left\|B_0^{(k)}(s)\right\|\le C_k$ in 4.3.(iii) can be seen from (B.1), (B.2), Assumption (a), and the Leibniz rule.

Induction step: Suppose now that the statement holds for all $n<n_o$ and all $k\in\mathbb{Z}_+$. Differentiating (4.9) k times with $n=n_o$ using the Leibniz rule and then using (4.2) and (4.4), we get that it also holds for $n=n_o$ and all $k\in\mathbb{Z}_+$. □

For localized spectral patches we modify the statement slightly.

Lemma 4.4. *Suppose that in addition to the assumptions of Lemma 4.3, there exists some set $\mathcal{A}$ and $M,c>0$ such that (4.7) holds. Let us also assume that*

$$\max_{s\in[0,1]}\left\|\mathrm{e}^{c\,\rho_{\mathcal{A}}^{\ell}}P(s)\right\|\le C,\quad \max_{s\in[0,1]}\left\|H^{(k)}(s)\right\|_{c,\ell}\le C_k\beta \text{ for } k\in\mathbb{N}. \tag{4.13}$$

Let

$$\nu=\min\Big(M^{-1}\left|\ln\Delta\right|^{-1},\Delta\Big),$$

and assume that $\beta\le\nu$. Then the operators B_n defined in Lemma 4.3 satisfy

$$\left\|\mathrm{e}^{c\rho_{\mathcal{A}}^{\ell}}B_n^{(k)}(s)\right\|\le C_{n,k}\nu^{-n},\quad k,n\in\mathbb{Z}_+. \tag{4.14}$$

Proof. We will suppress the s-dependence in the proof and use induction in n and k.

Induction base: For $n=0$ and arbitrary k, by the Leibniz rule we have

$$P^{(n)}=\left(P^{n+1}\right)^{(n)}=\sum_{k_1+k_2+\cdots+k_{n+1}=n}\binom{n}{k_1,k_2,\ldots,k_{n+1}}\prod_{1\le j\le n+1}P^{(k_j)}, \tag{4.15}$$

where the sum extends over all m-tuples $(k_1,\ldots,k_{n+1})$ of non-negative integers satisfying $\sum_{j=1}^{n+1}k_j=n$ (so at least for one value of j we have $k_j=0$).

Using the integral representation (B.1), the formula (B.2), the Leibniz rule, (4.7), (1.21), and Assumption (4.13), we can bound

$$\left\|P^{(k)}\right\|_{c,\ell}\le C_kM^k,\quad k\in\mathbb{N}.$$

We can now use (1.21) and (4.15) to deduce that

$$\left\|e^{c\,\rho^\ell_{\mathcal{A}}} P^{(n)}\right\|$$
$$\leq \sum_{k_1+k_2+\cdots+k_{n+1}=n} \binom{n}{k_1,k_2,\ldots,k_{n+1}} \prod_{1\leq j<j_o} \left\|P^{(k_j)}\right\|_{c,\ell} \left\|e^{c\,\rho^\ell_{\mathcal{A}}} P\right\| \prod_{j_o\leq j\leq n+1} \left\|P^{(k_j)}\right\|$$
$$\leq \sum_{k_1+k_2+\cdots+k_{n+1}=n} \binom{n}{k_1,k_2,\ldots,k_{n+1}} \prod_{1\leq j\leq n+1} C_{k_j} M^{k_j} = C_n M^n, \quad (4.16)$$

where j_o is the first value of the index j for which $k_j = 0$.

Induction step: Suppose now that the assertion holds for all $n < n_o$ and all k. Differentiating (4.9) k times with $n = n_o$ using the Leibniz rule and then using Lemma 4.2 (the assumption there is satisfied by Eq. (5.30)), we get the induction step. □

4.3. **Gapped adiabatic theorem.** An immediate consequence of Lemma 4.3 is

Lemma 4.5 (Gapped adiabatic theorem to all orders)**.** *In the setting of Lemma 4.3, let $P_N(s) := \sum_{n=0}^N \epsilon^n B_n(s)$. Then for all $N \in \mathbb{N}$,*

$$\|U_\epsilon(s)P(0)U_\epsilon(s)^* - P_N(s)\| \leq C_N \epsilon^N \Delta^{-N},$$

where U_ϵ was defined in (1.3)*.*

In particular, for $\epsilon < \Delta$, we have

$$\|U_\epsilon(s)P(0)U_\epsilon(s)^* - P(s)\| \leq C\epsilon\Delta^{-1}$$

and

$$\|U_\epsilon(1)P(0)U_\epsilon(1)^* - P(1)\| \leq C_N \epsilon^N \Delta^{-N}.$$

Proof. By Lemma 4.3,

$$\epsilon \dot{P}_N(s) = -i[H(s), P_N(s)] + \epsilon^{N+1} \dot{B}_N(s).$$

Using the fundamental theorem of calculus, we obtain

$$U_\epsilon(s)^* P_N(s) U_\epsilon(s) - P_N(0) = \epsilon^{-1} \int_0^s \epsilon^{N+1} \frac{d}{ds} \left(U_\epsilon(s)^* B_N(s) U_\epsilon(s)\right).$$

Using the unitarity of U_ϵ, Assumption 4.1, and Lemma 4.3.(iii), we obtain

$$\|U_\epsilon(s)^* P_N(s) U_\epsilon(s) - P_N(0)\| \leq C_N \epsilon^N \Delta^{-N}.$$

The assertion follows from $P_N(0) = P(0)$, $\|P_N(s) - P(s)\| \leq C\epsilon\Delta^{-1}$, and $P_N(1) = P(1)$. □

4.4. **Adiabatic theorem for a localized spectral patch.** The goal of this subsection is to prove the following assertion, which is of the independent interest.

Theorem 4.6 (Locabatic theorem on a torus)**.** *Suppose that the family $H(s)$ satisfies Assumption 1.2 and $H(0)$ satisfies Assumptions 1.3–1.4. Let $\mathcal{G}_\omega$ be an event $H^{\mathbb{T}}(0)$ possesses an ultralocal structure for the energy interval $J = (E-6\delta, E+6\delta)$. Then $\mathbb{P}(\mathcal{G}_\omega) > 1 - e^{-c\sqrt{\ell}}$. Moreover, for each $\omega \in \mathcal{G}_\omega$, the physical evolution $\psi_\epsilon(s)$ of each eigenvector $\psi = \psi_n$ with $E_n \in J$, given by* (1.2)*, satisfies*

$$\max_{s\in[0,1]} \left\|\bar{P}_{J_\gamma}\big(H^{\mathcal{T}_\gamma}(s)\big)\,\psi_\epsilon(s)\right\| \leq C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right) \quad (4.17)$$

for some γ. Furthermore, for any $N \in \mathbb{N}$, we can further improve (4.17) *for $s = 1$:*

$$\left\|\bar{P}_{J_\gamma}\big(H^{\mathcal{T}_\gamma}(1)\big)\,\psi_\epsilon(1)\right\| \leq C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \quad (4.18)$$

Proof of Theorem 4.6. We already established the first part of the assertion in Theorem 2.1. We now show the second part. We first note that $\mathcal{G} \subset \Omega_{loc,N}$ of the full configuration space for which $\mathbb{T}$ and all sets in $\{\mathcal{T}_\gamma\}$ are $\ell/10$-localizing, see Lemma 5.11 below. Thus Theorem 2.1.(ii) implies the existence of the patch $\mathcal{T}_\gamma$ such that $\left\|\bar{\chi}_{(\mathcal{T}_\gamma)_{8\ell}}\psi\right\| \le e^{-c\sqrt{\ell}}$. It then follows from Lemma B.2 below, specifically (B.6), that $E \in J_\gamma$ (see also (2.2)). Let $\hat{\mathcal{T}}_\gamma = (\mathcal{T}_\gamma)_{4\ell}$ and set

$$Q_\gamma(s) = \chi_{\hat{\mathcal{T}}_\gamma} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\hat{\mathcal{T}}_\gamma}. \tag{4.19}$$

By Lemma B.2, specifically (B.7), we know that (4.18) holds for $s = 0$ (with $\epsilon = 0$ on the right hand side). Let $\rho := Q_\gamma(0)$ be the (truncated) initial spectral patch. Then, since

$$\bar{\rho} = \chi_{\hat{\mathcal{T}}}\bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}(0))\chi_{\hat{\mathcal{T}}} + \bar{\chi}_{\hat{\mathcal{T}}},$$

we deduce that $\|\bar{\rho}\psi\| \le e^{-c\sqrt{\ell}}$. Hence by unitarity of the quantum evolution,

$$\|\bar{\rho}_\epsilon(s)\psi_\epsilon(s)\| \le e^{-c\sqrt{\ell}} \tag{4.20}$$

for all s, where ρ_ϵ denotes the (full) Heisenberg evolution of the (truncated) initial spectral patch $\rho := Q_\gamma(0)$, i.e.,

$$i\epsilon\dot{\rho}_\epsilon(s) = [H(s), \rho_\epsilon(s)], \quad \rho_\epsilon(0) = \rho. \tag{4.21}$$

Therefore the result follows from

Lemma 4.7. *(i) We can estimate*

$$\max_{s\in[0,1]} \|\rho_\epsilon(s) - Q_\gamma(s)\| \le C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right). \tag{4.22}$$

Moreover, for any $N \in \mathbb{N}$, we have

$$\max_{s=\{0,1\}} \|\rho_\epsilon(s) - Q_\gamma(s)\| \le C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \tag{4.23}$$

(ii) In addition,

$$\max_{s\in[0,1]} \left\|\bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\,\bar{Q}_\gamma(s)\right\| \le e^{-c\sqrt{\ell}}. \tag{4.24}$$

□

Proof of Lemma 4.7. We suppress the s dependence in the proof. The property (4.24) can be seen by decomposing

$$\bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}) = \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma})\,\bar{Q}_\gamma + \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma})\,Q_\gamma$$

and noticing that

$$\begin{aligned}\bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma})\,Q_\gamma &= \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma})\,\chi_{\hat{\mathcal{T}}_\gamma}P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\hat{\mathcal{T}}_\gamma}\\ &= \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma})\,P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\hat{\mathcal{T}}_\gamma} + O\!\left(e^{-c\sqrt{\ell}}\right) = O\!\left(e^{-c\sqrt{\ell}}\right),\end{aligned}$$

thanks to (2.4).

Lemma 4.7.(i): By our assumption, $H^{\mathcal{T}_\gamma}$ is a gapped Hamiltonian with a gap Δ. Following the argument in Section 4.2, we set B_n^γ the n-th order in the Nenciu's expansion. Explicitly, we use Lemma 4.3 with $B_0^\gamma = P_{J_\gamma}(H^{\mathcal{T}_\gamma})$. We set

$$Q_{\gamma,N} := \sum_{n=0}^{N} \epsilon^n \chi_{\hat{\mathcal{T}}} B_n^\gamma \chi_{\hat{\mathcal{T}}}. \tag{4.25}$$

and proceed to show that

$$\max_s \|\rho_\epsilon - Q_{\gamma,N}\| \le C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \tag{4.26}$$

The result then follows immediately from (4.26) by definition of $Q_{\gamma,N}$ and Lemma 4.3.(ii)–4.3.(iii) (we recall that $B_0^\gamma = P_{J_\gamma}(H^{\mathcal{T}_\gamma})$).

To get (4.26), we observe that by (4.27),

$$\begin{aligned}\epsilon\dot{Q}_{\gamma,N} &= -i\sum_{\gamma}\sum_{n=0}^{N}\epsilon^{n+1}\chi_{\hat{\mathcal{T}}}\left[H^{\mathcal{T}_\gamma}, B_{n+1}^{\gamma}\right]\chi_{\hat{\mathcal{T}}} \\ &= -i[H, Q_{\gamma,N}] - i\epsilon^{N+1}\chi_{\hat{\mathcal{T}}}\dot{B}_N^{\gamma}\chi_{\hat{\mathcal{T}}} \\ &\quad + \left(i\sum_{\gamma}\sum_{n=0}^{N}\epsilon^{n+1}\left[H^{\mathcal{T}_\gamma}, \chi_{\hat{\mathcal{T}}}\right]B_{n+1}^{\gamma}\chi_{\hat{\mathcal{T}}} + h.c.\right),\end{aligned}$$

where we have used $H^{\mathcal{T}}(s)\chi_{\hat{\mathcal{T}}} = H(s)\chi_{\hat{\mathcal{T}}}$. We bound the second term on the second line by $C_N\epsilon^{N+1}\Delta^{-N}$ using (4.12). For the term on the third line, we note that

$$\left\|\left[H^{\mathcal{T}_\gamma}(s), \chi_{\hat{\mathcal{T}}}\right]B_{n+1}^{\gamma}(s)\right\| \le \nu^{-n-1}\mathrm{e}^{-c\sqrt{\ell}}$$

using Lemma 4.4. Putting these bounds together, we get

$$\left\|\epsilon\dot{Q}_{\gamma,N} + i[H, Q_{\gamma,N}]\right\| \le C_N\epsilon^{N+1}\Delta^{-N} + Ce^{-c\sqrt{\ell}}. \tag{4.27}$$

Finally, we observe that

$$\partial_s\left(U_\epsilon(t,s)Q_{\gamma,N}(s)U_\epsilon(s,t)\right) = \epsilon^{-1}U_\epsilon(t,s)\left(\epsilon\dot{Q}_{\gamma,N}(s) + i[H(s), Q_{\gamma,N}(s)]\right)U_\epsilon(s,t).$$

where $U_\epsilon(t,s)$ was defined in (1.3).

Integrating over s and using (4.27), we deduce that

$$\|U_\epsilon(t,r)Q_{\gamma,N}(r)U_\epsilon(r,t) - Q_{\gamma,N}(t)\| \le \epsilon^{-1}\left(C_N\epsilon^{N+1}\Delta^{-N} + Ce^{-c\sqrt{\ell}}\right), \tag{4.28}$$

We now note that $Q_{\gamma,N}(0) = \rho$, so $U_\epsilon(t,0)Q_{\gamma,N}(0)U_\epsilon(0,t) = \rho_\epsilon(t)$ by uniqueness of a solution for the IVP (4.21). Combining it with (4.28) yields (4.26). □

4.5. **Adiabatic theorem for a thin spectral set near E.** In preparation for the proof of Theorem 2.2, we will investigate first the adiabatic behavior of a spectral data corresponding to a thin set of non-trivial thickness that contains energy E. It will play a role of a natural barrier that suppresses transitions between the spectral data below and above E which will make Theorem 2.2 works. The idea here is to combine localized spectral patches near E that we analyzed in the previous subsection into such a set. Specifically, we define

$$Q(s) := \sum_{\gamma} Q_\gamma(s), \tag{4.29}$$

where the spectral patch Q_γ was defined in (4.19). Our first assertion encapsulates the basic property of this operator.

Lemma 4.8. *For ℓ large enough, the operator $Q(s)$ satisfies:*

(i) If $H(s)$ is k times differentiable, so is $Q(s)$:

$$\max_{s\in[0,1]}\left\|\frac{d^jQ(s)}{d^js}\right\| \le C_j\beta, \quad j = 1, \ldots, k;$$

(ii) Near commutativity with $H(s)$:

$$\|[H(s), Q(s)]\| \le Ce^{-c\sqrt{\ell}}; \tag{4.30}$$

(iii) Almost projection:

$$\left\|\bar{Q}(s)Q(s)\right\| \le Ce^{-c\sqrt{\ell}}; \tag{4.31}$$

(iv) Spectrally thin but with non trivial thickness: Let $J_+ = (E - 6\delta, E + 6\delta)$, and $J_- = (E - \delta, E + \delta)$. Then

$$\left\|\bar{P}_{J_+}(s)Q(s)\right\| \le Ce^{-c\sqrt{\ell}}, \quad \left\|\bar{Q}(s)P_{J_-}(s)\right\| \le Ce^{-c\sqrt{\ell}}. \tag{4.32}$$

Proof. Lemma 4.8.(i): Note that for ℓ large enough, $\beta \ll \Delta$. The assertion follows from the integral representation (B.1) for $P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))$ with $E_{1,2} = E_\pm^\gamma$, the formula (B.2), (4.2), and the Leibniz rule.

Lemma 4.8.(ii): We compute

$$\begin{aligned}[H(s), Q_\gamma(s)] &= \left[H^{\mathcal{T}_\gamma}(s), Q_\gamma(s)\right] \\ &= \left[H^{\mathcal{T}_\gamma}(s), \chi_{\hat{\mathcal{T}}}\right] P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\hat{\mathcal{T}}} + \chi_{\hat{\mathcal{T}}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \left[H^{\mathcal{T}_\gamma}(s), \chi_{\hat{\mathcal{T}}}\right],\end{aligned}$$

and estimate both terms by $C\mathrm{e}^{-c\sqrt{\ell}}$ using Assumption 1.2 and Theorem 2.1.(ii).

Lemma 4.8.(iii): We note that for disjoint sets Ω_γ,

$$\|\sum_\gamma \chi_{\Omega_\gamma} A_\gamma \chi_{\Omega_\gamma}\| \le \max_\gamma \|\chi_{\Omega_\gamma} A_\gamma \chi_{\Omega_\gamma}\|. \tag{4.33}$$

Since $\mathcal{T}_\gamma$ are disjoint, we have

$$\left\|\bar{Q}(s)Q(s)\right\| = \left\|\sum_\gamma \chi_{\hat{\mathcal{T}}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\bar{\chi}_{\hat{\mathcal{T}}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\hat{\mathcal{T}}}\right\|.$$

The right hand side is bounded by $Ce^{-c\sqrt{\ell}}$ using Theorem 2.1.(ii).

Lemma 4.8.(iv): We apply Lemma B.1 with $H_1 = H(s)$, $H_2 = H^{\mathcal{T}}(s)$, and $R = \chi_{\hat{\mathcal{T}}}$ to bound

$$\left\|\bar{P}_{J_+}(s)\chi_{\hat{\mathcal{T}}} P_J(H^{\mathcal{T}}(s))\right\| \le Ce^{-c\sqrt{\ell}},$$

where we have used (2.4) and the fact that $H(s)$ has range r. Since

$$Q(s) \le \chi_{\hat{\mathcal{T}}} P_J(H^{\mathcal{T}}(s))\chi_{\hat{\mathcal{T}}}$$

by (2.2), we deduce that

$$\left\|\bar{P}_{J_+}(s)Q(s)\right\| \le \left\|\bar{P}_{J_+}(s)\chi_{\hat{\mathcal{T}}} P_J(H^{\mathcal{T}}(s))\right\| \le Ce^{-c\sqrt{\ell}}.$$

On the other hand, letting $J' = (E - 3\delta, E + 3\delta)$ and using Lemma B.1 with $H_1 = H^{\mathcal{T}}(s)$ and $H_2 = H(s)$, we get

$$\left\|\bar{P}_{J'}(H^{\mathcal{T}}(s))\chi_{\hat{\mathcal{T}}} P_{J_-}(s)\right\| \le Ce^{-c\sqrt{\ell}}$$

Since

$$\bar{Q}(s) \le \chi_{\Lambda\setminus\hat{\mathcal{T}}} + \chi_{\hat{\mathcal{T}}} \bar{P}_{J'}(H^{\mathcal{T}}(s))\chi_{\hat{\mathcal{T}}}$$

by (2.2), we deduce that

$$\left\|\bar{Q}(s)P_{J_-}(s)\right\| \le \left\|\chi_{\Lambda\setminus\hat{\mathcal{T}}} P_{J_-}(s)\right\| + \left\|\bar{P}_{J_+}(s)\chi_{\hat{\mathcal{T}}} P_{J_-}(s)\right\| \le Ce^{-c\sqrt{\ell}}$$

using (2.3) to bound the first term on the right hand side.

□

One disadvantage of working with Q is due to the fact that it is not a projection. We rectify this problem in the next assertion.

Lemma 4.9. *Let $N \in \mathbb{N}$. Suppose that ℓ large enough, then there exists a smooth family of projections Q_s with the following properties:*

(i)

$$\max_{s\in[0,1]} \|[Q_s, H(s)]\| \le C\left(\epsilon + e^{-c\sqrt{\ell}}\right) \tag{4.34}$$

and

$$\max_{s\in\{0,1\}} \|[Q_s, H(s)]\| \le C_N \epsilon^{N+1}\Delta^{-N} + Ce^{-c\sqrt{\ell}}; \tag{4.35}$$

(ii) Let $J_+ = (E-6\delta, E+6\delta)$*, and* $J_- = (E-\delta, E+\delta)$*. Then*

$$\max_{s\in[0,1]} \left(\left\|\bar{P}_{J_+}(s)Q_s\right\|, \left\|\bar{Q}_s P_{J_-}(s)\right\|\right) \le C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right) \tag{4.36}$$

and

$$\max_{s\in\{0,1\}} \left(\left\|\bar{P}_{J_+}(s)Q_s\right\|, \left\|\bar{Q}_s P_{J_-}(s)\right\|\right) \le Ce^{-c\sqrt{\ell}} \tag{4.37}$$

(iii) $Q_0^{(k)} = Q_1^{(k)} = 0$ *for all* $k \in \mathbb{Z}_+$*, and*

$$\max_{s\in[0,1]} \left\|Q_s^{(k)}\right\| \le C_k\beta, \quad k \in \mathbb{N};$$

(iv)

$$\left\|\epsilon\dot{Q}_s + i[H(s), Q_s]\right\| \le C_N\epsilon^{N+1}\Delta^{-N} + Ce^{-c\sqrt{\ell}}; \tag{4.38}$$

(v) If we denote by $Q_\epsilon(s)$ *the solution of IVP* $i\epsilon\dot{Q}_\epsilon(s) = [H(s), Q_\epsilon(s)]$*,* $Q_\epsilon(0) = Q_0$*, then we have*

$$\max_{s\in[0,1]} \|Q_\epsilon(s) - Q_s\| \le C_N\epsilon^N\Delta^{-N} + Ce^{-c\sqrt{\ell}}. \tag{4.39}$$

Proof. We set

$$Q_N(s) := \sum_\gamma Q_{\gamma,N}(s), \tag{4.40}$$

where was defined in (4.25), and first show the assertions of the lemma holds if we replace Q_s there with $Q_N(s)$. Note that the latter operator is not a projection.

It follows from Lemma 4.3 and the hypothesis $\epsilon \le \Delta$ that

$$\|Q_N(s) - Q_0(s)\| = \|Q_N(s) - Q(s)\| \le C_N\,\epsilon\Delta^{-1}. \tag{4.41}$$

Hence, combining this bound with Lemma 4.8, we conclude that $Q_N(s)$ satisfies the properties 4.9.(ii)–4.9.(iii).

We next observe that the property 4.9.(iv) holds for $Q_N(s)$ by (4.27), Assumption 1.2, and (4.33).

The property 4.9.(v) is established by replicating the argument employed in the proof of Lemma 4.7.(i).

Finally, the property 4.9.(i) holds for $Q_N(s)$ by the properties 4.9.(iii)–4.9.(iv) we already established.

We now note that $Q_N(0) = Q(0)$. Hence, defining $Q_\epsilon(t) := U_\epsilon(t,0)Q(0)U_\epsilon(0,t)$, we get $\left\|Q_\epsilon(t)\bar{Q}_\epsilon(t)\right\| = \left\|Q(0)\bar{Q}(0)\right\| \le Ce^{-c\sqrt{\ell}}$ by (4.31). Thus by the triangle inequality, we get

$$\begin{aligned}\left\|Q_N(t)\bar{Q}_N(t)\right\| &\le \left\|Q_N(t)\bar{Q}_N(t) - Q_\epsilon(t)\bar{Q}_\epsilon(t)\right\| + Ce^{-c\sqrt{\ell}}\\ &\le \left(\left\|\bar{Q}_N(t)\right\| + \|Q_\epsilon(t)\|\right)\|Q_N(t) - Q_\epsilon(t)\| + Ce^{-c\sqrt{\ell}}\\ &\le C_N\epsilon^N\Delta^{-N} + Ce^{-c\sqrt{\ell}},\end{aligned}$$

where in the last step we have used properties 4.9.(iii) and 4.9.(v) for Q_N.

It follows that

$$\max_s \operatorname{dist}\left(\sigma\left(Q_N(s)\right), \{0,1\}\right) \le C_N\epsilon^N\Delta^{-N} + Ce^{-c\sqrt{\ell}}.$$

If ϵ/Δ is small enough and ℓ large then the right hand side is smaller than $1/4$. We set Q_s to be the spectral projection for $Q_N(s)$ onto the interval $[\frac{1}{2}, \frac{3}{2}]$. Then by the functional calculus for self adjoint operators and triangle inequality, Lemma 4.9.(i), 4.9.(ii), and 4.9.(v) hold for this operator. To establish Lemma 4.9.(iii) we use the following integral representation for Q_s:

$$Q_s = (2\pi i)^{-1}\oint_\Gamma (Q_N(s) - z)^{-1}\,dz, \quad \Gamma = \{z \in \mathbb{C}:\ |z-1| = 1/2\}. \tag{4.42}$$

Since

$$\partial_s \left(Q_N(s) - z\right)^{-1} = -\left(Q_N(s) - z\right)^{-1} \partial_s Q_N(s) \left(Q_N(s) - z\right)^{-1},$$

and $\left\|\left(Q_N(s) - z\right)^{-1}\right\|$ is uniformly bounded for $z \in \Gamma$, (4.9.(iii)) follows by the Leibniz rule and the bounds on $Q_N^{(k)}(s)$.

Lemma 4.9.(iv):

$$\begin{aligned}\dot{Q}_s &= -(2\pi i)^{-1} \oint_\Gamma \left(Q_N(s) - z\right)^{-1} \dot{Q}_N(s) \left(Q_N(s) - z\right)^{-1} dz \\ &= -i\,(2\pi i)^{-1} \oint_\Gamma \left(Q_N(s) - z\right)^{-1} [H(s), Q_N(s)] \left(Q_N(s) - z\right)^{-1} dz \\ &\quad - (2\pi i)^{-1} \oint_\Gamma \left(Q_N(s) - z\right)^{-1} \Big(\dot{Q}_N(s) - i[H(s), Q_N(s)]\Big) \left(Q_N(s) - z\right)^{-1} dz,\end{aligned}$$

and the statement follows from the properties 4.9.(iv) and 4.9.(i) that we already proved for $Q_N(s)$.

For $s \in \{0, 1\}$, we have $Q_N(s) = Q(s)$, (4.35) and (4.37) follow from Lemma 4.8.

□

4.6. **Adiabatic behavior of distorted Fermi projection.** Here we prove Theorem 2.2. The idea behind it is that since the projection Q_s evolves adiabatically, it effectively induces a gap on its spectral support and decouples the energies separated by this induced gap.

Let $\bar{H}(s) = \bar{Q}_s H(s) \bar{Q}_s$. By Lemma 4.9, $\bar{Q}_s$ is close to a spectral projection of $H(s)$ and so the spectrum of $\bar{H}(s)$ is approximately a subset of the original spectrum and the point 0. To avoid discussing the position of 0 with respect to E, we assume without loss of generality that $E < 0$. We will need a pair of preparatory results.

Lemma 4.10. *Let $I = (E - \delta/2, E + \delta/2)$. Suppose that ℓ large enough, then we have $\sigma(\bar{H}(s)) \cap I = \emptyset$ for $s \in [0, 1]$. In addition, we have*

$$\max_{s \in [0,1]} \left\|\bar{H}(s)^{(k)}\right\| \le C_k \quad \text{for} \quad k = 1, \dots, N. \tag{4.43}$$

Proof. For ℓ large enough, $0 \notin I$. Hence it is enough to show the claim when $\bar{H}(s)$ is understood as an operator on the range of $\bar{Q}_s$. Let $w \in I$, we will show that $\left(\bar{H}(s) - w\right)^2 > 0$ from which the assertion follows. To this end, we suppress the s dependence and note that

$$\begin{aligned}\left(\bar{H} - w\right)^2 = \bar{Q}\,(H - w)\,\bar{Q}\,(H - w)\,\bar{Q} &= \bar{Q}\,(H - w)^2\,\bar{Q} - \bar{Q} H Q H \bar{Q} \\ &\ge \bar{Q}\bar{P}_{J_-}\,(H - w)^2\,\bar{Q} + \bar{Q}[H, Q][H, Q]\bar{Q},\end{aligned}$$

while we can bound

$$\bar{Q}\bar{P}_{J_-}\,(H - w)^2\,\bar{Q} \ge \frac{\delta^2}{4}\bar{Q}\bar{P}_{J_-}\bar{Q} = \frac{\delta^2}{4}\bar{Q} - \frac{\delta^2}{4}\bar{Q}P_{J_-}\bar{Q} \ge \frac{\delta^2}{4}\bar{Q} - \frac{\delta^2}{4}\left(C_N\epsilon + C\exp\left(-c\sqrt{\ell}\right)\right)^2\bar{Q},$$

using Lemma 4.9 4.9.(ii), and

$$\bar{Q}[H, Q][H, Q]\bar{Q} \le \left\|[H, \bar{Q}]\right\|^2 \bar{Q} \le \left(C_N\epsilon + C\exp\left(-c\sqrt{\ell}\right)\right)^2 \bar{Q}$$

using Lemma 4.9 4.9.(i). Hence

$$\left(\bar{H} - w\right)^2 \ge \left(\delta^2/4 - 2\left(C_N\epsilon + C\exp\left(-c\sqrt{\ell}\right)\right)^2\right)\bar{Q} > 0$$

on *Range* $\bar{Q}$.

The bound (4.43) follows from Lemma 4.9.(iii), Assumption 1.2, and the Leibniz rule. □

Lemma 4.11. *Let $T(s,s')$ be the unitary semigroup generated by $i[\dot{Q}_s, Q_s]$, i.e., $T(s,s')$ is the solution of the IVP*

$$i\partial_s T(s,s') = i[\dot{Q}_s, Q_s]T(s,s'), \quad T(s',s') = 1. \tag{4.44}$$

Then $T(s,s')$ satisfies

$$T(s,s')Q_{s'} = Q_s T(s,s'). \tag{4.45}$$

Suppose in addition that ϵ/Δ is small enough and ℓ large enough. Then

$$\max_s \left\| T^{(k)}(s,0) \right\| \le C_k \beta \quad \textit{for} \quad k = 1, \dots, N. \tag{4.46}$$

Proof. The interweaving relation (4.45) follows from observing that

$$\frac{d}{ds}\left(T(s',s)Q_s T(s,s')\right) = T(s',s)\left[Q_s, [\dot{Q}_s, Q_s]\right] T(s,s') + T(s',s)\dot{Q}_s T(s,s') = 0,$$

and $T(s',s')Q_{s'}T(s',s') = Q_{s'}$.

The bound (4.46) follows from Lemma 4.9.(iii), the unitarity of T, and the Leibniz rule. □

We now consider the evolution $U_\epsilon(s,s')$ generated by the equation

$$i\epsilon\partial_s U_\epsilon(s,s') = H(s)U_\epsilon(s,s'), \quad U_\epsilon(s',s') = 1.$$

Let Q_s^+ (Q_s^-) be the spectral projection of $\bar{H}_s$ associated with the interval (E,∞) (respectively $(-\infty, E)$).

Lemma 4.12. *Suppose that ℓ is large enough. Then we have*

$$\max_s \left\| Q_1^+ U_\epsilon(s,0) Q_0^- \right\| \le C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right) \tag{4.47}$$

and

$$\left\| Q_1^+ U_\epsilon(1,0) Q_0^- \right\| \le C_N\left(\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-2N-1}\right) + Ce^{-c\sqrt{\ell}}. \tag{4.48}$$

Proof. We first note that Lemma 4.9 implies that

$$\left\| Q_s U_\epsilon(s,s') \bar{Q}_{s'} \right\| \le C_N \epsilon^N \Delta^{-N} + Ce^{-c\sqrt{\ell}}. \tag{4.49}$$

Indeed, using the semigroup property for U_ϵ,

$$Q_s U_\epsilon(s,s')\bar{Q}_{s'} = Q_s(Q_s - Q_\epsilon(s))U_\epsilon(s,s') - Q_s U_\epsilon(s,s')(Q_{s'} - Q_\epsilon(s')),$$

and both terms on the right hand side can be now bounded using Lemma 4.9.(v).

Let $V_\epsilon(s) = \bar{Q}_s U_\epsilon(s,0)\bar{Q}_0$. Then a straightforward computation yields

$$\begin{aligned} i\epsilon\partial_s V_\epsilon(s) &= -i\epsilon\dot{Q}_s U_\epsilon(s,0)\bar{Q}_0 + \bar{Q}_s H(s) U_\epsilon(s,0)\bar{Q}_0 \\ &= i\epsilon[\dot{Q}_s, Q_s]V_\epsilon(s) + \bar{H}(s)V_\epsilon(s) + R_\epsilon(s), \end{aligned}$$

where

$$R_\epsilon(s) = -i\epsilon\dot{Q}_s Q_s U_\epsilon(s,0)\bar{Q}_0 + \bar{Q}_s H(s) Q_s U_\epsilon(s,0)\bar{Q}_0.$$

We note that

$$\|R_\epsilon(s)\| \le \left(\epsilon\left\|\dot{Q}_s\right\| + \|[H(s),Q_s]\|\right)\left\|Q_s U_\epsilon(s,0)\bar{Q}_0\right\| \le C_N \epsilon\epsilon^N \Delta^{-N} + Ce^{-c\sqrt{\ell}}, \tag{4.50}$$

by Lemma 4.9 and (4.49).

Let $W_\epsilon(s) = T(0,s)V_\epsilon(s)$, where T was defined in (4.44). Then

$$i\epsilon\partial_s W_\epsilon(s) = T(0,s)\bar{H}(s)T(s,0)W_\epsilon(s) + T(0,s)R_\epsilon(s).$$

By Lemma 4.10, the operator $\bar{H}(s)$ has a gap, δ, in its spectrum that separates the associated spectral projections $Q_s^\pm$. This implies that $T(0,s)\bar{H}(s)T(s,0)$ has the same gap with the associated projections given by $\mathcal{Q}_s^\pm := T(0,s)Q_s^\pm T(s,0)$. We can bound

$$\left\| \left(T(0,s)\bar{H}(s)T(s,0)\right)^{(k)} \right\| \le C_k\beta \quad \text{for} \quad k = 1,\dots,N,$$

using (4.43), (4.46), and the Leibniz rule.

Let $\tilde{W}_\epsilon(s)$ denote the evolution generated by $T(0,s)\bar{H}_s T(s,0)$:

$$i\epsilon\partial_s \tilde{W}_\epsilon(s) = T(0,s)\bar{H}(s)T(s,0)\tilde{W}_\epsilon(s), \quad \tilde{W}_\epsilon(0) = 1. \tag{4.51}$$

Then it follows from our previous analysis and the Leibniz rule that $T(0,s)\bar{H}(s)T(s,0)$ satisfies Assumption 4.1 and the gapped adiabatic theorem to all orders, Lemma 4.5, is applicable. Hence

$$\max_s \left\| \mathcal{Q}_1^+ \tilde{W}_\epsilon(s) \mathcal{Q}_0^- \right\| \le C\epsilon\delta^{-1}, \quad \left\| \mathcal{Q}_1^+ \tilde{W}_\epsilon(1) \mathcal{Q}_0^- \right\| \le C_N \epsilon^N \delta^{-N}. \tag{4.52}$$

We now observe that

$$W_\epsilon(s) = \tilde{W}_\epsilon(s) + i\epsilon^{-1} W_\epsilon(s) \int_0^s W_\epsilon^*(s') T(0,s') R_\epsilon(s') \tilde{W}_\epsilon(s') ds',$$

so

$$\left\| W_\epsilon(s) - \tilde{W}_\epsilon(s) \right\| \le \epsilon^{-1} \max_{s' \le s} \left\| R_\epsilon(s') \right\| \le C_N \epsilon^N \Delta^{-N} + Ce^{-c\sqrt{\ell}}, \tag{4.53}$$

using (4.50). We conclude that

$$\begin{aligned} \|Q_1^+ V_\epsilon(s) Q_0^-\| &= \|Q_1^+ T(s,0) W_\epsilon(s) Q_0^-\| = \|\mathcal{Q}_1^+ W_\epsilon(s) \mathcal{Q}_0^-\| \\ &\le \begin{cases} C_N \epsilon^N \Delta^{-N} + C\left(\epsilon\delta^{-1} + e^{-c\sqrt{\ell}}\right) & \text{uniformly in } s; \\ C_N\left(\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-N}\right) + Ce^{-c\sqrt{\ell}} & \text{if } s = 1. \end{cases} \end{aligned}$$

As $V_\epsilon(s) = \bar{Q}_s U_\epsilon(s,0) \bar{Q}_0$, and $\bar{Q}_0 Q_0^- = Q_0^-$, it follows that

$$\begin{aligned} \|Q_1^+ U_\epsilon(s,0) Q_0^-\| &\le \|Q_1^+ V_\epsilon(s) Q_0^-\| + \|Q_1 U_\epsilon(s,0)\bar{Q}_0\| \\ &\le \begin{cases} C_N \epsilon^N \Delta^{-N} + C\left(\epsilon\delta^{-1} + e^{-c\sqrt{\ell}}\right) & \text{uniformly in } s; \\ C_N\left(\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-N}\right) + Ce^{-c\sqrt{\ell}} & \text{if } s = 1, \end{cases} \end{aligned}$$

where in the last step we have used (4.49). □

Let $P^-(s)$ be the spectral projection of $H(s)$ on the interval $(-\infty, E - 6\delta)$ and $P^+(s)$ be the spectral projection on the interval $(E + 6\delta, \infty)$.

We are now ready to complete the proof.

Proof of Theorem 2.2. We pick $\mathcal{Q}(s) = Q_s^-$.

Theorem 2.2.(i): Using the integral representation (B.1),

$$Q_s^- = (2\pi i)^{-1} \oint_\Gamma \left(\bar{H}(s) - z\right)^{-1} dz,$$

we get

$$[\mathcal{Q}(s), H(s)] = (2\pi i)^{-1} \oint_\Gamma \left(\bar{H}(s) - z\right)^{-1} [H(s), \bar{H}(s)] \left(\bar{H}(s) - z\right)^{-1} dz,$$

and we can bound

$$\|[\mathcal{Q}(s), H(s)]\| \le C\delta^{-1} \left\| [H(s), \bar{H}(s)] \right\|.$$

But

$$[H(s), \bar{H}(s)] = [H(s), \bar{Q}_s H(s) \bar{Q}_s] = [H(s), \bar{Q}_s] H(s) \bar{Q}_s + h.c.,$$

which yields

$$\left\| [H(s), \bar{H}(s)] \right\| \le C_N \epsilon + Ce^{-c\sqrt{\ell}}.$$

by Lemma 4.9. Hence

$$\|[\mathcal{Q}(s), H(s)]\| \le C_N \epsilon\delta^{-1} + Ce^{-c\sqrt{\ell}},$$

and 2.2.(i) follows.

Theorem 2.2.(ii): Using (4.36) and $Q_s^- \bar{Q}_s = Q_s^-$, we deduce that

$$\left\| \left(H(s) - \bar{H}(s)\right) P_{<E-6\delta}(H(s)) \right\| + \left\| \left(H(s) - \bar{H}(s)\right) \mathcal{Q}(s) \right\| \le C_N \epsilon\Delta^{-1} + Ce^{-c\sqrt{\ell}}.$$

Hence we can use Lemma B.1 with $H_1 = \bar{H}(s)$, $H_2 = H(s)$, and $R = P_{<E-6\delta}(H(s))$ first to get

$$\left\|\bar{\mathcal{Q}}(s)P_{<E-6\delta}(H(s))\right\| \le C_N\epsilon\Delta^{-1} + Ce^{-c\sqrt{\ell}},$$

and then use the same lemma with $H_1 = H(s)$, $H_2 = \bar{H}(s)$, and $R = \mathcal{Q}(s)$ to get

$$\|P_{>E+6\delta}(H(s))\mathcal{Q}(s)\| \le C_N\epsilon\Delta^{-1} + Ce^{-c\sqrt{\ell}}.$$

Theorem 2.2.(iii): This part follows directly from Lemma 4.12 and the $\pm$ symmetry in the argument there, as

$$\|\mathcal{Q}_\epsilon(s) - \mathcal{Q}(s)\| = \left\|U_\epsilon(s,0)Q_0^- U_\epsilon(0,s) - Q_1^-\right\| \le \left\|Q_1^+ U_\epsilon(1,0)Q_0^-\right\| + \left\|Q_1^- U_\epsilon(1,0)Q_0^+\right\|.$$

□

5. Localization on a torus

5.1. **Consequences of Assumptions 1.2–1.4.** We first note that Assumptions 1.2–1.4 imply localization on torus, as well (e.g., [AW, Theorem 11.2]):

$$\sup_{E\in J_{loc}} \mathbb{E}\left(\left|(H^{\mathbb{T}} - E - i0)^{-1}(x,y)\right|^q\right) \le Ce^{-\mu d(x,y)} \text{ for all } x, y \in \mathbb{T}, \tag{5.1}$$

see Section 2 for notation.

Another consequence of these hypotheses is

Lemma 5.1 (The Wegner estimate)**.** *Let $\Theta \subset \mathbb{T}$. For all $E \in J_{loc}$,*

$$\mathbb{P}\left\{\operatorname{dist}\left\{E, \sigma(H^\Theta)\right\} \le \nu\right\} \le C\nu^q\,|\Theta|\,. \tag{5.2}$$

For the proof, see e.g., [ETV, Eq.(24)].

Together with Assumption 1.3, Lemma 5.1 yields

Lemma 5.2 (Distance between spectra)**.** *Let $\Theta, \Phi \subset \mathbb{T}$ be such that* $\operatorname{dist}(\Theta, \Phi) > r$*. Then*

$$\mathbb{P}\left\{\operatorname{dist}\left(\sigma(H^\Theta)\cap J_{loc}, \sigma(H^\Phi)\cap J_{loc}\right) \le \nu\right\} \le C\nu^q\,|\Theta|\,|\Phi|\,. \tag{5.3}$$

Moreover, if a collection $\{\Theta_i\}_{i=1}^n$ of subsets in $\mathbb{T}$ satisfies $\operatorname{dist}(\Theta_i, \Theta_j) > r$ *for $i \ne j$, $|\Theta_i| \le D$ for all i, and $E \in \mathbb{R}$, then*

$$\mathbb{P}\left\{\operatorname{dist}\left(E, \sigma(H^{\Theta_i})\right) \le \nu \textit{ for all } i\right\} \le (C\nu^q D)^n\,. \tag{5.4}$$

A more subtle implication of our assumptions is the fact that the associated eigenfunction correlator $Q(x, y; J_{loc})$, defined by

$$Q(x,y;J_{loc}) = \sum_{\lambda\in\sigma(H^\Theta)\cap J_{loc}} \left|P_{\{\lambda\}}(x,y)\right| \tag{5.5}$$

satisfies

$$\mathbb{E}Q(x,y;J_{loc}) \le C\mathrm{e}^{-c|x-y|_\Theta} \tag{5.6}$$

for some $c > 0$ that depends only on μ and q. For the non correlated randomness this assertion is known to hold with $c = \mu$, see, e.g. [AW, Theorem 7.7] (the proof relies on the so called spectral averaging procedure available in this case). For a more general class of correlated random models, this consequence of Assumption 1.4 was derived in [ESS, Theorem 4.2], with $c = \frac{q\mu}{9}$.

This statement implies that all eigenstates in $P_{J_{loc}}(H)$ are localized with large probability. We make this statement quantitative.

Definition 5.3. Let $c > 0$. We say that a set $\Theta \subset \mathbb{T}$ is (c,ℓ)*-localizing* for H in the interval $I \subset J_{loc}$ if for all eigenpairs $(E_n, \psi_n)_{E_n\in I}$ of H^Θ there exists a set $\{x_n\}$ in Θ such that

$$|\psi_n(y)| \le Ce^{-c|y-x_n|_\Theta} \text{ for any } y \in \Theta \text{ such that } |y - x_n|_\Theta \ge \sqrt{\ell}. \tag{5.7}$$

Then we have the following result:

Theorem 5.4. *Suppose that Assumption 1.4 holds. Then there exist $c > 0$ such that the probability that a set $\Theta \subset \mathbb{T}$ is (c, ℓ)-localizing for H in the interval J_{loc} is $\geq 1 - C|\Theta|^2 e^{-c\sqrt{\ell}}$.*

For the proof, see e.g., [AW, Theorem 7.4].

We recall that P_E (resp. $P_E^{\mathbb{T}}$) are spectral projections below energy E for H (resp. $H^{\mathbb{T}}$). The next assertion implies that deep inside $\mathbb{T}$, P_E and $P_E^{\mathbb{T}}$ are close.

Proposition 5.5. *Let $\mathcal{L}$, and $\mathbb{T}$ be as above. Then there exists $\bar{\mu} > 0$ such that the probability*

$$\mathbb{P}\left(\left\|\left(P_E - P_E^{\mathbb{T}}\right)\chi_{\Lambda_{\mathcal{L}/2}(0)}\right\| > \mathrm{e}^{-\bar{\mu}\mathcal{L}}\right) \leq \mathrm{e}^{-\bar{\mu}\mathcal{L}}. \tag{5.8}$$

For the proof, see [EPS, Lemma 4.11]. The argument is closely related to the one used in the proof of the following result, that establishes localization property of some bounded functions of H in the mobility gap.

Lemma 5.6. *For any $I := [E_1, E_2] \subset J_{loc}$ and any $\Theta \subset \mathbb{T}$,*

$$\mathbb{E}\left|\left(P_\sharp(H^\Theta)\right)(x, y)\right| \leq C\mathrm{e}^{-\mu|x-y|_\Theta}, \quad \sharp = I, E, \tag{5.9}$$

for all $x, y \in \Theta$. Moreover, for any $z \in \mathbb{C}$ with $Re(z) \in I/2$, we have

$$\mathbb{E}\left|\left(\bar{P}_I(H^\Theta)\left(H^\Theta - z\right)^{-1}\right)(x, y)\right| \leq \frac{C}{E_2 - E_1}\frac{\mathrm{e}^{-\mu|x-y|_\Theta}}{\langle Im\, z\rangle} \tag{5.10}$$

for all $x, y \in \Theta$.

Proof. Let $\sharp = I$. Since Θ is finite, the spectrum of H^Θ is a discrete set. By (1.9), $\{E_{1,2}\} \not\subset \sigma\left(H^\Theta\right)$ almost surely. Thus the spectral projection $P_I(H^\Theta)$ is equal to

$$P_I(H^\Theta) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{j=1}^{2}(-1)^j\left(H^\Theta - iu - E_j\right)^{-1}du \tag{5.11}$$

almost surely, see (B.1). Using $|\left(H^\Theta - iu - E_j\right)^{-1}(x, y)| \leq |u|^{-1}$, we get a bound

$$\left|\left(P_I(H^\Theta)\right)(x, y)\right| \leq \max_j \frac{1}{\pi}\int_{-\infty}^{\infty}\left|\left(H^\Theta - iu - E_j\right)^{-1}(x, y)\right|^q \frac{1}{|u|^{1-q}}du.$$

For $|u| \geq 1$, we use a decomposition

$$\left(H^\Theta - iu - E_j\right)^{-1} = -\left(iu + E_j\right)^{-1} + \left(iu + E_j\right)^{-1}H^\Theta\left(H^\Theta - iu - E_j\right)^{-1},$$

the range-r property for H, and $|H(x, y)| \leq C$ to estimate

$$\begin{aligned}\mathbb{E}\left|\left(P_I(H^\Theta)\right)(x, y)\right| \leq \frac{1}{\pi}\sup_{u\in\mathbb{R}}\max_j\Bigg(&\mathbb{E}\left|\left(H^\Theta - iu - E_j\right)^{-1}(x, y)\right|^q\int_{[-1,1]}\frac{du}{|u|^{1-q}} \\ &+ C\max_{\substack{z\in\mathbb{Z}^d:\\|z-x|\leq r}}\mathbb{E}\left|\left(H^\Theta - iu - E_j\right)^{-1}(z, y)\right|^q\int_{[-1,1]^c}\frac{du}{|u|^{2-q}}\Bigg) \\ \leq C\mathrm{e}^{-\mu|x-y|_\Theta}.&\end{aligned}$$

The argument for $\sharp = E$ is nearly identical.

To get the second assertion of the lemma, we use

$$\left(H^\Theta - z\right)^{-1} = -\left(iIm(z) + 1\right)^{-1} + \left(iIm(z) + 1\right)^{-1}\left(H^\Theta - Re(z) - 1\right)\left(H^\Theta - z\right)^{-1},$$

and

$$\bar{P}_I(H^\Theta)\left(H^\Theta - z\right)^{-1} = (2\pi)^{-1}\sum_{j=1}^{2}\int_{-\infty}^{\infty}\left(z - E_j - iu\right)^{-1}\left(H^\Theta - iu - E_j\right)^{-1}du.$$

It yields

$$\bar{P}_I(H^\Theta)\left(H^\Theta - z\right)^{-1} = -\left(iIm(z)+1\right)^{-1}\bar{P}_I(H^\Theta) +$$
$$(2\pi)^{-1}\sum_{j=1}^{2}\left(iIm(z)+1\right)^{-1}\left(H^\Theta - Re(z) - 1\right)\int_{-\infty}^{\infty}\left(z-E_j-iu\right)^{-1}\left(H^\Theta - iu - E_j\right)^{-1} du.$$

Since $\bar{P}_I = 1 - P_I$, $|iIm(z)+1| = \langle Im\, z\rangle$, and $\left|\frac{1}{z-E_j-iu}\right| \le \frac{4}{E_2-E_1}$ for any $Re(z)\in I/2$ and $u\in\mathbb{R}$, the remaining argument is identical to the one used in the proof of the first bound. □

We will be using the probabilistic version of Lemma 5.6 that follows from the previous statement by the Markov's inequality.

Lemma 5.7. *Let $J := [E_1, E_2] \subset J_{loc}$. Then there exist $c > 0$ such that for any $\Theta \subset \mathbb{T}$, the probability that for all x, y with $|x-y|_\Theta \ge \sqrt{\ell}$,*

$$\left|\left(P_J(H^\Theta)\right)(x,y)\right|, \left|\left(\bar{P}_J(H^\Theta)\left(H^\Theta - z\right)^{-1}\right)(x,y)\right| \le C\mathrm{e}^{-c|x-y|_\Theta} \tag{5.12}$$

is $\ge 1 - Ce^{-c\sqrt{\ell}}$.

5.2. **Local Structure of $H^{\mathbb{T}}$.** Given scales $\ell < \mathcal{L}$ with $\mathcal{L} \bmod \left(\frac{3}{2}\ell\right) = \ell$, and ℓ even, we cover the torus $\mathbb{T} = \mathbb{T}^d_{\mathcal{L}}$ by the collection of boxes Λ_ℓ of the form

$$\left\{\Lambda_\ell(a)/\mathcal{L}\mathbb{Z}^d\right\}_{a\in\Xi_\ell}, \tag{5.13}$$

where

$$\Xi_\ell := \left(\tfrac{3}{2}\ell\mathbb{Z}\right)^d/\mathcal{L}\mathbb{Z}^d. \tag{5.14}$$

We call the collection of boxes *a suitable l-cover* of $\mathbb{T}$.

The (trivial) properties of suitable covers are encapsulated by the following lemma. We recall that we use a max distance.

Lemma 5.8. *Let $r < \ell < \mathcal{L}$. Then a suitable ℓ-cover satisfies*

(i) $\mathbb{T} = \bigcup_{a\in\Xi_\ell}\Lambda_\ell(a)$;

(ii) for all $y \in \mathbb{T}$ there is $a = a(y) \in \Xi_\ell$ such that $\Lambda_{\ell/4}(y)/\mathcal{L}\mathbb{Z}^d \subset \Lambda_\ell(a)/\mathcal{L}\mathbb{Z}^d$. For such a, we will denote $\Lambda_\ell^{(y)} := \Lambda_\ell(a)$;

(iii) $\Lambda_{\ell/4}(a) \cap \Lambda_\ell(a') = \emptyset$ for all $a, a' \in \left(\frac{3}{2}\ell\mathbb{Z}\right)^d$, $a \ne a' \bmod \mathcal{L}$;

(iv) $\left(\frac{\mathcal{L}}{\ell}\right)^d \le |\Xi_\ell| \le \left(\frac{2\mathcal{L}}{\ell}\right)^d$.

Furthermore, any box $\Lambda_\ell(a)$ with $a \in \Xi_\ell$ overlaps with no more than $2d$ other boxes in the ℓ-cover, and any non overlapping boxes are separated by a distance $> r$.

Let $\ell < \mathcal{L}$ and let $\mathcal{S}$ be a subset of a suitable ℓ-cover such that boxes $\{\Lambda_\ell(a)\}_{\mathcal{S}}$ are separated by a distance r. Fix $E \in J_{loc}$, then, by Lemma 5.2, for all $\nu > 0$ we have

$$\mathbb{P}\left\{\operatorname{dist}\left(E, \sigma(H^{\Lambda_\ell(a)})\right) \le \nu \text{ for all } \Lambda_\ell(a)\in\mathcal{S}\right\} \le \left(C\nu^q\ell^d\right)^{|\mathcal{S}|}. \tag{5.15}$$

We now inspect the structure of $P_I(H^{\mathbb{T}})$. We will work with the scale ℓ and the interval $I \subset J_{loc}$ such that

$$\mathrm{e}^{c\sqrt{\ell}} = \mathcal{L} \gg \ell \gg 1, \quad |I| = c\ell^{-\frac{d}{q}}. \tag{5.16}$$

for an ℓ–independent constant c. We remind the reader that we are using a convention that c denotes a sufficiently small constant and C a sufficiently large constant. The value of these constants can change equation by equation.

We endow the set Ξ_ℓ with the usual graph structure, i.e., we will think of its elements as vertices and introduce the edges $\langle a, b\rangle$ between the neighboring elements $a, b \in \Xi_\ell$, separated by a distance $\frac{3}{2}\ell$ on the torus $\mathbb{T}$. By $\mathcal{R}_M$ we will denote a set of all connected subgraphs of Ξ_ℓ with cardinality M, and by $\mathcal{S}_M$ we will denote a collection of sets $\{\cup_{a\in R}\Lambda_\ell(a) :\ R \in \mathcal{R}_M\}$.

Lemma 5.9. *The cardinality of $\mathcal{R}_M$ is bounded by*

$$(2d\mathrm{e})^M |\Xi_\ell| \le \left(\tfrac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^M . \tag{5.17}$$

Proof of Lemma 5.9. We first note that each set S in $\mathcal{S}_M$ looks like a compressed d-dimensional polycube of size M, and we can bound the number of distinct S_M's using the same method as the one used for the latter, see e.g., [BBR]. To make the arguments self-contained, we reproduce it here.

A d-dimensional polycube of size n is a connected set of n cubical cells on the lattice $\mathbb{Z}^d$, where connectivity is through $(d-1)$ faces. Two fixed polycubes are equivalent if one can be transformed into the other by a translation.

Given S, assign the numbers $1, \ldots, M$ to the cubes of S in lexicographic order. Now search the cube-connectivity graph G of S, starting from cube 1. During the search, any cube $c \in S$ is reached through an edge e and connected by edges of G to at most $2d-1$ other cubes. Label such an outgoing edge e' with a pair (i,j): i is the number associated with c, and $1 \le j \le 2d-1$ identifies the orientation of e' with respect to e. By the end of the search, each of the $M-1$ edges of the resulting spanning tree is given a unique label from a set of $(2d-1)M$ possible labels. This is an injection from polycubes of size M to $M-1$-element subsets of a set of size $(2d-1)M$, and so a number of distinct shapes for S's is bounded by

$$\binom{(2d-1)M}{M-1} \le (2d\mathrm{e})^M . \tag{5.18}$$

The total number of sets S can be now bounded by noticing that they are contained in a set of all translates of the distinct shapes of S by elements of Ξ_ℓ, yielding (5.17).

□

For any given configuration ω, let $\tilde{\mathcal{T}}$ denote the union of boxes $\Lambda_\ell(a)$ with $a \in \Xi_\ell$ such that the Dirichlet restriction of H_ω to each box $\Lambda_\ell(a)$ has at least one eigenvalue in the interval $2I$. Let $\mathcal{T}$ denote the union of boxes $\Lambda_\ell(b)$ with $b \in \Xi_\ell$ that has a non trivial overlap with $\tilde{\mathcal{T}}$. We will enumerate by $\{\mathcal{T}_\gamma\}$ a set of connected components in $\mathcal{T}$, i.e.,

$$\mathcal{T} = \cup_\gamma \mathcal{T}_\gamma, \quad \mathcal{T}_\gamma \cap \mathcal{T}_{\gamma'} = \emptyset, \quad \mathcal{T}_\gamma \in \mathcal{S}_M \text{ for some } M \in \mathbb{N}.$$

For a given $\mathcal{T}$, we will denote by $M(\mathcal{T})$ the maximum

$$M(\mathcal{T}) = \max_\gamma M : \ \mathcal{T}_\gamma \in \mathcal{S}_M .$$

For an integer N, let Ω_N denote a subset of the full configuration space for which

$$M(\mathcal{T}) < N.$$

Lemma 5.10. *For $\ell > r$ and $I \subset J_{loc}$ with $|I|^q < c\ell^{-d}$. Then for c small enough we have*

$$\mathbb{P}(\Omega_N^c) \le \left(\tfrac{2\mathcal{L}}{\ell}\right)^d \mathrm{e}^{-N}. \tag{5.19}$$

Proof. For any $\omega \in \Omega_N^c$ there exists at least one cluster $\mathcal{T}_\gamma \in \mathcal{S}_M$ with $M \ge N$. Let $\tilde{\mathcal{T}}_\gamma \in S_{M_\gamma}$ denote the union of boxes that generates $\mathcal{T}_\gamma$, i.e., $\tilde{\mathcal{T}}_\gamma \subset \tilde{\mathcal{T}}$ and $\mathcal{T}_\gamma$ is formed by all boxes that overlap with a box in $\tilde{\mathcal{T}}_\gamma$. Note that any box $\Lambda_l(a) \subset \tilde{\mathcal{T}}_\gamma$ overlaps with 3^d boxes. Including each box and all its neighbors, the boxes generate $3^d M_\gamma$ boxes. Let U be a collection of vectors in $\mathbb{R}^d$ whose components are either zero or unit. Then $\Xi_\ell = \cup_{e \in U} \Xi_{\ell,e}$ where $\Xi_{\ell,e} = \frac{3}{2}e + (3\ell\mathbb{Z})^d / \mathcal{L}\mathbb{Z}^d$. Then $\Xi_{\ell,e} \cap \Xi_{\ell,e'} = \emptyset$ for $e \ne e'$ and

$$\Lambda_\ell(a) \cap \Lambda_\ell(a') = \emptyset \text{ for all } a \in \Xi_{\ell,e}, \quad a \in \Xi_{\ell,e'}, \tag{5.20}$$

using the fact that ℓ is even. Hence for any $S \subset \Xi_\ell$, there exists $e \in U$ such that $|S \cap \Xi_{\ell,e}| \ge 2^{-d} |S|$. In particular, the number of non overlapping boxes in $\tilde{\mathcal{T}}_\gamma$ is at least $6^{-d} M$ thanks to (5.20).

We are now in position to apply (5.15) to conclude that the probability that a *fixed* configuration $\mathcal{T}$ has at least one cluster $\mathcal{T}_\gamma \in \mathcal{S}_M$ with $M \geq N$ is bounded by $\left(C\,|I|^q\,\ell^d\right)^{6^{-d}M}$. It follows now from Lemma 5.9 that

$$\mathbb{P}(\Omega_N^c) \leq \sum_{M=N}^{\infty} \left(\tfrac{2\mathcal{L}}{\ell}\right)^d \left((2d\mathrm{e})^{(6^d)}\, C\,|I|^q\,\ell^d\right)^{6^{-d}M}. \tag{5.21}$$

This is less than equal to $\left(\frac{2\mathcal{L}}{\ell}\right)^d \mathrm{e}^{-N}$ provided c is small enough.

□

For an integer N, we now consider a subset $\Omega_{loc,N}$ of the full configuration space for which $\mathbb{T}$ and all sets in $\{S_M\}_{M=1}^N$ are $\ell/10$-localizing and satisfy (5.12).

Lemma 5.11. *There exists constants C, c such that*

$$\mathbb{P}(\Omega_{loc,N}^c) \leq CN^2\,(2\mathcal{L}\ell)^d\,(2d\mathrm{e})^N\,e^{-c\sqrt{\ell}}. \tag{5.22}$$

Proof. The total number of $\{S_M\}_{M=1}^N$ is bounded by

$$\sum_{M=1}^{N} \left(\tfrac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^M < 2\left(\tfrac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^N$$

thanks to Lemma 5.9. Their maximal volume is bounded by $N\ell^d$. So we can bound

$$\mathbb{P}(\Omega_{loc,N}^c) \leq C\left(\tfrac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^N \left(N\ell^d\right)^2 e^{-c\sqrt{\ell}} = CN^2\,(2\mathcal{L}\ell)^d\,(2d\mathrm{e})^N\,\mathrm{e}^{-c\sqrt{\ell}}, \tag{5.23}$$

using Theorem 5.4 and Lemma 5.7. □

We now optimize N in the previous two lemmas. To this end, we pick $N = \lfloor c\sqrt{\ell} \rfloor$. Then, using Lemmata 5.10–5.11, for ℓ large enough and intervals $I \subset J_{loc}$ satisfying $|I| < C\ell^{-d/q}$, we have

$$\mathbb{P}((\Omega_N \cap \Omega_{loc,N})^c) \leq \mathcal{L}^d \mathrm{e}^{-c\sqrt{\ell}}. \tag{5.24}$$

For $\omega \in \Omega_N \cap \Omega_{loc,N}$, the number of eigenvalues of $H^{\mathcal{T}_\gamma}$ cannot exceed $|\mathcal{T}_\gamma| \leq N\ell^d \leq C\ell^{d+1/2}$. Hence, for each γ, we can find $J_\gamma := [E_\gamma^-, E_\gamma^+]$ such that

$$I/2 \subset J_\gamma \subset I \quad \text{and} \quad \operatorname{dist}(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma})) \geq c\ell^{-d-1/2}\,|I|\,.$$

We note that

$$\max_\gamma \operatorname{diam}(\mathcal{T}_\gamma) \leq L := C\ell^{3/2}. \tag{5.25}$$

Let Ω_G be a subset of configuration set $\Omega_N \cap \Omega_{loc,N}$ such that, for c small enough,

$$\sup_{\mathcal{T}_\gamma} \left| \left(\left(H^{\mathcal{T}_\gamma} - z\right)^{-1} \right)(x,y) \right| \mathrm{e}^{c\ell^{-1/2}\,|x-y|_{\mathcal{T}_\gamma}} \leq C\ell^{d+\frac{1}{2}}\,|I|^{-1}\,\langle Im\, z\rangle^{-1} \tag{5.26}$$

for $\omega \in \Omega_G$, $z \in \mathbb{C}$ with $\mathrm{Re}(z) = E_\gamma^\pm$ and all $x, y \in \mathcal{T}_\gamma$.

Applying Lemma 5.7 with $J = E_\gamma^\pm + [-c\ell^{-d-1/2}\,|I|\,, c\ell^{-d-1/2}\,|I|]$, and $z \in \mathbb{C}$ with $\mathrm{Re}(z) = E_\gamma^\pm$ yields

$$\mathbb{P}\left(\Omega_G^c\right) \leq \mathcal{L}^d e^{-c\sqrt{\ell}}.$$

Proposition 5.12. *Let $\omega \in \Omega_G$, and let $I \subset J_{loc}$ be such that $|I| < c\ell^{-d/q}$. Suppose that ℓ is large enough, then*

(i) (Local Gap) There exists intervals $J_\gamma = [E_\gamma^-, E_\gamma^+]$ such that

$$I/2 \subset J_\gamma \subset I \quad \textit{and} \quad \operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma})\right) \geq c\ell^{-d-1/2}\,|I|\,; \tag{5.27}$$

(ii) (Support of spectral projections)

$$\left\|P_J(H^{\mathbb{T}})\delta_x\right\| \le e^{-c\sqrt{\ell}} \text{ for any } x \in \mathbb{T}\setminus\mathcal{T}_\ell, \tag{5.28}$$

and

$$\left\|P_{J_\gamma}(H^{\mathcal{T}_\gamma})\delta_x\right\| \le e^{-c\sqrt{\ell}} \text{ for any } x \notin \partial_{\ell/8}\mathcal{T}\cup\mathcal{T}_\ell; \tag{5.29}$$

(iii) (Exponential Decay of Correlations) Let $\mathcal{A}_o = \partial_\ell\mathcal{T}_\gamma \cup (\mathcal{T}_\gamma)_{8\ell}$, then (with $\mathcal{A} = \mathcal{A}_o$ in (1.18)–(1.20)*) we have*

$$\left\|\left(H^{\mathcal{T}_\gamma}(s)) - z\right)^{-1}\right\|_{c,\ell} \le \frac{\ell^{3d}}{\Delta}\frac{1}{\langle Im\, z\rangle}, \tag{5.30}$$

for $z \in \mathbb{C}$ with $\mathrm{Re}(z) = E_\gamma^\pm$*.*

Proof. The property 5.12.(i) we already established earlier and 5.12.(iii) is a consequence of (5.26). This leaves us with the property 5.12.(ii).

Let $\{\lambda_n, \psi_n\}$ be an eigenpair for $H^{\mathbb{T}}$ in I, and let x_n be its localization center. We first check that $x_n \in \tilde{\mathcal{T}}$. Indeed, suppose that $x_n \notin \tilde{\mathcal{T}}$. Then by the properties of the suitable cover, there exists a box $\Lambda_\ell(a) \notin \tilde{\mathcal{T}}$ such that $\Lambda_{\ell/4}(x_n) \subset \Lambda\setminus\tilde{\mathcal{T}}$ and $\Lambda_{\ell/4}(x_n) \subset \Lambda_\ell(a)$. Since

$$|\psi_n(y)| \le C\mathrm{e}^{-\mu|y-x_n|_{\Lambda_\ell(a)}} \text{ for } |y-x_n|_{\Lambda_\ell(a)} \ge \sqrt{\ell/10},$$

by localization of ψ_n, we can use Lemma B.2 below to conclude

$$\sigma\left(H^{\Lambda_\ell(a)}\right) \cap 2I \neq \emptyset, \tag{5.31}$$

which means that $\Lambda_\ell(a) \in \tilde{\mathcal{T}}$, a contradiction. This shows (5.28).

Let $\{\mu_n, \phi_n\}$ be an eigenpair for $H^{\mathcal{T}}$ in I. By the argument identical to the one for Item *(ii)*, its localization center y_n is either located in $\tilde{\mathcal{T}}$ or in $\partial_{C\sqrt{\ell}}\mathcal{T} \subset \partial_{\ell/8}\mathcal{T}$. This shows that

$$\|P_{J_\gamma}(H^{\mathcal{T}_\gamma}) - \chi_{\partial_{\ell/8}\mathcal{T}}P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\partial_{\ell/8}\mathcal{T}} - \chi_{\mathcal{T}_\ell}P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\mathcal{T}_\ell}\| \le e^{-c\sqrt{\ell}}, \tag{5.32}$$

and, in particular, shows (5.29). In fact it shows more, namely that (recall notation in Theorem 2.1.(iii))

$$\left\|P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\delta_x\right\| \le e^{-c\sqrt{\ell}} \text{ for any } x \notin \mathcal{A}_o. \tag{5.33}$$

□

This completes the proof that $H^{\mathbb{T}}$ possesses a local structure. Using perturbation theory methods we are now going to show that $H^{\mathbb{T}}(s)$ possesses a local structure as well.

5.3. **Proof of Theorem 2.1.** We denote by $\Delta = c\ell^{-d/q-d-1/2}$ the lower bound on the gap of $H^{\mathcal{T}_\gamma}$ that we established in the previous section, i.e. $\operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma})\right) \ge \Delta$. Let $a \in \mathbb{R}$ be such that $H_a^{\mathcal{T}_\gamma}(s) := H^{\mathcal{T}_\gamma}(s) + aP_{[E_\gamma^-, E_\gamma^+]}\left(H^{\mathcal{T}_\gamma}(s)\right)$ satisfy

$$\sigma\left(H_a^{\mathcal{T}_\gamma}(s)\right) \cap \left([-\tfrac{\Delta}{3}, \tfrac{\Delta}{3}] + [E_\gamma^-, E_\gamma^+]\right) = \emptyset, \tag{5.34}$$

provided $\beta < \frac{\Delta}{6}$ (e.g., $a = \Delta + E_\gamma^+ - E_\gamma^-$ will do).

For the next assertion, we recall the definition of a dilation and its norm, introduced in (1.18)–(1.20).

Lemma 5.13. *There exists $c > 0$ such that for $\beta < c\Delta\ell^{-d}$ and any $z \in \mathbb{C}$ with $Re(z) = E_\gamma^\pm$, we have*

$$\left\|\left(H^{\mathcal{T}_\gamma}(s) - z\right)^{-1}\right\|_{c,\ell} + \left\|\left(H_a^{\mathcal{T}_\gamma}(s) - z\right)^{-1}\right\|_{c,\ell} \le C\ell^{3d}\Delta^{-1}\langle Im\, z\rangle^{-1}, \tag{5.35}$$

where $\|\cdot\|_{c,\ell}$ is defined with $\mathcal{A} = \mathcal{A}_o$.

Proof. If we denote by $R^o_{z,a}$ and $R_{z,a}$ the resolvents

$$R^o_{z,a} = \left(H_a^{\mathcal{T}_\gamma}(0) - z\right)^{-1}, \quad R_{z,a} = \left(H_a^{\mathcal{T}_\gamma}(s) - z\right)^{-1}, \tag{5.36}$$

we have

$$\left\|R^o_{z,0}\right\|_{c,\ell} \le C\ell^{3d}\Delta^{-1}\langle Im\, z\rangle^{-1} \tag{5.37}$$

by (5.25)–(5.26).

Using (5.33), we deduce that

$$\left\|\mathrm{e}^{c\,\rho^\ell_{\mathcal{A}}}\, P_{[E^-_\gamma, E^+_\gamma]}\left(H^{\mathcal{T}_\gamma}(s)\right)\right\| \le C\ell^d. \tag{5.38}$$

Since

$$R^o_{z,a} = R^o_{z,0} - aP_{[E^-_\gamma, E^+_\gamma]}\left(H_o^{\mathcal{T}_\gamma}\right) R^o_{z,0} R^a_{z,a} P_{[E^-_\gamma, E^+_\gamma]}\left(H_o^{\mathcal{T}_\gamma}\right), \tag{5.39}$$

we obtain, using (5.37)–(5.39) and

$$\left\|P_{[E^-_\gamma, E^+_\gamma]}\left(H_o^{\mathcal{T}_\gamma}\right) R^o_{z,0} R^a_{z,a} P_{[E^-_\gamma, E^+_\gamma]}\left(H_o^{\mathcal{T}_\gamma}\right)\right\| \le C\langle Im\, z\rangle^{-2},$$

that

$$\left\|R^o_{z,a}\right\|_{c,\ell} \le C\ell^{3d}\Delta^{-1}\langle Im\, z\rangle^{-1}$$

as well.

We now expand $R_{z,a}$ into the Neumann series

$$R_{z,a} = R^o_{z,a}\sum_{n=0}^\infty \beta^n \left(WR^o_{z,a}\right)^n,$$

yielding, via (1.21),

$$\begin{aligned}\left\|R_{z,a}\right\|_{c,\ell} &\le \left\|R^o_{z,a}\right\|_{c,\ell}\sum_{n=0}^\infty \beta^n \left\|WR^o_{z,a}\right\|^n_{c,\ell} \\ &\le C\ell^{3d}\Delta^{-1}\langle Im\, z\rangle^{-1}\sum_{n=0}^\infty\left(\beta C\ell^d\right)^n\Delta^{-n} \le C\ell^{3d}\Delta^{-1}\langle Im\, z\rangle^{-1},\end{aligned}$$

provided $\beta \le c\Delta\ell^{-d}$. □

We are now ready to complete the proof. It will be convenient to relabel ℓ everywhere in the formulation of Theorem 2.1 by 8ℓ (this does not affect any of its implications). We first note that (5.30) is a direct consequence of Lemma 5.13. We recall $J_\gamma = [E^-_\gamma, E^+_\gamma]$ and we set $\hat{J}_\gamma = \left[-\frac{\Delta}{8}, \frac{\Delta}{8}\right] + [E^-_\gamma, E^+_\gamma]$. We will abbreviate $P_\gamma := P_{J_\gamma}\left(H^{\mathcal{T}_\gamma}(s)\right)$, $\hat{P}_\gamma := P_{\hat{J}_\gamma}\left(H^{\mathcal{T}_\gamma}(s)\right)$ and suppress the s dependence for this argument. We use the decomposion (5.11) with $E_1 = E^-_\gamma$ and $E_2 = E^+_\gamma$ to write (recall (5.36))

$$P_\gamma = P_\gamma\hat{P}_\gamma = (2\pi)^{-1}\int_{-\infty}^\infty\sum_{j=1}^2(-1)^j R_{iu+E_j,0}\hat{P}_\gamma du. \tag{5.40}$$

We note that the integrand can be bounded by

$$\max_{j=1,2}\left\|R_{iu+E_j,0}\hat{P}_\gamma\right\| \le 8\Delta^{-1}\langle u\rangle^{-1}, \quad u\in\mathbb{R}. \tag{5.41}$$

Using

$$R_{iu+E_j,0} = R_{iu+E_j,a} - R_{iu+E_j,a}P_{J_\gamma}(H_o^{\mathcal{T}_\gamma})R_{iu+E_j,0}$$

and

$$\int_{-\infty}^\infty\sum_{j=1}^2(-1)^j R_{iu+E_j,a}du = 0$$

which holds thanks to (5.34), we conclude that P_γ is equal to

$$-(2\pi)^{-1}\sum_{j=1}^{2}(-1)^j\int_{-\infty}^{\infty}R_{iu+E_j,a}P_{J_\gamma}(H_o^{\mathcal{T}_\gamma})R_{iu+E_j,0}\hat{P}_\gamma du. \tag{5.42}$$

Hence we can bound

$$\begin{aligned}\left\|\mathrm{e}^{\frac{c}{\sqrt{\ell}}\rho_{\mathcal{A}}}P_\gamma\right\| &\le \int_{-\infty}^{\infty}\max_j\left(\left\|R_{iu+E_j,a}\right\|_{c,\ell}\right)\left\|\mathrm{e}^{\frac{c}{\sqrt{\ell}}\rho_{\mathcal{A}}}P_{J_\gamma}(H_o^{\mathcal{T}_\gamma})\right\|\left\|R_{iu+E_j,0}\hat{P}_\gamma\right\| \\ &\le C\ell^{4d}\Delta^{-2}\int_{-\infty}^{\infty}\langle u\rangle^{-2}du \le C\ell^{4d}\Delta^{-2}, \end{aligned}\tag{5.43}$$

where we have used Lemma 5.13, (5.38), and (7.2) in the second step.

For $\beta \ll \Delta$, E_j does not belong to the spectrum of $H^{\mathcal{T}_\gamma}$ and we have

$$P_\gamma = \int_{-\infty}^{\infty}\sum_{j=1}^{2}(-1)^j R_{iu+E_j,0}\,du.$$

By perturbation expansion for the resolvent we then have

$$P_\gamma = P_{J_\gamma}(H_o^{\mathcal{T}_\gamma}) + \int_{-\infty}^{\infty}\sum_{j=1}^{2}\sum_{n=1}^{\infty}\beta^n R^0_{-iu+E_j}(WR^0_{-iu+E_j})^n.$$

Using the exponential decay of correlations of the resolvent (5.26), we have

$$\|\chi_{\partial_{\ell/8}\mathcal{T}}R^0_{-iu+E_j}\chi_{\mathcal{T}_\ell}\;| \le C\langle u\rangle^{-1}\,e^{-c\sqrt{\ell}}.$$

This together with (5.32) then implies $\left\|\chi_{\partial_{\ell/8}\mathcal{T}}P_I^\gamma\chi_{\mathcal{T}_\ell}\right\| \le e^{-c\sqrt{\ell}}$. Combining it with (5.43), we get (2.4).

The proof of (2.3) is essentially identical to the one above, and so is left out.

6. Uniformly localized eigenfunctions for $H(s)$

Let H_ω be the infinite volume operator that satisfies Assumptions 1.2–1.4. The starting point of our analysis here is

6.1. **Non-uniform bound on localization.**

Theorem 6.1 (Eigenfunction localization)**.** *There exists $c>0$ such that for $\mathbb{P}$-almost every ω and Lebesgue a.e. $\beta\in[0,1]$, $\sigma(H_\omega)$ is simple and for each $E\in\sigma(H_\omega)$ there is a localization center $x_E(\omega,\beta)$ such that the normalized eigenfunction $\psi_E(\cdot,\omega)$ satisfies, for all $y\in\mathbb{Z}$*

$$|\psi_E(y,\omega)|^2 \le A(\omega)\langle x_E(\omega)\rangle^2\mathrm{e}^{-c|y-x_E(\omega)|}, \tag{6.1}$$

with $A(\cdot)\in L^1(\Omega,\mathbb{P})$.

This statement is a consequence of [AW, Theorems 5.8, 7.4, and 12.11].

We now formulate the probabilistic version of this assertion. For this, we need a stronger concept of localizing Hamiltonian than the one introduced earlier in Definition 5.3.

Definition 6.2. For $\omega\in\Omega$ and a pair (c,θ) of positive parameters, we will say that H_ω is *non-uniformly (c,θ)-localizing* if for each $E\in\sigma(H_\omega)$

$$|\psi_E(y,\omega)|^2 \le \frac{1}{\theta}\langle x_E(\omega)\rangle^2\mathrm{e}^{-c|y-x_E(\omega)|}. \tag{6.2}$$

The quantifier "non-uniformly" here refers to the presence of the factor $\langle x_E(\omega)\rangle^2$ which prevents the uniform estimates on ψ_E.

Then Theorem A.1 implies, via Markov's inequality, that

$$\mathbb{P}\left(\{\omega\in\Omega:\ H_\omega \text{ is non-uniformly } (c,\theta)\text{-localizing}\}\right) \le 1-C\theta \tag{6.3}$$

for some $C>0$.

6.2. **From non-uniform to uniform estimates.** Our first goal in this section is to remove the "non-uniform" part from the above statement, at a price of the small fraction of the eigenstates for which the statement will fail to hold.

We first note that that the integrated density of states (IDOS) $\mathcal{N}_{J_{loc}}$ of H_o, associated with the interval J_{loc}, defined as

$$\mathcal{N}_{J_{loc}} = \lim_{R\to\infty} \frac{\mathrm{tr}\chi_{\Lambda_R(0)} P_{J_{loc}}(H_\omega)}{R^d} \tag{6.4}$$

is almost surely non random, see e.g., [AW, Theorem 3.15 and Corollary 3.16]. Moreover, if $\mathcal{N}_{J_{loc}} > 0$, then the convergence to the mean in (6.4) is exponentially fast, so in particular

$$\mathbb{P}\left(\frac{\mathrm{tr}\chi_{\Lambda_R(0)} P_{J_{loc}}(H_o)}{R^d} < \frac{\mathcal{N}_{J_{loc}}}{2}\right) \le \mathrm{e}^{-mR} \tag{6.5}$$

for some $m > 0$. This is a typical large deviations result, see e.g., [CL]. We will assume here that $\mathcal{N}_{J_{loc}} > 0$.

We now adjust the concept of localizing eigenvectors to make it uniform.

Definition 6.3. For $\omega \in \Omega$ and a pair (c, θ) of positive parameters, we will say that a normalized $\psi \in \ell^2(\mathbb{Z}^d)$ is (c, θ)-localized if there exists $x \in \mathbb{Z}^d$ such that

$$|\psi(x)|^2 \ge |\ln\theta|^{-d-1} \text{ and } |\psi(y)| \le \frac{|\ln\theta|^{\frac{d+1}{2}}}{\theta} \mathrm{e}^{-c|y-x|}, \quad y \in \mathbb{Z}^d. \tag{6.6}$$

Armed with this definition, we proceed in getting the uniform estimates, first for finite (albeit arbitrary large) systems, and then for an infinite volume one.

Let $H_L^{\mathbb{T}}$ denote the periodic restriction of H_ω to the torus $\mathbb{T}_L$. The following assertion follows from the judicious use of Markov's inequality and the elementary deterministic Lemma 6.5 below.

Theorem 6.4. *Suppose that Assumptions 1.2–1.4 hold and in addition $\sigma(H_L^{\mathbb{T}}) \cap J_{loc}$ is a.s. simple and $\mathcal{N}_{J_{loc}} > 0$. For a given configuration $\omega \in \Omega$, let $\mathbb{P}_E$ denote the normalized counting measure of eigenvalues of $H_L^{\mathbb{T}}$ in the interval J_{loc}. Let $\mathcal{G}$ be the set*

$$\mathcal{G} := \left\{E_n \in \sigma(H_L^{\mathbb{T}}) \cap J_{loc} : \ E_n \text{ is simple and } \psi_n \text{ is } (c,\theta)\text{-localized}\right\}.$$

Then there exist $c, C > 0$ such that for any L and θ sufficiently small we have a bound

$$\mathbb{P}\left(\mathbb{P}_E(\mathcal{G}) \ge 1 - \sqrt{\theta}\right) \ge 1 - C\sqrt{\theta}. \tag{6.7}$$

Proof. For an eigenpair (E_n, ψ_n), let

$$w_n = w(\omega, \psi_n) = \sum_{x,y} |\psi_n(x)|\,|\psi_n(y)|\,\mathrm{e}^{c|x-y|}. \tag{6.8}$$

We then have, by the bound (5.6) on the eigenvector correlator and $\mathcal{N}_{J_{loc}} > 0$,

$$\mathbb{E}_\omega \mathbb{E}_E[w_n] \le C.$$

Let $a, b > 0$, then by Markov's inequality, we then have

$$\mathbb{P}_\omega\left(\mathbb{E}_E[w_n] \le \theta^{-a}\right) \ge 1 - C\theta^a$$

Pick now an ω such that $\mathbb{E}_E[w_n] \le \theta^{-a}$. Then another application of Markov's inequality gives that

$$\mathbb{P}_E(w_n \le \theta^{-b}) \ge 1 - \theta^{b-a}. \tag{6.9}$$

For a normalized vector $\psi \in \ell^2(\mathbb{T}_L)$, let

$$M(\psi) := \max_{x\in\mathbb{T}_L} |\psi(x)|^2 =: |\psi(x_o)|^2 \text{ for some } x_o \in \mathbb{T}_L.$$

Clearly, we have $M(\psi_n) \le w_n$ for every n. Thus the assertion follows from (6.9) with $a = \frac{1}{2}$, $b = 1$, and

Lemma 6.5. *Suppose that the normalized vector $\psi \in \ell^2(\mathbb{T}_L)$ satisfies*

$$\max_{x,y\in\mathbb{T}_L}\left(|\psi(x)|\,|\psi(y)|\,\mathrm{e}^{c|x-y|}\right) \le \frac{1}{\theta}. \tag{6.10}$$

Then for any sufficiently small (but L-independent) θ we have $M(\psi) \ge |\ln\theta|^{-d-1}$ and, for the corresponding x_o,

$$|\psi(y)| \le \frac{|\ln\theta|^{\frac{d+1}{2}}}{\theta}\mathrm{e}^{-c|y-x_o|}, \quad y \in \mathbb{T}_L.$$

□

Proof of Lemma 6.5. The second bound is an immediate consequence of the first, so we only need to show that $M(\psi) \ge |\ln\theta|^{-d-1}$. Let $r = r(c,\theta) > 0$ be such that $\sum_{y\in\mathbb{Z}^d:\ |y|>r}\mathrm{e}^{-2c|y|} \le \frac{\theta^2M^2}{2}$. In particular, for a fixed c there exists C such that we can chose $r = -C\ln\left(\theta^2M^2\right)$ for θ sufficiently small. Then by (6.10) we can bound

$$1 \;=\; \sum_{x\in\mathbb{T}_L}|\psi(x)|^2 \;\le\; M(\psi)\sum_{\substack{x\in\mathbb{T}_L:\\ |x-x_o|\le r}}1 + \sum_{\substack{x\in\mathbb{T}_L:\\ |x-x_o|>r}}\frac{\mathrm{e}^{-2c|x-x_o|}}{M(\psi)^2\theta^2} \;\le\; M(\psi)(2r+1)^d + \frac{1}{2}. \tag{6.11}$$

This implies that $M(\psi) \ge \frac{1}{2(2r+1)^d}$ or, in view of what r is, $M(\psi) \ge u$, where u is a unique positive solution of

$$\mathrm{e}^{-Cu^{-\frac{1}{d}}} = \theta^2u^2.$$

Since $u > |\ln\theta|^{-d-1}$ for θ sufficiently small, we get $M(\psi) \ge |\ln\theta|^{-d-1}$. □

We now extend Theorem 6.4 to the infinite volume setting. To this end, we will use

Theorem 6.6. *Let x_n denote the localization center for ψ_n, an eigenvector of H_ω with energy E_n (i.e., $M(\psi_n) = |\psi_n(x_n)|^2$). Let $\mathcal{A}_\omega^R$ be the event*

$$\mathcal{A}_\omega^R = \{E_n \in \sigma(H_\omega)\cap J_{loc}:\ x_n \in \Lambda_R(0)\}.$$

Then

$$\lim_{R\to\infty}\mathbb{P}_E^R\left(\mathcal{A}_\omega^R\right) = 1 \tag{6.12}$$

$\mathbb{P}_\omega$-a.s., provided $\mathcal{N}_{J_{loc}} > 0$ and (A.3) *holds.*

Proof. For the case $J_{loc} = \mathbb{R}$, see [DJLS, Theorem 7.1]. The proof there can be readily adapted to the present situation. □

This assertion allows us to give an analogue of the normalized counting measure of eigenvalues of $H_L^{\mathbb{T}}$ in the interval J_{loc} for the infinite volume Hamiltonian H_ω. Namely, for a given configuration $\omega \in \Omega$, let $\mathbb{P}_E^R$ denote the normalized counting measure of eigenvalues of H_ω in the interval J_{loc} with $x_n \in \Lambda_R(0)$. We then have, by the bound (5.6) on the eigenvector correlator and $\mathcal{N}_{J_{loc}} > 0$,

$$\mathbb{E}_\omega\mathbb{E}_E^R[w_n] \le C,$$

see (6.8) for the definition of w_n. Then the argument identical to the one used in the proof of Theorem 6.4 yields

Theorem 6.7. *Suppose that Assumptions 1.2–1.4 hold and in addition $\sigma(H_L^{\mathbb{T}})\cap J_{loc}$ is a.s. simple and $\mathcal{N}_{J_{loc}} > 0$. Let $\mathcal{G}$ be the event*

$$\mathcal{G}_R := \{E_n \in \sigma(H_\omega)\cap J_{loc}:\ E_n \text{ is simple and } \psi_n \text{ is } (c,\theta)\text{-localized with } x_n \in \Lambda_R(0)\}.$$

Then there exist $c, C > 0$ such that for any R sufficiently large and θ sufficiently small we have a bound

$$\mathbb{P}\left(\mathbb{P}_E^R(\mathcal{G}_R) \ge 1-\sqrt{\theta}\right) \ge 1 - C\sqrt{\theta}. \tag{6.13}$$

We are now ready to complete

Proof of Theorem 1.5. Let $\theta = e^{-c\sqrt{\ell}}$ and pick $R = R(\ell)$ such that (6.13) holds. Let $\mathcal{L} = C\epsilon^{-1}$ and consider

$$\Xi_{\mathcal{L}} := \left(\tfrac{3}{2}\mathcal{L}\mathbb{Z}\right)^d, \tag{6.14}$$

cf. (5.14), and an $\mathcal{L}$-cover of $\mathbb{Z}^d$ of the form

$$\mathbb{Z}^d = \bigcup_{a \in \Xi_{\mathcal{L}}} \Lambda_{\mathcal{L}}(a).$$

We note that for any $x \in \mathbb{Z}^d$ we can find $a \in \Xi_{\mathcal{L}}$ such that $\operatorname{dist}(\Lambda^c_{\mathcal{L}}(a), x) \geq \mathcal{L}/4$.

We also cover J_{loc} with the overlapping intervals

$$J_{loc} = \cup_i J_i,$$

so that (a) length of each interval J_i is equal to $c\ell^{-\xi}$ and (b) for each $E \in J'_{loc}$ that satisfies $\operatorname{dist}(E, J^c_{loc}) \geq \ell^{-\xi}$ we can find J_i such that $\operatorname{dist}(E, J^c_i) \geq \ell^{-\xi}/3$. It is clear that one can always construct such a covering using $C\ell^{\xi}$ intervals J_i.

We will say that a property $\mathcal{A}$ is satisfied for a fraction $1-\theta$ of boxes $\Lambda_{\mathcal{L}}(a)$ (which we will be calling the good ones) if

$$\lim_{R\to\infty} \frac{\#\Lambda_{\mathcal{L}}(a) \subset \Lambda_R : \ \mathcal{A} \text{ is satisfied for } \Lambda_{\mathcal{L}}(a)}{\#\Lambda_{\mathcal{L}}(a) \subset \Lambda_R} = 1-\theta. \tag{6.15}$$

Identifying each box $\Lambda_{\mathcal{L}}(a)$ in the cover with the corresponding torus $\mathbb{T}_a$, we can use Theorem 2.1 to conclude, using the ergodicity, that $1 - e^{-c\sqrt{\ell}}$ fraction of such tori satisfy the conclusions of Theorem 4.6 for each interval J_i in the cover of J_{loc} (we recall that given N tori $\{\mathbb{T}_a\}$, we can chose at least $6^{-d}N$ of them to be disjoint, see the proof of Lemma 5.10).

We now pick any $\omega \in \mathcal{G}_R$ and conclude from the previous statement that the fraction $1-e^{-c\sqrt{\ell}}$ of eigenstates ψ_n for H_o with eigenvalues $E_n \in J_{loc}$ are (c,θ)-localized. Let ψ be one of such eigenfunctions. Then there exists a box $a \in \Xi_{\mathcal{L}}$ and an interval J_i such that

$$\operatorname{dist}(\Lambda^c_{\mathcal{L}}(a), x_n) \geq \mathcal{L}/4, \quad \|\bar{\chi}_\Lambda \psi\| \leq e^{-c\mathcal{L}}, \quad E \in J_i.$$

If this box happen to be a good box, then the first assertion of Theorem 1.5 holds for all s by Theorem 2.1, and the second assertion holds for $s = 0$ by Lemma B.2 below and the assertions of Theorem 2.1. It then follows from Theorem 4.6 that the second assertion holds for all $s \in [0,1]$. Since the fraction of good boxes is $1 - e^{-c\sqrt{\ell}}$, we get the result.

□

7. Derivation of Linear Response Theory

In this Section, we prove Theorem 1.6 and we assume the setting outlined in the section 1.4. The proof has three steps.

1) We approximate the $\mathbb{Z}^2$ geometry by the one of a (sufficiently large) torus. In particular, we replace P_{E_F} by $P^{\mathbb{T}}_{E_F}$.
2) We decompose $P^{\mathbb{T}}_{E_F}$ into the adiabatic projection $\mathcal{Q}$ corresponding to the energy $E = E_F + 6\delta$ (recall (2.1)) and a reminder term R. We then use the adiabatic theorem on torus to control the evolution of the adiabatic part.
3) We show that the reminder term does not contribute to the transport, and the adiabatic term gives the Kubo formula. The reminder term does not contribute because it consists of localized states.

Let $\mathcal{L} = C\epsilon^{-1}$ and let $\mathbb{T}$ be a torus of the linear size $\mathcal{L}$. For the first step, we consider a region $\mathcal{B} \subset \mathbb{T}$ be a region such that $\Lambda_{\mathcal{L}/4} \subset \mathcal{B} \subset \Lambda_{\mathcal{L}/3}$. The precise choice of $\mathcal{B}$ is made afterwards.

Lemma 7.1. *With probability* $1 - e^{-c\mathcal{L}}$, *the operator* $(P_\epsilon(s) - P)\, J$ *is trace class and*

$$\operatorname{tr}(P_\epsilon(s) - P)\, J = \operatorname{tr}\big(P^{\mathbb{T}}_\epsilon(s) - P^{\mathbb{T}}\big)\, \tilde{J} + \mathcal{O}(e^{-c\mathcal{L}}), \tag{7.1}$$

where $P^{\mathbb{T}} = P^{\mathbb{T}}_{E_F}(H)$ is a Fermi projection on the torus, and

$$\tilde{J} = \chi_{\mathcal{B}} J.$$

We note that $\tilde{J}$ is supported in a strip $|x_1| \leq r$.

Proof. We will show that, with probability $\geq 1 - \mathrm{e}^{-c\mathcal{L}}$,

$$\left\| \left(P^{\sharp}_{\epsilon}(s) - P^{\sharp} \right) \chi_{\{x\}} J \right\| \leq e^{-c|x|} \text{ for } |x| \geq \mathcal{L}/3, \quad P^{\sharp} = \left\{ P, P^{\mathbb{T}} \right\}. \tag{7.2}$$

This bound immediately implies the first assertion of the lemma. To get the second claim, we decompose

$$\operatorname{tr}(P_\epsilon(s) - P)\, J = \operatorname{tr}(P_\epsilon(s) - P)\, \tilde{J} + \operatorname{tr}(P_\epsilon(s) - P)\, \bar{\chi}_{\mathcal{B}} J, \tag{7.3}$$

and estimate them separately. For the first term, we use (3.1) to deduce that

$$\operatorname{tr}(P_\epsilon(s) - P)\, \tilde{J} = \operatorname{tr}\big(P^{\mathbb{T}}_\epsilon(s) - P^{\mathbb{T}}\big)\, \tilde{J} + \mathcal{O}(e^{-c\mathcal{L}}).$$

The second term on the right hand side of (7.3) is $\mathcal{O}(e^{-c\mathcal{L}})$ by (7.2), so we only need to prove the latter bound. For the time independent case, see [EGS, Lemma 5]. We provide the proof of this fact in our setting for completeness.

Since the argument is identical for both projections, we will only consider the case $P^{\sharp} = P$. Using the fundamental theorem of calculus, we write

$$\begin{aligned} P_\epsilon(s) - P &= -U_\epsilon(s) \left(\int_{-1}^{s} \partial_t \left(U^*_\epsilon(t) P U_\epsilon(t) \right) dt \right) U^*_\epsilon(s) \\ &= \frac{i}{\epsilon} U_\epsilon(s) \left(\int_{-1}^{s} U^*_\epsilon(t) [H(t), P] U_\epsilon(t) dt \right) U^*_\epsilon(s) \\ &= \frac{i\beta}{\epsilon} U_\epsilon(s) \left(\int_{-1}^{s} g(t) U^*_\epsilon(t) [\Lambda_2, P] U_\epsilon(t) dt \right) U^*_\epsilon(s). \end{aligned}$$

We next note that with probability $\geq 1 - \mathrm{e}^{-c\mathcal{L}}$,

$$\left\| [\Lambda_2, P] \chi_{\{x\}} \right\| \leq C e^{c\mathcal{L}} \, |x|^{d+1} \, e^{-c|x_2|}. \tag{7.4}$$

Indeed, (5.9) implies by Markov's inequality that

$$\sum_{x,y \in \mathbb{Z}^2} |x|^{-d-1} \, e^{c|x-y|} \, |P(x,y)| \leq C e^{c\mathcal{L}}$$

with this probability. It implies that on the same probabilistic set,

$$|P(x,y)| \leq C e^{c\mathcal{L}} \, |x|^{d+1} \, \mathrm{e}^{-c|x-y|}.$$

The relation (7.4) now follows by using $\| \Lambda_2 e^{cx_2} \| \leq 1$ for all x with $x_2 < 0$, and then using $\| \bar{\Lambda}_2 e^{cx_2} \| \leq 1$ for the remaining $x \in \mathbb{Z}^2$ together with $[\Lambda_2, P] = -[\bar{\Lambda}_2, P]$.

Combining (7.4) with Proposition B.3, we deduce that

$$\left\| [\Lambda_2, P] U_\epsilon(t) \chi_{\{x\}} \right\| \leq |x|^{d+1} \, e^{c\mathcal{L}} e^{-c|x_2|} \text{ for } |x| \geq \mathcal{L}/3. \tag{7.5}$$

The desired bound (7.2) now follows by combining (7.5) with $\left\| \chi_{\{x\}} e^{c|x_1|} J \right\| \leq C$ for all $x \in \mathbb{Z}^2$. □

This establishes the first step of the proof.

For the second step, we will consider configurations ω for which Theorem 2.1 (and consequently) are applicable. In particular, all bounds below hold with probability $\geq 1 - e^{-c\sqrt{\ell}}$. For a fixed ω, we consider a set $\mathcal{A} = \cup_\gamma \mathcal{T}_\gamma$, where the union is taken over all γ such that $\mathcal{T}_\gamma \cap \Lambda_{\mathcal{L}/4} \neq \emptyset$. Let $\mathcal{B} = \Lambda_{\mathcal{L}/4} \cup \mathcal{A}$. We note that by construction $\Lambda_{\mathcal{L}/4} \subset \mathcal{B} \subset \Lambda_{\mathcal{L}/4+L}$ and

$$\min_\gamma \operatorname{dist} \left(\partial\mathcal{B}, \hat{\mathcal{T}}_\gamma \right) \geq \ell/4 \tag{7.6}$$

(see a paragraph preceding (4.19) for notation), the fact that will be used often in the proof below.

We next decompose $P^{\mathbb{T}}$ into two components $P^{\mathbb{T}} = \mathcal{Q}(-1) + R$ where $\mathcal{Q}(s)$ is the smooth adiabatic projection constructed in Theorem 2.2 (adjusted to the interval $(-1,1)$) and $R := P^{\mathbb{T}} - \mathcal{Q}(-1)$. By Theorem 2.2 we then have that for $s \geq 0$ and $N \in \mathbb{N}$,

$$\|P_\epsilon^{\mathbb{T}}(s) - \mathcal{Q}(0) - R_\epsilon(s)\| \leq C_N \epsilon^N \left(\frac{1}{\Delta^N} + \frac{1}{\delta^{2N+1}}\right) + \mathcal{O}(e^{-c\sqrt{\ell}}),$$

with $R_\epsilon = U_\epsilon(s) R U_\epsilon^*(s)$, and where we used $\mathcal{Q}(s) = \mathcal{Q}(0)$ for $s \geq 0$. Hence

$$\sigma_m = \frac{1}{\beta}\mathrm{tr}((\mathcal{Q}(0)) - \mathcal{Q}(-1))\tilde{J}) + \frac{1}{\beta}\int_0^1 \mathrm{tr}(R_\epsilon(s) - R)\tilde{J})ds + \mathcal{O}(\epsilon^\infty + e^{-c\sqrt{\ell}}).$$

In Proposition 7.2 below we will show that the first term on the right hand side is equal to σ, up to corrections of order $\mathcal{O}(e^{-c\sqrt{\ell}})$.

Thus it remains to bound the second term. It will be convenient to introduce a new scale $\tilde{\ell}$ in addition to ℓ, defined by the modified value for δ, namely $\tilde{\delta} = 7\delta$. We consider the operator $\tilde{Q}_s$ constructed in Lemma 4.9. The important properties of $\tilde{Q}_s$ are that it covers the spectral support of R and that it allows us to control spatial support of R. Let $I = (E - 6\delta, E + 6\delta)$. Using Theorem 2.2.(ii), we have

$$\left\|R - P_I^{\mathbb{T}} R P_I^{\mathbb{T}}\right\| \leq \mathcal{O}(e^{-c\sqrt{\ell}}).$$

By definition of Q_s and the exponential decay of R, we then obtain

$$\left\|R - \sum_\gamma \tilde{Q}_{-1}^\gamma R \tilde{Q}_{-1}^\gamma\right\| \leq \mathcal{O}(e^{-c\sqrt{\ell}})$$

and, using Lemma 4.7.(i), we see that, for $s \geq 0$,

$$\|R_\epsilon(s) - \sum_\gamma Q_s^\gamma R_\epsilon(s) Q_s^\gamma\| \leq \mathcal{O}(\epsilon^\infty + e^{-c\sqrt{\ell}}). \tag{7.7}$$

Since Q_s^γ is supported in $\hat{\mathcal{T}}_\gamma$ (see a paragraph preceding (4.19) for notation), it follows that, up to a small error, $R_\epsilon(s)$ is a sum of terms each supported in a region $\hat{\mathcal{T}}_\gamma$. Let $\hat{U}_\epsilon$ denote the evolution generated by $H_{\mathcal{T}}(s)$, the restriction of $H^{\mathbb{T}}(s)$ to the union of all $\mathcal{T}_\gamma$. Then we have

$$\frac{d}{ds}\left(\hat{U}_\epsilon^*(s) R_\epsilon(s) \hat{U}_\epsilon(s)\right) = \frac{i}{\epsilon}\hat{U}_\epsilon^*(s)[H_{\mathcal{T}}(s) - H(s), R_\epsilon(s)]\hat{U}_\epsilon(s) = \mathcal{O}(\epsilon^\infty + e^{-c\sqrt{\ell}}),$$

thanks to (7.7) and Lemma 4.4. Thus we can approximate

$$\|R_\epsilon(s) - \sum_\gamma \tilde{Q}_s^\gamma \hat{R}_\epsilon(s) \tilde{Q}_s^\gamma\| \leq \mathcal{O}(\epsilon^\infty + e^{-c\sqrt{\ell}}),$$

where $\hat{R}_\epsilon(s) = \hat{U}_\epsilon^*(s) R \hat{U}_\epsilon(s)$.

Let $\mathcal{X}$ be a set

$$\mathcal{X} = \left\{\hat{\mathcal{T}}_\gamma : \left\{\hat{\mathcal{T}}_\gamma \cap \{|x_j| \leq r\right\} \neq \emptyset,\ j = 1,2\right\}, \tag{7.8}$$

then clearly $|\mathcal{X}| \leq CL^2$.

Consider now any $\hat{\mathcal{T}}_\gamma \notin \mathcal{X}$, then either $\mathrm{dist}\left(\hat{\mathcal{T}}_\gamma, \{x \in \mathbb{Z}^2 :\ x_1 = 0\}\right) \geq r$, in which case

$$Q_s^\gamma \tilde{J} = 0,$$

or $\mathrm{dist}\left(\hat{\mathcal{T}}_\gamma, \{x \in \mathbb{Z}^2 :\ x_2 = 0\}\right) \geq r$, in which case

$$Q_s^\gamma \hat{R}_\epsilon(s) Q_s^\gamma = Q_{-1}^\gamma R Q_{-1}^\gamma + \mathcal{O}(\mathrm{e}^{-c\sqrt{\ell}})$$

as the perturbation is constant in that region. Hence, using (7.7) and Lemma 4.4 again (recall that A^{Θ} stands for the restriction of the operator A to the set Θ),

$$\begin{aligned}\operatorname{tr}(R_\epsilon(s)-R)\,\tilde{J} &= \operatorname{tr}\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,\tilde{J} + \mathcal{O}(\epsilon^\infty + e^{-c\sqrt{\ell}})\\ &= \operatorname{tr}\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,J + \mathcal{O}(\epsilon^\infty + e^{-\#\sqrt{\ell}}).\end{aligned}$$

Next we observe, using the cyclicity of the trace for a trace class operator and (7.7), Lemma 2.2.(i), and Lemma 4.4 one more time, that

$$\operatorname{tr}\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,J = -i\operatorname{tr}\Big([H_{\mathcal{T}}(s),\hat{R}_\epsilon(s)]\Big)^{\mathcal{X}}\Lambda_1 + \mathcal{O}(e^{-\#\sqrt{\ell}}).$$

But

$$-i\operatorname{tr}\Big([H_{\mathcal{T}}(s),\hat{R}_\epsilon(s)]\Big)^{\mathcal{X}}\Lambda_1 = \epsilon\partial_s\operatorname{tr}\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}}\Lambda_1.$$

Hence by the fundamental theorem of calculus,

$$\frac{1}{\beta}\int_0^1 \operatorname{tr}\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,Jds = \frac{\epsilon}{\beta}\operatorname{tr}\Big(\Big(\hat{R}_\epsilon(1)\Big)^{\mathcal{X}} - \Big(\hat{R}_\epsilon(0)\Big)^{\mathcal{X}}\Big)\,\Lambda_1 + \mathcal{O}(e^{-c\sqrt{\ell}}),$$

so that we finally get

$$\left\|\frac{1}{\beta}\operatorname{tr}\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,Jds\right\| \le CL^2\frac{\epsilon}{\beta} + \mathcal{O}(e^{-c\ell}).$$

Picking

$$\ell = \beta^{-p}, \quad \text{for} \qquad p > 2d + \frac{1}{2} + \frac{d}{q},$$

in order to satisfy assumptions of Theorem 2.2, see the proof of Theorem 1.7, we get the statement with $p' = (6d + 5/2 + 3d/q)^{-1}$.

The next statement establishes the last step of the proof. The proof shows that the conductance is constant within the mobility gap, in the spirit of [AG].

Proposition 7.2. *We have*

$$\frac{1}{\beta}\operatorname{tr}(\mathcal{Q}(0)-\mathcal{Q}(-1))\,\tilde{J} = \sigma + \mathcal{O}(e^{-c\sqrt{\ell}}),$$

where σ is defined in (1.14).

Proof. We note that by locality of H, $\tilde{J} = i\chi_{\mathcal{B}}[H^{\mathbb{T}}(r),\Lambda_1]$. By the fundamental theorem of calculus,

$$\frac{1}{\beta}\operatorname{tr}(\mathcal{Q}(0)-\mathcal{Q}(-1))\,\tilde{J} = \frac{1}{\beta}\int_{-1}^0 \operatorname{tr}\left(\partial_r\mathcal{Q}(r)i\chi_{\mathcal{B}}[H^{\mathbb{T}}(r),\Lambda_1]\right)dr.$$

We claim that

$$\frac{1}{\beta}\operatorname{tr}\left(\partial_r\mathcal{Q}(r)i\chi_{\mathcal{B}}[H^{\mathbb{T}}(r),\Lambda_1]\right) = i\dot{g}(r)\operatorname{tr}\left(\mathcal{Q}(r)[[\mathcal{Q}(r),\Lambda_1],[\mathcal{Q}(r),\Lambda_2]]\chi_{\mathcal{B}}\right) + \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.9}$$

Indeed, let $\hat{\Lambda}_1(r) = \mathcal{Q}(r)\Lambda_1\bar{\mathcal{Q}}(r) + \bar{\mathcal{Q}}(r)\Lambda_1\mathcal{Q}(r)$. We have

$$\begin{aligned}\operatorname{tr}\left(\partial_r\mathcal{Q}(r)i\chi_{\mathcal{B}}[H^{\mathbb{T}}(r),\Lambda_1]\right) &= \operatorname{tr}\Big(\partial_r\mathcal{Q}(r)i\chi_{\mathcal{B}}[H^{\mathbb{T}}(r),\hat{\Lambda}_1(r)]\Big) + \mathcal{O}(e^{-c\sqrt{\ell}})\\ &= \operatorname{tr}\Big(-i[H^{\mathbb{T}},\partial_r\mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\Big) + \mathcal{O}(e^{-c\sqrt{\ell}})\\ &= \operatorname{tr}\Big(i[\dot{H}^{\mathbb{T}},\mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\Big) + \mathcal{O}(e^{-c\sqrt{\ell}})\\ &= \operatorname{tr}\Big(i[\beta\dot{g}(r)\Lambda_2,\mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\Big) + \mathcal{O}(e^{-c\sqrt{\ell}}),\end{aligned}$$

where in the first step we used $\mathcal{Q}(r)\partial_r\mathcal{Q}(r)\mathcal{Q}(r) = \bar{\mathcal{Q}}(r)\partial_r\mathcal{Q}(r)\bar{\mathcal{Q}}(r) = 0$ and in the third step we used $[H^{\mathbb{T}},\mathcal{Q}(r)] = \mathcal{O}(e^{-c\sqrt{\ell}})$. We also repeatedly used that commuting $\chi_{\mathcal{B}}$ with other operators under the trace contributes $\mathcal{O}(e^{-c\sqrt{\ell}})$ by virtue of (7.6) and the location of support of the involved operators. The relation (7.9) now follows, since $\hat{\Lambda}_1 = [\mathcal{Q}(r), [\mathcal{Q}(r), \Lambda_1]]$.

The implication is

$$\frac{1}{\beta}\mathrm{tr}\Big(\mathcal{Q}(0) - \mathcal{Q}(-1))\tilde{J}\Big) = i\int_{-1}^{0} \dot{g}(r)\mathrm{tr}(\mathcal{Q}(r)[[\mathcal{Q}(r), \Lambda_1], [\mathcal{Q}(r), \Lambda_2]])\,\chi_{\mathcal{B}} + \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.10}$$

We now define

$$\mathrm{Ind}_{\mathcal{L}}\,(\mathcal{Q}) = \mathrm{tr}(\mathcal{Q}[[\mathcal{Q}, \Lambda_1], [\mathcal{Q}, \Lambda_2]])\,\chi_{\mathcal{B}}. \tag{7.11}$$

For $\mathbb{Z}^2$ geometry without the cutoff function $\chi_{\mathcal{B}}$, the index (when it is well defined) is known to be integer valued and additive. I.e., for orthogonal projections Q, R with a compact R, $\mathrm{Ind}_\infty(Q + R) = \mathrm{Ind}_\infty(Q) + \mathrm{Ind}_\infty(R)$, provided $Q + R$ is a projection, [ASS, Proposition 2.5]. The argument in [ASS] assumes that the underlying projections are covariant and that their kernels enjoys good decay properties. The latter hold in the random setting; one can also relax the covariance requirement for such models, [EGS]. Moreover, $\lim_{\mathcal{L}\to\infty} \mathrm{Ind}_{\mathcal{L}}\,(P)$ exists and we have

$$\lim_{\mathcal{L}\to\infty} \mathrm{Ind}_{\mathcal{L}}\,(P) = \sigma, \tag{7.12}$$

[ASS, Section 6]. In fact, using (5.9) it is not hard to show that

$$|\sigma - \mathrm{Ind}_{\mathcal{L}}\,(P)| \le \mathcal{O}(e^{-c\mathcal{L}}) \text{ and } \left|\mathrm{Ind}_{\mathcal{L}}\,(P) - \mathrm{Ind}_{\mathcal{L}}\left(P^{\mathbb{T}}\right)\right| \le e^{-c\mathcal{L}}. \tag{7.13}$$

Next we observe that although $P^{\mathbb{T}}$ and $\mathcal{Q}(-1)$ do not commute, we have $\left\|[P^{\mathbb{T}}, \mathcal{Q}(-1)]\right\| \le e^{-c\sqrt{\ell}}$. Hence there exists a pair of operators $\hat{P}^{\mathbb{T}}, \hat{\mathcal{Q}}(-1)$ such that $[\hat{P}^{\mathbb{T}}, \hat{\mathcal{Q}}(-1)] = 0$ and $\left\|P^{\mathbb{T}} - \hat{P}^{\mathbb{T}}\right\| \le e^{-c\sqrt{\ell}}$, $\left\|\mathcal{Q}(-1) - \hat{\mathcal{Q}}(-1)\right\| \le e^{-c\sqrt{\ell}}$, [KaS]. Moreover, applying the compression procedure used to get a projection Q_s from near projection $Q_N(s)$ in the proof of Lemma 4.9, without loss of generality we can assume that $\hat{P}^{\mathbb{T}}, \hat{\mathcal{Q}}(-1)$ are in fact projections. Let $\check{R} = \hat{P}^{\mathbb{T}} - \hat{\mathcal{Q}}(-1)$. Since $\|\mathcal{Q}(-1)R\| \le e^{-c\sqrt{\ell}}$, we conclude that $\hat{\mathcal{Q}}(-1)\check{R} = 0$. In particular, the additivity of index is applicable for $\hat{\mathcal{Q}}(-1)$ and $\check{R}$ and yields

$$\left|\mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(-1)\right) + \mathrm{Ind}_{\mathcal{L}}\,(\check{R}) - \mathrm{Ind}_{\mathcal{L}}\left(\hat{P}^{\mathbb{T}}\right)\right| \le e^{-c\sqrt{\ell}}. \tag{7.14}$$

By construction, we deduce that

$$|\mathrm{Ind}_{\mathcal{L}}\,(Y_i) - \mathrm{Ind}_{\mathcal{L}}\,(Z_i)| \le e^{-c\sqrt{\ell}}, \quad i = 1, 2, 3, \tag{7.15}$$

where $Y_1 = \check{R}$, $Z_1 = R$, $Y_2 = \hat{\mathcal{Q}}(-1)$, $Z_2 = \mathcal{Q}(-1)$, $Y_3 = \hat{P}^{\mathbb{T}}$ and $Z_3 = P^{\mathbb{T}}$. In addition, since $\mathcal{Q}(r)$ is continuous, we conclude that

$$\mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(r)\right) = \mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(-1)\right) + \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.16}$$

Putting together (7.13)–(7.16) we see that the statement follows if we can show that

$$\mathrm{Ind}(R) = \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.17}$$

To establish this bound we observe that

$$\mathrm{Ind}(R) = \mathrm{Ind}(R^{\mathcal{X}}) + \mathcal{O}(e^{-c\sqrt{\ell}}),$$

where $\mathcal{X}$ was defined in (7.8), just as in the argument used in the second step above. But

$$\mathrm{Ind}(R^{\mathcal{X}}) = i\mathrm{tr}R^{\mathcal{X}}[[R^{\mathcal{X}}, \Lambda_1], [R^{\mathcal{X}}, \Lambda_2]],$$

and the right hand side is $\mathcal{O}(e^{-c\sqrt{\ell}})$ using $R^{\mathcal{X}}\left(\mathbb{I} - R^{\mathcal{X}}\right) = \mathcal{O}(e^{-c\sqrt{\ell}})$ and cyclicity of the trace. □

Appendix A. Hybridization in $1D$

Here we will consider the analytic family of Hamiltonians of the form

$$H_\beta = H_o + \beta W \tag{A.1}$$

acting on $\ell^2(\mathbb{Z})$ (we will only consider the phenomenon in one dimension).

The assumptions on H_o in this section will be stronger than elsewhere in the paper, too. In particular, the random operator H_o will be assumed to be the standard Anderson Hamiltonian, $H_A = \Delta + V_\omega$, with $V_\omega(i) = \omega_i$, the i.i. random coupling variables distributed according to the Borel probability measure $\mathbb{P} := \otimes_{\mathbb{Z}} P_0$. We will assume that the single-site distribution P_0 is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$. The corresponding Lebesgue density μ will be further assumed to be bounded with support $\mathrm{supp}(\mu) \subset [0,1]$. In addition we will assume that the single-site probability density is bounded away from zero on its support.

With such assumptions on H_o we can make stronger statements about the system's behavior. These are encapsulated in

A.1. **Background results.** Important properties concerning the behavior of H_o are encapsulated in Theorems A.1, A.4, A.6, and A.5 below.

Theorem A.1 (Eigenfunction localization)**.** *There exists $\nu > 0$ such that for $\mathbb{P}$-almost every ω and Lebesgue a.e. $\beta \in [0,1]$, $\sigma(H_\omega)$ is simple and for each $E \in \sigma(H_\omega)$ there is a localization center $x_E(\omega, \beta)$ such that the normalized eigenfunction $\psi_E(\cdot, \omega)$ satisfies, for all $y \in \mathbb{Z}$*

$$|\psi_E(y,\omega)|^2 \le A(\omega)\langle x_E(\omega)\rangle^2 \mathrm{e}^{-\nu|y-x_E(\omega)|}, \tag{A.2}$$

with $A(\cdot) \in L^1(\Omega, \mathbb{P})$.

This statement is a consequence of [AW, Theorems 5.8, 7.4, and 12.11].

We will need the probabilistic version of this assertion. For this, we introduce

Definition A.2. For $\omega \in \Omega$ and a pair (ν, θ) of positive parameters, we will say that H_ω is (ν, θ)-localized if for each $E \in \sigma(H_\omega)$

$$|\psi_E(y,\omega)|^2 \le \frac{1}{\theta}\langle x_E(\omega)\rangle^2 \mathrm{e}^{-\nu|y-x_E(\omega)|}. \tag{A.3}$$

Then Theorem A.1 implies, via Markov's inequality, that

$$\mathbb{P}\left(\{\omega \in \Omega :\ H_\omega \text{ is } (\nu,\theta)\text{-localized}\}\right) \le 1 - C\theta \tag{A.4}$$

for some $C > 0$.

We now state a consequence of (ν, θ)-localization. This is essentially repetition of the argument in [DJLS, Theorem 7.1 and Lemma 7.2] that tracks the dependence on θ. We omit the ω dependence in what follows since the result is deterministic.

Theorem A.3. *Assume that H is (ν, θ)-localized. Then there exists $C_\nu > 0$ and $E \in \sigma(H)$ such that $|\psi_E(0)|^2 \ge \frac{-C_\nu}{\ln\theta}$ and $|x_E| \le \frac{-\ln\theta}{9C_\nu}$.*

Proof. We first observe that for any $L \in \mathbb{N}$ and $E \in \sigma(H)$ we have

$$\begin{aligned}
\sum_{\substack{y\in\mathbb{Z}:\\ |y-x_E|\ge\frac12(|x_E|+L)}} |\psi_E(y)|^2 &\le \frac{\langle x_E\rangle^2}{\theta} \sum_{\substack{y\in\mathbb{Z}:\\ |y-x_E|\ge\frac12(|x_E|+L)}} \mathrm{e}^{-\nu|y-x_E|} \\
&= \frac{\langle x_E\rangle^2}{\theta} \sum_{\substack{u\in\mathbb{Z}:\\ |u|\ge\frac12(|x_E|+L)}} \mathrm{e}^{-\nu|u|} = \frac{\langle x_E\rangle^2}{\theta}\mathrm{e}^{-\frac{\nu}{2}(L+|x_E|)}\frac{2}{1-\mathrm{e}^{-\nu}} \le \frac{C_\nu}{\theta}\mathrm{e}^{-\frac{\nu}{2}(L+|x_E|)}
\end{aligned} \tag{A.5}$$

for some $C_\nu > 0$.

We next note that by orthonormality and completeness of $\{\psi_E\}$ we have

$$\sum_{E\in\sigma(H)} |\psi_E(y)|^2 = 1, \quad y \in \mathbb{Z}; \tag{A.6}$$

$$\sum_{y\in\mathbb{Z}} |\psi_E(y)|^2 = 1, \quad E \in \sigma(H); \tag{A.7}$$

Hence, using (A.5), there exists $K_\nu > 0$ such that

$$\begin{aligned} 4L+1 &= \sum_{|y|\le 2L} \sum_{E\in\sigma(H)} |\psi_E(y)|^2 \ge \sum_{|y|\le 2L} \sum_{\substack{E\in\sigma(H):\\ |x_E|\le L}} |\psi_E(y)|^2 = \sum_{\substack{E\in\sigma(H):\\ |x_E|\le L}} \left(1 - \sum_{|y|>2L} |\psi_E(y)|^2\right) \\ &\ge \#\left\{E\in\sigma(H):\ |x_E|\le L\right\}\left(1 - \frac{C_\nu}{\theta}\mathrm{e}^{-\frac{\nu}{2}L}\right) \ge \frac{1}{2}\#\left\{E\in\sigma(H):\ |x_E|\le L\right\} \end{aligned} \tag{A.8}$$

for $L \ge K_\nu |\ln\theta|$.

This bound together with (A.5) imply that for $L \ge K_\nu |\ln\theta|$ we have

$$\begin{aligned} \sum_{|y|\le L} \sum_{\substack{E\in\sigma(H):\\ |x_E|>3L}} |\psi_E(y)|^2 &\le \sum_{k=4}^{\infty} \#\left\{E\in\sigma(H):\ |x_E|\le kL\right\} \frac{C_\nu}{\theta}\mathrm{e}^{-\frac{\nu k L}{2}} \\ &\le \frac{9C_\nu}{\theta} L \sum_{k=4}^{\infty} k\mathrm{e}^{-\frac{\nu k L}{2}} < \frac{1}{2} \end{aligned} \tag{A.9}$$

for $L \ge M_\nu |\ln\theta|$ with some $M_\nu > 0$.

Using this estimate, we get

$$1 = \sum_{E\in\sigma(H)} |\psi_E(0)|^2 \le \sum_{\substack{E\in\sigma(H):\\ |x_E|\le 3L}} |\psi_E(0)|^2 + \frac{1}{2},$$

for $L \ge M_\nu |\ln\theta|$, so

$$\frac{1}{2} \le \sum_{E\in\sigma(H)} |\psi_E(0)|^2 \le \sum_{\substack{E\in\sigma(H):\\ |x_E|\le 3L}} |\psi_E(0)|^2,$$

and since $\#\left\{E\in\sigma(H):\ |x_E|\le 3L\right\} \le 13L$ by (A.9), we deduce that there exists $C_\nu > 0$ and $E \in \sigma(H)$ such that

$$|\psi_E(0)|^2 \ge \frac{1}{27L} = \frac{-C_\nu}{\ln\theta}, \quad |x_E| \le \frac{-\ln\theta}{9C_\nu}.$$

□

Let $\Lambda = [0, L] \cap \mathbb{Z}$. By H_β^Λ we will denote the (Dirichlet) restriction of H_β to Λ.

While Theorem A.1 asserts that the $\sigma(H_\beta^\Lambda)$ is a.s. simple, we will need a stronger statement about its spectrum.

Theorem A.4 (Two-sided Wegner estimate)**.** *There exist $L_0 > 0$ and constants $C_+ \ge C_- > 0$ such that for any subinterval J of $[0, 2]$ we have*

$$C_- |J| L \le \mathbb{E}\left(\mathrm{tr}\chi_J\left(H_\beta^\Lambda\right)\right) \le C_+ |J| L, \tag{A.10}$$

provided $L > L_0$.

The upper bound is well known, see e.g., [AW, Corollary 4.9]. The lower bound was recently established in [G, Theorem 1.1].

We will also need the following extension of the upper Wegner bound, known as the Minami estimate:

Theorem A.5 (Minami estimate)**.** *For any $n \in \mathbb{N}$, there exists $C_n > 0$ such that*

$$\mathbb{P}\left(\mathrm{tr}\chi_J\left(H_\beta^\Lambda\right) \ge n\right) \le C_n \left(C_+ |J| L\right)^n, \tag{A.11}$$

with C_+ defined in Theorem A.11.

See e.g., [AW, Theorem 17.11]. This assertion implies

Theorem A.6. *Let $\delta > 0$ and let $\mathcal{E}_\omega$ be an event*

$$\mathcal{E}_{\omega,\beta} := \left\{\sigma(H^{\Lambda}_{\beta}) \text{ is } \delta\text{-level spaced on } \Lambda\right\}.$$

Then there exists $C > 0$ such that

$$\mathbb{P}\left(\mathcal{E}_\omega\right) \geq 1 - C\delta L^2.$$

This statement is essentially [KM, Lemma 2], in the formulation given in [EK1, Lemma B.1].

A.2. **Hybridization.** We will henceforth assume that W in (A.1) satisfies $W \geq 0$, $\operatorname{supp}(W)$ is compact, and that there exists a point $z \in \mathbb{Z}$ such that $W \geq \chi_{\{z\}}$. Since H_ω is ergodic, without loss of generality we will also assume that $z = 0$.

Our main focus here will be on what we can show for a large but finite system size $\mathcal{L}$. Specifically, we will consider a box $\Lambda_{\mathcal{L}} := [-\mathcal{L}, \mathcal{L}] \cap \mathbb{Z}$, and $\mathcal{L}$ large enough so that we can decompose $\Lambda_{\mathcal{L}}$ into two distinct sets $\Lambda_{\mathcal{L}} = \Lambda_{in} \cup \Lambda_{out}$, such that $\operatorname{supp}(W) \subset \Lambda_{in}$ and $\operatorname{dist}(\operatorname{supp}(W), \Lambda_{out}) \geq \sqrt{\mathcal{L}}$. We will denote by $\langle l_\pm, r_\pm\rangle$ the pair of edges connecting Λ_{in} and Λ_{out} (i.e., $l_\pm \in \Lambda_{in}$, $r_\pm \in \Lambda_{out}$ and $|r_\pm - l_\pm| = 1$).

We will use the concept of n.u. (ν, θ)-localizing Hamiltonian, cf. Definition 6.2.

Definition A.7. Let $L \in \mathbb{N}$. We will say that H is *hybridization-susceptible* on the scale $\mathcal{L}$ if there exists $\theta \in (0, \frac{1}{4})$ and eigenpairs (E_{in}, ϕ_{in}), (E_{out}, ϕ_{out}) for $H_{in} := H^{\Lambda_{in}}$ and $H_{out} := H^{\Lambda_{out}}$, respectively, such that

$$|E_{out} - E_{in}| \leq \theta^3 L^{-1} \text{ and } \min\left(\operatorname{dist}\left(\sigma(H_{out}) \setminus E_{out}, E_{out}\right), \operatorname{dist}\left(\sigma(H_{in}) \setminus E_{in}, E_{in}\right)\right) \geq \theta^2 L^{-1}. \tag{A.12}$$

Definition A.8. Let $\ell \in \mathbb{N}$, $\nu > 0$, and let $\Psi \subset \mathbb{Z}$. We will say that an eigenvector ψ for H^Ψ is (ℓ, ν)-*bulk* if there exists $x \in \Psi$ such that $[-\ell, \ell] + x \subset \Psi$ and $\left\|\bar{\chi}_{[-\ell/2, \ell/2]+x}\psi\right\| \leq \mathrm{e}^{-\nu\ell/2}$.

We will shorthand $\Lambda_{\mathcal{L}}$ to Λ whenever the scale $\mathcal{L}$ is fixed. By Ω_Λ we will denote the restriction of Ω to Λ. By $\Lambda_\pm$ we will denote the sets $\Lambda_\pm = [-\frac{\sqrt{\mathcal{L}}}{10}, \frac{\sqrt{\mathcal{L}}}{10}] + l_\pm$. Our first result is

Proposition A.9. *Let $\theta \in (0, \frac{1}{4})$. Consider the following events $\mathcal{A}_\omega, \mathcal{B}_\omega, \mathcal{C}_\omega, \mathcal{D}_\omega, \mathcal{E}_\omega \subset \Omega_\Lambda$:*

$$\mathcal{A}_\omega = \left\{H_0^{\Lambda_\sharp} \text{ is n.u. } \left(\nu, \theta^4\right)\text{-localizing for } \sharp = in, out\right\};$$

$$\mathcal{B}_\omega = \left\{H_0^{\Lambda_\pm} \text{ are } \tfrac{\theta}{\mathcal{L}}\text{-gapped}\right\};$$

$$\mathcal{C}_\omega = \left\{H_0^{\Lambda} \text{ is hybridization-susceptible on the scale } \mathcal{L}\right\};$$

$$\mathcal{C}_\omega \supset \mathcal{D}_\omega = \left\{\text{The corresponding eigenvectors } \phi_\sharp \text{ are } \left(\tfrac{\sqrt{\mathcal{L}}}{10}, \tfrac{1}{\ln \mathcal{L}}\right)\text{-bulk}\right\};$$

$$\mathcal{D}_\omega \supset \mathcal{E}_\omega = \Big\{\text{the eigenpair } (E_{in}, \psi) \text{ with localization center } x_{in} \text{ satisfies } |x_{in}| < \ln^2\theta \text{ and } \left|\dot{E}_{in}(0)\right| > \ln^{-2}\theta\Big\}.$$

Let $\mathcal{F}_\omega = \mathcal{A}_\omega \cap \mathcal{B}_\omega \cap \mathcal{E}_\omega$. Then there exists $\theta_o \in (0,1)$ and $\mathcal{L}_o \in \mathbb{N}$ such that, for all $\theta < \theta_o$ and $\mathcal{L} > \mathcal{L}_o$, $\mathbb{P}\left(\mathcal{F}_\omega\right) \geq c\theta^3$.

Remark A.10. These conditions imply, via Hellmann-Feynman theorem, that for the decoupled system the eigenvalues $E_{in}(\beta)$ and $E_{out}(\beta)$ level cross for some $\beta \in I$, see the proof of Lemma A.18 below. This will play an important role in our analysis down the road.

We will now prove the above proposition. We first show a part of it, namely

Lemma A.11. *Let $\theta > 0$ be a small parameter. Consider the following events (recall Definition* **??***) $\tilde{\mathcal{A}}_\omega$ and $\tilde{\mathcal{B}}_\omega$:*

$$\tilde{\mathcal{A}}_\omega = \left\{\omega \in \Omega_{\Lambda_{in}} : \ H_0^{\Lambda_{in}} \text{ is n.u. } \left(\nu, \theta^4\right) \text{-localizing and } \theta^4\mathcal{L}^{-1}\text{-level spaced}\right\},$$

$$\tilde{\mathcal{B}}_\omega \quad = \quad \left\{\omega \in \Omega_{\Lambda_{in}} : \ \exists \ i \text{ s.t. } |x_{in}| < \ln^2\theta, \ \left|\dot{E}_{in}(0)\right| > c\ln^{-2}\theta\right\}$$

Then $\mathbb{P}\left(\tilde{\mathcal{A}}_\omega \cap \tilde{\mathcal{B}}_\omega\right) \geq 1 - K\theta^4$.

Proof. We first note that $\mathbb{P}\left(\tilde{\mathcal{A}}_\omega\right) \geq 1 - C\theta^4$ by (A.4) and Theorem A.6. We will now assume that $\omega \in \tilde{\mathcal{A}}_\omega$.

The existence of a level E_{in} supported around the origin and susceptible to W, follows from the Hellmann–Feynman theorem. Indeed, let $\mathcal{S}$ be a subset of $\sigma(H_{in})$ such that for each $E_{in} \in \mathcal{S}$ the corresponding localization center x_{in} satisfies $|x_{in}| \geq \ln^2\theta$. Then we have

$$\sum_{E_{in}\in\mathcal{S}} \operatorname{tr}\big(P_{E_{in}}\chi_{\{0\}}\big) \leq C\theta^2, \tag{A.13}$$

where $P_{E_{in}}$ denotes the spectral projection of H_{in} onto E_{in}. This implies that

$$\sum_{E_{in}\notin\mathcal{S}} \operatorname{tr}\big(P_{E_{in}}\chi_{\{0\}}\big) \geq \frac{1}{2} \tag{A.14}$$

for θ sufficiently small. We now note that $|\sigma(H_R) \setminus \mathcal{S}| \leq C\ln^2\theta$, see [DJLS, Lemma 7.2]. Hence, by the pigeonhole principle, there exists E_{in} such that

$$\operatorname{tr}(P_{E_{in}}W) \geq \operatorname{tr}\big(P_{E_{in}}\chi_{\{0\}}\big) \geq c\ln^{-2}\theta.$$

But, by the Hellmann–Feynman theorem, denoting by $E_{in}(\beta)$ the analytic family of eigenvalues associated with $H(\beta)$ and satisfying $E_{in}(0) = E_{in}$,

$$\dot{E}_{in}(0) = \operatorname{tr}(P_{E_{in}}W),$$

and the result follows.

□

Now we can complete

Proof of Proposition A.9. We first note that the lower Wegner estimate, Theorem A.4, and the statistical independence of H_{in} and H_{out} implies that

$$\operatorname{dist}\left(\sigma(H_{out}), E_{in}\right) \leq \theta^3\mathcal{L}^{-1}$$

with probability $c\theta^3$ (on $\Omega_{\Lambda_{out}}$), i.e., the first part of (A.12) holds with this probability. In fact, with the same probability we can ensure that the corresponding eigenvector ϕ_{out} is $(\frac{\sqrt{\mathcal{L}}}{10}, \frac{1}{\ln\mathcal{L}})$-bulk. Since $|x_{in}| < \ln^2\theta$, we deduce that ϕ_{in} is $(\frac{\sqrt{\mathcal{L}}}{10}, \frac{1}{\ln\mathcal{L}})$-bulk as well. Theorem A.6 ensures that $\mathbb{P}\left(\mathcal{B}_\omega\right) \geq 1 - \theta^4$ for $\mathcal{L}$ sufficiently large. On the other hand, the Minami estimate, Theorem A.5, applied for $H_o^{\Lambda_{out}}$ and the energy $E = E_{in}$, implies that the second part of (A.12) holds with probability $1 - C\theta^4$ (on $\Omega_{\Lambda_{out}}$). Putting everything together yields the desired result for θ sufficiently small (specifically we need to ensure $c\theta^3 > 2C\theta^4$). □

We recall that Λ_{out} consists of two disconnected parts $\Lambda_{out}^{\pm}$. Without loss of generality, we will henceforth assume that $E_{out} \in \sigma(H^{\Lambda_{out}^+})$ (i.e., the corresponding eigenvalue is supported on Λ_{out}^+), and will denote by $\langle l_+, r_+\rangle$ the edge connecting Λ_{in} and Λ_{out}^+.

Lemma A.12. *Let $\Psi = [-(\ln\mathcal{L})^2, (\ln\mathcal{L})^2] + l_+$ and let $\tilde{I} := E_{in} + [-\frac{\theta}{\sqrt{\mathcal{L}}}, \frac{\theta}{\sqrt{\mathcal{L}}}]$. In the notation of Proposition A.9 consider $\mathcal{F}_\omega \subset \mathcal{E}_\omega$ given by*

$$\mathcal{G}_\omega = \left\{\omega \in \mathcal{F}_\omega : \ \operatorname{tr}\chi_{\tilde{I}}(H_0^\Psi) = 0\right\}.$$

Then there exists $\theta_o \in (0,1)$ and $\mathcal{L}_o \in \mathbb{N}$ such that, for all $\theta < \theta_o$ and $\mathcal{L} > \mathcal{L}_o$, $\mathbb{P}\left(\mathcal{G}_\omega\right) \geq 1 - \theta^4$.

Proof. The result follows from the upper Wegner estimate, Theorem A.4. □

Lemma A.13. *Let*
$$\mathcal{H}_\omega = \left\{\omega \in \Omega : \ \left|\langle \delta_{r_+}, \left(H^\Psi\right)^{-1} \delta_{l_+}\rangle - 1\right| \geq \tfrac{1}{\mathcal{L}}\right\} \text{ and } H^\Psi \text{ is n.u. } \left(\nu, \theta^4\right)\text{-localizing.}$$
Then $\mathbb{P}\left(\mathcal{H}_\omega\right) \geq 1 - \theta^4$.

Proof of Lemma A.13. By (A.4), it suffices to show the first property.

Let $G(x,y) = \langle \delta_x, \left(H^\Psi\right)^{-1} \delta_y\rangle$. We first observe that thanks to the geometric resolvent identity (or just directly in [AW, Eq. 12.7]),
$$G(l_+, r_+) = \hat{G}^{-1}(l_+, r_+) G(r_+, r_+), \tag{A.15}$$
where $\hat{G}(x,y) = \langle \delta_x, \left(\hat{H}^\Psi\right)^{-1} \delta_y\rangle$ and $\hat{H}^\Psi$ is obtained from H^Ψ by the removal of the $\langle l_+, r_+\rangle$ bond, i.e., $\hat{H}^\Psi = H^\Psi - \Gamma_+$. We have
$$\frac{1}{\hat{G}(0,0)G(1,1) - 1} = \frac{1}{G(1,1)} \frac{1}{\hat{G}(0,0) - G^{-1}(1,1)} = -\frac{\tilde{G}(1,1)}{G(1,1)},$$
where $\tilde{G}(x,y) := \langle \delta_x, \left(\tilde{H}^\Psi\right)^{-1} \delta_y\rangle$ with $\tilde{H}^\Psi = H^\Psi - \hat{G}(0,0)\chi_{\{r\}}$. The important fact here is to note that $\hat{G}(0,0)$ is independent of ω_1 random variable. This independence allows us to conclude that
$$\mathbb{E}_{\omega_1} \left|\tilde{G}_\beta(1,1)\right|^s \leq C_s, \quad s \in (0,1).$$
On the other hand, under our conditions on probability distribution μ, we also have (see [AW, Theorem 12.8]
$$\mathbb{E}\left|G(0,0)\right|^{-s} \leq C_s, \quad s \in (0,1).$$
Combining these two bounds and using the Hölder inequality, we deduce that
$$\mathbb{E}\left|\frac{1}{\hat{G}(0,0)G(1,1) - 1}\right|^s \leq C_s, \quad s \in (0, 1/2),$$
from which the assertion follows. □

Proposition A.9 and Lemma A.12 let us conclude that the following holds for the coupled system. We will shorthand $H_\beta = H_\beta^{\Lambda_\mathcal{L}}$.

Proposition A.14. *Let* $P_{hyb}(\beta) = \sum_{\sharp=\{in,out\}} |\phi_\sharp\rangle\langle\phi_\sharp|$*, where* $\phi_{in,out}$ *were introduced in Definition A.7 and Lemma A.9. Let* $I := E_{in} + [-\frac{\theta^3}{\mathcal{L}}, \frac{\theta^3}{\mathcal{L}}]$*. Then for* $\omega \in \mathcal{G}_\omega \cap \mathcal{H}_\omega$ *and* $\beta \in J := [-\frac{\theta^2}{4\mathcal{L}}, \frac{\theta^2}{4\mathcal{L}}]$*, we have*

(i) $\sigma(H_\beta) \cap I = \{\lambda_-(\beta), \lambda_+(\beta)\}$ *where* $\lambda_\pm(\beta)$ *are real analytic in* β *for* $\beta \in J$*;*

(ii) $\chi_{\hat{I}}(H_\beta) = \chi_I(H_\beta)$ *for* $\hat{I} := E_{in} + [-\frac{\theta^2}{2\mathcal{L}}, \frac{\theta^2}{2\mathcal{L}}]$*;*

(iii) $\operatorname{tr}\chi_I(H_\beta) = 2$*;*

(iv) $\operatorname{tr}\left(\chi_{I^c}(H_\beta)P_{hyb}(\beta)\right) \leq \exp\left(-\frac{\sqrt{\mathcal{L}}}{\ln^2 \mathcal{L}}\right)$*;*

(v) $\operatorname{tr}\chi_I(H_\beta)W \geq \theta^2$ *for* $\beta \in J$*.*

Proof. By Lemma B.2 we deduce that
$$\operatorname{dist}\left(\sigma(H_o), E_\sharp\right) \leq \exp\left(-\frac{\sqrt{\mathcal{L}}}{\ln^2 \mathcal{L}}\right)$$
for $\sharp = in, out$. The property A.14.(ii) is proved by the same argument as one that is used in the proof of Theorem 2.1.(ii). The statements A.14.(i), A.14.(iii)–A.14.(v) now follow from the standard perturbation theory. □

We now draw some conclusions about the coupled system H_β with $\omega \in \mathcal{G}_\omega$.

Lemma A.15. *Under the same conditions as in the previous lemma, the operator*

$$R(\beta,\lambda) := \bar{P}_{hyb}(\beta)\,(H_\beta-\lambda)\,\bar{P}_{hyb}(\beta) \tag{A.16}$$

is invertible on the range of $\bar{P}_{hyb}(\beta)$ *for all* $\lambda \in I$*, with*

$$\left\|(R(\beta,\lambda))^{-1}\right\| \le \tfrac{8\mathcal{L}}{\theta^2} \tag{A.17}$$

for such λ*.*

Proof. This is proven using the same technique as in the proof of Lemma 4.10. □

Proposition A.14 allow us to introduce a concept of the eigenvalue hybridization:

Definition A.16. Suppose that H_β satisfies the assumptions of Proposition A.14. For β such that $\lambda_-(\beta) \neq \lambda_+(\beta)$, let $P_\pm(\beta) := \chi_{\{\lambda_\pm(\beta)\}}(H_\beta)$, and let $P_-^{dec}(\beta) := \chi_{\{E_{out}\}}(H_{out})$, $P_+^{dec}(\beta) := \chi_{\{E_{in})\}}(H_{in})$ (in particular, P_-^{dec} is in fact β-independent). We will say that the analytic families $\lambda_\pm(\beta)$, defined in Proposition A.14 do not hybridize if

$$\left\|P_-(\beta)P_+^{dec}(\beta)\right\| < \theta^2 \tag{A.18}$$

for all $\beta \in I$ for which $\lambda_-(\beta) \neq \lambda_+(\beta)$, and hybridize otherwise.

Remark A.17. Note that (A.18) implies, via Proposition A.14.(iv), that

$$\left\|P_+(\beta)P_-^{dec}(\beta)\right\| < \theta^2 + \exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right), \tag{A.19}$$

so (A.18) and (A.19) are essentially symmetrical.

We claim that in our setting, $\lambda_\pm(\beta)$ *have* to hybridize in J.

Theorem A.18. *In the notation of Proposition A.14, let* $\omega \in \mathcal{F}_\omega$*. Then* $\lambda_\pm(\beta)$ *hybridize in* I*. In addition,*

$$\min_{\beta\in I} |\lambda_+(\beta)-\lambda_-(\beta)| > \frac{\theta^4}{\mathcal{L}^2}. \tag{A.20}$$

Proof. Suppose in contradiction that the eigenvectors do not hybridize. Wlog let us assume that $\lambda_-(0) > \lambda_+(0)$. Since (A.18) holds, we deduce that in fact

$$\operatorname{tr}\left(P_\mp(0)P_\pm^{dec}(0)\right) < \exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right)$$

by Proposition A.14.(iv). By the Hellmann–Feynman theorem

$$\begin{aligned}\dot{\lambda}_-(\beta) &= \operatorname{tr}(P_-(\beta)W) \le \operatorname{tr}(P_-(\beta)P_{hyb}(\beta)W) + \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right)\\ &= \operatorname{tr}\left(P_-(\beta)P_+^{dec}(\beta)W\right) + \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right) \le \|W\|\,\theta^2 + \exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right) \le \theta \end{aligned}\tag{A.21}$$

for θ sufficiently small and $\mathcal{L}$ sufficiently large. On the other hand,

$$\begin{aligned}\dot{\lambda}_+(\beta) &= \operatorname{tr}(P_+(\beta)W) \ge \operatorname{tr}(P_+(\beta)P_{hyb}(\beta)W) - \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right)\\ &= \operatorname{tr}\left(P_+(\beta)P_+^{dec}(\beta)W\right) - \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right)\\ &\ge \operatorname{tr}\left(P_+^{dec}(\beta)W\right) - \|W\|\,\theta^3 - \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right)\\ &\ge \dot{E}_{in}(\beta) - \|W\|\,\theta^3 - \|W\|\exp\left(-\tfrac{\sqrt{\mathcal{L}}}{\ln^2\mathcal{L}}\right) \ge \frac{2}{3}\ln^{-2}\theta. \end{aligned}\tag{A.22}$$

Since

$$\tfrac{\theta^3}{\mathcal{L}} \ge \lambda_-(0)-\lambda_+(0) \ge 0,$$

we conclude that $\lambda_\pm(\beta)$ cross each other somewhere in the interval

$$\check{J} := \left[0, \tfrac{2\theta^3\ln^2\theta}{\mathcal{L}}\right] \subset J.$$

So, if we can show that these two eigenvalues never intersect, we will arrive at contradiction. Thus the result follows from

Lemma A.19. *The eigenvalues $\lambda_\pm(\beta)$ cannot intersect each other at the interval $\check{J}$ above. In fact, the relation* (A.20) *holds.*

Proof. We will shorthand H for H_β and P for $P_{hyb}(\beta)$ here. The idea here is to use Schur complementation. Namely, given $\lambda \in I$, we consider $M = M(\beta, \lambda)$, the Schur complement of H in $Range\,\bar{P}$, defined as

$$M := P\,(H-\lambda)\,P - PH\bar{P}\left(\bar{P}\,(H-\lambda)\,\bar{P}\right)^{-1}\bar{P}HP.$$

We note that by (A.17) M is well defined for our range of λ's and β's. The standard result in the theory of Schur complementation see e.g., [Z], is that $\mathrm{tr}\chi_{\{\lambda\}}(H) = 2$ (the sufficient and necessary conditions for the intersection of two eigenvalues) if and only if $M = 0$. In particular, the non-intersection property will follow if we can show that in this range we have $M_{12} \neq 0$. To this end, we note that

$$M_{12} = \langle \phi_{out}, (H-\lambda)\,\phi_{in}\rangle - \langle \phi_{out}, PH\bar{P}\left(\bar{P}\,(H-\lambda)\,\bar{P}\right)^{-1}\bar{P}HP\phi_{in}\rangle,$$

with $\phi_{in,out}$ defined in Proposition A.14 and where the right hand side is well defined thanks to (A.17). We now evaluate each of the terms on the right hand side. The first one is equal to

$$\langle \phi_{out}, H\phi_{in}\rangle = \langle \phi_{out}, \Gamma\phi_{in}\rangle = \phi_{out}(l_+)\phi_{in}(r_+),$$

where $\Gamma := H_0 - H_{out} - H_{in}$ is the coupling between vertices $l_\pm$ and $r_\pm$ (we recall that the bond $\langle l_\pm, r_\pm\rangle$ couples Λ_{out} and Λ_{in} and that $\mathrm{supp}\,(\phi_{out}) \subset \Lambda_{out}^+$). Let $\bar{H} = \bar{P}H\bar{P}$, and let $\left(\bar{H}-\lambda\right)^{-1}$ denote the inverse of $\bar{H}-\lambda$ on the $Range(\bar{P})$. To evaluate the second one, we use the identities

$$\bar{P}HP\phi_{in} = \Gamma\phi_{in} - \langle \phi_{out}, \Gamma\phi_{out}\rangle\phi_{out}, \quad \bar{P}HP\phi_{out} = \Gamma\phi_{out} - \langle \phi_{in}, \Gamma\phi_{out}\rangle\phi_{in},$$

Lemma A.15, localization of $\phi_{in,out}$ as well as position of their centers to get

$$\langle \phi_{out}, PH\bar{P}\left(\bar{H}-\lambda\right)^{-1}\bar{P}HP\phi_{in}\rangle = \sum_\pm \bar{\phi}_{out}(l_+)\phi_{in}(r_\pm)\langle \delta_{r_\pm}, \left(\bar{H}-\lambda\right)^{-1}\delta_{l_+}\rangle + O\left(\mathrm{e}^{-\sqrt{\mathcal{L}}/5}\right). \quad \text{(A.23)}$$

We next use the geometric resolvent identity to estimate the right hand side as

$$= \sum_\pm \bar{\phi}_{out}(l_+)\phi_{in}(r_\pm)\langle \delta_{r_\pm}, \left(H^\Psi-\lambda\right)^{-1}\delta_{l_+}\rangle + O\left(\left|\langle \delta_{r_\pm}, T\delta_{l_+}\rangle\right|\left|\phi_{out}(l_+)\phi_{in}(r_\pm)\right|\right) + O\left(\mathrm{e}^{-\sqrt{\mathcal{L}}/5}\right), \quad \text{(A.24)}$$

where Ψ was defined in Proposition A.12 (we note that $H^\Psi_\beta = H^\Psi$) and T is given by

$$T = \left(\bar{H}-\lambda\right)^{-1}\left(H - H^\Psi + \lambda P - \bar{P}\Gamma P - P\Gamma\bar{P} + PHP\right)\left(H^\Psi-\lambda\right)^{-1}.$$

We then bound

$$\left|\langle \delta_{r_\pm}, T\delta_{l_+}\rangle\right| \leq \mathrm{e}^{-\sqrt{\mathcal{L}}/5}, \quad \text{(A.25)}$$

using Proposition A.12, localization for H^Ψ, the properties of $\phi_{in,out}$, and position of their centers. Additionally, we note that the '$-$' summand in $\sum_\pm$ above vanishes. Finally, using the second resolvent identity and Proposition A.12, we obtain

$$\bar{\phi}_{out}(l_+)\phi_{in}(r_+)\langle \delta_{r_+}, \left(H^\Psi-\lambda\right)^{-1}\delta_{l_+}\rangle = \bar{\phi}_{out}(l_+)\phi_{in}(r)\langle \delta_{r_+}, \left(H^\Psi\right)^{-1}\delta_{l_+}\rangle + O\left(\frac{1}{\sqrt{\mathcal{L}}}\left|\phi_L(l)\phi_R(r)\right|\right). \quad \text{(A.26)}$$

Putting these bounds together, we get

$$|M_{12}| = |\phi_{out}(l_+)\phi_{in}(r_+)|\left|\langle \delta_{r_+}, \left(H^\Psi\right)^{-1}\delta_{l_+}\rangle - 1\right| + O\left(\tfrac{1}{\sqrt{\mathcal{L}}}\left|\phi_L(l)\phi_R(r)\right|\right). \quad \text{(A.27)}$$

Hence $M_{12} \neq 0$ as the eigenfunctions of $H_{in,out}$ cannot vanish at the respective boundary points and by Lemma A.13. More specifically, $|\phi_{out}(l_+)| \geq e^{-C\mathcal{L}}$, $|\phi_{in}(r)| \geq e^{-C\sqrt{\mathcal{L}}}$ using (a) $\phi_{in,out}$ are normalized; (b) $|\Lambda_{in}| = \sqrt{\mathcal{L}}$ and $\left|\Lambda_{out}^{+}\right| < \mathcal{L}$; [ESo, Lemma 3.1].

Now we prove the quantitative result, (A.20). This can be seen from the bound $\left\|\left(H^{\Psi}\right)^{-1}\right\| \leq \theta^{-1}\sqrt{\mathcal{L}}$, (A.17), and [ESo, Lemma 5.2]. □

□

Appendix B. Auxiliary results

Let H be a self-adjoint operator. Throughout the text we often use the integral representation

$$P_{[E_1,E_2]}(H) = -\frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{j=1}^{2}(-1)^j\left(H - ix - E_j\right)^{-1}dx, \tag{B.1}$$

which holds provided that E_1, E_2 are not in the spectrum $\sigma(H)$. If, in addition, $H(s)$ is a differentiable family of operators, the formula

$$\frac{d}{ds}\left(H(s) - ix - E_j\right)^{-1} = -\left(H(s) - ix - E_j\right)^{-1}\dot{H}(s)\left(H(s) - ix - E_j\right)^{-1} \tag{B.2}$$

holds. Similarly

$$[R, \frac{1}{H-z}] = -\frac{1}{H-z}[R,H]\frac{1}{H-z} \tag{B.3}$$

Lemma B.1. *Let H_1, H_2, R be bounded operators on $\ell^2(\Lambda)$, with H_1, H_2 self-adjoint. Let $J = [E_1, E_2]$ and denote by $J_{2\Delta}$ for $\Delta > 0$, the fattened interval $J + [-2\Delta, 2\Delta]$. Suppose that, for some ϵ_1, ϵ_2,*

(i) $\|(H_1 - H_2)R\| = \epsilon_1$

(ii) $\|[H_2, R]P_J(H_2)\| \leq \epsilon_2$.

Then

$$\left\|\bar{P}_{J_\Delta}(H_1)RP_J(H_2)\right\| \leq \frac{\epsilon_1 + \epsilon_2}{\Delta}.$$

Proof. Let $z_1 = E_1 - \Delta + ix$ and $z_2 = E_2 + \Delta + ix$ and write

$$G_{i,j} = \left(H_i - z_j\right)^{-1}.$$

We first establish the identity

$$\begin{aligned}\bar{P}_{J_\Delta}(H_1)RP_J(H_2) &= \frac{1}{2\pi}\sum_{j=1}^{2}(-1)^j\int_{-\infty}^{\infty}\bar{P}_{J_\Delta}(H_1)G_{1,j}[H_2,R]G_{2,j}P_J(H_2)dx \\ &+ \frac{1}{2\pi}\sum_{j=1}^{2}(-1)^j\int_{-\infty}^{\infty}\bar{P}_{J_\Delta}(H_1)G_{1,j}\left(H_2 - H_1\right)R\,G_{2,j}P_J(H_2)dx.\end{aligned}$$

Indeed, we start from

$$G_{1,j}[H_2,R]G_{2,j} = G_{1,j}(H_2 - H_1)RG_{2,j} + RG_{2,j} + G_{1,j}R.$$

Upon multiplying with $(-1)^j$, summing over $j = 1, 2$, integrating over x, and using (B.1) with $[E_1, E_2]$ replaced by $[E_1 - \Delta, E_2 + \Delta]$, we get the desired identity. Next, we bound

$$\max_{j=1,2}\left\|\bar{P}_{J_\Delta}(H_1)G_{1,j}\right\| \leq \frac{1}{\sqrt{x^2+\Delta^2}}, \quad \max_{j=1,2}\|G_{2,j}P_J(H_2)\| \leq \frac{1}{\sqrt{x^2+\Delta^2}}$$

to get

$$\left\|\bar{P}_{J_\Delta}(H_1)RP_J(H_2)\right\| \leq (\epsilon_1+\epsilon_2)\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{x^2+\Delta^2} = \frac{\epsilon_1+\epsilon_2}{\Delta}.$$

□

In the next two lemmas, we use the notation $J_a(\mu) = [\mu - a, \mu + a]$, and will let $P^\Theta_{J_a(\mu)}$ denote the spectral projection of H_o^Θ onto $J_a(\mu)$.

Lemma B.2. *Let Φ, Θ, with $\Phi \subset \Theta$ be finite subsets of $\mathbb{Z}^d$. Let (ϕ, μ) be an eigenpair for H_o^Φ. Then we have*

$$\operatorname{dist}\left(\mu, \sigma(H_o^\Theta)\right) \le C\, |\partial_r \Phi|\, \|\chi_{\partial_r\Phi}\phi\|_\infty , \tag{B.4}$$

and

$$\operatorname{dist}\left(\phi, \mathrm{R}ange\, P^\Theta_{J_a(\mu)}\right) \le \frac{C}{a}\, |\partial_r \Phi|\, \|\chi_{\partial_r\Phi}\phi\|_\infty . \tag{B.5}$$

Conversely, if (ψ, λ) is an eigenpair for H^Θ, then

$$\operatorname{dist}\left(\lambda, \sigma(H_o^\Phi)\right) \le C\, |\Theta \setminus \Phi|\, \left\|\chi_{\Theta\setminus\Phi}\psi\right\|_\infty , \tag{B.6}$$

and

$$\operatorname{dist}\left(\phi, \mathrm{R}ange\, P^\Phi_{J_a(\lambda)}\right) \le \frac{C}{a}\, |\Theta \setminus \Phi|\, \left\|\chi_{\Theta\setminus\Phi}\psi\right\|_\infty . \tag{B.7}$$

Proof. We have

$$\left(\left(H_o^\Theta - \mu\right)\phi\right)(y) = \begin{cases} \sum_{\substack{y' \in \Phi: \\ |y-y'| \le r}} H_o(y,y')\phi(y') & \text{if } y \in \Theta \setminus \Phi \text{ and } \operatorname{dist}(y,\Phi) \le r, \\ 0 & \text{otherwise.} \end{cases} \tag{B.8}$$

It follows that

$$\left\|\left(H_o^\Theta - \mu\right)\phi\right\| \le C\, |\partial_r \Phi|\, \|\chi_{\partial_r\Phi}\phi\|_\infty . \tag{B.9}$$

Thus, recall that ϕ is normalized,

$$\operatorname{dist}\left(\mu, \sigma(H_o^\Theta)\right) \le \left\|\left(H_o^\Theta - \mu\right)\phi\right\| \le C\, |\partial_r \Phi|\, \|\chi_{\partial_r\Phi}\phi\|_\infty . \tag{B.10}$$

On the other hand, we have

$$\left\|\bar{P}^\Theta_{J_a(\mu)}\phi\right\| \le \left\|\bar{P}^\Theta_{J_a(\mu)}\left(H_o^\Theta - \mu\right)^{-1}\right\| \left\|\left(H_o^\Theta - \mu\right)\phi\right\| \le \frac{C}{a} \left\|\chi_{\Theta\setminus\Phi}\psi\right\|_\infty , \tag{B.11}$$

from which the second assertion of the lemma follows.

Similar considerations yield

$$\left\|\left(H_o^\Phi - \lambda\right)\phi\right\| \le C\, |\Theta \setminus \Phi|\, \left\|\chi_{\Theta\setminus\Phi}\phi\right\|_\infty , \tag{B.12}$$

which in turn imply the bounds (B.6)–(B.7). □

In this paper we are interested in the evolution of the initial wave packet ψ_o supported near some $x \in \mathbb{Z}^d$ up to the (rescaled) times s of order 1. In this context we can always approximate the dynamics generated by $H(s)$ with the one generated by $\hat{H}^{\mathbb{T}}(s)$, where $H^{\mathbb{T}}(s)$ is understood as an operator on $\ell^2(\mathbb{Z}^d)$ by extending it by zero outside of the box Λ_L, in a following sense.

Proposition B.3 (The finite speed of propagation bound)**.** *Let $\mathbb{T}$ be a torus of the linear size R and let $U_\epsilon(s)$, $U_\epsilon^{\mathbb{T}}(s)$ be the dynamics generated by $H(s)$ and $H^{\mathbb{T}}(s)$, respectively, i.e.,*

$$\begin{aligned} i\epsilon\partial_s U_\epsilon(s) &= H(s)U_\epsilon(s), \quad U_\epsilon(0) = 1; && \text{(B.13)} \\ i\epsilon\partial_s U_\epsilon^{\mathbb{T}}(s) &= H^{\mathbb{T}}(s)U_\epsilon^{\mathbb{T}}(s), \quad U_\epsilon^{\mathbb{T}}(0) = 1. && \text{(B.14)} \end{aligned}$$

Then there exist $c > 0$ such that for any $\mathcal{L}$ satisfying $\mathcal{L} \ge C/\epsilon$ we have

$$\max_s \left|(U_\epsilon^\sharp(s))(y,x)\right| \le \mathrm{e}^{-c|x-y|}, \quad \text{for } |x-y| \ge \frac{\mathcal{L}}{4}, \tag{B.15}$$

where $U_\epsilon^\sharp$ is either U or $U^{\mathbb{T}}$.

Proof. This is a standard fact for (local) lattice Hamiltonians, see e.g., the proof of [EGS, Lemma 5] for the time independent case (which extends to the time dependent one without effort), or, for a more general approach, [LR]. □

References

[AG] Aizenman, M., and Graf, G.M.: Localization bounds for an electron gas. *J. Phys. A*, **31** 6783–6806 (1998).

[AW] Aizenman, M. and Warzel, S.: Random Operators: Disorder Effects on Quantum Spectra and Dynamics. *American Mathematical Society* (2015).

[ASS] J. E. Avron, R. Seiler, and B. Simon, "Charge deficiency, charge transport and comparison of dimensions", Commun. Math. Phys., 159, 399 (1994).

[AHS] J. E. Avron, J. S. Howland, and B. Simon, "Adiabatic Theorems for Dense Point Spectra", Commun. Math. Phys., 128, 497 (1990).

[BDF] S. Bachmann, W. De Roeck, and M. Fraas, The adiabatic theorem and linear response theory for extended quantum systems, Commun. Math. Phys. **361**, 997–1027 (2018).

[BBR] Barequet, R., Barequet, G., and Rote, G.: Formulae and growth rates of high-dimensional polycubes. Combinatorica **30**, 257–275 (2010).

[DE] A. Dietlein and A. Elgart,: Level spacing for continuum random Schrödinger operators with applications. J. Eur. Math. Soc. (JEMS) **23**, 1257–1293 (2021).

[EH] Elgart, A. and Hagedorn, G. A.: A note on the switching adiabatic theorem. Journal of Mathematical Physics, **53**, 102202 (2012).

[EK1] Elgart, A., Klein, A.: An eigensystem approach to Anderson localization. J. Funct. Anal. **271**, 3465–3512 (2016).

[EK2] Elgart, A., Klein, A.: Eigensystem multiscale analysis for the Anderson model via the Wegner estimate, Ann. Henri Poincaré , to appear.

[EPS] Elgart, A., Pastur, L., Shcherbina, M.: Large block properties of the entanglement entropy of free disordered fermions. J. Stat. Phys. **166**, 1092–1127 (2017).

[ES] A. Elgart and J.H. Schenker: A strong operator topology adiabatic theorem, Rev. Math. Phys., **14**, 569 (2002).

[ESS] Elgart, A., Shamis, M., and Sodin, S.: Localisation for non-monotone Schrödinger operators, J. Eur.Math. Soc. **16**, 909–924 (2014).

[AE] Avron J E and Elgart A. Adiabatic theorem without a gap condition. Commun. Math. Phys. 203 445–63, (1999).

[EGS] Elgart, A., G. M. Graf, and J. H. Schenker. Equality of the Bulk and Edge Hall Conductances in a Mobility Gap. Commun. Math. Phys. **259**, 185–221 (2005)

[ESo] Elgart, A. and Schmidt, D.: Eigenvalue counting inequalities, with applications to Schrödinger operators. J. Spectr. Theory, **5**, 251–278, (2015).

[ESo] Elgart, A. and Sodin, S.: The trimmed Anderson model at strong disorder: localisation and its breakup. J. Spectr. Theory, **7**, 87–110, (2017).

[ETV] Elgart, A., Tautenhahn, M., Veselić, I.: Anderson localization for a class of models with a sign-indefinite single-site potential via fractional moment method. Ann. Henri Poincaré 12, 1571–1599 (2010).

[G] Gevrey, M.: Sur la nature analytique des solutions des équations aux dérivés partielles. Premier mémoire., *Ann. de l'Ecole norm. sup.*, **3 35**, 129–190 (1918).

[Kl] Klarner, D.A.: Cell growth problems. Canadian J. of Mathematics, **19** 851–863 (1967).

[N] Nenciu, G., "Linear adiabatic theory: Exponential estimates," Commun. Math. Phys. **152**, 479–496 (1993).

[RS] M. Reed and B. Simon, Methods of Modern Mathematical Physics, I, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980. Func- tional analysis.

[BK] J. Bourgain and I. Kachkovskiy. Anderson localization for two interacting quasiperiodicparticles.Geom. Funct. Anal., 29(1):3–43, 2019.

[G] Gebert, M.: A lower Wegner estimate and bounds on the spectral shift function for continuum random Schrödinger operators. J. Funct. Anal. **277**, 108284 (2019)

[DJLS] R. del Rio, S. Jitomirskaya, Y. Last, B. Simon: Operators with singular continuous spectrumIV: Hausdorff dimensions, rank one perturbations and localization. J. d'Analyse Math. **69**, 153–200 (1996).

[KM] Klein, A., and Molchanov, S.: Simplicity of eigenvalues in the Anderson model. J. Stat. Phys. **122**, 95–99 (2006)

[Z] Zhang, F., ed. The Schur complement and its applications. Vol. 4. Springer Science & Business Media, 2006.

[CL] Carmona, R. and Lacroix, J.: Spectral theory of random Schrödinger operators. Springer Science & Business Media, 2012.

[KlS] F. Klopp and J. Schenker: On the spatial extent of localized eigenfunctions for random Schrödinger operators.

[B] Bornemann F. (1998) Homogenization in time of singularly perturbed mechanical systems (Lecture Notes in Mathematics vol. 1687) (Berlin: Springer)

[KaS] I. Kachkovskiy and Y. Safarov, Distance to normal elements in C^* algebras of real rank zero, J. Amer. Math. Soc. **29**, 61–80 (2016).

[Ka] Kato T.: On the adiabatic theorem of quantum mechanics Phys. Soc. Japan 5 435–9 (1950).

[LR] Lieb,E.H.,Robinson,D.W.:The finite group velocity of quantum spin systems. Commun.Math. Phys. **28**, 251–257 (1972).

[S] Simon, B. Fifteen problems in mathematical physics. Perspectives in mathematics, Birkhäuser, Basel **423** (1984).

Instituut Theoretische Fysica, KULeuven, 3001 Leuven, Belgium
Email address: wojciech.deroeck@kuleuven.be

Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA
Email address: aelgart@vt.edu

Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA
Email address: fraas@vt.edu