The topology of mobility-gapped insulators

Jacob Shapiro

Received: 27 November 2018 / Revised: 15 December 2019 / Accepted: 8 July 2020 / Published online: 18 July 2020
© Springer Nature B.V. 2020

Abstract

Studying deterministic operators, we define a topology on the space of mobility-gapped insulators such that topological invariants are continuous maps into discrete spaces, we prove that this is indeed the case for the integer quantum Hall effect, and lastly we show why our “insulator” condition makes sense from the point of view of the localization theory using the fractional moments method.

Keywords Topological insulators · Strong disorder · Integer quantum Hall effect · Random Schrödinger operators · Mobility gap

Mathematics Subject Classification 82C10 · 37B25 · 37H05

1 Introduction

Topological insulators [18] are usually studied in physics by assuming translation invariance, which allows for a topological description of Hamiltonians in terms of continuous maps from the Brillouin torus \( \mathbb{T}^d \to X \) where \( X \) is some smooth manifold which depends on the symmetry class under consideration (for example, for a system in class A (no symmetry) gapped after \( n \) levels, \( X = \text{Gr}_n(\mathbb{C}^\infty) \), the Grassmannian manifold). Such a description is extremely convenient because one may immediately apply classical results from algebraic topology, for example, that the set of homotopy classes \( \left[ \mathbb{T}^d \to \text{Gr}_n(\mathbb{C}^\infty) \right] \) is isomorphic to \( \mathbb{Z} \). This approach led to many classification results [26,27], which culminated in Kitaev’s periodic table of topological insulators [21], all the while ignoring the fact that the systems to be analyzed are actually not translation invariant, and in fact certain physical features of the phenomena demand strong disorder. Indeed, the plateaus of the integer quantum Hall effect (IQHE henceforth) are explained only when assuming the Fermi energy lies in a region of
localized states (*the mobility gap regime*) which cannot appear in a translation-invariant system.

Hence, a physically more realistic description calls for understanding *disordered* systems in which Bloch decomposition cannot be used. This has been done for the IQHE in [6] by applying ideas from non-commutative geometry and later generalized to the Kitaev table in [7,22] (and references therein). One problem with the application of non-commutative geometry is that it still required the Fermi energy to be placed in a spectral gap, which is why [6] goes beyond the C*-algebra generated by continuous functions of the Hamiltonian by defining a so-called non-commutative Sobolev spaces. Such an approach still uses crucially the translation invariance of the system; in contrast to the studies in physics, however, translation is used in terms of the probability distributions defining a random model. That is, a whole statistical ensemble of models is considered simultaneously, proofs use the covariance property, and statements are made either almost surely or about disorder averages. Furthermore, sometimes certain statements could also not be extended to these bigger Sobolev spaces, including the bulk-edge correspondence or the definition of edge invariants (which requires a certain regularization, see [11, Eq. (1.2)]).

Further explorations of the mobility gap regime using only one particular deterministic realization (i.e., *without* referring to a statistical ensemble) and without using covariance were pioneered in [11] for the IQHE and extended in [15,30] for chiral and Floquet topological systems, respectively. These studies demonstrate that some topological properties do not need translation invariance nor statistical averaging and should be associated with one particular mobility-gapped realization rather than a whole ensemble. These studies, however, do not use the algebraic framework of non-commutative geometry [6], and thus inherently topological features have been hard to establish. Such features include, for example, defining the ambient space in which the Hamiltonians live or establishing local constancy of the invariants w.r.t. deformations of the Hamiltonians. However, it is precisely these topological features that define what it means for a material to be a topological insulator, which motivated the present study still within the framework of deterministic operators.

In this note, we have the modest goal of continuing the deterministic line of research of [11,15,30] by defining a space of insulators and proving that a topological invariant is locally constant with respect to it. We use the IQHE as a case study since we understand it best, though future studies for other cases of topological insulators are certainly interesting. For example, chiral one-dimensional systems may relate to an extension of Fredholm theory to operators without closed range, but obey a localization estimate instead, which allows to salvage the local constancy of an integer-valued index.

This note is organized as follows. We begin by giving precise definition of what we mean by a “topological insulator” in our deterministic setting, define a topology on this space, and discuss its various properties. In the next section, we take up the IQHE as a case study and establish the deterministic local constancy. In the last section, we discuss why our deterministic definition of “insulator” makes sense for probabilistic models which exhibit localization about the Fermi energy.

In regard to existing literature, the question of the appropriate topology for topological insulators has already been raised in previous papers, see the discussion in [12, Appendix D] and [20, Introduction]. The mobility gap continuity has also been dealt
with before, in the context of probabilistic covariant models, see [24, Proposition 5.2] and references therein.

Since previous results of similar direction predate our result, one might ask why should one bother to prove the constancy of topological invariants in the deterministic mobility gap regime, which makes the analysis so much more complicated. Furthermore, it is of course judicious from the physical perspective to employ a random model in the analysis of macroscopic systems due to the inherently unknown nature of external perturbations, so that insisting on the deterministic setting calls for an explanation. Our main motivation was a philosophical one, rather than practical. Indeed, the kind of topology defined in the sequel is of little use in applications due to the fact that deterministic localization is highly unstable, and moreover, we do not currently have concrete examples of a one-parameter family of models which exhibits dynamical localization (and hence stability of its invariant). Rather, our perspective is as follows. When a striking experimental phenomenon (namely, the macroscopic quantization of the Hall conductivity) is discovered, it is worthwhile (in our opinion) to understand and identify exactly what kind of mathematical structure it relies on. Hence, we can at least say: the constancy of the Hall conductivity in the mobility gap regime is a purely topological phenomenon that does not rely on statistical averaging nor on covariance.

2 Deterministic topological insulators

Let $d \in \mathbb{N}$, the space dimension, $N \in \mathbb{N}$, the internal number of degrees of freedom, be given and fixed. We define our Hilbert space as $\mathcal{H} := \ell^2(\mathbb{Z}^d) \otimes \mathbb{C}^N$, and for an operator $A \in B(\mathcal{H})$, $A_{xy} \equiv \langle \delta_x, A \delta_y \rangle$ is an $N \times N$ matrix with $(\delta_x)_{x \in \mathbb{Z}^d}$ being the position basis of $\ell^2(\mathbb{Z}^d)$. The norm $\| \cdot \|$ is either a matrix norm on $\text{Mat}_N(\mathbb{C})$ or the 1-norm on $\mathbb{Z}^d$ induced by that of $\mathbb{R}^d$ and $\mathbb{Z}^d \subset \mathbb{R}^d$. Below, we also use $\| \cdot \|_1$ as the trace-class norm on operators.

We next define a metric on $B(\mathcal{H})$.

**Definition 2.1** For any $A, B \in B(\mathcal{H})$, define the local distance between them as

$$d_\ell(A, B) := \inf \left\{ t \in (0, \infty) \mid \exists C, \mu \in (0, \infty) \text{ s.t.} \right.$$  

$$t = \max(\{C, \mu^{-1}\}) \text{ and } \|(A - B)_{xy}\| \leq C e^{-\mu \|x-y\|} \forall x, y \in \mathbb{Z}^d \right\}. \tag{2.1}$$

This distance measures not only how close-by position-basis matrix elements are, but also their rate of off-diagonal exponential decay. The metric is mainly used as follows:

**Lemma 2.2** If $d_\ell(A, B) \in (0, \infty)$, then $\|(A - B)_{xy}\| \leq d_\ell(A, B) e^{-\mu \|x-y\|}/d_\ell(A, B) \forall x, y \in \mathbb{Z}^d$.

**Proof** By the approximation property of the infimum, we have $\forall \varepsilon > 0$ some $C_\varepsilon, \mu_\varepsilon \in (0, \infty)$ such that $d_\ell(A, B) \leq \max(C_\varepsilon, \mu_\varepsilon^{-1}) < d_\ell(A, B) + \varepsilon$ and $\|(A - B)_{xy}\| \leq C_\varepsilon e^{-\mu_\varepsilon \|x-y\|} \forall x, y \in \mathbb{Z}^d$. This implies that $C_\varepsilon \leq d_\ell(A, B) + \varepsilon$ and $\mu_\varepsilon \geq (d_\ell(A, B) + \varepsilon)^{-1}$. Hence,
\[(A - B)_{xy} \leq (d_\ell(A, B) + \varepsilon) e^{-\|x-y\|/(d_\ell(A, B) + \varepsilon)} \quad (2.2)\]

for all \(\varepsilon > 0\) which implies the result. \(\square\)

It is of course comforting to know that

**Lemma 2.3** \(d_\ell\) is a metric on \(B(\mathcal{H})\).

**Proof** By definition, it is obvious that \(d_\ell\) is nonnegative and symmetric.

Next, we want that \(d_\ell(A, A) = 0\). The second requirement in the set under the infimum in (2.1) becomes trivial for \(d_\ell(A, A) = 0\), so that taking \(C > 0\) arbitrarily small and \(\mu\) fixed, for example, we reach a zero infimum.

If we assume that \(d_\ell(A, B) = 0\), we want that \(A = B\). This assumption implies that \(\forall \varepsilon > 0 \exists C_\varepsilon, \mu_\varepsilon \in (0, \infty)\) with \(\max\{C_\varepsilon, \mu_\varepsilon^{-1}\} < \varepsilon\) and \(\|(A - B)_{xy}\| \leq C_\varepsilon e^{-\mu_\varepsilon \|x-y\|}\), which implies \(\|(A - B)_{xy}\| \leq C \leq \varepsilon\). This means \(A = B\).

Finally, we get to the triangle inequality. Let \(A, B, C \in B(\mathcal{H})\) be given. We have by the usual triangle inequality for the matrix norm that \(\|(A - B)_{xy}\| \leq \|(A - C)_{xy}\| + \|(C - B)_{xy}\|\). By Lemma 2.2, we then get

\[
\|(A - B)_{xy}\| \leq d_\ell(A, C) e^{-\|x-y\|/d_\ell(A, C)} + d_\ell(C, B) e^{-\|x-y\|/d_\ell(C, B)}
\]

\[
\leq (d_\ell(A, C) + d_\ell(C, B)) e^{-\|x-y\|/(\min(d_\ell(A, C), d_\ell(C, B)))}
\]

which means that \(d_\ell(A, B) \leq \max\{d_\ell(A, C) + d_\ell(C, B), \min(d_\ell(A, C), d_\ell(C, B))\}\) = \(d_\ell(A, C) + d_\ell(C, B)\), so we are finished. \(\square\)

**Remark 2.4** \(d_\ell\) is unfortunately not homogeneous, so it cannot induce a norm (it is translation invariant though). This is because it measures also the rate of exponential decay. Compare this with the local norm of [17] which has a fixed rate of decay.

**Remark 2.5** To require off-diagonal exponential decay is probably stronger than what is necessary for the topological invariants to be well-defined and continuous (as we see below). However, in the interest of keeping the calculations somewhat simpler, we prefer to stipulate one concrete form of off-diagonal decay. This means that the topology induced by \(d_\ell\) on the space insulating Hamiltonians (to be defined below) is finer than the initial topology corresponding to the topological invariant considered as a map from insulating Hamiltonians into \(\mathbb{Z}\). Since a continuous map may stop being so if the topology of its domain is made coarser, that means that using the topology induced by \(d_\ell\), a priori, we might not detect all path-connected components on the space of topological insulators (which might not be the correct object to study in any case, as the space of components may not form a group). Since in this paper we anyway do not concern ourselves with calculating the space of path-connected components of topological insulators, we ignore this problem.

**Lemma 2.6** If \(A_n \to A\) in the topology induced by \(d_\ell\), then \(A_n \to A\) in the norm operator topology.
**Proof** Recall Holmgren’s bound, Lemma 4.1,

\[ \|A\| \leq \max_{x \leftrightarrow y} \sup_{y \in \mathbb{Z}^d} \sum_{x \in \mathbb{Z}^d} \|A_{xy}\| \]

which together with Lemma 2.2 implies that

\[ \|A_n - A\| \leq d_l(A_n, A) \left( \coth \left( \frac{1}{2d_l(A_n, A)} \right) \right)^d. \]  (2.3)

We conclude by noting that \( \coth(t) \to 1 \) as \( t \to \infty \).

This shows also that \( A_n \to A \) in \( d_l \) does not imply the same in trace norm.

**Definition 2.7** A Hamiltonian \( H \) is a self-adjoint operator in \( B(\mathcal{H}) \) such that \( d_l(0, H) < \infty \).

Without loss of generality, we assume onward that the Fermi energy is always placed at zero (if this is not the case, replace the Hamiltonian by a shifted one).

**Definition 2.8** An insulator is a Hamiltonian \( H \) such that there is some open interval \( \Delta \subseteq \Delta_0 \ni 0 \) such that:

- All eigenvalues of \( H \) within \( \Delta \) are uniformly of finite degeneracy:
  \[ \sup_{\lambda \in \Delta} \|\chi_{\{\lambda\}}(H)\|_1 < \infty. \]

- With \( B_1(\Delta) \), the set of all Borel bounded functions \( f : \mathbb{R} \to \mathbb{C} \) which obey \( |f(\lambda)| \leq 1 \) for all \( \lambda \in \mathbb{R} \) and which are constant below and above \( \Delta \) (possibly with different constants) (see [11]), there is a constant \( \mu > 0 \) such that for any \( a \in \ell^1(\mathbb{Z}^d) \), there is a constant \( C < \infty \) such that
  \[ \sup_{f \in B_1(\Delta)} \|f(H)_{xy}\| \leq C|a(x)|^{-1} e^{-\mu \|x-y\|} \quad (x, y \in \mathbb{Z}^d). \]  (2.4)

The space of all insulators where the objects involved \( \Delta, \mu, C \) in estimate (2.4) is uniformly bounded by some fixed given worst objects \( \Delta_0, \mu_0, C_0 \) is denoted by \( \mathcal{I} \equiv I(\Delta_0, \mu_0, C_0) \), i.e., we have \( |\Delta_0| > 0, \mu_0 > 0, \) and \( C_0 < \infty \) (for all \( a \in \ell^1 \)) and for any insulator \( H \in \mathcal{I}(\Delta_0, \mu_0, C_0) \) we have: \( |\Delta| \geq |\Delta_0|, \mu \geq \mu_0, \) and \( C \leq C_0 \) (for all \( a \in \ell^1 \)). \( \mathcal{I} \) is a subset of \( B(\mathcal{H}) \). We give it the subspace topology induced by the metric topology from \( d_l \) on the space of all Hamiltonians.

This definition is not new. It encompasses essentially the same constraints as in [11, Eqs. (1.1)–(1.3)], has been used already in [15,30], and represents almost-sure consequences of a probabilistic model that exhibits strong dynamical localization (in the sense of [1]) about zero energy. See further remarks in Sect. 3. We note that in [30, Def 2.4], we required that estimate (2.4) hold for all \( a \in \ell^1 \), whereas in [11, Eq. (1.2)] one only asks for some particular \( a \in \ell^1 \) (polynomial). The former is of course also
a deterministic implication of dynamical localization, albeit a stronger one, and one could do with the fixed polynomial rate here as well.

We note that it is probably false that $\mathcal{I}$ is an open subset with respect to the topology induced by $d_\ell$. Indeed, in [13], it is observed that Anderson localization (in the sense of pure point spectrum, and hence Definition 2.8 which implies pure point spectrum within $\Delta$) of “generic” models breaks down by a rank-1 perturbation with arbitrarily small norm; see also [8].

One could hope to get rid of the uniform objects $\Delta_0, \mu_0, C_0$ on which $\mathcal{I}$ depends. This might be possible, but probably requires further specification of the details of the models considered and how their randomness arises (i.e., the probability distributions). In order to avoid this specificity, we use these rather unsatisfying uniform bounds. They allow us to conclude that if we have a sequence $(H_n)_n \subset \mathcal{I}$ such that $d_\ell(H_n, H) \to 0$ for some fixed $H \in \mathcal{I}$, the corresponding objects in the localization estimates $\Delta_n, \mu_n, C_n$ cannot explode. In analogy to the spectral gap regime, this is tantamount to assuming not only that there is a gap, but the whole collection of operators we consider has a uniform gap, which is unnecessary for spectrally gapped systems as the resolvent set is open in $\mathbb{C}$. Stated differently, the set of invertible operators is open in $B(H)$. In turn, this uniform restriction means we cannot detect the path-connected components of topological phases of insulators.

Another possible direction is to abandon Definition 2.8 and instead define an insulator as such a Hamiltonian whose Fermi projection has a well-defined topological invariant. The subspace of such Hamiltonians with the topology induced by the metric $d_\ell$ could be an open subset. There is some reason to believe this could be true, because even though infinitesimal rank-1 perturbations break localization in the spectral sense, they cannot destroy the estimates on moments of the time-evolved position operator too drastically (see [8, Theorem 8.1]).

Later on, we prove that

**Lemma 2.9** If $H$ is an insulator as above, then for any compact subinterval $\Delta' \subset \Delta$, there is some $s \in (0, 1)$, $b \in \ell^1(\mathbb{Z}^d)$ such that for all $\alpha \in \mathbb{N}$, we have some $D_\alpha < \infty$ with

$$\sup_{\eta \neq 0} \int_{E \in \Delta'} \|G(x, y; E + i \eta)\|^s \, dE \leq D_\alpha |b(x)|^{-1}(1 + \|x - y\|)^{-\alpha} \quad (x, y \in \mathbb{Z}^d)$$

(2.5)

where $G(x, y; z) \equiv R(z)_{xy} \equiv (H - z \mathbb{1})^{-1}_{xy}$ is the Greens function associated with $H$. Furthermore, one may choose $b(x) = (1 + \|x\|)^{-\nu}$ for some $\nu \in \mathbb{N}_{>d}$.

While this is anyway an almost-sure consequence of the probabilistic fractional moment condition (see Proposition 3.1), one can prove this directly from Definition 2.8. We will crucially use Lemma 2.9 to measure the proximity of spectral projections of two insulators.
2.1 The integer quantum Hall effect

We take the integer quantum Hall effect as a convenient case study for topological insulators, since much is known about it already; we rely mainly on \[11\]. For the IQHE, one takes \(d = 2\), and the topological invariant, physically the transversal (Hall) conductivity, is given by the Kubo formula as the Chern number

\[
\mathcal{I} \ni H \mapsto \text{Ch}(H) := \frac{2\pi i}{\epsilon} \text{tr} \epsilon_{\alpha\beta} PP_{\alpha} P_{\beta} \in \mathbb{Z}.
\]  \hspace{1cm} (2.6)

One finds in \[11\] a proof of the non-trivial fact that \(\text{Ch}\) is a well defined, \(\mathbb{Z}\)-valued map. Here, \(\epsilon_{\alpha\beta}\) is the antisymmetric tensor (with Einstein summation), \(A_{\alpha} := -i[\Lambda_{\alpha}, A]\) is the non-commutative derivative of an operator \(A\) with \(\Lambda_{\alpha} := \Lambda(X_{\alpha})\), \(X_{\alpha}\) the position operator in direction \(\alpha\) on \(\ell^2(\mathbb{Z}^2)\), and \(\Lambda : \mathbb{Z} \rightarrow \mathbb{R}\) is any \textit{switch function} (in the sense of \[10\], that is, any measurable interpolation from zero on negative values to one on positive values with bounded variation)—the choice of \(\Lambda\) does not influence the value of \(\text{Ch}\), but is fixed once and for all. Finally, \(P := \chi(-\infty,0)(H)\) is the Fermi projection associated with \(H\).

Remark 2.10 The formula in (2.6) is equivalent to the same formula where one replaces the switch functions with the position operators themselves and the trace with the trace per unit area (this equivalence is established without recourse to disorder averaging, e.g., in \[29\, Eq. (18)\]). In translation-invariant systems, after Bloch decomposition, the commutators with position operators become derivatives in momentum space and this formula then reduces to the usual integral (in momentum space) of the Brillouin 2-torus over the Berry curvature, as in, e.g., \[18\, Eq. (2)\].

Our main result is

\textbf{Theorem 2.11} The map \(\text{Ch} : \mathcal{I} \rightarrow \mathbb{Z}\) is continuous.

Consequently, since \(\text{Ch}\) is \(\mathbb{Z}\)-valued, it is locally constant. This gives a concrete criterion to be able to tell when two Hamiltonians will have the same Chern number, without having to actually calculate it. As noted in Remark 2.5, it is not the weakest possible criterion.

Since the Chern number of a Hamiltonian is defined through its associated Fermi projection, one would naively hope to bound \(d_\ell(P, P')\) by \(d_\ell(H, H')\). Not only does this turn out not to work, but the very definition of an insulator shows that we cannot even hope to have \(\|P_{xy}\|\) decaying in \(\|x - y\|\) uniformly. It is even false if we relax the condition to merely asking that \(\|(P - P')_{xy}\|\) is small and has some off-diagonal decay and diagonal blowup. Indeed, the problem is that \(\chi(-\infty,0)\) is not a continuous function, and we are considering operators that precisely have spectrum near zero, so even considering just the diagonal element \(\|(P - P')_{xx}\|\) for fixed \(x\), an arbitrarily small change from \(H\) to \(H'\) could make an eigenvalue jump over zero so that \(\|(P - P')_{xx}\|\) is one.

The way out is to mimic the probabilistic approach (which cures things by looking at averages, which has the effect of smoothening discontinuities), with a trick of averaging over the Fermi energy within the gap. Indeed, the point is that even though \(\chi(-\infty,0)\) is not a continuous function, \(\frac{1}{2\varepsilon} \int_{[-\varepsilon,\varepsilon]} \chi(-\infty,\lambda) \, d\lambda\) is (equal to a continuous...
interpolation of a step function with linear interpolation of width $2\varepsilon$ about zero). Our main effort below is to make this intuitive argument rigorous.

The Fermi energy averaging is permitted using the following key result:

**Proposition 2.12** ([11, Proposition 2]) Let $H \in \mathcal{I}$, i.e., there is some $\Delta \in \text{Open}(\mathbb{R})$ such that $0 \in \Delta$ and for which the mobility gap estimates are fulfilled as in Definition 2.8. Then, the following map is constant:

$$\Delta \ni E_F \mapsto \text{Ch}(H - E_F \mathbb{1}) \in \mathbb{Z}.$$ 

We can then formulate precisely in which sense is $P - P'$ small given that $H - H'$ is small:

**Proposition 2.13** Let $H, H' \in \mathcal{I}$ with $d_\ell(H, H') < \infty$. Let $\Delta'$ be a compact interval contained in the localization estimate interval (Definition 2.8) of both $H$ and $H'$.

Define

$$P_H := \chi_{(-\infty, 0)}(H - \lambda \mathbb{1}) = \chi_{(-\infty, \lambda)}(H).$$

Then, we have some $s \in (0, 1)$, $C < \infty$, $a \in \ell^1(\mathbb{Z}^d)$ (dependent on $H, H'$) such that for all $\alpha \in \mathbb{N},$

$$\int_{\lambda \in \Delta'} \|(P_H - P'_H)_{xy}\| d\lambda \leq C d_\ell(H, H')^s|a(x)|^{-1}(1 + \|x - y\|)^{-a} \quad (x, y \in \mathbb{Z}^d).$$  

(2.7)

**Proof** We replace the disorder averaging of [24, Proposition 5.2] with Fermi energy averaging. We start from formula [1]

$$P_H = \frac{i}{2\pi} \int_{\Gamma(\lambda)} R(z) \, dz$$

where $\Gamma(\lambda)$ is a rectangular curve in $\mathbb{C}$ going counterclockwise passing the points $\lambda + i, \lambda - i, -\|H\| - 1 - i, -\|H\| - 1 + i$. We divide the curve into two parts: $\Gamma_1(\lambda)$ which is the two horizontal segments and the left vertical segment, and $\Gamma_2(\lambda)$, the right vertical segment. On $\Gamma_1(\lambda)$, $z$ is always a minimum distance of 1 from $\sigma(H)$ so that one may use the Combes–Thomas estimate [2, Theorem 10.5]. On $\Gamma_2(\lambda)$, we must use localization, since we (possibly) cross the spectrum as we pass the real axis.

We thus find:

$$\|(P_H - P'_H)_{xy}\| = \left\|\left(\frac{i}{2\pi} \int_{\Gamma(\lambda)} R(z) \, dz - \frac{i}{2\pi} \int_{\Gamma(\lambda)} R'(z) \, dz\right)_{xy}\right\|$$

$$\leq \left\|\frac{i}{2\pi} \int_{\Gamma_1(\lambda)} (G(x, y; z) - G'(x, y; z)) \, dz\right\|$$

$$+ \left\|\frac{i}{2\pi} \int_{\Gamma_2(\lambda)} (G(x, y; z) - G'(x, y; z)) \, dz\right\|$$
Then, with the resolvent identity
\[
\| \frac{i}{2\pi} \int_{\Gamma_1(\lambda)} (G(x, y; z) - G'(x, y; z)) \, d\eta \| \\
\leq \frac{1}{2\pi} \int_{\Gamma_1(\lambda)} \| G(x, y; z) - G'(x, y; z) \| \, d\eta \\
\leq \frac{1}{2\pi} \int_{\Gamma_1(\lambda)} \sum_{x', x''} \| G(x, x'; z) \| \| (H' - H)_{x', x''} \| \| G'(x'', y; z) \| \, d\eta
\]

Using Lemma 2.2 and the Combes–Thomas estimate [2, Theorem 10.5] (for some universal \( \mu > 0 \))
\[
\| G(x, y; z) \| \leq \frac{2}{\text{dist}(z, \sigma(H))} e^{-\mu \text{dist}(z, \sigma(H)) \| x - y \|},
\]
we now estimate this (recalling that for \( z \in \Gamma_1(\lambda) \), \( \text{dist}(z, \sigma(H)) \geq 1 \) and using Lemma 2.2)
\[
\ldots \leq \frac{2}{\pi} \int_{\Gamma_1(\lambda)} \sum_{x', x''} e^{-\mu \| x - x' \|} \ell(H, H') e^{-\mu \| x' - x'' \|} d\ell(H, H') e^{-\mu \| x'' - y \|} \, d\eta \\
\leq \frac{2}{\pi} |\Gamma_1(\lambda)| (\coth(\mu/2))^{2d} \ell(H, H') e^{-\frac{1}{2} \min(\mu, \ell(H, H')^{-1}) \| x - y \|}
\]

We see that \( \Gamma_1(\lambda) \) does not require the averaging over energy.

On the other hand for \( \Gamma_2(\lambda) \), we do use the localization estimate, which needs the Fermi energy averaging, i.e., Lemma 2.9. Let \( s \in (0, 1) \). Then, using the basic estimate \( \| G(x, y; \lambda + i \eta) \| \leq |\eta|^{-1} \), we find
\[
\int_{\Delta'} \| \frac{1}{2\pi} \int_{\Gamma_2(\lambda)} (G(x, y; z) - G'(x, y; z)) \, d\eta \| \, d\lambda \\
= \int_{\Delta'} \| \frac{-1}{2\pi} \int_{-1}^{1} (G(x, y; \lambda + i \eta) - G'(x, y; \lambda + i \eta)) \, d\eta \| \, d\lambda \\
\leq \frac{1}{2\pi} \int_{-1}^{1} \int_{\Delta'} \| G(x, y; \lambda + i \eta) - G'(x, y; \lambda + i \eta) \| \, d\eta \, d\lambda \\
\leq \frac{1}{2\pi} \int_{-1}^{1} \int_{\Delta'} \left| \frac{2}{\eta} \right|^{1-s/2} \| G(x, y; \lambda + i \eta) - G'(x, y; \lambda + i \eta) \|^{s/2} \, d\eta \, d\lambda
\]

Only now, after pulling a fractional power of the imaginary energy (unlike in [24]), do we use the resolvent identity:
\[
\ldots \leq \frac{1}{2\pi} \int_{-1}^{1} \int_{\Delta'} \left| \frac{2}{\eta} \right|^{1-s/2} \sum_{x', x''} \| G(x, x'; \lambda + i \eta) \|^{s/2} \| (H' - H)_{x', x''} \|^{s/2} \| G'(x'', y; \lambda + i \eta) \|^{s/2} \, d\eta \, d\lambda
\]

We may pull out the \( \sum_{x', x''} \) sum out of the integrals using Fatou’s lemma. Next, we use the Cauchy–Schwarz inequality on the \( \int_{\Delta'} \cdot d\lambda \) integral to get

\( \text{Springer} \)
\[
\ldots \leq \frac{1}{2\pi} \sum_{x', x''} \int_{-1}^{1} \left| \frac{2}{\eta} \right|^{1-s/2} \, d \eta \| (H' - H)_{x', x''} \|^{s/2} (\sup_{\eta \neq 0} \int_{\lambda \in \Delta'} \| G(x, x'; \lambda + i \eta) \|^{s} \, d \lambda)^{1/2} \\
\times (\sup_{\eta \neq 0} \int_{\lambda \in \Delta'} \| G'(x'', y; \lambda + i \eta) \|^{s} \, d \lambda)^{1/2}
\]

At this point, we employ the part of the assumption on \( H, H' \) concerning (2.5), so using \( s > 0 \), we have for any \( \alpha \in \mathbb{N} \) (having performed the \( \eta \) integration and used Lemma 2.2):

\[
\ldots \leq \frac{2^{2-s/2}}{\pi s} D_{0, \alpha} d_{\ell}(H, H')^{s/2} \\
\sum_{x', x''} e^{-\frac{1}{2d_{\ell}(H, H')^{s}}} |b(x)|^{-1/2} (1 + \| x - x' \|)^{-\alpha/2} |b(y)|^{-1/2} (1 + \| x'' - y \|)^{-\alpha/2}.
\]

Having arrived at this line, we obtain the result by the analysis of [30, Section 3.1] on the so-called weakly local *-algebra, which shows this last sum has polynomial decay with \( \| x - y \| \) and blowup in \( |b(x)|^{-1} \) or \( |b(y)|^{-1} \) (it does not matter which, because there is the decay in \( \| x - y \| \)).

For the convenience of the reader, we recap how that would work below. We ignore the \( b \) factors for a moment. Looking at the square root of the factor within the sum, using the fact that any exponential decay is dominated by any polynomial decay,

\[
e^{-\frac{1}{2d_{\ell}(H, H')^{s}}} (1 + \| x - x' \|)^{-\alpha/4} (1 + \| x'' - y \|)^{-\alpha/4} \quad \text{(triangle inequality)}
\]

\[
\leq (1 + \| x - x' \|)^{-\alpha/4} (1 + \| x' - x'' \|)^{-\alpha/4} (1 + \| x'' - y \|)^{-\alpha/4}
\]

\[
\leq (1 + \| x - y \|)^{-\alpha/4}.
\]

When we perform the sum of what remains (another factor of the first line in the above displayed inequality), we get a finite constant uniformly bounded in \( x, y \).

So the entire sum is bounded by

\[
C (1 + \| x - y \|)^{-\alpha/4} |b(x)|^{-1/2} |b(y)|^{-1/2}.
\]

Now we use the fact that \( b \) may be chosen polynomial, so that using the triangle inequality at the cost of adjusting the rate \( \alpha \), we get rid of \( b(y) \):

\[
(1 + \| x - y \|)^{-\mu} (1 + \| y \|)^{+\nu} \leq (1 + \| x - y \|)^{-(\mu - \nu)} (1 + \| x \|)^{+\nu}.
\]

\[\Box\]

**Proof of Theorem 2.11** Let \( H \in \mathcal{I} \) be given. We seek some \( \varepsilon > 0 \) (dependent on \( H \)) such that if \( H' \in \mathcal{I} \) with \( d_{\ell}(H, H') < \varepsilon \), then \( \text{Ch}(H) = \text{Ch}(H') \).

By Proposition 2.12, we may replace the Chern number with its average within the mobility gap to get

\[\text{Springer}\]
\[ | \text{Ch}(H) - \text{Ch}(H') | \]
\[ \leq \frac{1}{|\Delta'|} \left| \int_{\Delta'} \text{Ch}(H - \lambda I) \ d \lambda - \int_{\Delta'} \text{Ch}(H' - \lambda I) \ d \lambda \right| \]
\[ \leq \frac{2\pi}{|\Delta'|} \int_{\Delta'} | \text{tr}(\varepsilon_{\alpha\beta}(P_{\lambda} P_{\lambda,\alpha} P_{\lambda,\beta} - P'_{\lambda} P'_{\lambda,\alpha} P'_{\lambda,\beta})) | \ d \lambda \]
\[ \leq \frac{2\pi}{|\Delta'|} \sum_{(\alpha,\beta) = (1,2),(2,1)} \int_{\Delta'} \| (P_{\lambda} - P'_{\lambda}) P_{\lambda,\alpha} P_{\lambda,\beta} \|_1 + \| P'_{\lambda} (P_{\lambda} - P'_{\lambda}) P_{\lambda,\alpha} P_{\lambda,\beta} \|_1 \\
+ \| P'_{\lambda} P'_{\lambda,\alpha} (P_{\lambda} - P'_{\lambda}) P_{\lambda,\beta} \|_1 \ d \lambda. \]

We will use the estimate \( \| AB \|_1 \leq \sum_{x,y,z} \| A_{xy} \| \| B_{yz} \| \) throughout. Consider the first term with \( \alpha \neq \beta \):
\[ \int_{\Delta'} \| (P_{\lambda} - P'_{\lambda}) P_{\lambda,\alpha} P_{\lambda,\beta} \|_1 \ d \lambda \leq \int_{\Delta'} \sum_{x,y,z} \| (P_{\lambda} - P'_{\lambda})_{xy} \| \| (P_{\lambda,\alpha} P_{\lambda,\beta})_{yz} \| \ d \lambda. \]

Using (2.4), we know that \( P_{\lambda} \) has off-diagonal decay with diagonal explosion, which was called “weakly-local” in [30]. Since \( \chi(-\infty,\lambda) \in B_1(\Delta) \), the “weakly local” estimate we get does not depend on \( \lambda \). Using [30, Remark 3.4], we estimate \( \| (P_{\lambda,\alpha} P_{\lambda,\beta})_{yz} \| \leq C (1 + \| y - z \|)^{-\alpha} (1 + \| z \|)^{-\alpha} \) for some \( \alpha \in \mathbb{N} \) as large as we want, and the constant \( C \) does not depend on \( \lambda \). We conclude now using Proposition 2.13 that
\[ \int_{\Delta'} \| (P_{\lambda} - P'_{\lambda}) P_{\lambda,\alpha} P_{\lambda,\beta} \|_1 \ d \lambda \]
\[ \leq \sum_{x,y,z} C (1 + \| y - z \|)^{-\alpha} (1 + \| z \|)^{-\alpha} \int_{\Delta'} \| (P_{\lambda} - P'_{\lambda})_{xy} \| \ d \lambda \]
\[ \leq \sum_{x,y,z} C (1 + \| y - z \|)^{-\alpha} (1 + \| z \|)^{-\alpha} \]
\[ \times \frac{2^{2-s/2} D_{0,\alpha} Q}{\pi S} d(H',H')^{s/2} |b(y)|^{-1/2} (1 + \| x - y \|)^{-\alpha/2}. \]

This last expression after the triple sum is summable so that we get some constant times \( d(H',H')^{s/2} \), which means we can make this term as small as we like by appropriate choice of \( H' \).

Consider now one of the derivative terms with \( \alpha \neq \beta \), where we again use (2.4) with estimates independent of \( \lambda \):
\[ \int_{\Delta'} \| P'_{\lambda} (P_{\lambda} - P'_{\lambda})_{\alpha P_{\lambda,\beta}} \|_1 \ d \lambda \]
\[ \leq \int_{\Delta'} \| (P_{\lambda} - P'_{\lambda})_{\alpha P_{\lambda,\beta}} \|_1 \ d \lambda \]
\[ \leq \int_{\Delta'} \sum_{x,y,z} \| ((P_{\lambda} - P'_{\lambda})_{\alpha x y} \| \| (P_{\lambda,\beta})_{yz} \| \ d \lambda. \]
Now we have \( \| (A_\alpha)_{xy} \| = \| [\Lambda_\alpha, A]_{xy} \| = \| (\Lambda(x_\alpha) - \Lambda(y_\alpha))A_{xy} \| \). We can now invoke [15, Proof of Lemma 2] to bound \( |\Lambda(x_\alpha) - \Lambda(y_\alpha)| \leq C_\Lambda (1 + |x_\alpha - y_\alpha|)^\mu (1 + \frac{1}{2}|x_\alpha|)^{-\mu} \) to get, using Proposition 2.13 again:

\[
\int_{\Delta'} \| P'_\lambda (P_\lambda - P'_\lambda)_{x\beta} \|_1 d\lambda \\
\leq \sum_{xyz} C(1 + \| y - z \|)^{-\alpha} |a(y)|^{-1} (1 + |y_\beta|)^{-\alpha} \\
\times C_\Lambda (1 + |x_\alpha - y_\alpha|)^\mu (1 + \frac{1}{2}|x_\alpha|)^{-\mu} \int_{\Delta'} \| (P_\lambda - P'_\lambda)_{xy} \| d\lambda \\
\leq \sum_{xyz} C(1 + \| y - z \|)^{-\alpha} |a(y)|^{-1} (1 + |y_\beta|)^{-\alpha} \\
\times C_\Lambda (1 + |x_\alpha - y_\alpha|)^\mu (1 + \frac{1}{2}|x_\alpha|)^{-\mu} \\
\times \frac{2^{2-s/2} D_{0,\alpha} Q}{\pi s} \frac{d\ell(H, H')^{s/2} |b(y)|^{-1/2} (1 + \| x - y \|)^{-\alpha/2}}.
\]

This last expression is unfortunately very long but the point is (when the dust settles) that it really is just a summable expression after the triple sum, so that we again get a constant times \( d_\ell(H, H')^{s/2} \), which we can make as small as we like (there will still be a coth dependence on \( d_\ell(H, H') \) coming from the sum, as in (2.3), and similarly that dependence approaches \( d_\ell(H, H') \to 0 \). The last derivative term is dealt with in the same manner, and we find our result. □

In concluding this proof of continuity of \( \text{Ch} \), we compare it to the probabilistic proof of [24, Proposition 5.2]. In short, the latter proof shows that if \([0, 1] \ni t \mapsto H(t)\) is a family of random ergodic Hamiltonians with \( t \) the parameter of deformations, then \( t \mapsto \mathbb{E}[\text{Ch}(H(t))] \) is locally constant. Since we know that \( \mathbb{E}[\text{Ch}(H(t))] \) is almost surely equal to \( \text{Ch}(H(t)) \) (by Birkhoff), we conclude that almost surely, if \( |t - s| \) is small, \( \text{Ch}(H(t)) = \text{Ch}(H(s)) \) (in the setting of [24], the family \( t \mapsto H(t) \) varies within one and the same random probability space so that it makes sense to compare random configurations at different values of \( t \).

In contrast, Theorem 2.11 shows that if the mobility gap property holds for two given realizations of nearby Hamiltonians (as measured by \( d_\ell \)), then their respective Chern numbers are necessarily equal. Hence, the key topological property, namely the local constancy of the Chern number, is unrelated and does not rely on translation invariance whether in real space or in probability space, nor on disorder averaging.

**Remark 2.14** One might wonder whether other strategies for proving the local constancy could work better, instead of the direct analytical approach presented here. The first alternative is to rewrite

\[
\text{Ch}(H) = \text{ind}(PUP + P^\perp)
\]
where $U = \exp(i \arg(X_1 + i X_2))$. This formula is perfectly valid in the mobility gap regime. In [11], it is shown that $\text{Ch}(H)$ may be rewritten as Avron–Seiler–Simon’s [5] index of pair of projections ($P$ and $U PU^*$), which in turn is known to equal the above Fredholm index via Fedosov’s tracial expression of the index. However, this does not bring one very far as one still has to demonstrate why perturbations of the Hamiltonian which keep the mobility gap open mean that $PU P$ is perturbed in either a norm continuous or (more likely) a compact way so that the index remains constant (it is clear that $P$ alone will, in general, be perturbed all throughout space and its norm could jump by 1).

Another possibility is to rewrite $P$ as a strong limit of smooth functions approximating the step function applied on the Hamiltonian and hope to be able to pull this limit out of the formula for $\text{Ch}(H)$, which we have not been able to accomplish.

### 2.2 Other symmetry classes and dimensions

In [30], it is shown that the topological invariant for mobility-gapped Floquet 2D systems with no symmetry constraints is invariant under selection of the logarithm branch cut within the mobility gap. This means that it might be possible for the proof above to be adapted for such systems. Indeed, apparently the crucial ingredients are statements such as Proposition 2.12 and a rewriting of the invariant in terms of contour integrals on resolvents, which allows for resolvent identities to be used. Part of [30, Theorem 2.1] is the analog of Proposition 2.12, though it is only through [30, Theorem 2.6] that this is really established.

Coincidentally, an analog of Proposition 2.12 is precisely what we do not have for the chiral 1D systems studied in [15]. Indeed, for such chiral systems, the Fermi energy must be fixed at zero and so there is not an obvious parameter to average over. However (using the notation of [15]), it is conceivable that the chiral invariant of the Hamiltonian

$$H + \begin{bmatrix} 0 & \tilde{a} \\ a & 0 \end{bmatrix}$$

is independent of $\alpha \in \mathbb{C}$ small as long the mobility gap of the sum Hamiltonian stays open.

In contrast, there is little hope to control the local constancy of $\mathbb{Z}_2$ invariants associated with time-reversal invariant systems [19] using the approach presented here. Indeed, the formulas we have for invariant [28, Section 5] (which remain well defined in the mobility gap regime) are defined as an integer modulo 2. For example,

$$\dim \ker PU P + P^\perp \mod 2$$

where $U$ is as in (2.8), so that jumps of even integers in $\dim \ker PU P + P^\perp$ should be allowed via continuous deformations. Moreover, there still no (and no reason to expect) a local trace formula for the $\mathbb{Z}_2$ invariant analogous to (2.6).
3 Signatures of localization

In this section, we justify the somewhat technical Definition 2.8. Let $H \in \mathcal{B}(\mathcal{H})$ be a given random (ergodic) Hamiltonian. Let $\Delta \in \text{Open}(\mathbb{R})$ be a given bounded interval. The fractional moments condition on $\Delta$ [3, Lemma 2.1] says that for Lebesgue-almost-all $E \in \Delta$, there are some fraction $s_E \in (0, 1)$ and constants $C_E < \infty, \mu_E > 0$ such that

$$\sup_{\eta \neq 0} \mathbb{E}[\|G(x, y; E + i \eta)\|^{s_E}] \leq C_E e^{-\mu_E \|x-y\|} \quad (x, y \in \mathbb{Z}^d). \quad (3.1)$$

Here, $\mathbb{E}$ is the disorder averaging and the other notation symbols are as in the preceding sections.

A further condition [16, Eq. (4)] that does not seem to follow automatically from (3.1) (see [23]), but rather requires more input from $H$ is that for all $E \in \Delta$ there are constants $C_E < \infty, \mu_E > 0$ (here and below, these constants differ for each constraint) such that

$$\sup_{\eta \neq 0} \mathbb{E}[\|G(x, y; E + i \eta)\|^2] \leq C_E e^{-\mu_E \|x-y\|} \quad (x, y \in \mathbb{Z}^d). \quad (3.2)$$

Two very important consequences of these two conditions for topological insulators appeared in [1]. Equation (3.1) was shown to imply that for all $E \in \Delta$, there are constants $C_E < \infty, \mu_E > 0$ such that

$$\mathbb{E}[\chi_{(-\infty, E)}(H)_{xy}] \leq C_E e^{-\mu_E \|x-y\|} \quad (x, y \in \mathbb{Z}^d). \quad (3.3)$$

Equation (3.2) in turn was shown to imply that there are constants $C < \infty, \mu > 0$ such that

$$\mathbb{E}[\sup_{f \in B_1(\Delta)} \|f(H)_{xy}\|] \leq C e^{-\mu \|x-y\|} \quad (x, y \in \mathbb{Z}^d), \quad (3.4)$$

with $B_1(\Delta)$ as in Definition 2.8. We note that since $\chi_{(-\infty, E)} \in B_1(\Delta)$, (3.4) implies (3.3). Also, this implies (see, e.g., [30, Proposition A1]) that there is a deterministic constant $\mu > 0$ such that for all $a \in \ell^1$, almost surely, there is a (random) constant $C_a < \infty$ such that

$$\sup_{f \in B_1(\Delta)} \|f(H)_{xy}\| \leq C_a |a(x)|^{-1} e^{-\mu \|x-y\|} \quad (x, y \in \mathbb{Z}^d). \quad (3.5)$$

as in, e.g., [11, Eq. (1.2)].

Moreover, the Kubo formula (2.6) clearly shows that (3.5) implies the longitudinal DC conductivity is finite and zero; hence, it makes sense to include it as the first part of the definition of an insulator, Definition 2.8. The second part, which concerns the finite degeneracy of localized eigenvalues within the mobility gap, is again a standard almost-sure consequence of models where localization is established [31], and we include it here for technical reasons, following [11].

Springer
Proof of Lemma 2.9 First, we note that via the RAGE theorem, Definition 2.8 implies that $H$ has pure point spectrum within $\Delta$.

In [11, Proof of Lemma 4], it is proven that Definition 2.8 implies that $H$ has a SULE eigenbasis within $\Delta$, in the sense of [9, Eq. (7.1)], i.e., we will use that there is some set of normalized eigenvectors $\{\psi_n\}_{n \in \mathbb{N}}$ with eigenvalues $\{\lambda_n\}_{n \in \mathbb{N}} \subseteq \Delta$ such that there is a constant $\mu > 0$ such that for any $\varepsilon > 0$ there is some $C_\varepsilon < \infty$ with

$$
\|\psi_n(x)\| \leq C_\varepsilon \exp(-\mu \|x - x_n\| + \varepsilon \|x_n\|) \quad (x \in \mathbb{Z}^d)
$$

for some $x_n \in \mathbb{Z}^d$ (the localization center of $\psi_n$). Furthermore, in the proof of [9, Corollary 7.3], it is shown that

$$
\|x_n\| \geq \frac{1}{3} n^{\frac{1}{d}} - C_0 \quad (n \in \mathbb{N})
$$

for some $C_0 < \infty$. This is the estimate that one obtains on

$$
\psi_n(x) := (\chi_{\{\lambda_n\}}(H)_{x,x})^{-1/2} \chi_{\{\lambda_n\}}(H)_{x,x},
$$

where $x_n$ is defined as the point of maximum of $x \mapsto \chi_{\{\lambda_n\}}(H)_{x,x}$, by using Definition 2.8 with $f := \chi_{\{\lambda_n\}}$, $a(x) := \exp(-\varepsilon \|x\|)$. But we can also get (3.6) with polynomial explosion in $\|x\|$ by using $a(x) := (1 + \|x\|)^{-\nu}$ for $\nu \in \mathbb{N}_{>d}$ instead, in order to (eventually) get the second part of the statement of Lemma 2.9.

To focus on the region of localization and be able to use the SULE basis only in that region, we break up $\Delta'$ as follows: let $\Delta''$ be an open interval such that $\Delta' \subseteq \Delta'' \subseteq \Delta$, and pick some smooth $\varphi : \mathbb{R} \to [0, 1]$ such that $\varphi |_{\Delta'} = 0$ and $\varphi |_{\Delta''} = 1$. Then, for all $z \in \mathbb{C}$ such that $\Re(z) \in \Delta'$, $\mathbb{R} \ni \lambda \mapsto (1 - \varphi(\lambda))(\lambda - z)^{-1}$ is a smooth function with compact support and hence by the Helffer–Sjöstrand smooth functional calculus [10, Appendix A] combined with the Combes–Thomas estimate [2, Thm 10.5], $(1 - \varphi(H))R(z)$ has position-basis matrix elements with any rate polynomial off-diagonal decay, uniformly as $\Im(z) \to 0$. We note that the constant depends on the rate of decay chosen, which is why in (2.5), $D_\alpha$ depends on $\alpha$.

Hence, we concentrate on the off-diagonal decay of matrix elements of the operator $\varphi(H)R(z)$. Since $\text{im}(\chi_{\Delta}(H))$ is spanned by $\{\psi_n\}_{n \in \mathbb{N}}$ and $\text{supp}(\varphi) \subseteq \Delta$, we may write

$$
\varphi(H)R(z) = \sum_{n \in \mathbb{N}} \frac{\varphi(\lambda_n)}{\lambda_n - z} \psi_n \otimes \psi_n^*,
$$

so that using $(\sum_n a_n)^s \leq \sum_n a_n^s$, valid for all $s \leq 1$, applying the polynomial version of (3.6) for any $\alpha \in \mathbb{N}$ large and some $\nu \in \mathbb{N}_{>d}$, we get for some $C'_s < \infty$,
\[
\int_{E \in \Delta'} \| (\varphi(H) R(E + i\eta))_{xy} \|^s \, dE \\
\leq \int_{E \in \Delta'} \sum_{n \in \mathbb{N}} |\lambda_n - E - i\eta|^{-s} \| \psi_n(x) \| \| \psi_n(y) \|^s \, dE \\
\leq C_s' \sum_{n \in \mathbb{N}} C_v^{2s} (1 + \|x - x_n\|)^{-s\alpha} (1 + \|y - x_n\|)^{-s\alpha} (1 + \|x_n\|)^{2su} \\
\leq C_s' C_v^{2s} (1 + \|x - y\|)^{-s\alpha/2} \sum_{n \in \mathbb{N}} (1 + \|x - x_n\|)^{-s\alpha/2} (1 + \|x_n\|)^{2su}. \]

To get the last inequality, we applied the triangle inequality on a square root of the factor \((1 + \|x - x_n\|)^{-s\alpha} (1 + \|y - x_n\|)^{-s\alpha}\) and estimated \((1 + \|y - x_n\|)^{-s\alpha/2} \leq 1\). Next, we use another consequence of the triangle inequality, namely if \(\mu \geq 2\nu\),

\[
(1 + \|x - x_n\|)^{-\mu} (1 + \|x_n\|)^{+\nu} \leq (1 + \|x - x_n\|)^{-2\nu} (1 + \|x_n\|)^{+\nu} \\
\leq (1 + \|x\|)^{2\nu} (1 + \|x_n\|)^{-\nu}. \]

Finally, we use (3.7) to get that the \(n\) sum is finite and hence our result, assuming \(s, \nu, \alpha\) are appropriately chosen.

Combining the two estimates using \(1 = \varphi + (1 - \varphi)\) yields the result. We emphasize that the downgrade to polynomial from exponential decay comes from application of the smooth functional calculus. \(\Box\)

### 3.1 Further consequences from the probabilistic fractional moment condition

Now we present the following additional deterministic consequence of (3.1), which however—this is Lemma 2.9—is also a consequence of (3.5) together with the assumption of uniform finite degeneracy of the localized eigenvalues, as in Definition 2.8. Hence, we include the following result mainly for general interest (as it does not involve the finite degeneracy of eigenvalues), but it is not strictly necessary for the logical flow.

**Proposition 3.1** If \(H\) is localized on \(\Delta\) in the sense of (3.1), then almost surely, for any \(\alpha \in \mathbb{N}\), there is a (random) constant \(C_\alpha < \infty\) and (deterministic) \(a \in \ell^1(\mathbb{Z}^d)\) such that

\[
\sup_{\eta \neq 0} \int_{E \in \Delta'} \| G(x, y; E + i\eta) \|^s \, dE \leq C_\alpha |a(x)|^{-1} (1 + \|x - y\|)^{-\alpha} \quad (x, y \in \mathbb{Z}^d) \tag{3.8}
\]

where \(\Delta' \subset \Delta\) is a compact subinterval.

To calibrate (3.8), let us first make the following
Remark 3.2  Regardless of the fate of localization for $H$, for all disorder configurations, for any fixed interval $I \subset \mathbb{R}$, $s \in (0, 1)$ and $x, y \in \mathbb{Z}^d$,

$$\sup_{\eta \neq 0} \int_{E \in I} \|G(x, y; E + i \eta)\|^s \, dE < \infty \quad (3.9)$$

as shown in [2, Eq. (8.2)].

To prove Proposition 3.1, we start by making the weaker statement about what happens not in $\sup_{\eta \neq 0}$ but only in the $\lim_{\eta \to 0^+}$.

Lemma 3.3 If $H$ obeys the fractional moment condition on $\Delta$, (3.1), then for any compact subinterval $\Delta' \subset \Delta$, almost surely, there is a (random) constant $C < \infty$ and (deterministic) $a \in \ell^1(\mathbb{Z}^d)$, $\mu > 0$, $s \in (0, 1)$ such that

$$\lim_{\eta \to 0^+} \sup_{\eta \to 0^+} \int_{E \in \Delta'} \|G(x, y; E + i \eta)\|^s \, dE \leq C |a(x)|^{-1} e^{-\mu \|x-y\|} \quad (x, y \in \mathbb{Z}^d) \quad (3.10)$$

Proof  Relating to the notations above (3.1), let $s := \min(|s_E| \in E \in \Delta')$,

$$C' := |\Delta'| \max(|\frac{s}{s_E} E \in \Delta') \quad \mu := \frac{s}{2 \max(|\frac{s}{s_E} E \in \Delta') \min(|\mu_E \in E \in \Delta')}$$

and pick any $a \in \ell^1(\mathbb{Z}^d)$. By Remark 3.2, we know that $\lim_{\eta \to 0^+} \int_{\Delta'} \|G(x, y; E + i \eta)\|^s < \infty$, so that we may use the dominated convergence theorem in the second line below. We get

$$\star := \mathbb{E} \left[ \sum_{x, y \in \mathbb{Z}^d} \lim_{\eta \to 0^+} \sup_{\eta \to 0^+} \int_{\Delta'} \|G(x, y; q + i \eta)\|^s \, dq |a(x)| e^{+\mu \|x-y\|} \right]$$

(Fatou on $\lim_{\eta \to 0^+}$)

$$\leq \liminf_{\Lambda \to \mathbb{Z}^d} \sum_{x, y \in \Lambda} \mathbb{E} \left[ \limsup_{\eta \to 0^+} \int_{\Delta'} \|G(x, y; q + i \eta)\|^s \, dq |a(x)| e^{+\mu \|x-y\|} \right]$$

$$= \liminf_{\Lambda \to \mathbb{Z}^d} \sum_{x, y \in \Lambda} \limsup_{\eta \to 0^+} \mathbb{E} \left[ \int_{\Delta'} \|G(x, y; q + i \eta)\|^s \, dq |a(x)| e^{+\mu \|x-y\|} \right]$$

(Dominated convergence)

$$= \liminf_{\Lambda \to \mathbb{Z}^d} \sum_{x, y \in \Lambda} \limsup_{\eta \to 0^+} \int_{\Delta'} \mathbb{E} \left[ \|G(x, y; q + i \eta)\|^s \right] \, dq |a(x)| e^{+\mu \|x-y\|} \quad (\text{Fubini})$$

Now we note that if (3.1) holds for some $s_E \in (0, 1)$, then the same holds with $s_E$ replaced by $\sigma$ for all $\sigma \in (0, s_E)$ by Jensen’s inequality [25, Thm 3.3]:

$$\sup_{\eta \neq 0} \mathbb{E} \left[ \|G(x, y; E + i \eta)\|^s \right] \leq C_E^{\frac{s}{s_E}} e^{-\frac{s}{s_E} \mu_E \|x-y\|} \quad (x, y \in \mathbb{Z}^d).$$
If we restrict to \( E \in \Delta' \), then the RHS is bounded by \( \frac{1}{|\Delta'|} C' e^{-2\mu \|x-y\|} \). Hence,

\[
\star \leq \liminf_{\Lambda \to \mathbb{Z}^d} \sum_{x, y \in \Lambda} \limsup_{\eta \to 0^+} \int_{\Delta'} \frac{1}{|\Delta'|} C' e^{-2\mu \|x-y\|} |a(x)| e^{+\mu \|x-y\|}
\]

\[
= C' \sum_{x, y \in \mathbb{Z}^d} e^{-\mu \|x-y\|} |a(x)|
\]

\[
= C' \|a\|_1 \sum_{x \in \mathbb{Z}^d} e^{-\mu \|x\|}
\]

\[
< \infty.
\]

Hence, there is some random \( C < \infty \) such that

\[
\sum_{x, y \in \mathbb{Z}^d} \limsup_{\eta \to 0^+} \int_{\Delta'} \|G(x, y; q + i \eta)\|^s d \eta |a(x)| e^{+\mu \|x-y\|} \leq C
\]

and so also

\[
\limsup_{\eta \to 0^+} \int_{\Delta'} \|G(x, y; q + i \eta)\|^s d \eta \leq C |a(x)|^{-1} e^{-\mu \|x-y\|} (x, y \in \mathbb{Z}^d)
\]

which is what we wanted to prove. \( \square \)

Our next task is to upgrade the lim sup to a sup. We first have to establish a certain subharmonicity:

**Lemma 3.4** With \( \mathbb{C}_+ := \{z| \Im\{z\} > 0\} \) we have that for any \( a > 0, s \in (0, 1), \)

\[
\mathbb{C}_+ \ni z \mapsto \int_{\Re [z] - a}^{\Re [z] + a} \|G(x, y; q + i \Im \{z\})\|^s d \eta =: \varphi_{x, y, a, s} (z) \in \mathbb{R}
\]

is a subharmonic function.

**Proof** First, note that \( \mathbb{C}_+ \ni z \mapsto G(x, y; z) \in \text{Mat}_N (\mathbb{C}) \) is analytic, so that \( \mathbb{C}_+ \ni z \mapsto \|G(x, y; z)\|^s =: g(z) \) is subharmonic. To verify that \( \varphi \) is subharmonic, we pick any \( z \in \mathbb{C}_+ \) and \( r > 0 \) such that \( \overline{B_r (z)} \subseteq \mathbb{C}_+ \) (so \( r < \Im \{z\} \)) and aim to prove \( \varphi (z) \leq \frac{1}{2\pi} \int_{0}^{2\pi} \varphi (z + r e^{i\theta}) d \theta; \)

\( \square \) Springer
\( \varphi(z) \equiv \int_{\Re[z] + a}^{\Re[z] - a} g(\lambda + i \Im\{z\}) \, d\lambda \quad \text{(Subharmonicity of } g) \)

\[ \leq \int_{\Re[z] - a}^{\Re[z] + a} \frac{1}{2\pi} \int_{\theta=0}^{2\pi} g(\lambda + i \Im\{z\} + r e^{i\theta}) \, d\theta \, d\lambda \]

\[ \leq \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \int_{\Re[z] + a}^{\Re[z] - a} g(\lambda + i \Im\{z\} + r e^{i\theta}) \, d\lambda \, d\theta \quad \text{(Fubini)} \]

\[ = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \varphi(\Re\{z\} + r \cos(\theta) + i (\Im\{z\} + r \sin(\theta))) \, d\theta \]

\[ = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \varphi(z + r e^{i\theta}) \, d\theta. \]

\[ = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} \varphi(z + r e^{i\theta}) \, d\theta. \]

**Proof of Proposition 3.1** Since \( \varphi_{x, y, a, s} \) is a subharmonic function which decays in \( \|x - y\| \) above the real axis (via the Combes–Thomas estimate) and whose \( \lim sup \) on the real axis decays in \( \|x - y\| \) via Lemma 3.3, we find our result using either [4, Theorem 4.2] or [14, Proposition 25].

**Acknowledgements** It is a pleasure to thank Gian Michele Graf for many useful discussions. I am grateful to Jeffrey Schenker for Lemma 2.9, which allows one to avoid Proposition 3.1 and hence simplify Definition 2.8. Finally, I thank two anonymous referees for many useful comments. This research is supported in part by Simons Foundation Math + X Investigator Award #376319 (Michael I. Weinstein).

### 4 Appendix

The following basic lemma, known as Holmgren’s bound, is a well-known result for both continuum and discrete operators. Instead of giving a reference to the general result in the context of PDEs, we provide the simple proof in the present setting below.

**Lemma 4.1 (Holmgren’s bound)** Let \( \mathcal{H} \) be a separable Hilbert space and \( A \) a bounded linear operator on \( \mathcal{H} \). Then,

\[ \|A\| \leq \max \left( \left\{ \sup_n \sum_m |A_{nm}|, \sup_m \sum_n |A_{nm}| \right\} \right) \]

where \( A_{nm} \equiv \langle \delta_n, A \delta_m \rangle \) and \( \{\delta_n\}_n \) is any ONB. We call the first element in the above set the \( \infty \), 1 norm of \( A \) and the second one the 1, \( \infty \) norm of \( A \).
Proof We use the characterization \( \|A\| = \sup \left( \left\{ \| \langle \varphi, A \psi \rangle \| : \| \varphi \| = \| \psi \| = 1 \right\} \right) \). Then

\[
| \langle \varphi, A \psi \rangle | \leq \sum_{n,m} | \varphi_n | | A_{nm} | \| \psi_m \| = \sum_{n,m} (| \varphi_n | | A_{nm} | ) (| A_{nm} | \| \psi_m \|)
\]

(Cauchy–Schwarz on \( \sum_{n,m} \))

\[
\leq \left( \sum_{n,m} | \varphi_n |^2 | A_{nm} | \right)^{1/2} \left( \sum_{n,m} | A_{nm} | \| \psi_m \| ^2 \right)^{1/2}
\]

\[
\leq \left( \| \varphi \| ^2 \sup_n \sum_m | A_{nm} | \right)^{1/2} \left( \sup_m \sum_n | A_{nm} | \| \psi \| ^2 \right)^{1/2}.
\]

\[\square\]

References

1. Aizenman, M., Graf, G.M.: Localization bounds for an electron gas. J. Phys. A Math. Gen. 31, 6783–6806 (1998)
2. Aizenman, M., Warzel, S.: Random Operators. American Mathematical Society, Providence (2015)
3. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. 157(2), 245–278 (1993)
4. Aizenman, M., et al.: Finite-volume fractional-moment criteria for Anderson localization. Commun. Math. Phys. 224(1), 219–253 (2001)
5. Avron, J.E., Seiler, R., Simon, B.: Homotopy and quantization in condensed matter physics. Phys. Rev. Lett. 51, 51–53 (1983)
6. Bellissard, J., van Elst, A., Schulz-Baldes, H.: The noncommutative geometry of the quantum Hall effect. J. Math. Phys. 35, 5373–5451 (1994)
7. Bourne, C., Kellendonk, J., Rennie, A.: The K-theoretic bulk-edge correspondence for topological insulators. Ann. Henri Poincaré 18(5), 1833–1866 (2017)
8. Del Rio, R., Makarov, N., Simon, B.: Operators with singular continuous spectrum. II. Rank one operators. Commun. Math. Phys. 165(1), 59–67 (1994)
9. del Rio, R., et al.: Operators with singular continuous spectrum, IV. Hausdorff dimensions, rank one perturbations, and localization. J. d’Anal. Mathématique 69(1), 153–200 (1996)
10. Elbau, P., Graf, G.M.: Equality of bulk and edge Hall conductance revisited. Commun. Math. Phys. 229(3), 415–432 (2002)
11. Elgart, A., Graf, G.M., Schenker, J.: Equality of the bulk and edge Hall conductances in a mobility gap. Commun. Math. Phys. 259(1), 185–221 (2005)
12. Freed, D.S., Moore, G.W.: Twisted Equivariant Matter. Ann. Henri Poincaré 14(18), 1927–2023 (2013)
13. Gordon, A.Y.: Pure point spectrum under 1-parameter perturbations and instability of Anderson localization. Commun. Math. Phys. 164(3), 489–505 (1994)
14. Graf, G. M., Shapiro, J.: Complete localization for disordered Chiral chains (in preparation) (2019)
15. Graf, G.M., Shapiro, J.: The bulk-edge correspondence for disordered chiral chains. Commun. Math. Phys. 363, 829–846 (2018)
16. Graf, G.M.: Anderson localization and the space-time characteristic of continuum states. J. Stat. Phys. 75(1), 337–346 (1994)
17. Graf, G.M., Tauber, C.: Bulk-edge correspondence for two-dimensional Floquet topological insulators. Ann. Henri Poincaré 19(3), 709–741 (2018)
18. Hasan, M.Z., Kane, C.L.: Colloquium: topological insulators. Rev. Mod. Phys. 82, 3045–3067 (2010)
19. Kane, C.L., Mele, E.J.: \( Z_2 \) topological order and the quantum spin hall effect. Phys. Rev. Lett. 95, 146802 (2005)
20. Kellendonk, J.: On the C-star-algebraic approach to topological phases for insulators. Ann. Henri Poincaré. 18(7), 2251–2300 (2017)
21. Kitaev, A.: Periodic table for topological insulators and superconductors. AIP Conf. Proc. 1134(1), 22–30 (2009)
22. Prodan, E., Schulz-Baldes, H.: Bulk and Boundary Invariants for Complex Topological Insulators: From K-Theory to Physics. Springer (2016). https://doi.org/10.1007/978-3-319-29351-6_4
23. Mavi, R., Schenker, J.: Resonant tunneling in a system with correlated pure point spectrum. (2017). arXiv: 1705.03039 [math-ph]
24. Prodan, E., Schulz-Baldes, H.: Non-commutative odd Chern numbers and topological phases of disordered chiral systems. J. Funct. Anal. 271(5), 1150–1176 (2016)
25. Rudin, W.: Real and Complex Analysis (Higher Mathematics Series). McGraw-Hill Education, New York City (1986)
26. Ryu, S., et al.: Topological insulators and superconductors: tenfold way and dimensional hierarchy. New J. Phys. 12(6), 065010 (2010)
27. Schnyder, A.P., et al.: Classification of topological insulators and superconductors in three spatial dimensions. Phys. Rev. B. 78, 195125 (2008)
28. Schulz-Baldes, H.: Z2 indices and factorization properties of odd symmetric Fredholm operators. Doc. Math. 20, 1500 (2015)
29. Shapiro, J.: Notes on “Topological Aspects in Condensed Matter Physics” following a course taught by G. M. Graf in the Fall semester of 2015 (2016). https://people.phys.ethz.ch/~jshapiro/PDFs/Top_SSP_Lecture_Notes.pdf Online. Accessed 30 July 2018
30. Shapiro, J., Tauber, C.: Strongly disordered floquet topological systems. Ann. Henri Poincaré. 20(6), 1837–1875 (2019)
31. Simon, B.: Cyclic vectors in the Anderson model. Rev. Math. Phys. 06(05a), 1183–1185 (1994)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.