Inventiones mathematicae
https://doi.org/10.1007/s00222-023-01227-z

Derivation of Kubo's formula for disordered systems at zero temperature

Wojciech De Roeck[1] · Alexander Elgart[2] · Martin Fraas[3]

In memoriam: Rachel Vaiman

Received: 10 March 2022 / Accepted: 6 November 2023

Abstract
This work justifies the linear response formula for the Hall conductance of a two-dimensional disordered system. The proof rests on controlling the dynamics associated with a random time-dependent Hamiltonian.

The principal challenge is related to the fact that spectral and dynamical localization are intrinsically unstable under perturbation, and the exact spectral flow - the tool used previously to control the dynamics in this context - does not exist. We resolve this problem by proving a local adiabatic theorem: With high probability, the physical evolution of a localized eigenstate ψ associated with a random system remains close to the spectral flow for a restriction of the instantaneous Hamiltonian to a region R where the bulk of ψ is supported. Allowing R to grow at most logarithmically in time ensures that the deviation of the physical evolution from this spectral flow is small.

To substantiate our claim on the failure of the global spectral flow in disordered systems, we prove eigenvector hybridization in a one-dimensional Anderson model at all scales.

Mathematics Subject Classification (2000) 82B44 · 35Q41 · 82C44 · 82C70

✉ A. Elgart
aelgart@vt.edu

W. De Roeck
wojciech.deroeck@kuleuven.be

M. Fraas
mfraas@ucdavis.edu

1 Instituut Theoretische Fysica, KU Leuven, 3001, Leuven, Belgium

2 Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA

3 Department of Mathematics, Davis, CA 95616, USA

Published online: 28 November 2023

1 Introduction

In this work we examine the response of a disordered quantum system, described by a random self-adjoint operator H, to a weak time-dependent external perturbation $W(t)$, with the interaction strength modulated by the parameter β. This produces a family of self-adjoint operators

$$H(t) = H + \beta W(t), \quad t \in \mathbb{R}. \tag{1.1}$$

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

This article provides a microscopic derivation of the Kubo formula for Hall conductance, a problem that arises in theoretical condensed matter physics and pertains to the dynamics generated by $H(t)$. It lies in the intersection of two broader problems in mathematical physics: microscopic justification of linear response theory and justification of quantization of Hall conductance.

1.1 Quantum Hall effect

In the early 1980s, von Klitzing and his collaborators [46] made a remarkable discovery: At low temperatures, the Hall conductance for the $2D$ electron gas in a strong magnetic field was found to be a staircase-like function of the electron density. The plateaus take values in $\mathbb{Z} \times q^2/h$ with such incredible precision (one part in a billion) that this effect is used in the metrological definitions of the kilogram and the ampere. Further experimentation revealed that the stairs vanish in very clean samples, strongly indicating that the effect requires disorder. To comprehend the effect, the physical and mathematical theory thus has to address three fundamental questions:

(i) Why is the Hall conductance quantized in the units of q^2/h?
(ii) What is the role of the disorder?
(iii) What explains the precision of this quantization?

We first discuss Question (ii). Many aspects of Hall conductance can be encapsulated by translation-invariant magnetic Hamiltonians, characterized by bands of the absolutely continuous spectrum separated by the spectral gaps. The conductance in such models is quantized when the *Fermi energy* E_F falls into the spectral gap, and transitions to a different value as E_F crosses a conducting band. In what follows, we refer to this intensively-studied class of models as the *disorder-free case*. However, the critical feature of QHE that cannot be explained within such a framework is the existence of plateaus, as the electron density remains constant within the spectral gap. An appropriate Hamiltonian modeling this aspect of the effect must instead have a spectrum consisting of interlacing intervals of conducting and insulating bands, with quantized conductance for the values of E_F that lie in an insulating band. The role of disorder is precisely to create such a structure. The physics community universally accepts that a suitable H, namely a random magnetic Schrödinger operator, is the correct operator to describe this phenomenon. One of the long-standing open problems in mathematical physics is proving that the spectrum of H consists of intervals of alternating absolutely continuous (conducting) and dense pure point (insulating) spectra. The only progress in this direction, namely the proof that the spectrum cannot be entirely pure point, has been made using the topological structure associated with the plateaus in QHE, [32], which brings us back to the first question.

The mechanism explaining Question (i) above was suggested shortly after the discovery of QHE and is associated with the Kubo formula σ_H for the Hall conductance, which was proven to be a topological invariant. In the disorder-free case, σ_H is linked to a Chern number of the ground state bundle whenever E_F lies in the spectral gap. This is now well understood both in the absence [6, 65] and presence [5, 11, 33, 37, 56] of interactions between the electrons. For disordered systems, σ_H has been linked to a Fredholm index using both non-commutative geometrical [15] and analytical [9] methods. The microscopic derivation of the Fredholm index for an Anderson-type Hamiltonian assuming the Kubo formula and that E_F lies in the dense point spectrum was first supplied in [2].

The theory associated with Question (iii) aims to justify the Kubo formula for conductance when the Fermi energy is in the insulating band. The Kubo formula is a standard expression for conductances, or more broadly for response coefficients, obtained by a formal first-order perturbation theory in the strength of a driving field β. To explain the precision, the theory must validate the formal calculation and demonstrate that all higher-order terms in β vanish. In the disorder-free case, this was achieved for non-interacting [7, 26] and interacting [12, 13, 53, 64] models. This work establishes the microscopic proof of this formula for disordered systems.

1.2 Linear response theory

LRT explores the behavior of macroscopic variables in response to small perturbations. In the field of condensed matter physics, it serves as an essential and versatile tool with numerous variants applicable to a wide range of physical variables and models. To ground the discussion in the application we have in mind, we will discuss the response of the current **J** to an electric field **E** with a finite voltage **V** applied across the system in a given direction. *Ohm's law* states that for small **V** the current is proportional to the voltage,

$$\mathbf{J} = \sigma \mathbf{V},$$

where the constant of proportionality is called conductance. The purpose of LRT is to provide a microscopic expression for σ.

LRT was first developed by Kubo, [48]. The expressions for σ corresponding to nonzero and zero temperatures are known as the Green-Kubo and Kubo-Středa formulas, accordingly, [35, 62]. In this work we consider the latter case. LRT has a wide range of settings, [52]; we have chosen one guided by simplicity and convenience.

The theory computes the response from a time-dependent Hamiltonian model of the form (1.1). In the context of electrical conductance, $W(t) = e^t V(x)$, where $V(x)$ is an electric potential of unit voltage. At $t = -\infty$, the system is initiated in an equilibrium state ρ of the unperturbed Hamiltonian H and then evolves according to the Heisenberg equation

$$\dot{\rho}_t = -i[H(t), \rho_t], \quad H(t) = H + \beta e^{\epsilon t} V(x) \tag{1.2}$$

with the adiabatic parameter ϵ. The expected value of the measured current at $t = 0$ is $\mathbf{J} = \mathrm{tr}(\rho_0 J)$, where J is the current operator, and the measured conductance is

$$\sigma_m(\epsilon, \beta) = \beta^{-1}\mathrm{tr}(\rho_0 J).$$

In a typical experiment that measures conductance, the time scales involved are such that both ϵ and β are small parameters. However, ϵ is significantly smaller than β by several orders of magnitude. For a standard experimental setup $\epsilon/\beta < 10^{-9}$ (based on experimental time longer than 1 milisecond and electric potential greater in magnitude than $10^{-3}V$; [66] estimates that linear approximation in β would be justified only for electric fields of order $10^{-16}\frac{V}{m}$). This relationship between timescales ensures that the system will produce a *non-trivial* steady current. On the other hand, the Kubo formula σ_H for conductance is obtained by taking the limit $\beta \ll \epsilon$,

$$\sigma_H = \lim_{\epsilon\to 0}\lim_{\beta\to 0}\sigma_m(\beta, \epsilon) = \lim_{\epsilon\to 0} i\int_{-\infty}^{0} e^{\epsilon t}\mathrm{tr}\left(\rho[e^{iHt}Je^{-iHt}, V]\right)dt, \tag{1.3}$$

and only depends on the spectral data for the unperturbed Hamiltonian H. Nevertheless, the formula is spectacularly successful in matching available experimental data. This raises the question of how the Kubo formula not only works at all in this context but also predicts the experimentally observed conductance with astonishing precision. The *problem of linear response* is to either prove that the joint limit

$$\lim_{\epsilon\ll\beta\to 0}\sigma_m(\beta, \epsilon)$$

exists and is equal to σ_H, or to provide an alternative explanation for the validity of expression (1.3).

Although the focus of our attention is on the response of the current to the electric field, i.e., Ohm's law, the same question can be posed for Fourier's law, Fick's law, and other phenomena. The justifications of the Kubo formula for these various physics laws are long-standing open problems in mathematical physics, each posing a unique mathematical challenge, see, e.g., [59, Problem 4B]. Our work provides the first proof of the Kubo formula in a disordered system.

1.3 Microscopic derivation of the Kubo formula for Hall conductance

The Hall conductance σ_H is defined in 2D as the proportionality constant between the applied potential difference and the current flowing in the perpendicular direction. In what follows, we make a specific choice for the applied electric potential $V(x)$ and the current operator J. We will assume that the Fermi energy E_F lies in the mobility gap for H, where the latter concept will be formally defined in Sect. 2.1.2.

We denote by (x_1, x_2) the coordinates of points in $\mathbb{Z}^2$ and by Λ_n the characteristic function of the subset $\{x_n \geq 0\}$, $n = 1, 2$. These functions are examples of so-called *switches*, i.e., functions h of one variable that are real valued, monotone, and non-decreasing, with $h(-\infty) = 0$ and $h(\infty) = 1$.

We consider an electric potential $V = \Lambda_2$, which has a unit voltage drop across the x_2 direction. The (Hall) current flowing in the perpendicular direction across the fiducial line $x_1 = 0$ corresponds to the operator $J = i[H, \Lambda_1]$. The equilibrium state ρ is given by the Fermi projection $P_F := \chi_{<E_F}(H)$. The Kubo-Středa formula (1.3) is then given by

$$\sigma_H = \mathrm{tr}(P_F[[P_F, \Lambda_1], [P_F, \Lambda_2]]), \tag{1.4}$$

see e.g. [2]. We make two changes to the linear response setup explained above. We replace e^t by a compactly supported switch g, and average the current over a time window of order ϵ^{-1}. More specifically, we consider a Hamiltonian of a form

$$H(t) = H + \beta g(\epsilon t)\Lambda_2,$$

where the function g satisfies

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

We (re)define the measured conductance as

$$\sigma_m(\beta, \epsilon) := \beta^{-1}\epsilon \int_0^{1/\epsilon} \mathrm{tr}\,(J(\rho_t - \rho))\,dt. \tag{1.5}$$

There are no equilibrium currents [10], i.e., $\mathrm{tr}(J\rho) = 0$ when the trace is properly defined. However, in infinite volume $J\rho_t$ is not a trace class operator and subtracting $J\rho$ is a physically correct way to regularize it. We stress again that our goal is to understand the behavior of $\sigma_m(\beta, \epsilon)$ for $\epsilon \ll \beta \to 0$.

Our main result on the problem of linear response establishes the existence of the joint limit under the constraint $\epsilon = e^{-\beta^{-p}}$ with the positive exponent p.

Theorem 1.1 *Suppose that H satisfies Assumptions 2.3–2.4 below with E_F lying in the interior of a mobility gap. Then there exist $p > 0$ such that*

$$\mathbb{E}\,|\sigma_H - \sigma_m| \le e^{-\beta^{-p/2}},$$

provided $\epsilon = e^{-\beta^{-p}}$.

Remark 1.2

(i) The use of a compactly supported switch function $g(t)$ instead of the exponential is a natural choice from a mathematical point of view. That being said, Theorem 1.1 could also be established for $g(t) = e^t$.
(ii) Some form of the current averaging is likely needed for the result to hold. We did not try to minimize the size of the time window over which the average is performed.
(iii) The choice of profiles for switches Λ_i and g does not affect the result. This is related to the fact that the expression for σ_H is universal in the sense that the value of σ_H (almost surely) does not change upon modifying the switches or changing E_F within the same interval J_{loc}, see [27].

(iv) One can also study conductivity instead of conductance, where the switch functions Λ_i are replaced by the linear relations $X_i(x) = x_i$ and the trace in (1.4)–(1.5) is replaced by the trace per unit volume. While working with conductivity simplifies some of the analysis (e.g., one no longer needs to regularize $\mathrm{tr}(J\rho_t)$ in (1.5)), it also offers different technical challenges (e.g., the corresponding Hamiltonian $H(t)$ is no longer bounded and even if H has spectral gaps, they close for $H(t)$). In particular, even the justification of the Kubo formula for conductivity when the limit $\beta \to 0$ is taken first requires non-trivial effort for disordered systems, [18]. We refer the reader to [38, 53] for state-of-the-art articles on the conductivity approach in the disorder-free case. It would be interesting to see whether the techniques developed in our work can also be extended to handle this choice.

(v) Using Theorem 1.1, we can bound the finite temperature corrections to σ_m by $\frac{1}{\epsilon}e^{-d_\mu/T}$, where T is the absolute temperature, μ is the chemical potential, and d_μ is a distance from μ to the boundary of the insulating band. Let us mention that the finite temperature correction has been recently addressed for the gapped systems in the many-body context [36].

The majority of the mathematical work related to the Kubo formula in disordered systems, with or without an application to QHE, falls into two categories: In the first one, the Kubo formula is taken for granted (or at least the order of limits $\beta \ll \epsilon$ is assumed) and its various consequences in different settings, such as the mathematical proof of Mott's formula, [45], are studied. The second category aims to justify the Kubo formula itself with the correct order of limits. Since our work lies firmly in the second category, we primarily focus our attention on past works in this direction. For a recent review of efforts pertaining to both categories, we refer the reader to [38].

The Kubo formula has been validated in systems with a *spectral gap* ($\mathrm{dist}(\sigma(H), E_F) > 0$), under various sets of assumptions on H and the underlying geometry, [7, 12, 13, 26, 53, 64]. In this scenario, the weak field $\beta \to 0$ and adiabatic $\epsilon \to 0$ limits commute. On the technical level, this can be linked with the stability of the spectral gaps under small perturbations (i.e., $\mathrm{dist}\,(\sigma(H(t)), E_F) > 0$ holds), ensuring that the adiabatic theorem of quantum mechanics could be used. The latter implies that ρ_t is the zero temperature equilibrium state of $H(t)$ up to uniformly small corrections of order ϵ. However, in the disordered case, there is no spectral gap to begin with, and the pure point spectrum associated with a mobility gap is unstable under small perturbations, [22]. Consequently, in this scenario, the limits are not expected to coincide on physical grounds [66]. In this sense, the result presented above with the joint limit is optimal.

The prior mathematical results in this direction for disordered systems are scarce. As previously mentioned, [18] established the existence of the limit $\beta \to 0$ at fixed ϵ. For $\epsilon \to 0$, the only available result, namely the absence of transport, $\sigma_m = o(1)$, was proven in the case $\beta = \epsilon$ in [54] under the assumption of *complete localization* (i.e., there are no conducting bands). Under this assumption, the dynamics of the perturbed system can be controlled for long timescales using the one associated with the unperturbed operator H, e.g., [1, 19, 24, 54, 61]. Beyond this, despite general interest in the mathematical physics community from the moment that the problem was identified in [2, 15], it remained completely open, [38].

We have had to develop new concepts in order to handle conducting bands and explore the regime $\epsilon \ll \beta$. In particular, our proof rests on the construction of the *local gap structure* for disordered systems, which is more robust than the standard description of the localization and, in particular, survives the time-dependent perturbations described by (1.1). This is the content of Theorem 3.2 below. We then build an adiabatic theory associated with this structure for the dynamics of $H(t)$, characterized by local rather than global adiabatic behavior. We believe that this new result (Theorem 2.8 below), which we will refer to as the *local adiabatic theorem*, is of independent interest. The derivation of the Kubo formula then follows via more standard (albeit technically involved) methods.

The rest of the paper is organized as follows: We formulate our core technical result, the local adiabatic theorem, Theorem 2.8, in Sect. 2. This result relies on the dynamical properties associated with the local gap structure for the time-dependent Hamiltonian $H(s)$, presented in Sect. 3. The origin of this structure can be traced back to the time-independent random system H on a torus, which is studied in Sect. 4. We then study the local adiabatic behavior of disordered systems in Sect. 5 and complete the proof of Theorem 2.8 in Sect. 6. This theory is used to prove our principal result on the Kubo formula, Theorem 1.1, in Sect. 7. Appendices A–B contain results of independent interest, namely hybridization delocalization in dimension one and the construction of a Wannier-type basis for disordered systems, respectively. Various auxiliary results are included in Appendix C.

2 Local adiabatic theorem

In this section, we unveil our core technical result - the local adiabatic theorem, specifically designed to work with disordered systems. Our starting point here is a brief discussion of the localization phenomenon.

2.1 Localization and delocalization for time-dependent systems

The presence of disorder in quantum mechanical systems leads to the phenomenon of localization. *Spectral* localization manifests in the emergence of energy interval(s) $J_{loc} \subset \mathbb{R}$ such that, for almost all random configurations ω, $\sigma(H) \cap J_{loc}$ is pure point. Moreover, the eigenvectors of H in J_{loc} are (spatially) exponentially localized in the sense of (2.1) below.

Spectral localization is not stable under perturbation: The rank one perturbation family $H_A(\beta)$ of the form $H_A(\beta) = H_A + \beta\chi_{\{0\}}$ exhibits almost sure singular continuous spectrum for a G_δ-dense set of β's, [22, 34]. Although there are no rigorous results beyond rank 1 perturbation, one should not expect much uniformity of the localization properties as a function of t or β of the Hamiltonian (1.1), provided that W is sufficiently non-trivial.

2.1.1 Dynamical localization

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 concept is still well-defined for a time-dependent Hamiltonian $H(t)$, and a natural question is whether it is still dynamically localized for at least small perturbations $\beta \ll 1$.

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 [67], 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 [61] and the time-quasi-periodic [19] settings. The case of spatially extensive periodic systems with few frequencies was considered in [24]. In the $\beta = \epsilon$ adiabatic setting, it was considered in [54]. For time periodic systems, one can also consider the spectral localization of the associated Floquet operator, [1, 24, 61]. On a heuristic level [24, Sect. 1], there should be a transition from a localized regime to a non-localized regime when $\nu \sim \beta \exp\left(-c_d \beta^{-p_d}\right)$, where ν is the Floquet frequency[1] for W and c_d, p_d are dimension-dependent parameters. For $\nu \gg \beta \exp\left(-c_d \beta^{-p_d}\right)$ only a small fraction of Floquet eigenstates delocalizes. Apart from constraints on β, ϵ, 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 $\mathbb{R}$.

The instability of spectral and dynamical localization is due to the phenomena of resonant hybridization that we will describe next.

2.1.2 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, ψ_i), $E_i \in J_{loc}$ one can find $x_i \in \mathbb{Z}^d$, called a *localization center* for ψ_i, satisfying

$$|\psi_i(x)| \leq C \langle x \rangle^{d+1} e^{-c|x-x_i|}, \tag{2.1}$$

where $\langle x \rangle := \sqrt{|x|^2 + 1}$. 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 μ) and in dimension $d = 1$, for almost all configurations ω. We will not attempt to cite the extensive literature of history, reviews, results and open problems concerning this model and its variants. We will instead refer the interested reader to a recent monograph [3] on the subject.

The instability of such uniform localization properties with respect to perturbations can be linked to a mechanism known as *resonant hybridization*, see, e.g., [3, Chap. 15]. This concept can be illustrated by considering a two-level system with a Hamiltonian $H(s)$ of the form

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

[1] The parameter ν in [24] plays the same role as ϵ in our setting.

When $s = 0$, the canonical basis e_1, e_2 is an eigenbasis for $H(s)$. 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.

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 of approximate length g. As we show in Appendix A, this behavior also occurs in the extended disordered system. The hybridization implies that the spectral flow is very nonlocal, as disordered analogues of $e_{1,2}$ can be localized arbitrarily far away from each other.

More precisely, if we consider a finite volume restriction of H, say to a box with side length $\mathcal{L}$, we can then label the eigenstates $\psi_{i,s}$ so that for each i, $s \mapsto \psi_{i,s}$ is continuous, [43]. However, we do expect the modulus of continuity to diverge badly as $\mathcal{L} \to \infty$.

We are not aware of any prior rigorous results making the two-level heuristics exact for $\mathbb{Z}^d$ systems for any d (however, see [3, Chap. 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 A.2, which informally can be expressed as

Theorem 2.1 *Let H 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.*

2.2 Adiabatic theory

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

$$i\dot{\psi}(t) = H(t)\psi(t), \quad \psi(0) = \psi_o, \tag{2.2}$$

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

A more interesting and physically realistic situation arises when the dependence on time in $H(t)$ is present but is 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 (2.2)). To make this statement more quantitative, it is convenient to consider the family $H(\epsilon t)$, where ϵ is a small (adiabatic) parameter, and the physical time t runs over the long interval $[0, 1/\epsilon]$. After a change of variables

$s = \epsilon t$ where s is a rescaled time, the relevant IVP becomes

$$i\epsilon\dot{\psi}_\epsilon(s) = H(s)\psi_\epsilon(s), \quad \psi_\epsilon(0) = \psi_o, \quad s \in [0,1]. \tag{2.3}$$

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

$$i\epsilon\partial_s U_\epsilon(s) = H(s)U_\epsilon(s), \quad U_\epsilon(0) = \mathbb{1}. \tag{2.4}$$

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

$$\lim_{\epsilon\to 0} \|\psi_\epsilon(s) - P(s)\psi_\epsilon(s)\| = 0, \tag{2.5}$$

provided $H(s)$ is smooth. A stronger statement holds, namely

$$\lim_{\epsilon\to 0} \left\|U_\epsilon(s)P(0)U_\epsilon^*(s) - P(s)\right\| = 0, \tag{2.6}$$

and one can make the error estimate for the norm above explicit in terms of its ϵ and g dependencies, see e.g., Lemma 5.5 below.

As mentioned above, we can label the eigenstates $\psi_{i,s}$ of a finite system in such a way that the spectral flow $s \mapsto \psi_{i,s}$ is continuous for each i. Suppose there are no degeneracies, which is the generic case. Then each eigenvalue is gapped, and the adiabatic theorem says that in the limit $\epsilon \to 0$, the solution of (2.3) is the spectral flow. Combined with Theorem 2.1 this implies that dynamical localization fails for $\epsilon \to 0$ as the spectral flow is extremely nonlocal. However, for $\epsilon > 0$, 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 $\psi_{i,s}$, even 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 ϕ_s 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 *local adiabatic theorem*, and state it quantitatively as Theorem 2.8 below. One can interpret this result as meta-stability of ϕ_s with a very long lifetime.

The adiabatic theorem and its derivatives play an fundamental 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, [16]. The modern adiabatic theory was initiated by Kato in 1950, [42], and has since been studied intensively in the mathematical physics literature. The adiabatic theorem has been extended to a situation where the family $P(s)$ is smooth, but no gap is present, [4, 17]. This situation usually occurs for a ground state in the threshold of the continuous spectrum. The other possible scenario occurs in rank one perturbed completely localized system, where one can show that the Fermi projection $P_F(t)$ is a continuous function for a set of the full Lebesgue measure, even when $\sigma(H(t))$ is not pure point, [8]. In space-adiabatic perturbation theory [57], the gap is closed by a locally small but globally large perturbation (for related work in field theory, see [63]). More recently, the adiabatic theorem was established for certain systems with a spectral gap but non-smooth $P(s)$, [12, 53]. This situation arises in the context of the thermodynamic limit for many-body systems.

For the disordered systems that are not entirely localized, it is necessary to consider a scenario where *both* conditions fail to hold.

2.3 Local adiabatic theorem

To properly formulate this assertion, we must first establish the appropriate framework.

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

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

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

Assumption 2.2 The operators $H(s)$ are uniformly bounded, smooth, r-local, self-adjoint operators acting on $\ell^2\left(\mathbb{Z}^d\right)$, of the form (1.1) that satisfy $\|H(s)\| \le C$. In addition, for all $k \in \mathbb{N}_0$, $W^{(k+1)}(0) = W^{(k+1)}(1) = 0$, and there exists a constant C_k such that $\|W^{(k)}(s)\| \le C_k$.

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

Assumption 2.3 (Finite range of disorder correlations) For any pair of subsets Θ, Φ of $\mathbb{Z}^d$ that satisfy $\operatorname{dist}(\Theta, \Phi) > r$, the operators H^Θ and H^Φ are statistically independent.

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

$$|x-y|_\Theta = \min\left(|x-y|, (\operatorname{dist}(x, \partial_1\Theta) + \operatorname{dist}(y, \partial_1\Theta))\right), \tag{2.7}$$

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

Assumption 2.4 (Fractional moment condition on J_{loc}) There exist $q \in (0, 1)$ and $C_q, c > 0$ such that, for any subset Θ of $\mathbb{Z}^d$, for any $E \in J_{loc}$, and any $\eta \neq 0$, we have

$$\sup_{E \in J_{loc}} \mathbb{E}\left(\left|(H^\Theta - E - i\eta)^{-1}(x, y)\right|^q\right) \le C_q e^{-c|x-y|_\Theta} \text{ for all } x, y \in \Theta, \tag{2.8}$$

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

For some of our results we will also need

Assumption 2.5 (Finite spectral multiplicity) There exists $m \in \mathbb{N}$ such that, for any $\Theta \subset \mathbb{Z}^d$, the multiplicity of eigenvalues of H^Θ does not exceed m almost surely.

Remark 2.6 For the standard Anderson model with absolutely continuous random distributions $m = 1$, [60]. This type of result can be extended to a larger class of discrete models, see, e.g., [3, Theorem 5.8] and [23]. While the simplicity of the spectrum is, in general, not known to hold for models that satisfy Assumptions 2.3–2.4, in practice, a majority of them are generated using finite-rank operators for which Assumption 2.5 does hold, [39].

Remark 2.7 Surprisingly, the basic localization property (2.1) has only been proven in existing literature under the assumption of spectrum simplicity (i.e., $m = 1$ in Assumption 2.5 above), cf. [3, 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. The argument there relies on the construction of the so-called generalized Wannier basis for an eigenprojection of the localized Hamiltonian, consisting of exponentially localized functions.

The local adiabatic theorem is easier stated in finite volume for a bulk system, we introduce a periodized restriction of $H(s)$ to a discrete torus $\mathbb{T} = \mathbb{T}^d_M$, which we associate with the hypercube $[1, M]^d$ whose opposite faces are identified. This restriction is defined as

$$H^{\mathbb{T}}(x, y) = \frac{1}{2} \sum_{n \in M\mathbb{Z}^d} H(x, y+n) + H(x+n, y), \quad x, y \in \mathbb{T}. \tag{2.9}$$

Our two main parameters are the adiabaticity parameter ϵ and the driving strength β, introduced earlier in (2.3) and (1.1), respectively. In our results we will use four exponents,

$$\xi = \tfrac{d}{q}, \quad \xi' = d + \tfrac{1}{2} + \xi, \quad p_1 > d + \xi', \quad p_2 > \max\left(\xi', 2\xi\right), \tag{2.10}$$

with fixed p_1, p_2 satisfying the last two inequalities. Throughout this paper, we will assume that $\beta \ll 1$ and $\epsilon \ll 1$ satisfy

$$e^{-\beta^{-1/(2p_1)}} < \epsilon < \beta^{p_2 p_1}. \tag{2.11}$$

It will be convenient to work with a (generally flexible) scale parameter $\ell \in \mathbb{N}$ satisfying

$$\ell^{-p_2} \geq \epsilon \geq e^{-c\sqrt{\ell}}, \qquad \beta \leq \ell^{-p_1}, \tag{2.12}$$

whose existence is guaranteed by (2.11).

We will use generic, M, ϵ, β, ℓ-independent constants C, c 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 μ, as well as on the constants C_q, C_k, etc.). We allow for the system size

M to be arbitrarily large, and all of our estimates will be uniform in M. We will use C to indicate that the constant should be sufficiently large for a bound to hold, and c to indicate that the constant should be sufficiently small.

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

Theorem 2.8 (Local adiabatic theorem) *Suppose that Assumptions 2.2–2.5 hold for $H(0)$ and the integrated density of states $\mathcal{N}_{J_{loc}}$ is a.s. positive. Let β, ϵ, ℓ satisfy (2.11)–(2.12) and J'_{loc} be any closed interval contained in J_{loc}. Assuming that ℓ is large enough, with probability at least $1 - e^{-c\sqrt{\ell}}$, the following holds for a fraction of at least $1 - e^{-c\sqrt{\ell}}$ of eigenstates ψ of $H^{\mathbb{T}}$ with eigenvalues $E \in J'_{loc}$: There is a region $R \subset \mathbb{T}$ with* $\operatorname{diam}(R) \leq c\ell^{3/2}$ *such that*

(i) *For all $s \in [0.1]$, $H^R(s)$ possesses the spectral patch $S(s) \subset \sigma(H^R(s))$ which is isolated from the rest of the spectrum $\sigma(H^R(s))$. We denote the associated spectral projector by $P(s)$.*
(ii) *The solution $\psi_\epsilon(s)$ of (2.3) with $\psi_\epsilon(0) = \psi$ satisfies*

$$\max_{s\in[0,1]} \|(1 - P(s))\psi_\epsilon(s)\| \leq C\left(\epsilon\ell^{\xi'} + e^{-c\sqrt{\ell}}\right). \tag{2.13}$$

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

$$\|(1 - P(1))\psi_\epsilon(1)\| \leq C_N\left(\epsilon^N\left(\ell^{N\xi'} + \ell^{(2N+1)\xi}\right) + e^{-c\sqrt{\ell}}\right). \tag{2.14}$$

This statement will be proved in Sect. 6.

Remark 2.9 While the assertion is formulated for tori of the arbitrary size M, in applications (e.g., in the proof of our main result, Theorem 1.1), we often have $M \ll e^{-c\sqrt{\ell}}$. In this case, the statement holds for *all* eigenstates rather than their fraction, with the same probability.

Remark 2.10 Let us note that both the upper and lower bounds on ϵ in (2.12) have to do with the faithfulness of our approximation of the actual eigenstate for $H^{\mathbb{T}}$ 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 $H^{\mathbb{T}}$ (so the spatial faithfulness of our approximation is destroyed). On the other hand, if R is too big, the gaps in the spectrum of $H^{\mathbb{T}}$ become smaller than the size β 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 2.11 If the spectrum of $H^{\mathcal{R}}$ is *level-spaced*, i.e., if the probability of a spacing significantly smaller than $|\mathcal{R}|^{-1}$ is small (as one can prove, e.g., for the standard

Anderson model [44] and, at the bottom of the spectrum, for more general random models, [23]), then with large probability the spectral patch $S(s)$ can be chosen to consist of a simple eigenvalue, making $P(s)$ rank-one. 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 $H^{\mathbb{T}}(s)$. However, we do not expect this property to hold for all times s on the basis of the hybridization result, Theorem 2.1, which shows that physical evolution cannot follow the non-local spectral flow.

Remark 2.12 It follows from the relationship between ϵ and ℓ in (2.12) that for the times $\sim \epsilon^{-1}$ the dynamical localization length is $O(\ln(\epsilon^{-1}))$. It is consistent with the estimates for rank-one perturbation of completely localized (time-independent) systems, where the growth of localization length is sub-polynomial in ϵ, [49], due to the zero Hausdorff dimensionality of the spectrum, [22].

3 Local gap structure

Analyzing the spatial structure of spectral gaps is crucial to proving the adiabatic theorem described above. We introduce relevant concepts and state the corresponding results in this section.

We start with some supplementary notation. By $\Lambda_R(y) \subset \mathbb{Z}^d$ we will denote a cube $\Lambda_R = \Lambda_R(y) := \left([-R, R]^d + y\right) \cap \mathbb{Z}^d$ 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 ℓ-extended boundary, i.e.,

$$\partial_\ell \Phi = \left\{x \in \Phi : \operatorname{dist}\left(x, \Phi^c\right) \leq \ell\right\}. \tag{3.1}$$

By Φ_ℓ we will denote

$$\Phi_\ell = \Phi \setminus \partial_\ell \Phi. \tag{3.2}$$

For a Hermitian operator H, we denote by $P_J(H)$ the spectral projection of H on the set $J \subset \mathbb{R}$. For a positive real number a, aJ denotes the interval obtained from J by scaling the interval with respect to its midpoint by a factor of a. For an operator X, we denote $\bar{X} := 1 - X$. For $\mathcal{A} \subset \mathbb{T}$, $c \in \mathbb{R}_+$, and $\ell \in \mathbb{N}$, let $\rho^\ell_{\mathcal{A}}$ be a (scaled) distance function

$$\rho^\ell_{\mathcal{A}}(x) = \frac{\operatorname{dist}(\mathcal{A}, \{x\})}{\sqrt{\ell}}. \tag{3.3}$$

We set

$$\|K\|_{c,\ell} = \left\| \mathrm{e}^{-c\,\rho^\ell_{\mathcal{A}}} K \,\mathrm{e}^{c\,\rho^\ell_{\mathcal{A}}} \right\| \tag{3.4}$$

This norm is multiplicative, i.e.,

$$\|AB\|_{c,\ell} \leq \|A\|_{c,\ell} \, \|B\|_{c,\ell} \tag{3.5}$$

for a pair of operators A, B.

We now introduce the concepts of local and ultra-local gap 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 2.8. 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 $J \subset J_{loc}$ 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 gap structure* in J if there exists a disjoint collection $\{\mathcal{T}_\gamma\}$ of subsets of $\mathbb{T}$ with $\operatorname{diam}(\mathcal{T}_\gamma) \le C\ell^{3/2}$ such that the following property holds: For each ψ_n, there exists γ such that

$$\left\| \psi_n - P_{\hat{J}}\left(H^{\mathcal{T}_\gamma}(0)\right) \psi_n) \right\| \le e^{-c\sqrt{\ell}}, \tag{3.6}$$

where $\hat{J} := \left\{ x \in \mathbb{R} : \operatorname{dist}(x, J) \le e^{-c\sqrt{\ell}} \right\}$. Let us note that the random Schrödinger operators $H(0)$ satisfying Assumptions 2.4 possess the ultra-local property with probability $\ge 1 - e^{-c\sqrt{\ell}}$ provided the length of the interval J is of order $\ell^{-\xi}$ (in fact, a stronger statement holds, see Theorem 4.4 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 3.1 We will say that $H^{\mathbb{T}}(s)$ possesses a *local gap structure* in $J \subset J_{loc}$ if there exists a disjoint collection $\{\mathcal{T}_\gamma\}$ of subsets of $\mathbb{T}$ such that $\operatorname{diam}(\mathcal{T}_\gamma) \le \ell^{3/2}$ for each γ with the following properties:

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

$$J_\gamma \subset J \text{ and } \operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma}(s))\right) \ge \Delta; \tag{3.7}$$

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

$$\| P_J(s) \chi_{\mathbb{T}\setminus\mathcal{T}_{8\ell}} \| \le e^{-c\sqrt{\ell}}, \tag{3.8}$$

and

$$\| P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \chi_{\partial_\ell \mathcal{T}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\partial_\ell \mathcal{T}} - \chi_{\mathcal{T}_{8\ell}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\mathcal{T}_{8\ell}} \| \le e^{-c\sqrt{\ell}}. \tag{3.9}$$

The unperturbed Hamiltonian possesses a local gap structure for small, but not too small, Δ. As we shall see in the proof of Theorem 3.2, the local gap structure is stable under perturbation, i.e., if the Hamiltonian possesses a local gap structure for $s = 0$ on J, it possesses it for *all* s on a slightly smaller interval J', provided β is sufficiently small. The reason for this stability is related to the fact that, under small local perturbations, an eigenstate with energy E is 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 $\mathcal{T}_\gamma$, so is the eigenstate. The locality property is fully compatible with the hybridization effect: Even if initially the state is ultra-local (concentrated in a single patch $\mathcal{T}_{\gamma_o}$), it can hybridize to a number of different patches $\mathcal{T}_\gamma$ as s increases.

The scaling of various objects with ℓ depends on q, d and our choice of stretch-exponential error $\exp(-c\sqrt{\ell})$. The correct scaling of Δ and β to ensure the existence of local gap structure is given in Theorem 3.2.

Once the local gap structure for the family $H(s)$ is established, one can use an (enhanced) version of the standard, gapped adiabatic theorem (Lemma 5.5) to control the behavior of the individual spectral patches $P_{J_\gamma}\left(H^{\mathcal{T}_\gamma}(s)\right)$, invoking Definition 3.1(i). This in turn allows us to control the physical evolution of spectral data $Q(s)$ for $H^{\mathbb{T}}(s)$ near the energy E (see Sect. 5.5 for details). Finally, we show that this translates to the adiabatic theorem for the (distorted) Fermi projection, Theorem 3.3. The principal idea here is that the removal of the spectral data $Q(s)$ on one hand creates a spectral gap for H (making the standard adiabatic theorem applicable) and on the other does not distort the adiabatic behavior of the system too much since $Q(s)$ itself evolves adiabatically, a feature verified in the previous step.

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

We will show in Sect. 4 that Anderson-type models possess a local gap structure in the sense of Definition 3.1. In fact, a stronger statement holds:

Theorem 3.2 (Local gap structure of $H^{\mathbb{T}}(s)$) *Suppose that H satisfies Assumptions 2.3–2.4 and the family $H(s)$ satisfies Assumption 2.2. We consider a torus $\mathbb{T}$ whose linear dimension is $\mathcal{L}$. Then, there exist constants c, $\{c_i\}_{i=1}^6$ such that for any $a \le c_1$,*

$$\mathcal{L} = e^{a\sqrt{\ell}}, \quad V_\ell = \ell^{d+1/2}, \quad \delta = c_2 \ell^{-\xi}, \quad \Delta = c_3 V_\ell^{-1} \ell^{-\xi}, \tag{3.10}$$

ℓ large enough, and $\beta \le \ell^{-p_1}$, $H^{\mathbb{T}}(s)$ possesses a local gap structure for the energy interval $J = (E - 6\delta, E + 6\delta)$: One can find a disjoint collection $\left\{\mathcal{T}_\gamma\right\}$ of subsets of Λ such that $|\mathcal{T}_\gamma| \le c_4 V_\ell$, $\operatorname{diam}\left(\mathcal{T}_\gamma\right) \le c_5 \ell^{3/2}$ for each γ and the following conditions hold true with probability $> 1 - e^{-c_6\sqrt{\ell}}$:

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

$$(E - 3\delta, E + 3\delta) \subset J_\gamma \subset J \text{ and } \operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma}(s))\right) \ge \Delta; \tag{3.11}$$

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

$$\left\| P_J(s) \chi_{\Lambda \setminus \mathcal{T}_{8\ell}} \right\| \le e^{-c\sqrt{\ell}}, \tag{3.12}$$

and

$$\| P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \chi_{\partial_\ell \mathcal{T}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\partial_\ell \mathcal{T}} - \chi_{\mathcal{T}_{8\ell}} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) \chi_{\mathcal{T}_{8\ell}} \| \le e^{-c\sqrt{\ell}}. \tag{3.13}$$

(iii) (*Exponential Decay of Correlations*) *Let $\mathcal{A}_o = \partial_\ell \mathcal{T}_\gamma \cup (\mathcal{T}_\gamma)_{8\ell}$, then (with $\mathcal{A} = \mathcal{A}_o$ in (3.3)–(3.4)) we have*

$$\left\| \left(H^{\mathcal{T}_\gamma}(s)) - z\right)^{-1} \right\|_{c,\ell} \le \frac{\ell^{3d}}{\Delta} \frac{1}{\langle Im\, z\rangle}, \tag{3.14}$$

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

The dependence on β 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 $\beta \le \ell^{-p_1}$.

This assertion will be proved in Sect. 4.3.

An additional statement that we will need in our proof of Theorem 1.1 is

Theorem 3.3 (Local adiabatic theorem for distorted Fermi projection) *In the setting of Theorem 3.2, assume in addition that (2.12) holds and fix $N \in \mathbb{N}$. Then for ℓ large enough, there exists a smooth family of orthogonal projections $\mathcal{Q}(s)$ with the following properties:*

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

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

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

$$\|\mathcal{Q}_\epsilon(s) - \mathcal{Q}(s)\| \le C_N \left(\epsilon^N \left(\frac{1}{\Delta^N} + \frac{1}{\delta^{2N+1}}\right) + e^{-c\sqrt{\ell}}\right). \tag{3.15}$$

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

This assertion will be proved in Sect. 5.6.

4 Localization on a torus

4.1 Consequences of Assumptions 2.2–2.4

We first note that Assumptions 2.2–2.4 imply localization on a torus as well (e.g., [3, Theorem 11.2]):

$$\sup_{E \in J_{loc}} \mathbb{E}\left(\left|(H^{\mathbb{T}} - E - i\eta)^{-1}(x, y)\right|^q\right) \le Ce^{-cd_{\mathbb{T}}(x,y)} \text{ for all } x, y \in \mathbb{T}, \tag{4.1}$$

where $d_{\mathbb{T}}(x, y)$ 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 $E \in J_{loc}$,*

$$\mathbb{P}\left\{\operatorname{dist}\left\{E, \sigma(H^{\Theta})\right\} \le \nu\right\} \le C\nu^q |\Theta|. \tag{4.2}$$

For a proof, see e.g., [28, the proof of Proposition 5.1].

Together with Assumption 2.3, Lemma 4.1 yields

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

$$\mathbb{P}\left\{\operatorname{dist}\left(\sigma(H^{\Theta}) \cap J_{loc}, \sigma(H^{\Phi}) \cap J_{loc}\right) \leq \nu\right\} \leq C\nu^{2q} |\Theta| |\Phi| . \tag{4.3}$$

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

$$\mathbb{P}\left\{\operatorname{dist}\left(E, \sigma(H^{\Theta_i})\right) \leq \nu \text{ for all } i\right\} \leq \left(C\nu^q D\right)^n . \tag{4.4}$$

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^{Θ} a projection $P_I(H^{\Theta})$ and analogously for $P_I(H^{\mathbb{T}})$.

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

$$Q^{\Theta}(x, y; J_{loc}) = \sum_{\lambda \in \sigma(H^{\Theta}) \cap J_{loc}} \left| P_{\{\lambda\}}^{\Theta}(x, y) \right| \tag{4.5}$$

satisfies

$$\mathbb{E} Q^{\Theta}(x, y; J_{loc}) \leq \mathrm{e}^{-c|x-y|_{\Theta}} \tag{4.6}$$

for some $c > 0$ that depends only on μ and q. For the non correlated randomness, see, e.g. [3, 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 [29, Theorem 4.2].

The relation (4.6) implies that all eigenstates in $P_{J_{loc}}^{\Theta}$ 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, ℓ)-*localizing* for $H^{\mathbb{T}}$ in the interval $I \subset J_{loc}$ if for all eigenpairs $(E_n, \psi_n)_{E_n \in I}$ of H^{Θ} there exists a set $\{x_n\}$ in Θ such that

$$|\psi_n(y)| \leq e^{-c|y-x_n|_{\Theta}} \text{ for any } y \in \Theta \text{ such that } |y - x_n|_{\Theta} \geq \sqrt{\ell}. \tag{4.7}$$

We then have the following result:

Theorem 4.4 *Suppose that Assumptions 2.4–2.5 hold. Then there exist* $c > 0$ *such that the probability that a set* $\Theta \subset \mathbb{T}$ *is* (c, ℓ)*-localizing for* $H^{\mathbb{T}}$ *in the interval* J_{loc} *is* $\geq 1 - C|\Theta|^2 e^{-c\sqrt{\ell}}$.

The proof of this statement can be found in Appendix B (Theorem B.2).

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

$$\tilde{P}_E^{\mathbb{T}}(x, y) = \begin{cases} P_E^{\mathbb{T}}(x \bmod \mathcal{L}Z^d, y \bmod \mathcal{L}Z^d) & x - y \in \mathbb{T} \\ 0 & x - y \notin \mathbb{T} \end{cases}$$

The next assertion implies that deep inside $\mathbb{T}$, P_E and $\tilde{P}_E^{\mathbb{T}}$ are close.

Proposition 4.5 *Suppose that Assumptions 2.2–2.4 hold. Then there exists $c > 0$ such that the probability*

$$\mathbb{P}\left(\left\| \left(P_E - \tilde{P}_E^{\mathbb{T}}\right) \chi_{\Lambda_{\mathcal{L}/2}(0)} \right\| > \mathrm{e}^{-c\mathcal{L}}\right) \le \mathrm{e}^{-c\mathcal{L}}. \tag{4.8}$$

For a proof, see [30, 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 2.2–2.4 hold. Then for any $I := [E_1, E_2] \subset J_{loc}$ and any $\Theta \subset \mathbb{T}$, there exists $c > 0$ such that*

$$\mathbb{E}\left| P_\sharp^\Theta(x, y) \right| \le \mathrm{e}^{-c|x-y|_\Theta}, \quad \sharp = I, E, \tag{4.9}$$

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

$$\mathbb{E}\left|\left(\bar{P}_I^\Theta \left(H^\Theta - z\right)^{-1}\right)(x, y)\right| \le \frac{1}{E_2 - E_1} \frac{\mathrm{e}^{-c|x-y|_\Theta}}{\langle Im\, z\rangle} \tag{4.10}$$

Proof Let $\sharp = I$. Since Θ is finite, the spectrum of H^Θ is a discrete set. By (2.8),

$$\{E_1, E_2\} \not\subset \sigma\left(H^\Theta\right)$$

almost surely. Thus the spectral projection P_I^Θ is equal to

$$P_I^\Theta = -(2\pi)^{-1} \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j \left(H^\Theta - iu - E_j\right)^{-1} du \tag{4.11}$$

almost surely, see (C.7). Using $|\left(H^\Theta - iu - E_j\right)^{-1}(x, y)| \le |u|^{-1}$, we get a bound

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

We note that for $|u| \ge 1$, we can decompose

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

Thus, using (4.1), r-locality of H, and $|H(x, y)| \le C$,

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

Since $\left|P_I^{\Theta}(x,y)\right| \le 1$ for all $x, y \in \Theta$, by modifying c if necessary we get (4.9) for $\sharp = I$. The argument for $\sharp = E$ is nearly identical.

To get the second assertion of the lemma, we use

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

and

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

They yield

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

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

We will be using the probabilistic version of Lemma 4.6, which follows from the previous statement by Markov's inequality.

Lemma 4.7 *Suppose that Assumptions 2.2–2.4 hold. Let $J := [E_1, E_2] \subset J_{loc}$. Then, there exists $c > 0$ such that for any $\Theta \subset \mathbb{T}$ with $|\Theta| \le \ell^{3/4}$, the probability that for all x, y with $|x - y|_{\Theta} \ge \sqrt{\ell}$,*

$$\left|\left(P_J^{\Theta}\right)(x,y)\right|, \left|\left(\bar{P}_J^{\Theta}\left(H^{\Theta} - z\right)^{-1}\right)(x,y)\right| \le \mathrm{e}^{-c|x-y|_{\Theta}} \tag{4.12}$$

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

4.2 Local gap structure of $H^{\mathbb{T}}$

Here we will again suppose that Assumptions 2.2–2.4 hold.

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

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

where

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

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., $\Lambda_\ell(a) = \{x \in \mathbb{T} : d_{\mathbb{T}}(x, a) \le \ell\}$. We recall that we use a max distance throughout this paper. We will refer to this collection of boxes as a *suitable ℓ-cover* of $\mathbb{T}$.

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

Lemma 4.8 *Let $r < \ell < \mathcal{L}$. Then, a suitable ℓ-cover satisfies*

(i) $\mathbb{T} = \bigcup_{a \in \Xi_\ell} \Lambda_\ell(a)$;
(ii) *For all $y \in \mathbb{T}$ there is $a = a(y) \in \Xi_\ell$ such that $\Lambda_{\ell/4}(y) \subset \Lambda_\ell(a)$. 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) $\left(\frac{\mathcal{L}}{\ell}\right)^d \le |\Xi_\ell| \le \left(\frac{2\mathcal{L}}{\ell}\right)^d$.

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

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

$$\mathbb{P}\left\{\text{dist}\left(E, \sigma(H^{\Lambda_\ell(a)})\right) \le \nu \text{ for all } \Lambda_\ell(a) \in \mathcal{S}\right\} \le \left(C\nu^q \ell^d\right)^{|\mathcal{S}|}. \tag{4.15}$$

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

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

for an ℓ–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 Ξ_ℓ 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 $\frac{3}{2}\ell$ on the torus $\mathbb{T}$. By $\mathcal{R}_M$ we will denote a set of all connected subgraphs of Ξ_ℓ with cardinality M, and by $\mathcal{S}_M$ we will denote a collection of sets $\{\cup_{a \in R} \Lambda_\ell(a) : R \in \mathcal{R}_M\}$.

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

$$(2d\mathrm{e})^M |\Xi_\ell| \le \left(\frac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^M. \tag{4.17}$$

Proof of Lemma 4.9 We first note that each set S in $\mathcal{S}_M$ looks like a compressed d-dimensional polycube of size M, and that we can bound the number of distinct S_Ms using the same method as for the regular polycubes, see e.g., [14]. 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 $1, \dots, M$ 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 $1 \le j \le 2d-1$ 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

$$\binom{(2d-1)M}{M-1} \le (2d\mathrm{e})^M . \tag{4.18}$$

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 Ξ_ℓ, yielding (4.17). □

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

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

For a given $\mathcal{T}$, we will denote by $M(\mathcal{T})$ the size of the largest connected component,

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

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

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

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

$$\mathbb{P}(\Omega_N^c) \le \left(\tfrac{2\mathcal{L}}{\ell}\right)^d \mathrm{e}^{-N}. \tag{4.19}$$

Proof For any $\omega \in \Omega_N^c$, there exists at least one cluster $\mathcal{T}_\gamma \in \mathcal{S}_M$ with $M \ge N$. Let $\tilde{\mathcal{T}}_\gamma$ denote the union of boxes that generates $\mathcal{T}_\gamma$, i.e., $\mathcal{T}_\gamma$ is formed by all boxes that overlap with at least one box in $\tilde{\mathcal{T}}_\gamma$. We note that $\tilde{\mathcal{T}}_\gamma$ is in general not uniquely defined, but this will not play a role in our argument. We also remark that any box $\Lambda_l(a) \subset \tilde{\mathcal{T}}_\gamma$ overlaps with 3^d boxes, so $\left|\tilde{\mathcal{T}}_\gamma\right| \le 3^d \left|\mathcal{T}_\gamma\right|$. Let U be a collection of vectors in $\mathbb{R}^d$ whose components take binary values. Then $\Xi_\ell = \cup_{e\in U} \Xi_{\ell,e}$, where $\Xi_{\ell,e} = \frac{3}{2}e + (3\ell\mathbb{Z})^d / \mathcal{L}\mathbb{Z}^d$, 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 ℓ is even. Hence, for any $S \subset \Xi_\ell$, there exists $e \in U$ such that $\left|S \cap \Xi_{\ell,e}\right| \ge 2^{-d} |S|$. In particular, the number of non-overlapping boxes in $\tilde{\mathcal{T}}_\gamma$ 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 $\mathcal{T}$ has at least one cluster $\mathcal{T}_\gamma \in \mathcal{S}_M$ with $M \ge N$ is bounded by $\left(C\,|I|^q\,\ell^d\right)^{6^{-d}M}$. It follows now from Lemma 4.9 that

$$\mathbb{P}(\Omega_N^c) \le \sum_{M=N}^{\infty} \left(\tfrac{2\mathcal{L}}{\ell}\right)^d \left((2d\mathrm{e})^{(6^d)}\, C\,|I|^q\,\ell^d\right)^{6^{-d}M}. \tag{4.21}$$

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

For an integer N, we now consider a subset $\Omega_{loc,N}$ of the full configuration space for which $\mathbb{T}$ and all 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*

$$\mathbb{P}(\Omega_{loc,N}^c) \le C N^2\,(2\mathcal{L}\ell)^d\,(2d\mathrm{e})^N\, e^{-c\sqrt{\ell}}. \tag{4.22}$$

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

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

thanks to Lemma 4.9. Their maximal volume is bounded by $N\ell^d$. Thus, we can bound

$$\mathbb{P}(\Omega_{loc,N}^c) \le C\left(\tfrac{2\mathcal{L}}{\ell}\right)^d (2d\mathrm{e})^N \left(N\ell^d\right)^2 e^{-c\sqrt{\ell}} = C N^2\,(2\mathcal{L}\ell)^d\,(2d\mathrm{e})^N\,\mathrm{e}^{-c\sqrt{\ell}} \tag{4.23}$$

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 4.10–4.11, for ℓ large enough and intervals $I \subset J_{loc}$ satisfying $|I| < c\ell^{-d/q}$, we have

$$\mathbb{P}(\left(\Omega_N \cap \Omega_{loc,N}\right)^c) \le \mathcal{L}^d \mathrm{e}^{-c\sqrt{\ell}}. \tag{4.24}$$

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

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

We note that

$$\max_\gamma \operatorname{diam}\left(\mathcal{T}_\gamma\right) \leq L := C\ell^{3/2}. \tag{4.25}$$

Let Ω_G be a subset of the configuration set $\Omega_N \cap \Omega_{loc,N}$ such that, for c small enough, $\omega \in \Omega_G$, $z \in \mathbb{C}$ with $\mathrm{Re}(z) = E_\gamma^\pm$, and all $x, y \in \mathcal{T}_\gamma$, the following bound holds:

$$\sup_{\mathcal{T}_\gamma}\left|\left(\left(H^{\mathcal{T}_\gamma} - z\right)^{-1}\right)(x,y)\right| \mathrm{e}^{c\ell^{-1/2}|x-y|_{\mathcal{T}_\gamma}} \leq C\ell^{d+\frac{1}{2}}\,|I|^{-1}\,\langle Im\, z\rangle^{-1}. \tag{4.26}$$

Applying Lemma 4.7 with $J = E_\gamma^\pm + [-c\ell^{-d-1/2}\,|I|\,, c\ell^{-d-1/2}\,|I|]$ and $z \in \mathbb{C}$ with $\mathrm{Re}(z) = E_\gamma^\pm$ yields

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

Proposition 4.12 *Let $\omega \in \Omega_G$, and let $I \subset J_{loc}$ be such that $|I| < c\ell^{-d/q}$. Suppose that ℓ is large enough, then*

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

$$I/2 \subset J_\gamma \subset I \quad \text{and} \quad \operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma})\right) \geq c\ell^{-d-1/2}\,|I|; \tag{4.27}$$

(ii) (*Support of spectral projections*)

$$\left\|P_I(H^{\mathbb{T}})\delta_x\right\| \leq e^{-c\sqrt{\ell}} \text{ for any } x \in \mathbb{T} \setminus \mathcal{T}_\ell \tag{4.28}$$

(*recall* (3.2)), *and*

$$\left\|P_{J_\gamma}(H^{\mathcal{T}_\gamma})\delta_x\right\| \leq e^{-c\sqrt{\ell}} \text{ for any } x \notin \partial_{\ell/8}\mathcal{T} \cup \mathcal{T}_\ell; \tag{4.29}$$

(iii) (*Exponential Decay of Correlations*) *Let $\mathcal{A}_o$ be any subset of $\mathcal{T}_\gamma$, then* (*with $\mathcal{A}_o$ in* (3.3)–(3.4)) *we have*

$$\left\|\left(H^{\mathcal{T}_\gamma} - z\right)^{-1}\right\|_{c,\ell} \leq \ell^{4d+1/2}|I|^{-1}\langle Im\, z\rangle^{-1} \tag{4.30}$$

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

Proof Proposition 4.12(i) has been established earlier, and Proposition 4.12(iii) is a consequence of (4.26). This leaves us with the task of proving Proposition 4.12(ii).

Let $\{\lambda_n, \psi_n\}$ be an eigenpair for $H^{\mathbb{T}}$ in I, and let x_n be its localization center. We first check that $x_n \in \tilde{\mathcal{T}}$. Indeed, suppose that $x_n \notin \tilde{\mathcal{T}}$. Then, by the properties of the suitable cover, there exists a box $\Lambda_\ell(a) \not\subset \tilde{\mathcal{T}}$ such that $\Lambda_{\ell/4}(x_n) \subset \Lambda_\ell(a)$. Moreover, $\omega \in \Omega_G \subset \Omega_{loc,N}$ implies that $\mathbb{T}$ is $\ell/10$-localizing, so in particular

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

We can now use Lemma C.4 below to conclude

$$\sigma\left(H^{\Lambda_\ell(a)}\right) \cap 2I \ne \emptyset, \tag{4.31}$$

which means that $\Lambda_\ell(a) \subset \tilde{\mathcal{T}}$, a contradiction. This establishes (4.28), since for any $x \in \mathbb{T} \setminus \mathcal{T}_\ell$ we have $\operatorname{dist}\left(x, \tilde{\mathcal{T}}\right) \ge \ell/8$.

Let $\{\mu_n, \phi_n\}$ be an eigenpair for $H^{\mathcal{T}}$ in I. By the argument identical to the one used earlier, its localization center y_n is located either in $\tilde{\mathcal{T}}$ or in $\partial_{C\sqrt{\ell}}\mathcal{T} \subset \partial_{\ell/8}\mathcal{T}$. Hence

$$\|P_{J_\gamma}(H^{\mathcal{T}_\gamma}) - \chi_{\partial_{\ell/8}\mathcal{T}} P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\partial_{\ell/8}\mathcal{T}} - \chi_{\mathcal{T}_\ell} P_{J_\gamma}(H^{\mathcal{T}_\gamma})\chi_{\mathcal{T}_\ell}\| \le e^{-c\sqrt{\ell}}, \tag{4.32}$$

which in particular establishes (4.29). In fact, the above argument shows more, namely that, recalling the notation in Theorem 3.2(iii),

$$\left\| P_{J_\gamma}(H^{\mathcal{T}_\gamma})\delta_x \right\| \le e^{-c\sqrt{\ell}} \text{ for any } x \notin \mathcal{A}_o. \tag{4.33}$$

The latter bound will be of use to us momentarily. □

This completes the proof that $H^{\mathbb{T}}$ possesses a local gap structure in the sense defined by Theorem 3.2. Using perturbation theory, we are now going to show that $H^{\mathbb{T}}(s)$ possesses a local gap structure as well.

4.3 Proof of Theorem 3.2

It suffices to establish the assertion for $a = c_1$ as probabilities only improve as the system size decreases. We note that Proposition 4.12 is applicable here with $I = c\ell^{-\xi}$. In particular, for $\omega \in \Omega_G$, we have $\operatorname{dist}\left(E_\gamma^\pm, \sigma(H^{\mathcal{T}_\gamma})\right) \ge \Delta$. Let

$$\tilde{H}^{\mathcal{T}_\gamma}(s) := H^{\mathcal{T}_\gamma}(0) + P_{[E_\gamma^-, E_\gamma^+]}\left(H^{\mathcal{T}_\gamma}(0)\right) + \beta W(s).$$

Then, for ℓ sufficiently small

$$\sigma\left(\tilde{H}^{\mathcal{T}_\gamma}(s)\right) \cap \left([-\tfrac{\Delta}{3}, \tfrac{\Delta}{3}] + [E_\gamma^-, E_\gamma^+]\right) = \emptyset, \tag{4.34}$$

provided that $\beta < \frac{\Delta}{6}$.

For the next assertion, we recall the definition of a dilation and its norm, introduced in (3.3)–(3.4).

Lemma 4.13 *There exists* $c>0$ *such that for any* $z\in\mathbb{C}$ *with* $Re(z)=E_\gamma^\pm$ *and for any* $\beta<c\Delta\ell^{-3d}$, *we have*

$$\left\|\left(H^{\mathcal{T}_\gamma}(s)-z\right)^{-1}\right\|_{c,\ell}+\left\|\left(\tilde{H}^{\mathcal{T}_\gamma}(s)-z\right)^{-1}\right\|_{c,\ell}\leq C\ell^{3d}\Delta^{-1}\langle Im\,z\rangle^{-1}, \tag{4.35}$$

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

Proof If we denote

$$R_z^o=\left(H^{\mathcal{T}_\gamma}(0)-z\right)^{-1},\ \tilde{R}_z^o=\left(\tilde{H}^{\mathcal{T}_\gamma}(0)-z\right)^{-1},\ R_z=\left(H^{\mathcal{T}_\gamma}(s)-z\right)^{-1},$$
$$\tilde{R}_z=\left(\tilde{H}^{\mathcal{T}_\gamma}(s)-z\right)^{-1}, \tag{4.36}$$

we have

$$\left\|R_z^o\right\|_{c,\ell}\leq C\ell^{3d}\Delta^{-1}\langle Im\,z\rangle^{-1} \tag{4.37}$$

by (4.26).

We now expand R_z into the Neumann series

$$R_z=R_z^o\sum_{n=0}^\infty\beta^n\left(-WR_z^o\right)^n,$$

yielding, via (3.5),

$$\|R_z\|_{c,\ell}\leq\left\|R_z^o\right\|_{c,\ell}\sum_{n=0}^\infty\beta^n\left\|WR_z^o\right\|_{c,\ell}^n$$
$$\leq C\ell^{3d}\Delta^{-1}\langle Im\,z\rangle^{-1}\sum_{n=0}^\infty\left(\beta C\ell^{3d}\right)^n\Delta^{-n}\leq C\ell^{3d}\Delta^{-1}\langle Im\,z\rangle^{-1}, \tag{4.38}$$

provided $\beta\leq c\Delta\ell^{-3d}$.

Using (4.33), we deduce that

$$\left\|e^{c\rho_{\mathcal{A}}^\ell}P_{[E_\gamma^-,E_\gamma^+]}\left(H_o^{\mathcal{T}_\gamma}\right)\right\|\leq C\ell^d. \tag{4.39}$$

Since

$$\tilde{R}_z^o=R_z^o-P_{[E_\gamma^-,E_\gamma^+]}\left(H_o^{\mathcal{T}_\gamma}\right)R_z^o\tilde{R}_z^o, \tag{4.40}$$

we obtain, using (4.37)–(4.39) and

$$\left\|R_z^o\right\|\leq C\Delta^{-1}\langle Im\,z\rangle^{-1},\quad\left\|\tilde{R}_z^oP_{[E_\gamma^-,E_\gamma^+]}\left(H_o^{\mathcal{T}_\gamma}\right)\right\|\leq 2,$$

that

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

We now expand $\tilde{R}_z$ into the Neumann series

$$\tilde{R}_z = \tilde{R}^o_z \sum_{n=0}^{\infty} \beta^n \left(-W\tilde{R}^o_z\right)^n,$$

and repeat the argument in (4.38) to complete the proof. □

We are now ready to finish the proof. For this, we will show that conditions 3.2(i)–3.2(iii) in Theorem 3.2 hold on Ω_G, ensuring the desired probability for these events.

We first note that Theorem 3.2(i) follows from Proposition 4.12(i) (with $I = c\ell^{-\xi}$) by standard perturbation theory for allowable values of β. On the other hand, Theorem 3.2(iii) is a direct consequence of Lemma 4.13.

This leaves us with the task of proving Theorem 3.2(ii). We recall that $J_\gamma = [E^-_\gamma, E^+_\gamma]$ and set $\hat{J}_\gamma = \left[-\frac{\Delta}{8}, \frac{\Delta}{8}\right] + [E^-_\gamma, E^+_\gamma]$. We will abbreviate $P_\gamma := P_{J_\gamma}\left(H^{\mathcal{T}_\gamma}(s)\right)$ 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 (4.11) with $E_1 = E^-_\gamma$ and $E_2 = E^+_\gamma$ to write

$$P_\gamma = -(2\pi)^{-1}\int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j R_{iu+E_j}\, du. \tag{4.41}$$

We note that the integrand can be bounded, using Theorem 3.2(i), by

$$\max_{j=1,2} \left\| R_{iu+E_j} \right\| \le \Delta^{-1}\langle u\rangle^{-1}, \quad u \in \mathbb{R}. \tag{4.42}$$

Using (recall (4.36))

$$R_{iu+E_j} = \tilde{R}_{iu+E_j} - \tilde{R}_{iu+E_j} P_{J_\gamma}(H^{\mathcal{T}_\gamma}_o) R_{iu+E_j}$$

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_γ is equal to

$$(2\pi)^{-1} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} \tilde{R}_{iu+E_j} P_{J_\gamma}(H^{\mathcal{T}_\gamma}_o) R_{iu+E_j} \hat{P}_\gamma\, du. \tag{4.43}$$

Hence we can bound

$$\left\| \mathrm{e}^{\frac{c}{\sqrt{\ell}}\rho_{\mathcal{A}}} P_\gamma \right\| \leq \int_{-\infty}^{\infty} \max_j \left(\left\| \tilde{R}_{iu+E_j} \right\|_{c,\ell} \right) \left\| \mathrm{e}^{\frac{c}{\sqrt{\ell}}\rho_{\mathcal{A}}} P_{J_\gamma}(H_o^{\mathcal{T}_\gamma}) \right\| \left\| R_{iu+E_j} \hat{P}_\gamma \right\|$$
$$\leq C\ell^{4d}\Delta^{-2} \int_{-\infty}^{\infty} \langle u \rangle^{-2} du \leq C\ell^{4d}\Delta^{-2}, \quad (4.44)$$

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

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

We first observe that, due to (4.32), $\left\| \chi_{\partial_\ell \mathcal{T}} P_{J_\gamma}(H_o^{\mathcal{T}_\gamma}) \chi_{\mathcal{T}_{8\ell}} \right\| \leq e^{-c\sqrt{\ell}}$.

Next, letting $\mathcal{A}_o = (\mathcal{T}_\gamma)_{8\ell}$, we can estimate, using Lemma 4.13 and (3.5), that

$$\left\| \chi_{\partial_\ell \mathcal{T}} R^o_{iu+E_j} (W R^o_{iu+E_j})^n \chi_{\mathcal{T}_{8\ell}} \right\| \leq C^n \left\| \chi_{\partial_\ell \mathcal{T}} \mathrm{e}^{-\frac{c}{\sqrt{\ell}}\rho_{\mathcal{A}_o}} \right\| \left\| R^o_{iu+E_j} \right\|_{c,\ell}^{n+1}$$
$$\leq C^n \ell^{3dn} \Delta^{-n} \langle Im\, z \rangle^{-2} e^{-c\sqrt{\ell}}.$$

Hence

$$\left\| \chi_{\partial_\ell \mathcal{T}} \sum_{n=1}^{\infty} \beta^n R^o_{iu+E_j} (-W R^o_{iu+E_j})^n \chi_{\mathcal{T}_{8\ell}} \right\| \leq e^{-c\sqrt{\ell}} \langle Im\, z \rangle^{-2} \sum_{n=1}^{\infty} \beta^n C^n \ell^{3dn} \Delta^{-n}$$
$$\leq e^{-c\sqrt{\ell}} \langle Im\, z \rangle^{-2}.$$

Integrating over the u variable, we see that $\left\| \chi_{\partial_\ell \mathcal{T}} P_I^\gamma \chi_{\mathcal{T}_{8\ell}} \right\| \leq e^{-c\sqrt{\ell}}$ holds. Combining this bound with (4.44), we get (3.13).

The proof of (3.12) is essentially identical to the one above, and so is left out.

5 Adiabatic theory for localized spectral patches

Throughout this section we continue to work on a torus, in the setting of Theorem 3.2. To simplify the notation, we will shorthand $H(s) := H^{\mathbb{T}}(s)$ in this section.

We note that for β, ϵ, ℓ satisfying (2.11)–(2.12) and the exponents in (2.10) and (3.10), the conditions $\epsilon, \beta \ll 1$ imply that $\epsilon^{-1} e^{-c\sqrt{\ell}} \leq e^{-c\sqrt{\ell}}$ and $\epsilon/\Delta \ll 1$. We will use this repeatedly. We will also assume that $1 \geq \Delta \geq \beta > 0$ (in fact, (2.11)–(2.12) imply $\Delta \gg \beta$ for large ℓ, but this will only matter later on).

5.1 Kato's operator

In this subsection, we will consider the general adiabatic framework, keeping the notation consistent with that in (1.1). 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 5.1 (a) $\|H(s)\| \le C$ and $\left\|H^{(k)}(s)\right\| \le \beta C_k$ for $k \in \mathbb{N}$, where $H^{(k)}(s)$ stands for the k-th derivative of $H(s)$ with respect to the s variable;
(b) There exist $E_{1,2} \in \mathbb{R}$ and $\Delta > 0$ such that $\min_{s\in[0,1]} \operatorname{dist}(\sigma(H(s)), \{E_1, E_2\}) \ge 2\Delta$;
(c) $H^{(k)}(s) = 0$ for $s = \{0, 1\}$ and $k \in \mathbb{N}$.

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

$$X_A(s) = \frac{1}{2\pi} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} R_{ix+E_j}(s)\, A\, R_{ix+E_j}(s)\, dx. \tag{5.1}$$

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 5.1,

$$\max_{j=1,2} \left\| R_{ix+E_j}(s) \right\| \le \left(x^2 + \Delta^2\right)^{-1/2} \tag{5.2}$$

and consequently

$$\|X_A(s)\| \le \frac{\|A\|}{\pi} \int_{-\infty}^{\infty} \left(x^2 + \Delta^2\right)^{-1} dx \le \Delta^{-1} \|A\|. \tag{5.3}$$

Using the Leibniz rule and (C.8), it is straightforward to see that, more generally,

$$\left\| X_A^{(k)}(s) \right\| \le C_k \, \|A\|_k\,, \quad k \in \mathbb{Z}_+, \tag{5.4}$$

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

$$\|A\|_k = \sum_{j=0}^{k} \left\| A^{(j)}(s) \right\|. \tag{5.5}$$

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

$$[H(s), X_A(s)] = [P(s), A], \tag{5.6}$$

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 5.2 *Let $A(s)$ be a smooth family of operators on* $[0, 1]$. *Suppose that in addition to Assumption* 5.1, *there exists some set $\mathcal{A}$ and $M, c > 0$ such that*

$$\left\| R_{ix+E_j}(s) \right\|_{c,\ell} \le M \langle x \rangle^{-1}, \quad j = 1, 2. \tag{5.7}$$

Then,

$$\left\| \mathrm{e}^{c\rho^{\ell}_{\mathcal{A}}} X^{(1)}_{A(s)}(s) \right\| \leq C \left(\beta M^2 |\ln \Delta| + \beta M \Delta^{-1} \right) \left\| \mathrm{e}^{c\rho^{\ell}_{\mathcal{A}}} A(s) \right\| + CM |\ln \Delta| \left\| \mathrm{e}^{c\rho^{\ell}_{\mathcal{A}}} A^{(1)}(s) \right\| . \tag{5.8}$$

Proof We will suppress the s-dependence in the proof below. Using (C.8) and (3.5), we can bound

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

Using (5.7) and Assumption 5.1.*(b)*, we get (5.8). □

5.2 Nenciu's expansion

An elegant approach for the analysis of the adiabatic behavior of gapped systems was discovered by Nenciu [55]. We will use it as a starting point for our construction.

Lemma 5.3 (Nenciu's expansion) *Let $H(s)$ be a smooth family of self-adjoint operators on* $[0, 1]$ *that satisfies Assumption* 5.1. *Let $B_n(s)$ be a smooth family defined recursively as follows*: $B_0(s) = P(s)$ *and, for* $n \in \mathbb{N}$,

$$B_n(s) = \left(\bar{P}(s) X_{\dot{B}_{n-1}(s)}(s) P(s) + h.c. \right) + S_n(s) - 2P(s)S_n(s)P(s), \tag{5.9}$$

where

$$S_n(s) = \sum_{j=1}^{n-1} B_j(s) B_{n-j}(s). \tag{5.10}$$

We then have

(i)

$$\dot{B}_n(s) = -i \, [H(s), B_{n+1}(s)] \tag{5.11}$$

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

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

(iii) *We have*

$$\sup_s \left\| B_n^{(k)}(s) \right\| \leq C_{n,k} \Delta^{-n}, \quad k, n \in \mathbb{Z}_+. \tag{5.12}$$

Proof Property 5.3(i) is due to Nenciu, [55]. Property 5.3(ii) follows directly from the recursive definition of B_ns. We establish 5.3(iii) by induction:

Induction base: For $n = 0$ and an arbitrary k, the bound $\left\| B_0^{(k)}(s) \right\| \le C_k$ in 5.3(iii) 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 (5.9) k times with $n = n_o$ using the Leibniz rule and then using (5.2) and (5.4), we get that it also holds for $n = n_o$ and all $k \in \mathbb{Z}_+$. □

For localized spectral patches, we slightly modify the statement.

Lemma 5.4 *Suppose that in addition to the assumptions of Lemma 5.3, there exists some set $\mathcal{A}$ and $M, c > 0$ such that (5.7) holds. Let us also assume that*

$$\max_{s\in[0,1]} \left\| \mathrm{e}^{c\,\rho_{\mathcal{A}}^{\ell}} P(s) \right\| \le C, \qquad \max_{s\in[0,1]} \left\| H^{(k)}(s) \right\|_{c,\ell} \le C_k \beta \textit{ for } k \in \mathbb{N}. \tag{5.13}$$

Let

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

and assume that $\beta \le \nu$. Then the operators B_n defined in Lemma 5.3 satisfy

$$\left\| \mathrm{e}^{c\rho_{\mathcal{A}}^{\ell}} B_n^{(k)}(s) \right\| \le C_{n,k} \nu^{-n}, \quad k, n \in \mathbb{Z}_+. \tag{5.14}$$

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

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

$$P^{(n)} = \left(P^{n+1}\right)^{(n)} = \sum_{k_1+k_2+\cdots+k_{n+1}=n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1\le j\le n+1} P^{(k_j)}, \tag{5.15}$$

where the sum extends over all m-tuples $(k_1, \ldots, k_{n+1})$ of non-negative integers satisfying $\sum_{j=1}^{n+1} k_j = n$ (so 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, (5.7), (3.5), and Assumption (5.13), we can bound

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

We can now use (3.5) and (5.15) to deduce that

$$\begin{aligned} &\left\| \mathrm{e}^{c\,\rho_{\mathcal{A}}^{\ell}} P^{(n)} \right\| \\ &\le \sum_{k_1+k_2+\cdots+k_{n+1}=n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1\le j<j_o} \left\| P^{(k_j)} \right\|_{c,\ell} \left\| \mathrm{e}^{c\,\rho_{\mathcal{A}}^{\ell}} P \right\| \prod_{j_o\le j\le n+1} \left\| P^{(k_j)} \right\| \\ &\le \sum_{k_1+k_2+\cdots+k_{n+1}=n} \binom{n}{k_1, k_2, \ldots, k_{n+1}} \prod_{1\le j\le n+1} C_{k_j} M^{k_j} = C_n M^n, \end{aligned} \tag{5.16}$$

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

Induction step: Suppose now that the assertion holds for all $n < n_o$ and all k. Differentiating (5.9) k times with $n = n_o$ using the Leibniz rule and then using Lemma 5.2 (the assumption there is satisfied by Eq. (3.14)), we get the induction step. □

5.3 Gapped adiabatic theorem

An immediate consequence of Lemma 5.3 is

Lemma 5.5 (Gapped adiabatic theorem to all orders) *In the setting of Lemma 5.3, let $P_N(s) := \sum_{n=0}^{N} \epsilon^n B_n(s)$. 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_ϵ was defined in (2.4).

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

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

and

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

Proof By Lemma 5.3,

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

Using the fundamental theorem of calculus, we obtain

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

Using the unitarity of U_ϵ, Assumption 5.1, and Lemma 5.3(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 $P_N(1) = P(1)$. □

5.4 Adiabatic theorem for a localized spectral patch

The goal of this subsection is to prove the following assertion, which is of independent interest.

Theorem 5.6 (Local adiabatic theorem on a torus) *Suppose that the family $H(s)$ satisfies Assumption 2.2 and $H(0)$ satisfies Assumptions 2.3–2.4. Let $\mathcal{G}_\omega$ be the event that $H^{\mathbb{T}}(0)$ possesses a local gap structure for the energy interval $J = (E - 6\delta, E + 6\delta)$ in the sense of Definition 3.1. Then $\mathbb{P}(\mathcal{G}_\omega) > 1 - e^{-c\sqrt{\ell}}$. Moreover, for each*

$\omega \in \mathcal{G}_\omega$, the physical evolution $\psi_\epsilon(s)$ of each eigenvector $\psi = \psi_n$ with $E_n \in J$ given by (2.3), satisfies

$$\max_{s\in[0,1]} \left\| \bar{P}_{J_\gamma}\left(H^{\mathcal{T}_\gamma}(s)\right) \psi_\epsilon(s)\right\| \leq C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right) \tag{5.17}$$

for some γ. For any $N \in \mathbb{N}$, we can further improve (5.17) for $s = 1$:

$$\left\| \bar{P}_{J_\gamma}\left(H^{\mathcal{T}_\gamma}(1)\right) \psi_\epsilon(1)\right\| \leq C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \tag{5.18}$$

Proof of Theorem 5.6 The first part of Theorem 3.2 has already been established. We now show the second part. We first note that $\mathcal{G}$ is a subset of $\Omega_{loc,N}$, the portion of the configuration space for which $\mathbb{T}$ and all sets in $\{\mathcal{T}_\gamma\}$ are $\ell/10$-localizing, see Lemma 4.11 below. Thus, Theorem 3.2(ii) implies the existence of a patch $\mathcal{T}_\gamma$ such that $\left\|\bar{\chi}_{(\mathcal{T}_\gamma)_{8\ell}}\psi\right\| \leq e^{-c\sqrt{\ell}}$. It then follows from Lemma C.4 below, specifically (C.12), that $E \in J_\gamma$ (see also (3.11)). Let $\hat{\mathcal{T}}_\gamma = \left(\mathcal{T}_\gamma\right)_{4\ell}$ and set

$$Q_\gamma(s) = \chi_{\hat{\mathcal{T}}_\gamma} P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\chi_{\hat{\mathcal{T}}_\gamma}. \tag{5.19}$$

By Lemma C.4, specifically (C.13), we know that (5.18) holds for $s = 0$ (with $\epsilon = 0$ on the right hand side). Let $\rho := Q_\gamma(0)$ be the (truncated) initial spectral patch. Then, since

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

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

$$\|\bar{\rho}_\epsilon(s)\psi_\epsilon(s)\| \leq e^{-c\sqrt{\ell}} \tag{5.20}$$

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

$$i\epsilon\dot{\rho}_\epsilon(s) = [H(s), \rho_\epsilon(s)], \quad \rho_\epsilon(0) = \rho. \tag{5.21}$$

Therefore the result follows from □

Lemma 5.7 (i) *We can estimate*

$$\max_{s\in[0,1]} \|\rho_\epsilon(s) - Q_\gamma(s)\| \leq C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right). \tag{5.22}$$

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

$$\max_{s=\{0,1\}} \|\rho_\epsilon(s) - Q_\gamma(s)\| \leq C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \tag{5.23}$$

(ii) *In addition,*

$$\max_{s\in[0,1]} \left\|\bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}(s)) - \bar{P}_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))\,\bar{Q}_\gamma(s)\right\| \leq e^{-c\sqrt{\ell}}. \tag{5.24}$$

Remark 5.8 We note that in the proof of Theorem 5.6, the initial spectral data ψ_n can be $\psi \in Ran(P_{[}E-\delta, E+\delta)$ that satisfies $\left\| \bar{\chi}_{(\mathcal{T}_\gamma)_{8\ell}} \psi \right\| \le e^{-c\sqrt{\ell}}$ for some patch $\mathcal{T}_\gamma$.

Proof of Lemma 5.7 We suppress the s dependence in the proof below. The property (5.24) can be seen by decomposing

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

and noticing that

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

by (3.13).

Lemma 5.7(i): By our assumption, $H^{\mathcal{T}_\gamma}$ is a gapped Hamiltonian with gap Δ. Following the argument in Sect. 5.2, we denote by B_n^γ the n-th order in Nenciu's expansion and use Lemma 5.3 with $B_0^\gamma = P_{J_\gamma}(H^{\mathcal{T}_\gamma})$. We set

$$Q_{\gamma,N} := \sum_{n=0}^{N} \epsilon^n \chi_{\hat{\mathcal{T}}} B_n^\gamma \chi_{\hat{\mathcal{T}}}. \tag{5.25}$$

and proceed to show that

$$\max_s \left\|\rho_\epsilon - Q_{\gamma,N}\right\| \le C_N\left(\epsilon^N\left(\Delta^{-N} + \delta^{-2N-1}\right) + e^{-c\sqrt{\ell}}\right). \tag{5.26}$$

The result then follows immediately from (5.26) by the definition of $Q_{\gamma,N}$ and Lemma 5.3(ii)–5.3(iii) (we recall that $B_0^\gamma = P_{J_\gamma}(H^{\mathcal{T}_\gamma})$).

To get (5.26), we observe that by (5.11),

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

where we have used $H^{\mathcal{T}}(s)\chi_{\hat{\mathcal{T}}} = H(s)\chi_{\hat{\mathcal{T}}}$. We bound the second term on the second line by $C_N\epsilon^{N+1}\Delta^{-N}$ using (5.12). For the term on the third line, we note that

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

using Lemma 5.4. Putting these bounds together, we get

$$\left\|\epsilon\dot{Q}_{\gamma,N} + i[H, Q_{\gamma,N}]\right\| \le C_N\epsilon^{N+1}\Delta^{-N} + Ce^{-c\sqrt{\ell}}. \tag{5.27}$$

Finally, we observe that

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

where $U_\epsilon(t,s)$ was defined in (2.4).

Integrating over s and using (5.27), we deduce that

$$\left\| U_\epsilon(t,r) Q_{\gamma,N}(r) U_\epsilon(r,t) - Q_{\gamma,N}(t) \right\| \le \epsilon^{-1} \left(C_N \epsilon^{N+1} \Delta^{-N} + C e^{-c\sqrt{\ell}} \right), \quad (5.28)$$

We now note that $Q_{\gamma,N}(0) = \rho$, so $U_\epsilon(t,0) Q_{\gamma,N}(0) U_\epsilon(0,t) = \rho_\epsilon(t)$ by uniqueness of the solution for the IVP (5.21). Combining this with (5.28) yields (5.26). □

5.5 Adiabatic theorem for a thin spectral set near E

In preparation for the proof of Theorem 3.3, 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 3.3 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

$$Q(s) := \sum_\gamma Q_\gamma(s), \quad (5.29)$$

where the spectral patch Q_γ was defined in (5.19). Our first assertion encapsulates the basic properties of this operator.

Lemma 5.9 *For ℓ large enough, the operator $Q(s)$ satisfies the following properties:*

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

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

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

$$\|[H(s), Q(s)]\| \le C e^{-c\sqrt{\ell}}; \quad (5.30)$$

(iii) *Almost projection:*

$$\left\| \bar{Q}(s) Q(s) \right\| \le C e^{-c\sqrt{\ell}}; \quad (5.31)$$

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

$$\left\| \bar{P}_{J_+}(s) Q(s) \right\| \le C e^{-c\sqrt{\ell}}, \quad \left\| \bar{Q}(s) P_{J_-}(s) \right\| \le C e^{-c\sqrt{\ell}}. \quad (5.32)$$

Proof Lemma 5.9(i): Note that, for ℓ large enough, $\beta \ll \Delta$. The assertion follows from the integral representation (C.7) for $P_{J_\gamma}(H^{\mathcal{T}_\gamma}(s))$ with $E_{1,2} = E_\pm^\gamma$, the formula (C.8), (5.2), and the Leibniz rule.

Lemma 5.9(ii): We compute

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

and estimate both terms by $C\mathrm{e}^{-c\sqrt{\ell}}$ using Assumption 2.2 and Theorem 3.2(ii).

Lemma 5.9(iii): We note that, for disjoint sets Ω_γ,

$$\| \sum_\gamma \chi_{\Omega_\gamma} A_\gamma \chi_{\Omega_\gamma} \| \le \max_\gamma \| \chi_{\Omega_\gamma} A_\gamma \chi_{\Omega_\gamma} \|. \tag{5.33}$$

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

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

The right hand side is bounded by $Ce^{-c\sqrt{\ell}}$ using Theorem 3.2(ii).

Lemma 5.9(iv): We apply Lemma C.3 with $H_1 = H(s)$, $H_2 = H^{\mathcal{T}}(s)$, and $R = \chi_{\hat{\mathcal{T}}}$ to bound

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

where we have used (3.13) and the fact that $H(s)$ has range r. Since

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

by (3.11), we deduce that

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

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

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

Since

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

by (3.11), we deduce that

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

using (3.12) to bound the first term on the right hand side. □

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

Lemma 5.10 *Let $N \in \mathbb{N}$. Suppose that ℓ is sufficiently large. Then there exists a smooth family of projections Q_s with the following properties*:

(i)

$$\max_{s\in[0,1]} \|[Q_s, H(s)]\| \le C\left(\epsilon + e^{-c\sqrt{\ell}}\right) \tag{5.34}$$

and

$$\max_{s\in\{0,1\}} \|[Q_s, H(s)]\| \le C_N \epsilon^{N+1} \Delta^{-N} + C e^{-c\sqrt{\ell}}; \tag{5.35}$$

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

$$\max_{s\in[0,1]} \left(\left\| \bar{P}_{J_+}(s) Q_s \right\|, \left\| \bar{Q}_s P_{J_-}(s) \right\| \right) \le C\left(\epsilon \Delta^{-1} + e^{-c\sqrt{\ell}}\right) \tag{5.36}$$

and

$$\max_{s\in\{0,1\}} \left(\left\| \bar{P}_{J_+}(s) Q_s \right\|, \left\| \bar{Q}_s P_{J_-}(s) \right\| \right) \le C e^{-c\sqrt{\ell}} \tag{5.37}$$

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

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

(iv)

$$\left\| \epsilon \dot{Q}_s + i[H(s), Q_s] \right\| \le C_N \epsilon^{N+1} \Delta^{-N} + C e^{-c\sqrt{\ell}}; \tag{5.38}$$

(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*

$$\max_{s\in[0,1]} \|Q_\epsilon(s) - Q_s\| \le C_N \epsilon^N \Delta^{-N} + C e^{-c\sqrt{\ell}}. \tag{5.39}$$

Proof We set

$$Q_N(s) := \sum_{\gamma} Q_{\gamma,N}(s), \tag{5.40}$$

where $Q_{\gamma,N}$ was defined in (5.25), 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 5.3 and the hypothesis $\epsilon \le \Delta$ that

$$\|Q_N(s) - Q_0(s)\| = \|Q_N(s) - Q(s)\| \le C_N \epsilon \Delta^{-1}. \tag{5.41}$$

Hence, combining this bound with Lemma 5.9, we conclude that $Q_N(s)$ satisfies the properties 5.10(ii)–5.10(iii).

We next observe that the property 5.10(iv) holds for $Q_N(s)$ by (5.27), Assumption 2.2, and (5.33).

The property 5.10(v) is established by replicating the argument employed in the proof of Lemma 5.7(i).

Finally, the property 5.10(i) holds for $Q_N(s)$ by the properties 5.10(iii)–5.10(iv) we already established.

We now note that $Q_N(0) = Q(0)$. Hence, defining $Q_\epsilon(t) := U_\epsilon(t,0)Q(0)U_\epsilon(0,t)$, we get $\left\| Q_\epsilon(t)\bar{Q}_\epsilon(t) \right\| = \left\| Q(0)\bar{Q}(0) \right\| \le Ce^{-c\sqrt{\ell}}$ by (5.31). Thus, by the triangle inequality, we get

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

where in the last step we have used the properties 5.10(iii) and 5.10(v) for Q_N.

It follows that

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

If ϵ/Δ is small enough and ℓ large 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 $[\frac{1}{2}, \frac{3}{2}]$. Then by functional calculus for self-adjoint operators and the triangle inequality, Lemma 5.10(i), 5.10(ii), and 5.10(v) hold for this operator. To establish Lemma 5.10(iii), we use the following integral representation for Q_s:

$$Q_s = (2\pi i)^{-1} \oint_\Gamma (Q_N(s) - z)^{-1}\, dz, \quad \Gamma = \{z \in \mathbb{C} : |z - 1| = 1/2\}. \tag{5.42}$$

Since

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

and $\left\| (Q_N(s) - z)^{-1} \right\|$ is uniformly bounded for $z \in \Gamma$, the property 5.10(iii) follows by the Leibniz rule and the bounds on $Q_N^{(k)}(s)$.

Lemma 5.10(iv):

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

and the statement follows from the properties 5.10(iv) and 5.10(i) already proved for $Q_N(s)$.

For $s \in \{0, 1\}$, we have $Q_N(s) = Q(s)$, so (5.35) and (5.37) follow from Lemma 5.9. □

5.6 Adiabatic behavior of the distorted Fermi projection

The idea behind the proof of Theorem 3.3 is that, since the projection Q_s evolves adiabatically, it effectively induces a gap on its spectral support and decouples the energies separated by this induced gap.

Let $\bar{H}(s) = \bar{Q}_s H(s) \bar{Q}_s$. By Lemma 5.10, $\bar{Q}_s$ is close to a spectral projection of $H(s)$ and so the spectrum of $\bar{H}(s)$ is approximately a subset of the original spectrum and the point 0. To avoid discussing the position of 0 with respect to E, we assume without loss of generality that $E < 0$. We will need a pair of preparatory results.

Lemma 5.11 *Let $I = (E - \delta/2, E + \delta/2)$. Suppose that ℓ is large enough. Then we have $\sigma(\bar{H}(s)) \cap I = \emptyset$ for $s \in [0, 1]$. In addition, we have*

$$\max_{s \in [0,1]} \left\| \bar{H}(s)^{(k)} \right\| \le C_k \quad \textit{for} \quad k = 1, \ldots, N. \tag{5.43}$$

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

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

while we can bound

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

using Lemma 5.10(ii), and

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

using Lemma 5.10(i). Hence

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

on $Ran\left(\bar{Q}\right)$.

The bound (5.43) follows from Lemma 5.10(iii), Assumption 2.2, and the Leibniz rule. □

Lemma 5.12 *Let $T(s,s')$ be the unitary semigroup generated by $i[\dot{Q}_s, Q_s]$, i.e., $T(s,s')$ is the solution of the IVP*

$$i\partial_s T(s,s') = i[\dot{Q}_s, Q_s]T(s,s'), \quad T(s',s') = 1. \tag{5.44}$$

Then $T(s,s')$ satisfies

$$T(s,s')Q_{s'} = Q_s T(s,s'). \tag{5.45}$$

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

$$\max_s \left\|T^{(k)}(s,0)\right\| \leq C_k\beta \quad \text{for} \quad k = 1,\ldots,N. \tag{5.46}$$

Proof The interweaving relation (5.45) follows from observing that

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

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

The bound (5.46) follows from Lemma 5.10(iii), the unitarity of T, and the Leibniz rule. □

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

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

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

Lemma 5.13 *Suppose that ℓ is large enough. Then we have*

$$\max_s \left\| Q_1^+ U_\epsilon(s,0) Q_0^- \right\| \leq C\left(\epsilon\Delta^{-1} + e^{-c\sqrt{\ell}}\right) \tag{5.47}$$

and

$$\left\| Q_1^+ U_\epsilon(1,0) Q_0^- \right\| \leq C_N\left(\epsilon^N\Delta^{-N} + \epsilon^N\delta^{-2N-1}\right) + Ce^{-c\sqrt{\ell}}. \tag{5.48}$$

Proof We first note that Lemma 5.10 implies that

$$\left\| Q_s U_\epsilon(s,s')\bar{Q}_{s'} \right\| \leq C_N \epsilon^N \Delta^{-N} + C e^{-c\sqrt{\ell}}. \tag{5.49}$$

Indeed, using the semigroup property for U_ϵ,

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

and both terms on the right hand side can now be bounded using Lemma 5.10(v).

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

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

where

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

We note that

$$\begin{aligned} \|R_\epsilon(s)\| &\leq \left(\epsilon \left\|\dot{Q}_s\right\| + \|[H(s), Q_s]\|\right) \left\|Q_s U_\epsilon(s,0)\bar{Q}_0\right\| \\ &\leq C_N \epsilon\epsilon^N \Delta^{-N} + C e^{-c\sqrt{\ell}} \end{aligned} \tag{5.50}$$

by Lemma 5.10 and (5.49).

Let $W_\epsilon(s) = T(0,s)V_\epsilon(s)$, where T was defined in (5.44). Then,

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

By Lemma 5.11, the operator $\bar{H}(s)$ has a gap δ in its spectrum that separates the associated spectral projections $Q_s^\pm$. This implies that $T(0,s)\bar{H}(s)T(s,0)$ has the same gap with the associated projections given by $\mathcal{Q}_s^\pm := T(0,s)Q_s^\pm T(s,0)$. We can bound

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

using (5.43), (5.46), and the Leibniz rule.

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

$$i\epsilon\partial_s \tilde{W}_\epsilon(s) = T(0,s)\bar{H}(s)T(s,0)\tilde{W}_\epsilon(s), \quad \tilde{W}_\epsilon(0) = 1. \tag{5.51}$$

Then, it follows from our previous analysis and the Leibniz rule that $T(0,s)\bar{H}(s) \times T(s,0)$ satisfies Assumption 5.1 and the gapped adiabatic theorem to all orders, Lemma 5.5, is applicable. Hence

$$\max_s \left\| \mathcal{Q}_1^+ \tilde{W}_\epsilon(s) \mathcal{Q}_0^- \right\| \leq C\epsilon\delta^{-1}, \quad \left\| \mathcal{Q}_1^+ \tilde{W}_\epsilon(1) \mathcal{Q}_0^- \right\| \leq C_N \epsilon^N \delta^{-N}. \tag{5.52}$$

We now observe that

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

so

$$\left\| W_\epsilon(s) - \tilde{W}_\epsilon(s) \right\| \leq \epsilon^{-1} \max_{s' \leq s} \left\| R_\epsilon(s') \right\| \leq C_N \epsilon^N \Delta^{-N} + C e^{-c\sqrt{\ell}}, \tag{5.53}$$

using (5.50). We conclude that

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

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

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

where in the last step we have used (5.49). □

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 3.3. We pick $Q(s) = Q_s^-$.

Theorem 3.3(i): Using the integral representation (C.7),

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

we get

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

and we can bound

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

But

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

which yields

$$\left\| [H(s), \bar{H}(s)] \right\| \leq C_N \epsilon + C e^{-c\sqrt{\ell}}$$

by Lemma 5.10. Hence

$$\| [Q(s), H(s)] \| \leq C_N \epsilon \delta^{-1} + C e^{-c\sqrt{\ell}},$$

and 3.3(i) follows.

Theorem 3.3(ii): Using (5.36) and $Q_s^- \bar{Q}_s = Q_s^-$, we deduce that

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

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

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

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

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

Theorem 3.3(iii): This part follows directly from Lemma 5.13 and the $\pm$ symmetry in the argument there, as

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

□

6 Uniformly localized eigenfunctions for $H(s)$ and the proof of Theorem 2.8

Disclaimer: In the process of completing this paper, we learned about a recent paper [47], which has a significant thematic overlap with the results presented here.

6.1 Non-uniform bound on localization

Let H_ω be an infinite volume operator satisfying Assumptions 2.2–2.5. We will need a stronger concept of a localizing Hamiltonian than the one introduced earlier in Definition 4.3.

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

$$|\psi_i(y)|^2 \leq \frac{1}{\theta} \langle x_i \rangle^{d+1} \mathrm{e}^{-c|y-x_i|} \text{ for some } x_i \in \mathbb{Z}^d. \tag{6.1}$$

Here, the quantifier "non-uniformly" refers to the presence of the factor $\langle x_i \rangle^{d+1}$.

Theorem 6.2 (Non-uniform eigenfunction localization) *Let H_ω be an infinite volume operator satisfying Assumptions 2.2–2.5 with $m = 1$. Then*

$$\mathbb{P}\left(\{\omega \in \Omega : \ H_\omega \textit{ is non-uniformly } (c, \theta)\textit{-localizing}\}\right) \geq 1 - C\theta \tag{6.2}$$

for some $C > 0$.

Proof The assertion above follows from [3, 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) $\mathcal{N}_{J_{loc}}$ of H_o, associated with the interval J_{loc}, given by

$$\mathcal{N}_{J_{loc}} = \lim_{R\to\infty} \frac{\operatorname{tr}\left(\chi_{\Lambda_R(0)} P_{J_{loc}}(H_\omega)\right)}{R^d}, \tag{6.3}$$

is well-defined and almost surely non-random, see e.g., [3, Theorem 3.15 and Corollary 3.16]. Moreover, if $\mathcal{N}_{J_{loc}} > 0$, the convergence to the mean in (6.3) is exponentially fast, so in particular

$$\mathbb{P}\left(\frac{\operatorname{tr}\left(\chi_{\Lambda_R(0)} P_{J_{loc}}(H_o)\right)}{R^d} < \frac{\mathcal{N}_{J_{loc}}}{2}\right) \le \mathrm{e}^{-mR} \tag{6.4}$$

for some $m > 0$. This is a typical large deviations result, see e.g., [20].

We now adjust the concept of localized eigenvectors to make it uniform. We will assume here that $\mathcal{N}_{J_{loc}} > 0$.

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

$$|\psi(x)|^2 \ge |\ln\theta|^{-d-1} \text{ and } |\psi(y)| \le \frac{|\ln\theta|^{\frac{d+1}{2}}}{\theta}\mathrm{e}^{-c|y-x|}, \quad y \in \mathbb{Z}^d. \tag{6.5}$$

We will say that the orthogonal projection $P \in \mathcal{L}(\ell^2(\mathbb{Z}^d))$ is (c, θ)-*Wannier decomposable* if there exists an orthonormal basis $\{\psi_i\}$ for $Ran(P)$ such that each ψ_i is (c, θ)-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 $H_L^{\mathbb{T}}$ denote the periodic restriction of H_ω 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 B.2 below.

Theorem 6.4 *Suppose that Assumptions 2.2–2.5 hold and that in addition* $\mathcal{N}_{J_{loc}} > 0$. *For a given configuration* $\omega \in \Omega$, *let* $\mathbb{P}_E$ *denote the normalized counting measure of eigenvalues of* $H_L^{\mathbb{T}}$ *in the interval* J_{loc} *(counting multiplicities). Let* $\mathcal{G}$ *be the set*

$$\mathcal{G} := \left\{E_n \in \sigma(H_L^{\mathbb{T}}) \cap J_{loc} : \ P_{\{E_n\}} \text{ is } \left(\tfrac{c}{m}, \theta^2\right)\text{-Wannier decomposable}\right\}.$$

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

$$\mathbb{P}\left(\mathbb{P}_E(\mathcal{G}) \ge 1 - \sqrt{\theta}\right) \ge 1 - C\sqrt{\theta}. \tag{6.6}$$

Proof For a pair $(E_n, P_{\{E_n\}})$, let

$$w_n = w(\omega, P_{\{E_n\}}) = \sum_{x,y} \left| P_{\{E_n\}}(x, y) \right| \mathrm{e}^{c|x-y|}. \tag{6.7}$$

We then have, by the bound (4.6) on the eigenvector correlator and $\mathcal{N}_{J_{loc}} > 0$,

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

Letting $a, b > 0$, we have by Markov's inequality that

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

We now pick an ω such that $\mathbb{E}_E[w_n] \leq \theta^{-a}$. Another application of Markov's inequality then gives

$$\mathbb{P}_E (w_n \leq \theta^{-b}) \geq 1 - \theta^{b-a}. \tag{6.8}$$

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 2.8 Here we will use $\theta = e^{-c\sqrt{\ell}}$.

Let $\mathcal{L} = C\epsilon^{-1}$ and consider

$$\Xi_{\mathcal{L}} := \left(\tfrac{3}{2} \mathcal{L} \mathbb{Z} \right)^d, \tag{6.9}$$

cf. (4.14), and an $\mathcal{L}$-cover of $\mathbb{Z}^d$ of the form

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

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

We also cover J'_{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 $E \in J'_{loc}$ that satisfies $\operatorname{dist}\left(E, \left(J'_{loc}\right)^c\right) \geq \ell^{-\xi}$ we can find J_i such that $\operatorname{dist}\left(E, (J_i)^c\right) \geq c\ell^{-\xi}/3$;
(iii) $\cup_i J_i \subset J_{loc}$.

One can always construct such a covering using $C\ell^{\xi}$ intervals J_i for ℓ sufficiently large.

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

$$\lim_{R \to \infty} \frac{\# \Lambda_{\mathcal{L}}(a) \subset \Lambda_R : \ \mathcal{A} \text{ is satisfied for } \Lambda_{\mathcal{L}}(a)}{\# \Lambda_{\mathcal{L}}(a) \subset \Lambda_R} \geq 1 - \sqrt{\theta}. \tag{6.10}$$

For a given box $\Lambda_{\mathcal{L}}(a)$ in the cover we construct the corresponding torus $\mathbb{T}_a$ and pick any J_i from the cover of J'_{loc}. It follows that the conclusions of Theorem 5.6 are satisfied with probability $\geq 1 - e^{-c\sqrt{\ell}}$. Moreover, as the number of J_is in the cover is $C\ell^{\xi}$, we deduce that with the same probability the conclusions of Theorem 5.6 hold for *all* J_is in the cover. We next note that, given N tori $\{\mathbb{T}_a\}$, 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 2.3 and ergodicity, we obtain that the fraction $1 - e^{-c\sqrt{\ell}}$ of tori $\{\mathbb{T}_a\}_{a\in\Xi_{\mathcal{L}}}$ satisfy the conclusions of Theorem 5.6 for each interval J_i in the cover of J_{loc}.

Let $\Omega_1 \subset \Omega$ be a collection of ω such that $\mathbb{P}_E(\mathcal{G}) \geq 1 - \sqrt{\theta}$ for all $R \geq R_o$ (in particular, $\mathbb{P}\left(\Omega_1^c\right) \leq e^{-c\sqrt{\ell}}$ holds by (6.6)).

We now pick any $\omega \in \Omega_1$ and conclude from Theorem 6.4 that the fraction $1 - e^{-c\sqrt{\ell}}$ of eigenstates ψ_n for $H^{\mathbb{T}}$ with eigenvalues $E_n \in J_{loc}$ are $\left(c/m, \theta^2\right)$-localized. Let ψ be 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

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

If this box happens to be a good box, then the first assertion of Theorem 2.8 holds for all s by Theorem 3.2 while the second assertion holds for ψ at $s = 0$ by Lemma C.4 below and by the assertions of Theorem 3.2. It then follows from Theorem 5.6 (see Remark 5.8 there) that the second assertion holds for all $s \in [0, 1]$. Since the fraction of good boxes is $1 - e^{-c\sqrt{\ell}}$, we get the result. □

7 Derivation of linear response theory

In this section, we prove Theorem 1.1 assuming the setting described in Sect. 2. The proof rests on several technical results proven at the end of the section. Since the methods used here are sufficiently standard, our arguments will be somewhat abbreviated for the most part.

Proof of Theorem 1.1 In the rescaled variable $s = \epsilon t$ and for the zero temperature case ($\rho = P := P_F$, the Fermi projection at $s = -1$), (1.5) assumes form

$$\sigma_m = \beta^{-1} \int_0^1 \operatorname{tr}\left((P_\epsilon(s) - P)\, J\right) ds,$$

see Sect. 1.3.

It is a standard fact in the theory of quantum Hall effect, often referred to as "cross geometry", that the operator $(P_\epsilon(s) - P)\, J$ is supported (in an appropriate sense) around the origin. We make this precise in Lemma 7.1 and use it to show that $(P_\epsilon(s) - P)\, J$ is trace class and that we can replace the plane by a torus of linear size $\mathcal{L}$ up to exponentially small errors. Explicitly, let $\mathcal{L} = C\epsilon^{-1}$ and let $\mathbb{T}$ be a torus of linear size $\mathcal{L}$. Then we show that

$$\mathbb{E}\left(\sup_{\mathcal{B}} \left|\operatorname{tr}(P_\epsilon(s) - P)\, J - \operatorname{tr}\left(P_\epsilon^{\mathbb{T}}(s) - P^{\mathbb{T}}\right) \tilde{J}\right|\right) \leq Ce^{-c\mathcal{L}}, \tag{7.1}$$

where $P^{\mathbb{T}} = P_{E_F}(H^{\mathbb{T}})$ is a Fermi projection on the torus, $\tilde{J} = \chi_{\mathcal{B}} J$, and the supremum is taken over $\mathcal{B} \subset \mathbb{T}$ satisfying $\Lambda_{\mathcal{L}/4} \subset \mathcal{B} \subset \Lambda_{\mathcal{L}/3}$.

In the torus geometry we can apply the local adiabatic theorem. For this we fix $\epsilon = e^{-a\sqrt{\ell}}$ and $\ell = (\beta/a)^{-2p}$ with $2p < 1/p_1$ so that $\epsilon = e^{-\beta^{-p}}$. Then for a small enough (but β-independent) the assumptions of Theorem 3.3 hold, i.e. there exists an event $\mathcal{E}$ for which Theorem 3.2 (and consequently Theorem 3.3) is applicable, and $\mathbb{P}(\mathcal{E}) \geq 1 - e^{-c\sqrt{\ell}}$.

We next decompose $P^{\mathbb{T}}$ into two components $P^{\mathbb{T}} = \mathcal{Q}(-1) + R$ where $\mathcal{Q}(s)$ is the smooth adiabatic projection constructed in Theorem 3.3 (adjusted to the interval $(-1, 1)$) and $R := P^{\mathbb{T}} - \mathcal{Q}(-1)$. By Theorem 3.3 we then have that for $s \geq 0$ and $N \in \mathbb{N}$,

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

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

$$\sigma_m = \frac{1}{\beta} \mathrm{tr}((\mathcal{Q}(0)) - \mathcal{Q}(-1))\tilde{J}) + \frac{1}{\beta} \int_0^1 \mathrm{tr}(R_{\epsilon}(s) - R)\tilde{J}) ds + \mathcal{O}(e^{-a\sqrt{\ell}}). \qquad (7.2)$$

For each $\omega \in \mathcal{E}$, we will construct a suitable set $\mathcal{B} = \mathcal{B}_{\omega}$ that will be used in the analysis below. In Proposition 7.4 we will establish that for such $\mathcal{B}$ we have

$$\frac{1}{\beta} \mathrm{tr}\Big((\mathcal{Q}(0) - \mathcal{Q}(-1)) \, \tilde{J} \Big) = \sigma_H + \mathcal{O}(e^{-c\sqrt{\ell}}), \qquad (7.3)$$

where σ_H was defined in (1.4). The principle idea here is that $\mathcal{Q}$ differs from the Fermi projection by localized states that do not contribute to the Hall conductance.

Finally, in Proposition 7.5 we will show that for the same $\mathcal{B}$, the remainder can be estimated as

$$\left| \frac{1}{\beta} \int_0^1 \mathrm{tr} \Big((R_{\epsilon}(s) - R)\tilde{J} \Big) \, ds \right| \leq C L^2 \frac{\epsilon}{\beta} + e^{-c\sqrt{\ell}} \leq C e^{-a\sqrt{\ell}}. \qquad (7.4)$$

Combining the bounds (7.1)–(7.4), we obtain

$$\mathbb{E} \, |\sigma_m - \sigma_H| \leq C e^{-a\sqrt{\ell}} + C e^{-c\sqrt{\ell}} + \mathbb{P}(\mathcal{E}^c) C \mathcal{L} \leq C e^{-a\sqrt{\ell}},$$

where in the last step we used the rough deterministic estimate

$$\left| \mathrm{tr}\Big(\Big(P_{\epsilon}^{\mathbb{T}}(s) - P^{\mathbb{T}} \Big) \, \tilde{J} \Big) \right| \leq C \mathcal{L}. \qquad (7.5)$$

This completes the proof of Theorem 1.1.

The statement of Remark 1.2(iv) can be now verified as follows: We first use Remark 7.3 below to reduce the finite temperature problem to the torus, just as for the $T = 0$ case. We then use the spectral theorem for self-adjoint operators to decompose

$$\rho_T(H) = -\int_{-\infty}^{\infty} P_E \rho_T'(E) dE = -\int_{J_{loc}} P_E \rho_T'(E) dE + O(e^{-d_{\mu}/T}), \qquad (7.6)$$

where J_{loc} is the mobility gap that contains μ. Using Theorem 1.1 and the fact that $\sigma_H = \sigma_H(E)$ is almost surely ω-independent constant within J_{loc}, we deduce that

$$\mathbb{E}\left|\sigma_m + \sigma_H \int_{J_{loc}} \rho_T'(E)dE\right| \leq C\left(e^{-a\sqrt{\ell}} + \epsilon^{-1}e^{-d_\mu/T}\right).$$

But

$$\int_{J_{loc}} \rho_T'(E)dE = -1 + O(e^{-d_\mu/T}),$$

and the result follows. □

We now present the technical statements used in the proof.

Lemma 7.1 *The operator* $(P_\epsilon(s) - P)J$ *is trace class almost surely, and* (7.1) *holds.*

Remark 7.2 We note that $\tilde{J}$ is supported on a strip $|x_1| \leq r$.

Remark 7.3 If one replaces P by the Fermi-Dirac distribution $\rho_T(H)$ with $\rho_T(E) = \frac{1}{e^{(E-\mu)/T}+1}$, where T is the absolute temperature and μ is the chemical potential, then (7.1) holds *deterministically* with $c = 1/T$ for $\epsilon \ll T$.

Proof We first note that (4.9) holds with $\Theta = \mathbb{Z}^2$ as well (the argument is only a slight modification of the one used in the proof of (4.9) but is also an explicit content of [3, Theorem 13.6]). Hence we have

$$\sum_{x,y\in\mathbb{Z}^2} \langle x\rangle^{-3} e^{4c|x-y|}\mathbb{E}\,|P(x,y)| \leq C \tag{7.7}$$

for some $c > 0$. Let

$$A(\omega) := \sum_{x,y\in\mathbb{Z}^2} \langle x\rangle^{-3} e^{4c|x-y|}\,|P(x,y)|\,, \tag{7.8}$$

then it follows that $A(\omega) \in L_1(\mathbb{P})$. We will only consider configurations ω for which $A(\omega) < \infty$ (the set of full measure in Ω) from now on.

Using the fundamental theorem of calculus, we write

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

We next note that $\left\|\Lambda_2 e^{4cx_2}\chi_{x_2<0}\right\| \leq 1$ and $\left\|\bar{\Lambda}_2 e^{4cx_2}\chi_{x_2\geq 0}\right\| \leq 1$. Thus, using (7.8) together with $[\Lambda_2, P] = -[\bar{\Lambda}_2, P]$, we get

$$\left\|[\Lambda_2, P]\chi_{\{x\}}\right\| \leq 2A(\omega)\,\langle x\rangle^3 e^{-4c|x_2|}. \tag{7.9}$$

Combining (7.9) with Proposition C.5, we deduce that

$$\left\|[\Lambda_2, P]U_\epsilon(t)\chi_{\{x\}}\right\| \leq CA(\omega)\,\langle x\rangle^3 e^{-c|x_2|} \text{ for } |x_2| \geq \mathcal{L}/3. \tag{7.10}$$

Since $\left\|\chi_{\{x\}}e^{c|x_1|}J\right\| \leq C$ for all $x \in \mathbb{Z}^2$, we arrive to the bound

$$\left\|(P_\epsilon(s) - P)\,\chi_{\{x\}}J\right\| \leq CA(\omega)\,\langle x\rangle^3 e^{-c|x|} \leq A(\omega)\,e^{-c|x|} \text{ for } |x| \geq \mathcal{L}/3. \tag{7.11}$$

This bound immediately implies the first assertion of the lemma. We also observe that by the identical argument, one can also replace P and $P_\epsilon(s)$) in the equation above with $P^{\mathbb{T}}$ and $P_\epsilon^{\mathbb{T}}(s)$, respectively.

To get the second claim of the lemma, we first bound

$$\begin{aligned}&\mathbb{E}\left(\sup_{\mathcal{B}}\left|\mathrm{tr}((P_\epsilon(s) - P)\,J) - \mathrm{tr}\Big((P_\epsilon(s) - P)\,\tilde{J}\Big)\right|\right)\\ &\quad\leq \mathbb{E}\left(\sup_{\mathcal{B}}\left|\mathrm{tr}((P_\epsilon(s) - P)\,\bar{\chi}_{\mathcal{B}}J)\right|\right) \leq Ce^{-c\mathcal{L}}\end{aligned} \tag{7.12}$$

using (7.11) and $A(\omega) \in L_1(\mathbb{P})$.

The comparison between the plane and torus spectral projection will be established using the bound

$$\mathbb{E}\left\|\left(P - P^{\mathbb{T}}\right)\chi_{\Lambda_{\mathcal{L}/2}(0)}\right\| \leq e^{-c\mathcal{L}}, \tag{7.13}$$

see [30, Lemma 4.11]. Using it together with Proposition C.5 (repeatedly) in the same vein as in the proof of the first part of the assertion, we obtain

$$\mathbb{E}\left\|\left(U_\epsilon(s,0)P_E U_\epsilon(0,s) - U_\epsilon^{\mathbb{T}}(s,0)P_E^{\mathbb{T}}U_\epsilon^{\mathbb{T}}(0,s)\right)\chi_{\Lambda_{\mathcal{L}/3}}\right\| \leq e^{-c\mathcal{L}}. \tag{7.14}$$

It implies that

$$\mathbb{E}\left(\sup_{\mathcal{B}}\left|\mathrm{tr}\Big((P_\epsilon(s) - P)\,\tilde{J}\Big) - \mathrm{tr}\Big(\Big(P_\epsilon^{\mathbb{T}}(s) - P^{\mathbb{T}}\Big)\,\tilde{J}\Big)\right|\right) \leq e^{-c\mathcal{L}},$$

and the result follows.

The statement of Remark 7.3 can be verified in the similar fashion, using Proposition C.5 and quasi-locality of analytic functions for local Hamiltonians,

$$|\rho_T(y,x)| \leq C_T e^{-|x-y|/T}, \tag{7.15}$$

see e.g., [58, Corollary 5.2] for the latter property. □

We construct the suitable set $\mathcal{B}$ for the next two assertions, given $\omega \in \mathcal{E}$. Let $\mathcal{A} = \cup_\gamma \mathcal{T}_\gamma$, where the union is taken over all γ such that $\mathcal{T}_\gamma \cap \Lambda_{\mathcal{L}/4} \neq \emptyset$, and let $\mathcal{B} = \Lambda_{\mathcal{L}/4} \cup \mathcal{A}$. We note that by construction $\Lambda_{\mathcal{L}/4} \subset \mathcal{B} \subset \Lambda_{\mathcal{L}/4+L}$ and

$$\min_\gamma \operatorname{dist}\left(\partial\mathcal{B}, \hat{\mathcal{T}}_\gamma\right) \geq \ell/4 \tag{7.16}$$

(see the paragraph preceding (5.19) for notation). These two facts will be used often in the proofs below.

We will also need a set $\mathcal{X}$ defined by

$$\mathcal{X} = \left\{\hat{\mathcal{T}}_\gamma : \left\{\hat{\mathcal{T}}_\gamma \cap \{|x_j| \leq r\right\} \neq \emptyset,\ j = 1, 2\right\}. \tag{7.17}$$

We note that $|\mathcal{X}| \leq CL^2$.

Proposition 7.4 *For any $\omega \in \mathcal{E}$, the relation (7.3) holds.*

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

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

We claim that

$$\begin{aligned}
&\frac{1}{\beta}\int_{-1}^{0} \operatorname{tr}\left(\partial_r \mathcal{Q}(r)\chi_{\mathcal{B}}[H^{\mathbb{T}}(r), \Lambda_1]\right) dr \\
&\quad = \int_{-1}^{0} \dot{g}(r)\operatorname{tr}\left(\mathcal{Q}(r)[[\mathcal{Q}(r), \Lambda_1], [\mathcal{Q}(r), \Lambda_2]]\chi_{\mathcal{B}}\right) dr + \mathcal{O}(e^{-c\sqrt{\ell}}).
\end{aligned} \tag{7.18}$$

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

$$\begin{aligned}
&\int_{-1}^{0} \operatorname{tr}\left(\partial_r \mathcal{Q}(r)\chi_{\mathcal{B}}[H^{\mathbb{T}}(r), \Lambda_1]\right) dr \\
&\quad = \int_{-1}^{0} \operatorname{tr}\left(\partial_r \mathcal{Q}(r)\chi_{\mathcal{B}}[H^{\mathbb{T}}(r), \hat{\Lambda}_1(r)]\right) + \mathcal{O}(e^{-c\sqrt{\ell}}) dr \\
&\quad = \int_{-1}^{0} \operatorname{tr}\left(-[H^{\mathbb{T}}, \partial_r \mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\right) dr + \mathcal{O}(e^{-c\sqrt{\ell}}) \\
&\quad = \int_{-1}^{0} \operatorname{tr}\left([\dot{H}^{\mathbb{T}}, \mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\right) dr + \mathcal{O}(e^{-c\sqrt{\ell}}) \\
&\quad = \int_{-1}^{0} \operatorname{tr}\left([\beta\dot{g}(r)\Lambda_2, \mathcal{Q}(r)]\chi_{\mathcal{B}}\hat{\Lambda}_1(r)\right) dr + \mathcal{O}(e^{-c\sqrt{\ell}}),
\end{aligned}$$

where in the first step we have used $\mathcal{Q}(r)\partial_r\mathcal{Q}(r)\mathcal{Q}(r) = \bar{\mathcal{Q}}(r)\partial_r\mathcal{Q}(r)\bar{\mathcal{Q}}(r) = 0$ and in the third step we employed $[H^{\mathbb{T}}, \mathcal{Q}(r)] = \mathcal{O}(e^{-c\sqrt{\ell}})$ and integration by parts. We have also repeatedly used the fact that commuting $\chi_{\mathcal{B}}$ with other operators under the trace contributes $\mathcal{O}(e^{-c\sqrt{\ell}})$ by virtue of (7.16) and the location of support of the involved operators. The relation (7.18) now follows, since $\hat{\Lambda}_1 = [\mathcal{Q}(r), [\mathcal{Q}(r), \Lambda_1]]$.

The implication is that

$$\frac{1}{\beta}\mathrm{tr}\Big(\mathcal{Q}(0) - \mathcal{Q}(-1))\tilde{J}\Big)$$
$$= i\int_{-1}^{0} \dot{g}(r)\mathrm{tr}(\mathcal{Q}(r)[[\mathcal{Q}(r), \Lambda_1], [\mathcal{Q}(r), \Lambda_2]]\chi_{\mathcal{B}}) + \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.19}$$

We now define

$$\mathrm{Ind}_{\mathcal{L}}(\mathcal{Q}) = \mathrm{tr}(\mathcal{Q}[[\mathcal{Q}, \Lambda_1], [\mathcal{Q}, \Lambda_2]]\chi_{\mathcal{B}}). \tag{7.20}$$

For $\mathbb{Z}^2$ geometry without the cutoff function $\chi_{\mathcal{B}}$, the index (when it is well-defined) is known to be integer valued and additive. I.e., for orthogonal projections Q, R with a compact R, $\mathrm{Ind}_\infty(Q+R) = \mathrm{Ind}_\infty(Q) + \mathrm{Ind}_\infty(R)$, provided $Q+R$ is a projection, [9, Proposition 2.5]. The argument in [9] 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 [27]. Moreover, $\lim_{\mathcal{L}\to\infty} \mathrm{Ind}_{\mathcal{L}}(P)$ exists and we have

$$\lim_{\mathcal{L}\to\infty} \mathrm{Ind}_{\mathcal{L}}(P) = \sigma, \tag{7.21}$$

[9, Sect. 6]. In fact, using (4.9) one can readily show that

$$|\sigma - \mathrm{Ind}_{\mathcal{L}}(P)| \le e^{-c\mathcal{L}} \text{ and } \left|\mathrm{Ind}_{\mathcal{L}}(P) - \mathrm{Ind}_{\mathcal{L}}\left(P^{\mathbb{T}}\right)\right| \le e^{-c\mathcal{L}}. \tag{7.22}$$

We next observe that, although $P^{\mathbb{T}}$ and $\mathcal{Q}(-1)$ do not commute, we have $\|[P^{\mathbb{T}}, \mathcal{Q}(-1)]\| \le e^{-c\sqrt{\ell}}$. Hence there exists a pair of self-adjoint operators $\hat{P}^{\mathbb{T}}$, $\hat{\mathcal{Q}}(-1)$ such that $[\hat{P}^{\mathbb{T}}, \hat{\mathcal{Q}}(-1)] = 0$ and $\left\|P^{\mathbb{T}} - \hat{P}^{\mathbb{T}}\right\| \le e^{-c\sqrt{\ell}}$, $\left\|\mathcal{Q}(-1) - \hat{\mathcal{Q}}(-1)\right\| \le e^{-c\sqrt{\ell}}$, [41]. Moreover, applying the compression procedure used to get a projection Q_s from a near-projection $Q_N(s)$ in the proof of Lemma 5.10, without loss of generality we can assume that $\hat{P}^{\mathbb{T}}$, $\hat{\mathcal{Q}}(-1)$ are projections. Let $\check{R} = \hat{P}^{\mathbb{T}} - \hat{\mathcal{Q}}(-1)$. Since $\|\mathcal{Q}(-1)R\| \le e^{-c\sqrt{\ell}}$, we conclude that $\hat{\mathcal{Q}}(-1)\check{R} = 0$. In particular, the additivity of index is applicable for $\hat{\mathcal{Q}}(-1)$ and $\check{R}$, and yields

$$\left|\mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(-1)\right) + \mathrm{Ind}_{\mathcal{L}}\left(\check{R}\right) - \mathrm{Ind}_{\mathcal{L}}\left(\hat{P}^{\mathbb{T}}\right)\right| \le e^{-c\sqrt{\ell}}. \tag{7.23}$$

By construction, we deduce that

$$|\mathrm{Ind}_{\mathcal{L}}(Y_i) - \mathrm{Ind}_{\mathcal{L}}(Z_i)| \le e^{-c\sqrt{\ell}}, \quad i = 1, 2, 3, \tag{7.24}$$

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

$$\mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(r)\right) = \mathrm{Ind}_{\mathcal{L}}\left(\hat{\mathcal{Q}}(-1)\right) + \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.25}$$

Putting together (7.22)–(7.25), we see that the statement follows if we can show that

$$\mathrm{Ind}_{\mathcal{L}}(R) = \mathcal{O}(e^{-c\sqrt{\ell}}). \tag{7.26}$$

To establish this bound we observe that

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

where $\mathcal{X}$ was defined in (7.17), just as in the argument used in the second step above. But

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

and the right hand side is $\mathcal{O}(e^{-c\sqrt{\ell}})$ using $R^{\mathcal{X}}\left(\mathbb{1} - R^{\mathcal{X}}\right) = \mathcal{O}(e^{-c\sqrt{\ell}})$ and the cyclicity of the trace. □

Proposition 7.5 *For any $\omega \in \mathcal{E}$, the relation (7.4) holds.*

Proof of Proposition 7.5 It will be convenient to introduce a new scale $\tilde{\ell}$ in addition to ℓ, defined by the modified value for δ, namely $\tilde{\delta} = 7\delta$. We consider the operator $\tilde{Q}_s$ constructed in Lemma 5.10. The important properties of $\tilde{Q}_s$ 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 3.3(ii), we have (recall that $R = P^{\mathbb{T}} - Q(-1)$)

$$\left\| R - P_I^{\mathbb{T}} R P_I^{\mathbb{T}} \right\| \le e^{-c\sqrt{\ell}}.$$

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

$$\left\| R - \sum_{\gamma} \tilde{Q}_{-1}^{\gamma} R \tilde{Q}_{-1}^{\gamma} \right\| \le e^{-c\sqrt{\ell}}$$

and, using Lemma 5.7(i), we see that for $s \ge 0$,

$$\left\| R_\epsilon(s) - \sum_{\gamma} Q_s^{\gamma} R_\epsilon(s) Q_s^{\gamma} \right\| = \mathcal{O}\left(\epsilon^{\infty} + e^{-c\sqrt{\ell}}\right). \tag{7.27}$$

Since Q_s^{γ} is supported in $\hat{\mathcal{T}}_\gamma$ (see the paragraph preceding (5.19) for notation), it follows that, up to a small error, $R_\epsilon(s)$ is the sum of terms supported in the region $\hat{\mathcal{T}}_\gamma$. Let $\hat{U}_\epsilon$ denote the evolution generated by $H_{\mathcal{T}}(s)$, the restriction of $H^{\mathbb{T}}(s)$ to the union of all $\mathcal{T}_\gamma$. Then we have

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

thanks to (7.27) and Lemma 5.4. Thus we can approximate

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

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

Considering now any $\hat{\mathcal{T}}_\gamma \notin \mathcal{X}$ (recall (7.17)), either $\operatorname{dist}\left(\hat{\mathcal{T}}_\gamma, \left\{x \in \mathbb{Z}^2 :\ x_1 = 0\right\}\right) \geq r$, in which case

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

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

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

as the perturbation is constant in that region. Hence, using (7.27) and Lemma 5.4 again (recall that A^Θ stands for the restriction of the operator A to the set Θ),

$$\begin{aligned} \operatorname{tr}\Big((R_\epsilon(s) - R)\,\tilde{J}\Big) &= \operatorname{tr}\Big(\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,\tilde{J}\Big) + \mathcal{O}\left(\epsilon^\infty + e^{-c\sqrt{\ell}}\right) \\ &= \operatorname{tr}\Big(\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,J\Big) + \mathcal{O}\left(\epsilon^\infty + e^{-c\sqrt{\ell}}\right). \end{aligned} \tag{7.28}$$

Next we observe, using the cyclicity of the trace for a trace class operator and (7.27), Lemma 3.3(i), and Lemma 5.4 one more time, that

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

However,

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

Hence by the fundamental theorem of calculus,

$$\begin{aligned} &\frac{1}{\beta}\int_0^1 \operatorname{tr}\Big(\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,J\Big)ds \\ &\quad = \frac{\epsilon}{\beta}\operatorname{tr}\Big(\Big(\Big(\hat{R}_\epsilon(1)\Big)^{\mathcal{X}} - \Big(\hat{R}_\epsilon(0)\Big)^{\mathcal{X}}\Big)\,\Lambda_1\Big) + \mathcal{O}\left(e^{-c\sqrt{\ell}}\right), \end{aligned}$$

so we finally get

$$\left|\frac{1}{\beta}\int_0^1 \operatorname{tr}\Big(\Big(\Big(\hat{R}_\epsilon(s)\Big)^{\mathcal{X}} - R^{\mathcal{X}}\Big)\,J\Big)ds\right| \leq CL^2\frac{\epsilon}{\beta} + e^{-c\ell}. \tag{7.29}$$

□

Appendix A: Hybridization in $1D$

In this appendix, we show eigenvector hybridization 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 Δ 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 that $|x-y| = 1$. We recall that $\sigma(-\Delta) = [0,4]$. We will use a decomposition $\Delta = \sum_{x\sim y}\Gamma_{xy} - 2$, where Γ_{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_ω is a random potential, with $V_\omega(x) = \omega_x$ the i.i.d. random coupling variables distributed according to the Borel probability measure $\mathbb{P} := \otimes_{\mathbb{Z}} P_0$. We will assume that the single-site distribution P_0 is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$. We assume that the corresponding Lebesgue density μ is bounded with $\mathrm{supp}(\mu) \subset [0,1]$, and that the single-site probability density is bounded away from zero on its support. We denote the configuration space by Ω. 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 $\|H(\beta)\| \leq 6$ in our setup. We remark that $\sigma(H(0))$ is a $\mathbb{P}$-a.s. deterministic set (see e.g., [3, Theorem 3.10]), which we denote by Σ, and that $\Sigma \supset [0,5]$.

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 $\Lambda_{full} := (-\mathcal{L},\mathcal{L})$, 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 Λ_R and by l the rightmost point of Λ_L, so by construction $l \sim r$. We consider the Hamiltonians associated with these regions, $H_{full} := H^{\Lambda_{full}}$, $H := H^{\Lambda}$, $H_L := H^{\Lambda_L}$, and $H_R := H^{\Lambda_R}$, as well as the decoupled Hamiltonian H_{dec} obtained by erasing the coupling between the left and right regions, i.e. $H_{dec} = H_L + H_R = H - \Gamma_{lr} - \Gamma_{rl}$. All of these Hamiltonians a priori depend on β. Here and later, we only stress the dependence on β in some equations, and suppress the dependence in others. We will assume henceforth that $\mathcal{L}$ is large enough so that $\mathrm{supp}(W) \subset \Lambda_R$. In particular, H_L does not depend on β.

We consider an eigenvector φ_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 levels cross, i.e. $E - E_R(\beta)$ changes sign as β varies. In Sect. 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 Λ_R. Let P_{dec} be the orthogonal projection onto $Span(\varphi_L, \varphi_R)$. Let us consider the rank two operator $\mathbb{H} := P_{dec} H P_{dec}$ acting on $Ran(P_{dec})$. We note that the matrix representation for $\mathbb{H}$ with respect to the $\{\varphi_L, \varphi_R\}$ basis is given by a 2×2 matrix

$$M_\beta := \begin{pmatrix} E & gap \\ gap & E_R + \beta \langle W \rangle_{\varphi_R} \end{pmatrix} \tag{A.2}$$

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

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

(i) $\|\chi_{|x| \geq \sqrt{\mathcal{L}}/\ln \mathcal{L}} \varphi(0)\| \leq e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}$,
(ii) There exists β such that $\|\chi_{\Lambda_L} \varphi(\beta)\|^2 \geq F$, and $\|\chi_{|x| < \sqrt{\mathcal{L}}/\ln \mathcal{L}} \varphi(\beta)\|^2 \geq F$.

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

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

If we now consider an infinite volume operator $H(\beta)$ (i.e., $\Lambda_{full} = \mathbb{Z}$), any $F < \frac{1}{2}$, 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 potentially be 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 this is the only possible mechanism of hybridization, but in this work we have not tried 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.

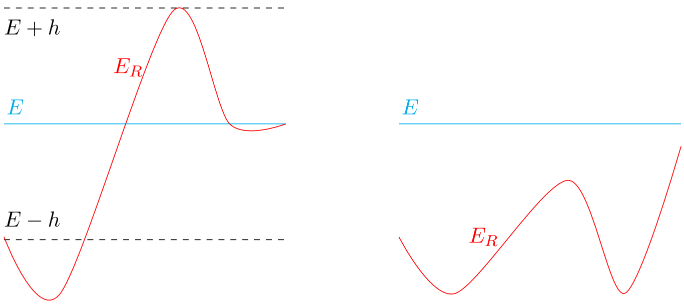

Fig. 1 The left panel shows the crossing of E and E_R (colored in cyan and red, respectively) as β varies in $(-1,1)$. The parameter $h>0$ captures the crossing width. The right panel shows the avoided crossing, $h=0$. (Color figure online)

A.1 Perturbation of a non-avoided crossing

We consider the eigenvalues $E_L \equiv E$, $E_R(\beta)$ of $H_{dec}(\beta)$ for β in a compact interval J associated with the (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 will stand 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

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

where $(x)_+$ is equal to x for positive x 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 $E_R(\beta)$, see Figure 1 for an illustration.

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

$$\|P - P_{dec}\| \le \varepsilon, \quad \|H_{dec}P_{dec} - HP\| \le \varepsilon, \tag{A.3}$$

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 $c_L^\pm, c_R^\pm$ be the Fourier coefficients of $\varphi_\pm$ with respect to the elements φ_L, φ_R of an eigenbasis of H_{dec}, i.e.,

$$\varphi_\pm = c_L^\pm \varphi_L + c_R^\pm \varphi_R + \varphi_\perp^\pm,$$

with $\langle \varphi_\perp^\pm, \varphi_L\rangle = \langle \varphi_\perp^\pm, \varphi_R\rangle = 0$, 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 $\epsilon = 0$, F equals zero by the continuity of β dependence in H, H_{dec}, P, and P_{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 ϵ. Indeed, in this example $\epsilon > 0$ corresponds to $gap > 0$ and $F = 1/2$ is achieved for β that solves $E_L = E_R + \beta\langle W\rangle_{\varphi_R}$.

Our principle indicator of hybridization will be the fact that F has to be close to $1/2$ whenever the level crossing for H_{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 $\|\varphi_\perp^+\| \leq \varepsilon$, hence

$$|c_L^+(\beta)|^2 + |c_R^+(\beta)|^2 \geq 1 - \varepsilon^2. \tag{A.4}$$

By the same equation,

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

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

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

Using the second equation in (A.3) and Weyl's theorem, [40, Theorem 4.3.1] we get

$$\mathrm{dist}_H(\{0, E_L, E_R\}, \{0, E_-, E_+\}) = \mathrm{dist}_H(\sigma(H_{dec}P_{dec}), \sigma(HP)) \leq \varepsilon, \tag{A.7}$$

where dist_H stands for the Hausdorff distance between a pair of sets. Hence,

$$\mathrm{dist}_H(\{E_L, E_R\}, \{E_-, E_+\}) \leq 2\varepsilon. \tag{A.8}$$

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

$$\begin{aligned} &\max\left(|E_R(\beta_1) - E_-(\beta_1)|, |E_L - E_+(\beta_1)|, |E_R(\beta_2) - E_+(\beta_2)|, |E_L - E_-(\beta_2)|\right) \\ &\quad \leq 2\varepsilon. \end{aligned}$$

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 $|c_R^+(\beta_2)|^2 \geq \frac{3}{4} - \varepsilon > \frac{1}{2}$ by (A.5). It follows from the continuity of the coefficient c_R^+ that there exists $\beta \in (\beta_1, \beta_2)$ such that $|c_R^+(\beta)|^2 = \frac{1-\epsilon^2}{2}$. Hence, by (A.4) we also have $|c_L^+(\beta)|^2 \geq \frac{1-\epsilon^2}{2}$, completing the proof. □

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 (ν, θ) of positive parameters, we will say that H is (ν, θ)-localized if all eigenvalues of H are simple and for each $E \in \sigma(H)$, the corresponding eigenvector ψ_E satisfies

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

We call x_E the localization center of the eigenvector ψ_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*

$$\mathbb{P}\left(\left\{\omega \in \Omega : \ H^\sharp(0) \textit{ is } (\nu,\theta)\textit{-localized}\right\}\right) \leq 1 - C\theta, \quad \sharp = \Lambda, \Lambda_L, \Lambda_R. \tag{A.10}$$

Proof This is a consequence of [3, Theorems 5.8, 7.4, and 12.11] and Markov's inequality. □

We will fix this value of ν henceforth.

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

There exists eigenvalues $E_R(0)$ (resp. E_L) of $H_R(0)$ (resp. H_L) with eigenvectors φ_L, φ_R such that

(i) H_L, $H_R(0)$ are (ν, θ)-localizing; In particular, φ_L, φ_R are localized;
(ii) $|E_L - E_R(0)| \leq b\theta/\mathcal{L}$;
(iii) Let

$$J := \{\lambda \in \mathbb{R} : \ \mathrm{dist}(\lambda, \{E_L, E_R(0)\}) \leq \sqrt{\theta}/\mathcal{L}\} \tag{A.11}$$

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_ν 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 θ is small enough. Then $E_R(\beta)$, E_L intersect for some $\beta \in I$, where*

$$I := [-a, a], \quad a = 4\frac{b}{C_\nu}\frac{\theta \ln\theta}{\mathcal{L}}, \tag{A.12}$$

and the associated function h satisfies $h \geq b\theta/\mathcal{L}$.

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

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

Since $\|H_R(\beta) - H_R(0)\| \le \beta$, by Weyl's theorem

$$\mathrm{dist}\,(E_R(\beta), \sigma(H_R(\beta)) \setminus \{E_R(\beta)\}) \ge \mathrm{dist}\,(E_R(0), \sigma(H_R(0)) \setminus \{E_R(0)\}) - 2\beta \ge \frac{\sqrt{\theta}}{2\mathcal{L}}$$

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

$$\|\dot{P}_R(\beta)\| \le \beta/\mathrm{dist}\,(E_R(\beta), \sigma(H_R(\beta)) \setminus \{E_R(\beta)\}) \le 2\beta\frac{\mathcal{L}}{\sqrt{\theta}}.$$

We now estimate

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

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

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

Using Definition A.6(ii), we see that $h \ge b\theta/\mathcal{L}$, completing the proof. □

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

Proof Let $\mathcal{C}_k$ denote the event that the property (k) with $k = i, ii, iii, iv$ in Definition A.6 holds. By (A.10), $\mathbb{P}(\mathcal{C}_i) \ge 1 - C\theta$.

If $H_R(0)$ is (ν, θ) 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 φ_R such that $|\varphi_R(0)|^2 > -C_\nu/\ln(\theta)$. As shown in Lemma C.1, (C.2) indeed holds deterministically with the choice $J = [\frac{1}{4}, \frac{15}{4}]$, $c = \frac{1}{49}$. Thus we can pick $\mathcal{C}_{iv} := \mathcal{C}_i$.

To bound $\mathbb{P}(\mathcal{C}_{ii})$ 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 $C_+ \ge C_- > 0$ such that we have*

$$C_-\,|J|\,|K| \le \mathbb{E}\left(\mathrm{tr}\left(\chi_J(H^K)\right)\right) \le C_+\,|J|\,|K|, \tag{A.13}$$

provided $|K| > L_0$.

Proof The upper bound is well known, see e.g., [3, Corollary 4.9]. The lower bound was recently established in [31, Theorem 1.1] in the continuum setting, but the same proof works for the lattice systems considered here as well. □

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*

$$\mathbb{P}\left(\operatorname{tr}\left(\chi_J(H^K)\right) \geq n\right) \leq \frac{1}{n!}\left(C_+|J||K|\right)^n . \tag{A.14}$$

Proof In this generality, the bound goes back to [21], see also [3, Theorem 17.11]. □

Let $\check{I} := [E_R(0) - b\theta/\mathcal{L}, E_R(0) + b\theta/\mathcal{L}]$. 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

$$\mathbb{P}\left(\operatorname{tr}\left(\chi_{\check{I}}(H_L)\right) \geq 1\right) \geq \mathbb{E}\left(\operatorname{tr}\chi_{\check{I}}(H_L)\right) - \sum_{n=2}^{\infty}(n-1)\,\mathbb{P}\left(\operatorname{tr}\left(\chi_{\check{I}}(H_L)\right) \geq n\right) \geq cb\theta \tag{A.15}$$

for some b-independent constant $c > 0$. This implies that $\mathbb{P}(\mathcal{C}_{ii}) \geq cb\theta$ for such b.

This leaves us with estimating $\mathbb{P}(\mathcal{C}_{iii})$. Let $\hat{J} := \{\lambda \in \mathbb{R} : |\lambda - E_R(0)| \leq 2\sqrt{\theta}/\mathcal{L}\}$. Then $J \subset \hat{J}$ for J specified in (A.11) and, using the statistical independence of H_L and $H_R(0)$, by (A.14)

$$\mathbb{P}\left(\operatorname{tr}\left(\chi_J(H_L)\right) \geq 2\right) \leq \mathbb{P}\left(\operatorname{tr}\left(\chi_{\hat{J}}(H_L)\right) \geq 2\right) \leq C\theta. \tag{A.16}$$

To complete the argument, we will use the following consequence of Theorem A.10.

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

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

Then there exists $C > 0$ such that

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

Proof This statement is essentially [44, Lemma 2], in the formulation given in [25, Lemma B.1]. □

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

$$\mathbb{P}\left(\operatorname{tr}\left(\chi_J(H_R)\right) \geq 2\right) \leq C\sqrt{\theta}/\ln^2\mathcal{L} \leq \theta \tag{A.17}$$

for $\mathcal{L}$ large enough. This yields $\mathbb{P}(\mathcal{C}_{iii}) \geq 1 - C\theta$.

Putting our bounds on $(\mathcal{C}_i)$–$(\mathcal{C}_{iv})$ together, we see that for b large enough $\mathbb{P}(\mathcal{C}) \geq cb\theta$ for some constant $c > 0$. □

A.3 Construction of the avoided crossing

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

Definition A.12 Let Λ_B be a region of size $\mathcal{L}^{1/8}$, centered at the boundary between Λ_L and Λ_R, i.e. (recall (3.1)–(3.2)) $\Lambda_B = (\partial \Lambda_R)_{\mathcal{L}^{1/8}}$. We pick $b_L, b_R \in \mathbb{Z}$ so that $\Lambda_B = (b_L, b_R)$. We denote by $H_B := H^{\Lambda_B}$ 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 $\hat{J} := (E_L - \theta^{-1}\mathcal{L}^{-1/2}, E_L + \theta^{-1}\mathcal{L}^{-1/2})$.
(ii) There are at most two eigenvalues of $H(0)$ in the interval J defined in (A.11).
(iii) The centers of φ_L and φ_R are a distance of order $\sqrt{\mathcal{L}}/\ln \mathcal{L}$ away from the boundary of Λ_R. Specifically,

$$\left\| \chi_{\{|x| > \sqrt{\mathcal{L}}/(4 \ln \mathcal{L})\}} \varphi_R \right\| + \left\| \chi_{\{x > -3\sqrt{\mathcal{L}}/\ln \mathcal{L}\}} \varphi_L \right\| \le e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}. \tag{A.18}$$

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

$$\left\| \chi_{\{|x-l| \ge \mathcal{L}^{1/8}\}} (H_B - \lambda)^{-1} \delta_l \right\| \le e^{-c\mathcal{L}^{1/8}}.$$

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

$$\left| \langle \delta_r, \left(H_B - E_\sharp\right)^{-1} \delta_l \rangle - 1 \right| \ge 2\theta^{\frac{1}{4}}.$$

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

The dependence on the parameter θ in the above definition is chosen so that $\mathbb{P}(\mathcal{A}) = O(\theta)$. We will establish this at the end of the 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. [43]). We recall that $H_L(\beta)$ is β-independent, so $\varphi_L(\beta) \equiv \varphi_L$. We first show that the analogue of (A.18) holds if we replace $\varphi_R(0)$ with $\varphi_R(\beta)$. For an interval J, we set $J_a := aJ$.

Lemma A.13 *Assume that $\omega \in \mathcal{A}$. For $\beta \in I$ defined in* (A.12),

$$\left\| \chi_{\{|x| > \sqrt{\mathcal{L}}/(2 \ln \mathcal{L})\}} \varphi_R(\beta) \right\| \le e^{-c\sqrt{\mathcal{L}}/\ln \mathcal{L}}. \tag{A.19}$$

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 $Span(\varphi_R(0))$. We observe that by Definition A.6(iii) and $|E_R(\beta) - E_R(\beta)| \le \beta$,

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

We have

$$\begin{aligned}
&\chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln\mathcal{L})\}}\varphi_R(\beta) \\
&\quad = \chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln\mathcal{L})\}}\left(\hat{H}_R(0)-E_R(\beta)\right)^{-1}\left(\hat{H}_R(0)-E_R(\beta)\right)\varphi_R(\beta) \\
&\quad = \chi_{\{|x|>\sqrt{\mathcal{L}}/(2\ln\mathcal{L})\}}\left(\hat{H}_R(0)-E_R(\beta)\right)^{-1}((1-E_R(0))P_R(0)+\beta W)\varphi_R(\beta).
\end{aligned}$$

To estimate the right hand side, we note that

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

by (A.18) and the compactness of $\mathrm{supp}(W)$. Hence (A.19) will follow once we show that

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

The latter 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 (ν,θ)-localizing for $\omega\in\mathcal{A}$. □

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

Proposition A.14 *Assume that $\omega\in\mathcal{A}$. 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 β;*
- (ii) $\mathrm{dist}(\{E_-(\beta),E_+(\beta)\},\{E_L,E_R(\beta)\})\le \mathrm{e}^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}}$;
- (iii) *Let $P(\beta)$ be the spectral projection on $E_\pm(\beta)$, then* $\|P(\beta)-P_{dec}(\beta)\|\le \mathrm{e}^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}}$;
- (iv) *We can label $E_\pm(\beta)$ so that the associated eigenfunctions $\varphi_\pm(\beta)$ satisfy*

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

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

$$\mathrm{dist}\left(\sigma(H(\beta)),E_\sharp(\beta)\right)\le \mathrm{e}^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}},\quad \sharp=L,R.$$

It follows that $H(\beta)$ has at least two eigenvalues in the interval I. Combined with standard perturbation theory and the fact that for $\omega\in\mathcal{A}$ 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) holds. The last statement follows from (A.18), Lemma A.13, and Lemma C.4. □

Proposition A.15 *Suppose that $\omega\in\mathcal{A}$, 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 $\bar{P}_{dec}(\beta)\,(H(\beta)-\lambda)\,\bar{P}_{dec}(\beta)$ is invertible on the range of $\bar{P}_{dec}(\beta)$ 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

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

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

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

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

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

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

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

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

We note that by Lemma A.16, M is well-defined for our range of λs and βs. M is a rank-two operator whose range is spanned by (φ_R, φ_L). Using the Guttman rank additivity formula, [68, Sect. 0.9], we see that $\mathrm{tr}\big(\chi_{\{\lambda\}}(H)\big) = 2$ (a sufficient and necessary condition 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

$$M_{LR} = \varphi_L(l)\varphi_R(r)\left(1 - \langle \delta_r, (H_B - E_-)^{-1}\delta_l\rangle + Error\right), \tag{A.21}$$

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

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

Hence, for sufficiently large $\mathcal{L}$, $M_{LR} \neq 0$ as the eigenfunctions of $H_{L,R}$ 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 Λ_R to the region Λ_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

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

where we denote $\bar{H} = \bar{P}H\bar{P}$, and let $\left(\bar{H}-\lambda\right)^{-1}$ denote the inverse of $\bar{H}-\lambda$ on the $Ran\left(\bar{P}\right)$. 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 $\bar{P}HP = \bar{P}\Gamma P$ to get

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

We next use the resolvent identity

$$\left(\bar{H}-\lambda\right)^{-1} = (H_B-\lambda)^{-1} + T, \quad T := \left(\bar{H}-\lambda\right)^{-1}(H_B - H + \bar{P}HP)(H_B-\lambda)^{-1}.$$

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

$$\begin{aligned}\|T\delta_l\| &\le 5\left\|\left(\bar{H}-\lambda\right)^{-1}\right\|\left\|\chi_{\{|x-l|\ge\mathcal{L}^{1/8}\}}(H_B-\lambda)^{-1}\delta_l\right\| \\ &\quad + \left\|\left(\bar{H}-\lambda\right)^{-1}\right\|\left\|P\chi_{\{|x-l|<\mathcal{L}^{1/8}\}}\right\|\left\|(H_B-\lambda)^{-1}\right\| \le Ce^{-c\mathcal{L}^{1/8}},\end{aligned}$$

which implies that

$$|\langle\delta_r, T\delta_l\rangle| \le Ce^{-c\mathcal{L}^{1/8}}.$$

Furthermore, by standard perturbation theory and Definition A.12(iii),

$$\left\|(H_B-\lambda)^{-1} - (H_B-E_-)^{-1}\right\| \le C|E_- - \lambda|\theta^2\mathcal{L}.$$

Since $E_- - \lambda$ is of order $\mathcal{L}^{-1}$ for $\lambda \in J$, we get (A.21). □

We now show

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

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

Using the upper bound in (A.13), we get $\mathbb{P}(\mathcal{A}_i) \ge \mathbb{P}(\mathcal{C}) - C\theta^{-1}\mathcal{L}^{-1/2}\mathcal{L}^{1/8} \ge cb\theta$ for $\mathcal{L}$ large enough. On the other hand, using (A.14), we deduce that

$$\mathbb{P}(\mathcal{A}_{ii} \cap \mathcal{A}_i) \ge \mathbb{P}(\mathcal{A}_i) - \mathbb{P}\left(\mathrm{tr}\chi_J\left(H(0)\right) \ge 3\right) \ge \mathbb{P}(\mathcal{A}_i) - C\theta^{3/2} \le cb\theta.$$

Let $\hat{\Lambda}_L = [-4\sqrt{\mathcal{L}}/\ln\mathcal{L}, l]$. Then, using the upper bound in (A.13), for $\mathcal{L}$ large enough,

$$\mathbb{P}\left(\operatorname{tr}\left(\chi_{\hat{J}}(H^{\hat{\Lambda}_L})\right) = 0\right) \geq 1 - C(\sqrt{\mathcal{L}}/\ln\mathcal{L})\theta^{-1}\mathcal{L}^{-1/2} \geq 1 - \theta^2.$$

Let $\mathcal{E} := \mathcal{A}_{ii} \cap \mathcal{D}$, where $\mathcal{D}$ is the event $\operatorname{tr}\left(\chi_{\hat{J}}(H^{\hat{\Lambda}_L})\right) = 0$. Then we see that

$$\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(\mathcal{A}_{ii}) - \theta^2 \geq cb\theta.$$

We claim that (A.18) holds for $\omega \in \mathcal{E}$, implying that $\mathbb{P}(\mathcal{A}_{iii} \cap \mathcal{A}_{ii}) \geq cb\theta$. Indeed, the bound $\left\|\chi_{\{|x|>\sqrt{\mathcal{L}}/(4\ln\mathcal{L})\}}\varphi_R\right\| \leq e^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}}$ 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 φ_L were located in $[-\frac{7}{2}\sqrt{\mathcal{L}}/\ln\mathcal{L}, l]$, Definition A.6(i) would imply that $\left\|\chi_{x<-4\sqrt{\mathcal{L}}/\ln\mathcal{L}}\varphi_L\right\| \leq \mathrm{e}^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}}$. But then we would have $\operatorname{dist}\left(E_L, \sigma(H^{\hat{\Lambda}_L})\right) \leq \mathrm{e}^{-c\sqrt{\mathcal{L}}/\ln\mathcal{L}}$ thanks to Lemma C.4, contradicting $\operatorname{tr}\left(\chi_{\hat{J}}(H^{\hat{\Lambda}_L})\right) = 0$. This implies that the localization center for φ_L is located in $\Lambda_L \setminus [-\frac{7}{2}\sqrt{\mathcal{L}}/\ln\mathcal{L}, l]$, 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

$$\sup_{\lambda\in\mathbb{R}} \mathbb{E}\left\|\chi_{\{|x-l|\geq\mathcal{L}^{1/8}\}}(H_B - \lambda - i0)^{-1}\delta_l\right\| \leq Ce^{-c\mathcal{L}^{1/8}},$$

[3, Theorem 12.11]. Hence, denoting

$$\mathcal{F} := \left\{\omega \in \Omega : \left\|\chi_{\{|x-l|\geq\mathcal{L}^{1/8}\}}(H_B - \lambda)^{-1}\delta_l\right\| \leq e^{-c\mathcal{L}^{1/8}}\right\},$$

we see that $\mathbb{P}(\mathcal{A}_{iv} \cap \mathcal{A}_{iii}) \geq cb\theta$ for $\mathcal{L}$ large enough by Markov's inequality.

Finally, the bound $\mathbb{P}(\mathcal{A}_v \cap \mathcal{A}_{iv}) \geq cb\theta$ is a direct consequence of □

Lemma A.18 *For a fixed $s \in (0, 1/2)$ and $\lambda \in I$, we have*

$$\mathbb{P}\left(\left\{\omega \in \Omega : \left|\langle\delta_r, (H_B - E)^{-1}\delta_l\rangle - 1\right| \geq \theta^{\frac{1}{s}}\right\}\right) \geq 1 - C_s\theta.$$

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 directly by [3, Eq. 12.7]),

$$G(l, r) = \hat{G}(l, l)G(r, r), \tag{A.22}$$

where $\hat{G}(x, y) = \langle\delta_x, \left(\hat{H}_B - \lambda\right)^{-1}\delta_y\rangle$ 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

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

to obtain

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

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

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

On the other hand, under our conditions on the probability distribution μ, we also have (see [3, Theorem 12.8]

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

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

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

from which the assertion follows by the Markov inequality. □

A.4 Proof of Theorem A.2

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

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

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

$$\varphi_+(\beta) = c_L^+(\beta)\varphi_L + c_R^+(\beta)\varphi_R + \varphi^\perp,$$

with

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

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

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

$$\liminf_{\mathcal{L}\to\infty} \mathbb{P}(\tilde{\Omega}_{F,\mathcal{L}}) > 0.$$

The assertion of Theorem A.2 is established completely analogously, by splitting Λ_{full} into Λ_L, Λ_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 Λ_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. □

Appendix B: A Wannier basis for quasi-local projections

Here, we show the existence of a (generalized) Wannier basis, consisting of exponentially localized functions, for a rank m orthogonal projection P on $\ell^2(\mathbb{Z}^d)$ that satisfies the quasi-locality property (B.4) below. The motivation for constructing such a basis is related to the fact that it allows showing the localization property (2.1) without assuming spectrum simplicity.

To illustrate the idea behind this construction, we start with the case $m = 1$.

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

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

Then, for any sufficiently small (but L-independent) θ, we have $\|\psi\|_\infty^2 \ge |\ln\theta|^{-d-1}$, and there exists $x_o \in \mathbb{T}_L$ such that

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

Proof of Lemma B.1 The second bound is an immediate consequence of the first with a (non unique, in general) choice of x_o such that $|\psi(x_o)| = \|\psi\|_\infty$, so we only need to show that $\|\psi\|_\infty^2 \ge |\ln\theta|^{-d-1}$. Let $r = r(c,\theta) > 0$ be such that

$$\sum_{y\in\mathbb{Z}^d:\,|y|>r} \mathrm{e}^{-2c|y|} \le \frac{\theta^2\|\psi\|_\infty^2}{2}.$$

In particular, for a fixed c there exists C such that we can choose $r = -C\ln\left(\theta\|\psi\|_\infty^2\right)$ for θ sufficiently small. Then by (B.1) we can bound

$$\begin{aligned}1 = \sum_{x\in\mathbb{T}_L} |\psi(x)|^2 &\le \|\psi\|_\infty^2 \sum_{\substack{x\in\mathbb{T}_L:\\|x-x_o|\le r}} 1 + \sum_{\substack{x\in\mathbb{T}_L:\\|x-x_o|>r}} \frac{\mathrm{e}^{-2c|x-x_o|}}{\|\psi\|_\infty^2\theta^2}\\ &\le \|\psi\|_\infty^2(2r+1)^d + \frac{1}{2}.\end{aligned} \tag{B.2}$$

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

$$\mathrm{e}^{-Cu^{1/d}} = \theta u^2. \tag{B.3}$$

Since $u > |\ln\theta|^{-d-1}$ for θ sufficiently small, we get $\|\psi\|_\infty^2 \geq |\ln\theta|^{-d-1}$. □

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 2.3–2.4 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., [51]. 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 $P \in \mathcal{L}(\mathcal{H})$, $\mathcal{H} = \ell^2\left(\mathbb{Z}^d\right)$ satisfies*

$$\max_{x,y\in\mathbb{Z}^d}\left(|P(x,y)|\,\mathrm{e}^{c|x-y|}\right) \leq \theta^{-1}. \tag{B.4}$$

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 $\|\psi_i\|_\infty \geq |\ln\theta|^{-d-1}$ and, for some $x_i \in \mathbb{Z}^d$,

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

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

Proof We will need some preparatory results. Using the argument identical to the one used in Lemma B.1 we obtain

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 (B.3). In particular, for θ sufficiently small, $M \geq |\ln\theta|^{-d-1}$.*

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

$$\sum_{\Lambda^c_{L/4}(0)} \mathrm{e}^{-2c|y|} \leq \theta^6 M \tag{B.5}$$

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

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

for θ sufficiently small. Consider

$$\Xi_L := \left(\tfrac{3}{2} L \mathbb{Z}\right)^d, \tag{B.7}$$

cf. (4.14), 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 $\operatorname{dist}\left(\Lambda_L^c(a), x\right) \geq L/4$.

Lemma B.4 *For L as above, let $T = \max_{a \in \Xi_L} \operatorname{tr}\left(P \chi_{\Lambda_L(a)}\right)$. Then $T \geq 1/2$ for θ sufficiently small.*

Proof Suppose in contradiction that $\operatorname{tr}\big(P\chi_{\Lambda_L(a)}\big) < 1/2$ for any $a \in \Xi_L$. Picking x_o as in the previous lemma and letting $a \in \Xi_L$ be such that $\operatorname{dist}\left(\Lambda_L^c(a), x_o\right) \geq L/4$, we have

$$\begin{aligned} M &\leq P(x_o, x_o) \sum_{y \in \Lambda_L(a)} P(y, y) + \sum_{y \in \Lambda_L^c(a)} |P(x_o, y)|^2 \\ &\leq M \sum_{y \in \Lambda_L(a)} P(y, y) + \theta^4 M < 2M/3, \end{aligned}$$

a contradiction. □

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

$$\mathcal{S} := \left\{a \in \Xi_L : \operatorname{tr}\big(P\chi_{\Lambda_L(a)}\big) \geq 1/2\right\}$$

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

Let $\mathcal{R} := \cup \Lambda_L(a)$, 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

$$P(y, y) < 2M\theta^4 \tag{B.8}$$

for θ sufficiently small. Indeed, if $y \notin \mathcal{R}$, then $\operatorname{dist}(y, \cup_{a\in\mathcal{S}} \Lambda_L(a)) \geq L/2$. In particular,

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

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 θ sufficiently small. In particular, Q is invertible as an operator on $Ran(P)$, with $Q \geq 1 - \theta^3$.*

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

$$\|\chi_{\mathcal{R}^c} P \chi_{\mathcal{R}}\|_{HS} = \sum_{y\in\mathcal{R}_1^c, x\in\mathcal{R}} |P(x,y)|^2 = \sum_{0<\mathrm{dist}(y,\mathcal{R})\le L/2, x\in\mathcal{R}} |P(x,y)|^2 + \sum_{\mathrm{dist}(y,\mathcal{R})>L/2, x\in\mathcal{R}} |P(x,y)|^2 .$$

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

$$\|\chi_{\mathcal{R}^c} P \chi_{\mathcal{R}}\|_{HS} \le \theta^3, \tag{B.9}$$

so $\left\| Q^2 - Q \right\|_{HS} \le \theta^3$ for θ sufficiently small.

We next observe that, in view of (B.4),

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

by the properties of exponential sums. Let $\bar{Q} = P - Q$. Then $\bar{Q}$ is (a) close to be a projection on $Ran(P)$ and (b) $\left|\bar{Q}(x,y)\right| \le C\theta^{-2}\mathrm{e}^{-c|x-y|}$. 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 $\bar{Q}$ is close to zero, which implies the result. Indeed, suppose in contradiction that $\bar{Q}$ is close to a non-trivial projection, i.e., $\mathrm{dist}\left(\sigma(\bar{Q}), 1\right) = O(\theta^3)$. Let $y_o \in \mathbb{Z}^d$ be such that $\bar{M} := \max \bar{Q}(x,x) = \bar{Q}(y_o, y_o)$ for some y_o which is not necessary unique. Just as in the proof of Lemma B.3, let $\bar{r} = \bar{r}(c,\theta) > 0$ be such that $\sum_{y\in\mathbb{Z}^d:\,|y|>\bar{r}} \mathrm{e}^{-2c|y|} \le \theta^4 \bar{M}^2$. In particular, there exists C such that we can choose $r = -C\ln\left(\theta^2\bar{M}\right)$ for θ sufficiently small.

Essentially repeating the argument of Lemma B.3, we have

$$\begin{aligned}
\bar{M} = \bar{Q}(y_o,y_o) &= \left(\bar{Q} - \bar{Q}^2\right)(y_o,y_o) + (\bar{Q}^2)(y_o,y_o) \\
&= O(\theta^3) + \sum_{y\in\mathbb{Z}^d} \left|\bar{Q}(y_o,y)\right|^2 \\
&= O(\theta^3) + \sum_{y\in\Lambda_r(y_o)} |Q(y_o,y)|^2 + \sum_{y\in\Lambda_r^c(x_o)} \left|\bar{Q}(y_o,y)\right|^2 \\
&\le O(\theta^3) + 3^d\bar{M}^2 r^d.
\end{aligned}$$

This yields $\bar{M} \le 3^{d+1}\bar{M}^2 r^d$, which in turn yields $\bar{M} \ge \bar{u}$, where u is implicitly given by the analogue of (B.3). Since $\bar{u} > |\ln\theta|^{-d-1}$ for θ sufficiently small, we get $\bar{M} \ge$

$|\ln\theta|^{-d-1}$. But then (B.8) implies

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

a contradiction. □

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

$$\mathrm{dist}_{i\neq j}\left(\mathcal{R}_i, \mathcal{R}_j\right) \geq L/2 \tag{B.11}$$

We now introduce the operator

$$X = \sum_{j=1}^{n} j P\chi_{\mathcal{R}_j}P, \tag{B.12}$$

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 θ sufficiently small.*

Proof For any $\lambda \in \sigma(X)$, we have

$$(X-\lambda)^2 = \sum_{j=1}^{n}(j-\lambda)^2 P\chi_{\mathcal{R}_j}P + \sum_{j\neq j'}(j-\lambda)\left(j'-\lambda\right)P\chi_{\mathcal{R}_j}P\chi_{\mathcal{R}_{j'}}P.$$

The second sum can be bounded in norm by $n^2\theta^3$ using (B.11) and (B.5), while the first one satisfies

$$\sum_{j=1}^{n}(j-\lambda)^2 P\chi_{\mathcal{R}_j}P \geq \min_j (j-\lambda)^2 Q \geq \min_j (j-\lambda)^2\left(1-\theta^3\right)$$

using Lemma B.5. But $0 \in \sigma\left((X-\lambda)^2\right)$, from which the result follows. □

The assertion of Theorem B.2 will follow from

Lemma B.7 *Let (λ, ψ_λ) be an eigenpair for X with normalized ψ_λ. Then*

$$|\psi_\lambda(x)| \leq C\theta^{-2}\mathrm{e}^{-c\,\mathrm{dist}\left(x,\mathcal{R}_{j_o}\right)}, \tag{B.13}$$

where j_o is chosen so that $|\lambda - j_o| \leq \theta$.

Proof Let

$$Y_\lambda := P\chi_{\mathcal{R}_{j_o}}P + \sum_{j\neq j_o}(j-\lambda)P\chi_{\mathcal{R}_j}P,\; Z_\lambda := P\chi_{\mathcal{R}_{j_o}}P + \sum_{j\neq j_o}(j-\lambda)^{-1}P\chi_{\mathcal{R}_j}P.$$

We have

$$Y_\lambda Z_\lambda = P + \sum_{j \neq j'} f(j, j') \left(j' - \lambda\right) P \chi_{\mathcal{R}_j} P \chi_{\mathcal{R}_{j'}} P =: P + W,$$

where $\left|f(j, j')\right| \le 2n$ for all $j \neq j'$. We have $\|W\| \le n^3\theta^3$ using (B.9). Hence by standard perturbation theory, the operator Y_λ is invertible on $Ran(P)$, with

$$Y_\lambda^{-1} = Z_\lambda (P + W)^{-1} = Z_\lambda \sum_{i=0}^{\infty} (-W)^i. \tag{B.14}$$

We now note that, analogously to (B.10),

$$|Z_\lambda(x, y)| \le C\theta^{-2} \mathrm{e}^{-c|x-y|},$$

while

$$\begin{aligned} |W(x, y)| &\le n^3 \max_{j \neq j'} \left| \sum_{z \in \mathcal{R}_j, w \in \mathcal{R}_{j'}} P(x, z) P(z, w) P(w, y) \right| \\ &\le C n^3 \theta^{-3} \mathrm{e}^{-c|x-y|/2} \max_{j \neq j'} \sum_{z \in \mathcal{R}_j, w \in \mathcal{R}_{j'}} \mathrm{e}^{-c|z-w|/2} \le \theta^2 \mathrm{e}^{-c|x-y|/2} \end{aligned}$$

using (B.11), (B.5), and (B.6). This in turn implies that

$$\left|W^i(x, y)\right| \le \theta^i \mathrm{e}^{-c|x-y|/2}, \quad i \in \mathbb{N}.$$

Using these bounds in (B.14), we deduce that

$$\left|Y_\lambda^{-1}(x, y)\right| \le C\theta^{-2} \mathrm{e}^{-c|x-y|/2}.$$

Hence we have

$$\begin{aligned} |\psi_\lambda(x)| &= \left\|\chi_{\{x\}} \psi_\lambda\right\| = \left\|\chi_{\{x\}} Y_\lambda^{-1} Y_\lambda \psi_\lambda\right\| \\ &= \left\|\chi_{\{x\}} Y_\lambda^{-1} (Y_\lambda - X + \lambda) \psi_\lambda\right\| \\ &= |1 - j_o + \lambda| \left\|\chi_{\{x\}} Y_\lambda^{-1} P \chi_{\mathcal{R}_{j_o}} P \psi_\lambda\right\| \le C\theta^{-2} \mathrm{e}^{-c \operatorname{dist}(x, \mathcal{R}_{j_o})}. \quad \square \end{aligned}$$

We are now ready to complete the proof of Theorem B.2. We pick the set $\{\psi_i\}$ to be $\{\psi_\lambda\}_{\lambda \in \sigma(X)}$, which is an orthonormal basis for $Ran(P)$ since X is hermitian. Since

$$\max_j \operatorname{diam}(\mathcal{R}_j) \le 2mL = -mC \ln \theta,$$

picking some $x_j \in \mathcal{R}_j$, we have

$$\mathrm{e}^{-c\,\mathrm{dist}(x,\mathcal{R}_{j_o})} \leq \mathrm{e}^{-c(|x-x_j|-2mL)} \leq \mathrm{e}^{-c|x-x_j|/m} \text{ for } |x-x_j| \geq 3mL.$$

On the other hand, since $|\psi(x)| \leq 1$ for all x, we can pick c sufficiently small so that

$$\mathrm{e}^{-c|x-x_j|/m} \geq \theta^2 \text{ for } |x-x_j| < 3mL,$$

and the assertion follows. □

Appendix C: Auxiliary results

Lemma C.1 *Let $H = -\Delta + V_\omega$ be the random operator on $\ell^2(\mathbb{Z})$ with V_ω that satisfies assumptions introduced in Appendix A. Let $J = [\frac{1}{4}, \frac{15}{4}]$ and $c = \frac{1}{49}$. Then*

$$\sum_{E\in\sigma(H)\cap J} |\psi_E(y)|^2 \geq c, \quad y \in \mathbb{Z}, \tag{C.1}$$

and the same bound holds for any Dirichlet restriction H^Λ of H.

Proof Let $P_J := P_J(H)$. Suppose in contradiction that $\mathrm{tr}(\chi_{\{y\}} P_J) < c$ for some $y \in \mathbb{Z}$. Then we have

$$\mathrm{tr}\Big(\chi_{\{y\}}(H-2)^2\Big) \geq \mathrm{tr}\Big(\chi_{\{y\}}(H-2)^2 \bar{P}_J\Big) \geq \tfrac{49}{16}\,\mathrm{tr}\big(\chi_{\{y\}}\bar{P}_J\big) > 3.$$

However, the left hand side can be computed explicitly: $\mathrm{tr}\big(\chi_{\{y\}}(H-2)^2\big) = 2 + V_\omega^2(y) \leq 3$, a contradiction. The proof for H^Λ is identical. □

Theorem C.2 *Assume that H is (ν,θ)-localized on $\mathbb{Z}$ and that there exists $c > 0$ and a compact interval J such that*

$$\sum_{E\in\sigma(H)\cap J} |\psi_E(y)|^2 \geq c, \quad y \in \mathbb{Z}. \tag{C.2}$$

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^Λ, 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

$$\sum_{\substack{y\in\mathbb{Z}:\\ |y-x_E|\geq\frac{1}{2}(|x_E|+L)}} |\psi_E(y)|^2 \leq \frac{\langle x_E\rangle^2}{\theta} \sum_{\substack{y\in\mathbb{Z}:\\ |y-x_E|\geq\frac{1}{2}(|x_E|+L)}} \mathrm{e}^{-\nu|y-x_E|}$$

$$= \frac{\langle x_E \rangle^2}{\theta} \sum_{\substack{u \in \mathbb{Z}: \\ |u| \geq \frac{1}{2}(|x_E|+L)}} \mathrm{e}^{-\nu|u|}$$

$$= \frac{\langle x_E \rangle^2}{\theta} \mathrm{e}^{-\frac{\nu}{2}(L+|x_E|)} \frac{2}{1-\mathrm{e}^{-\nu}} \leq \frac{C_\nu}{\theta} \mathrm{e}^{-\frac{\nu}{2}(L+|x_E|)} \tag{C.3}$$

for some $C_\nu > 0$.

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

$$\sum_{y \in \mathbb{Z}} |\psi_E(y)|^2 = 1, \quad E \in \sigma(H). \tag{C.4}$$

Hence, using (C.2) and (C.3), there exists $K_\nu > 0$ such that

$$4L + 1 \geq \sum_{|y| \leq 2L} \sum_{E \in \sigma(H) \cap J} |\psi_E(y)|^2 \geq \sum_{|y| \leq 2L} \sum_{\substack{E \in \sigma(H): \\ |x_E| \leq L}} |\psi_E(y)|^2$$

$$= \sum_{\substack{E \in \sigma(H) \cap J: \\ |x_E| \leq L}} \left(1 - \sum_{|y| > 2L} |\psi_E(y)|^2 \right)$$

$$\geq \#\{E \in \sigma(H) \cap J : |x_E| \leq L\} \left(1 - \frac{C_\nu}{\theta} \mathrm{e}^{-\frac{\nu}{2}L} \right)$$

$$\geq \frac{1}{2} \#\{E \in \sigma(H) \cap J : |x_E| \leq L\} \tag{C.5}$$

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

This bound together with (C.3) imply that for $L \geq K_\nu |\ln \theta|$ we have

$$\sum_{|y| \leq L} \sum_{\substack{E \in \sigma(H) \cap J: \\ |x_E| > 3L}} |\psi_E(y)|^2 \leq \sum_{k=4}^{\infty} \#\{E \in \sigma(H) \cap J : |x_E| \leq kL\} \frac{C_\nu}{\theta} \mathrm{e}^{-\frac{\nu k L}{2}}$$

$$\leq \frac{9 C_\nu}{\theta} L \sum_{k=4}^{\infty} k \mathrm{e}^{-\frac{\nu k L}{2}} < \frac{c}{2} \tag{C.6}$$

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

Using this estimate, we get

$$c \leq \sum_{E \in \sigma(H) \cap J} |\psi_E(0)|^2 \leq \sum_{\substack{E \in \sigma(H) \cap J: \\ |x_E| \leq 3L}} |\psi_E(0)|^2 + \frac{c}{2},$$

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

$$\frac{c}{2} \leq \sum_{\substack{E \in \sigma(H) \cap J: \\ |x_E| \leq 3L}} |\psi_E(0)|^2,$$

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

$$|\psi_E(0)|^2 \geq \frac{c}{26L} = \frac{-C_\nu}{\ln \theta}, \quad |x_E| \leq \frac{-\ln \theta}{C_\nu}. \qquad \square$$

Let H be a self-adjoint operator. Here we will often use the integral representation

$$P_{[E_1, E_2]}(H) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{j=1}^{2} (-1)^j \left(H - ix - E_j\right)^{-1} dx, \tag{C.7}$$

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

$$\frac{d}{ds}\left(H(s) - ix - E_j\right)^{-1} = -\left(H(s) - ix - E_j\right)^{-1} \dot{H}(s) \left(H(s) - ix - E_j\right)^{-1} \tag{C.8}$$

holds. Furthermore, for any operator R, we have

$$[R, \frac{1}{H - z}] = -\frac{1}{H - z}[R, H]\frac{1}{H - z}. \tag{C.9}$$

Lemma C.3 *Let H_1, H_2, R be bounded operators on $\ell^2(\Lambda)$, with H_1, H_2 self-adjoint. Let $J = [E_1, E_2]$ and denote by $J_{2\Delta}$ for $\Delta > 0$, the widened interval $J + [-2\Delta, 2\Delta]$. Suppose that for some ϵ_1, ϵ_2,*

(i) $\|(H_1 - H_2) R\| = \epsilon_1$
(ii) $\|[H_2, R] P_J(H_2)\| \leq \epsilon_2$.

Then

$$\left\| \bar{P}_{J_\Delta}(H_1) R P_J(H_2) \right\| \leq \frac{\epsilon_1 + \epsilon_2}{\Delta}.$$

Proof Let $z_1 = E_1 - \Delta + ix$, $z_2 = E_2 + \Delta + ix$ and write

$$G_{i,j} = \left(H_i - z_j\right)^{-1}.$$

We first establish the identity

$$\bar{P}_{J_\Delta}(H_1) R P_J(H_2) = \frac{1}{2\pi} \sum_{j=1}^{2} (-1)^j \int_{-\infty}^{\infty} \bar{P}_{J_\Delta}(H_1) G_{1,j} [H_2, R] G_{2,j} P_J(H_2) dx$$

$$+\frac{1}{2\pi}\sum_{j=1}^{2}(-1)^j\int_{-\infty}^{\infty}\bar{P}_{J_\Delta}(H_1)G_{1,j}\,(H_2-H_1)\,R\,G_{2,j}\,P_J(H_2)dx.$$

Indeed, we start from

$$G_{1,j}[H_2,R]G_{2,j}=G_{1,j}(H_2-H_1)RG_{2,j}+RG_{2,j}+G_{1,j}R.$$

Upon multiplying with $(-1)^j$, summing over $j=1,2$, integrating over x, and using (C.7) with $[E_1,E_2]$ replaced by $[E_1-\Delta,E_2+\Delta]$, we get the desired identity. We next bound

$$\max_{j=1,2}\left\|\bar{P}_{J_\Delta}(H_1)G_{1,j}\right\|\le\frac{1}{\sqrt{x^2+\Delta^2}},\qquad\max_{j=1,2}\left\|G_{2,j}P_J(H_2)\right\|\le\frac{1}{\sqrt{x^2+\Delta^2}}$$

to get

$$\left\|\bar{P}_{J_\Delta}(H_1)RP_J(H_2)\right\|\le(\epsilon_1+\epsilon_2)\,\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{x^2+\Delta^2}=\frac{\epsilon_1+\epsilon_2}{\Delta}.\qquad\square$$

For the next lemma, we will use the notation $J_a(\mu)=[\mu-a,\mu+a]$, and will let $P^{\Theta}_{J_a(\mu)}$ denote the spectral projection of H_o^{Θ} onto $J_a(\mu)$.

Lemma C.4 *Let Φ and Θ, with $\Phi\subset\Theta$, be finite subsets of $\mathbb{Z}^d$. Let (ϕ,μ) be an eigenpair for H_o^{Φ}. Then we have*

$$\operatorname{dist}\left(\mu,\sigma(H_o^{\Theta})\right)\le C\,|\partial_r\Phi|\,\left\|\chi_{\partial_r\Phi}\phi\right\|_{\infty},\tag{C.10}$$

and

$$\operatorname{dist}\left(\phi,Ran\left(P^{\Theta}_{J_a(\mu)}\right)\right)\le\frac{C}{a}\,|\partial_r\Phi|\,\left\|\chi_{\partial_r\Phi}\phi\right\|_{\infty}.\tag{C.11}$$

Conversely, if (ψ,λ) is an eigenpair for H^{Θ}, then

$$\operatorname{dist}\left(\lambda,\sigma(H_o^{\Phi})\right)\le C\,|\Theta\setminus\Phi|\,\left\|\chi_{\Theta\setminus\Phi}\psi\right\|_{\infty}\tag{C.12}$$

and

$$\operatorname{dist}\left(\phi,Ran\left(P^{\Phi}_{J_a(\lambda)}\right)\right)\le\frac{C}{a}\,|\Theta\setminus\Phi|\,\left\|\chi_{\Theta\setminus\Phi}\psi\right\|_{\infty}.\tag{C.13}$$

Proof We have

$$\left(\left(H_o^{\Theta}-\mu\right)\phi\right)(y)=\begin{cases}\sum_{\substack{y'\in\Phi:\\|y-y'|\le r}}H_o(y,y')\phi(y') & \text{if } y\in\Theta\setminus\Phi \text{ and } \operatorname{dist}(y,\Phi)\le r,\\ 0 & \text{otherwise.}\end{cases}\tag{C.14}$$

It follows that

$$\left\|\left(H_o^{\Theta}-\mu\right)\phi\right\|\le C\,|\partial_r\Phi|\,\left\|\chi_{\partial_r\Phi}\phi\right\|_{\infty}.\tag{C.15}$$

Thus, recalling that ϕ is normalized,

$$\operatorname{dist}\left(\mu, \sigma(H_o^{\Theta})\right) \leq \left\| \left(H_o^{\Theta} - \mu\right) \phi \right\| \leq C \left|\partial_r \Phi\right| \left\| \chi_{\partial_r \Phi} \phi \right\|_{\infty}. \tag{C.16}$$

On the other hand, we have

$$\left\| \bar{P}^{\Theta}_{J_a(\mu)} \phi \right\| \leq \left\| \bar{P}^{\Theta}_{J_a(\mu)} \left(H_o^{\Theta} - \mu\right)^{-1} \right\| \left\| \left(H_o^{\Theta} - \mu\right) \phi \right\| \leq \frac{C}{a} \left\| \chi_{\Theta \setminus \Phi} \psi \right\|_{\infty}, \tag{C.17}$$

from which the second assertion of the lemma follows.

Similar considerations yield

$$\left\| \left(H_o^{\Phi} - \lambda\right) \phi \right\| \leq C \left|\Theta \setminus \Phi\right| \left\| \chi_{\Theta \setminus \Phi} \phi \right\|_{\infty}, \tag{C.18}$$

which in turn imply the bounds (C.12)–(C.13). □

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

Proposition C.5 (The finite speed of propagation bound) *Let $\mathbb{T}$ be a torus of linear size R and let $U_\epsilon(s)$, $U_\epsilon^{\mathbb{T}}(s)$ be the dynamics generated by $H(s)$ and $H^{\mathbb{T}}(s)$, respectively, i.e.,*

$$i\epsilon \partial_s U_\epsilon(s) = H(s) U_\epsilon(s), \quad U_\epsilon(0) = 1; \tag{C.19}$$

$$i\epsilon \partial_s U_\epsilon^{\mathbb{T}}(s) = H^{\mathbb{T}}(s) U_\epsilon^{\mathbb{T}}(s), \quad U_\epsilon^{\mathbb{T}}(0) = 1. \tag{C.20}$$

Then there exists $c > 0$ such that for any $\mathcal{L}$ satisfying $\mathcal{L} \geq C/\epsilon$ we have

$$\max_s \left| (U_\epsilon^{\sharp}(s))(y, x) \right| \leq \mathrm{e}^{-c|x-y|}, \quad \textit{for } |x - y| \geq \frac{\mathcal{L}}{4}, \tag{C.21}$$

where $U_\epsilon^{\sharp}$ is either U or $U^{\mathbb{T}}$.

Proof This is a standard fact for (local) lattice Hamiltonians, see e.g., the proof of [27, Lemma 5] for the time-independent case (which extends to the time-dependent one without effort), or, for a more general approach, [50]. □

Acknowledgements We are grateful to Gian Michele Graf for helpful discussions. We would also like to thank the referees for providing insightful comments and suggestions that helped to improve this paper.

Funding W.D.R. was supported in part by the Fonds Wetenschappelijk Onderzoek under grant G098919N. A.E. and M.F. were supported in part by the National Science Foundation under grant DMS-1907435. A.E. was supported in part by the Simons Fellowship in Mathematics Grant 522404.

Declarations

Competing Interests We declare no competing interests.

References

1. Abanin, D.A., De Roeck, W., Huveneers, F.: Theory of many-body localization in periodically driven systems. Ann. Phys. **372**, 1–11 (2016)
2. Aizenman, M., Graf, G.M.: Localization bounds for an electron gas. J. Phys. A **31**, 6783 (1998)
3. Aizenman, M., Warzel, S.: Random Operators. vol. 168. AMS, Providence (2015)
4. Avron, J.E., Elgart, A.: Adiabatic theorem without a gap condition. Commun. Math. Phys. **203**, 445–463 (1999)
5. Avron, J.E., Seiler, R.: Quantization of the Hall conductance for general, multiparticle Schrödinger Hamiltonians. Phys. Rev. Lett. **54**, 259 (1985)
6. Avron, J.E., Seiler, R., Simon, B.: Homotopy and quantization in condensed matter physics. Phys. Rev. Lett. **51**, 51 (1983)
7. Avron, J., Seiler, R., Yaffe, L.: Adiabatic theorems and applications to the quantum Hall effect. Commun. Math. Phys. **110**, 33–49 (1987)
8. Avron, J., Howland, J., Simon, B.: Adiabatic theorems for dense point spectra. Commun. Math. Phys. **128**, 497–507 (1990)
9. Avron, J.E., Seiler, R., Simon, B.: Charge deficiency, charge transport and comparison of dimensions. Commun. Math. Phys. **159**, 399–422 (1994)
10. Bachmann, S., Fraas, M.: On the absence of stationary currents. Rev. Math. Phys. **33**, 2060011 (2021)
11. Bachmann, S., Bols, A., De Roeck, W., Fraas, M.: Quantization of conductance in gapped interacting systems. In: Annales Henri Poincaré., vol. 19, pp. 695–708. Springer, Berlin (2018)
12. Bachmann, S., De Roeck, W., Fraas, M.: The adiabatic theorem and linear response theory for extended quantum systems. Commun. Math. Phys. **361**, 997–1027 (2018)
13. Bachmann, S., De Roeck, W., Fraas, M., Lange, M.: Exactness of linear response in the quantum Hall effect. In: Annales Henri Poincaré., vol. 22, pp. 1113–1132. Springer, Berlin (2021)
14. Barequet, R., Barequet, G., Rote, G.: Formulae and growth rates of high-dimensional polycubes. Combinatorica **30**, 257–275 (2010)
15. Bellissard, J., van Elst, A., Schulz-Baldes, H.: The noncommutative geometry of the quantum Hall effect. J. Math. Phys. **35**, 5373–5451 (1994)
16. Born, M., Fock, V.: Beweis des adiabatensatzes. Z. Phys. **51**, 165–180 (1928)
17. Bornemann, F.: Homogenization in Time of Singularly Perturbed Mechanical Systems. Springer, Berlin (1998)
18. Bouclet, J.-M., Germinet, F., Klein, A., Schenker, J.H.: Linear response theory for magnetic Schrödinger operators in disordered media. J. Funct. Anal. **226**, 301–372 (2005)
19. Bourgain, J., Wang, W.-M.: Anderson localization for time quasi-periodic random Schrödinger and wave equations. Commun. Math. Phys. **248**, 429–466 (2004)
20. Carmona, R., Lacroix, J.: Spectral Theory of Random Schrödinger Operators. Springer, Berlin (2012)
21. Combes, J.-M., Germinet, F., Klein, A.: Generalized eigenvalue-counting estimates for the Anderson model. J. Stat. Phys. **135**, 201–216 (2009)
22. del Rio, R., Makarov, N., Simon, B.: Operators with singular continuous spectrum: II. Rank one operators. Commun. Math. Phys. **165**, 59–67 (1994)
23. Dietlein, A., Elgart, A.: Level spacing and Poisson statistics for continuum random Schrödinger operators. J. Eur. Math. Soc. **23**, 1257–1293 (2021)
24. Ducatez, R., Huveneers, F.: Anderson localization for periodically driven systems. Ann. Henri Poincaré **18**, 2415–2446 (2017)
25. Elgart, A., Klein, A.: An eigensystem approach to Anderson localization. J. Funct. Anal. **271**, 3465–3512 (2016)
26. Elgart, A., Schlein, B.: Adiabatic charge transport and the Kubo formula for Landau-type Hamiltonians. Commun. Pure Appl. Math. **57**, 590–615 (2004)
27. Elgart, A., Graf, G., Schenker, J.: Equality of the bulk and edge Hall conductances in a mobility gap. Commun. Math. Phys. **259**, 185–221 (2005)
28. Elgart, A., Tautenhahn, M., Veselić, I.: Anderson localization for a class of models with a sign-indefinite single-site potential via fractional moment method. Ann. Henri Poincaré **12**, 1571–1599 (2011)
29. Elgart, A., Shamis, M., Sodin, S.: Localisation for non-monotone Schrödinger operators. J. Eur. Math. Soc. **16**, 909–924 (2014)
30. Elgart, A., Pastur, L., Shcherbina, M.: Large block properties of the entanglement entropy of free disordered fermions. J. Stat. Phys. **166**, 1092–1127 (2017)

31. Gebert, M.: A lower Wegner estimate and bounds on the spectral shift function for continuum random Schrödinger operators. J. Funct. Anal. **277**, 108284 (2019)
32. Germinet, F., Klein, A., Schenker, J.H.: Dynamical delocalization in random Landau Hamiltonians. Ann. Math. **166**, 215–244 (2007)
33. Giuliani, A., Mastropietro, V., Porta, M.: Universality of the Hall conductivity in interacting electron systems. Commun. Math. Phys. **349**, 1107–1161 (2017)
34. Gordon, A.Y.: Pure point spectrum under 1-parameter perturbations and instability of Anderson localization. Commun. Math. Phys. **164**, 489–505 (1994)
35. Green, M.S.: Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids. J. Chem. Phys. **22**, 398–413 (1954)
36. Greenblatt, R.L., Lange, M., Marcelli, G., Porta, M.: Adiabatic Evolution of Low-Temperature Many-Body Systems (2022). ArXiv preprint arXiv:2211.16836
37. Hastings, M.B., Michalakis, S.: Quantization of Hall conductance for interacting electrons on a torus. Commun. Math. Phys. **334**, 433–471 (2015)
38. Henheik, J., Teufel, S.: Justifying Kubo's formula for gapped systems at zero temperature: a brief review and some new results. Rev. Math. Phys. **33**, 2060004 (2021)
39. Hislop, P.D., Krishna, M.: Eigenvalue statistics for random Schrödinger operators with non rank one perturbations. Commun. Math. Phys. **340**, 125–143 (2015)
40. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (2012)
41. Kachkovskiy, I., Safarov, Y.: Distance to normal elements in C^*-algebras of real rank zero. J. Am. Math. Soc. **29**, 61–80 (2016)
42. Kato, T.: On the adiabatic theorem of quantum mechanics. J. Phys. Soc. Jpn. **5**, 435–439 (1950)
43. Kato, T.: Perturbation Theory for Linear Operators. 132. Springer, Berlin (2013)
44. Klein, A., Molchanov, S.: Simplicity of eigenvalues in the Anderson model. J. Stat. Phys. **122**, 95–99 (2006)
45. Klein, A., Lenoble, O., Müller, P.: On Mott's formula for the ac-conductivity in the Anderson model. Ann. Math. **166**, 549–577 (2007)
46. Klitzing, K.v., Dorda, G., Pepper, M.: New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. Phys. Rev. Lett. **45**, 494 (1980)
47. Klopp, F., Schenker, J.: On the spatial extent of localized eigenfunctions for random Schrödinger operators. Commun. Math. Phys. **394**, 679–710 (2022)
48. Kubo, R.: Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. J. Phys. Soc. Jpn. **12**, 570–586 (1957)
49. Last, Y.: Quantum dynamics and decompositions of singular continuous spectra. J. Funct. Anal. **142**, 406–445 (1996)
50. Lieb, E.H., Robinson, D.W.: The finite group velocity of quantum spin systems. Commun. Math. Phys. **28**, 251–257 (1972)
51. Marcelli, G., Moscolari, M., Panati, G.: Localization implies Chern triviality in non-periodic insulators. Ann. Henri Poincaré **24**, 895–930 (2023)
52. Marconi, U.M.B., Puglisi, A., Rondoni, L., Vulpiani, A.: Fluctuation–dissipation: response theory in statistical physics. Phys. Rep. **461**, 111–195 (2008)
53. Monaco, D., Teufel, S.: Adiabatic currents for interacting fermions on a lattice. Rev. Math. Phys. **31**, 1950009 (2019)
54. Nakano, F., Kaminaga, M.: Absence of transport under a slowly varying potential in disordered systems. J. Stat. Phys. **97**, 917–940 (1999)
55. Nenciu, G.: Linear adiabatic theory. Exponential estimates. Commun. Math. Phys. **152**, 479–496 (1993)
56. Niu, Q., Thouless, D.J., Wu, Y.-S.: Quantized Hall conductance as a topological invariant. Phys. Rev. B **31**, 3372 (1985)
57. Panati, G., Spohn, H., Teufel, S.: Effective dynamics for Bloch electrons: peierls substitution and beyond. Commun. Math. Phys. **242**, 547–578 (2003)
58. Remling, C.: Finite propagation speed and kernel estimates for Schrödinger operators. Proc. Am. Math. Soc. **135**, 3329–3340 (2007)
59. Simon, B.: Fifteen Problems in Mathematical Physics. Perspectives in Mathematics, vol. 423. Birkhäuser, Basel (1984)
60. Simon, B.: Cyclic vectors in the Anderson model. Rev. Math. Phys. **6**, 1183–1185 (1994)
61. Soffer, A., Wang, W.-M.: Anderson localization for time periodic random Schrödinger operators. Commun. Partial Differ. Equ. **28**, 333–347 (2003)

62. Streda, P.: Theory of quantised Hall conductivity in two dimensions. J. Phys. C **15**, L717 (1982)
63. Tenuta, L., Teufel, S.: Effective dynamics for particles coupled to a quantized scalar field. Commun. Math. Phys. **280**, 751–805 (2008)
64. Teufel, S.: Non-equilibrium almost-stationary states and linear response for gapped quantum systems. Commun. Math. Phys. **373**, 621–653 (2020)
65. Thouless, D., Kohmoto, M., Nightingale, M., Den Nijs, M.: Quantized Hall conductance in a two-dimensional periodic potential. Phys. Rev. Lett. **49**, 405 (1982)
66. van Kampen, N.G.: The case against linear response theory. Phys. Norv. **5**, 279 (1971)
67. Wilkinson, M.: Statistical aspects of dissipation by Landau-Zener transitions. J. Phys. A **21**, 4021 (1988)
68. Zhang, F.: The Schur Complement and Its Applications. 4. Springer, Berlin (2006)