LOCABATIC THEOREM FOR DISORDERED MEDIA AND VALIDITY OF LINEAR RESPONSE

WOJCIECH DE ROECK, ALEXANDER ELGART, AND MARTIN FRAAS

In memoriam: Rachel Vaiman

Abstract. Spectral localization is intrinsically unstable under perturbation. As a result, the adiabatic theorem of quantum mechanics cannot generally hold for localized eigenstates. However, it turns out that a remnant of the adiabatic theorem, which we name the “locobatic theorem”, survives: The physical evolution of a typical eigenstate \( \psi \) for a random system remains close, with high probability, to the spectral flow for \( \psi \) associated with a restriction of the full Hamiltonian to a region where \( \psi \) is supported. We make the above statement precise for a class of Hamiltonians describing a particle in a disordered background. Our argument relies on finding a local structure that remains stable under the small perturbation of a random system.

An application of this work is the justification of the linear response formula for the Hall conductivity of a two-dimensional system with the Fermi energy lying in a mobility gap. Additional results are concerned with eigenvector hybridization in a one-dimensional Anderson model and the construction of a Wannier basis for underlying spectral projections.

1. Introduction

In this work, we examine the dynamical properties of a disordered quantum system, described by a random self-adjoint operator \( H \) on a Hilbert space \( \mathcal{H} \), interacting weakly with a time-dependent external perturbation, \( W(t) \), with the interaction strength modulated by the parameter \( \beta \). It produces a family of self-adjoint operators

\[
H(t) = H + \beta W(t), \quad t \in \mathbb{R}_+.
\]

\[\text{(1.1)}\]

A typical example of such an \( H \) is the Anderson Hamiltonian \( H_A \) acting on \( \mathcal{H} = \ell^2(\mathbb{Z}^d) \), with \( H_A := \Delta + \omega \). Here, \( \Delta \) is the discrete Laplacian and \( \omega \) is a multiplication operator, \( \omega(x) \psi(x) = \omega_x \psi(x) \) for \( \psi \in \mathcal{H} \), where \( \omega_x \) are i.i.d. random variables with some joint probability distribution \( \mu \).

The dynamical properties of the (not necessarily random) family \( H(t) \) are of special interest in transport theory (where time-dependent perturbation describes the driving of an equilibrium system \( H \) away from the equilibrium). They also play an important role in the problem of thermalization, in the studies of time-quasi-periodic, non-linear Schrödinger (NLS) operators, and in other areas of research.

The presence of disorder in quantum mechanical systems leads to the phenomenon of localization. Spectral localization manifests in the emergence of energy intervals \( J_{loc} \subset \mathbb{R} \) such that, for almost all random configurations \( \omega \), \( \sigma(H) \cap J_{loc} \) is pure point. Moreover, the eigenvectors of \( H \) in \( J_{loc} \) are (spatially) exponentially localized in the sense of (1.2) below. Dynamical localization is concerned with the non-spreading of wave packets during time evolution. It is expressed as the (uniform in time) exponential decay of the matrix elements of \( e^{-itH} P_{J_{loc}} \), the unitary semigroup generated by \( H \) and restricted to the energy interval \( J_{loc} \) (here, \( P_{J_{loc}} \) denotes the spectral projection of \( H \) onto \( J_{loc} \)).

The latter concept of localization is still well-defined for the full system \( H(t) \), and a natural question is whether it is still dynamically localized for at least small perturbations \( \beta \ll 1 \). What
makes this question difficult is the fact that localization for even time-independent systems is not stable under perturbation: The rank one perturbation family \( H_A(\beta) \) of the form \( H_A(\beta) = H_0 + \beta \chi \) exhibits almost sure singular continuous spectrum for a \( G_\beta \)-dense set of \( \beta \)'s, [16, 27]. This only shows that spectral localization fails for such \( \beta \), but it turns out that dynamical localization fails as well, due to the resonant hybridization phenomenon. We will discuss the latter in detail in subsection 1.2 and Appendix A below. For a time-dependent system, one can then expect dynamical de-localization if it spends a long time near such values of \( \beta \), particularly if the dependence on time is very slow. One can in fact characterize such a transition more precisely on a heuristic level, [18, Section 1], specifically \( \nu \sim \beta \exp(-c_d \beta^{-p_d}) \), where \( \nu \) is a typical time frequency for \( W \) and \( c_d, p_d \) are dimension-dependent parameters.

The properties of the system (1.1) have been studied before under various assumptions. In physics literature, one of the earliest works in this direction goes back to [50], which analyzes the behavior of a random matrix model. On a mathematical footing, compact (in space) perturbations \( W \) have been studied in the time-periodic [48] and the time-quasi-periodic [13] settings. The case of spatially extensive periodic systems with few frequencies was considered in [18]. In the \( \beta = \nu \) adiabatic setting, it was considered in [43]. For time periodic systems, one can also consider the spectral localization of the associated Floquet operator, [48, 18, 1].

In all these works, the analysis heavily depends on the assumption of strong disorder, under which the interval \( J_{loc} \) can be replaced by the whole \( \mathbb{R} \). As a result, it is reasonable to expect that the perturbed system also exhibits almost sure dynamical localization on \( \mathbb{R} \), which is indeed the conclusion of these works. One of the purposes of this paper is to develop an approach that works for any interval \( J_{loc} \) and in particular for any strength of the disorder. This relies on underpinning the local structure for the disordered systems introduced here, which is more robust than the standard description of the localization and, in particular, survives the time-dependent perturbations described by (1.1). We consider this problem in the adiabatic setting.

### 1.1. Adiabatic theory

The Schrödinger dynamics associated with \( H(t) \) in (1.1) are given by the linear initial value problem (IVP):

\[
\dot{\psi}(t) = H(t)\psi(t), \quad \psi(0) = \psi_o, \tag{1.2}
\]

where \( \psi_o \) 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 \( H(t) = H_0 \) and the initial state \( \psi_o \) being an eigenvector for \( H_o \). In this case, the evolution \( \psi(t) \) coincides with \( \psi_o \) up to an acquired phase.

As we have seen above, 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 a parameter is given by the spectral gap in \( H(t) \) (note that energy has units time \(^{-1} \) in (1.2)).

To make this statement more quantitative, it is convenient to consider the family \( H(\epsilon t) \), where \( \epsilon \) 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

\[
\dot{\psi}_\epsilon(s) = H(s)\psi_\epsilon(s), \quad \psi_\epsilon(0) = \psi_o, \quad s \in [0, 1]. \tag{1.3}
\]

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

\[
\dot{U}_\epsilon(s) = H(s)U_\epsilon(s), \quad U_\epsilon(0) = 1. \tag{1.4}
\]

Let us assume that the spectrum \( \sigma(H(s)) \) of the operator \( H(s) \) contains a set \( S(s) \) 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 \( S(s) \), and assuming that \( P(0)\psi_o = \psi_o \), the (qualitative) adiabatic theorem states that

\[
\lim_{\epsilon \to 0} \| \psi_\epsilon(s) - P(s)\psi_\epsilon(s) \| = 0, \tag{1.5}
\]
provided $H(s)$ is smooth. In fact, a stronger statement holds true, namely
\[
\lim_{\epsilon \to 0} \|U_\epsilon(s)P(0)U_\epsilon^*(s) - P(s)\| = 0, \tag{1.6}
\]
and one can make the error estimate for the norm above explicit in terms of its $\epsilon$ and $g$ dependencies, see e.g., Lemma 5.9 below.

The adiabatic theorem and its derivatives play an important role in the various branches of quantum and statistical mechanics. The first results on adiabatic behavior go back to the dawn of quantum mechanics and are due to Born and Fock in 1928, \cite{10}. Modern adiabatic theory was initiated by Kato in 1950, \cite{34}, and has since been 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, \cite{11, 16} (we also mention a related work \cite{5}). This situation usually occurs for a ground state in the threshold of the continuous spectrum. In space-adiabatic perturbation theory \cite{45}, the gap is closed by a locally small but globally large perturbation (for related work in field theory see \cite{49}). More recently, the adiabatic theorem was established for certain systems characterized by a spectral gap but non-smooth $P(s)$, \cite{8, 42}. 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 Hamiltonian $H$.

**Adiabatic theorem for a finite system.** The evolution considered in (1.2) is equivalent via a unitary transformation to the solution of
\[
i\dot{\phi}(t) = \tilde{H}(t)\phi(t), \quad \phi(0) = \psi_o, \quad \tilde{H}(t) = \mathcal{W}_\beta(t)HW_\beta^*(t), \tag{1.7}
\]
where $\mathcal{W}_\beta(t)$ is a unitary generated by $\beta W(t)$:
\[
i\dot{\mathcal{W}}_\beta(t) = \beta W(t)\mathcal{W}_\beta(t), \quad \mathcal{W}_\beta(0) = 1.
\]
If $H$ is gapped, so is $\tilde{H}(t)$, and in particular (1.7) can be studied using the gapped adiabatic framework (e.g., \cite{23}) with adiabatic parameter $\beta$. To this end, one needs the gap $g$ in $H$ to satisfy $\frac{\beta}{g} \ll 1$. Since for a typical finite system $H^\Lambda$ the generic gap size is $O(|\Lambda|^{-1})$ (see also Lemma A.9 below for the random case), this type of construction works well for $\Lambda$ of the linear dimension $o(\beta^{-1/d})$. In fact, one can also consider the interacting systems and take into account the effects of the edge states that can close the gap in $H$, due to the boundary conditions using this framework implicitly. Physically, the reason for why this approach works is related to the fact that a bulk eigenstate of the system constitutes a meta-stable (resonance) state for the finite system, and the lifetime for such a state is at least comparable with $\beta^{-1}$. One then expects that the adiabatic theory will remain valid for times comparable with the lifetime of the resonance, \cite{2, 20}. For state of the art assertions on this topic, we refer the reader to \cite{29} and the references therein.

Let us note that the above discussion is not disordered systems-specific. As we shall see, in the presence of disorder, these results can be drastically improved. Specifically, we will be able to consider finite systems $H^\Lambda$ with $|\Lambda|$ that is stretched exponentially large in $\beta^{-1}$. Roughly speaking, this improvement is possible due to the fact that the metastable states in the random systems have much longer lifetimes than the a-priori estimate $\beta^{-1}$ mentioned earlier. We first discuss what exactly we mean by localization.

### 1.2. Localized systems and resonant hybridization

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 $0 < C, c, m < \infty$, 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 $\psi_i$, satisfying
\[
|\psi_i(x)| \leq C|x|^{d+1} e^{-c|x-x_i|}. \tag{1.8}
\]

The prototypical example of such an $H$ is the Anderson model $H_A$ described earlier. The Anderson Hamiltonian is known to display exponential localization in the vicinity of spectral edges, at large values of disorder (for a sufficiently regular distribution $\mu$) and in dimension
To make any meaningful statements about continuum systems, we will not attempt to cite the extensive literature of history, reviews, results and open problems concerning this model and its variants, but will rather refer the interested reader to a recent monograph [4] on the subject.

For the family of operators \( H(s) \) described by (1.1), one should not expect much uniformity of the localization properties as a function of \( s \) or \( \beta \), provided that \( W \) is sufficiently non-trivial, c.f. the family \( H_A(\beta) \) above.

The destruction of such uniform localization properties can be linked to a mechanism known as resonant hybridization, see e.g. [4, Chapter 15]. The latter concept can be illustrated by considering a 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 for \( H(s) \) is \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \). These remain approximate eigenvectors for \( H(s) \) provided that \( |s| \ll g \). However, the picture is different for the case where 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 separate eigenstates and as such is present at all scales. 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 in our extended disordered system as well, 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 \( L \), we can then label the eigenstates \( \psi_{i,s} \) so that for each \( i \) \( t \mapsto \psi_{i,s} \) is continuous. However, we do expect the modulus of continuity to diverge badly as \( L \to \infty \).

We are not aware of any prior rigorous results making the two-level heuristics precise for \( Z^d \) systems for any \( d \) (however, see [4, Chapter 15] for the results on regular trees). In Appendix \( A \) we show the emergence of hybridization rigorously for a one-dimensional system. Specifically, we prove Theorem \( \ref{t:main} \) which qualitatively can be formulated as

**Theorem 1.1.** Let \( H_0 \) 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 scale-independent probability. The corresponding eigenvalues exhibit avoided level crossings.

### 1.3. The locobatic behaviour.

Theorem \( \ref{t:main} \) 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 \( \epsilon \) is small enough. However, as we have mentioned above, this spectral flow is extremely nonlocal, whereas the physical evolution cannot be arbitrarily nonlocal. We believe that the way that this dilemma is resolved is that the physical evolution of an initial eigenvector, for most values of \( s \), stays close to one of the global eigenvectors \( \psi_{i,s} \), though the index \( i \) 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 \( \phi_\epsilon \) of the restriction of \( H(s) \) to a local box around the support of the initial eigenvector \( \psi_{i,0} \). We will refer to this statement as a locobatic theorem, and state it quantitatively as Theorem \( \ref{t:locobatic} \) below. In other words, \( \phi_\epsilon \) can be interpreted as a meta-stable state for \( H(s) \) with a very long lifetime.

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

---

4
An operator $K$ acting on $\ell^2 (\mathbb{Z}^d)$ 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 $\ell^\infty$ distance in $\mathbb{Z}^d$.

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

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

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

For any $\Theta \subset \mathbb{Z}^d$, we denote by $H^\Theta$ 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 $\Theta, \Phi$ of $\mathbb{Z}^d$ that satisfy $\text{dist}(\Theta, \Phi) > r$, the operators $H^\Theta$ and $H^\Phi$ are statistically independent.

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

$$|x - y|_\Theta = \min (|x - y|, (\text{dist}(x, \partial_1 \Theta) + \text{dist}(y, \partial_1 \Theta))),$$

with the interior boundary $\partial_1 \Theta = \{ x \in \Theta, \text{dist}(x, \Theta^C) = 1 \}$. 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 $J_{\text{loc}}$ for $H$ reads

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

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

(1.10)

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

For some of our results we will also need

**Assumption 1.5** (Finite spectral multiplicity). There exists $m \in \mathbb{N}$ such that, for any $\Theta \subset \mathbb{Z}^d$, the multiplicity of eigenvalues of $H^\Theta$ does not exceed $m$ almost surely.

**Remark 1.6.** For the standard Anderson model with absolutely continuous random distributions, Simon noted that $m = 1$, [17]. This type of result can be extended to a larger class of discrete models, see, e.g., [3, Theorem 5.8] and [17]. While the simplicity of the spectrum is in general not known to hold for models that satisfy Assumptions [1.3][1.4] in practice a majority of them are generated using finite-rank operators for which Assumption [1.5] does hold. [31].

**Remark 1.7.** Surprisingly, the basic localization property (1.8) has only been proven in existing literature under the assumption of spectrum simplicity (i.e., $m = 1$ in Assumption [1.3] above), see [4, Theorem 7.4]. In order to avoid this rather restrictive condition, we obtain its analogue for a more general case of finite $m$ in Appendix [B] below. This relies on the construction of the so-called Wannier basis for an eigenprojection of the localized Hamiltonian.

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

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

(1.11)

Our two main parameters are the adiabaticity parameter $\epsilon$ and the driving strength $\beta$, introduced earlier in [1.3] and [1.4], 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 of our estimates will be uniform in $M$.

The following then is the *[locobatic theorem]*. It is based on the emergence of a local structure for the spectral data associated with a torus, once partitioned into smaller boxes of linear size $\ell$. To make its presentation more accessible, we will use an extra assumption on the integrated density of states $N_{S_{\text{loc}}}$ (see (6.3) below) in addition to our standard hypotheses on the model.

**Theorem 1.8** *(Locobatic theorem on a torus).* Suppose that Assumption [1.23 and Assumptions [1.3] hold for $H(0)$ and the integrated density of states $N_{S_{\text{loc}}}$ is a.s. positive. We introduce a scale parameter $\ell \in \mathbb{N}$ satisfying

$$\ell^p \leq \epsilon^{-1} \leq Ce^{\sqrt{\ell}}, \quad \beta^{1/p} \leq 1/\ell. \quad (1.12)$$

Let $J_{\text{loc}}'$ be any closed interval contained in $J_{\text{loc}}$. With probability at least $1 - e^{-\sqrt{\ell}}$, the following holds true for a fraction of at least $1 - e^{-c\sqrt{\ell}}$ of eigenstates $\psi$ of $H^\text{T}$ with eigenvalue $E \in J_{\text{loc}}'$:

There is a region $R \subset \mathbb{T}$ with $\text{diam}(R) \leq c\ell^{3/2}$ 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 $\sigma(H^R(s))$.

(ii) The solution $\psi(s)$ of the IVP with $\psi(0) = \psi$ satisfies

$$\max_{s \in [0,1]} \|(1 - P(s))\psi(s)\| \leq C \left( \epsilon \ell^{d+1/2+\xi} + e^{-c\sqrt{\ell}} \right). \quad (1.13)$$

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

$$\|(1 - P(1))\psi_s(1)\| \leq C_N \left( \epsilon^N \left( \ell^{N(d+1/2+\xi)} + \ell^{(2N+1)\xi} \right) + e^{-c\sqrt{\ell}} \right). \quad (1.14)$$

**Remark 1.9.** Let us note that both the upper and lower bounds on $\epsilon$ in (1.12) have to do with the faithfulness of our approximation of the actual eigenstate for $H^\text{T}$ by the local spectral patch for $H^R$. If $R$ is too small, then there is no reason for its eigenvectors (even the bulk ones) to be close to the eigenvectors of $H^\text{T}$ (so the spatial faithfulness of our approximation is destroyed). On the other hand, if $R$ is too big, the gaps in the spectrum of $H^\text{T}$ become smaller than the size $\beta$ of the perturbation, allowing for transition between eigenstates that are energetically far apart from one another (so the energetic faithfulness of our approximation is destroyed). In particular, one can think of these constraints as a consequence of the uncertainty principle for disordered systems.

**Remark 1.10.** If the spectrum of $H^R$ is *level-spaced*, i.e. if the probability of a spacing significantly smaller than $|R|^{-1}$ is small (as one can prove, e.g., for the standard Anderson model [37] and, at the bottom of the spectrum, for more general random models. [17]), 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, for a large fraction of times $s$, the range of $P(s)$ stays close to an eigenprojection of the global Hamiltonian $H^\text{T}(s)$. However, we do not expect this property to hold for all times $s$ on the basis of the hybridization result, Theorem 1.11 which shows that physical evolution cannot follow the non-local spectral flow.

We will use generic, $M, \epsilon, \beta, \ell$-independent constants $C, c$ (the scale parameter $\ell$ will be introduced below), whose values can change from line to line. They will, however, in general depend on the other parameters and constants introduced above (such as the range $r$ and the probability distribution $\mu$, as well as on the constants $C_q, C_k$, etc.).

### 1.4. A problem of linear response

Here, we will continue to consider the family $H(t)$ described in (1.1) above. Assuming that $W(t) = 0$ (no driving at time 0), an initial equilibrium state of the system at time $t = 0$ is described by the unperturbed density matrix $\rho \in \mathcal{L}(\mathcal{H})$ (a positive definite, trace one operator) that commutes with $H$, $[\rho, H] = 0$. 

The time evolution of the initial state \( \rho \) associated with the family \( H(t) \) is then given by its dynamical evolution, the perturbed density matrix \( \rho(t) \), satisfying the Heisenberg equation

\[
i\dot{\rho}(t) = [H(t), \rho(t)], \quad \rho(0) = \rho. \tag{1.15}\]

One can then formally compute \( \rho(t) \) in (1.15) to the first order in \( \beta \), obtaining

\[
\rho - \rho(t) = -i\beta \int_{0}^{t} e^{iH[s]} [W(s), \rho] e^{-iH[s]} ds + O(\beta^2) =: A(t)\beta + O(\beta^2).
\]

The operator \( A(t) \) in front of \( \beta \) on the right hand side is called the (quantum) linear response function.

The linear response theory (LRT) was originally developed by Green \(^{28}\) and Kubo \(^{39}\) and is surprisingly effective in applications, in particular in the explanation of the Quantum Hall Effect (QHE). The predictions of a system’s behavior based on LRT in QHE far surpass the naive range of its validity. Specifically, the system size \( \text{dim}(\mathcal{H}) \) can be taken to infinity (the so-called thermodynamic limit) independently of the values of \( \beta \). Moreover, the times \( t \) for which the LRT is applicable far exceed \( \text{max} \left( \frac{1}{\beta}, \frac{1}{\epsilon} \right) \) (we recall that \( \epsilon \) is an adiabatic parameter). This leads naturally to the question as to why LRT is justified in such regimes. It has been actively studied in mathematical physics for the last two decades, with various degrees of success. For a recent review of these efforts, we refer the reader to \(^{30}\). We postpone the discussion of results most relevant to the current work until the end of the next subsection.

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

**Setup:** We consider \( d = 2 \) and let \( \mathbb{Z}^2 \ni x = (x_1, x_2) \). We will denote by \( \Lambda_n \) the characteristic function of the set \( \{ x \in \mathbb{Z}^2 : x_n \geq 0 \}, n = 1, 2 \). 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 \( x_2 \)-direction. The function \( g \) satisfies

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

\[
Q = \int_{0}^{T} j(t) dt = \int_{0}^{T} \text{tr}(P_\epsilon(t) - P) J) dt,
\]

where \( j \) is the current, \( P = P_{\text{Fermi}}(H_0) \) 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 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

\[
Q = \frac{1}{\epsilon} \int_{0}^{1} \text{tr}(P_\epsilon(s) - P) J) ds,
\]

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

\[
\epsilon \partial_s P_\epsilon(s) = [H(s), P_\epsilon(s)], \quad P_\epsilon(-1) = P.
\]
The Hall conductance is defined as a proportionality constant between the applied potential difference (the spatial integral of the electric field) and the current flowing in the perpendicular direction. I.e., the measured conductance \( \sigma_m \) is defined by a relation

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

which gives

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

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

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

exists for a protocol such that \( \epsilon \ll \beta \) and is equal to the conductance \( \sigma \) obtained from the Kubo formula, see e.g., [3],

\[
\sigma := \text{tr}(P[[P, \Lambda_1], [P, \Lambda_2]]).
\]

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. [3]). 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 \( \sigma \).

Our result reads

**Theorem 1.11.** Suppose that \( H \) satisfies Assumptions 1.3–1.4 with \( E_F \) lying in the interior of \( J_{\text{loc}} \). Assume, moreover, that \( e^{-\beta - p_1/2} < \epsilon < \beta^{p_2p_1} \). Then

\[
|\sigma - \sigma_m| \leq C \frac{\epsilon}{\beta^p} + O(\beta^\infty + \epsilon^\infty)
\]

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

**Comparison with existing results.** In the context of QHE, the LRT has been justified for disordered systems if one takes the thermodynamic limit first, followed by the limit \( \beta \to 0 \) (keeping \( \epsilon \) fixed), [12]. One can also study the ac conductivity of a sample in the same order of limits, reproducing Mott’s formula, [36]. For a completely localized system, \( J_{\text{loc}} = \mathbb{R} \), the absence of transport has been established in the case \( \beta = \epsilon \), [43].

### 1.6. Adiabatic theorem in the strong operator topology.

An application of our results for finite geometry is the following assertion that holds for \( \mathbb{Z}^d \).

**Theorem 1.12.** Let \( E_F \in J_{\text{loc}} \) and suppose that Assumptions 1.2–1.4 hold. Then (cf. (1.6))

\[
s\lim_{\epsilon, \beta \to 0} U_\epsilon(s) P_{E_F}(0) U_\epsilon^*(s) = P_{E_F}(s), \quad s \in [0, 1]
\]

almost surely, provided we take a limit maintaining the relation

\[
e^{-1/\beta^p} < \epsilon < \beta^{p_2p_1},
\]

with \( p = p_1/2 \).

**Remark 1.13.** For \( \epsilon = \beta \), this statement was earlier proven in [23]. Let us also mention that the choice of topology (the strong operator topology) is essential here - in the absence of the spectral gap, the result is not expected to hold in the norm operator topology, see [23] for a counterexample. A major difference between prior work and our result is that the (formal) total variation of the perturbation, \( \epsilon^{-1} \beta \| W \| \), blows up as \( \epsilon, \beta \to 0 \), due to the constraints on \( \epsilon \) and \( \beta \).

\[^1\text{It is very likely that the evolution at very small } \epsilon \text{ is delocalized, but we do not quite grasp the implications of this for our problem.}\]
**Additional notation.** By \( \Lambda_R(y) \subset \mathbb{Z}^d \) we will denote a cube \( \Lambda_R = \Lambda_R(y) := ([-R, R]^d + y) \cap \mathbb{Z}^d \) for \( y \in \mathbb{Z}^d \), with side length \( 2R \). For a subset \( \Phi \subset \mathbb{Z}^d \), we will denote by \( \partial_\ell \Phi \) its \( \ell \)-extended boundary, i.e.,

\[
\partial_\ell \Phi = \{ x \in \Phi : \text{dist}(x, \Phi^c) \leq \ell \}.
\]

By \( \Phi_\ell \) we will denote

\[
\Phi_\ell = \Phi \setminus \partial_\ell \Phi.
\]

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 \( X := 1 - X \). For \( \mathcal{A} \subset \mathbb{T}, c \in \mathbb{R}_+ \), and \( \ell \in \mathbb{N} \), let \( \rho^\ell_\mathcal{A} \) be a (scaled) distance function

\[
\rho^\ell_\mathcal{A} := \frac{\text{dist}(\mathcal{A}, \{ x \})}{\sqrt{\ell}}.
\]

We denote the associated norm by

\[
\| K \|_{c, \ell} = \left\| e^{-c \rho^\ell_\mathcal{A}} K e^{c \rho^\ell_\mathcal{A}} \right\|
\]

This norm is multiplicative, i.e.,

\[
\| AB \|_{c, \ell} \leq \| A \|_{c, \ell} \| B \|_{c, \ell}
\]

for a pair of operators \( A, B \).

**1.7. Outline of the proofs.** We will now comment on the arguments pertaining to the proofs of our core assertions, namely Theorems 1.8, 1.14 below. We will not comment on the derivations of the remaining results, as they follow via 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 1.8. It will be convenient to formulate the concepts on a torus \( \mathbb{T} \) whose linear dimension is \( \mathcal{L} = e^{c \sqrt{\ell}} \), but this condition can be relaxed.

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

\[
\| P_{J_\gamma}(H^{T_\gamma}(0)) \psi_n \| \leq e^{-c \sqrt{\mathcal{L}}}.
\]

Let us note that the random Schrödinger operators \( H(0) \) satisfying Assumption 1.4 possess the ultra-local property with probability \( \geq 1 - e^{-c \sqrt{\mathcal{L}}} \) provided the length of the interval \( J \) is of order \( \ell^{-\zeta} \) (in fact, a stronger statement holds true, see Theorem 1.14 below). Unfortunately, localization in the usual sense (or in an ultra-local sense for that matter) breaks down under perturbations due to the hybridization phenomenon. As a result, the first step is to identify a weaker notion than ultra-locality that however remains stable under small perturbations.

**Definition 1.14.** We will say that \( H^\mathbb{T}(s) \) possesses a *local structure* in \( J \subset J_{\text{loc}} \) if there exists a disjoint collection \( \{T_\gamma\} \) of subsets of \( \mathbb{T} \) such that \( \text{diam}(T_\gamma) \leq \ell^{3/2} \) for each \( \gamma \) with the following properties:

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

\[
J_\gamma \subset J \quad \text{and} \quad \text{dist}(E_\gamma^\pm, \sigma(H^{T_\gamma}(s))) \geq \Delta;
\]

(ii) (Support of spectral projections) Let \( T := \cup_\gamma T_\gamma \). Then

\[
\| P_{J}(s) \chi_{T \setminus T_\ell} \| \leq e^{-c \sqrt{\mathcal{L}}},
\]

and

\[
\| P_{J_\gamma}(H^{T_\gamma}(s)) - \chi_{\partial T} P_{J_\gamma}(H^{T_\gamma}(s)) \chi_{\partial T} - \chi_{T_\ell} P_{J_\gamma}(H^{T_\gamma}(s)) \chi_{T_\ell} \| \leq e^{-c \sqrt{\mathcal{L}}}.
\]
The unperturbed Hamiltonian possesses a local structure for small, but not too small, $\Delta$. As we shall see in the proof of Theorem 2.1 the local structure is stable under 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 $\beta$ is sufficiently small. The reason for this stability is related to the fact that, under small local perturbations, an eigenstate with energy $E$ is close to the range of a thin spectral projection of the unperturbed operator centered at $E$. Since the latter is supported in the localized patches $T_s$, 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 $T_{\alpha}$), it can hybridize to a number of different patches $T_s$ as $s$ increases.

The scaling of various objects with $\ell$ depends on $q, d$ and our choice of sub-exponential error $\exp(-c\sqrt{\ell})$. The correct scaling of $\Delta$ and $\beta$ to ensure the 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 3.5) to control the behavior of the individual spectral patches $P_{T_s}(H^T(\gamma))$, invoking Definition 1.14. This in turn allows us to control the physical evolution of spectral data $Q(s)$ for $H^T(\gamma)$ near the energy $E$ (see Section 3.3 for details). Finally, we show that this 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 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, a feature verified in the previous step.

2. Adiabatic theorem on a torus

We will use the shorthand $P_J(s) := P_J(H^T(s))$ and $P_J := P_J(0)$ in this section.

We will show in Section 4 that Anderson-type models possess a local structure in the sense of Definition 1.14. In fact, a stronger statement holds true:

**Theorem 2.1 (Local structure of $H^T(s)$).** Suppose that $H$ satisfies Assumptions 1.3, 1.4 and the family $H(s)$ satisfies Assumption 1.1. Let

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

(2.1)

and suppose that $\beta \leq \ell^{-p_1}$. Then, there exist constants $c, c_1, c_2, c_3, c_4, c_5, c_6$ such that for $\ell$ large enough $H^T(s)$ possesses a local structure for the energy interval $J = (E - 6\delta, E + 6\delta)$: One can find a disjoint collection $\{T_{\gamma}\}$ of subsets of $\Lambda$ such that $|T_{\gamma}| \leq c_4 \ell^d, \text{diam}(T_{\gamma}) \leq c_5 \ell^{3/2}$ for each $\gamma$ and the following conditions hold true with probability $> 1 - e^{-c_6 \sqrt{\ell}}$:

(i) (Local Gap) There exist intervals $T_{\gamma} = [E^-_{\gamma}, E^+_{\gamma}]$ such that

$$\text{(2.2)}$$

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

$$\|P_J(s)\chi_{\Lambda \setminus T_{\gamma}}\| \leq e^{-c\sqrt{\ell}},$$

(2.3)

and

$$\|P_{\ell}(s) - \chi_{\partial T} P_{\ell}(s) \chi_{\partial T} - \chi_{T_{\gamma}} P_{\ell}(s) \chi_{T_{\gamma}}\| \leq e^{-c\sqrt{\ell}}.$$  

(2.4)

(iii) (Exponential Decay of Correlations) Let $A_{\alpha} = \partial T_{\gamma} \cup (T_{\gamma})_{s\ell}$, then (with $A = A_{\alpha}$ in 1.20–1.21) we have

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

(2.5)

for $z \in \mathbb{C}$ with $\text{Re}(z) = E^\pm_{\gamma}$.

The dependence on $\beta$ here is deterministic, i.e., there exists 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 that we will establish is

**Theorem 2.2 (Locabatic theorem for distorted Fermi projection).** In the setting of Theorem 2.2, assume in addition that\[ e^{-c\sqrt{T}} \leq \epsilon \leq \epsilon^{-p_2}, \] and fix \( N \in \mathbb{N} \). Then for \( \ell \) large enough, there exists a smooth family of orthogonal projections \( Q(s) \) with the following properties:

(i) \( \|Q(s), H^T(s)\| \leq C_N \left( \epsilon \Delta^{-1} + e^{-c\sqrt{T}} \right) \);
(ii) \( \|P_{s-E-6\delta}(H^T(s))Q(s)\| + \|Q(s)P_{s+E+6\delta}(H^T(s))\| \leq C_N \left( \epsilon \Delta^{-1} + e^{-c\sqrt{T}} \right) \);
(iii) If we denote by \( Q_c(s) \) the solution of IVP \( i\epsilon \dot{Q}_c(s) = [Q_c(s), H^T(s)] \), \( Q_c(0) = Q(0) \), we have
\[
\|Q_c(s) - Q(s)\| \leq C_N \left( \frac{1}{\Delta N} + \frac{1}{\delta^2 N+1} \right) + e^{-c\sqrt{T}}.
\] Furthermore, for \( s = 0 \) and \( s = 1 \), the inequalities in (i) and (ii) hold without the terms proportional to \( \epsilon \).

### 3. Adiabatic theory for localized spectral patches

Throughout this section we will work on the torus, in the setting of Theorem 2.2. To simplify the notation, we will shorthand \( H(s) := H^T(s) \) in this section.

In addition, we will use the assumption of Theorem 2.2, namely that
\[ e^{-c\sqrt{T}} \leq \epsilon \leq \epsilon^{-p_2}, \quad p_2 > \max\{d + \frac{1}{2}, \frac{d}{q}, \frac{d}{2q} \}. \]
This, in particular, implies that for \( \ell \) large enough \( \epsilon^{-1}e^{-c\sqrt{T}} \leq e^{-c\sqrt{T}} \) and that \( \epsilon/\Delta \ll 1 \). We will use this repeatedly. We will also assume that \( 1 \geq \Delta \geq \beta > 0 \) (in fact, the above conditions imply \( \Delta \gg \beta \) for large \( \ell \), but this will only matter later on).

#### 3.1. 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 3.1.**

(a) \( \|H(s)\| \leq C \) and \( \|H^{(k)}(s)\| \leq \beta C_k \) 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 \( \min_{s \in [0,1]} \text{dist}(\sigma(H(s)), \{E_1, E_2\}) \geq 2\Delta \);
(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
\[
X_A(s) = \frac{1}{2\pi} \sum_{j=1}^2 (-1)^j \int_{-\infty}^{\infty} R_{ix+E_j}(s) A R_{ix+E_j}(s) \, dx.
\] (3.1)

This operator was introduced by Kato in his work on the adiabatic theorem, and henceforth we will refer to it as Kato’s operator.

We note that, for \( H(s) \) satisfying Assumption 3.1,
\[
\max_{j=1,2} \left\| R_{ix+E_j}(s) \right\| \leq (x^2 + \Delta^2)^{-1/2}
\] (3.2)
and consequently
\[
\left\| X_A(s) \right\| \leq \frac{\|A\|}{\pi} \int_{-\infty}^{\infty} (x^2 + \Delta^2)^{-1} \, dx \leq \Delta^{-1} \|A\|.
\] (3.3)

Using the Leibniz rule and (C8), it is straightforward to see that, more generally,
\[
\left\| X_A^{(k)}(s) \right\| \leq C_k \|A\|, \quad k \in \mathbb{Z}_+,
\] (3.4)
Lemma 3.3

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

$$[H(s), X_A(s)] = [P(s), A],$$

which plays a role in a construction of adiabatic theory for gapped Hamiltonians, particularly 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 Kato’s operator.

Lemma 3.2. Let $A(s)$ be a smooth family of operators on $[0, 1]$. Suppose that in addition to Assumption 3.1 there exists some set $A$ and $M, c > 0$ such that

$$\|R_{ix+E_j}(s)\|_{c,\ell} \leq M(x)^{-1}, \quad j = 1, 2.$$  \hspace{1cm} (3.7)

Then,

$$\left\|e^{\phi_A} X_A^{(1)}(s) \right\| \leq C \left( \beta M^2 |\ln \Delta| + \beta M \Delta^{-1} \right) \left\|e^{\phi_A} A(s) \right\| + CM |\ln \Delta| \left\|e^{\phi_A} A^{(1)}(s) \right\|. \hspace{1cm} (3.8)$$

Proof. We will suppress the $s$-dependence in the proof below. Using (3.8) and (1.22), we can bound

$$\left\|e^{\phi_A} X_A^{(1)} \right\| \leq \sum_{j=1}^{2} \left( \frac{C\beta}{\pi} \int_{-\infty}^{\infty} \|R_{ix+E_j}\|_{c,\ell} \left\|e^{\phi_A} A \right\| \|R_{ix+E_j}\| dx \right.$$

$$+ \frac{1}{\pi} \int_{-\infty}^{\infty} \|R_{ix+E_j}\|_{c,\ell} \left\|e^{\phi_A} A^{(1)} \right\| \|R_{ix+E_j}\| dx$$

$$+ \frac{C\beta}{\pi} \int_{-\infty}^{\infty} \|R_{ix+E_j}\|_{c,\ell} \left\|e^{\phi_A} A \right\| \|R_{ix+E_j}\|^2 dx \right).$$

Using (3.7) and Assumption 3.1(b) we get (3.8). \hfill \Box

3.2. Nenciu’s expansion. An elegant approach for the analysis of the adiabatic behavior of gapped systems was discovered by Nenciu [44]. We will use it as a starting point for our construction.

Lemma 3.3 (Nenciu’s expansion). Let $H(s)$ be a smooth family of self-adjoint operators on $[0, 1]$ that satisfies Assumption 3.1. Let $B_n(s)$ be a smooth family defined recursively as follows: $B_0(s) = P(s)$ and, for $n \in \mathbb{N}$,

$$B_n(s) = \left( P(s) X_{B_{n-1}(s)}(s) P(s) + h.c. \right) + S_n(s) - 2P(s)S_n(s)P(s), \hspace{1cm} (3.9)$$

where

$$S_n(s) = \sum_{j=1}^{n-1} B_j(s) B_{n-j}(s). \hspace{1cm} (3.10)$$

We then have

(i) \hspace{1cm} $\dot{B}_n(s) = -i [H(s), B_{n+1}(s)]$ \hspace{1cm} (3.11)

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

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

(iii) We have

$$\sup_s \left\|B_n^{(k)}(s) \right\| \leq C_{n,k} \Delta^{-n}, \quad k, n \in \mathbb{Z}_+. \hspace{1cm} (3.12)$$
Proof. Property \(3.3.1\) is due to Nenciu, \[44\]. Property \(3.3.2\) follows directly from the recursive definition of \(B_n\)'s. We establish \(3.3.3\) by induction:

**Induction base:** For \(n = 0\) and an arbitrary \(k\), the bound \(\|B_0^{(k)}(s)\| \leq Ck\) in \(3.3.3\) can be seen from \([C.7], [C.8]\). 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 \(3.9\) \(k\) times with \(n = n_o\) using the Leibniz rule and then using \(3.2\) and \(3.4\), we get that it also holds for \(n = n_o\) and all \(k \in \mathbb{Z}_+\). \(\square\)

For localized spectral patches, we slightly modify the statement.

**Lemma 3.4.** Suppose that in addition to the assumptions of Lemma \(3.3\) there exists some set \(A\) and \(M, c > 0\) such that \(3.7\) holds. Let us also assume that

\[
\max_{s \in [0,1]} \left\| e^{\rho A} P(s) \right\| \leq C, \quad \max_{s \in [0,1]} \left\| H^{(k)}(s) \right\|_{c \ell} \leq Ck\beta \text{ for } k \in \mathbb{N}.
\]

Let

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

and assume that \(\beta \leq \nu\). Then the operators \(B_n\) defined in Lemma \(3.3\) satisfy

\[
\left\| e^{\rho A} B_n^{(k)}(s) \right\| \leq C_{n,k} \nu^{-n}, \quad k, n \in \mathbb{Z}_+.
\]

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

**Induction base:** For \(n = 0\) and an arbitrary \(k\), by the Leibniz rule we have

\[
P(n) = (P_{n+1})^{(n)} = \sum_{k_1+k_2+\cdots+k_{n+1} = n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1 \leq j \leq n+1} P^{(k_j)},
\]

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

Using the integral representation \([C.7]\), the formula \([C.8]\), the Leibniz rule, \(3.7\), \(1.22\), and Assumption \(3.13\), we can bound

\[
\left\| P^{(k)} \right\|_{c \ell} \leq C_k M^k, \quad k \in \mathbb{N}.
\]

We can now use \(1.22\) and \(3.15\) to deduce that

\[
\left\| e^{\rho A} P^{(n)} \right\| \leq \sum_{k_1+k_2+\cdots+k_{n+1} = n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1 \leq j < j_0} P^{(k_j)} \left\| e^{\rho A} P \right\| \prod_{j_0 \leq j \leq n+1} \left\| P^{(k_j)} \right\|
\]

\[
\leq \sum_{k_1+k_2+\cdots+k_{n+1} = n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1 \leq j \leq n+1} C_{k_j} M^{k_j} = C_n M^n,
\]

where \(j_0\) 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 \(3.9\) \(k\) times with \(n = n_o\) using the Leibniz rule and then using Lemma \(3.2\) (the assumption there is satisfied by Eq. \(2.5\)), we get the induction step. \(\square\)

### 3.3. Gapped adiabatic theorem

An immediate consequence of Lemma \(3.3\) is

**Lemma 3.5** (Gapped adiabatic theorem to all orders). In the setting of Lemma \(3.3\), let \(P_N(s) := \sum_{n=0}^N e^n B_n(s)\). Then for all \(N \in \mathbb{N}\),

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

where \(U_\epsilon\) was defined in \(1.4\).
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 3.3,
\[
\epsilon \dot{P}_N(s) = -i[H(s), P_N(s)] + \epsilon^{N+1} \dot{B}_N(s).
\]
Using the fundamental theorem of calculus, we obtain
\[
U_\epsilon(s)^* P_N(s) U_\epsilon(s) - P_N(0) = \epsilon^{-1} \int_0^s \epsilon^{N+1} \frac{d}{ds} (U_\epsilon(s)^* B_N(s) U_\epsilon(s)).
\]
Using the unitarity of $U_\epsilon$, Assumption 3.1 and Lemma 3.3.(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 $P_N(1) = P(1)$. \hfill $\square$

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

**Theorem 3.6 (Locabatic theorem on a torus).** Suppose that the family $H(s)$ satisfies Assumption \ref{assumption1} and $H(0)$ satisfies Assumptions \ref{assumption1} \ref{assumption2}. Let $\mathcal{G}_\omega$ be the event that $H^T(0)$ possesses an ultra-local structure for the energy interval $J = \{E - 6\delta, E + 6\delta\}$. Then $P(\mathcal{G}_\omega) > 1 - e^{-c\sqrt{T}}$.

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 \ref{eq:1.2}, satisfies
\[
\max_{s \leq [0,1]} \| \tilde{P}_{J_\gamma}(H^{T_\gamma}(s)) \psi(s) \| \leq C \left( \epsilon^{N-1} + e^{-c\sqrt{T}} \right)
\]
for some $\gamma$. Furthermore, for any $N \in \mathbb{N}$, we can further improve \ref{eq:3.17} for $s = 1$:
\[
\| \tilde{P}_{J_\gamma}(H^{T_\gamma}(1)) \psi(1) \| \leq C_N \left( \epsilon^N (\Delta^{-N} + \delta^{-2N-1}) + e^{-c\sqrt{T}} \right).
\]

**Proof of Theorem 3.6.** We have already established the first part of the assertion in Theorem \ref{thm:adiabatic}. We now show the second part. We first note that $\mathcal{G} \subset \Omega_{\text{loc},N}$ of the full configuration space for which $\mathbb{T}$ and all sets in $\{ \mathcal{T}_\gamma \}$ are $\ell/10$-localizing, see Lemma \ref{lem:localization} below. Thus, Theorem \ref{thm:adiabatic} \ref{item:adiabatic2} implies the existence of the patch $\mathcal{T}_\gamma$ such that $\| \chi(\mathcal{T}_\gamma_{\omega}) \psi \| \leq e^{-c\sqrt{T}}$. It then follows from Lemma \ref{lem:localization} below, specifically \ref{lem:localization2}, that $E \in J_\gamma$ (see also \ref{eq:2.2}). Let $\mathcal{T}_\gamma = (\mathcal{T}_\gamma)_{\Delta l}$ and set
\[
Q_\gamma(s) = \chi_{\mathcal{T}_\gamma} P_{J_\gamma}(H^{T_\gamma}(s)) \chi_{\mathcal{T}_\gamma}.
\]

By Lemma \ref{lem:localization} specifically \ref{lem:localization3}, we know that \ref{eq:3.18} 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
\[
\rho = \chi_{\mathcal{T}} \tilde{P}_{J_\gamma}(H^{T_\gamma}(0)) \chi_{\mathcal{T}} + \chi_{\mathcal{T}},
\]
we deduce that $\| \rho \psi \| \leq e^{-c\sqrt{T}}$. Hence, by the unitarity of the quantum evolution,
\[
\| \tilde{\rho}_\epsilon(s) \psi(s) \| \leq e^{-c\sqrt{T}}
\]
for all $s$, where $\rho_\epsilon(s)$ denotes the (full) Heisenberg evolution of the (truncated) initial spectral patch $\rho := Q_\gamma(0)$, i.e.,
\[
i \epsilon \tilde{\rho}_\epsilon(s) = [H(s), \rho_\epsilon(s)], \quad \rho_\epsilon(0) = \rho.
\]
Therefore the result follows from
Lemma 3.7. (i) We can estimate
\[
\max_{s \in [0,1]} \|\rho_e(s) - Q_{\gamma}(s)\| \leq C \left( \epsilon \Delta^{-N} + e^{-c\sqrt{\gamma}} \right).
\] (3.22)
Moreover, for any \( N \in \mathbb{N} \), we have
\[
\max_{s \in \{0,1\}} \|\rho_e(s) - Q_{\gamma}(s)\| \leq C_N \left( \epsilon^N \left( \Delta^{-N} + \delta^{-2N-1} \right) + e^{-c\sqrt{\gamma}} \right).
\] (3.23)
(ii) In addition,
\[
\max_{s \in [0,1]} \left\| \bar{P}_{J_{\gamma}}(H^{T_\gamma}(s)) - \bar{P}_{J_{\gamma}}(H^{T_\gamma}(s)) \right\| \bar{Q}_{\gamma}(s) \right\| \leq e^{-c\sqrt{\gamma}}.
\] (3.24)
\[\square\]

Remark 3.8. We note that in the proof of Theorem 3.6, the initial spectral data \( \psi_n \) can be replaced by any vector \( \psi \in \text{Ran}(PE - \delta, E + \delta) \) that satisfies \( \left\| \bar{\chi}(\gamma,\psi) \right\| \leq e^{-c\sqrt{\gamma}} \) for some patch \( T_\gamma \).

Proof of Lemma 3.7. We suppress the \( s \) dependence in the proof. The property (3.24) can be seen by decomposing \( \bar{P}_{J_{\gamma}}(H^{T_\gamma}) = \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \bar{Q}_{\gamma} + \bar{P}_{J_{\gamma}}(H^{T_\gamma}) Q_{\gamma} \) and noticing that
\[
\bar{P}_{J_{\gamma}}(H^{T_\gamma}) Q_{\gamma} = \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \chi_{T_\gamma} \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \chi_{T_\gamma} \chi_{T_\gamma} \gamma
\]
\[
= \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \chi_{T_\gamma} \gamma + O\left( e^{-c\sqrt{\gamma}} \right),
\]
thanks to (2.4).

Lemma 3.7.(i). By our assumption, \( H^{T_\gamma} \) is a gapped Hamiltonian with gap \( \Delta \). Following the argument in Section 3.2, we set \( B_0^n \) the n-th order in the Nenciu's expansion. Explicitly, we use Lemma 3.3 with \( B_0^n = \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \).

We set
\[
Q_{\gamma,N} := \sum_{n=0}^N \epsilon^n \chi_{T_\gamma} B_0^n \chi_{T_\gamma}.
\] (3.25)
and proceed to show that
\[
\max_{s} \|\rho_e - Q_{\gamma,N}\| \leq C_N \left( \epsilon^N \left( \Delta^{-N} + \delta^{-2N-1} \right) + e^{-c\sqrt{\gamma}} \right).
\] (3.26)
The result then follows immediately from (3.20) by the definition of \( Q_{\gamma,N} \) and Lemma 3.3.(ii) (we recall that \( B_0^n = \bar{P}_{J_{\gamma}}(H^{T_\gamma}) \)).

To get (3.20), we observe that by (3.11),
\[
\epsilon \hat{Q}_{\gamma,N} = -i \sum_{\gamma} \sum_{n=0}^N \epsilon^{n+1} \chi_{T_\gamma} \left[ H^{T_\gamma}, B_0^{n+1} \right] \chi_{T_\gamma}
\]
\[
= -i \left[ H, Q_{\gamma,N} \right] - i \epsilon^{N+1} \chi_{T_\gamma} B_0^N \chi_{T_\gamma}
\]
\[+ \left( i \sum_{\gamma} \sum_{n=0}^N \epsilon^{n+1} \left[ H^{T_\gamma}, \chi_{T_\gamma} \right] B_0^N \chi_{T_\gamma} + \text{h.c.} \right),
\]
where we have used \( H^{T_\gamma}(s) \chi_{T_\gamma} = H(s) \chi_{T_\gamma} \). We bound the second term on the second line by \( C_N \epsilon^{N+1} \Delta^{-N} \) using (3.12). For the term on the third line, we note that
\[
\left\| \left[ H^{T_\gamma}(s), \chi_{T_\gamma} \right] B_0^N(s) \right\| \leq \nu^{-n-1} e^{-c\sqrt{\gamma}}
\]
using Lemma 3.3. Putting these bounds together, we get
\[
\left\| \epsilon \hat{Q}_{\gamma,N} + i \left[ H, Q_{\gamma,N} \right] \right\| \leq C_N \epsilon^{N+1} \Delta^{-N} + C e^{-c\sqrt{\gamma}}.
\] (3.27)
Finally, we observe that

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

where $U_\epsilon(t,s)$ was defined in (1.4).

Integrating over $s$ and using (3.27), we deduce that

$$\|U_\epsilon(t,r)Q_{\gamma,N}(r)U_\epsilon(r,t) - Q_{\gamma,N}(t)\| \leq \epsilon^{-1} \left( C_N \epsilon^{N+1} \Delta^{-N} + C \epsilon^{-C \sqrt{T}} \right),$$

(3.28)

We now note that $Q_{\gamma,N}(0) = \rho_s$, so $U_\epsilon(t,0)Q_{\gamma,N}(0)U_\epsilon(0,t) = \rho_\epsilon(t)$ by uniqueness of the solution for the IVP (3.21). Combining this with (3.28) yields (3.26).

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

$$Q(s) := \sum_\gamma Q_\gamma(s),$$

(3.29)

where the spectral patch $Q_\gamma$ was defined in (3.19). Our first assertion encapsulates the basic properties of this operator.

**Lemma 3.9.** For $\ell$ large enough, the operator $Q(s)$ satisfies:

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

$$\max_{s \in [0,1]} \left\| \frac{d^j Q(s)}{ds^j} \right\| \leq C j \beta, \quad j = 1, \ldots, k;$$

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

$$\| [H(s), Q(s)] \| \leq C \epsilon^{-C \sqrt{T}};$$

(3.30)

(iii) Almost projection:

$$\left\| \bar{Q}(s) Q(s) \right\| \leq C \epsilon^{-C \sqrt{T}};$$

(3.31)

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

$$\| \bar{P}_{J_+}(s) Q(s) \| \leq C \epsilon^{-C \sqrt{T}}, \quad \| \bar{Q}(s) P_{J_-}(s) \| \leq C \epsilon^{-C \sqrt{T}}.$$  

(3.32)

**Proof.** Lemma 3.9.(i) Note that, for $\ell$ large enough, $\beta \ll \Delta$. The assertion follows from the integral representation (C.7) for $P_{J_+}(H^{T_\gamma}(s))$ with $E_{1,2} = E^\gamma_{1,2}$, the formula (C.8), (3.2), and the Leibniz rule.

Lemma 3.9.(ii) We compute

$$[H(s), Q_{\gamma}(s)] = \left[ H^{T_\gamma}(s), Q_{\gamma}(s) \right] = \left[ H^{T_\gamma}(s), \chi_{\gamma} \right] P_{\gamma}(s) H^{T_{\gamma}}(s) \chi_{\gamma} + \chi_{\gamma} P_{\gamma}(s) H^{T_{\gamma}}(s) \left[ H^{T_\gamma}(s), \chi_{\gamma} \right],$$

and estimate both terms by $C \epsilon^{-C \sqrt{T}}$ using Assumption 1.2 and Theorem 2.1.(ii).

Lemma 3.9.(iii) We note that, for disjoint sets $\Omega_\gamma$,

$$\| \sum_\gamma \chi_{\gamma} A_{\gamma} \chi_{\gamma} \| \leq \max_\gamma \| \chi_{\gamma} A_{\gamma} \chi_{\gamma} \|.$$  

(3.33)

Since $T_\gamma$ are disjoint, we have

$$\| \bar{Q}(s) Q(s) \| = \left\| \sum_\gamma \chi_{\gamma} P_{\gamma}(s) H^{T_\gamma}(s) \chi_{\gamma} P_{\gamma}(s) H^{T_\gamma}(s) \chi_{\gamma} \right\|.$$
The right hand side is bounded by \( C e^{-c \sqrt{T}} \) using Theorem 2.1.(ii).

Lemma 3.9.(iv) We apply Lemma C.3 with \( H_1 = H(s) \), \( H_2 = H^T(s) \), and \( R = \chi_\mathcal{P} \) to bound
\[
\| \hat{P}_{J_+}(s) \chi_\mathcal{P} P_J(H^T(s)) \| \leq C e^{-c \sqrt{T}},
\]
where we have used (2.2) and the fact that \( H(s) \) has range \( r \). Since
\[
Q(s) \leq \chi_\mathcal{P} P_J(H^T(s))\chi_\mathcal{P}
\]
by (2.2), we deduce that
\[
\| \hat{P}_{J_+}(s) Q(s) \| \leq \| \hat{P}_{J_+}(s) \chi_\mathcal{P} P_J(H^T(s)) \| \leq C e^{-c \sqrt{T}}.
\]

On the other hand, letting \( J' = (E - 3\delta, E + 3\delta) \) and using Lemma C.3 with \( H_1 = H^T(s) \) and \( H_2 = H(s) \), we get
\[
\| \hat{P}_{J'}(H^T(s)) \chi_\mathcal{P} P_{J_-}(s) \| \leq C e^{-c \sqrt{T}}
\]
Since
\[
\tilde{Q}(s) \leq \chi_{\mathcal{A} \setminus \mathcal{P}} + \chi_\mathcal{P} \hat{P}_{J'}(H^T(s))\chi_\mathcal{P}
\]
by (2.2), we deduce that
\[
\| \tilde{Q}(s) P_{J_-}(s) \| \leq \| \chi_{\mathcal{A} \setminus \mathcal{P}} P_{J_-}(s) \| + \| \hat{P}_{J_+}(s) \chi_\mathcal{P} P_{J_-}(s) \| \leq C e^{-c \sqrt{T}},
\]
using (2.3) to bound the first term on the right hand side.

\[ \Box \]

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

**Lemma 3.10.** Let \( N \in \mathbb{N} \). Suppose that \( \ell \) is sufficiently large. Then there exists a smooth family of projections \( Q_s \) with the following properties:

(i) \[
\max_{s \in [0,1]} \| [Q_s, H(s)] \| \leq C \left( \epsilon + e^{-c \sqrt{T}} \right)
\] (3.34)

and
\[
\max_{s \in [0,1]} \| [Q_s, H(s)] \| \leq C N \epsilon^{N+1} \Delta^{-N} + C e^{-c \sqrt{T}}; \quad (3.35)
\]

(ii) Let \( J_+ = (E - 6\delta, E + 6\delta) \) and \( J_- = (E - \delta, E + \delta) \). Then
\[
\max_{s \in [0,1]} \left( \| \hat{P}_{J_+}(s) Q_s \|, \| \hat{Q}_s P_{J_-}(s) \| \right) \leq C \left( \epsilon \Delta^{-1} + e^{-c \sqrt{T}} \right)
\] (3.36)

and
\[
\max_{s \in [0,1]} \left( \| \hat{P}_{J_+}(s) Q_s \|, \| \hat{Q}_s P_{J_-}(s) \| \right) \leq C e^{-c \sqrt{T}}
\] (3.37)

(iii) \( Q_0^{(k)} = Q_1^{(k)} = 0 \) for all \( k \in \mathbb{Z}_+ \) and
\[
\max_{s \in [0,1]} \| Q_s^{(k)} \| \leq C_k \beta, \quad k \in \mathbb{N};
\]

(iv) \[
\| \epsilon \hat{Q}_s + i [H(s), Q_s] \| \leq C N \epsilon^{N+1} \Delta^{-N} + C e^{-c \sqrt{T}}; \quad (3.38)
\]

(v) If we denote by \( Q_\epsilon(s) \) the solution of the IVP \( i \epsilon \dot{Q}_\epsilon(s) = [H(s), Q_\epsilon(s)] \), \( Q_\epsilon(0) = Q_0 \), then we have
\[
\max_{s \in [0,1]} \| Q_\epsilon(s) - Q_s \| \leq C N \epsilon^{N} \Delta^{-N} + C e^{-c \sqrt{T}}.
\] (3.39)
Proof. We set
\[ Q_N(s) := \sum_\gamma Q_{\gamma,N}(s), \quad (3.40) \]
where \( Q_{\gamma,N} \) was defined in \( 3.25 \), and first show that the assertions of the lemma hold if we replace \( Q_s \) with \( Q_N(s) \) there. Note that the latter operator is not a projection.

It follows from Lemma \( 3.3 \) and the hypothesis \( \epsilon \leq \Delta \) that
\[ \|Q_N(s) - Q_0(s)\| = \|Q_N(s) - Q(s)\| \leq C_N \epsilon \Delta^{-1}. \quad (3.41) \]
Hence, combining this bound with Lemma \( 3.9 \) we conclude that \( Q_N(s) \) satisfies the properties \( 3.10.(iii) \) and \( 3.10.(i) \).

We next observe that the property \( 3.10.(iv) \) holds for \( Q_N(s) \) by \( 3.27 \), Assumption \( 1.2 \) and \( 3.33 \).

The property \( 3.10.(v) \) is established by replicating the argument employed in the proof of Lemma \( 3.7.(i) \).

Finally, the property \( 3.10.(i) \) holds for \( Q_N(s) \) by the properties \( 3.10.(iii) \) and \( 3.10.(iv) \) 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
\[ \|Q_\epsilon(t)Q_\epsilon(t)\| = \|Q(0)Q(0)\| \leq Ce^{-c\sqrt{t}} \]
Thus, by the triangle inequality, we get
\[
\|Q_N(t)Q_N(t)\| \leq \|Q_N(t)Q_N(t) - Q_\epsilon(t)Q_\epsilon(t)\| + Ce^{-c\sqrt{t}}
\leq (\|Q_N(t)\| + \|Q_\epsilon(t)\|) \|Q_N(t) - Q_\epsilon(t)\| + Ce^{-c\sqrt{t}}
\leq C_N e^{c\Delta^{-N}} + Ce^{-c\sqrt{t}},
\]
where in the last step we have used the properties \( 3.10.(iii) \) and \( 3.10.(v) \) for \( Q_N \).

It follows that
\[ \max_\sigma \text{dist} (\sigma(Q_N(s)), \{0, 1\}) \leq C_N e^{c\Delta^{-N}} + Ce^{-c\sqrt{t}}. \]

If \( \epsilon/\Delta \) is small enough and \( \ell \) large enough, 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 \([1, \frac{3}{2}]\). Then by functional calculus for self-adjoint operators and the triangle inequality, Lemma \( 3.10.(i) \) and \( 3.10.(ii) \) hold for this operator. To establish Lemma \( 3.10.(iii) \) we use the following integral representation for \( Q_s \):
\[ Q_s = (2\pi i)^{-1} \oint_\Gamma (Q_N(s) - z)^{-1} \, dz, \quad \Gamma = \{ z \in C : |z - 1| = 1/2 \}. \quad (3.42) \]
Since
\[ \partial_s (Q_N(s) - z)^{-1} = - (Q_N(s) - z)^{-1} \partial_s Q_N(s) (Q_N(s) - z)^{-1}, \]
and \( \| (Q_N(s) - z)^{-1} \| \) is uniformly bounded for \( z \in \Gamma \), the property \( 3.10.(iii) \) follows by the Leibniz rule and the bounds on \( Q_N^{(k)}(s) \).

Lemma \( 3.10.(iv) \)
\[
\hat{Q}_s = - (2\pi i)^{-1} \oint_\Gamma (Q_N(s) - z)^{-1} \hat{Q}_N(s) (Q_N(s) - z)^{-1} \, dz
\]
\[ = -i (2\pi i)^{-1} \oint_\Gamma (Q_N(s) - z)^{-1} [H(s), Q_N(s)] (Q_N(s) - z)^{-1} \, dz
\]
\[ - (2\pi i)^{-1} \oint_\Gamma (Q_N(s) - z)^{-1} \left( \hat{Q}_N(s) - i[H(s), Q_N(s)] \right) (Q_N(s) - z)^{-1} \, dz,
\]
and the statement follows from the properties \( 3.10.(iv) \) and \( 3.10.(i) \) already proved for \( Q_N(s) \).

For \( s \in \{0, 1\} \), we have \( Q_N(s) = Q(s) \), so \( 3.35 \) and \( 3.37 \) follow from Lemma \( 3.9 \). \( \square \)
3.6. Adiabatic behavior of the distorted Fermi projection. The idea behind the proof of Theorem 2.2 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 \( H(s) = Q_s H(s) Q_s \). By Lemma 3.10, \( Q_s \) is close to a spectral projection of \( H(s) \) and so the spectrum of \( H(s) \) 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 3.11. Let \( I = (E - \delta/2, E + \delta/2) \). Suppose that \( \ell \) is large enough. Then we have

\[
\max_{s \in [0,1]} \left\| H(s)^{(k)} \right\| \leq C_k \quad \text{for} \quad k = 1, \ldots, N.
\]

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

\[
(H - w)^2 = \bar{Q} (H - w) \bar{Q} (H - w) \bar{Q} = \bar{Q} (H - w)^2 \bar{Q} - \bar{Q} HQH \bar{Q} \geq \bar{Q} P_{J^c} (H - w)^2 \bar{Q} + \bar{Q}[H, Q][H, Q] \bar{Q},
\]

while we can bound

\[
\bar{Q} P_{J^c} (H - w)^2 \bar{Q} \geq \frac{\delta^2}{4} \bar{Q} P_{J^c} \bar{Q} = \frac{\delta^2}{4} \bar{Q} - \frac{\delta^2}{4} \bar{Q} P_{J^c} \bar{Q} \geq \frac{\delta^2}{4} \bar{Q} - \frac{\delta^2}{4} \left( C_N \epsilon + C \exp \left( -c \sqrt{T} \right) \right)^2 \bar{Q},
\]

using Lemma 3.10 (3.10(ii)) and

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

using Lemma 3.10 (3.10(i)). Hence

\[
(H - w)^2 \geq \left( \frac{\delta^2}{4} / 2 \left( C_N \epsilon + C \exp \left( -c \sqrt{T} \right) \right)^2 \bar{Q} \right) > 0
\]
on \text{Ran} (\bar{Q}).

The bound (3.43) follows from Lemma 3.10 (3.10(iii)), Assumption 1.2 and the Leibniz rule.

Lemma 3.12. Let \( T(s, s') \) be the unitary semigroup generated by \( i(\bar{Q}_s, Q_s) \), i.e., \( T(s) \) is the solution of the IVP

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

Then \( T(s, s') \) satisfies

\[
T(s, s') Q s' = Q s T(s, s').
\]

Suppose in addition that \( \epsilon / \Delta \) is small enough and \( \ell \) is sufficiently large. Then

\[
\max_s \left\| T^{(k)}(s, 0) \right\| \leq C_k \beta \quad \text{for} \quad k = 1, \ldots, N.
\]

Proof. The intertwining relation (3.45) follows from observing that

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

and \( T(s', s') Q s T(s', s') = Q s' \).

The bound (3.46) follows from Lemma 3.10 (3.10(iii)) the unitarity of \( T \), and the Leibniz rule.

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

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

Let \( Q^+_\epsilon (Q^-_\epsilon) \) be the spectral projection of \( H_\epsilon \) associated with the interval \((E, \infty) ((-\infty, E) \) respectively).
Lemma 3.13. Suppose that $\ell$ is large enough. Then we have
\[
\max_s \|Q^+_1 U_\epsilon(s,0)Q_0^+\| \leq C \left(\epsilon^{\Delta^{-1}} + e^{-\ell^{\frac{1}{2}}}\right) \tag{3.47}
\]
and
\[
\|Q^+_1 U_\epsilon(1,0)Q_0^+\| \leq C_N \left(\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-2N-1}\right) + Ce^{-\ell^{\frac{1}{2}}}. \tag{3.48}
\]

**Proof.** We first note that Lemma 3.1 implies that
\[
\|Q_s U_\epsilon(s,s')Q_s'\| \leq C_N \epsilon^N \Delta^{-N} + Ce^{-\ell^{\frac{1}{2}}}. \tag{3.49}
\]
Indeed, using the semigroup property for $U_\epsilon$,
\[
Q_s U_\epsilon(s,s')Q_s' = Q_s (Q_s - Q_\epsilon(s))U_\epsilon(s,s') - Q_s U_\epsilon(s,s')(Q_s' - Q_\epsilon(s)'),
\]
and both terms on the right hand side can now be bounded using Lemma 3.10.(v).

Let $V_\epsilon(s) = Q_s U_\epsilon(s,0)Q_0$. Then a straightforward computation yields
\[
\begin{align*}
\delta e \partial_s V_\epsilon(s) &= -\delta e Q_\epsilon U_\epsilon(s,0)Q_0 + \delta e \delta H(s) U_\epsilon(s,0)Q_0 \\
&= \delta e [Q_\epsilon, Q_\delta] V_\epsilon(s) + \delta H(s) V_\epsilon(s) + R_\epsilon(s),
\end{align*}
\]
where
\[
R_\epsilon(s) = -\delta e Q_\epsilon Q_\delta U_\epsilon(s,0)Q_0 + \delta H(s) Q_\epsilon U_\epsilon(s,0)Q_0.
\]
We note that
\[
\|R_\epsilon(s)\| \leq C \left(\delta e \|Q_\epsilon\| + \|H(s), Q_\delta\|\right) \|Q_s U_\epsilon(s,0)Q_0\| \leq C_N \epsilon^N \Delta^{-N} + Ce^{-\ell^{\frac{1}{2}}}
\]
by Lemma 3.10 and 3.49.

Let $W_\epsilon(s) = T(0,s)V_\epsilon(s)$, where $T$ was defined in (3.44). Then,
\[
\delta e \partial_s W_\epsilon(s) = T(0,s) \delta H(s) T(s,0)W_\epsilon(s) + T(0,s) R_\epsilon(s).
\]
By Lemma 3.11 the operator $H(s)$ has a gap $\delta$ in its spectrum that separates the associated spectral projections $Q^\pm_s$. This implies that $T(0,s) \delta H(s) T(s,0)$ has the same gap with the associated projections given by $Q^\pm_s := T(0,s) Q^\pm_s T(s,0)$. We can bound
\[
\left\|\left(T(0,s) \delta H(s) T(s,0)\right)^{(k)}\right\| \leq C_k \beta \quad \text{for } k = 1, \ldots, N,
\]
using (3.43), (3.46), and the Leibniz rule.

Let $\tilde{W}_\epsilon(s)$ denote the evolution generated by $T(0,s) \delta H(s) T(s,0)$:
\[
\delta e \partial_s \tilde{W}_\epsilon(s) = T(0,s) \delta H(s) T(s,0) \tilde{W}_\epsilon(s), \quad \tilde{W}_\epsilon(0) = 1. \tag{3.51}
\]
Then, it follows from our previous analysis and the Leibniz rule that $T(0,s) \delta H(s) T(s,0)$ satisfies Assumption 3.1 and the gapped adiabatic theorem to all orders, Lemma 3.5 is applicable. Hence
\[
\max_s \|Q^+_1 \tilde{W}_\epsilon(s)Q_0^+\| \leq C \delta e^{-1}, \quad \|Q^+_1 \tilde{W}_\epsilon(1)Q_0^+\| \leq C_N \epsilon^N \delta^{-N}. \tag{3.52}
\]
We now observe that
\[
W_\epsilon(s) = \tilde{W}_\epsilon(s) + \epsilon^{-1} W_\epsilon(s) \int_0^s W^*_\epsilon(s') T(0,s') R_\epsilon(s') \tilde{W}_\epsilon(s') ds',
\]
so
\[
\|W_\epsilon(s) - \tilde{W}_\epsilon(s)\| \leq \epsilon^{-1} \max_{s' \leq s} \|R_\epsilon(s')\| \leq C_N \epsilon^N \Delta^{-N} + Ce^{-\ell^{\frac{1}{2}}}, \tag{3.53}
\]
using (3.50). We conclude that
\[
\|Q^+_1 V_\epsilon(s)Q_0^+\| = \|Q^+_1 T(s,0) W_\epsilon(s)Q_0^+\| = \|Q^+_1 W_\epsilon(s)Q_0^-\|
\leq \begin{cases} 
C_N \epsilon^N \Delta^{-N} + C \left(\epsilon \delta^{-1} + e^{-\ell^{\frac{1}{2}}}\right) \quad \text{uniformly in } s; \\
C_N \left(\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-N}\right) + Ce^{-\ell^{\frac{1}{2}}} \quad \text{if } s = 1.
\end{cases}
\]

As \( V_\epsilon(s) = Q_s U_\epsilon(s, 0) Q_0 \), and \( \bar{Q}_0 Q_0^- = Q_0^- \), it follows that
\[
\| Q_1^+ U_\epsilon(s, 0) Q_0^- \| \leq \| Q_1^+ V_\epsilon(s) Q_0^- \| + \| Q_1 U_\epsilon(s, 0) \bar{Q}_0 \|
\]
\[
\leq \begin{cases} 
C_N \epsilon^N \Delta^{-N} + C (\epsilon \delta^{-1} + \epsilon^{-c_3}) & \text{uniformly in } s; \\
C_N (\epsilon^N \Delta^{-N} + \epsilon^N \delta^{-N}) + C e^{-c_3} & \text{if } s = 1,
\end{cases}
\]
where in the last step we have used (3.49).

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 \( Q(s) = Q_s^- \).

**Theorem 2.2(i).** Using the integral representation (C.7),
\[
Q_s^- = (2\pi i)^{-1} \oint_{\Gamma} (\bar{H}(s) - z)^{-1} \, dz,
\]
we get
\[
[Q(s), H(s)] = (2\pi i)^{-1} \oint_{\Gamma} (\bar{H}(s) - z)^{-1} [H(s), \bar{H}(s)] (\bar{H}(s) - z)^{-1} \, dz,
\]
and we can bound
\[
\|[Q(s), H(s)]\| \leq C \delta^{-1} \|[H(s), \bar{H}(s)]\|.
\]
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
\[
\|[H(s), \bar{H}(s)]\| \leq C_N \epsilon + C e^{-c_3} \sqrt{\ell}
\]
by Lemma 3.10. Hence
\[
\|[Q(s), H(s)]\| \leq C_N \epsilon \delta^{-1} + C e^{-c_3} \sqrt{\ell},
\]
and 2.2.(i) follows.

**Theorem 2.2(ii).** Using (3.37) and \( Q_s^- \bar{Q}_s = Q_s^- \), we deduce that
\[
\||H(s) - \bar{H}(s)|| P_{E - 6\delta}(H(s))\| + \||H(s) - \bar{H}(s)|| Q(s)\| \leq C_N \epsilon \Delta^{-1} + C e^{-c_3} \sqrt{\ell}.
\]
Hence, we can use Lemma C.3 with \( H_1 = \bar{H}(s), H_2 = H(s), \) and \( R = P_{E - 6\delta}(H(s)) \) to first get
\[
\|[\bar{Q}(s) P_{E - 6\delta}(H(s))]\| \leq C_N \epsilon \Delta^{-1} + C e^{-c_3} \sqrt{\ell},
\]
and then use the same lemma with \( H_1 = H(s), H_2 = \bar{H}(s) \), and \( R = Q(s) \) to get
\[
\|P_{E + 6\delta}(H(s)) Q(s)\| \leq C_N \epsilon \Delta^{-1} + C e^{-c_3} \sqrt{\ell}.
\]

**Theorem 2.2(iii).** This part follows directly from Lemma 3.13 and the \( \pm \) symmetry in the argument there, as
\[
\|Q_\epsilon(s) - Q(s)\| = \|U_\epsilon(s, 0) Q_0^- U_\epsilon(0, s) - Q_0^-\| \leq \|Q_1^+ U_\epsilon(1, 0) Q_0^-\| + \|Q_1^- U_\epsilon(1, 0) Q_0^+\|.
\]

\( \square \)
4. Localization on a Torus

4.1. Consequences of Assumptions \[1.2,1.3\] We first note that Assumptions \[1.2,1.3\] imply localization on a torus as well (e.g., \[4, \text{Theorem 11.2}\]):

\[
sup_{E \in J_{\text{loc}}} \mathbb{E} \left( |(H_{\mathbb{T}} - E - i0)^{-1}(x,y)|^q \right) \leq Ce^{-cd_T(x,y)} \text{ for all } x, y \in \mathbb{T},
\]

(4.1)

where \(d_T(x,y)\) represents the usual distance function on a torus.

Another consequence of these hypotheses is

**Lemma 4.1** (The Wegner estimate). Let \(\Theta \subset \mathbb{T}\). For all \(E \in J_{\text{loc}}\),

\[
\mathbb{P}\{\text{dist}\{E, \sigma(H^\Theta)\} \leq \nu\} \leq C\nu^\alpha|\Theta|.
\]

(4.2)

For a proof, see e.g., \[25\], the proof of Proposition 5.1.

Together with Assumption \[1.3\], Lemma 4.1 yields

**Lemma 4.2** (Distance between spectra). Let \(\Theta, \Phi \subset \mathbb{T}\) be such that \(\text{dist}(\Theta, \Phi) > r\). Then

\[
\mathbb{P}\{\text{dist}\{\sigma(H^\Theta) \cap J_{\text{loc}}, \sigma(H^\Phi) \cap J_{\text{loc}}\} \leq \nu\} \leq C\nu^\alpha|\Theta|.|\Phi|.
\]

(4.3)

More generally, if a collection \(\{\Theta_i\}_{i=1}^n\) of subsets in \(\mathbb{T}\) satisfies \(\text{dist}(\Theta_i, \Theta_j) > r\) for \(i \neq j\), \(|\Theta_i| \leq D\) for all \(i\), and \(E, \sigma \in \mathbb{R}\), then

\[
\mathbb{P}\{\text{dist}\{E, \sigma(H^\Theta_i)\} \leq \nu\text{ for all } i\} \leq (C\nu^\alpha D)^n.
\]

(4.4)

We recall that by \(P_I(H)\) we denote the spectral projection of \(H\) onto a set \(I\), and that \(P_E(H)\) stands for \(P_{(-\infty,E]}(H)\). We will often suppress the \(H\) dependence in this notation, denoting by \(P_{I}^\Theta\) a projection \(P_I(H^\Theta)\) and analogously for \(P_I(H^\mathbb{T})\).

A subtler implication of our assumptions on \(H^\Theta\) is the fact that the associated eigenfunction correlator \(Q^\Theta(x,y;J_{\text{loc}})\) for \(x, y \in \Theta\), defined by

\[
Q^\Theta(x,y;J_{\text{loc}}) = \sum_{\lambda \in \sigma(H^\Theta) \cap J_{\text{loc}}} |P_{\lambda}^\Theta(x,y)|
\]

(4.5)

satisfies

\[
\mathbb{E}Q^\Theta(x,y;J_{\text{loc}}) \leq e^{-c|x-y|_\Theta}
\]

(4.6)

for some \(c > 0\) that depends only on \(\mu\) and \(q\). For the non correlated randomness, see, e.g., \[4\], 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, such an assertion was derived in \[24\], Theorem 4.2.

The relation (4.6) implies that all eigenstates in \(P_{J_{\text{loc}}}^\Theta\) are localized with large probability. We make this statement quantitative below.

**Definition 4.3.** Let \(c, \ell > 0\) be fixed. We say that a set \(\Theta \subset \mathbb{T}\) is \((c, \ell)\)-localizing for \(H^\mathbb{T}\) in the interval \(I \subset J_{\text{loc}}\) if for all eigenpairs \((E_n, \psi_n)_{E_n \in I}\) of \(H^\Theta\) there exists a set \(\{x_n\}_i\) in \(\Theta\) such that

\[
|\psi_n(y)| \leq e^{-c|y-x_n|_\Theta} \text{ for any } y \in \Theta \text{ such that } |y-x_n|_\Theta \geq \sqrt{\ell}.
\]

(4.7)

We then have the following result:

**Theorem 4.4.** Suppose that Assumption \[1.4\] holds. Then there exist \(c > 0\) such that the probability that a set \(\Theta \subset \mathbb{T}\) is \((c, \ell)\)-localizing for \(H^\mathbb{T}\) in the interval \(J_{\text{loc}}\) is \(\geq 1 - C|\Theta|^2e^{-c\ell}\).

For a proof, see e.g., \[4\], Theorem 7.4.

Sometimes it will be useful to compare a finite volume projection \(P_{J_{\text{loc}}}^\mathbb{T}\) with the infinite volume one \(P_E\). To be able to do so, we will use the periodic extension \(\hat{P}_E^\mathbb{T}\) of \(P_E^\mathbb{T}\) to \(\mathbb{Z}^d\), i.e.,

\[
\hat{P}_E^\mathbb{T}(x,y) = \begin{cases} P_E^\mathbb{T}(x \text{ mod } \mathbb{L}Z^d, y \text{ mod } \mathbb{L}Z^d) & x - y \in \mathbb{T} \\ 0 & x - y \notin \mathbb{T} \end{cases}
\]

The next assertion implies that deep inside \(\mathbb{T}\), \(P_E\) and \(\hat{P}_E^\mathbb{T}\) are close.
Proposition 4.5. Suppose that Assumptions I,J,K hold. Then there exists \( c > 0 \) such that the probability
\[
P \left( \left\| \left( P_{\mathcal{E}} - \tilde{P}_{\mathcal{E}}^{\mathcal{T}} \right) \chi_{\lambda \mathcal{E}/3(0)} \right\| > e^{-c\mathcal{E}} \right) \leq e^{-c\mathcal{E}}. \tag{4.8} \]

For a proof, see [22, Lemma 4.11]. The argument is closely related to the one used in the proof of the following result that establishes the localization property of some bounded functions of \( H \) in the mobility gap.

Lemma 4.6. Suppose that Assumptions I,J,K hold. Then for any \( I := [E_1, E_2] \subset J_{\text{loc}} \) and any \( \Theta \subset \mathbb{T} \), there exists \( c > 0 \) such that
\[
\mathbb{E} \left| P_{\mathcal{I}}^{\Theta}(x, y) \right| \leq e^{-c|x-y|\Theta}, \quad \mathcal{I} = I, E, \tag{4.9} \]
for all \( x, y \in \Theta \). Moreover, for any \( z \in \mathbb{C} \) with \( \text{Re}(z) \in I/2 \), we have
\[
\mathbb{E} \left| \left( P_{\mathcal{I}}^{\Theta} (H^{\Theta} - z)^{-1} \right)(x, y) \right| \leq \frac{1}{E_2 - E_1} \frac{e^{-c|x-y|\Theta}}{\text{Im} z}. \tag{4.10} \]

Proof. Let \( \mathcal{I} = I \). Since \( \Theta \) is finite, the spectrum of \( H^{\Theta} \) is a discrete set. By (1.10),
\[
\{E_1, E_2\} \nsubseteq \sigma (H^{\Theta}) \]
almost surely. Thus the spectral projection \( P_{\mathcal{I}}^{\Theta} \) is equal to
\[
P_{\mathcal{I}}^{\Theta} = - (2\pi)^{-1} \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j (H^{\Theta} - iu - E_j)^{-1} du \tag{4.11} \]
almost surely, see (C.7). Using \( |(H^{\Theta} - iu - E_j)^{-1}(x, y)| \leq |u|^{-1} \), we get a bound
\[
|P_{\mathcal{I}}^{\Theta}(x, y)| \leq \max_j \pi^{-1} \int_{-\infty}^{\infty} \left| (H^{\Theta} - iu - E_j)^{-1} (x, y) \right|^q |u|^{-1} du. \]
For \( |u| \geq 1 \), we use a decomposition
\[
(H^{\Theta} - iu - E_j)^{-1} = - (iu + E_j)^{-1} + (iu + E_j)^{-1} H^{\Theta} (H^{\Theta} - iu - E_j)^{-1}, \]
the range-\( r \) property for \( H \), and \( |H(x, y)| \leq C \) to estimate
\[
\mathbb{E} \left| P_{\mathcal{I}}^{\Theta}(x, y) \right| \leq \pi^{-1} \sup_{u \in \mathbb{R}} \max_j \left( \mathbb{E} \left| (H^{\Theta} - iu - E_j)^{-1} (x, y) \right|^q \int_{[-1, 1]} |u|^{-1} du \right. \]
\[
+ C \max_{z \in \mathbb{Z}^d} \max_{|z-x| \leq r} \mathbb{E} \left| (H^{\Theta} - iu - E_j)^{-1} (z, y) \right|^q \int_{[-1, 1]^r} |u|^{-2} du \right) \]
\[
\leq C e^{-c|x-y|\Theta}. \]
Since \( |P_{\mathcal{I}}^{\Theta}(x, y)| \leq 1 \) for all \( x, y \in \Theta \), by modifying \( c \) if necessary we get (4.3) for \( \mathcal{I} = I \). The argument for \( \mathcal{I} = E \) is nearly identical.

To get the second assertion of the lemma, we use
\[
(H^{\Theta} - z)^{-1} = -(i \text{Im}(z) + 1)^{-1} + (i \text{Im}(z) + 1)^{-1} (H^{\Theta} - \text{Re}(z) - 1) (H^{\Theta} - z)^{-1} \]
and
\[
\tilde{P}_{\mathcal{I}}^{\Theta} (H^{\Theta} - z)^{-1} = -(2\pi)^{-1} \sum_{j=1}^{2} \int_{-\infty}^{\infty} (z - E_j - iu)^{-1} (H^{\Theta} - iu - E_j)^{-1} du. \]
They yield
\[
\tilde{P}_{\mathcal{I}}^{\Theta} (H^{\Theta} - z)^{-1} = -(i \text{Im}(z) + 1)^{-1} \tilde{P}_{\mathcal{I}} (H^{\Theta}) + \]
\[
(2\pi)^{-1} \sum_{j=1}^{2} (i \text{Im}(z) + 1)^{-1} (H^{\Theta} - \text{Re}(z) - 1) \int_{-\infty}^{\infty} (z - E_j - iu)^{-1} (H^{\Theta} - iu - E_j)^{-1} du. \]
Since $P_\Theta^\| = 1 - P_\Theta^0$, $|i\text{Im}(z) + 1| = (\text{Im} z)$, and $|z - E_j - iu|^{-1} \leq (E_2 - E_1)^{-1}$ for any $\text{Re}(z) \in I/2$ 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 4.6, which follows from the previous statement by Markov’s inequality.

**Lemma 4.7.** Suppose that Assumptions 4.3–4.4 hold. Let $J := [E_1, E_2] \subset J_{loc}$. Then, there exists $c > 0$ such that for any $\Theta \subset \mathbb{T}$, the probability that for all $x, y$ with $|x - y|_\Theta \geq \sqrt{T}$,

$$\left| \left( \frac{P_\Theta^2}{(H^\Theta - z)^{-1}} \right)(x, y) \right| \leq e^{-c|x - y|_\Theta} \quad (4.12)$$

is $\geq 1 - e^{-c\sqrt{T}}$.

### 4.2. Local Structure of $H^\mathbb{T}$

Here we will again suppose that Assumptions 4.2–4.4 hold.

Given scales $\ell < L$ with $L \mod (\frac{3}{2} \ell) = \ell$, and $\ell$ even, we cover the torus $\mathbb{T} = \mathbb{T}^d_L$ with the collection of boxes

$$\{\Lambda_\ell(a)\}_{a \in \Xi_\ell}, \quad (4.13)$$

where

$$\Xi_\ell := \left( \left( \frac{3}{2} \ell \mathbb{Z} \right)^d / L \mathbb{Z}^d \right). \quad (4.14)$$

Here the boxes $\Lambda_\ell(a)$ (defined earlier as a subset of $\mathbb{Z}^d$) are understood, with a slight abuse of notation, as subsets of $\mathbb{T}$, i.e., $\Lambda_\ell(a) = \{ x \in \mathbb{T} : d_\ell(x, a) \leq \ell \}$. We recall that we use a max distance throughout this paper. We will refer to this collection of boxes as a *suitable* $\ell$-cover of $\mathbb{T}$.

The (trivial) properties of suitable covers are encapsulated by the following lemma.

**Lemma 4.8.** Let $r < \ell < L$. Then, a suitable $\ell$-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 $\Lambda_{\ell/4}(y) \subset \Lambda_\ell(a)$. For such a value of $a$ we will denote $\Lambda_\ell(y) := \Lambda_\ell(a)$;

(iii) $\Lambda_{\ell/4}(a) \cap \Lambda_\ell(a') = \emptyset$ for all $a, a' \in \Xi_\ell, a \neq a'$;

(iv) $(\frac{\ell}{2})^d \leq |\Xi_\ell| \leq (\frac{2L}{\ell})^d$.

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

Let $\ell < L$ and let $S$ be a subset of a suitable $\ell$-cover such that the boxes $\{ \Lambda_\ell(a) \}_{a \in S}$ are separated by a distance $r$. Fix $E \in J_{loc}$, then, by Lemma 4.2 for all $\nu > 0$ we have

$$\mathbb{P}\left\{ \text{dist} \left( E, \sigma(H^{\Lambda_\ell(a)}) \right) \leq \nu \text{ for all } \Lambda_\ell(a) \in S \right\} \leq \left( C \nu^{\frac{d}{2}} \right)^{|S|}. \quad (4.15)$$

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

$$\mathcal{L} \gg \ell \gg 1, \quad |I| = c\ell^{-\frac{d}{4}}. \quad (4.16)$$

for an $\ell$–independent constant $c$. We recall that we are using a convention where $c$ denotes a sufficiently small constant and $C$ a sufficiently large constant. The values of these constants can change equation by equation.

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

**Lemma 4.9.** The cardinality of $\mathcal{R}_M$ is bounded by

$$|\Xi_\ell| \leq (2de)^M \leq (\frac{2L}{\ell})^d (2de)^M. \quad (4.17)$$
Proof of Lemma 4.9. We first note that each set $S$ in $S_M$ looks like a compressed $d$-dimensional polycube of size $M$, and that we can bound the number of distinct $S_M$ using the same method as for the regular polycubes, see e.g., [9]. To make the argument 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 a pair of polycubes is considered adjoint if they share a $((d-1)$-dimensional) face. Two fixed polycubes are equivalent if one can be transformed into the other by a translation.

Given $S$, we assign the numbers $1, \ldots, M$ to the cubes of $S$ in lexicographic order. We now search for the (cube) connectivity graph $G$ of $S$, beginning with cube 1. During the search, any cube $c \in S$ is reached through an edge $e$ and connected by the edges of $G$ to at most $2d-1$ other cubes. We label each outgoing edge $e'$ with a pair $(i,j)$, where $i$ is the number associated with $c$, and $1 \leq j \leq 2d-1$ is determined by the orientation of $e'$ with respect to $e$. By the end of the search, each of the $M-1$ edges in 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 the number of distinct shapes for $S$ is bounded by

$$\left(\frac{(2d-1)M}{M-1}\right) \leq (2dM)^M. \tag{4.18}$$

The total number of sets $S$ can be now bounded by noticing that they are contained in the set of all translates of the distinct shapes of $S$ by elements of $\Xi$, yielding (4.17).

For any given configuration $\omega$, let $\tilde{T}$ denote the union of the boxes $\Lambda_{\ell}(a)$ with $a \in \Xi$, such that the restricted Hamiltonian $H^{\omega}_{\Lambda_{\ell}(a)}$ has at least one eigenvalue in the interval $2J$. Let $T$ denote the union of boxes $\Lambda_{\ell}(b)$ with $b \in \Xi$ that has a non-trivial overlap with $\tilde{T}$. We will enumerate by $\{T_{\gamma}\}$ a set of connected (with respect to the graph structure of $\tilde{T}$) components in $T$, i.e.,

$$T = \bigcup_{\gamma} T_{\gamma}, \quad T_{\gamma} \cap T_{\nu} = \emptyset, \quad T_{\gamma} \in S_M \text{ for } M \in \mathbb{N}.$$ 

For a given $T$, we will denote by $M(T)$ the maximum

$$M(T) = \max_{\gamma} \{M : T_{\gamma} \in S_M\}.$$ 

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

$$M(T) < N.$$

Lemma 4.10. Let $\ell > r$ and $I \subset J_{\text{loc}}$ with $|I|^q < \ell^{-d}$. Then for $c$ small enough we have

$$\mathbb{P}(\Omega_N^c) \leq \left(\frac{2e^c}{\ell}\right)^d e^{-N}. \tag{4.19}$$

Proof. For any $\omega \in \Omega_N^c$, there exists at least one cluster $T_{\gamma} \in S_M$ with $M \geq N$. Let $\tilde{T}_{\gamma}$ denote the union of boxes that generate $T_{\gamma}$, i.e., $T_{\gamma}$ is formed by all boxes that overlap with at least one box in $\tilde{T}_{\gamma}$. We note that $\tilde{T}_{\gamma}$ is in general not uniquely defined, but this will not play a role in our argument. We also remark that any box $\Lambda_{\ell}(a) \subset \tilde{T}_{\gamma}$ overlaps with $3^d$ boxes, so $|\tilde{T}_{\gamma}| \leq 3^d |T_{\gamma}|$. Let $U$ be a collection of vectors in $\mathbb{R}^d$ whose components take binary values. Then $\Xi_{\ell} = \bigcup_{e \in U} \Xi_{\ell,e}$, where $\Xi_{\ell,e} = \frac{2}{3}e + (3\ell \mathbb{Z})^d / \mathbb{Z}^d$, and $\Xi_{\ell,e} \cap \Xi_{\ell,e'} = \emptyset$ for $e \neq e'$ and

$$\Lambda_{\ell}(a) \cap \Lambda_{\ell}(a') = \emptyset \text{ for all } a \in \Xi_{\ell,e}, \quad a \in \Xi_{\ell,e'}, \tag{4.20}$$

using the fact that $\ell$ is even. Hence, for any $S \subset \Xi_{\ell}$, there exists $e \in U$ such that $|S \cap \Xi_{\ell,e}| \geq 2^{-d} |S|$. In particular, the number of non-overlapping boxes in $\tilde{T}_{\gamma}$ is at least $6^{-d}M$ due to (4.20).

We are now in a position to apply (4.15) to conclude that the probability that a fixed configuration $T$ has at least one cluster $T_{\gamma} \in S_M$ with $M \geq N$ is bounded by $(C |I|^q \ell^d)^{6^{-d}M}$. It
follows now from Lemma 4.9 that
\[
\mathbb{P}(\Omega_N^c) \leq \sum_{M=N}^{\infty} \left( \frac{2c}{\ell^2} \right)^d \left( (2de)^{(6^d)} |I|^q \ell^d \right)^{6-dM}.
\]
(4.21)

This is less than or equal to \( \left( \frac{2c}{\ell^2} \right)^d e^{-N} \) provided that \( c \) in (4.16) is small enough.

□

For an integer \( N \), we now consider a subset \( \Omega_{\text{loc},N} \) of the full configuration space for which \( \mathbb{T} \) and all of the sets in \( \{ S_M \}_{M=1}^{N} \) are \( \ell/10 \)-localizing and satisfy (4.12).

**Lemma 4.11.** There exists constants \( C, c > 0 \) such that
\[
\mathbb{P}(\Omega_{\text{loc},N}^c) \leq CN^2 (2L\ell)^d (2de)^N e^{-c\sqrt{\ell}}.
\]
(4.22)

**Proof.** The total number of \( \{ S_M \}_{M=1}^{N} \) is bounded by
\[
\sum_{M=1}^{N} \left( \frac{2c}{\ell^2} \right)^d (2de)^M < 2 \left( \frac{2c}{\ell^2} \right)^d (2de)^N
\]
thanks to Lemma 4.9. Their maximal volume is bounded by \( N\ell^d \). Thus, we can bound
\[
\mathbb{P}(\Omega_{\text{loc},N}^c) \leq C \left( \frac{2c}{\ell^2} \right)^d (2de)^N \left( N\ell^d \right)^2 e^{-c\sqrt{\ell}} = CN^2 (2L\ell)^d (2de)^N e^{-c\sqrt{\ell}}
\]
(4.23)
using Theorem 4.4 and Lemma 4.7. □

We now optimize \( N \) from the previous two lemmas. To this end, we pick \( N = \lfloor c\sqrt{\ell} \rfloor \). Then, using Lemmata 4.10 and 4.11 for \( \ell \) large enough and intervals \( I \subset J_{\text{loc}} \) satisfying \( |I| < cf^{-d/q} \), we have
\[
\mathbb{P}(\Omega_N^c \cap \Omega_{\text{loc},N}^c) \leq L^d e^{-c\sqrt{\ell}}.
\]
(4.24)

For \( \omega \in \Omega_N \cap \Omega_{\text{loc},N} \), the number of eigenvalues of \( H^T_{\gamma} \) cannot exceed \( |T_{\gamma}| \leq N\ell^d \leq C\ell^{d+1/2} \).

Hence, for each \( \gamma \), we can find \( J_{\gamma} := [E^-_{\gamma}, E^+_{\gamma}] \) such that
\[
I/2 \subset J_{\gamma} \subset I \quad \text{and} \quad \text{dist}(E^+_\gamma, \sigma(H^T_{\gamma})) \geq c\ell^{-d-1/2} |I|.
\]

We note that
\[
\max_{\gamma} \text{diam}(T_{\gamma}) \leq L := C\ell^{3/2}.
\]
(4.25)

Let \( \Omega_G \) be a subset of the configuration set \( \Omega_N \cap \Omega_{\text{loc},N} \) such that, for \( c \) small enough, \( \omega \in \Omega_G \), \( z \in \mathbb{C} \) with \( \text{Re}(z) = E^+_\gamma \), and all \( x, y \in \partial T_{\gamma} \), the following bound holds:
\[
\sup_{T_{\gamma}} \left| \left( (H^T_{\gamma} - z)^{-1} \right)(x,y) e^{c\ell^{-d-1/2} |x-y|_{\gamma}} \right| \leq C\ell^{d+1/2} |I|^{-1} |\text{Im} z|^{-1}.
\]
(4.26)

Applying Lemma 4.7 with \( J = E^+_\gamma + [-c\ell^{-d-1/2} |I|, c\ell^{-d-1/2} |I|] \) and \( z \in \mathbb{C} \) with \( \text{Re}(z) = E^+_\gamma \) yields
\[
\mathbb{P}(\Omega_G^c) \leq L^d e^{-c\sqrt{\ell}}.
\]

**Proposition 4.12.** Let \( \omega \in \Omega_G \), and let \( I \subset J_{\text{loc}} \) be such that \( |I| < cf^{-d/q} \). Suppose that \( \ell \) is large enough, then
(i) (Local Gap) There exist intervals \( J_{\gamma} = [E^-_{\gamma}, E^+_{\gamma}] \) such that
\[
I/2 \subset J_{\gamma} \subset I \quad \text{and} \quad \text{dist}(E^+_\gamma, \sigma(H^T_{\gamma})) \geq c\ell^{-d-1/2} |I|;
\]
(4.27)

(ii) (Support of spectral projections)
\[
\| P_{J_{\gamma}} (H^T_{\gamma}) \delta_x \| \leq e^{-c\sqrt{\ell}} \text{ for any } x \in \mathbb{T} \setminus T_{\ell}
\]
(4.28)
(recall (4.19)), and
\[
\| P_{J_{\gamma}} (H^T_{\gamma}) \delta_x \| \leq e^{-c\sqrt{\ell}} \text{ for any } x \notin \partial_{\ell/8} \mathbb{T} \cup T_{\ell};
\]
(4.29)
(iii) (Exponential Decay of Correlations) Let \(A_o\) be any subset of \(T_\gamma\), then (with \(A_o\) in (4.20)–(4.21)) we have
\[
\left\| (H^{T_\gamma} - z)^{-1} \right\|_{c,\ell} \leq \ell^{d+1/2}|I|^{-1}(Im\ z)^{-1}
\]
for \(z \in \mathbb{C}\) with \(\text{Re}(z) = E_\gamma^+\).

Proof. Proposition 4.12.(i) has been established earlier, and Proposition 4.12.(iii) is a consequence of (4.20). This leaves us with the task of proving Proposition 4.12.(ii).

Let \(\{n, \psi_n\}\) be an eigenpair for \(H^{T_\gamma}\) in \(I\), and let \(x_n\) be its localization center. We first check that \(x_n \in \bar{T}\). Indeed, suppose that \(x_n \notin T\). Then, by the properties of the suitable cover, there exists a box \(\Lambda_{\ell}(a) \notin T\) such that \(\Lambda_{\ell/4}(x_n) \subset \Lambda_{\ell}(a)\). Moreover, \(\omega \in \Omega_G \subset \Omega_{\text{loc},N}\) implies that \(T\) is \(\ell/10\)-localizing, so in particular
\[
|\psi_n(y)| \leq Ce^{-\mu|y-x_n|\Lambda_{\ell}(a)} \text{ for } |y-x_n|\Lambda_{\ell}(a) \geq \sqrt{\ell/10}.
\]
We can now use Lemma [C.4] below to conclude
\[
\sigma \left( H^{\Lambda_{\ell}(a)} \right) \cap 2I \neq \emptyset,
\]
which means that \(\Lambda_{\ell}(a) \subset \bar{T}\), a contradiction. This establishes (4.28), since for any \(x \in \bar{T} \setminus T_\ell\) we have \(\text{dist} (x, \bar{T}) \geq \ell/8\).

Let \(\{\mu_n, \phi_n\}\) be an eigenpair for \(H^{T_\gamma}\) in \(I\). By the argument identical to the one used earlier, its localization center \(y_n\) is located either in \(\bar{T}\) or in \(\partial C\sqrt{2}T \subset \partial \ell/8T\). Hence
\[
\|P_{\mu_n}(H^{T_\gamma}) - \chi_{\partial C\sqrt{2}T} P_{\mu_n}(H^{T_\gamma})\| \leq e^{-c\sqrt{T}},
\]
which in particular establishes (4.29). In fact, the above argument shows more, namely that, recalling the notation in Theorem 2.1.(iii)
\[
\|P_{\mu_n}(H^{T_\gamma})\delta_x\| \leq e^{-c\sqrt{T}} \text{ for any } x \notin A_o.
\]
(4.33)
The latter bound will be of use to us momentarily. \(\Box\)

This completes the proof that \(H^{T_\gamma}\) possesses a local structure in the sense defined by Theorem 2.1. Using perturbation theory, we are now going to show that \(H^{T_\gamma}(s)\) possesses a local structure as well.

4.3. Proof of Theorem 2.1. We note that Proposition [1.12] is applicable here with \(I = c\ell^{-\xi}\). In particular, for \(\omega \in \Omega_G\), we have dist \((E_\gamma^+\), \(\sigma(H^{T_\gamma})\)) \(\geq \Delta\). Let
\[
\tilde{H}^{T_\gamma}(s) := H^{T_\gamma}(0) + P_{[E_\gamma^-, E_\gamma^+]}(H^{T_\gamma}(0)) + \beta W(s).
\]
Then, for \(\ell\) sufficiently small
\[
\sigma \left( \tilde{H}^{T_\gamma}(s) \right) \cap \left( \left[ -\frac{\Delta}{3}, \frac{\Delta}{3} \right] + [E_\gamma^-, E_\gamma^+] \right) = \emptyset,
\]
provided that \(\beta < \frac{\Delta}{c}\)
(4.34)

For the next assertion, we recall the definition of a dilation and its norm, introduced in (1.20)–(1.21).

Lemma 4.13. There exists \(c > 0\) such that for any \(z \in \mathbb{C}\) with \(\text{Re}(z) = E_\gamma^+\) and for any \(\beta < c\Delta\ell^{-3d}\), we have
\[
\left\| (H^{T_\gamma}(s) - z)^{-1} \right\|_{c,\ell} + \left\| \tilde{H}^{T_\gamma}(s) - z \right\|_{c,\ell} \leq C\ell^{3d}\Delta^{-1}(Im\ z)^{-1},
\]
(4.35)
where \(\| \cdot \|_{c,\ell}\) is defined with \(A = A_o\).
Proof. If we denote
\[ R_z^o = \left( H^{T_s}(0) - z \right)^{-1}, \tilde{R}_z^o = \left( H^{T_s}(0) - z \right)^{-1}, R_z = \left( H^{T_s}(s) - z \right)^{-1}, \tilde{R}_z = \left( H^{T_s}(s) - z \right)^{-1}, \]
we have
\[ \| R_z^o \|_{c,\ell} \leq C \ell^d \Delta^{-1} \langle Im z \rangle^{-1} \]
by (4.20).
We now expand \( R_z \) into the Neumann series
\[ R_z = R_z^o \sum_{n=0}^{\infty} \beta^n \left( -W R_z^o \right)^n, \]
yielding, via (1.22),
\[ \| R_z \|_{c,\ell} \leq \| R_z^o \|_{c,\ell} \sum_{n=0}^{\infty} \beta^n \| W R_z^o \|_{c,\ell} \]
\[ \leq C \ell^d \Delta^{-1} \langle Im z \rangle^{-1} \sum_{n=0}^{\infty} \left( \beta \ell^d \right)^n \Delta^{-n} \leq C \ell^d \Delta^{-1} \langle Im z \rangle^{-1}, \]
provided \( \beta \leq c \Delta^{-3d} \).
Using (4.33), we deduce that
\[ \left\| e^{e^{p_a} \langle P_{E^-, E^+} \rangle \left( H_o^{T_s} \right)} \right\| \leq C \ell^d. \]
Since
\[ \tilde{R}_z^o = R_z^o - P_{\left[ E^{-}, E^{+} \right]} \left( H_o^{T_s} \right) R_z^o \tilde{R}_z^o, \]
we obtain, using (1.37), (1.39) and
\[ \| R_z \| \leq C \Delta^{-1} \langle Im z \rangle^{-1}, \quad \left\| \tilde{R}_z^o P_{\left[ E^{-}, E^{+} \right]} \left( H_o^{T_s} \right) \right\| \leq 2, \]
that
\[ \left\| \tilde{R}_z \right\|_{c,\ell} \leq C \ell^d \Delta^{-1} \langle Im z \rangle^{-1}. \]
We now expand \( \tilde{R}_z \) into the Neumann series
\[ \tilde{R}_z = \tilde{R}_z^o \sum_{n=0}^{\infty} \beta^n \left( -W \tilde{R}_z^o \right)^n, \]
and repeat the argument in (4.33) to complete the proof.

We are now ready to finish the proof. For this, we will show that conditions (2.1.(i), 2.1.(iii)) in Theorem 2.1 hold on \( \Omega_G \), ensuring the desired probability for these events. This leaves us to prove Theorem 2.1.(ii).

We first note that Theorem 2.1.(i) follows from Proposition 4.12.(i) (with \( I = e\ell^{-5} \)) by standard perturbation theory for allowable values of \( \beta \). On the other hand, Theorem 2.1.(iii) is a direct consequence of Lemma 4.13. We recall that \( J_\gamma = \left[ E^{-}, E^{+} \right] \) and set \( \tilde{J}_\gamma = \left[ \frac{1}{8}, \frac{3}{8} \right] + \left[ E^{-}, E^{+} \right] \). We will abbreviate \( P_\gamma := P_{\tilde{J}_\gamma} \left( H^{T_s}(s) \right) \) and suppress the \( s \)-dependence for this argument, indicating by the subscript (or superscript) \( o \) the value \( s = 0 \), if needed. We use the decomposition (4.41) with \( E_1 = E^{-} \) and \( E_2 = E^{+} \) to write
\[ P_\gamma = -\left( 2\pi \right)^{-1} \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j R_{iu+E_j} du. \]
We note that the integrand can be bounded, using Theorem 2.1.(i), by
\[ \max_{j=1,2} \| R_{iu+E_j} \| \leq \Delta^{-1} \langle u \rangle^{-1}, \quad u \in \mathbb{R}. \]
Using (recall (4.39))
\[ R_{iu+E_j} = \tilde{R}_{iu+E_j} - \tilde{R}_{iu+E_j} P_{\gamma}(H_o^T) R_{iu+E_j} \]
and
\[ \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j \tilde{R}_{iu+E_j} du = 0, \]
which holds thanks to (4.34), we conclude that \( P_{\gamma} \) is equal to
\[ (2\pi)^{-1} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} \tilde{R}_{iu+E_j} P_{\gamma}(H_o^T) R_{iu+E_j} \hat{P}_{\gamma} du. \]  
(4.43)

Hence we can bound
\[ \left\| \sqrt{\rho} \gamma P_{\gamma} \right\| \leq \int_{-\infty}^{\infty} \max_{j} \left( \left\| \tilde{R}_{iu+E_j} \right\| \right) \left\| \sqrt{\rho} \gamma P_{\gamma}(H_o^T) \right\| \left\| R_{iu+E_j} \hat{P}_{\gamma} \right\| \]
\[ \leq C \ell^d \Delta^{-2} \int_{-\infty}^{\infty} (u)^{-2} du \leq C \ell^d \Delta^{-2}, \]  
(4.44)

where we have used Lemma 4.13, 4.39, and 4.42 in the second step.

By perturbation expansion for the resolvent and (4.41), we have
\[ P_{\gamma} = P_{\gamma}(H_o^T) - (2\pi)^{-1} \int_{-\infty}^{\infty} \sum_{j=1}^{2} \sum_{n=1}^{\infty} \beta^n R_{iu+E_j} (-W R_{iu+E_j})^n. \]

We first observe that, due to (4.39), \( \left\| \chi_{\theta_0} T P_{\gamma}(H_o^T) \chi_{T^c} \right\| \leq e^{-c\sqrt{T}}. \)

Next, letting \( A_o = (T^c)_{8\ell} \), we can estimate, using Lemma 4.13 and (1.22), that
\[ \left\| \chi_{\theta_0} T P_{\gamma}o R_{iu+E_j} (-W R_{iu+E_j})^n \chi_{T^c} \right\| \leq C^n \left\| \chi_{\theta_0} T e^{-\sqrt{\rho} \gamma A_o} \right\| \left\| R_{iu+E_j} \right\|^n \leq C^n \ell^{3dn} \Delta^{-n} (Im z)^{-2} e^{-c\sqrt{T}}. \]

Hence
\[ \left\| \chi_{\theta_0} T \sum_{n=1}^{\infty} \beta^n R_{iu+E_j} (-W R_{iu+E_j})^n \chi_{T^c} \right\| \leq e^{-c\sqrt{T}} (Im z)^{-2} \sum_{n=1}^{\infty} \beta^n C^n \ell^{3dn} \Delta^{-n} \leq e^{-c\sqrt{T}} (Im z)^{-2}. \]

Integrating over the \( u \) variable, we see that \( \left\| \chi_{\theta_0} T P_{\gamma} \chi_{T^c} \right\| \leq e^{-c\sqrt{T}} \) holds. Combining this bound with (4.44), we get (2.4).

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

5. 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 2.2 We first set \( \ell = \lfloor \beta^{-p_1} \rfloor \). Using Theorem 2.1 Proposition 4.5 and Lemma 4.6 the probability of the event that conditions (2.2)–(2.5),
\[ \left\| (P_E - P_E^T) \right\| < e^{-c\ell}, \text{ and } |P_E(H^T)(x, y)| < e^{-c|x-y|} \]
hold on the scale \( \ell \) for all \( x, y \) is \( \geq 1 - e^{-c\sqrt{T}} \). Thus, by Borel-Cantelli’s lemma, for almost all random configurations \( \omega \in \Omega \), there exists \( \beta_0 \) such that the event holds for all \( \beta < \beta_0 \). Furthermore, \( P-a.s., E \) is not an eigenvalue of \( H \) by Lemma 4.1. From now on, we will fix the configuration \( \omega \) for which all of these conditions are satisfied and will assume that \( \beta \) is below the corresponding threshold value \( \beta_0 \).

The condition
\[ e^{-1/\beta_{\omega}} < \epsilon < \beta^{p_1 p_2} \]
implies that $e^{-c\sqrt{T}} < \epsilon < t^{-p_2}$. Hence for $\beta$ sufficiently small, the conclusions of Theorem 2.2.(i)
2.2.(iii) are satisfied.

We note that (repeatedly) using Proposition 4.3, \((P_E - P_{E_\beta})\) \(< e^{-cL}\), and \(|P_E(H^T)(x,y)| < e^{-c|x-y|}\), we have
\[
\left\| (U_\epsilon(s,0)P_EU_\epsilon(0,s) - U_\epsilon^T(s,0)P_{E_\beta}U_\epsilon^T(0,s)) \chi_{\Lambda_\epsilon}/\beta \right\| \leq Ce^{-cL},
\]
where we have also used (4.3) and (4.9) (which hold deterministically by virtue of our assumptions on $\omega$ and $\beta$).

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

**Lemma 5.1.** Let $E_{\pm} \in J_{loc}$ be such that $|E_+ - E_-| \leq \delta$. Let $P_{\pm}(s) := P_{E_{\pm}}(s)$. Then
\[
\left\| \tilde{P}_+(0)(U_\epsilon^T)^*_T(s)P_- (s) \chi_{\Lambda_\epsilon} \right\| \leq D_n \left( \epsilon^n + e^{-1/\beta p} \right),
\]
where $D_n$ is some $n$-dependent coefficient depending only on $C, C_k$ in Assumption 1.2.

**Proof of Lemma 5.1.** Using Theorem 2.2, we have
\[
\| \tilde{P}_+(0)(U_\epsilon^T)^*_T(s)P_- (s) - \tilde{P}_+(0)\tilde{Q}(0)(U_\epsilon^T)^*_T(s)\tilde{Q}(s)P_- (s) \| \leq Ce^{-1/\beta p},
\]
and by the same assertion
\[
\| \tilde{P}_+(0)\tilde{Q}(0)(U_\epsilon^T)^*_T(s)\tilde{Q}(s)P_- (s) \| \leq C_N \epsilon^N + e^{-1/\beta p}.
\]

We are now ready to complete the proof. We first observe that it suffices to establish that
\[
\lim_{\epsilon, \beta \to 0} \left\| (U_\epsilon(s,0)P_E(0)U_\epsilon(0,s) - P_E(s)) \chi_{\Lambda_\epsilon} \right\| = 0.
\]
Using 5.1, we deduce that we can further reduce it to showing
\[
\lim_{\epsilon, \beta \to 0} \left\| (U_\epsilon^T(s,0)P_E(0)U_\epsilon^T(0,s) - P_{E_\beta}(s)) \chi_{\Lambda_\epsilon} \right\| = 0.
\]

We next decompose
\[
U_\epsilon^T(s,0)P_E(0)U_\epsilon^T(0,s) - P_{E_\beta}(s) = U_\epsilon^T(s,0)P_E(0)U_\epsilon^T(0,s)\tilde{P}_{E_\epsilon+\delta}(s)
- U_\epsilon^T(s,0)P_E(0)U_\epsilon^T(0,s)\tilde{P}_{E_\epsilon-\delta}(s)
+ (U_\epsilon^T(s,0)P_E(0)U_\epsilon^T(0,s) - P_{E_\beta}(s))P_{E_\epsilon+\delta}(s),
\]
and bound the first term using Lemma 5.1. To bound the last term, we note that \(\{H^\delta(s)\}\) (extended to an operator on $L^2(\mathbb{R}^d)$) converges to $H(0)$ in the strong resolvent sense as $\epsilon \to 0$. Hence for any interval $(a, b)$, we have
\[
\text{s-lim}_{\epsilon, \beta \to 0}P_{(a,b)}(s) = P_{(a,b)}(0),
\]
16. Theorem VIII.24]. Hence $P_{E_\epsilon+\delta}(s)$ $\overset{SOF}{\to} P_{(E)}(0) = 0$, and (5.4) follows.

**Remark 5.2.** For $\epsilon \geq \beta$ Theorem 1.12 follows from the result in 23 (where it is proven for $\epsilon = \beta$, but the argument there remains valid for $\epsilon \geq \beta$ as well). Here we are focused on the regime where $\epsilon \ll \beta$. While we have not attempted to extend the theorem to the remaining interval $\beta \geq \epsilon \geq \beta^{p_1p_2}$, we do expect it to hold in this parameter range as well.

6. **Uniformly localized eigenfunctions for $H(s)$ and the proof of Theorem 1.8**

Disclaimer: In the process of completing this paper, we learned about a recent preprint 38, which has a significant thematic overlap with the results presented here.
6.1. Non-uniform bound on localization. Let $H_\omega$ be an infinite volume operator satisfying Assumptions [12,13]. We will need a stronger concept of a localizing Hamiltonian than the one introduced earlier in Definition [13].

**Definition 6.1.** For $\omega \in \Omega$ and a pair $(c, \theta)$ of positive valued parameters, we will say that $H_\omega$ is non-uniformly $(c, \theta)$-localizing if there exists an eigenbasis $\{\psi_i\}$ for $H_\omega$ such that

$$|\psi_i(y)|^2 \leq \frac{1}{\theta} (x_i)^{d-1} e^{-c|y-x_i|}$$

for some $x_i \in \mathbb{Z}^d$. (6.1)

Here, the quantifier "non-uniformly" refers to the presence of the factor $\frac{1}{\theta}$.

**Theorem 6.2** (Non-uniform eigenfunction localization). Let $H_\omega$ be an infinite volume operator satisfying Assumptions [12,13] with $m = 1$. Then

$$\mathbb{P}\left( \left\{ \omega \in \Omega : \text{H}_\omega \text{ is non-uniformly } (c, \theta)-\text{localizing} \right\} \right) \leq 1 - C \theta$$

for some $C > 0$.

**Proof.** The assertion above follows from [13, Theorem 7.4] by Markov’s inequality. □

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 the price of a small fraction of eigenstates for which the statement will fail to hold.

We first note that the integrated density of states (IDOS) $N_{loc}$ of $H_\omega$, associated with the interval $J_{loc}$ and defined as

$$N_{loc} = \lim_{R \to \infty} \frac{\text{tr} \chi_{\Lambda R(0)} P_{J_{loc}} (H_\omega)}{R^d}$$

is almost surely non-random, see e.g., [13, Theorem 3.15 and Corollary 3.16]. Moreover, if $N_{loc} > 0$, the convergence to the mean in (6.3) is exponentially fast, so in particular

$$\mathbb{P}\left( \frac{\text{tr} \chi_{\Lambda R(0)} P_{J_{loc}} (H_\omega)}{R^d} < \frac{N_{loc}}{2} \right) \leq e^{-mR}$$

for some $m > 0$. This is a typical large deviations result, see e.g., [14].

We now adjust the concept of localized eigenvectors to make it uniform. We will assume here that $N_{loc} > 0$.

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

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

(6.5)

We will say that the orthogonal projection $P \in \mathcal{L}(\ell^2(\mathbb{Z}^d))$ is $(c, \theta)$-Wannier decomposable if there exists an orthonormal basis $\{\psi_i\}$ for $\text{Ran}(P)$ such that each $\psi_i$ is $(c, \theta)$-localized.

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

Let $H^\upharpoonright_\omega$ denote the periodic restriction of $H_\omega$ to the torus $\mathbb{T}_L$ of a linear size $L$. The following assertion follows from the judicious use of Markov’s inequality and the deterministic Lemma [13,2] below.

**Theorem 6.4.** Suppose that Assumptions [12,13] hold and that in addition $N_{loc} > 0$. For a given configuration $\omega \in \Omega$, let $\mathbb{P}_E$ denote the normalized counting measure of eigenvalues of $H^\upharpoonright_\omega$ in the interval $J_{loc}$ (counting multiplicities). Let $\mathcal{G}$ be the set

$$\mathcal{G} := \left\{ E_n \in \sigma(H^\upharpoonright_\omega) \cap J_{loc} : P(E_n) \text{ is } \left(\frac{\omega}{m}, \theta^2 \right) \text{-Wannier decomposable} \right\}.$$ 

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

$$\mathbb{P}\left( \mathbb{P}_E (\mathcal{G}) \geq 1 - \sqrt{\theta} \right) \geq 1 - C \sqrt{\theta}.$$
Proof. For a pair $(E_n, P_{\{E_n\}})$, let
\[ w_n = w(\omega, P_{\{E_n\}}) = \sum_{x,y} |P_{\{E_n\}}(x,y)| e^{\|x-y\|}. \] (6.7)
We then have, by the bound (4.6) on the eigenvector correlator and $\mathcal{N}_{\text{loc}} > 0$,
\[ \mathbb{E}_\omega \mathcal{E}[w_n] \leq C. \]
Letting $a, b > 0$, we have by Markov’s inequality that
\[ P_\omega (\mathcal{E}[w_n] \leq \theta^{-a}) \geq 1 - C\theta^a. \]
We now pick an $\omega$ such that $\mathbb{E}_\omega \mathcal{E}[w_n] \leq \theta^{-a}$. Another application of Markov’s inequality then gives
\[ P_E(w_n \leq \theta^{-b}) \geq 1 - \theta^{b-a}. \] (6.8)
The assertion now follows from (6.8) with $a = \frac{1}{2}, b = 1$, and Lemma B.2  
We are now ready to complete

**Proof of Theorem 1.8**. Here we will use $\theta = e^{-c\sqrt{r}}$.

Let $\mathcal{L} = C\epsilon^{-1}$ and consider
\[ \Xi_\mathcal{L} := \left( \frac{3}{4} \mathcal{L} \mathcal{Z} \right)^d, \] (6.9)
and an $\mathcal{L}$-cover of $\mathcal{Z}^d$ of the form
\[ \mathcal{Z}^d = \bigcup_{a \in \Xi_\mathcal{L}} \Lambda_\mathcal{L}(a). \]

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

We also cover $J_{\text{loc}}$ with the overlapping intervals $\{ J_i \}$ so that
(i) The length of each interval $J_i$ is equal to $c \ell^{-\xi}$;
(ii) For each $E \in J_{\text{loc}}$ that satisfies $\text{dist}(E, (J_{\text{loc}})^c) \geq \ell^{-\xi}$ we can find $J_i$ such that $\text{dist}(E, (J_i)^c) \geq c \ell^{-\xi}/3$;
(iii) $\bigcup_i J_i \subset J_{\text{loc}}$.

One can always construct such a covering using $C\ell^4$ intervals $J_i$ for $\ell$ sufficiently large.

We will say that a property $\mathcal{A}$ is satisfied for at least a fraction $1 - \sqrt[3]{\theta}$ of boxes $\Lambda_\mathcal{L}(a)$ (which we will be calling good boxes) if
\[ \lim_{R \to \infty} \frac{\# \Lambda_\mathcal{L}(a) \subset \Lambda_R : \mathcal{A} \text{ is satisfied for } \Lambda_\mathcal{L}(a)}{\# \Lambda_\mathcal{L}(a) \subset \Lambda_R} \geq 1 - \sqrt[3]{\theta}. \] (6.10)

For a given box $\Lambda_\mathcal{L}(a)$ in the cover we construct the corresponding torus $\mathcal{T}_a$ and pick any $J_i$ from the cover of $J_{\text{loc}}$. It follows that the conclusions of Theorem 3.6 are satisfied with probability $\geq 1 - e^{-c\sqrt{r}}$. Moreover, as the number of $J_i$s in the cover is $C\ell^4$, we deduce that with the same probability the conclusions of Theorem 3.6 hold for all $J_i$s in the cover. We next note that, given $N$ tori $\{ \mathcal{T}_a \}$, we can choose at least $6^{-d} N$ of them to be separated by a distance greater than $r$, see the proof of Lemma 4.10. Hence, using Assumption 1.3 and ergodicity, we obtain that the fraction $1 - e^{-c\sqrt{r}}$ of tori $\{ \mathcal{T}_a \}_{a \in \Xi_\mathcal{L}}$ satisfy the conclusions of Theorem 3.6 for each interval $J_i$ in the cover of $J_{\text{loc}}$.

Let $\Omega_1 \subset \Omega$ be a collection of $\omega$ such that $P_E(G) \geq 1 - \sqrt[3]{\theta}$ for all $R \geq R_0$ (in particular, $P(G) \leq e^{-c\sqrt{r}}$ holds by (5.10)).

We now pick any $\omega \in \Omega_1$ and conclude from Theorem 6.4 that the fraction $1 - e^{-c\sqrt{r}}$ of eigenstates $\psi_n$ for $H^G$ with eigenvalues $E_n \in J_{\text{loc}}$ are $(c/m, \theta^2)$-localized. Let $\psi$ such one such eigenfunction, with energy $E$ and a localization center at $x$. Then there exists a box $a \in \Xi_\mathcal{L}$ and an interval $J_i$ such that
\[ \text{dist} (\Lambda_\mathcal{L}(a), x) \geq \mathcal{L}/4, \quad \| \chi_a \psi \| \leq e^{-c\mathcal{L}}, \quad E \in J_i. \]
If this box happens to be a good box, then the first assertion of Theorem 1.8 holds for all s by Theorem 2.1 while the second assertion holds for \( \psi \) at \( s = 0 \) by Lemma 3.4 below and by the assertions of Theorem 2.1. It then follows from Theorem 3.6 (see Remark 3.8 there) that the second assertion holds for all \( s \in [0, 1] \). Since the fraction of good boxes is \( 1 - e^{-c\sqrt{t}} \), we get the result.

\[ \square \]

### 7. Derivation of Linear Response Theory

In this section, we prove Theorem 1.11 assuming the setting outlined in Section 1.3. The proof consists of three steps:

1. We approximate the \( \mathbb{Z}^2 \) geometry by one of a (sufficiently large) torus. In particular, we replace \( P_{E_F} \) by \( P_{E_F}^T \).

2. We decompose \( P_{E_F}^T \) into the adiabatic projection \( Q \), corresponding to the energy \( E = E_F + 6\delta \) (recall (2.1)), and a remainder term \( R \). We then use the adiabatic theorem on the torus to control the evolution of the adiabatic portion.

3. We show that the remainder term does not contribute to the transport as it consists of localized states, and that the adiabatic term gives the Kubo formula.

Let \( \mathcal{L} = C\epsilon^{-1} \) and let \( \mathbb{T} \) be a torus of linear size \( \mathcal{L} \). For the first step, we let \( \mathcal{B} \subset \mathbb{T} \) be a region satisfying \( \Lambda_{\mathcal{L}/4} \subset \mathcal{B} \subset \Lambda_{\mathcal{L}/3} \). A precise choice for \( \mathcal{B} \) will be made afterwards.

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

\[
\text{tr}(P_\epsilon(s) - P) J = \text{tr}(P_\epsilon^T(s) - P^T) \tilde{J} + \mathcal{O}(e^{-c\mathcal{L}}),
\]

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

\[
\tilde{J} = \chi_\mathcal{B} J.
\]

We note that \( \tilde{J} \) is supported on a strip \( |x_1| \leq r \).

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

\[
\| (P_\epsilon^T(s) - P^T) \chi_{\{x\}} \| \leq e^{-c|x|} \text{ for } |x| \geq \mathcal{L}/3, \quad P^T = \{P, P^\ast\}.
\]

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

\[
\text{tr}(P_\epsilon(s) - P) J = \text{tr}(P_\epsilon(s) - P) \tilde{J} + \text{tr}(P_\epsilon(s) - P) \chi_\mathcal{B} J,
\]

and estimate each contribution separately. The second term is \( \mathcal{O}(e^{-c\mathcal{L}}) \) by (7.2). For the first term, we use (5.3) to deduce that

\[
\text{tr}(P_\epsilon(s) - P) \tilde{J} = \text{tr}(P_\epsilon^T(s) - P^T) \tilde{J} + \mathcal{O}(e^{-c\mathcal{L}}).
\]

It thus remains to show that (7.2) holds. For the proof in a time-independent situation, see [10, Lemma 5]. The proof below follows similar ideas.

Since the argument is identical for both projections, we will only consider the case \( P^T = P \).

Using the fundamental theorem of calculus, we write

\[
P_\epsilon(s) - P = -U_\epsilon(s) \left( \int_{-1}^{s} \partial_t (U_\epsilon^*)(t) U_\epsilon(t) \right) U_\epsilon^* (s)
\]

\[
= i\epsilon U_\epsilon(s) \left( \int_{-1}^{s} U_\epsilon^* (t) [H(t), P] U_\epsilon(t) \right) U_\epsilon^* (s)
\]

\[
= i\beta \epsilon U_\epsilon(s) \left( \int_{-1}^{s} g(t) U_\epsilon^* (t) [A_2, P] U_\epsilon(t) \right) U_\epsilon^* (s).
\]

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

\[
\| [A_2, P] \chi_{\{x\}} \| \leq C e^{c\mathcal{L}} |x|^{d+1} e^{-c|x_2|}.
\]

(7.4)
Indeed, (6.19) implies by Markov’s inequality that
\[
\sum_{x,y \in \mathbb{Z}^2} |x|^{-d-1} e^{c|x-y|} |P(x, y)| \leq C e^{cL}
\]
with the same probability. This implies that on the same probabilistic set,
\[
|P(x, y)| \leq C e^{cL} |x|^{d+1} e^{-c|x-y|}.
\]
The relation (7.4) now follows by using \( \|A_2 e^{c\omega_x}\| \leq 1 \) for all \( x \) with \( x_2 < 0 \), and then using \( \|A_2 e^{c\omega_x}\| \leq 1 \) for the remaining \( x \in \mathbb{Z}^2 \) together with \( [A_2, P] = -[A_2, P] \).

Combining (7.4) with Proposition C.5, we deduce that
\[
\text{Combining (7.4) with Proposition C.5, we deduce that}
\]
\[
\|A_2, P\| \leq \|e^{c\omega_x}\| |x|^{\ell} e^{c|x|2} \text{ for } |x| \geq L/3.
\]
The desired bound (7.2) now follows by combining (7.3) with \( \|\omega_x e^{c|x|} J\| \leq C \) for all \( x \in \mathbb{Z}^2 \).

This establishes the first step of the proof.

For the second step, we will consider configurations \( \omega \) for which Theorem 2.1 (and consequently Theorem 2.2) is applicable. In particular, all bounds below hold with probability \( \geq 1 - e^{-c\sqrt{T}} \). For a fixed \( \omega \), we consider a set \( \mathcal{A} = \bigcup \mathcal{T}_\gamma \), where the union is taken over all \( \gamma \) such that \( \mathcal{T}_\gamma \cap \Lambda_{\ell/4} \neq \emptyset \). Let \( \mathcal{B} = \Lambda_{\ell/4} \cup \mathcal{A} \). We note that by construction \( \Lambda_{\ell/4} \subset \mathcal{B} \subset \Lambda_{\ell/4+L} \) and
\[
\min \text{dist} \left( \partial \mathcal{B}, \mathcal{T}_\gamma \right) \geq \ell/4
\]
(see the paragraph preceding (8.19) for notation), two facts that will be used often in the proof below.

We next decompose \( P^T \) into two components \( P^T = \mathcal{Q}(-1) + R \) where \( \mathcal{Q}(s) \) is the smooth adiabatic projection constructed in Theorem 2.2 (adjusted to the interval \((-1, 1)\)) and \( R := P^T - \mathcal{Q}(-1) \). By Theorem 2.2 we then have that for \( s \geq 0 \) and \( N \in \mathbb{N} \),
\[
\|P^T_s(s) - \mathcal{Q}(0) - R_s(s)\| \leq C_N e^{\gamma \sqrt{T}} \left( \frac{1}{\delta^N} + \frac{1}{\delta^{2N+1}} \right) + \mathcal{O}(e^{-c\sqrt{T}}),
\]
with \( R_s = U_s()RU^*_s(s) \), where we have used \( \mathcal{Q}(s) = \mathcal{Q}(0) \) for \( s \geq 0 \). Hence
\[
\sigma_m = \frac{1}{\beta} \text{tr}((\mathcal{Q}(0) - \mathcal{Q}(-1) J) + \frac{1}{\beta} \int_0^1 \text{tr}(R_s(s) - R) J) ds + \mathcal{O}(e^{\gamma \sqrt{T}}).
\]
In Proposition 7.2 below we will show that the first term on the right hand side is equal to \( \sigma \), up to corrections of order \( \mathcal{O}(e^{-c\sqrt{T}}) \).

Thus it remains to bound the second term. It will be convenient to introduce a new scale \( \bar{\ell} \) in addition to \( \ell \), defined by the modified value for \( \delta \), namely \( \bar{\delta} = 7\delta \). We consider the operator \( \mathcal{Q}_s \) constructed in Lemma 3.10. The important properties of \( \mathcal{Q}_s \) are that it covers the spectral support of \( R \) and that it allows us to control the spatial support of \( R \). Let \( I = (E - 6\delta, E + 6\delta) \). Using Theorem 2.2.(ii) we have
\[
\|R - P^T_I R P^T_I\| \leq \mathcal{O}(e^{-c\sqrt{T}}).
\]
By the definition of \( \mathcal{Q}_s \) and the exponential decay of \( R \), we then obtain
\[
\left\| R - \sum_\gamma \tilde{Q}_s \right\| \leq \mathcal{O}(e^{-c\sqrt{T}})
\]
and, using Lemma 3.7.(i) we see that for \( s \geq 0 \),
\[
\|R_s(s) - \sum_\gamma Q_s \| \leq \mathcal{O}(e^c + e^{-c\sqrt{T}}).
\]
Since $Q_r^2$ is supported in $\hat{T}_\gamma$ (see the paragraph preceding (3.19) for notation), it follows that, up to a small error, $R_\epsilon(s)$ is the sum of terms supported in the region $\hat{T}_\gamma$. Let $\hat{U}_\epsilon$ denote the evolution generated by $H_T(s)$, the restriction of $H^{\tilde{T}}(s)$ to the union of all $T_\gamma$. Then we have
\[
\frac{d}{ds} \left( \hat{U}_\epsilon^* R_\epsilon(s) \hat{U}_\epsilon(s) \right) = \frac{i}{\epsilon} \hat{U}_\epsilon^* (s) [H_T(s) - H(s), R_\epsilon(s)] \hat{U}_\epsilon(s) = O(\epsilon^\infty + e^{-cv^{\tilde{T}}}),
\]
thanks to (7.7) and Lemma 3.4. Thus we can approximate
\[
\|R_\epsilon(s) - \sum_\gamma \hat{Q}_r^2 \hat{R}_\epsilon(s) \hat{Q}_r^2 \| \leq O(\epsilon^\infty + e^{-cv^{\tilde{T}}}),
\]
where $\hat{R}_\epsilon(s) = \hat{U}_\epsilon^* (s) R_\epsilon(s)$.

Letting $X$ be the set
\[
X = \left\{ \hat{T}_\gamma : \{ \hat{T}_\gamma \cap \{|x| \leq r\} \neq \emptyset, j = 1, 2 \} \right\},
\]
we have $|X| \leq CL^2$.

Considering now any $\hat{T}_\gamma \notin X$, either $\text{dist} \left( \hat{T}_\gamma, \{ x \in \mathbb{Z}^2 : x_1 = 0 \} \right) \geq r$, in which case
\[
Q_r^2 \hat{J} = 0,
\]
or $\text{dist} \left( \hat{T}_\gamma, \{ x \in \mathbb{Z}^2 : x_2 = 0 \} \right) \geq r$, in which case
\[
Q_r^2 \hat{R}_\epsilon(s) Q_r^2 = Q_{r-1} R Q_{r-1} + O(e^{-cv\tilde{T}}),
\]
as the perturbation is constant in that region. Hence, using (7.7) and Lemma 3.4 again (recall that $A^\Theta$ stands for the restriction of the operator $A$ to the set $\Theta$),
\[
\text{tr}(R_\epsilon(s) - R) \hat{J} = \text{tr} \left( \left( \hat{R}_\epsilon(s) \right)^X - R^X \right) \hat{J} + O(\epsilon^\infty + e^{-cv\tilde{T}})
\]
\[
= \text{tr} \left( \left( \hat{R}_\epsilon(s) \right)^X - R^X \right) J + O(\epsilon^\infty + e^{-\#\tilde{T}}).
\]

Next we observe, using the cyclicity of the trace for a trace class operator and (7.7), Lemma 2.2.(i) and Lemma 3.4 one more time, that
\[
\text{tr} \left( \left( \hat{R}_\epsilon(s) \right)^X - R^X \right) J = -i \text{tr} \left( \left[ H_T(s), \hat{R}_\epsilon(s) \right] \right)^X J + O(e^{-\#\tilde{T}}).
\]

However,
\[
-\text{tr} \left( \left[ H_T(s), \hat{R}_\epsilon(s) \right] \right)^X J = \epsilon \partial_t \text{tr} \left( \hat{R}_\epsilon(s) \right)^X J.
\]

Hence by the fundamental theorem of calculus,
\[
\frac{1}{\beta} \int_0^1 \text{tr} \left( \left( \hat{R}_\epsilon(s) \right)^X - R^X \right) J ds = \frac{\epsilon}{\beta} \text{tr} \left( \left( \hat{R}_\epsilon(1) \right)^X - \left( \hat{R}_\epsilon(0) \right)^X \right) 1_1 + O(e^{-cv\tilde{T}}),
\]
so we finally get
\[
\left\| \frac{1}{\beta} \text{tr} \left( \left( \hat{R}_\epsilon(s) \right)^X - R^X \right) J ds \right\| \leq CL^2 \frac{\epsilon}{\beta} + O(e^{-\ell}).
\]
Picking
\[
\ell = \beta^{-p}, \quad \text{for} \quad p > 2d + \frac{1}{2} + \frac{d}{q}
\]
in order to satisfy the assumptions of Theorem 2.2 (see the proof of Theorem 1.12) we get the statement with $p' = (6d + 5/2 + 3d/q)^{-1}$.

The next statement establishes the last step of the proof, showing that the conductance is constant within the mobility gap, similarly to the corresponding argument in [3].
Proposition 7.2. We have
\[
\frac{1}{\beta} \text{tr}(Q(0) - Q(-1)) \hat{J} = \sigma + O(e^{-c\sqrt{\beta}}),
\]
where \(\sigma\) was defined in (1.10).

Proof. We note that by locality of \(H\), \(\hat{J} = i\chi_B[H^T(r), \Lambda_1]\). By the fundamental theorem of calculus,
\[
\frac{1}{\beta} \text{tr}(Q(0) - Q(-1)) \hat{J} = \frac{1}{\beta} \int_{-1}^{0} \text{tr}\left( \partial_r Q(r) i\chi_B[H^T(r), \Lambda_1] \right) dr.
\]
We claim that
\[
\frac{1}{\beta} \text{tr}\left( \partial_r Q(r) i\chi_B[H^T(r), \Lambda_1] \right) = \frac{1}{\beta} \int_{-1}^{0} \text{tr}\left( \partial_r Q(r) i\chi_B[H^T(r), \Lambda_1] \right) dr.
\]
Indeed, let \(\hat{\Lambda}_1(r) = Q(r)\Lambda_1 Q(r) + \hat{Q}(r)\Lambda_1 Q(r)\). We have
\[
\text{tr}\left( \partial_r Q(r) i\chi_B[H^T(r), \Lambda_1] \right) = \text{tr}\left( \partial_r Q(r) i\chi_B[H^T(r), \hat{\Lambda}_1(r)] \right) + O(e^{-c\sqrt{\beta}})
\]
where in the first step we have used \(Q(r)\partial_r Q(r) Q(r) = \hat{Q}(r)\partial_r Q(r)\hat{Q}(r) = 0\) and in the third step we employed \([H^T, Q(r)] = O(e^{-c\sqrt{\beta}})\). We have also repeatedly used the fact that commuting \(\chi_B\) with other operators under the trace contributes \(O(e^{-c\sqrt{\beta}})\) by virtue of (7.6) and the location of support of the involved operators. The relation (7.9) now follows, since
\(\hat{\Lambda}_1 = [Q(r), [Q(r), \Lambda_1]]\).

The implication is that
\[
\frac{1}{\beta} \text{tr}(Q(0) - Q(-1)) \hat{J} = i \int_{-1}^{0} \dot{g}(r) \text{tr}(Q(r) [[Q(r), \Lambda_1], [Q(r), \Lambda_2]]\chi_B) + O(e^{-c\sqrt{\beta}}).
\]
We now define
\[
\text{Ind}_L(Q) = \text{tr}(Q[[Q, \Lambda_1], [Q, \Lambda_2]])\chi_B.
\]
For \(Z^2\) geometry without the cutoff function \(\chi_B\), 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 \(R\), \(\text{Ind}_\infty(Q + R) = \text{Ind}_\infty(Q) + \text{Ind}_\infty(R)\), provided \(Q + R\) is a projection. [2] Proposition 2.5. The argument in [2] assumes that the underlying projections are covariant and that their kernels satisfy good decay properties. The latter hold in a random setting and one can relax the covariance requirement for such models as well, see [19]. Moreover, \(\lim_{L \to \infty} \text{Ind}_L(P)\) exists and we have
\[
\lim_{L \to \infty} \text{Ind}_L(P) = \sigma.
\]
[7, Section 6]. In fact, using [19] it is not difficult to show that
\[
|\sigma - \text{Ind}_L(P)| \leq O(e^{-cL})\quad \text{and} \quad |\text{Ind}_L(P) - \text{Ind}_L(P^T)| \leq e^{-c\sqrt{\beta}}.
\]
We next observe that, although \(P^T\) and \(Q(-1)\) do not commute, we have \(\|P^T, Q(-1)\| \leq e^{-c\sqrt{\beta}}\). Hence there exists a pair of operators \(\tilde{P}^T, \tilde{Q}(-1)\) such that \(\tilde{P}^T, \tilde{Q}(-1) = 0\) and \(\|P^T - \tilde{P}^T\| \leq e^{-c\sqrt{\beta}}, \|Q(-1) - \tilde{Q}(-1)\| \leq e^{-c\sqrt{\beta}}\). Moreover, applying the compression procedure used to get a projection \(Q_N\) from a near-projection \(Q_N(s)\) in the proof of Lemma 3.10 without loss of generality we can assume that \(\tilde{P}^T, \tilde{Q}(-1)\) are in fact projections. Let
\( \hat{R} = \hat{P}^T - \hat{Q}(-1) \). Since \( \|Q(-1)R\| \leq e^{-c\sqrt{T}} \), we conclude that \( \hat{Q}(-1)\hat{R} = 0 \). In particular, the additivity of index is applicable for \( \hat{Q}(-1) \) and \( \hat{R} \), and yields

\[
\left| \text{Ind}_\mathcal{L} \left( \hat{Q}(-1) \right) + \text{Ind}_\mathcal{L} \left( \hat{R} \right) - \text{Ind}_\mathcal{L} \left( \hat{P}^T \right) \right| \leq e^{-c\sqrt{T}}. \tag{7.14}
\]

By construction, we deduce that

\[
\left| \text{Ind}_\mathcal{L} \left( Y_i \right) - \text{Ind}_\mathcal{L} \left( Z_i \right) \right| \leq e^{-c\sqrt{T}}, \quad i = 1, 2, 3, \tag{7.15}
\]

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

\[
\text{Ind}_\mathcal{L} \left( \hat{Q}(r) \right) = \text{Ind}_\mathcal{L} \left( Q(-1) \right) + \mathcal{O}(e^{-c\sqrt{T}}). \tag{7.16}
\]

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

\[
\text{Ind}(R) = \mathcal{O}(e^{-c\sqrt{T}}). \tag{7.17}
\]

To establish this bound we observe that

\[
\text{Ind}(R) = \text{Ind}(R^\mathcal{X}) + \mathcal{O}(e^{-c\sqrt{T}}),
\]

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

\[
\text{Ind}(R^\mathcal{X}) = \text{itr} R^\mathcal{X} [\mathcal{X}, \Lambda_1], [\mathcal{X}, \Lambda_2],
\]

and the right hand side is \( \mathcal{O}(e^{-c\sqrt{T}}) \) using \( R^\mathcal{X} (1 - R^\mathcal{X}) = \mathcal{O}(e^{-c\sqrt{T}}) \) and the cyclicity of the trace. \( \square \)

**Appendix A. Hybridization in 1D**

In this appendix, we show hybridization of eigenvectors for a family of 1D Anderson Hamiltonians. Apart from an occasional reference to a definition or a technical lemma, this appendix is self-contained. In some places, the notation used here conflicts with the notation used in the main text.

We consider the Hilbert space \( \ell^2(\mathbb{Z}) \) and denote its scalar product by \( \langle \cdot, \cdot \rangle \). Delta functions \( \{\delta_x\}_{x \in \mathbb{Z}} \), equal to 1 at \( x \) and 0 elsewhere, form a basis for the Hilbert space. The discrete Laplacian \( \Delta \) is the operator given by

\[
\langle \delta_x, \Delta \delta_y \rangle = \begin{cases} -2 & x = y, \\ 1 & x \sim y, \\ 0 & \text{otherwise}, \end{cases}
\]

where \( x \sim y \) denotes \( |x - y| = 1 \). We recall that \( \sigma(-\Delta) = [0, 4] \). We will use a decomposition \( \Delta = \sum_{x \sim y} \Gamma_{xy} - 2 \), where \( \Gamma_{xy} \) is a rank one operator defined by \( \Gamma_{xy} f = f(x) \delta_y \) for \( f \in \ell^2(\mathbb{Z}) \). For a set \( Z \subset \mathbb{Z} \), we let \( \chi_Z = \sum_{x \in Z} \Gamma_{xx} \) be the orthogonal projection onto \( Z \).

Our results concern the analytic family of Hamiltonians \( H(\beta) \) with \( \beta \in (-1, 1) \) of the form

\[
H(\beta) = -\Delta + V_\omega + \beta W \tag{A.1}
\]

acting on \( \ell^2(\mathbb{Z}) \). Here, \( V_\omega \) is a random potential, with \( V_\omega(x) = \omega_x \) the i.i.d. random coupling variables distributed according to the Borel probability measure \( \overline{P} := \otimes_x P_0 \). We will assume that the single-site distribution \( P_0 \) is absolutely continuous with respect to Lebesgue measure on \( \mathbb{R} \). We assume that the corresponding Lebesgue density \( \mu \) is bounded with \( \text{supp}(\mu) \subset [0, 1] \), and that the single-site probability density is bounded away from zero on its support. We denote the configuration space by \( \Omega \). The perturbation \( W \) is a compactly supported non-negative potential. For concreteness, we anchor \( W \) at the origin by assuming that \( W(0) = 1 \) and \( \|W\| = 1 \), in particular \( \|H(0)\| \leq 4 \) in our setup. We remark that \( \sigma(H(0)) \) is a \( \mathbb{P} \)-a.s. deterministic set (see e.g., \( [3, \text{Theorem 3.10}] \)), which we denote by \( \Sigma \), and that \( \Sigma \supset [0, 4] \).
For a region $Z \subset \mathbb{Z}$, we write $H^Z = \chi_Z H \chi_Z$, understood as an operator acting on $\ell^2(Z)$. We will use the natural embedding $\ell^2(Z) \subset \ell^2(\mathbb{Z})$ without further comment. With some slight abuse of notation, $(a,b)$ denotes $(a,b) \cap \mathbb{Z}$ whenever it signifies a subset of the lattice.

We consider a length scale $\mathcal{L}$, a symmetric region $\Lambda_{\text{full}} := (-\mathcal{L}, \mathcal{L})$, and an asymmetric region $\Lambda := (-\mathcal{L}, 2\sqrt{\mathcal{L}}/\ln \mathcal{L})$ that we divide into a right region $\Lambda_R = (-2\sqrt{\mathcal{L}}/\ln \mathcal{L}, 2\sqrt{\mathcal{L}}/\ln \mathcal{L})$, and a left region $\Lambda_L = \Lambda \setminus \Lambda_R$. (The reasons for this asymmetry will be clear later on.) We denote by $r$ the leftmost point of $\Lambda_R$ and by $l$ the rightmost point of $\Lambda_L$, so by construction $l \sim r$. We consider the Hamiltonians associated with these regions, $H_{\text{full}} := H^{\Lambda_{\text{full}}}, H := H^\Lambda, H_L := H^{\Lambda_L}, \text{ and } H_R := H^{\Lambda_R}$, as well as the decoupled Hamiltonian $H_{\text{dec}}$ obtained by erasing the coupling between left and right regions, i.e. $H_{\text{dec}} = H_L + H_R = H - \Gamma_L - \Gamma_R$. All of these Hamiltonians a priori depend on $\beta$. Here and later, we only stress the dependence on $\beta$ in some equations, and suppress the dependence in others. We will assume henceforth that $\mathcal{L}$ is large enough so that $\text{supp}(W) \subset (-\mathcal{L}, \mathcal{L}) \subset \Lambda_R$. In particular, $H_L$ does not depend on $\beta$.

We consider an eigenvector $\varphi_L$ of $H_L$ with eigenvalue $E_L \equiv E$ and a continuous family of eigenvectors $\varphi_R(\beta)$ of $H_R(\beta)$ with eigenvalues $E_R(\beta)$. We will assume that these two energy level cross, i.e. $E - E_R(\beta)$ changes sign as $\beta$ varies. In Section A.2, we will show that such levels exist with large probability thanks to two-sided Wegner estimates.

For a typical realization of the disorder, the eigenvectors $\varphi_L, \varphi_R := \varphi_R(0)$ are well localized with localization centers $x_L, x_R$, respectively (we will make this statement quantitative later on). We pick the eigenvectors in such a way that $x_R$ is close to the origin and $x_L$ is located at least a distance of $\sqrt{\mathcal{L}}$ away from $\Lambda_R$. Let $P_{\text{dec}}$ be the orthogonal projection onto $\text{Span}(\varphi_L, \varphi_R)$. Let us consider the rank two operator $\mathbb{H} := P_{\text{dec}} H P_{\text{dec}}$ acting on $\text{Ran}(P_{\text{dec}})$. We note that the matrix representation for $\mathbb{H}$ with respect to $\{\varphi_L, \varphi_R\}$ basis is given by a $2 \times 2$ matrix

$$M_\beta := \begin{pmatrix} E & \text{gap} \\ \text{gap} & E_R + \beta \langle W \rangle_{\varphi_R} \end{pmatrix}$$

(A.2)

with $\text{gap} := \langle \varphi_L, H(0) \varphi_R \rangle = \varphi_L(l) \varphi_R(r), \langle W \rangle_{\varphi_R} := \langle \varphi_R, W \varphi_R \rangle$. Moreover, $\text{gap} \neq 0$ since eigenfunctions of a Schrödinger operator restricted to an interval do not vanish on its boundary. We now note that for $\beta$ such that $E_L = E_R + \beta \langle W \rangle_{\varphi_R}$, the eigenvectors $\varphi_L := \varphi_R \pm \varphi_L$ of $\mathbb{H}$ are delocalized in a sense that these functions are not small at both points $x_R$ and $x_L$ that are separated by a distance comparable with the system’s size. We call this phenomena a hybridization across lengthscale $\mathcal{L}$. We are going to show that such hybridization also occurs for eigenvectors of the full Hamiltonian $H_{\text{full}}(\beta)$.

**Definition A.1.** Let $F \in (0, 1/2]$ be a parameter. We say that $H_{\text{full}}(\beta)$ $F$-hybridize on a length scale $\mathcal{L}$ if there exists an analytical family of eigenvectors $\varphi(\beta)$ of $H_{\text{full}}(\beta)$ for $\beta \in (-1, 1)$ such that

$$\begin{align*}
(i) & \| \chi_{[x] \geq \sqrt{\mathcal{L}}/\ln \mathcal{L}} \varphi(0) \| \leq e^{-c \sqrt{\mathcal{L}}/\ln \mathcal{L}}, \\
(ii) & \text{There exists } \beta \text{ such that } \| \chi_{[x] \geq \sqrt{\mathcal{L}}/\ln \mathcal{L}} \varphi(\beta) \|^2 \geq F, \text{ and } \| \chi_{[x] \geq \sqrt{\mathcal{L}}/\ln \mathcal{L}} \varphi(\beta) \|^2 \geq F.
\end{align*}$$

We call $F$ a hybridization strength and denote by $\Omega_{F, \mathcal{L}} \subset \Omega$ all realizations for which $H_{\text{full}}(\beta)$ $F$-hybridize.

**Theorem A.2.** For any $F < 1/2$, $\liminf_{\mathcal{L} \to \infty} P(\Omega_{F, \mathcal{L}}) > 0$.

If we now consider an infinite volume operator $H(\beta)$ (i.e., $\Lambda_{\text{full}} = \mathbb{Z}$), any $F < \frac{1}{7}$, and an arbitrary sequence $\mathcal{L}_n \to \infty$, then by the Borel-Cantelli lemma, for almost all random configurations $\omega \in \Omega$ we can find a subsequence $\mathcal{L}_{n_k} \to \infty$ such that $H^{\Lambda_{\mathcal{L}_{n_k}}}(\beta)$ $F$-hybridizes.

While there could be potentially different mechanisms leading to the hybridization phenomenon, our construction below hinges on the behavior of the simple two-level system (characterized by the avoided eigenvalue crossing) discussed above. Since the probability of multiple level crossings is much smaller than that of two-level ones, we expect that in fact this is the only possible mechanism of hybridization, but in this work we did not try to formalize this statement. We chose this definition for its simplicity; our construction of the hybridization event is more detailed and exactly matches the underlying motivation.
A.1. Perturbation of a non-avoided crossing. We consider eigenvalues $E_L \equiv E, E_R(\beta)$ of $H_{\text{dec}}(\beta)$ for $\beta$ in a compact interval $J$, associated to (normalized) eigenvectors $\varphi_L, \varphi_R(\beta)$. Later, the notation $E_L$ will stand, more generally, for an eigenvalue of $H_L$ and $E_R$ for an eigenvalue of $H_R$, but this is not important at the moment. We assume that $\varphi_R(\beta)$ is continuous, which implies that $E_R(\beta)$ is continuous.

Let

$$h := \min\left\{ \max_{\beta \in (-1,1)} (E - E_R(\beta))_+, \max_{\beta \in (-1,1)} (E_R(\beta) - E)_+ \right\},$$

where $(x)_+$ is equal to $x$ for $x$ positive and zero otherwise. If the eigenvalues $E_R(\beta)$ do not intersect $E$ then $h = 0$, otherwise $h$ is a maximal number such that both $E - h$ and $E + h$ intersect $E_R(\beta)$.

Suppose now that the Hamiltonian $H(\beta)$ has a continuous family of spectral projections $P(\beta)$ such that (suppressing the $\beta$ dependence)

$$\|P - P_{\text{dec}}\| \leq \varepsilon, \quad \|H_{\text{dec}} P_{\text{dec}} - H P\| \leq \varepsilon,$$

(A.3)

for some $\varepsilon \ll 1$. The range of $P$ is then two dimensional and is spanned by (normalized) eigenvectors of $H$ that we denote $\varphi_{\pm}$. We denote the associated eigenvalues $E_{\pm}$. Let $c^+_L, c^+_R$ be the Fourier coefficients of $\varphi_{\pm}$ with respect to the elements $\varphi_L, \varphi_R$ of an eigenbasis of $H_{\text{dec}}$, i.e.,

$$\begin{align*}
\varphi_{\pm} &= c^+_L \varphi_L + c^+_R \varphi_R + \varphi_{\pm}, \\
\langle \varphi_{\pm}^+, \varphi_L \rangle &= \langle \varphi_{\pm}^+, \varphi_R \rangle = 0,
\end{align*}$$

and let

$$F := \max_{\beta \in (-1,1)} \min \left( |c^+_L(\beta)|^2, |c^+_R(\beta)|^2 \right).$$

This value will be used for the parameter $F$ introduced in Definition A.1.

Since $|c^+_L(\beta)|^2 + |c^+_R(\beta)|^2 \leq 1$, we know that $F \leq 1/2$. For $\varepsilon = 0$, $F$ equals zero by the continuity of $\beta$ dependence in $H, H_{\text{dec}}, P,$ and $P_{\text{dec}}$, so there is no hybridization. As can be seen from the two level system described in (A.2), $F$ can be equal to $1/2$ for an arbitrarily small but non-zero value of $\varepsilon$. Indeed, in this example $\varepsilon > 0$ corresponds to $\text{gap} > 0$ and $F = 1/2$ is achieved for $\beta$ that solves $E_L = E_R + \beta(W)_{\varphi_R}$.

Our principle indicator of hybridization will be the fact that $F$ has to be close to $1/2$ whenever the level crossing for $H_{\text{dec}}$ is avoided for the full $H$.

**Lemma A.3.** Suppose that $E_+(\beta), E_-(\beta)$ do not intersect in $J$, and $h \geq 4\varepsilon$. Then

$$F \geq \frac{1 - \varepsilon^2}{2}.$$
Proof. By the continuity of $E_{\pm}$ and the non-crossing condition, we may assume without loss of generality that $E_{+}(\beta) - E_{-}(\beta) > 0$ for $\beta \in (-1,1)$. By the first equation in (A.3), we know that $\|\varphi_{L}^{+}\| \leq \varepsilon$, hence

$$|c_{L}^{+}(\beta)|^{2} + |c_{R}^{+}(\beta)|^{2} \geq 1 - \varepsilon^{2}. \tag{A.4}$$

By the same equation,

$$\varepsilon \geq \langle \varphi_{L}, \varphi_{L} - P\varphi_{L} \rangle = 1 - \left( |c_{L}^{-}|^{2} + |c_{L}^{+}|^{2} \right), \quad \sharp = L, R. \tag{A.5}$$

On the other hand, the second equation in (A.3) implies

$$|c_{L}^{+}|^{2}(E_{-} - E_{\pm})^{2} \leq \varepsilon^{2}, \quad \sharp = L, R. \tag{A.6}$$

Using the second equation in (A.3) and Weyl’s theorem, [32, Theorem 4.3.1] we get

$$\text{dist}_{H}([0, E_{L}, E_{R}], \{0, E_{-}, E_{+}\}) = \text{dist}_{H}(\sigma(H_{dec}P_{dec}), \sigma(HP)) \leq \varepsilon, \tag{A.7}$$

where $\text{dist}_{H}$ stands for the Hausdorff distance between a pair of sets. Hence,

$$\text{dist}_{H}([E_{L}, E_{R}], \{E_{-}, E_{+}\}) \leq 2\varepsilon. \tag{A.8}$$

The definition of $h$ implies that there exist $\beta_{1}, \beta_{2} \in (-1,1)$ such that $E_{L} - E_{R}(\beta_{1}) = h$ and $E_{R}(\beta_{2}) - E_{L} = h$. Thus, it follows from (A.8) and $E_{+}(\beta) - E_{-}(\beta) > 0$ for $\beta \in (-1,1)$ that

$$\max(|E_{R}(\beta_{1}) - E_{-}(\beta_{1})|, |E_{L} - E_{R}(\beta_{1})|, |E_{R}(\beta_{2}) - E_{+}(\beta_{2})|, |E_{L} - E_{-}(\beta_{2})|) \leq 2\varepsilon. \tag{A.9}$$

Using (A.6) at $\beta_{1,2}$ with $\sharp = R$, we get $|c_{R}^{+}(\beta_{1})|^{2}(h - 2\varepsilon)^{2} \leq \varepsilon^{2}$ and $|c_{R}^{-}(\beta_{2})|^{2}(h - 2\varepsilon)^{2} \leq \varepsilon^{2}$, which imply $|c_{R}^{+}(\beta_{1})|^{2} \leq \frac{1}{4}$ and $|c_{R}^{-}(\beta_{2})|^{2} \leq \frac{1}{4}$. The latter relation yields $|c_{R}^{+}(\beta_{2})|^{2} \geq \frac{1}{4} - \varepsilon > \frac{1}{2}$ by (A.6). It follows from the continuity of the coefficient $c_{L}^{+}$ that there exists $\beta \in (\beta_{1}, \beta_{2})$ such that $|c_{R}^{+}(\beta)|^{2} = \frac{1}{2}\varepsilon$. Hence, by (A.4) we also have $|c_{L}^{+}(\beta)|^{2} \geq \frac{1}{2}\varepsilon^{2}$, completing the proof. \hfill \Box

A.2. Construction of the non-avoided crossing. We first give a precise notion of eigenvector localization.

**Definition A.4.** For $\omega \in \Omega$ and a pair $(\nu, \theta)$ of positive parameters, we will say that $H$ is $(\nu, \theta)$-localized if all eigenvalues of $H$ are simple and for each $\nu \in \sigma(H)$, the corresponding eigenvector $\psi_{E}$ satisfies

$$|\psi_{E}(y, \omega)|^{2} \leq \frac{1}{\theta} \langle x_{E}(\omega) \rangle^{2} e^{-\nu|y - x_{E}(\omega)|}. \tag{A.9}$$

We call $x_{E}$ the localization center of the eigenvector $\psi_{E}$.

One of the key results we will use in this appendix is

**Theorem A.5** (Eigenfunctions localization). There exist $C, \nu > 0$ such that

$$\mathbb{P}\left( \{ \omega \in \Omega : H_{0}^{2}(0) \text{ is } (\nu, \theta)\text{-localized} \} \right) \leq 1 - C\theta, \quad \sharp = \Lambda, \Lambda_{L}, \Lambda_{R}. \tag{A.10}$$

**Proof.** It is a consequence of [4, Theorems 5.8, 7.4, and 12.11] and Markov’s inequality. \hfill \Box

We will fix this value of $\nu$ henceforth.

**Definition A.6.** In this definition we gather requirements on $\omega \in \Omega$ that we use in our construction. The requirements depend on a small parameter $\theta < 1$, and a large parameter $b$.

There exists eigenvalues $E_{R}(0) \text{ (resp. } E_{L}(0) \text{) of } H_{R}(0) \text{ (resp. } H_{L}) \text{ with eigenvectors } \varphi_{L}, \varphi_{R} \text{ such that}

(i) $H_{L}, H_{R}(0)$ are $(\nu, \theta)$-localizing; In particular $\varphi_{L}, \varphi_{R}$ are localized;
(ii) $|E_{L} - E_{R}(0)| \leq b\theta/\mathcal{L}$;
(iii) Let

$$J := \{ \lambda \in \mathbb{R} : \text{dist}(\lambda, \{E_{L}, E_{R}(0)\}) \leq \sqrt[3]{\theta}/\mathcal{L} \} \tag{A.11}$$

Then $\sigma(H_{L}) \cap J = \{E_{L}\}$ and $\sigma(H_{R})(0) \cap J = \{E_{R}(0)\}$.

(iv) $|\varphi_{R}(0)|^{2} \geq -C_{\nu}/\ln \theta$. Here $C_{\nu}$ is an explicit constant given in Theorem C.2.

We will denote by $\mathcal{C}$ a set of all $\omega \in \Omega$ for which (i)-(iv) hold true.
For $\omega \in \mathcal{C}$, let $(E_R(\beta), \varphi_R(\beta))$ be the eigenpair of $H_R(\beta)$ that depends smoothly on $\beta \in J$.

**Proposition A.7.** Suppose that $\omega \in \mathcal{C}$, and that $\theta$ is small enough. Then $E_R(\beta), E_L$ intersect for some $\beta \in I$, where

$$I := [-a, a], \quad a = \frac{b}{C\nu} \frac{\theta \ln \theta}{L},$$

(A.12)

and the associated $h$ satisfies $h \geq b\theta / L$.

**Proof.** Let $P_R(\beta)$ be the projection on $\varphi_R(\beta)$. By the Hellmann-Feynman theorem

$$\dot{E}_R(\beta) = \text{tr} P_R(\beta) W; \quad \ddot{E}_R(\beta) = \text{tr} \dot{P}_R(\beta) W.$$  

Since $\|H_R(\beta) - H_R(0)\| \leq \beta$, by Weyl’s theorem

$$\text{dist} (E_R(\beta), \sigma(H_R(\beta))) \leq \text{dist} (E_R(0), \sigma(H_R(0))) \leq 2\beta \geq \frac{\sqrt{\theta}}{2L}$$

for $\beta \in I$ and $\theta$ sufficiently small. Hence, by the standard perturbation theory,

$$\|\dot{P}_R(\beta)\| \leq \frac{\beta}{\text{dist} (E_R(\beta), \sigma(H_R(\beta)))} \leq 2\beta \frac{L}{\sqrt{\theta}}.$$  

We now estimate

$$\dot{E}_R(\beta) = \dot{E}_R(0) + \int_0^\beta \ddot{E}_R(s) ds \geq -\frac{C\nu}{\ln \theta} - 2\beta^2 \frac{L}{\sqrt{\theta}} \geq -\frac{C\nu}{2\ln \theta}, \quad \beta \in I,$$

using Definition A.6(iv), $\text{Rank}(P_R) = 1$, and $\|W\| \leq 1$ in the second step. Hence

$$E_R(a) - E_R(0), E_R(0) - E_R(-a) \geq 2b \frac{\theta}{L}.$$  

Using Definition A.6(ii), we see that $h \geq b\theta / L$, completing the proof. \hfill \Box

**Lemma A.8.** For $b$ large enough, $\mathcal{P}(\mathcal{C}) \geq cb\theta$ for some constant $c$ independent of $\theta$ and $b$.

**Proof.** Let $C_k$ denote the event that the property (k) in Definition A.6 holds true. By (A.10), $\mathcal{P}(C_k) \geq 1 - C\theta$.

If $H_R(0)$ is $\nu, \theta$ localizing and (C.2) is satisfied for some interval $J$ and constant $c$, it follows from Lemma C.2 that there exists an eigenvalue $E_R$ of $H_R(0)$ and the associated eigenvector $\varphi_R$ such that $|\varphi_R(0)|^2 > -C\nu / \ln(\theta)$. As we show in Lemma C.1 (C.2) indeed holds deterministically with a choice $J = [\frac{1}{L}, \frac{L}{10}]$, $c = \frac{1}{10}$. Thus we can pick $C_{iv} := C_i$.

To bound $\mathcal{P}(C_{ii})$ we will invoke

**Theorem A.9** (Two-sided Wegner estimate). Let $K \subset \mathbb{Z}$ be an interval. Then for any compact subinterval $J$ of $(0, 4)$ there exist $L_0 > 0$ and constants $C_+ \geq C_- > 0$ such that we have

$$C_- |J| |K| \leq E \left( \text{tr}_{\chi_J}(H^K) \right) \leq C_+ |J| |K|,$$

(A.13)

provided $|K| > L_0$.

**Proof.** The upper bound is well known, see e.g., [4, Corollary 4.9]. The lower bound was recently established in [26, Theorem 1.1] in the continuum setting, but the same proof works for lattice systems considered here just as well. \hfill \Box

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

**Theorem A.10** (Minami estimate). Under the same assumptions as in Theorem A.9 for any $n \in \mathbb{N}$ we have

$$\mathcal{P} \left( \text{tr}_{\chi_J}(H^K) \geq n \right) \leq \frac{1}{n!} \left( C_+ |J| |K| \right)^n.$$  

(A.14)

**Proof.** In this generality, the bound goes back to [15], see also [4, Theorem 17.11]. \hfill \Box
Let \( \hat{I} := [E_R(0) - b\theta/L, E_R(0) + b\theta/L] \). Combining the lower bound in (A.13) with (A.14) and using the statistical independence of \( H_L \) and \( H_R(0) \), we see that

\[
\mathbb{P}(\text{tr} \chi_J(H_L) \geq 1) \geq \mathbb{E}(\text{tr} \chi_J(H_L)) - \sum_{n=2}^{\infty} (n-1) \mathbb{P}(\text{tr} \chi_J(H_L) \geq n) \geq c b \theta
\]  

(A.15)

for some \( b \)-independent constant \( c > 0 \). This implies \( \mathbb{P}(C_{ii}) \geq c b \theta \) for such \( b \).

This leaves us with the estimation of \( \mathbb{P}(C_{ii}) \). Let \( \hat{J} := \{ \lambda \in \mathbb{R} : |\lambda - E_R(0)| \leq 2\sqrt{\theta}/L \} \).

Then \( J \subset \hat{J} \) for \( J \) specified in \( \text{(A.11)} \) and, using the statistical independence of \( H_L \) and \( H_R(0) \), by \( \text{(A.14)} \)

\[
\mathbb{P}(\text{tr} \chi_J(H_L) \geq 1) \leq \mathbb{P}(\text{tr} \chi_J(H_L) \geq 2) \leq C \theta. 
\]  

(A.16)

To complete the argument, we will use the following consequence of Theorem \( \text{A.10} \).

**Theorem A.11.** Let \( \delta > 0 \) and let \( \mathcal{E}_\omega \) be an event

\[ \mathcal{E}_\omega := \{ \sigma(H^K) \text{ is } \delta \text{-level spaced on } \Lambda \}. \]

Then there exists \( C > 0 \) such that

\[ \mathbb{P}(\mathcal{E}_\omega) \geq 1 - C \delta |K|^2. \]

**Proof.** This statement is essentially [37, Lemma 2], in the formulation given in [21, Lemma B.1]. \( \square \)

Applying this with the choice \( K = \Lambda_R \), we deduce that

\[ \mathbb{P}(\text{tr} \chi_J(H_R) \geq 1) \leq C \sqrt{\theta}/\ln^2 L \leq \theta \]

for \( L \) large enough. This yields \( \mathbb{P}(C_{ii}) \geq 1 - C \theta \).

Putting our bounds on \( C_i \)–\( C_{iv} \) together, we see that for \( b \) large enough \( \mathbb{P}(C) \geq c b \theta \) for some constant \( c > 0 \). \( \square \)

### A.3. Construction of the avoided crossing.

In addition to \( \omega \in \mathcal{C} \), we will assume further properties of \( \omega \) that will allow us to use perturbation theory to study the crossing.

**Definition A.12.** Let \( \Lambda_B \) be a region of size \( L^{1/8} \), centered at the boundary between \( \Lambda_L \) and \( \Lambda_R \), i.e. (recall (1.18)–(1.19)) \( \Lambda_B = (\theta \Lambda_R)L^{1/8} \). We pick \( b_L, b_R \in \mathbb{Z} \) so that \( \Lambda_B = (b_L, b_R) \). We denote \( H_B := H^{\Lambda_B} \), the Hamiltonian restricted to this region. We will say that \( \omega \in \mathcal{A} \) if \( \omega \in \mathcal{C} \) and the following items hold true:

(i) \( H_B \) has no spectrum in the interval \( \hat{J} := (E_L - \theta^{-1} L^{-1/2}, E_L + \theta^{-1} L^{-1/2}) \).

(ii) There are at most two eigenvalues of \( H(0) \) in the interval \( J \) defined in (A.11).

(iii) Centers of \( \varphi_L \) and \( \varphi_R \) are distance of order \( \sqrt{\mathcal{L}}/\ln \mathcal{L} \) away from the boundary of \( \Lambda_R \).

Specifically,

\[
\left\| \chi_{\{x > \sqrt{\mathcal{L}}/(4 \ln \mathcal{L})\}} \varphi_R \right\| + \left\| \chi_{\{x > -3 \sqrt{\mathcal{L}}/(4 \ln \mathcal{L})\}} \varphi_L \right\| \leq e^{-c \sqrt{\mathcal{L}}/(4 \ln \mathcal{L})}. 
\]  

(A.18)

(iv) For \( \lambda \in \hat{J} \),

\[
\left\| \chi_{\{x - l \geq \mathcal{L}^{1/8}\}} (H_B - \lambda)^{-1} \right\| \leq e^{-c \mathcal{L}^{1/8}}. 
\]

(v) For \( \sharp = L, R \),

\[
\left\| \delta t, (H_B - E_0)^{-1} \right\| \geq 2 \theta \delta. 
\]

We note that condition (i) above ensures that resolvents in (iv)-(v) are well defined.

The dependence on parameter \( \theta \) in the above definition is chosen so that \( \mathbb{P}(\mathcal{A}) = O(\theta) \). We will establish this at the end of this section.

Let \( \varphi_R(\beta) \) be an eigenvector of \( H_R(\beta) \), which is an analytic continuation of \( \varphi_R(0) \). (Note that \( H_R(\beta) \) is a finite rank operator, so its eigenvectors do have analytical continuation on the real line, c.f. [32]). We recall that \( H_L(\beta) \) is \( \beta \)-independent, hence \( \varphi_L(\beta) \equiv \varphi_L \). We first show
Lemma A.13. Assume that \( \omega \in A \). For \( \beta \in I \) defined in (A.12),

\[
\left\| \chi_{\{|x|>\sqrt{\mathcal{L}}/(4\ln \mathcal{L})\}} \varphi_R(\beta) \right\| \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}. \tag{A.19}
\]

Proof. Let \( \hat{H}_R(0) = H_R(0) + (1 - E_R(0))P_R(0) \), where \( P_R(0) \) is an orthogonal projection onto \( \text{Span}(\varphi_R(0)) \). We observe that by Definition A.6(ii) and \( |E_R(\beta) - E_R(\beta)| \leq \beta \),

\[
\left\| (\hat{H}_R(0) - E_R(\beta))^{-1} \right\| \leq \frac{2\mathcal{L}}{\sqrt{\beta}}, \quad \beta \in I. \tag{A.20}
\]

We have

\[
\chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln \mathcal{L})\}} \varphi_R(\beta) = \chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln \mathcal{L})\}} (\hat{H}_R(0) - E_R(\beta))^{-1} (\hat{H}_R(0) - E_R(\beta)) \varphi_R(\beta)
\]

\[
= \chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln \mathcal{L})\}} (\hat{H}_R(0) - E_R(\beta))^{-1} ((1 - E_R(0))P_R(0) + \beta W) \varphi_R(\beta)
\]

To estimate the right hand side, we note that

\[
\left\| \chi_{\{|x|>\sqrt{\mathcal{L}}/(4\ln \mathcal{L})\}} (P_R(0) + \beta W) \right\| \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}
\]

by (A.18) and compactness of supp(W). Hence (A.19) will follow once we show that

\[
\left\| \chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln \mathcal{L})\}} (\hat{H}_R(0) - E_R(\beta))^{-1} \chi_{\{|x|\leq \sqrt{\mathcal{L}}/(4\ln \mathcal{L})\}} \right\| \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}.
\]

The later bound is a consequence of the spectral theorem, the estimate (A.20), and the fact that \( H_R(0) \) (and hence \( \hat{H}_R(0) \)) is \((\nu,\theta)\)-localizing for \( \omega \in A \).

We recall that \( P_{\text{dec}}(\beta) \) denotes the orthogonal projection onto \( \text{Span}(\varphi_L, \varphi_R(\beta)) \). By standard perturbation theory, \( P_{\text{dec}}(\beta) \) is a spectral projection of \( H_{\text{dec}}(\beta) \) for all \( \beta \in I \). We first establish that \( P_{\text{dec}}(\beta) \) is close to a spectral projection of \( H(\beta) \).

Proposition A.14. Assume that \( \omega \in A \). Then for \( \beta \in I \) (recall (A.11) and (A.12)) we have

(i) \( \sigma(H(\beta)) \cap J = \{ E_-(\beta), E_+(\beta) \} \) where \( E_\pm(\beta) \) are real analytic in \( \beta \);

(ii) \( \text{dist} \{E_-(:\beta),E_+(\beta)\},\{E_L,E_R(\beta)\} \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}} \);

(iii) Let \( P(\beta) \) be the spectral projection on \( E_\pm(\beta) \), then \( \|P(\beta) - P_{\text{dec}}(\beta)\| \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}} \);

(iv) We can label \( E_\pm(\beta) \) so that the associated eigenfunctions \( \varphi_{\pm}(\beta) \) satisfy

\[
|\langle \varphi_-(0), \varphi_R(0) \rangle|^2 \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}, \quad |\langle \varphi_+(0), \varphi_L \rangle|^2 \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}.
\]

Proof. By Lemma C.4, (A.18), and Lemma A.13 we deduce that

\[
\text{dist}(\sigma(H(\beta)), E_\pm(\beta)) \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}, \quad \sharp = L,R.
\]

It follows that \( H(\beta) \) has at least two eigenvalues in the interval \( I \). Combined with standard perturbation theory and the fact that for \( \omega \in A \) the operator \( H(0) \) has at most two eigenvalues in \( J \), see Definition A.12(ii), we see that Proposition A.14.(i)–A.14.(iii) hold. The last statement follows from (A.18), Lemma A.13 and Lemma C.4.

Proposition A.15. Suppose that \( \omega \in A \), then the eigenvalues \( E_\pm(\beta) \) cannot intersect each other in the interval \( I \).

We start with the following preliminary observation.

Lemma A.16. The operator \( \hat{P}_{\text{dec}}(\beta) = (H(\beta) - \lambda)^{-1} \hat{P}_{\text{dec}}(\beta) \) is invertible on the range of \( \hat{P}_{\text{dec}}(\beta) \)

for all \( \lambda \in J \) and \( \beta \in I \), and the norm of the inverse is bounded by \( C\mathcal{L}/\sqrt{\theta} \).
Proof. It is a standard result in perturbation theory that if $B$ is invertible and $\|B^{-1}\|(A-B)\| < 1$, then $A$ is invertible and

$$\|A^{-1}\| \leq \frac{\|B^{-1}\|}{1 - \|B^{-1}\|\|A-B\|}.$$  

To prove Lemma A.16, we combine this observation with

$$B = \bar{P}(\beta)(H(\beta) - \lambda)\bar{P}(\beta) + P(\beta), \quad A = \bar{P}_{\text{dec}}(\beta)(H(\beta) - \lambda)\bar{P}_{\text{dec}}(\beta) + P_{\text{dec}}(\beta).$$

By Proposition A.14, $\|A-B\| \leq e^{-c\sqrt{\mathcal{L}/\ln \mathcal{L}}}$. By $\omega \in \mathcal{A}$, $B^{-1}$ is invertible with

$$\|B^{-1}\| \leq C\frac{\mathcal{L}}{\sqrt{\theta}}.$$ 

Now, we note that $A$ is block diagonal with respect to $P_{\text{dec}}(\beta), \bar{P}_{\text{dec}}(\beta)$, and its inverse exists if and only if each associated block has an inverse. \hfill \Box

Proof of Proposition A.15: We will suppress the $\beta$ dependence and use the shorthand $P$ for $P_{\text{dec}}(\beta)$ for the proof. The idea here is to use Schur complementation. Namely, given $\lambda \in J$, we consider $M = M(\beta, \lambda)$, the Schur complement of $H$ in $\text{Ran}(\bar{P})$, defined as

$$M := P(H(\lambda) - \lambda)P - PH\bar{P}(H(\lambda) - \lambda)\bar{P}H P^{-1} \bar{P}H P.$$  

We note that by Lemma A.16, $M$ is well-defined for our range of $\lambda$'s and $\beta$'s. $M$ is a rank 2 operator whose range is spanned by $(\varphi_\beta, \varphi_\lambda)$. Using the Guttman rank additivity formula, p. 14, we see that $\text{tr}(\chi_{\Lambda}(H)) = 2$ (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 $M_{LR} = \langle \varphi_L, M\varphi_R \rangle \neq 0$. We claim that

$$M_{LR} = \langle \varphi_L(l)\varphi_R(r) - (\delta_r, (H_B - E_-)^{-1}\delta_l) + \text{Error} \rangle,$$  

(A.21)

where $|\text{Error}| \leq \theta^2$. Since $\omega \in \mathcal{A}$, by Definition A.12(v) we have

$$\left|\langle \delta_r, (H_B - E_-)^{-1}\delta_l \rangle - 1 \right| \geq \theta^4.$$ 

Hence, for sufficiently large $\mathcal{L}$, $M_{LR} \neq 0$ as the eigenfunctions of $H_{LR}$ cannot vanish at the respective boundary points.

It remains to derive (A.21). We recall that $\Gamma := \Gamma_{lr} + \Gamma_{rl}$ is the hopping term connecting the region $\Lambda_R$ to the region $\Lambda_L$. In particular, $\Gamma\varphi_L = \varphi_L(l)\delta_r$ and $\Gamma\varphi_R = \varphi_R(r)\delta_l$. We use these equations to evaluate the terms in

$$M_{LR} = \langle \varphi_L, (H - \lambda)\varphi_R \rangle - \langle \varphi_L, PHP(H - \lambda)^{-1}\bar{P}HP\varphi_R \rangle,$$ 

where we denote $\bar{H} = \bar{P}H \bar{P}$, and let $(H - \lambda)^{-1}$ denote the inverse of $H - \lambda$ on the $\text{Ran}(\bar{P})$. The first term is equal to

$$\langle \varphi_L, H\varphi_R \rangle = \langle \varphi_L, \Gamma\varphi_R \rangle = \varphi_L(l)\varphi_R(r).$$

To evaluate the second term, we use the identity $PHP = P\Gamma P$ to get

$$\langle \varphi_L, PHP(H - \lambda)^{-1}\bar{P}HP\varphi_R \rangle = \varphi_L(l)\varphi_R(r)\langle \delta_r, (H - \lambda)^{-1}\delta_l \rangle.$$ 

We next use the resolvent identity

$$(H - \lambda)^{-1} = (H_B - \lambda)^{-1} + T, \quad T := (H - \lambda)^{-1}(H_B - H + PHP)(H_B - \lambda)^{-1}.$$ 

We note that since $\omega \in \mathcal{A}$, by Definition A.12 iii) the resolvent $(H_B - \lambda)^{-1}$ is well defined and its norm is bounded by $C\mathcal{L}^{1/4}$. Moreover, since $(H_B - H)\chi_{\{\|x-l\|<\mathcal{L}^{1/8}\}} = 0$, by Definition A.12 iv), A.18, and Lemma A.13 we get

$$\|T\delta_l\| \leq 5 \left\|\bar{H} - \lambda\right\|^{-1} \left\|\chi_{\{\|x-l\|<\mathcal{L}^{1/8}\}}(H_B - \lambda)^{-1}\delta_l\right\| + \left\|\bar{H} - \lambda\right\|^{-1} \left\|P\chi_{\{\|x-l\|<\mathcal{L}^{1/8}\}}\right\| \left\|(H_B - \lambda)^{-1}\right\| \leq Ce^{-c\mathcal{L}^{1/8}};$$
which implies, in particular, that
\[|\langle \delta_r, T \delta_l \rangle| \leq Ce^{-c \mathcal{L}^{1/8}}.\]
Furthermore, by the standard perturbation theory and Definition [A.12(iii)]
\[\left\| (H_B - \lambda)^{-1} - (H_B - E_-)^{-1} \right\| \leq C|E_- - \lambda| \theta^2 \mathcal{L}.\]
Since \(E_- - \lambda\) is of order \(\mathcal{L}^{-1}\) for \(\lambda \in J\), we get ([A.21]).

We now show

**Lemma A.17.** \(\mathbb{P}(\mathcal{A}) \geq c \theta\) for some constant \(c\).

**Proof.** Let \(\mathcal{A}_k\) denote the event that property (k) in Definition [A.12] holds true.

Using the upper bound in ([A.13]), we get \(\mathbb{P}(\mathcal{A}_i) \geq \mathbb{P}(\mathcal{C}) - C \theta^{-1} \mathcal{L}^{-1/2} \mathcal{L}^{1/8} \geq c \theta\) for \(\mathcal{L}\) large enough. On the other hand, using ([A.14]), we deduce that
\[\mathbb{P}(\mathcal{A}_{ii} \cap \mathcal{A}_i) \geq \mathbb{P}(\mathcal{A}_i) - \mathbb{P}(\mathrm{tr}_{\mathcal{J}}(H(0)) \geq 3) \geq \mathbb{P}(\mathcal{A}_i) - C \theta^{3/2} \leq c \theta.\]

Let \(\mathcal{A}_L = [-4\sqrt{\mathcal{L}}/\ln \mathcal{L}, \ell]\). Then using the upper bound in ([A.13]), for \(\mathcal{L}\) large enough,
\[\mathbb{P}\left(\mathrm{tr}_{\mathcal{J}}\left(H_{\mathcal{L}}^\Lambda\right) = 0\right) \geq 1 - C(\sqrt{\mathcal{L}}/\ln \mathcal{L}) \theta^{-1} \mathcal{L}^{-1/2} \geq 1 - \theta^2.\]

Let \(\mathcal{E} := \mathcal{A}_{ii} \cap \mathcal{D}\), where \(\mathcal{D}\) is the event \(\mathrm{tr}_{\mathcal{J}}\left(H_{\mathcal{L}}^\Lambda\right) = 0\). Then we see that
\[\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(\mathcal{A}_{ii}) - \theta^2 \geq c \theta.\]

We claim that ([A.18]) holds for \(\omega \in \mathcal{E}\), implying that \(\mathbb{P}(\mathcal{A}_{iii} \cap \mathcal{A}_{ii}) \geq c \theta\). Indeed, the bound
\[\left\| \chi_{\{|x| \geq \sqrt{\mathcal{L}}/(4 \ln \mathcal{L})\}} f \right\| \leq e^{-c \sqrt{\mathcal{L}}/\ln \mathcal{L}}\]
follows directly from Definition [A.6] parts (i,iv) (we recall that \(\mathcal{A} \subset \mathcal{C}\)). On the other hand, if the localization center for \(\varphi_L\) were located in \([-\frac{7}{8}\sqrt{\mathcal{L}}/\ln \mathcal{L}, \ell]\), Definition [A.6(i)] would imply that
\[\left\| \chi_{\{|x| < -4\sqrt{\mathcal{L}}/\ln \mathcal{L}\}} \varphi_L \right\| \leq e^{-c \sqrt{\mathcal{L}}/\ln \mathcal{L}}.\]
But then we would have
\[\text{dist}(E_L, \sigma(H_{\mathcal{L}}^\Lambda)) \leq e^{-c \sqrt{\mathcal{L}}/\ln \mathcal{L}}\]
thanks to Lemma [C.3] contradicting \(\mathrm{tr}_{\mathcal{J}}\left(H_{\mathcal{L}}^\Lambda\right) = 0\). This implies that the localization center for \(\varphi_L\) is located in \(\Lambda_L \setminus [-\frac{7}{8}\sqrt{\mathcal{L}}/\ln \mathcal{L}, \ell]\), which in turn implies that
\[\left\| \chi_{\{|x| > 3\sqrt{\mathcal{L}}/\ln \mathcal{L}\}} \varphi_L \right\| \leq e^{-c \sqrt{\mathcal{L}}/\ln \mathcal{L}}\]
by Definition [A.6(i)].

To estimate \(\mathbb{P}(\mathcal{A}_{iv} \cap \mathcal{A}_{iii})\), we note that our assumptions on randomness imply
\[\sup_{\lambda \in \mathbb{R}} \mathbb{E}_{\chi_{\{|x| \geq \mathcal{L}^{1/8}\}}}(H_B - \lambda - i0)\left(\frac{\lambda}{\theta}\right)^{-1}\delta_l \right\| \leq Ce^{-c \mathcal{L}^{1/8}}.\]

[4. Theorem 12.11]. Hence, denoting
\[\mathcal{F} := \left\{ \omega \in \Omega : \left\| \chi_{\{|x| \geq \mathcal{L}^{1/8}\}}(H_B - \lambda - i0)\left(\frac{\lambda}{\theta}\right)^{-1}\delta_l \right\| \leq e^{-c \mathcal{L}^{1/8}} \right\},\]
we see that \(\mathbb{P}(\mathcal{A}_{iv} \cap \mathcal{A}_{iii}) \geq c \theta\) for \(\mathcal{L}\) large enough by Markov’s inequality.

Finally, the bound \(\mathbb{P}(\mathcal{A}_{v} \cap \mathcal{A}_{iv}) \geq c \theta\) is a direct consequence of

**Lemma A.18.** For a fixed \(s \in (0, 1/2)\) and \(\lambda \in I\), we have
\[\mathbb{P}\left(\left\{ \omega \in \Omega : \left| \langle \delta_r, (H_B - E)^{-1} \delta_l \rangle - 1 \right| \geq \theta^s \mathcal{L} \right\} \right) \geq 1 - C_s \theta.\]

**Proof of Lemma A.18.** Let \(G(x, y) := \langle \delta_x, (H_B - \lambda)^{-1} \delta_y \rangle\). We first observe that, thanks to the geometric resolvent identity (or just directly by [4, Eq. 12.7]),
\[G(l, r) = \hat{G}(l, l)G(r, r),\]
where $\hat{G}(x, y) = \langle \delta_x, (\hat{H}_B - \lambda)^{-1} \delta_y \rangle$ and $\hat{H}_B$ is obtained from $H_B$ by the removal of the $(l, r)$ bond, i.e., $\hat{H}_B = H_B - \Gamma_{(l, r)} - \Gamma_{(r, l)}$. We use the resolvent identity

$$\hat{G}(r, r) = G(r, r) - \hat{G}(r, r)\hat{G}(l, l)G(r, r)$$

to obtain

$$\frac{1}{G(l, l)G(r, r) - 1} = -\frac{\hat{G}(r, r)}{G(r, r)},$$

where $\hat{G}(x, y) := \langle \delta_x, (H_B + \hat{G}(l, l)\chi_{(r)} - \lambda)^{-1} \delta_y \rangle$. The important fact here is to note that $\hat{G}(l, l)$ is independent of the $\omega_r$ random variable. This independence allows us to conclude that

$$\mathbb{E}_{\omega_1} \left| \hat{G}(r, r) \right|^s \leq C_s, \quad s \in (0, 1).$$

On the other hand, under our conditions on the probability distribution $\mu$, we also have (see \[4, Theorem 12.8\])

$$\mathbb{E} \left| G(r, r) \right|^{-s} \leq C_s, \quad s \in (0, 1).$$

Combining these two bounds and using the Hölder inequality, we deduce that

$$\mathbb{E} \left| \frac{1}{G(l, l)G(r, r) - 1} \right|^s \leq C_s, \quad s \in (0, 1/2),$$

from which the assertion follows by the Markov inequality. \hfill \Box

A.4. Proof of Theorem A.2

Theorem A.19. Let us denote by $\tilde{\Omega}_{F, L} \subset \Omega$ all realizations for which $H(\beta)$ $F$-hybridize. Let $\omega \in A$ and $F < 1/2$. Then, for $L$ large enough, $\omega \in \tilde{\Omega}_{F, L}$.

Proof. Consider an analytical family of eigenvectors $\varphi_R(\beta)$, $\varphi_L$ of $H_{\text{dec}}(\beta)$ and the analytical family $\varphi_{\pm}(\beta)$ of eigenvectors of $H(\beta)$. We are going to show that $\varphi(\beta) := \varphi_{\pm}(\beta)$ is an analytical family whose existence is required in the Definition A.1 of $\Omega_{F, L}$. We recall that the families are labeled in such a way that at $\beta = 0$, $\varphi_{\pm}$ has exponentially small overlap with $\varphi_{\pm}$. In particular, $\varphi_{\pm}(0)$ satisfies item (i) in Definition A.1

By Proposition A.14, the families satisfy (A.3) with $\varepsilon = e^{-c\sqrt{L}/\ln L}$. Proposition A.7 implies that the bandwidth of the crossing satisfies $h > 4\varepsilon$. It then follows from Lemma A.3 that there exists $\beta$ such that

$$\varphi_{\pm}(\beta) = c_L^+(\beta)\varphi_L + c_R^+(\beta)\varphi_R + \varphi_{\pm}^1,$$

with

$$|c_L^+(\beta)|^2 = |c_R^+(\beta)|^2 \geq \frac{1 - c^2}{2}.$$

It follows that item (ii) of Definition A.1 is satisfied for any $F < 1/2$, provided $L$ is large enough. \hfill \Box

As a corollary of the above result and Lemma A.17 we get that for any $F < 1/2$,

$$\liminf_{L \to \infty} \mathbb{P}(\Omega_{F, L}) > 0.$$

The assertion of Theorem A.2 is established completely analogously, by splitting $\Lambda_{\text{full}}$ into $\Lambda_L$, $\Lambda_R$, and $-\Lambda_L$, and then repeating the same steps as above. The reason that we present a proof for the asymmetric region is related to the fact that, in this case, the boundary of $\Lambda_R$ consists of a single point $r$, whereas in the symmetric case it consists of two points $\pm r$, making the presentation slightly more cumbersome. \hfill \Box
Appendix B. A Wannier basis for quasi-local projections

Here, we show the existence of a (generalized) Wannier basis for a rank \( m \) orthogonal projection \( P \) on \( \ell^2(\mathbb{Z}^d) \) that satisfies the quasi-locality property (B.3) below. To illustrate this concept, we start with the case \( m = 1 \).

**Lemma B.1.** Suppose that the normalized vector \( \psi \in \ell^2(\mathbb{T}_L) \) satisfies

\[
\max_{x,y \in \mathbb{T}_L} \left( |\psi(x)| \left| \psi(y) \right| e^{c|x-y|} \right) \leq \frac{1}{\theta}.
\] (B.1)

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

\[
|\psi(y)| \leq |\ln \theta|^{d+1} \theta^{-2c|y-x_o|}, \quad y \in \mathbb{T}_L.
\]

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

\[
1 = \sum_{x \in \mathbb{T}_L} |\psi(x)|^2 \leq M(\psi) \sum_{x \in \mathbb{T}_L: |x - x_o| \leq r} 1 + \sum_{x \in \mathbb{T}_L: |x - x_o| > r} \frac{e^{-2c|x-x_o|}}{M(\psi)^2 \theta^2} \leq M(\psi)(2r + 1)^d + \frac{1}{2}.
\] (B.2)

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

\[
e^{-Cu^{-\frac{1}{d}}} = \theta^2 u^2.
\]

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

While considering the rank one projection \( P \) is sometimes enough for random operators (e.g., for the randomness given by the rank one single site potential as in the standard Anderson model), in general it is not known whether the spectrum of a random operator that satisfies Assumptions 1.3–1.4 is a.s. simple or even has finite multiplicities. For our applications, one needs to be able to decompose \( P \) into a sum of rank one mutually orthogonal projections that individually exhibit exponential decay. Such a decomposition is called a (generalized) Wannier basis for \( P \). In general, finding a Wannier basis is a hard problem, due to a topological obstruction, see e.g., [4]. Here, we assert its existence for a finite rank \( P \) with explicit control over its rank \( m \), which is sufficient for our purposes.

**Theorem B.2.** Let \( m \in \mathbb{N}, \theta > 0 \) be such that \( m^3 \theta \ll 1 \). Suppose that a rank \( m \) orthonormal projection \( P \in \mathcal{L}(\mathcal{H}), \mathcal{H} = \ell^2(\mathbb{Z}^d) \) satisfies

\[
\max_{x,y \in \mathbb{Z}^d} \left( |P(x,y)| e^{c|x-y|} \right) \leq \theta^{-1}.
\] (B.3)

Then we can decompose \( P \) as \( P = \sum_{i=1}^{m} P_i \), where \( P_i = |\psi_i\rangle \langle \psi_i| \) are rank one mutually orthogonal projections that satisfy \( M(\psi_i) \geq |\ln \theta|^{-d-1} \) and, for some \( x_i \in \mathbb{Z}^d \),

\[
|\psi_i(y)| \leq \theta^{-2} e^{-c|y-x_i|/m}, \quad y \in \mathbb{Z}^d.
\]

We stress that the constant \( c \) here is \( m \)-independent.

**Proof.** We will need some preparatory results.

**Lemma B.3.** Let \( M = \max_{x \in \mathbb{Z}^d} P(x,x) \). Then there exists a (\( \theta \)-independent) \( C > 0 \) such that \( M \geq u \), where \( u \) is a unique positive solution of

\[
e^{-Cu^d} = \theta u^2.
\] (B.4)

In particular, for \( \theta \) sufficiently small, \( M \geq |\ln \theta|^{-d-1} \).

\[47\]
Let \( r = r(c, \theta) > 0 \) be such that \( \sum_{y \in \mathbb{Z}^d : |y| > r} e^{-2c|y|} \leq \theta^2 M^2 \). In particular, there exists \( C \) such that we can chose \( r = -C \ln (\theta M) \) for \( \theta \) sufficiently small. Let \( x_o \in \mathbb{Z}^d \) be such that \( M = \mathbb{P}(x_o, x_o) \) (\( x_o \) is not necessary unique). We then have

\[
M = \mathbb{P}(x_o, x_o) = \sum_{y \in \mathbb{Z}^d} |\mathbb{P}(x_o, y)|^2 = \sum_{y \in L^o(x_o)} |\mathbb{P}(x_o, y)|^2 + \sum_{y \in L^e(x_o)} |\mathbb{P}(x_o, y)|^2
\]

and continue to estimate the two sums on the right hand side separately. For the first sum, we use the positivity of \( \mathbb{P} \) to deduce that \( |\mathbb{P}(x_o, y)|^2 \leq \mathbb{P}(x_o, x_o) P(y, y) \), which yields

\[
\sum_{y \in L^o(x_o)} |\mathbb{P}(x_o, y)|^2 \leq M^2 |L^o(x_o)| < (3^d - 1)M^2 r^d.
\]

On the other hand, using \( (B.3) \) we get

\[
\sum_{y \in L^e(x_o)} |\mathbb{P}(x_o, y)|^2 \leq \theta^{-2} \sum_{y \in L^e(x_o)} e^{-c|x_o - y|} \leq M^2.
\]

Putting the two bounds together, we obtain \( M \leq 3^d M^2 r^d \), which in turn yields \( M \geq u \), where \( u \) is implicitly given by \( (B.4) \). Since \( u > |\ln \theta|^{-d-1} \) for \( \theta \) sufficiently small, we get \( M \geq |\ln \theta|^{-d-1} \).

Let \( L = L(c, \theta) > 0 \) be such that

\[
\sum_{\Lambda_L^{e}(a)} e^{-2c|y|} \leq \theta^6 M \tag{B.5}
\]

with \( M \) as above. In particular, there exists \( C \) such that we can choose

\[
L = -C \ln \theta \tag{B.6}
\]

for \( \theta \) sufficiently small. Consider

\[
\Xi_L := \left( \frac{3}{2} L \mathbb{Z} \right)^d \tag{B.7}
\]

cf. \( (1.12) \), and an \( L \)-cover of \( \mathbb{Z}^d \) of the form

\[
\mathbb{Z}^d = \bigcup_{a \in \Xi_L} \Lambda_L(a).
\]

We note that for any \( x \in \mathbb{Z}^d \) we can find \( a \in \Xi_L \) such that \( \text{dist}(\Lambda_L^e(a), x) \geq L/4 \).

**Lemma B.4.** For \( L \) as above, let \( T = \max_{a \in \Xi_L} \text{tr} \mathbb{P}_{\chi \Lambda_L(a)} \). Then \( T \geq 1/2 \) for \( \theta \) sufficiently small.

**Proof.** Suppose in contradiction that \( \text{tr} \mathbb{P}_{\chi \Lambda_L(a)} < 1/2 \) for any \( a \in \Xi_L \). Picking \( x_o \) as in the previous lemma and letting \( a \in \Xi_L \) be such that \( \text{dist}(\Lambda_L^e(a), x_o) \geq L/4 \), we have

\[
M \leq \mathbb{P}(x_o, x_o) \sum_{y \in \Lambda_L(a)} \mathbb{P}(y, y) + \sum_{y \in \Lambda_L^e(a)} |\mathbb{P}(x_o, y)|^2 \leq M \sum_{y \in \Lambda_L(a)} \mathbb{P}(y, y) + \theta^4 M < 2M/3,
\]

a contradiction.

We now observe that since \( \text{tr} \mathbb{P} = m \), the cardinality of a set

\[
\mathcal{S} := \{ a \in \Xi_L : \text{tr} \mathbb{P}_{\chi \Lambda_L(a)} \geq 1/2 \}
\]

cannot exceed \( 2 \cdot 3^d m \) as each box \( \Lambda_L(a) \) can overlap with at most \( 3^d \) other boxes.

Let \( \mathcal{R} := \bigcup \Lambda_L(a) \), where the union is taken over boxes with \( a \in \mathcal{S} \) and boxes that overlap with them. We note that if \( y \notin \mathcal{R} \), then

\[
\mathbb{P}(y, y) < \theta^4 \tag{B.8}
\]
for $\theta$ sufficiently small. Indeed, if $y \notin \mathcal{R}$, then $\text{dist}(y, \cup_{a \in S} \Lambda_L(a)) \geq L/2$. In particular,

$$P(y, y) \leq P(y, y) \sum_{z \in \Lambda_{L/2}(y)} P(z, z) + \sum_{z \in \Lambda_{L/2}(y)} |P(z, y)|^2 \leq \frac{1}{2} P(y, y) + \theta^4 M,$$

which yields (B.8).

**Lemma B.5.** Let $Q = P\chi_{\mathcal{R}}P$. Then $Q$ is close to $P$, namely $\|P - Q\| \leq \theta^3$ for $\theta$ sufficiently small. In particular, $Q$ is invertible as an operator on $\text{Ran}(P)$, with $Q \geq 1 - \theta^3$.

**Proof.** We have $Q^2 = Q - P\chi_{\mathcal{R}}P\chi_{\mathcal{R}}P$ and

$$\|P\chi_{\mathcal{R}}P\chi_{\mathcal{R}}P\|_{HS} = \sum_{y \in \mathcal{R}, x \in \mathcal{R}} |P(x, y)|^2 = \sum_{0 < \text{dist}(y, \mathcal{R}) \leq L/2, x \in \mathcal{R}} |P(x, y)|^2 + \sum_{\text{dist}(y, \mathcal{R}) > L/2, x \in \mathcal{R}} |P(x, y)|^2. $$

The first term can be estimated by $Cm\theta^4 |\ln \theta|^d \leq \theta^3/2$ using (B.8). For the second sum, we use (B.5) to bound it by $Cm\theta^4 |\ln \theta|^d M < \theta^3/2$. This shows that

$$\|P\chi_{\mathcal{R}}P\chi_{\mathcal{R}}P\|_{HS} \leq \theta^3,$$

so $\|Q^2 - Q\|_{HS} \leq \theta^3$ for $\theta$ sufficiently small.

We next observe that, in view of (B.3),

$$|Q(x, y)| = \left| \sum_{z \in \mathcal{R}} P(z, z)P(z, y) \right| \leq C\theta^{-2} e^{-d|x-y|} \quad (B.10)$$

by the properties of exponential sums. Let $\tilde{Q} = P - Q$. Then $\tilde{Q}$ is (A) close to be a projection on $\text{Ran}(P)$ and (B) $|\tilde{Q}(x, y)| \leq C\theta^{-2} e^{-d|x-y|}$. Indeed, (A) follows from

$$\tilde{Q}^2 = P - 2Q + Q^2 = \tilde{Q} - (Q - Q^2) = \tilde{Q} + O(\theta^3),$$

while (B) follows directly from the decay properties of $P(x, y)$ and $Q(x, y)$.

We next show that $\tilde{Q}$ is in fact close to zero, which implies the result. Indeed, suppose in contradiction that $\tilde{Q}$ is close to a non-trivial projection, i.e., $\text{dist}(\sigma(\tilde{Q}), 1) = O(\theta^3)$. Let $y_o \in \mathbb{Z}^d$ be such that $\bar{M} = \tilde{Q}(y_o, y_o)$ ($y_o$ is not necessary unique). Just as in the proof of Lemma B.3, let $\tilde{r} = \tilde{r}(c, \theta) > 0$ be such that $\sum_{y \in \mathbb{Z}^d : |y| > \tilde{r}} e^{-2|x|} \leq \theta^4 \bar{M}^2$. In particular, there exists $C$ such that we can choose $r = -C \ln (\theta^4 \bar{M})$ for $\theta$ sufficiently small.

Essentially repeating the argument of Lemma B.3 we have

$$\bar{M} = \tilde{Q}(y_o, y_o) = (\tilde{Q} - \tilde{Q}^2)(y_o, y_o) + (\tilde{Q})^2(y_o, y_o) = O(\theta^3) + \sum_{y \in \Lambda^d} |\tilde{Q}(y_o, y)|^2 = O(\theta^3) + \sum_{y \in \Lambda^d} |\tilde{Q}(y_o, y)|^2 + \sum_{y \in \Lambda^d} |\tilde{Q}(y_o, y)|^2 \leq O(\theta^3) + 3d \bar{M}^2 r^d. $$

This yields $\bar{M} \leq 3d \bar{M}^2 r^d$, which in turn yields $\bar{M} \geq u$, where $u$ is implicitly given by the analogue of (B.4). Since $\bar{u} > |\ln \theta|^{-d-1}$ for $\theta$ sufficiently small, we get $\bar{M} \geq |\ln \theta|^{-d-1}$. But then (B.8) implies

$$\theta^4 > P(y_o, y_o) = (P\chi_{\mathcal{R}}P)(y_o, y_o) + (P\chi_{\mathcal{R}}P)(y_o, y_o) \geq (P\chi_{\mathcal{R}}P)(y_o, y_o) = \tilde{Q}(y_o, y_o) > \theta,$$

a contradiction.

\[\square\]

Let $\mathcal{R} = \bigcup_{i=1}^{n} \mathcal{R}_i$ be a partition of $\mathcal{R}$ into connected components. We note that $n \leq 2m$, and that by construction,

$$\text{dist}_{i \neq j}(\mathcal{R}_i, \mathcal{R}_j) \geq L/2 \quad (B.11)$$
We now introduce the operator
\[ X = \sum_{j=1}^{n} j P_{\mathcal{R}_j} P, \quad \text{(B.12)} \]
which acts on Ran\( (P) \). Clearly, \( X \) is hermitian.

**Lemma B.6.** Let \( \lambda \in \sigma(X) \). Then there exists \( j \in \{1, \ldots, n\} \) such that \( |\lambda - j| \leq \theta \) for \( \theta \) sufficiently small.

**Proof.** For any such \( \lambda \), we have
\[ (X - \lambda)^2 = \sum_{j=1}^{n} (j - \lambda)^2 P_{\mathcal{R}_j} P + \sum_{j \neq j'} (j - \lambda)(j' - \lambda) P_{\mathcal{R}_j} P_{\mathcal{R}_{j'}} P. \]
The second sum can be bounded in norm by \( n^2 \theta^3 \) using (B.11) and (B.5), while the first one satisfies
\[ \sum_{j=1}^{n} (j - \lambda)^2 P_{\mathcal{R}_j} P \geq \min_{j} (j - \lambda)^2 Q \geq \min_{j} (j - \lambda)^2 (1 - \theta^2) \]
using Lemma B.6. But \( 0 \in \sigma((X - \lambda)^2) \), from which the result follows. \( \square \)

The assertion of the theorem follows from

**Lemma B.7.** Let \( (\lambda, \psi_\lambda) \) be an eigenpair for \( X \) with normalized \( \psi_\lambda \). Then
\[ |\psi_\lambda(x)| \leq C \theta^{-2} e^{-c \text{dist}(x, \mathcal{R}_{j_0})}, \quad \text{(B.13)} \]
where \( j_0 \) is chosen so that \( |\lambda - j_0| \leq \theta \).

**Proof.** Let
\[ Y_\lambda := P_{\mathcal{R}_{j_0}} P + \sum_{j \neq j_0} (j - \lambda) P_{\mathcal{R}_j} P, \quad Z_\lambda := P_{\mathcal{R}_{j_0}} P + \sum_{j \neq j_0} (j - \lambda)^{-1} P_{\mathcal{R}_j} P. \]
We have
\[ Y_\lambda Z_\lambda = P + \sum_{j \neq j'} f(j, j') (j' - \lambda) P_{\mathcal{R}_j} P_{\mathcal{R}_{j'}} P =: P + W, \]
where \( |f(j, j')| \leq 2n \) for all \( j \neq j' \). We have \( \|W\| \leq n^3 \theta^3 \) using (B.39). Hence by standard perturbation theory, the operator \( Y_\lambda \) is invertible on Ran\( (P) \), with
\[ Y_\lambda^{-1} = Z_\lambda (P + W)^{-1} = Z_\lambda \sum_{i=0}^{\infty} (-W)^i. \quad \text{(B.14)} \]
We now note that, analogously to (B.10),
\[ |Z_\lambda(x, y)| \leq C \theta^{-2} e^{-c|x-y|}, \]
while
\[ |W(x, y)| \leq n^3 \max_{j \neq j'} \sum_{z \in \mathcal{R}_j, w \in \mathcal{R}_{j'}} P(z, x) P(z, w) P(w, y) \leq C n^3 \theta^{-3} e^{-c|x-y|/2} \max_{j \neq j'} \sum_{z \in \mathcal{R}_j, w \in \mathcal{R}_{j'}} e^{-c|z-w|/2} \leq \theta^2 e^{-c|x-y|/2} \]
using (B.11), (B.5), and (B.6). This in turn implies that
\[ |W^i(x, y)| \leq \theta^i e^{-c|x-y|/2}, \quad i \in \mathbb{N}. \]
Using these bounds in (B.14), we deduce that
\[ |Y_\lambda^{-1}(x, y)| \leq C \theta^{-2} e^{-c|x-y|/2}. \]
Hence we have
\[|\psi_\Lambda(x)| = ||\chi_{\{x\}}\psi_\Lambda|| = ||\chi_{\{x\}}Y^{-1}_\Lambda\psi_\Lambda|| = ||\chi_{\{x\}}Y^{-1}_\Lambda(Y_\Lambda - X + \lambda)\psi_\Lambda|| = |1 - j_0 + \lambda||\chi_{\{x\}}Y^{-1}_\Lambda P_{|R|} P\psi_\Lambda|| \leq C\theta^{-2}e^{-c\text{dist}(x,R_{j_0})}.
\]

We are now ready to complete the proof of Theorem [B.2]. We pick the set \(\{\psi_\lambda\}_{\lambda \in \sigma(X)}\), which is an orthonormal basis for \(\text{Ran}(P)\) since \(X\) is hermitian. Since
\[
\max_j \text{diam}(R_j) \leq 2mL = -mC\ln \theta,
\]
picking some \(x_j \in R_j\), we have
\[e^{-c\text{dist}(x,R_{j_0})} \leq e^{-c|x-x_j|-2mL} \leq e^{-c|x-x_j|/m} \quad \text{for } |x-x_j| \geq 3mL.
\]
On the other hand, since \(|\psi(x)| \leq 1\) for all \(x\), we can pick \(c\) sufficiently small so that
\[e^{-c|x-x_j|/m} \geq \theta^2 \quad \text{for } |x-x_j| < 3mL,
\]
and the assertion follows.

**Appendix C. Auxiliary results**

**Lemma C.1.** Let \(H = -\Delta + V_\omega\) be the random operator on \(\ell^2(Z)\) with \(V_\omega\) that satisfies assumptions introduced in Appendix [A]. Let \(J = [\frac{1}{8}, \frac{15}{8}]\) and \(c = \frac{1}{12}\). Then
\[
\sum_{E \in \sigma(H) \cap J} |\psi_E(y)|^2 \geq c, \quad y \in Z, \quad (C.1)
\]
and the same bound holds for any Dirichlet restriction \(H^\Lambda\) of \(H\).

**Proof.** Let \(P_J := P_J(H)\). Suppose in contradiction that \(\text{tr}\chi_{\{y\}}P_J < c\) for some \(y \in Z\). Then we have
\[\text{tr}\chi_{\{y\}}(H - 2)^2 \geq \text{tr}\chi_{\{y\}}(H - 2)^2 P_J \geq \frac{49}{16} \text{tr}\chi_{\{y\}}P_J > 3.
\]
However, the left hand side can be computed explicitly: \(\text{tr}\chi_{\{y\}}(H - 2)^2 = 2 + V_\omega^2(y) \leq 3\), a contradiction. The proof for \(H^\Lambda\) is identical.

**Theorem C.2.** Assume that \(H\) is \((\nu, \theta)\)-localized on \(Z\) and that there exists \(c > 0\) and a compact interval \(J\) such that
\[
\sum_{E \in \sigma(H) \cap J} |\psi_E(y)|^2 \geq c, \quad y \in Z. \quad (C.2)
\]

Then there exists \(C_\nu > 0\) and \(E \in \sigma(H) \cap J\) such that \(|\psi_E(0)|^2 \geq \frac{C_\nu}{\ln \theta}\) and \(|x_E| \leq \frac{\ln \theta}{C_\nu}\). The same result holds for \(H\) replaced by the finite volume Hamiltonian \(H^\Lambda\), provided that \(|\Lambda|\) is sufficiently large, namely \(|\Lambda| \gg |\ln \theta|\).

**Proof.** We first observe that for any \(L \in \mathbb{N}\) and \(E \in \sigma(H)\) we have
\[
\sum_{y \in Z: |y-x_E| \geq \frac{1}{2}(|x_E|+L)} |\psi_E(y)|^2 \leq \frac{\langle x_E \rangle^2}{\theta} \cdot \sum_{y \in Z: |y-x_E| \geq \frac{1}{2}(|x_E|+L)} e^{-\nu|y-x_E|} = \frac{\langle x_E \rangle^2}{\theta} \cdot \sum_{y \in Z: |u| \geq \frac{1}{2}(|x_E|+L)} e^{-\nu|u|} = \frac{\langle x_E \rangle^2}{\theta} \cdot \frac{2}{1-e^{-\nu}} \leq \frac{C_\nu}{\theta} e^{-\frac{\nu}{2}(L+|x_E|)} \quad (C.3)
\]
for some \(C_\nu > 0\).
We next note that by the orthonormality of \( \{ \psi_E \} \) we have
\[
\sum_{y \in \mathcal{Y}} |\psi_E(y)|^2 = 1, \quad E \in \sigma(H).
\] (C.4)

Hence, using (C.2) and (C.3), there exists \( K_\nu > 0 \) such that
\[
4L + 1 \geq \sum_{|y| \leq 2L} \sum_{E \in \sigma(H) \cap J} |\psi_E(y)|^2 \geq \sum_{|y| \leq 2L} \sum_{E \in \sigma(H)} |\psi_E(y)|^2 = \sum_{E \in \sigma(H) \cap J: |x_E| \leq L} \left( 1 - \frac{C_\nu}{\theta} e^{-\frac{|y|}{2L}} \right) \geq \frac{1}{2} \# \{ E \in \sigma(H) \cap J : |x_E| \leq L \}
\] (C.5)

for \( L \geq K_\nu |\ln \theta| \).

This bound together with (C.3) imply that for \( L \geq K_\nu |\ln \theta| \) we have
\[
\sum_{|y| \leq L} \sum_{E \in \sigma(H) \cap J: |x_E| \geq 3L} |\psi_E(y)|^2 \leq \sum_{k=4}^{\infty} \# \{ E \in \sigma(H) \cap J : |x_E| \leq kL \} \frac{C_\nu e^{-\frac{k|y|}{2L}}}{\theta} \leq \frac{9C_\nu}{\theta} L \sum_{k=4}^{\infty} ke^{-\frac{k|y|}{2L}} < \frac{c}{2}
\] (C.6)

for \( L \geq M_\nu |\ln \theta| \) with some \( M_\nu > 0 \).

Using this estimate, we get
\[
c \leq \sum_{E \in \sigma(H) \cap J} |\psi_E(0)|^2 \leq \sum_{E \in \sigma(H) \cap J: |x_E| \leq 3L} |\psi_E(0)|^2 + \frac{c}{2}
\]
for \( L \geq M_\nu |\ln \theta| \), so
\[
c \leq \sum_{E \in \sigma(H) \cap J: |x_E| \leq 3L} |\psi_E(0)|^2,
\]
and since \( \# \{ E \in \sigma(H) : |x_E| \leq 3L \} \leq 13L \) by (C.5), we deduce that there exists \( C_\nu > 0 \) and \( E \in \sigma(H) \cap J \) such that
\[
|\psi_E(0)|^2 \geq \frac{c}{26L} = \frac{C_\nu}{\ln \theta}, \quad |x_E| \leq \frac{-\ln \theta}{C_\nu}.
\]
\[
\square
\]

Let \( H \) be a self-adjoint operator. Here we will often use the integral representation
\[
P_{[E_1, E_2]}(H) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j (H - ix - E_j)^{-1} dx,
\] (C.7)
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
\[
d \frac{d}{ds} (H(s) - ix - E_j)^{-1} = -(H(s) - ix - E_j)^{-1} \dot{H}(s) (H(s) - ix - E_j)^{-1}
\] (C.8)
holds. Furthermore, for any operator \( R \), we have
\[
[R, \frac{1}{H - z}] = -\frac{1}{H - z} [R, H] \frac{1}{H - z}.
\] (C.9)

**Lemma C.3.** Let \( H_1, H_2, R \) be bounded operators on \( l^2(\Lambda) \), with \( H_1, H_2 \) self-adjoint. Let \( J = [E_1, E_2] \) and denote by \( J_{2\Delta} \) for \( \Delta > 0 \), the widened interval \( J + [-2\Delta, 2\Delta] \). Suppose that for some \( \epsilon_1, \epsilon_2 \),
It follows that
\[(i) \, \|(H_1 - H_2) R\| = \epsilon_1 \]
\[(ii) \, \|(H_2, R) P_j(H_2)\| \leq \epsilon_2.\]

Then
\[
\left\| \bar{P}_{\Delta}(H_1) R P_j(H_2) \right\| \leq \frac{\epsilon_1 + \epsilon_2}{\Delta}.
\]

Proof. Let \( z_1 = E_1 - \Delta + ix, \ z_2 = E_2 + \Delta + ix \) and write
\[
G_{i,j} = (H_1 - z_j)^{-1}.
\]

We first establish the identity
\[
\bar{P}_{\Delta}(H_1) R P_j(H_2) = \frac{1}{2\pi} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} \bar{P}_{\Delta}(H_1) G_{i,j}(H_2) P_j(H_2) dx
\]
\[
+ \frac{1}{2\pi} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} \bar{P}_{\Delta}(H_1) G_{i,j}(H_2 - H_1) R G_{i,j}(H_2) P_j(H_2) dx.
\]

Indeed, we start from
\[
G_{i,j}(H_2, R) G_{i,j} = G_{i,j}(H_2 - H_1) R G_{i,j} + R G_{i,j} + G_{i,j} R.
\]

Upon multiplying with \((-1)^j\), summing over \( j = 1, 2 \), integrating over \( x \), and using \([C.7] \) with \([E_1, E_2] \) replaced by \([E_1 - \Delta, E_2 + \Delta] \), we get the desired identity. We next bound
\[
\max_{j=1,2} \| \bar{P}_{\Delta}(H_1) G_{i,j} \| \leq \frac{1}{\sqrt{x^2 + \Delta^2}}, \quad \max_{j=1,2} \| G_{i,j} P_j(H_2) \| \leq \frac{1}{\sqrt{x^2 + \Delta^2}}
\]
to get
\[
\left\| \bar{P}_{\Delta}(H_1) R P_j(H_2) \right\| \leq (\epsilon_1 + \epsilon_2) \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dx}{x^2 + \Delta^2} = \frac{\epsilon_1 + \epsilon_2}{\Delta}.
\]

\( \square \)

For the next lemma, we will use the notation \( J_\alpha(\mu) = [\mu - a, \mu + a] \), and will let \( P_{J_\alpha(\mu)}^\Theta \) denote the spectral projection of \( H_\alpha^\Theta \) onto \( J_\alpha(\mu) \).

Lemma C.4. Let \( \Phi \) and \( \Theta \), with \( \Phi \subset \Theta \), be finite subsets of \( \mathbb{Z}^d \). Let \((\phi, \mu)\) be an eigenpair for \( H_\alpha^\Theta \). Then we have
\[
\dist (\mu, \sigma(H_\alpha^\Theta)) \leq C |\partial_\alpha \Phi| \|\chi_{\partial_\alpha \Phi} \phi\|_{\infty}, \quad (C.10)
\]
and
\[
\dist \left( \phi, \text{Ran} \left( P_{J_\alpha(\mu)}^\Theta \right) \right) \leq \frac{C}{a} |\partial_\alpha \Phi| \|\chi_{\partial_\alpha \Phi} \phi\|_{\infty}. \quad (C.11)
\]

Conversely, if \((\psi, \lambda)\) is an eigenpair for \( H_\Theta^\Theta \), then
\[
\dist (\lambda, \sigma(H_\alpha^\Theta)) \leq C |\Theta \setminus \Phi| \|\chi_\Theta \psi\|_{\infty} \quad (C.12)
\]
and
\[
\dist \left( \phi, \text{Ran} \left( P_{J_\alpha(\lambda)}^\Theta \right) \right) \leq \frac{C}{a} |\Theta \setminus \Phi| \|\chi_\Theta \psi\|_{\infty}. \quad (C.13)
\]

Proof. We have
\[
((H_\alpha^\Theta - \mu) \phi) (y) = \begin{cases} \sum_{y' \in \Phi : \ |y - y'| \leq r} H_\alpha(y, y')\phi(y') & \text{if } y \in \Theta \setminus \Phi \text{ and } \dist (y, \Phi) \leq r, \\ 0 & \text{otherwise}. \end{cases} \quad (C.14)
\]

It follows that
\[
\left\| (H_\alpha^\Theta - \mu) \phi \right\| \leq C |\partial_\alpha \Phi| \|\chi_{\partial_\alpha \Phi} \phi\|_{\infty}. \quad (C.15)
\]

Thus, recalling that \( \phi \) is normalized,
\[
\dist (\mu, \sigma(H_\alpha^\Theta)) \leq \left\| (H_\alpha^\Theta - \mu) \phi \right\| \leq C |\partial_\alpha \Phi| \|\chi_{\partial_\alpha \Phi} \phi\|_{\infty}. \quad (C.16)
\]
On the other hand, we have
\[
\left\| \tilde{P}_s J_a(\mu) \phi \right\| \leq \left\| \tilde{P}_s J_a(\mu) (H_o^\Theta - \mu)^{-1} \right\| \left\| (H_o^\Theta - \mu) \phi \right\| \leq \frac{C}{a} \left\| \chi_{\Theta \setminus \Phi} \psi \right\|_\infty, \tag{C.17}
\]
from which the second assertion of the lemma follows.

Similar considerations yield
\[
\left\| (H_o^\Phi - \lambda) \phi \right\| \leq C \left\| \Theta \setminus \Phi \right\| \left\| \chi_{\Theta \setminus \Phi} \phi \right\|_\infty, \tag{C.18}
\]
which in turn imply the bounds (C.12)–(C.13).

\(\Box\)

In this paper we are interested in the evolution of the initial wave packet \(\psi_o\) supported near some \(x \in \mathbb{Z}^d\) up to the (rescaled) time \(s\) of order 1. In this context, we can always approximate the dynamics generated by \(H(s)\) with the one generated by \(\hat{H}^T(s)\), where \(H^T(s)\) is understood as an operator on \(L^2(\mathbb{Z}^d)\) (extending it by zero outside of the box \(\Lambda_L\)), in the following sense.

**Proposition C.5 (The finite speed of propagation bound).** Let \(T\) be a torus of linear size \(R\) and let \(U_t(s), U^T_t(s)\) be the dynamics generated by \(H(s)\) and \(H^T(s)\), respectively, i.e.,
\[
\begin{align*}
&ic\partial_s U_t(s) = H(s)U_t(s), \quad U_t(0) = 1; \tag{C.19} \\
&ic\partial_s U^T_t(s) = H^T(s)U^T_t(s), \quad U^T_t(0) = 1. \tag{C.20}
\end{align*}
\]

Then there exists \(c > 0\) such that for any \(\mathcal{L}\) satisfying \(\mathcal{L} \geq C/\epsilon\) we have
\[
\max_s \left| (U^T_t(s))(y,x) \right| \leq e^{-c|x-y|}, \quad \text{for } |x - y| \geq \frac{\mathcal{L}}{4}, \tag{C.21}
\]
where \(U^T_t\) is either \(U\) or \(U^T\).

**Proof.** This is a standard fact for (local) lattice Hamiltonians, see e.g., the proof of [19, Lemma 5] for the time-independent case (which extends to the time-dependent one without effort), or, for a more general approach, [40]. \(\Box\)

Acknowledgment. We are grateful to Gian Michele Graf for helpful discussions.

**References**

[1] D. A. Abanin, W. De Roeck, and F. Huveneers. “Theory of many-body localization in periodically driven systems”. *Ann. Phys.* 372 (2016), 1–11.

[2] W. K. Abou-Salem and J. Fröhlich. “Adiabatic theorems and reversible isothermal processes”. *Lett. Math. Phys.* 72 (2005), 153–163.

[3] M. Aizenman and G. M. Graf. “Localization bounds for an electron gas”. *J. Phys. A* 31 (1998), 6783.

[4] M. Aizenman and S. Warzel. “Localization bounds for multiparticle systems”. *Comm. Math. Phys.* 290 (2009), 903–934.

[5] J. Avron, J. Howland, and B Simon. “Adiabatic theorems for dense point spectra”. *Comm. Math. Phys.* 128 (1990), 497–507.

[6] J. E. Avron and A. Elgart. “Adiabatic theorem without a gap condition”. *Comm. Math. Phys.* 203 (1999), 445–463.

[7] J. E. Avron, R. Seiler, and B. Simon. “Charge deficiency, charge transport and comparison of dimensions”. *Comm. Math. Phys.* 159 (1994), 399–422.

[8] S. Bachmann, W. De Roeck, and M. Fraas. “The adiabatic theorem and linear response theory for extended quantum systems”. *Comm. Math. Phys.* 361 (2018), 997–1027.

[9] R. Barequet, G. Barequet, and G. Rote. “Formulae and growth rates of high-dimensional polycubes”. *Combinatorica* 30 (2010), 257–275.

[10] M. Born and V. Fock. “Beweis des adiabatsatzes”. *Z. Phys.* 51 (1928), 165–180.

[11] F. Bornemann. *Homogenization in time of singularly perturbed mechanical systems*. Springer, 2006.
[12] J.-M. Bouclet, F. Germinet, A. Klein, and J. H. Schenker. “Linear response theory for magnetic Schrödinger operators in disordered media”. *J. Funct. Anal.* **226** (2005), 301–372.

[13] J. Bourgain and W.-M. Wang. “Anderson localization for time quasi-periodic random Schrödinger and wave equations”. *Comm. Math. Phys.* **248** (2004), 429–466.

[14] R. Carmona and J. Lacroix. *Spectral theory of random Schrödinger operators*. Springer Science & Business Media, 2012.

[15] J.-M. Combes, F. Germinet, and A. Klein. “Generalized eigenvalue-counting estimates for the Anderson model”. *J. Stat. Phys.* **135** (2009), 201–216.

[16] R del Rio, N Makarov, and B Simon. “Operators with singular continuous spectrum: II. Rank one operators”. *Comm. Math. Phys.* **165** (1994), 59–67.

[17] A. Dietlein and A. Elgart. “Level spacing and Poisson statistics for continuum random Schrödinger operators”. *J. Eur. Math. Soc.* **23** (2021), 1257–1293.

[18] R. Ducatez and F. Huveneers. “Anderson localization for periodically driven systems”. *Ann. Henri Poincaré* **18** (2017), 2415–2446.

[19] A Elgart, G. Graf, and J. Schenker. “Equality of the bulk and edge Hall conductances in a mobility gap”. *Comm. Math. Phys.* **259** (2005), 185–221.

[20] A. Elgart and G. A. Hagedorn. “An adiabatic theorem for resonances”. *Comm. Pure Appl. Math.* **64** (2011), 1029–1058.

[21] A. Elgart and A. Klein. “An eigensystem approach to Anderson localization”. *J. Funct. Anal.* **271** (2016), 3465–3512.

[22] A. Elgart, L. Pastur, and M. Shcherbina. “Large block properties of the entanglement entropy of free disordered fermions”. *J. Stat. Phys.* **166** (2017), 1092–1127.

[23] A. Elgart and J. H. Schenker. “A strong operator topology adiabatic theorem”. *Rev. Math. Phys.* **14** (2002), 569–584.

[24] A. Elgart, M. Shamis, and S. Sodin. “Localisation for non-monotone Schrödinger operators”. *J. Eur. Math. Soc.* **16** (2014), 909–924.

[25] A. Elgart, M. Tautenhahn, and I. Veselić. “Anderson localization for a class of models with a sign-indefinite single-site potential via fractional moment method”. *Ann. Henri Poincaré* **12** (2011), 1571–1599.

[26] M. Gebert. “A lower Wegner estimate and bounds on the spectral shift function for continuum random Schrödinger operators”. *J. Funct. Anal.* **277** (2019), 108284.

[27] A. Y. Gordon. “Pure point spectrum under 1-parameter perturbations and instability of Anderson localization”. *Comm. Math. Phys.* **164** (1994), 489–505.

[28] M. S. Green. “Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids”. *J. Chem. Phys.* **22** (1954), 398–413.

[29] J. Henheik and S. Teufel. “Adiabatic theorem in the thermodynamic limit: Systems with a gap in the bulk”. *arXiv preprint arXiv:2012.15239* (2020).

[30] J. Henheik and S. Teufel. “Justifying Kubo’s formula for gapped systems at zero temperature: A brief review and some new results”. *Rev. Math. Phys.* **33** (2021), 2060004.

[31] P. D. Hislop and M Krishna. “Eigenvalue statistics for random Schrödinger operators with non rank one perturbations”. *Comm. Math. Phys.* **340** (2015), 125–143.

[32] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[33] I. Kachkovskiy and Y. Safarov. “Distance to normal elements in $C^*$-algebras of real rank zero”. *J. Amer. Math. Soc.* **29** (2016), 61–80.

[34] T. Kato. “On the adiabatic theorem of quantum mechanics”. *J. Phys. Soc. Japan* **5** (1950), 435–439.

[35] T. Kato. *Perturbation theory for linear operators*. **132**. Springer Science & Business Media, 2013.

[36] A. Klein, O. Lenoble, and P. Müller. “On Mott’s formula for the ac-conductivity in the Anderson model”. *Annals of Math.* (2007), 549–577.

[37] A. Klein and S. Molchanov. “Simplicity of eigenvalues in the Anderson model”. *J. Stat. Phys.* **122** (2006), 95–99.
[38] F. Klopp and J. Schenker. “On the spatial extent of localized eigenfunctions for random Schrödinger operators”. arXiv preprint arXiv:2105.13215 (2021).

[39] R. Kubo. “Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems”. J. Phys. Soc. Japan 12 (1957), 570–586.

[40] E. H. Lieb and D. W. Robinson. “The finite group velocity of quantum spin systems”. Comm. Math. Phys. (1972), 251–257.

[41] G. Marcelli, M. Moscolari, and G. Panati. “Localization implies Chern triviality in non-periodic insulators”. arXiv preprint arXiv:2012.14407 (2020).

[42] D. Monaco and S. Teufel. “Adiabatic currents for interacting fermions on a lattice”. Rev. Math. Phys. 31 (2019), 1950009.

[43] F. Nakano and M. Kaminaga. “Absence of transport under a slowly varying potential in disordered systems”. J. Stat. Phys. 97 (1999), 917–940.

[44] G. Nenciu. “Linear adiabatic theory. Exponential estimates”. Comm. Math. Phys. 152 (1993), 479–496.

[45] G. Panati, H. Spohn, and S. Teufel. “Effective dynamics for Bloch electrons: Peierls substitution and beyond”. Comm. Math. Phys. 242 (2003), 547–578.

[46] M. Reed and B. Simon. Methods of modern mathematical physics: Functional analysis. Elsevier, 2012.

[47] B. Simon. “Cyclic vectors in the Anderson model”. Rev. Math. Phys. 6 (1994), 1183–1185.

[48] A. Soffer and W.-M. Wang. “Anderson localization for time periodic random Schrödinger operators”. Comm. in PDE 28 (2003), 333–347.

[49] L Tenuta and S Teufel. “Effective dynamics for particles coupled to a quantized scalar field”. Comm. Math. Phys. 280 (2008), 751–805.

[50] M. Wilkinson. “Statistical aspects of dissipation by Landau-Zener transitions”. J. Phys. A 21 (1988), 4021.

[51] F. Zhang. The Schur complement and its applications. 4. Springer Science & Business Media, 2006.