



J. Bourgain · W.-M. Wang

Quasi-periodic solutions of nonlinear random Schrödinger equations

Received August 1, 2006

Contents

1. Introduction and statement of the theorem	1
2. Hamiltonian representation and Lyapunov–Schmidt decomposition	8
3. Newton scheme	11
4. P -equations and statement in θ	14
5. Invertibility of $T(y_i)$, Q -equations and determination of ω	28
6. Construction of y_{i+1} and completion of the assemblage	37
7. Proof of the Theorem	40
Appendix: Localization results for random Schrödinger operators	41

1. Introduction and statement of the theorem

The nonlinear random Schrödinger equation

We seek time quasi-periodic solutions to the nonlinear random Schrödinger equation

$$i \frac{\partial}{\partial t} u = (\epsilon \Delta + V)u + \delta |u|^{2p} u \quad (p > 0) \tag{1.1}$$

on $\mathbb{Z}^d \times [0, \infty)$, where $0 < \epsilon, \delta \ll 1$, Δ is the discrete Laplacian:

$$\Delta_{ij} = \begin{cases} 1, & |i - j|_{\ell^1} = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{1.2}$$

and $V = \{v_j\}_{j \in \mathbb{Z}^d}$, the potential, is a family of *time independent* independent identically distributed (i.i.d.) random variables with common distribution $g = \tilde{g}(v_j)dv_j$, $\tilde{g} \in L^\infty$.

We also assume $\text{supp } g$ is a bounded set. The probability space is taken to be $\mathbb{R}^{\mathbb{Z}^d}$ with measure

$$\prod_{j \in \mathbb{Z}^d} g(v_j) = \prod_{j \in \mathbb{Z}^d} \tilde{g}(v_j)dv_j, \quad \tilde{g} \in L^\infty. \tag{1.3}$$

$V = \{v_j\}_{j \in \mathbb{Z}^d}$ serve as parameters for the nonlinear problem in (1.1).

J. Bourgain: Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA;
e-mail: bourgain@math.ias.edu

W.-M. Wang: Département de Mathématique, Université Paris Sud, 91405 Orsay Cedex, France;
Department of Mathematics, University of Massachusetts, Amherst, MA 01003, USA;
e-mail: wei-min.wang@math.u-psud.fr, weimin@math.umass.edu

Given an initial condition $u(0)$ in $\ell^2(\mathbb{Z}^d)$, one of the central questions is whether $u(t)$ remains localized for all t , i.e., if $u(0) \in \ell^2(\mathbb{Z}^d)$, for all κ , can one find R such that

$$\|u(t)\|_{\ell^2(\{\mathbb{Z} \setminus [-R, R]\}^d)} < \kappa, \quad \forall t? \quad (1.4)$$

(From now on, we write $|\cdot|$ for $|\cdot|_{\ell^1}$, and $\|\cdot\|$ for $\|\cdot\|_{\ell^2}$.) When $\epsilon = \delta = 0$, the answer to (1.4) is affirmative. Since $u(0) = \sum_{j \in \mathbb{Z}^d} a_j \delta_j$ with $a_j \rightarrow 0$ as $|j| \rightarrow \infty$, $u(t) = \sum_{j \in \mathbb{Z}^d} a_j \delta_j e^{-i v_j t}$ is almost-periodic (infinite number of frequencies) and the upper bound in (1.4) is trivially satisfied.

In this paper, for appropriate initial conditions $u(0)$, we construct time quasi-periodic solutions to (1.1). So the answer to (1.4) is affirmative for such $u(0)$'s. This is the content of the Theorem and its Corollary.

Before we enter into the heart of the matter, we first address question (1.4) for

The linear random Schrödinger equation

When $\delta = 0$, (1.1) reduces to the linear random Schrödinger equation:

$$i \frac{\partial}{\partial t} u = (\epsilon \Delta + V)u =: Hu \quad (1.5)$$

on $\mathbb{Z}^d \times [0, \infty)$. When $0 < \epsilon \ll 1$, it is well known from the works in [AFHS, AM, vDK, FMSS, FS, GB, GK, GMP] etc. that the upper bound in (1.4) is satisfied. This is customarily called *Anderson localization* (A.L.) after the physicist P. Anderson [An]. Since the potential is time independent: $V(j, t) = V(j)$, properties of time evolution can be deduced from the spectral properties of H , which we summarize below. For more details, see the Appendix.

Let $\sigma(H)$ be the spectrum of H . For H defined in (1.5),

$$\sigma(H) = [-2\epsilon d, 2\epsilon d] + \text{supp } g, \quad \text{a.s.} \quad (1.6)$$

(recall the probability space defined in (1.3)) [CFKS, PF]. If $0 < \epsilon \ll 1$ and the probability measure satisfies (1.3), then almost surely the spectrum of H is (dense) pure point, $\sigma(H) = \sigma_{\text{pp}}(H)$, with exponentially localized eigenfunctions ϕ_j , $j \in \mathbb{Z}^d$.

Given $u(0) \in \ell^2(\mathbb{Z}^d)$, we decompose $u(0)$ as $u(0) = \sum_{j \in \mathbb{Z}^d} a_j \phi_j$. So

$$u(t) = \sum_{j \in \mathbb{Z}^d} a_j \phi_j e^{-i \omega_j t}, \quad (1.7)$$

where ω_j are the eigenvalues for the eigenfunctions ϕ_j . Thus $u(t)$ is almost-periodic and satisfies the upper bound in (1.4). So equation (1.5) has A.L.

Some motivations for studying equation (1.1)

Schrödinger equations describe physical systems which typically correspond to an n -body problem. The linear equation in (1.5) is a 0th order approximation, where the n -body interaction is lumped into the effective potential V . Quantum mechanically, $|u|^2$ is interpreted

as particle density, so the nonlinear term in (1.1) can be interpreted as modeling particle-particle interaction. (The nonlinear term in (1.1) can be more general and of convolution type. It will not affect our construction below.) This is sometimes called the Hartree–Fock approximation (cf. [LL, O, Sh]) and is a first order approximation to the original n -body problem. This is our first motivation to study (1.1). Other physical motivations along this line appear in [FSW].

In particular, our method permits us to construct quasi-periodic solutions for the Landau–Lifschitz equations for nonlinear classical spin waves with a large random external magnetic field,

$$\dot{S}_j = S_j \times [(\Delta S)_j + h_j] \quad (j \in \mathbb{Z}^d)$$

where S_j are unit vectors in \mathbb{R}^3 and $h_j = V_j \vec{e}_3$ say, with $V = (V_j)_{j \in \mathbb{Z}^d}$ a large random potential.

As explained in [FSW], we may then seek for a solution $S_j \approx e_3$ and the perturbation is subject to an equation of the form (1.1), but with a nearest neighbor convolution nonlinearity instead of the local one $|u|^{2p}u$ (see [FSW] for details). As mentioned before, (1.1) was chosen as a model but the method described in the paper is sufficiently robust to cover in particular any nonlinearity with finite range interactions.

Our second motivation originates from KAM type stability questions for infinite-dimensional dynamical systems. (For results in the standard KAM context, see e.g. [E].) (1.1) is a Hamiltonian PDE. It can be recast as the equation of motion corresponding to a Hamiltonian of a perturbed \mathbb{Z}^d -system of coupled harmonic oscillators with i.i.d. random frequencies (see (2.2, 2.3)). When $\delta = 0$, the linear system has pure point spectrum: $\sigma(H) = \sigma_{\text{pp}}(H)$. This corresponds to the KAM tori scenario. A natural question is the stability of such invariant tori under small ($0 < \delta \ll 1$) perturbations, which leads to construction of quasi-periodic or almost-periodic solutions to (1.1).

Remark. Previously in [AF, AFS], solutions to the nonlinear eigenvalue problem

$$(\epsilon \Delta + V)\phi + \delta |\phi|^{2p} \phi = E\phi \quad \text{on } \ell^2(\mathbb{Z}^d)$$

were found, which give the time periodic solutions to (1.1) of the particular form

$$u(j, t) = \phi(j) e^{-iEt}.$$

A sketch of the construction

We expand in the Fourier basis $e^{in \cdot \omega t} \delta_k(j)$, and as an ansatz, seek solutions of the form

$$u(\ell, t) = \sum_{(j,n) \in \mathbb{Z}^{d+v}} \hat{u}(j, n) e^{in \cdot \omega t} \delta_j(\ell), \quad (1.8)$$

with the initial condition

$$u(\ell, 0) = \sum_{k=1}^v a_k \delta_k(\ell) \quad \text{satisfying} \quad \sum_{k=1}^v |a_k| \ll 1, \quad (1.9)$$

where we have identified δ_k with δ_{j_k} ($k = 1, \dots, v$), $j_k \in \mathbb{Z}^d$. The unperturbed frequen-

cies are therefore $\omega = \omega(\mathcal{V}) = \mathcal{V} \in \mathbb{R}^\nu$, where $\mathcal{V} := \{v_{j_k}\}_{k=1}^\nu$ are the random potentials at sites $j_k \in \mathbb{Z}^d$.

Substituting (1.8) into (1.1), we obtain the following equation for the Fourier coefficients:

$$(n \cdot \omega + \epsilon \Delta_j + V_j) \hat{u}(j, n) + \delta[(\hat{u} * \hat{v})^{*p} * \hat{u}](j, n) = 0, \quad (1.10)$$

where $\hat{v}(j, n) = \tilde{\hat{u}}(j, -n)$, the convolution $*$ is in the n variable only, $*p$ denotes the p -fold convolution and we added the subscript j to operators that originated from $\ell^2(\mathbb{Z}^d)$. We also write the equation for \hat{v} :

$$(-n \cdot \omega + \epsilon \Delta_j + V_j) \hat{v}(j, n) + \delta[(\hat{u} * \hat{v})^{*p} * \hat{v}](j, n) = 0. \quad (1.11)$$

Combining (1.10, 1.11), we then have a closed system of equations for $y = \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}$, which we write as

$$F(y) = 0. \quad (1.12)$$

Equation (1.12) is a $\mathbb{Z}^{d+\nu}$ system of equations. Let $y_0 = y(t=0)$. Then

$$\text{supp } y_0 = \{j_k, -e_k\}_{k=1}^\nu \cup \{j_k, e_k\}_{k=1}^\nu, \quad (1.13)$$

where e_k are the unit vectors of \mathbb{Z}^ν . We seek solutions to (1.12) with y fixed at the initial condition on $\text{supp } y_0$, i.e., $\hat{u}(j_k, -e_k) = a_k$, $\hat{v}(j_k, e_k) = \bar{a}_k$, $k = 1, \dots, \nu$ (cf. (2.8)). We make a Lyapunov–Schmidt decomposition as in [B1,3, CW1,2]. Let $y_0 = y(t=0)$. The equations

$$F(y) = 0|_{\mathbb{Z}^{d+\nu} \setminus \text{supp } y_0} \quad \text{on } \ell^2(\mathbb{Z}^{d+\nu} \setminus \text{supp } y_0)$$

are the so called P -equations, the rest are the Q -equations. The P -equations are used to determine $y(j, n)$ on $\{\text{supp } y_0\}^c$. On $\text{supp } y_0$, $y(j, n)$ are held fixed at the initial condition from (1.9). Instead the ν Q -equations determine $\omega = \omega(\mathcal{V})$.

We use a Newton scheme to solve the P -equations (for more details, see Section 3). This leads to investigate the invertibility of the linearized operators $F'_i(y_i)$, where y_i is the i^{th} approximate solution, and F'_i is F' restricted to $[-M^{i+1}, M^{i+1}]^{d+\nu}$ ($i \geq 0$) for appropriate M .

The random potentials $\mathcal{V} = \{v_{j_k}\}_{k=1}^\nu \in \mathbb{R}^\nu$ are the parameters in the problem. Invertibility of $F'_i(y_i)$ is ensured by appropriate incisions in the probability space \mathbb{R}^ν . Similar to the linear case in [BW], this is done by using semi-algebraic set techniques to control the complexity of the singular sets and a Cartan type theorem for analytic matrix-valued functions to control the measure.

The main difference from the linear case in [BW] is that F'_i are evaluated at different y_i . But due to rapid convergence of the Newton scheme, made possible by estimates on $F'_i(y_{i'})$ for $i' < i$, this is within the margin of estimates.

Solving the P -equations iteratively is the main part of the work. The solutions to the P -equations are then substituted into the Q -equations to determine $\omega = \omega(\mathcal{V})$ iteratively by using the implicit function theorem. We obtain time quasi-periodic solutions to (1.1) of the form (1.8), which are exponentially localized (both in the spatial and Fourier space) to the initial condition (1.9), with modified frequencies $\omega = \omega(\mathcal{V})$, which are $(\epsilon + \delta)$ -close to the unperturbed frequencies $\mathcal{V} = \{v_{j_k}\}_{k=1}^\nu$.

We therefore have

Statement of the Theorem

Theorem. Consider the nonlinear random Schrödinger equation

$$i \frac{\partial}{\partial t} u = (\epsilon \Delta + V)u + \delta |u|^{2p} u \quad (p \in \mathbb{N}^+), \quad (1.14)$$

where Δ is the discrete Laplacian defined in (1.2), and $V = \{v_j\}_{j \in \mathbb{Z}^d}$ is a family of i.i.d. random variables with common distribution g satisfying (1.3). Fix $j_k \in \mathbb{Z}^d$, $k = 1, \dots, v$. Let $\mathcal{R} = \{j_k\}_{k=1}^v \subset \mathbb{Z}^d$ and $\mathcal{V} = \{v_\alpha\}_{\alpha \in \mathcal{R}} \in \mathbb{R}^v$. Consider an unperturbed solution of (1.14) with $\epsilon, \delta = 0$,

$$u_0(y, t) = \sum_{k=1}^v a_k e^{-i v_{j_k} t} \delta_{j_k}(y),$$

with $\sum_{k=1}^v |a_k|$ sufficiently small. Let $a = \{a_k\}_{k=1}^v$.

For $0 < \epsilon \ll 1$, there exists $X_\epsilon \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^v$ of positive probability such that for $0 < \delta \ll 1$, if we fix $x \in X_\epsilon$, there exist a Cantor set $\mathcal{G}_{\epsilon, \delta}(x; a) \subset \mathbb{R}^v$ of positive measure and a smooth function $\omega = \omega_{\epsilon, \delta}(\mathcal{V}; a)$ defined on $\mathcal{G}_{\epsilon, \delta}(x; a)$ such that if $\mathcal{V} \in \mathcal{G}_{\epsilon, \delta}(x; a)$, then

$$u_{\epsilon, \delta, x}(y, t) = \sum_{(j, n) \in \mathbb{Z}^{d+v}} \hat{u}(j, n) e^{i n \cdot \omega t} \delta_j(y) \quad (1.15)$$

is a solution to (1.14) satisfying

$$\begin{aligned} \hat{u}(j_k, -e_k) &= a_k, \quad k = 1, \dots, v, \\ \sum_{(j, n) \notin \mathcal{S}} e^{c(|n|+|j|)} |\hat{u}(j, n)| &< \sqrt{\epsilon + \delta}, \\ |\omega - \mathcal{V}| &< c(\epsilon + \delta), \end{aligned} \quad (1.16)$$

for some $c > 0$, where $\{e_k\}_{k=1}^v$ are the basis vectors for \mathbb{Z}^v and $\mathcal{S} = \{j_k, -e_k\}_{k=1}^v \subset \mathbb{Z}^{d+v}$. The sets X_ϵ and $\mathcal{G}_{\epsilon, \delta}(x; a)$ satisfy

$$\text{Prob } X_\epsilon \rightarrow 1, \quad \text{mes } \mathbb{R}^v \setminus \mathcal{G}_{\epsilon, \delta}(x; a) \rightarrow 0 \quad \text{as } \epsilon + \delta \rightarrow 0.$$

Remark. The set $X_\epsilon \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^v$ only depends on ϵ ; while $\mathcal{G}_{\epsilon, \delta}(x; a) \subset \mathbb{R}^v$ depends on $\epsilon, \delta, x \in X_\epsilon$ (the random potentials in X_ϵ) and a (the initial amplitude).

Corollary. For $0 < \epsilon, \delta \ll 1$, there exists $X_{\epsilon, \delta} \subset \mathbb{R}^{\mathbb{Z}^d}$ of positive probability, satisfying

$$\text{Prob } X_{\epsilon, \delta} \rightarrow 1 \quad \text{as } \epsilon + \delta \rightarrow 0,$$

such that for initial amplitudes a sufficiently small, there are quasi-periodic solutions to (1.14).

Comments on the family of parameters $\{v_j\}_{j \in \mathbb{Z}^d}$

In solving (1.14), we use the basis $e^{i n \cdot \omega t} \delta_j$, $(j, n) \in \mathbb{Z}^{d+v}$ (cf. (1.8)). In the \mathbb{Z}^d basis δ_j ($j \in \mathbb{Z}^d$), the linear operator $H = \epsilon \Delta + V$ is not diagonalized. Hence $\{v_j\}_{j \in \mathbb{Z}^d}$ is not

a family of independent parameters. This is a slight variation from the “usual” scenario, where the linear operator is diagonalized and the parameters are independent, which is the case in, e.g., [B3].

Here it is convenient to work with the \mathbb{Z}^d basis δ_j instead of the basis provided by the eigenfunctions ψ_j of H , as ψ_j depends on $\{v_k\}_{k \in \mathbb{Z}^d}$. More precisely, as $\{v_k\}_{k \in \mathbb{Z}^d \setminus \mathcal{R}}$ is held fixed on the appropriate probability subspace, ψ_j depends on $\{v_{k'}\}_{k' \in \mathcal{R}}$, which serve as parameters for the construction and are therefore varying (see the statement of the Theorem).

From the KAM perspective, the normal frequencies are provided by the eigenvalues μ_j of H . Since $\{v_k\}_{k \in \mathbb{Z}^d \setminus \mathcal{R}}$ is fixed, the strong localization property (A8) (see Appendix) of ψ_j implies that the normal frequencies μ_j for $|j| > \rho$, where ρ only depends on the radius of \mathcal{R} , can in fact be held fixed. This is close to the usual terrain, where the normal frequencies are fixed, while the tangential frequencies vary to avoid small divisors, either via the parameters or via amplitude-frequency modulation (cf. [B3, KP]).

Insertion into a larger picture

The Theorem presented above is proven for i.i.d. random potentials $V = \{v_j\}_{j \in \mathbb{Z}^d}$. The construction used to prove the theorem is, however, general. The essential ingredient is a *spectral separation property* of the linearized operator $\tilde{H} = n \cdot \omega + H$, where ω are the tangential frequencies, H is the original linear operator (corresponding to the quadratic part of the Hamiltonian, cf. (2.2)). In the present case, $H = \epsilon \Delta + V$. Assume H has pure point spectrum and look at initial conditions localized about the origin. Below is a tentative formulation of this spectral property.

Let μ_j be the eigenvalues of H . For $i = (j, n)$, let $\lambda_i = n \cdot \omega + \mu_j$ be the eigenvalues of \tilde{H} . Let χ be an appropriate function, which depends essentially only on the initial condition, localized about the origin, $|\chi| \lesssim 1$. Let ϕ_i, ϕ'_i be eigenfunctions of \tilde{H} (i.e., products of eigenfunctions of H and the exponentials). Define

$$K(i, i') = \int \phi_i \chi \phi_{i'}.$$

\tilde{H} has the *spectral separation property* if for each scale L , there exists $\ell \ll L$ such that

$$|\lambda_i - \lambda_{i'}| \gg K(i, i') \quad (1.17)$$

for $\ell \leq |i - i'| \leq L$ ($i \neq i'$). The μ_j, λ_i, ϕ_i can be replaced by their local versions whenever appropriate.

In the present case, $H = \epsilon \Delta + V$, we use the local version. Assume ϵ is small so that H has A.L. (1.17) is provided by using (A5–7) and restricting to the appropriate probability subspace, (2.10) and a direct incision in the frequency space. Related spectral separation properties seem to hold in [B3, W]. (Compare (1.17) with the nondegeneracy condition in [KP, p. 164], where eigenfunctions do not seem to play an explicit role.)

Remark. For the random Schrödinger operator $H = \epsilon \Delta + V$ ($\epsilon \ll 1$), no Diophantine property of the eigenvalues seems to be known at present. So a possible extension of the standard KAM method, as outlined in, e.g., [FSW] is not feasible. It is known from [Mi], however, that the eigenvalue statistics is Poisson and that in a box of size N , the eigenvalue spacing is N^{-p} ($p \geq d$). From general considerations, the spectrum $\sigma(H)$ is simple [Si].

The construction of time quasi-periodic (or almost-periodic) solutions needs a parameter. This parameter can sometimes be extracted from amplitude-frequency modulation (see, e.g., [KP]). The nonlinear random Schrödinger equation is an equation endowed with a family of parameters, where the separation property (1.17) can be obtained from A.L. of the linear operator. So it is a natural candidate for the construction of KAM type solutions.

The continuum Schrödinger equations (linear or nonlinear) are a more frequently studied subject. The discrete nonlinear Schrödinger equation presented here should be seen as the analogue of the continuum nonlinear Schrödinger equation in a compact domain, e.g., on a torus. The \mathbb{Z}^d lattice can therefore be seen as the indices of the eigenvalues or eigenfunctions for the underlying linear Schrödinger operator.

Time quasi-periodic solutions have been constructed for the continuum nonlinear Schrödinger or wave equation in 1-D, on a finite interval with either Dirichlet or periodic boundary conditions. See for example the works of Bourgain, Kuksin, Pöschel and Wayne in [B1, KP, Wy]. In [B3], time quasi-periodic solutions are constructed for the 2-D nonlinear Schrödinger equation on \mathbb{T}^2 . In arbitrary dimension, time quasi-periodic solutions for nonlinear Schrödinger and wave equations are treated in [B5, EK].

The construction presented here is related to those in [B1–5], which use a Newton scheme directly on the equations. This direct approach is originated by Craig and Wayne in [CW1,2]. It has the advantage of not relying on the underlying Hamiltonian structure. The Hamiltonian structure does ensure, however, that the frequency ω is real during the iteration (see Section 2 and [B3]).

We end this section by remarking that the present method, as it stands, does not yet extend to a construction of almost-periodic solutions. This is because the linear equation that serves as the starting point of our perturbation is

$$i \frac{\partial}{\partial t} u = V u,$$

and *not*

$$i \frac{\partial}{\partial t} u = (\epsilon \Delta + V) u.$$

In order to construct almost-periodic solutions, we will need more information on the spectrum of the linear operator $H = \epsilon \Delta + V$.

In [B2], the construction of almost-periodic solutions for 1-D nonlinear Schrödinger and wave equations under Dirichlet boundary conditions was made possible by the precise knowledge of the spectrum of the linear operator and the fact that the perturbation is quartic (in the Hamiltonian). In the present case it is quadratic. In [B6], almost-periodic

solutions for a 1-D nonlinear Schrödinger equation under periodic boundary conditions and realistic decay conditions were constructed. In particular this applies in the real analytic category. Almost-periodic solutions have also been constructed by Pöschel [Pö2] in the case of a nonlinear Schrödinger equation, where the nonlinearity is “nonlocal”.

PDE’s (such as (1.1)) typically correspond to the so called “short range” (but not finite range) case. In the “finite range” case, which typically corresponds to perturbation of integrable Hamiltonian systems, almost-periodic solutions have been constructed in [CP, FSW, Pö1] among others.

2. Hamiltonian representation and Lyapunov–Schmidt decomposition

Recall from Section 1 the nonlinear random Schrödinger equation

$$i \frac{\partial}{\partial t} u = (\epsilon \Delta + V)u + \delta |u|^{2p} u \quad (p \in \mathbb{N}^+), \quad (2.1)$$

where $0 < \epsilon, \delta \ll 1$, Δ is the discrete Laplacian as defined in (1.2), and $V = \{v_j\}_{j \in \mathbb{Z}^d}$ are i.i.d. random variables with common distribution g as in (1.3). The solutions are $u = \{u(j, t)\}_{j \in \mathbb{Z}^d, t \in [0, \infty)}$.

Equation (2.1) can be recast as (infinite-dimensional) Hamiltonian equations of motion, with canonical variables (u, \bar{u}) and the Hamiltonian

$$\begin{aligned} H(u, \bar{u}) &= \frac{1}{2} \left[\sum_{j, j' \in \mathbb{Z}^d \times \mathbb{Z}^d} (\epsilon \Delta + V)_{jj'} u_j \bar{u}_{j'} + \frac{\delta}{p+1} \sum_j u_j^{p+1} \bar{u}_j^{p+1} \right] \\ &=: H_0(u, \bar{u}) + \delta H_1(u, \bar{u}). \end{aligned} \quad (2.2)$$

Equation (2.1) can then be written as

$$i \dot{u} = 2 \frac{\partial H}{\partial \bar{u}}. \quad (2.3)$$

Remark. The connection with the usual canonical variables (p, q) is $u = p + iq$, $\bar{u} = p - iq$. The equation of motion in the (p, q) coordinates is

$$\dot{p} = \frac{\partial H}{\partial q}, \quad \dot{q} = -\frac{\partial H}{\partial p},$$

which can be rewritten as a single equation (2.3). (This also explains the factors i and 2 .)

Equations (2.2, 2.3) show that (2.1) can be viewed as a perturbed \mathbb{Z}^d system of coupled harmonic oscillators with i.i.d. random frequencies. The perturbation H_1 can be of a more general type, e.g.,

$$H_1(u, \bar{u}) = \sum_{j, j' \in \mathbb{Z}^d \times \mathbb{Z}^d} a_{jj'} u_j^{p+1} \bar{u}_{j'}^{p+1} \quad (2.4)$$

with $a_{jj'} = a_{j'j}$ decaying exponentially or polynomially of sufficiently high degree as $|j - j'| \rightarrow \infty$. The reason we mention (2.4) is to stress that the construction we present

below *does not* rely on integrability of the system. It also carries through for H_1 of type (2.4), although we only present it for $a_{jj'} = \delta_{jj'}$.

The goal of the rest of the paper is to seek time quasi-periodic solutions to (2.1) for appropriately chosen *localized* initial conditions. We hence expand u in the basis

$$e^{in \cdot \omega t} \delta_k(j), \quad (2.5)$$

where $n \in \mathbb{Z}^\nu$, $\omega \in \mathbb{R}^\nu$, $k, j \in \mathbb{Z}^d$, $\delta_k(j)$ is the canonical basis for \mathbb{Z}^d . $\delta_k(j)$ is a natural basis here due to smallness of ϵ . (In [B1–3], the spatial basis is given by the eigenfunctions of the linear operator. The k -labeling there is the eigenvalue labeling.)

In the basis (2.5), (2.1) becomes

$$(n \cdot \omega + \epsilon \Delta_j + V_j) \hat{u}(j, n) + \delta \frac{\partial \widehat{H}_1}{\partial \bar{u}}(j, n) = 0, \quad (2.6)$$

where $n \in \mathbb{Z}^\nu$, $j \in \mathbb{Z}^d$, H_1 is defined in (2.2) and \hat{u} are the Fourier coefficients of u :

$$u(k, t) = \sum_{(j, n)} \hat{u}(j, n) e^{in \cdot \omega t} \delta_j(k). \quad (2.7)$$

We have also put the subscript j on operators that operate in the spatial (\mathbb{Z}^d) variable only. (This is the same notation as in [BW].)

In view of the Theorem, we seek solutions to (2.6) with the constraint

$$\hat{u}(j_k, -e_k) = a_k \quad (k = 1, \dots, \nu), \quad (2.8)$$

where $j_k \in \mathbb{Z}^d$, e_k are unit vectors in \mathbb{Z}^ν , a_k are fixed. Assume $\omega_1, \dots, \omega_\nu$ are rationally independent, i.e., $\omega = \{\omega_i\}_{i=1}^\nu \in \mathbb{R}^\nu$ is a Diophantine vector, which will be the case when the Theorem applies. A time shift and a limiting argument (since the Kronecker flow is dense) permit us to assume the a_k are real. Hence from now on $a_k \in \mathbb{R}$, $k = 1, \dots, \nu$.

Due to the smallness of ϵ , we take our initial unperturbed linear equation to be

$$i \frac{\partial}{\partial t} u = V u. \quad (2.9)$$

The conditions in (2.8) thus correspond to the initial unperturbed solution

$$u_0(k, t) = \sum_{\ell=1}^\nu a_\ell e^{-iv_\ell t} \delta_{j_\ell}(k). \quad (2.10)$$

Let

$$a = \{a_k\}_{k=1}^\nu \in \mathbb{R}^\nu, \quad \mathcal{R} = \{j_k\}_{k=1}^\nu \subset \mathbb{Z}^d, \quad \mathcal{V} = \{v_\alpha\}_{\alpha \in \mathcal{R}} \in \mathbb{R}^\nu.$$

We constructively show that for ϵ small enough, there exists $X_\epsilon \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^\nu$ of positive probability, satisfying $\text{Prob } X_\epsilon \rightarrow 1$ as $\epsilon \rightarrow 0$, such that if we fix $x \in X_\epsilon$, then for δ, a small enough, there exists a Cantor set $\mathcal{G}_{\epsilon, \delta}(x; a) \subset \mathbb{R}^\nu$ of positive measure, satisfying $\text{mes } \mathbb{R}^\nu \setminus \mathcal{G}_{\epsilon, \delta}(x; a) \rightarrow 0$ as $\epsilon + \delta \rightarrow 0$. We can find a smooth function $\omega = \omega(\mathcal{V}; a)$

defined on $\mathcal{G}_{\epsilon, \delta}(x; a)$ and \hat{u} such that (2.6) holds. ω and \hat{u} are determined simultaneously in an iterative way.

Toward that end, we first perform a Lyapunov–Schmidt type decomposition (see [B1-3, CW1,2]) of (2.6). Let

$$\mathcal{S} = \{(j_k, -e_k) \mid k = 1, \dots, v\} \subset \mathbb{Z}^{d+v}. \quad (2.11)$$

From (2.10), $\mathcal{S} = \text{supp } u_0$, u_0 is a solution to (1.14) when $\epsilon = \delta = 0$. We call \mathcal{S} the *resonant set* and consider the v equations

$$[(n \cdot \omega + \epsilon \Delta_j + V_j)\hat{u}](j_k, -e_k) + \delta \frac{\partial \widehat{H}_1}{\partial \bar{u}}(j_k, -e_k) = 0 \quad (k = 1, \dots, v) \quad (2.12)$$

obtained by taking $(j, n) \in \mathcal{S}$.

They form the finite system of *Q-equations*. The remaining infinite system of equations are called the *P-equations*

$$(n \cdot \omega + \epsilon \Delta_j + V_j)\hat{u}(j, n) + \delta \frac{\partial \widehat{H}_1}{\partial \bar{u}}(j, n) = 0, \quad (j, n) \notin \mathcal{S}. \quad (2.13)$$

The *P-equations* are used to determine $\hat{u}(j, n)$ for $(j, n) \notin \mathcal{S}$. (Recall from (2.8) that $\{\hat{u}(j, n), (j, n) \in \mathcal{S}\} = a$ are given.)

Once $\hat{u}(j, n)$ are determined, the *Q-equations* in (2.12) are used to determine $\omega = \omega(\mathcal{V}, a)$ via the implicit function theorem. Since a is real, H_1 is a polynomial in u, \bar{u} with real coefficients, the solution \hat{u} to (2.13) will be real and hence also ω determined from (2.12). (For more details, see the comment after (2.3) of [B3].)

To solve (2.13), we duplicate the equation for \bar{u} to form a closed system. Let

$$\begin{aligned} v &= \bar{u}, \\ \hat{v}(j, n) &= \bar{\hat{u}}(j, -n), \\ -\mathcal{S} &= \{(j_k, +e_k) \mid k = 1, \dots, v\} \subset \mathbb{Z}^{d+v}. \end{aligned} \quad (2.14)$$

(The flip in sign in the second equation of (2.14) is solely in order that the convolution coming from the nonlinearity obeys the usual sign convention.)

We then have the closed system of *P-equations*

$$\begin{cases} (n \cdot \omega + \epsilon \Delta_j + V_j)\hat{u}(j, n) + \delta \frac{\partial \widehat{H}_1}{\partial v}(j, n) = 0, & (j, n) \notin \mathcal{S}, \\ (-n \cdot \omega + \epsilon \Delta_j + V_j)\hat{v}(j, n) + \delta \frac{\partial \widehat{H}_1}{\partial u}(j, n) = 0, & (j, n) \notin -\mathcal{S}. \end{cases} \quad (2.15)$$

For H_1 as in (2.2), (2.15) takes the explicit form

$$\begin{cases} [(n \cdot \omega + \epsilon \Delta_j + V_j)\hat{u}](j, n) + \delta[(\hat{u} * \hat{v})^{*p} * \hat{u}](j, n) = 0, \\ [(-n \cdot \omega + \epsilon \Delta_j + V_j)\hat{v}](j, n) + \delta[(\hat{u} * \hat{v})^{*p} * \hat{v}](j, n) = 0, \end{cases} \quad (j, n) \in \mathbb{Z}^{d+v} \setminus (\mathcal{S} \cup -\mathcal{S}), \quad (2.16)$$

where the convolution $*$ is in the n variable only. We solve (2.16) by using a Newton iteration scheme to be amplified in the next section. We also identify \hat{u} with u , \hat{v} with v and write y for $\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}$.

3. Newton scheme

Let F denote the left hand side (LHS) of (2.16). Our task is to restrict the set of (ω, \mathcal{V}) in $\mathbb{R}^{2\nu}$ in order to find y such that

$$F(y) = 0, \quad (3.1)$$

so that (2.16) is resolved. We use a Newton iteration. Recall first the formal scheme.

Starting from the initial approximant y_0 , a solution to (1.14), and its conjugate when $\epsilon = \delta = 0$, the successive approximants y_i are defined by

$$\Delta_{i+1}y := y_{i+1} - y_i = -[F'(y_i)]^{-1}F(y_i). \quad (3.2)$$

Let T denote the linearized operator F' . From (2.16),

$$T = D + \delta S, \quad (3.3)$$

where D is diagonal (in the $n \in \mathbb{Z}^\nu$ variables)

$$D = \begin{pmatrix} n \cdot \omega + \epsilon \Delta_j + V_j & 0 \\ 0 & -n \cdot \omega + \epsilon \Delta_j + V_j \end{pmatrix} =: \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix} \quad (3.4)$$

and

$$S = S(u, v) = \begin{pmatrix} (p+1)(u * v)^{*p} & p(u * v)^{*p-1} * u * u \\ p(u * v)^{*p-1} * v * v & (p+1)(u * v)^{*p} \end{pmatrix} \quad (p \in \mathbb{N}^+) \quad (3.5)$$

evaluated along the previous approximant. We note that S is self-adjoint, although this does not play a role in our construction.

Denote by $\| \cdot \|$ the ℓ^2 norm of a vector or an operator on $\ell^2(\mathbb{Z}^{d+\nu})$. By (3.2), the error of the approximation at stage $i+1$ can be estimated:

$$F(y_{i+1}) = F(y_i) + F'(y_i)(y_{i+1} - y_i) + \mathcal{O}(\|y_{i+1} - y_i\|^2) = \mathcal{O}(\|y_{i+1} - y_i\|^2). \quad (3.6)$$

So by (3.2),

$$\|F(y_{i+1})\| = \mathcal{O}(\|[F'(y_i)]^{-1}\|^2 \|F(y_i)\|^2). \quad (3.7)$$

The crux of the matter is thus to control $\|[F'(y_i)]^{-1}\|$ in order that

$$\|F(y_{i+1})\| \ll \|F(y_i)\|. \quad (3.8)$$

(Note the squaring of the norm of $F(y_i)$ on the RHS of (3.7), which makes this feasible.)

Since (3.1) represents an infinite system of equations and the initial condition (2.10) is localized in a compact region in $\mathbb{Z}^{d+\nu}$, to control the norm of $[F'(y_i)]^{-1}$ we implement the Newton scheme in a slightly modified way, gradually increasing the size of the system that we consider.

Let $M \in \mathbb{N}^+$, which can be assumed large in order that $[-M, M]^{d+\nu} \supset 2p \operatorname{supp} y_0$, in view of (3.4, 3.5) (see also (3.10) below). At stage i , let $N = M^{i+1}$ and let $T_N(y_i)$ be the restriction of $T(y_i)$ to $[-N, N]^{d+\nu}$. We define

$$\Delta_{i+1}y = y_{i+1} - y_i := -[T_N(y_i)]^{-1}F(y_i). \quad (3.9)$$

So

$$\begin{aligned}
F(y_{i+1}) &= F(y_i) + F'(y_i)(y_{i+1} - y_i) + \mathcal{O}(\|y_{i+1} - y_i\|^2) \\
&= (T - T_N)(y_{i+1} - y_i) + \mathcal{O}(\|y_{i+1} - y_i\|^2) \\
&= -[(T - T_N)T_N^{-1}]F(y_i) + \mathcal{O}(\|T_N^{-1}\|^2\|F(y_i)\|^2), \tag{3.10}
\end{aligned}$$

where we used (3.9). Compared to (3.6) the first term on the RHS of (3.10) is new. Moreover it is only *linear* in $F(y_i)$. This necessitates the control of off-diagonal decay of T and T_N^{-1} evaluated at y_i , in addition to that of $\|T_N^{-1}\|$.

The control of T_N^{-1}

Recall that y_0 , the 0th approximant (initial condition) to (3.1), has compact support, $\text{supp } y_0 = \mathcal{S} \cup -\mathcal{S}$, where \mathcal{S} and $-\mathcal{S}$ are defined in (2.11, 2.14). From (3.3–3.5), $T(y_0)$ is a diagonal dominated matrix with finite range off-diagonal elements. So off-diagonal decay of $T(y_0)$ is automatically satisfied.

Assume the successive approximants y_i are *uniformly* (in i) exponentially localized about $\mathcal{S} \cup -\mathcal{S}$ (cf. (1.16)). This assumption will be verified later from the construction itself in view of (3.9, 3.10). From (3.5) the successive $S(y_i)$ have *uniformly* exponentially decaying off-diagonal elements in the n direction, and are diagonal in the j direction, with a prefactor which decays exponentially in j . (The exponential decay of the prefactor stems from the uniform exponential decay assumption on y_i .) Hence $T(y_i)$ are of the type (although more complicated) of the matrix operator studied in [BW].

To study the T 's, we introduce, as in [BW], an auxiliary parameter $\theta \in \mathbb{R}$. We consider instead

$$T^\theta = D^\theta + \delta S, \tag{3.11}$$

where

$$\begin{aligned}
D^\theta &= \begin{pmatrix} n \cdot \omega + \theta + \epsilon \Delta_j + V_j & 0 \\ 0 & -(n \cdot \omega + \theta) + \epsilon \Delta_j + V_j \end{pmatrix} \\
&=: \begin{pmatrix} D_+^\theta & 0 \\ 0 & D_-^\theta \end{pmatrix} \tag{3.12}
\end{aligned}$$

and S is as before in (3.5). Similarly we define $T_N^\theta(y_i)$, where $N = M^{i+1}$ as in (3.9).

In Section 4, we fix x in a good set of probability space, where there is Anderson localization for the linear random Schrödinger operator $H_j = \epsilon \Delta_j + V_j$, so that (A6) holds. (For precise details see the Appendix.) Assuming ω Diophantine, y_i uniformly (in i) exponentially localized about $\mathcal{S} \cup -\mathcal{S}$, we bound the norm of $[T_N^\theta(y_{i'})]^{-1}$, where $N = M^i$, $i' > i$ as in (4.9) (the precise relation between i and i' is dictated by the construction in Section 5) and establish exponential decay properties of its off-diagonal matrix elements on a set of θ of small complementary measure.

In Lemma 4.1, we fix any y_k , and we bound $[T_N^\theta(y_k)]^{-1}$ for all N . We then use it to obtain estimates on $[T_N^\theta(y_{i'})]^{-1}$, where $N = M^i$, $i' > i$ satisfying the restriction in the third line of (4.9). This bound is abstract in the sense that ω , \mathcal{V} , $y_{i'}$ are viewed as *independent* parameters for the time being.

As in [BW], this is an iteration process, using semi-algebraic set techniques and a Cartan-type theorem. To start the iteration, we neglect δS and exclude a set of θ such that $D_N^\theta = D_M^\theta$ has a small diagonal element. To continue the iteration, we also need to exclude a set of ω of small measure. It is important to remark that this set of ω is *independent* of \mathcal{V} , y_k . It only depends on $x \in \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^{\mathbb{V}}$. We stress that for fixed x in the good probability set, Lemma 4.1 holds for *any* fixed ω in the good frequency set, *any* $\mathcal{V} \in \mathbb{R}^{\mathbb{V}}$ and *any* y_k which satisfy (H1–3) in Section 4. The set $\mathcal{B}(N)$ of excluded θ depends, of course, on x , ω , \mathcal{V} , y_k .

In Section 5, we iteratively transfer the estimates on $T_N^\theta(y_{i'})$ in θ into estimates on $T_{\bar{N}}(y_{\bar{i}})$ in (ω, \mathcal{V}) , where $N = M^{i+1}$, $\bar{N} = M^{\bar{i}+1}$, $\bar{i} > i' > i$ to be made precise, using the resolvent equation and taking into account the Q -equations, which are implicit functions relating ω , \mathcal{V} , y_i . (Recall that θ is an auxiliary variable. In the original problem (3.3), θ is fixed at 0.) x is fixed in the good set of probability space as in Section 4.

For the first i iterations, we treat δS as a perturbation and use a direct ϵ, δ perturbation series. Instead of excluding a set of θ as in Section 4, we exclude a set of $(\omega, \mathcal{V}) \in \mathbb{R}^{2\mathbb{V}}$, so that in the complement, $T_N(y_i)$ are invertible with exponentially decaying off-diagonal elements.

This generates an initial set of “good” intervals: $\mathbb{R}^{2\mathbb{V}} \supset \Lambda_1 \supset \cdots \supset \Lambda_i$ in the (ω, \mathcal{V}) space. The use of the Newton scheme in (3.9) also shows that y_0, y_1, \dots, y_i are exponentially localized about $\mathcal{S} \cup -\mathcal{S}$. (Recall that y_0 is the initial condition, $\text{supp } y_0 = \mathcal{S} \cup -\mathcal{S}$.)

Starting from the $(i+1)^{\text{th}}$ iteration, aside from direct δ perturbation series, for certain parts of the estimates (which concerns the regions far from the origin in the \mathbb{Z}^d direction), we need to keep δS . This is the heart of the matter. In technical terms, we need to deal with more general semi-algebraic sets, which are not solely defined by products of monomials. For such semi-algebraic sets, we use Q -equations and a decomposition lemma (Lemma 9.9 in [B5], restated as Lemma 5.3 in Section 5) to transfer the measure estimates in θ in Lemma 4.1 into measure estimates in $\omega = \omega(\mathcal{V}) \in \mathbb{R}^{\mathbb{V}}$. Using perturbation, this in turn generates a new set of intervals $\Lambda_{i+1} \subset \Lambda_i \subset \cdots \subset \mathbb{R}^{2\mathbb{V}}$, in the (ω, \mathcal{V}) space, on which $T_{M^{i+1}}(y_i)$ is invertible and whose inverse has uniformly (in i) exponentially decaying off-diagonal elements.

In Section 6, using the Newton scheme (3.9), we construct y_{i+1} . The (uniform in i) exponential localization about $\mathcal{S} \cup -\mathcal{S}$ is preserved. Hence Lemma 4.1 is now available at $y_k = y_{i+1}$ for future iterations. We generate $\Lambda_{i+2}, y_{i+2}, \dots$

Section 7 summarizes the entire construction. It is merely meant as a recapitulation of the sequence of events leading to the Theorem.

Two technical subtleties

- The estimates in Section 4 are obtained following the construction devised in [BW]. However, for the application to (4.9), $T_N^\theta(y_{i'})$ need to be evaluated at different $y_{i'}$ at different scales N . Due to the *uniform* exponential decay estimates on $y_{i'}$, Lemma 4.1 can be applied as explained after its statement.

- From the P -equations, the y_i 's are constructed on a *good* set of $(\omega, \mathcal{V}) \in \mathbb{R}^{2\nu}$. (This set eventually becomes a Cantor set.) On the same set of (ω, \mathcal{V}) we also have estimates on ∂y_i , where ∂ is with respect to ω or \mathcal{V} . Using this, we can construct y_i which is smoothly defined on the whole (ω, \mathcal{V}) parameter space. (Note that outside the good set of (ω, \mathcal{V}) , y_i is no longer close to a solution to $F(y_i) = 0$.) Substituting for y_i , the Q -equations are therefore defined on the whole (ω, \mathcal{V}) parameter space. We can then use the standard implicit function theorem to determine $\omega = \omega(\mathcal{V})$.

4. P -equations and statement in θ

Recall the system of P -equations in (2.16):

$$\begin{cases} (n \cdot \omega + \epsilon \Delta_j + V_j) \hat{u} + \delta(\hat{u} * \hat{v})^{*p} * \hat{u} = 0, \\ (-n \cdot \omega + \epsilon \Delta_j + V_j) \hat{v} + \delta(\hat{u} * \hat{v})^{*p} * \hat{v} = 0, \end{cases} \quad (4.1)$$

on $\ell^2(\mathbb{Z}^{d+\nu} \setminus (\mathcal{S} \cup -\mathcal{S}))$, where $\mathcal{S}, -\mathcal{S}$ are as defined in (2.11, 2.14), which are collectively written as $F(y) = 0$, with $y = \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$.

We solve (4.1) using a Newton scheme, with the family of linearized operators $T(y_i)$, evaluated at the i^{th} approximant y_i ,

$$T(y_i) = D + \delta S(y_i), \quad (4.2)$$

where

$$D = \begin{pmatrix} n \cdot \omega + \epsilon \Delta_j + V_j & 0 \\ 0 & -n \cdot \omega + \epsilon \Delta_j + V_j \end{pmatrix} = \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix} \quad (4.3)$$

and

$$S(y_i) = \begin{pmatrix} (p+1)(u_i * v_i)^{*p} & p(u_i * v_i)^{*p-1} * u_i * u_i \\ p(u_i * v_i)^{*p-1} * v_i * v_i & (p+1)(u_i * v_i)^{*p} \end{pmatrix} \quad (4.4)$$

as in (3.4, 3.5).

In view of the Newton scheme in (3.9), we need to study the family of restricted operators $T_N(y_i)$, $N = M^{i+1}$, M assumed large depending on p ,

$$T_N(y_i) = R_N T(y_i) R_N, \quad (4.5)$$

where R_N is the characteristic function of the set $[-N, N]^{d+\nu}$. This will be achieved in Section 5 by using the resolvent identity, covering $[-M^{i+1}, M^{i+1}]^{d+\nu}$ with the interval

$$I = [-M^i, M^i]^{d+\nu} \text{ and smaller intervals } J = [-M_0, M_0]^{d+\nu} + k, \quad (4.6)$$

$$\frac{1}{2}M^i < |k| < M^{i+1}, \quad M_0 \sim (\log N)^{C/2}$$

(see (5.5)), and restricting the set of $(\omega, \mathcal{V}) \in \mathbb{R}^{2\nu}$.

Toward that end, as previously mentioned in Section 3, we introduce an additional parameter $\theta \in \mathbb{R}$ and let

$$D^\theta = \begin{pmatrix} D_+ + \theta & 0 \\ 0 & D_- - \theta \end{pmatrix}, \quad (4.7)$$

$$T^\theta(y_i) = D^\theta + \delta S(y_i), \quad T_N^\theta(y_i) = R_N T^\theta(y_i) R_N.$$

As mentioned there, we temporarily view $\omega, \mathcal{V} \in \mathbb{R}^v$ as independent parameters in this section. In the same vein, we also dissociate y_i from ω, \mathcal{V} , assuming only that they satisfy (H1–3) below. It is only in Section 5 that we restrict to $\omega = \omega(\mathcal{V})$, determined from the Q -equations and y_i , the i^{th} approximate solution to (4.1), which depends on ω, \mathcal{V} .

In the rest of the section ω, \mathcal{V} are held *fixed*, only θ is varying. Note that T^θ is of the form $T^\theta := T(\theta) = T'(\theta + n \cdot \omega)$. Later in Section 5, we transfer the estimate in θ into estimates in ω , hence \mathcal{V} by restricting θ to be of the form $\theta = n \cdot \omega$, thereby resolving (4.1) (which is at $\theta = 0$) on the good set of ω, \mathcal{V} .

The Newton scheme is an iterative scheme. The estimate on $[T_I(y_i)]^{-1}$ for I defined in (4.6) is easily obtained by perturbation arguments on $[T_I(y_{i-1})]^{-1}$ known from the previous step, which is the step to construct y_i (see (3.9)). The main task is to estimate $[T_J(y_i)]^{-1}$ for J defined in (4.6). We therefore study $[T_{[-M_0, M_0]^{d+v}}^\theta(y_i)]^{-1}$ (and later in Section 5, we restrict to $\theta = k \cdot \omega, \frac{1}{2}M^i < |k| < M^{i+1}$). This is the subject of Lemma 4.1 and its application. Note that M_0 corresponds to the size of the interval at a stage $[\log M_0 / \log M]$, which precedes i , while the linearized operator T is evaluated at y_i : $T = T(y_i)$.

Assume y_i satisfies

$$(H1) \quad \text{supp } y_i \subseteq [-M^i, M^i]^{d+v} \quad (i \geq 1).$$

(This is by construction, see (3.9).)

$$(H2') \quad \|\Delta_i y\| = \|y_i - y_{i-1}\| < \sqrt{\epsilon + \delta} M^{-b^i} \quad (i \geq 1)$$

for some $1 < b < 2$ in view of (3.6, 3.9); b will be specified in (6.20). (Recall $y_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$, u_0 defined in (2.10), $v_0 = \bar{u}_0$.)

$$(H3) \quad |y_i(k)| < e^{-\alpha|k|} \quad \text{for some } \alpha > 0 \text{ (uniform in } i).$$

(There is no constant in front of the exponential, as we assume small initial data: $|a_\ell| < 1, \ell = 1, \dots, v$. See (1.9).)

Remark. Using (3.6) in (3.9) we get $\|\Delta_i y\|^2 < \|\Delta_{i+1} y\| < \|\Delta_i y\|$, assuming an appropriate condition on T_N^{-1} . This is consistent with $1 < b < 2$ in (H2').

(H1–3) will be verified along the iteration in Sections 5, 6 using Lemma 4.1 below.

Let $\Lambda_N = [-N, N]^d$, and let $X_N \subset \mathbb{R}^{\Lambda_N} \setminus \mathbb{R}^v$ be a set, where $\epsilon \Delta_j + V_j$ has A.L. at scale N , in a sense to be made explicit in the process of the proof; X_N is asymptotically ($\epsilon \rightarrow 0$) of full measure. (Recall also that $\mathcal{V} = \{v_{jk}\}_{k=1}^v \in \mathbb{R}^v$ with measure $\prod_{k=1}^v g(v_{jk}) dv_{jk}$ is the parameter set.)

Definition. For $A, c > 0$, $DC_{A,c}(N) \subset \mathbb{R}^v$ is the set of ω such that

$$\|n \cdot \omega\|_{\mathbb{T}} \geq \frac{c}{|n|^A}, \quad n \in [-N, N]^v \setminus \{0\}. \quad (4.8)$$

$DC_{A,c} \subset \mathbb{R}^v$ is the set of ω such that (4.8) is satisfied for all N .

$\mathcal{B}_{\beta,\gamma}(N) \subset \mathbb{R}$ is the complement of the set of $\theta \in \mathbb{R}$ such that

$$\begin{aligned} \|[T_N^\theta(y_{i'})]^{-1}\| &< e^{N^\beta} \quad (0 < \beta < 1), \\ |[T_N^\theta(y_{i'})]^{-1}(k, k')| &< e^{-\gamma|k-k'|} \quad (\gamma > 0), \\ \text{for } |k-k'| > N/10, N = M^{i'+1}, i' &\simeq \frac{\log M}{\log b} i > i \text{ (b as in (H2'))}, \end{aligned} \quad (4.9)$$

$y_{i'}$ satisfies (H1–3); i' is chosen in view of a later construction in Lemma 5.1 (see in particular (5.14, 5.15)). This means of course that at least the first $\mathcal{O}(\log M / \log b)$ approximants are obtained by direct perturbation series in ϵ, δ . So $\alpha = \mathcal{O}(1)|\log(\epsilon + \delta)|$. In general we write $\mathcal{B}(N)$ for $\mathcal{B}_{\beta,\gamma}(N)$, unless the parameters β, γ need to be emphasized.

Remark. At this stage of the construction, it is sufficient to have a lower bound on b . This can be easily obtained in the first few perturbation series by adjusting ϵ, δ . For later purposes, we mention that the Diophantine condition (4.8) will be used for $\omega = \omega_{i'}$, the i' th approximation.

The inequalities in (4.9) are proven iteratively as in [BW]. The rate of decay γ will deteriorate with iteration. So $\gamma = \gamma_N$. But the decrease will decrease with increasing scales and we have $\gamma_N > \gamma/2$ for all N (cf. Lemma 4.1 and the paragraph following it). This rate of decay determines the rate of decay of y_i . So this is consistent with the assumption (H3).

Inspecting the definition of D_θ in (4.7, 4.3), we see that θ is *not* equivalent to a spectral parameter. Hence we need to resort to Cartan-type theorems as in the wave case in [BW]. This necessitates that we obtain estimate (4.9) for more general regions than cubes at each scale N , the elementary regions to be defined below.

Remark. The various approximants y_i are still evaluated using cubes $[-N, N]^{d+v}$, $N = M^{i'+1}$, as in (3.9). It is only at each scale N that we also look at restrictions of $T(y_{i'})$ ($i' \simeq \frac{\log M}{\log b} i$) to these more general regions.

Elementary regions

An *elementary region* is a set Λ of the form

$$\Lambda := R \setminus (R + k), \quad k \in \mathbb{Z}^{d+v} \text{ is arbitrary}, \quad (4.10)$$

and R is a hyper-rectangle

$$\begin{aligned} R = \{\ell' \in \mathbb{Z}^{d+v} \mid |\ell'_i - \ell_i| \leq N_i, i = 1, \dots, d, d+1, \dots, d+v, \\ \ell' = \{\ell'_i\}_{i=1}^{v+d} \in \mathbb{Z}^{d+v}, \ell = \{\ell_i\}_{i=1}^{v+d} \in \mathbb{Z}^{d+v}\} \end{aligned} \quad (4.11)$$

Let $N = \max_i N_i =: N_{i_{\max}}$. Assume $\ell \in \mathbb{Z}^{d+v}$ is fixed. We call ℓ the *center* of R . Then $\mathcal{ER}(N)$ (at a fixed center) is defined to be the set of all regions obtained by varying $k \in \mathbb{Z}^{d+v}$ and N_i ($i \neq i_{\max}$), keeping $N_i \leq N$. We say $2N$ is the *diameter* of the elementary regions.

To be economical, we extend the notation T_N to mean $T_{\Lambda(N)} = R_{\Lambda(N)} T R_{\Lambda(N)}$ for any $\Lambda(N) \in \mathcal{ER}(N)$, where $R_{\Lambda(N)}$ is the characteristic function of the set $\Lambda(N)$; $\mathcal{B}(N)$ is then the corresponding bad set, on which (4.9) are violated. For the purpose of constructing approximate solutions, we only need to specialize to $N = M^{i+1}$ ($i \geq 0$). However, to state the various intermediate technical lemmas, it is more convenient to let N be any integer.

Fix any y_k satisfying (H1–3). Let T_N be the linearized operator evaluated at y_k for all N , i.e., $T_N = T_N(y_k)$. With a slight abuse of notation, we also let $\mathcal{B}_{\beta,\gamma}(N)$ be the corresponding bad set. The main goal of this section is to prove

Lemma 4.1. *Fix $0 < \sigma < 1/6(d + \nu)$, $\sigma < \beta < 1$, \bar{N}_0 sufficiently large, and $\max(1/\sigma, 6(d + \nu)) \ll C < \bar{N}_0^{\sigma/2}$. There exist $\epsilon_0, \delta_0 > 0$ such that for all $0 < \epsilon < \epsilon_0$, $0 < \delta < \delta_0$, there exists $X \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^\nu$ with*

$$\text{mes } X \geq 1 - \mathcal{O}(1)\bar{N}_0^{-\kappa}, \quad (4.12)$$

where $\kappa = \kappa(C, p', d) > 0$ and p' is as in (A2). Fix $x \in X$. Then there exists $\Omega \subset \mathbb{R}^\nu$ (independent of $\mathcal{V} \in \mathbb{R}^\nu$ and y_k), with

$$\text{mes } \Omega \leq e^{-\bar{N}_0^{\kappa'}},$$

where $\kappa' = \kappa'(C, \beta) > 0$, such that if

$$\omega \in DC_{A,c} \setminus \Omega \quad (4.13)$$

then for any $\Lambda(N) \in \mathcal{ER}(N)$ with $N \geq \bar{N}_0$,

$$\text{mes } \mathcal{B}_{\beta,\gamma_N}(N) \leq e^{-N^\sigma}, \quad (4.14)$$

where $\gamma_N \geq \alpha - \bar{N}_0^{-\delta'}$ ($\delta' > 0$) for all N , with α as in (H3), $\alpha = \mathcal{O}(1)|\log(\epsilon + \delta)|$.

Remark. $\mathcal{B}_{\beta,\gamma_N}(N)$ depends only on y_k, ω, \mathcal{V} as x is fixed. In the proof of Lemma 4.1, only (H3) on y_k is needed.

In order to obtain (4.9) at all scales, we apply Lemma 4.1 as follows. From the third line of (4.9), for any fixed y_k , we only need the lemma at scale $N = M^i$ with $i = \frac{\log b}{\log M} k$. To go to scale $N' = N^C = M^{iC}$ with the corresponding $y_{k'}, k' = \frac{\log M}{\log b} iC$, we first use (H2'), which gives

$$\|T_N^\theta(y_k) - T_N^\theta(y_{k'})\| \leq \mathcal{O}_{d+\nu}(1)\|y_k - y_{k'}\| \leq \mathcal{O}_{d+\nu}(1)M^{-b^k}. \quad (4.15)$$

(4.15) shows that we have essentially the same estimate on $\|T_N^\theta(y_{k'})\|$ as for $\|T_N^\theta(y_k)\|$.

We use N as the initial scale instead of \bar{N}_0 ; the proof of the inductive step in Lemma 4.1 then gives instead

$$\gamma_{N'}(y_{k'}) = \gamma_N(y_k) - N^{-\delta'} \quad (\delta' > 0), \quad N' = N^C. \quad (4.16)$$

From (3.9), the decay rate α_{i+1} of y_{i+1} is governed by the decay rate of $[T_N(y_i)]^{-1}$, where $N = M^{i+1}$. (Note that θ is fixed at 0.) This operator is treated in Section 5 using several

considerations including Lemma 4.1, (4.9), the resolvent equation and semi-algebraic sets. The decay rate α_{i+1} depends on γ_{M_0} , where $M_0 \ll N = M^{i+1}$ is determined in (5.5) (cf. also (6.1)–(6.8)). So (4.16) prevents the deterioration of α_i as $i \rightarrow \infty$ and we will have $\alpha_i > \alpha/2$ for all i in Section 6.

We prove Lemma 4.1 using iteration. The two pillars of this iteration are semi-algebraic set techniques and a Cartan-type theorem for analytic matrix-valued functions (see [B5, BGS]). The general construction of the iteration is the same as in our previous paper [BW].

The initial estimate (0th step)

In view of (4.4) and the Newton scheme (3.9), choose $M > 2p$ such that

$$\mathcal{S} \cup -\mathcal{S} \subset [-M, M]^{d+v}. \quad (4.17)$$

Lemma 4.2. *Fix $0 < \sigma < \beta < 1$. Then there exists $M_0 = M_0(d + v, \sigma, \beta)$ such that for all $M \geq M_0$, there exist ϵ_0, δ_0 such that*

$$\begin{aligned} \| [T_M^\theta]^{-1} \| &< e^{M^\beta} \quad (0 < \beta < 1), \\ | [T_M^\theta]^{-1}(\ell, \ell') | &< e^{-\alpha|\ell-\ell'|} \quad (\alpha \text{ as in (H3)}), \\ |\ell - \ell'| &> M/10, \quad M \geq M_0, \end{aligned}$$

for $0 < \epsilon \leq \epsilon_0$, $0 < \delta \leq \delta_0$, $\theta \in \mathbb{R} \setminus \mathcal{B}_{\beta, \alpha}(M)$, $\text{mes } \mathcal{B}_{\beta, \alpha}(M) \leq e^{-M^\sigma}$, all $x \in \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^v$, and all $\omega \in [0, 1]^v$. (Recall from (4.3–4.5) the x and ω dependence of $T_N^\theta(y'_i)$.)

Proof. We use Neumann series in ϵ, δ to estimate T_M^{-1} . We require

$$\begin{cases} |\theta + n \cdot \omega + v_j| > 2 \max(e^{-M^\beta}, (\epsilon + \delta)^{1/2}), \\ |-\theta - n \cdot \omega + v_j| > 2 \max(e^{-M^\beta}, (\epsilon + \delta)^{1/2}), \end{cases} \quad \forall (j, n) \in [-M, M]^{d+v} \setminus \{\mathcal{S} \cup -\mathcal{S}\}. \quad (4.18)$$

Clearly, (4.18) holds away from a set of $\theta \in \mathbb{R}$ of measure at most

$$4(2M + 1)^{d+v} \max(e^{-M^\beta}, (\epsilon + \delta)^{1/2}). \quad (4.19)$$

Choose ϵ, δ such that $(\epsilon + \delta)^{1/2} \leq e^{-M^\beta}$, which can be satisfied if $0 < \epsilon \leq \epsilon_0$, $0 < \delta \leq \delta_0$ with $\epsilon_0 = \delta_0 = \frac{1}{2}e^{-2M^\beta}$. From (4.19), we need $4(2M + 1)^{d+v} e^{-M^\beta} \leq e^{-M^\sigma}$ ($0 < \sigma < \beta < 1$). This leads to $M \geq M_0(d + v, \sigma, \beta)$.

On the complement of the set defined in (4.18), using Neumann series in ϵ, δ for T_M^{-1} and (H3), we verify that

$$\begin{aligned} \| [T_M^\theta]^{-1} \| &< e^{M^\beta} \quad (0 < \beta < 1), \\ | [T_M^\theta]^{-1}(\ell, \ell') | &< e^{-\alpha|\ell-\ell'|} \quad (\alpha \text{ as in (H3)}), \\ |\ell - \ell'| &> M/10, \quad M \geq M_0, \end{aligned}$$

for $0 < \epsilon \leq \epsilon_0$, $0 < \delta \leq \delta_0$. The probability set at this scale, X_M , and the frequency set at this scale, Ω_M , on which and on the complement of which (4.14) holds, satisfy

$$\text{mes } X_M = 1, \quad \text{mes } \Omega_M = 0. \quad (4.20)$$

□

This direct perturbation argument is the same as in [BW]. Note that (4.20) entails that (4.14) holds for all ω, v_j , as in the 0th step, the invertibility is entirely provided by shifting in θ . There is *no* bad site. We will only use the above lemma for the initial set of scales.

The iteration

We now prove Lemma 4.1 using iteration from scale N_0 to $N_0^C = N$ (C assumed large). We call an elementary region $\Lambda(N_0)$ *bad* if the first two inequalities in (4.9) are violated. As in [BW], we need to perform an incision in the frequency space, in order that inside any $\Lambda(N) \in \mathcal{ER}(N)$, there are at most N^{1-} pairwise disjoint bad elementary regions at scale N_0 , where N^{1-} means N^a ($0 < a < 1$). For technical reasons (cf. [BGS, BW]), this requirement pertains to all elementary regions $\Lambda(N')$, $N_0 \leq N' \leq 2N_0$, and not simply at $N' = N_0$. For later constructions in Section 5, it is important to note once again that this set is independent of $\mathcal{V} = \{v_{j_k}\}_{k=1}^v$.

Let $\Lambda(N) \in \mathcal{ER}(N)$, and let $\bar{\Lambda}(N)$ be its projection onto \mathbb{Z}^d ; define

$$\mathcal{T}(N) = \{[-N_0, N_0]^d \times [-N, N]^v\} \cap \Lambda(N), \quad (4.21)$$

and let $\bar{\mathcal{T}}(N)$ be its projection onto \mathbb{Z}^d . Denote by $\overline{\mathcal{ER}}(N)$ the projection of $\mathcal{ER}(N)$ onto \mathbb{Z}^d . Note that $\bar{\Lambda}(N) \in \overline{\mathcal{ER}}(N)$ can be of much smaller diameter than $2N$.

Number of bad elementary regions at scale N_0 disjoint from $\mathcal{T}(N)$

By using (H3) in (4.2–4.4), the region $\Lambda(N) \setminus \mathcal{T}(N)$ can be treated perturbatively. We make separate incisions in the probability space and the frequency space. We first make incisions in the probability space. Toward that end, we look at

$$\Lambda(N') \in \mathcal{ER}(N'), \quad \Lambda(N') \subset \Lambda(N), \quad \Lambda(N') \cap \mathcal{T}(N) = \emptyset \quad (N_0 \leq N' \leq 2N_0). \quad (4.22)$$

Let

$$X'_N \subset \mathbb{R}^{\{\mathbb{Z} \setminus [-N_0, N_0]\}^d} \quad (4.23)$$

be the probability set such that there is at most **one** (pairwise disjoint) $\bar{\Lambda}(N')$ ($N_0 \leq N' \leq 2N_0$) satisfying

$$\begin{cases} \bar{\Lambda}(N') \in \overline{\mathcal{ER}}(N'), \\ \bar{\Lambda}(N') \subset [-N, N]^d, \\ \bar{\Lambda}(N') \cap [-N_0, N_0]^d = \emptyset, \end{cases} \quad (4.24)$$

where (A1) is violated for some $E \in I$, $I = \sigma(H)$, the set defined in (1.6).

Theorem A with $m = \gamma$ yields

$$\begin{aligned} \text{mes } X'_N &> \left(1 - \frac{1}{N_0^{2p'}}\right)^{\mathcal{O}(1)(N^d \times N_0^{d+1})^2} \\ &\geq 1 - \frac{\mathcal{O}(1)}{N^{\frac{2}{\bar{c}}(p'-d(C+1)-1)}} \quad (p' > d(C+1) + 1), \end{aligned} \quad (4.25)$$

where $\mathcal{O}(1)$ is a universal geometric constant; p' is to be determined at the conclusion of the proof of Lemma 4.1. We used $N = N_0^C$; the second factor N_0^{d+1} comes from the estimate on the number of elementary regions of sizes 1 to $2N_0$ associated to each lattice site: $\mathcal{O}(1) \sum_{N'_0=1}^{2N_0} N_0'^d$ (cf. (4.11)); the exponent is an upper bound on the number of pairs of elementary regions of sizes up to $2N_0$ in $\Lambda(N)$.

Given two elementary regions $\Lambda_1(N')$, $\Lambda_2(N')$, we say that they are *convex-disjoint* if their convex envelopes are disjoint. (This is in order that we have (4.23–4.25) at our disposal.) To control the number of bad elementary regions at scale N_0 , we now make additional incisions in the frequency space. Recall that (4.23, 4.24) pertain only to the projected elementary regions in \mathbb{Z}^d .

We are now ready to prove

Lemma 4.3. *Fix $x \in X'_N \cap \tilde{X}_{N_0}$, where X'_N is the set defined in (4.23–4.24) and \tilde{X}_{N_0} is defined as in (4.23–4.24), but with $N_0^{1/C}$ replacing N_0 . There exists a set Ω'_N with*

$$\text{mes } \Omega'_N \leq e^{-N^{\beta/2C}} \quad (4.26)$$

such that if $\omega \notin \Omega'_N$, then for any $\Lambda(N) \in \mathcal{ER}(N)$ any fixed θ , there are at most **two** convex-disjoint bad $\Lambda(N') \in \mathcal{ER}(N')$, $\Lambda(N') \cap \mathcal{T}(N) = \emptyset$, $N_0 \leq N' \leq 2N_0$ in $\Lambda(N)$ ($N = N_0^C$). Moreover Ω'_N is semi-algebraic with degree bounded above by $\mathcal{O}(1)N^{6d+v}$ and it is contained in the union of at most $\mathcal{O}(1)N^{6d+v}$ connected components.

Remark. Ω'_N is independent of \mathcal{V}, y_k . Observe also that we need localization information on the random Schrödinger operators at two scales, N_0 and $N_0^{1/C}$.

Proof. In view of (4.3, 4.4, 4.7, H3), for $\Lambda(N_0)$ such that $\Lambda(N_0) \cap \mathcal{T}(N) = \emptyset$, δS can be treated as a small perturbation. We only need to ensure the invertibility of $D_{\Lambda(N_0)}^\theta$. Assume $\Lambda(N_0)$, $\Lambda'(N_0)$, and $\Lambda''(N_0)$ are three convex-disjoint *bad* elementary regions. So there exist $(n, j) \in \Lambda(N_0)$, $(n', j') \in \Lambda'(N_0)$, and $(n'', j'') \in \Lambda''(N_0)$ such that

$$\begin{aligned} |\theta + n \cdot \omega + \mu_j| < 2e^{-N_0^\beta} \quad \text{or} \quad |-\theta - n \cdot \omega + \mu_j| < 2e^{-N_0^\beta}, \\ |\theta + n' \cdot \omega + \mu_{j'}| < 2e^{-N_0^\beta} \quad \text{or} \quad |-\theta - n' \cdot \omega + \mu_{j'}| < 2e^{-N_0^\beta}, \end{aligned} \quad (4.27)$$

and

$$|\theta + n'' \cdot \omega + \mu_{j''}| < 2e^{-N_0^\beta} \quad \text{or} \quad |-\theta - n'' \cdot \omega + \mu_{j''}| < 2e^{-N_0^\beta},$$

where $\mu_j, \mu_{j'}, \mu_{j''}$ are eigenvalues of $\bar{\Lambda}(N_0), \bar{\Lambda}'(N_0), \bar{\Lambda}''(N_0)$ respectively.

(4.27) implies that there exist m, λ such that

$$|m \cdot \omega + \lambda| < 4e^{-N_0^\beta}, \quad (4.28)$$

where $m = \pm(n - n')$ or $\pm(n' - n'')$ or $\pm(n - n'')$,

$$\lambda = \mu_j - \mu_{j'} \text{ or } \mu_{j'} - \mu_{j''} \text{ or } \mu_{j''} - \mu_j. \quad (4.29)$$

We use the same argument as in the proof of Lemma 2.3 of [BW], which we summarize below.

There are two possibilities: $m = 0$, $m \neq 0$. When $m = 0$, from pairwise disjointness (4.24, 4.25), (A6) implies

$$|\lambda| > e^{-N_0^{\beta'}} \quad (0 < \beta' < \beta), \quad (4.30)$$

which contradicts (4.28).

When $m \neq 0$, (4.28) corresponds to at most

$$\mathcal{O}(1)N^\nu \cdot (N^d \cdot N_0^{2d+1})^2 < \mathcal{O}(1)N^{6d+\nu} \quad (4.31)$$

equations. Since each equation in (4.28) involves a monomial of degree 1 in ω , the excluded set Ω'_N is of degree less than $\mathcal{O}(1)N^{6d+\nu}$. Since $|\omega|$ may be assumed to be bounded for each such equation, we exclude a set of ω of measure $\mathcal{O}(1)e^{-N_0^\beta}$. It is simple to see that for each equation in (4.28), the excluded set of ω has a single component.

So in conclusion, for fixed $x \in X'_N$, $\omega \in \Omega'_N$, there are at most two convex-disjoint $\Lambda(N') \in \mathcal{ER}(N')$ with $\Lambda(N') \cap \mathcal{T}(N) = \emptyset$ such that the first inequality in (4.9) is violated by using (H3) and a simple perturbative argument.

Assume $\Lambda(N')$ is such that the first inequality of (4.9) is satisfied. So $|\pm\theta \pm n \cdot \omega + \mu_j| \geq 2e^{-N_0^\beta}$ for $n, j \in \Lambda(N')$ from the above considerations. To obtain the second inequality we proceed as follows. Let $\bar{\Lambda}(N')$ be the projection of $\Lambda(N')$ onto \mathbb{Z}^d . In view of the restriction in the third expression in (4.9), we may assume $\text{diam } \bar{\Lambda}(N') \geq N_0/10$. We cover $\bar{\Lambda}(N')$ with elementary regions $\bar{\Lambda}(N_0^{1/C})$ of diameter $2N_0^{1/C}$.

Since $x \in X'_N \cap \tilde{X}_{N_0}$ and on \tilde{X}_{N_0} for all E there is at most one (pairwise disjoint) $\bar{\Lambda}(N_0^{1/C})$ in $\bar{\Lambda}(N')$ where (A1) (with $m = \gamma$) is violated, using the resolvent equation, (A1) and the estimates $|\pm\theta \pm n \cdot \omega + \mu_j| \geq 2e^{-N_0^\beta}$ for $n, j \in \Lambda(N')$ for the bad $\bar{\Lambda}(N_0^{1/C})$, we obtain exponential decay in the j direction for $\bar{\Lambda}(N')$ for all E (of the form $E = \pm\theta \pm n \cdot \omega$).

We obtain the second inequality in (4.9) for this $\Lambda(N')$ by another application of resolvent series in δS and using (H3). This holds for all $\Lambda(N')$ such that the first inequality of (4.9) is satisfied, in view of the definition of \tilde{X}_{N_0} . Using $N_0 = N^{1/C}$, we obtain the lemma. \square

Number of bad elementary regions at scale N_0 intersecting $\mathcal{T}(N)$

We now estimate the number of bad $\Lambda(N') \in \mathcal{ER}(N')$ such that $\Lambda(N') \subset \Lambda(N)$, $\Lambda(N') \cap \mathcal{T}(N) \neq \emptyset$, $\mathcal{T}(N)$ as in (4.21), $N_0 \leq N' \leq 2N_0$, using semi-algebraic set techniques. Here it is important to emphasize the \mathbb{Z}^d coordinate of the center of elementary regions, as the linearized operator is not a Toeplitz operator in the \mathbb{Z}^d variable. We look at elementary regions with centers in $\{0\} \times \mathbb{Z}^d$. We write $\mathcal{ER}(N', j)$ for the set of elementary regions centered at $j \in \mathbb{Z}^d$. For any $\Lambda(N', j) \in \mathcal{ER}(N', j)$, let $\mathcal{B}(N', j) := \mathcal{B}_{\beta, \gamma}(N', j)$ be a set such that on $\mathcal{B}(N', j)^c$, (4.9) hold. (Later for more general elementary regions centered at $i \in \mathbb{Z}^{d+v}$, we will use the same notations.)

Assume that there are $X_{N', j}, \Omega_{N', j}$ such that for $x \in X_{N', j}$ and $\omega \in DC_{A, c}(2N) \setminus \Omega_{N', j}$,

$$\text{mes } \mathcal{B}(N', j) \leq e^{-N_0^\sigma} \quad (N_0 \leq N' \leq 2N_0, 0 < \sigma < 1). \quad (4.32)$$

Let

$$\begin{aligned} X''_{N_0} &:= \bigcap_{j \in [-3N_0, 3N_0]^d} \bigcap_{N_0 \leq N' \leq 2N_0} X_{N', j}, \\ \Omega''_{N_0} &:= \bigcup_{j \in [-3N_0, 3N_0]^d} \bigcup_{N_0 \leq N' \leq 2N_0} \Omega_{N', j}, \\ \mathcal{A} &:= \bigcup_{j \in [-3N_0, 3N_0]^d} \bigcup_{N_0 \leq N' \leq 2N_0} \bigcup_{\mathcal{ER}(N', j)} \mathcal{B}(N', j). \end{aligned} \quad (4.33)$$

We have $\text{mes } \mathcal{A} \leq \mathcal{O}(1)N_0^{2d+v}e^{-N_0^\sigma}$ from (4.32).

Lemma 4.4. *Let $N = N_0^C$. For any fixed $\theta \in \mathbb{R}$, let*

$$I = \{n \in [-N, N]^v \mid n \cdot \omega + \theta \in \mathcal{A}\}. \quad (4.34)$$

Fix $x \in X''_{N_0}$. Then for $\omega \in DC_{A, c}(2N) \setminus \Omega''_{N_0}$,

$$|I| \leq \mathcal{O}_{d, v}(1)N_0^{6(d+v)} = \mathcal{O}_{d, v}(1)N^{6(d+v)/C} := N^{1-b_0}, \quad (4.35)$$

$0 < b_0 < 1$ and assuming $6(d+v) < C < N_0^{\sigma/2}$. Hence there are at most $\mathcal{O}(1)N^{1-b_0}$ ($\mathcal{O}(1)$ a universal geometric constant) pairwise disjoint bad $\Lambda(N') \in \mathcal{ER}(N')$ with $\Lambda(N') \cap \mathcal{T}(N) \neq \emptyset$, $N_0 \leq N' \leq 2N_0$ in $\Lambda(N)$ ($N = N_0^C$).

Proof. Since the Green's function is the ratio of two determinants and the norm of the Green's function can be replaced by its Hilbert–Schmidt norm, (4.9) can be reexpressed as polynomial inequalities in θ . Therefore \mathcal{A} is a semi-algebraic set. (See [Ba, Section 7 of BGS].)

\mathcal{A} is defined by

$$\mathcal{O}(1)N_0^d \cdot N_0^{d+v} \cdot N_0^{2(d+v)} \cong \mathcal{O}(1)N_0^{4d+3v} \quad (4.36)$$

polynomials: N_0^d for the number of centers, N_0^{d+v} the number of elementary regions per center, and $N_0^{2(d+v)}$ the number of matrix elements. Each polynomial is of degree $N_0^{2(d+v)}$ in θ (as one squares the matrix elements).

Basu's theorem [Ba], restated as Theorem 7.3 in [BGS], then shows that the number of connected components in \mathcal{A} does not exceed $N_0^{6(d+v)}$. If there are $n, n', n \neq n'$, such that n, n' belong to the same connected component of \mathcal{A} , then from the last inequality in (4.33),

$$|(n - n') \cdot \omega| \leq \mathcal{O}(1)N_0^{2d+v}e^{-N_0^\sigma}. \quad (4.37)$$

Since $n, n' \in [-N, N]^v$, $n - n' \in [-2N, 2N]^v$, $N = N_0^C$, the membership $\omega \in DC_{A,c}(2N)$ is in contradiction with (4.37) for $C < N_0^{\sigma/2} < N_0^\sigma/2A \log N_0$ (assuming $N_0 \gg 1$), so there can be at most one integral point in a connected component of \mathcal{A} . We therefore obtain (4.35).

Let $i = (j, n) \in \mathbb{Z}^{d+v}$. Since $\Lambda(N', i) = \Lambda(N', j) + n$ and T is a Toeplitz operator in the \mathbb{Z}^v variable, we obtain the second conclusion of the lemma. \square

Remark. C will be a *fixed* expansion factor. So the upper bound on C will be satisfied for all $N \geq N_0$ as soon as it is satisfied for an initial N_0 .

A large deviation estimate in θ

Lemmas 4.3 and 4.4 combined imply that the number of bad elementary regions at scale N_0 in $\Lambda(N)$ is at most N^{1-b_0} , where

$$N = N_0^C, \quad b_0 = 1 - 6(d+v)/C \quad (6(d+v) < C < N_0^{\sigma/2})$$

from (4.35). This enables us to use a Cartan-type theorem for analytic matrix-valued functions (see [B5]) to prove a large deviation estimate on $\|[T_N^\theta(y_{i'})]^{-1}\|$, necessary for the proof of Lemma 4.1. The proof of the lemma is very similar to the one in [BW] (see also [BGS]), after using (4.15) to appropriately adjust $y_{i'}$ according to the scale N . So we state (without details of the proof)

Lemma 4.5. *Let $b_0, \beta, \sigma, \gamma$ be fixed positive numbers so that*

$$0 < b_0, \beta, \sigma < 1 \quad \text{and} \quad \beta + b_0 > 1 + 3\sigma. \quad (4.38)$$

Let $N_0 \leq N_1$ be positive integers satisfying

$$\bar{N}_0(\beta, \sigma, \gamma) \leq 100N_0 \leq N_1^\sigma \quad (4.39)$$

with some large constant \bar{N}_0 depending only on β, σ, γ . Assume that for any $N_0 \leq L \leq N_1$, and any $\Lambda(L) \in \mathcal{ER}(L, i)$, $i \in \mathbb{Z}^{d+v}$,

$$\text{mes } \mathcal{B}_{\beta, \gamma}(L, i) \leq e^{-L^\sigma}. \quad (4.40)$$

Let $\bar{X}_N, \bar{\Omega}_N$ be the sets such that on \bar{X}_N and $DC_{A,c}(2N) \setminus \bar{\Omega}_N$, (4.40) holds for all $i \in [-N, N]^{d+v}$, $L \in [N_0, N_1]$. Let $X'_N, \tilde{X}_{N^{1/C}}$ be the sets defined in (4.23, 4.24) and Lemma 4.3, and Ω'_N the corresponding frequency set as in Lemma 4.3. If

$$x \in X'_N \cap \tilde{X}_{N^{1/C}} \cap \bar{X}_N, \quad \omega \in DC_{A,c}(2N) \setminus \{\Omega'_N \cup \bar{\Omega}_N\}, \quad (4.41)$$

then

$$\text{mes}\{\theta \mid \| [T_N^\theta]^{-1} \| > e^{N^\beta} \} < e^{-N^{3\sigma}}, \quad (4.42)$$

where T_N is the restriction of T to any $\Lambda \in \mathcal{ER}(N)$, the elementary regions centered at 0, provided $N_0^{C_1} \leq N \leq N_1^{\sigma C_1}$, with $C_1 \gg \max(1/\sigma, 6(d+v))$ depending only on β, σ .

Remark. We note that from (4.23, 4.24), the dependence of the probability set $X_{L,i}$ (on which (4.40) holds) on $i \in \mathbb{Z}^{d+v}$ is only through the \mathbb{Z}^d coordinate. For simplicity, we keep the notation $X_{L,i}$. The set $\Omega_{L,i}$ (on the complement of which (4.40) holds), on the other hand, does have full dependence on i .

A summary of the proof

We use analytic and harmonic function theory together with a 2-scale (in the range $[N_0, N_1]$) analysis to control the measure of the set in (4.42) at scale $N \gg N_0$. (See the proofs of Lemma 4.4 in [BGS] and Lemma 6.2 in [BW].) For a given N , let these two scales $L_1, L_2 \in [N_0, N_1]$, $L_1 < L_2$, satisfy

$$\log L_1 \sim \frac{1}{C_1} \log N, \quad \log L_2 \sim \frac{1}{\sigma} \log L_1 \sim \frac{1}{C_1 \sigma} \log N,$$

with C_1 as in the lemma, $C_1 \gg 1/\sigma$. We reiterate the main line of arguments below.

- Fix θ . Let

$$\Lambda_*^c = \{m \in \Lambda(N) \mid \exists \Lambda_1 \in \mathcal{ER}(L_1), N_0 \leq L_1 \leq 2N_0, \\ \Lambda_1 \subset m + [-L_1, L_1]^{d+v}, \Lambda_1 \text{ is bad}\}. \quad (4.43)$$

For x, ω satisfying (4.41),

$$|\Lambda_*^c| \leq N^{1-b_0} \quad (b_0 > 0), \quad (4.44)$$

by Lemmas 4.3 and 4.4. Since $X''_{N_0} \supset \bar{X}_N$, we have $\Omega''_{N_0} \subset \bar{\Omega}_N$, where $X''_{N_0}, \Omega''_{N_0}$ are as defined by the first two equations of (4.33).

- Let Λ_* be, roughly speaking, the complement of the set in (4.43). For a more precise definition, which requires a partition of Λ , see the beginning of the proof of Lemma 4.4 in [BGS]. Using an elementary resolvent expansion (Lemma B in the appendix, which is Lemma 2.2 of [BGS] reiterated), we obtain an upper bound on $\| [T_{\Lambda_*}^\theta]^{-1} \|$ at fixed θ by using the decay estimate on the Λ_1 's, elementary regions at scale L_1 , in Λ_* . By definition they are all *good*. By standard Neumann series arguments, this bound is preserved inside the disk $B(\theta, e^{-N_0}) \subset \mathbb{C}$.

Remark. We have control over the size of Λ_* via (4.43, 4.44), but not its geometry. Typically Λ_* is *nonconvex*, hence the need for elementary regions which are more general than cubes, in particular L-shaped regions, in view of Lemma B.

- Define a matrix-valued analytic function $A(\theta')$ on $B(\theta, e^{-N_0})$ as

$$A(\theta') = R_{\Lambda_*^c} T_N^{\theta'} R_{\Lambda_*^c} - R_{\Lambda_*^c} T_N^{\theta'} R_{\Lambda_*} [T_{\Lambda_*}^{\theta'}]^{-1} R_{\Lambda_*} T_N^{\theta'} R_{\Lambda_*^c}, \quad (4.45)$$

where $\Lambda_*^c = \Lambda \setminus \Lambda_*$ and $R_{\Lambda_*}, R_{\Lambda_*^c}$ are projections. From (4.44), $A(\theta')$ is an $\mathcal{O}(N^{1-b_0}) \times \mathcal{O}(N^{1-b_0})$ matrix. The raison d'être of introducing $A(\theta')$ is the following inequality:

$$\|[A(\theta')]^{-1}\| \lesssim \|[T_N^{\theta'}]^{-1}\| \lesssim e^{2N_0} \|[A(\theta')]\|$$

(see Lemma 4.8 of [BGS]). So to bound $\|[T_N^{\theta'}]^{-1}\|$, it is sufficient to bound $\|[A(\theta')]^{-1}\|$, which is of smaller dimension.

- Toward that end, we introduce an intermediate scale L_2 with $\log L_2 \sim (\log L_1)/\sigma > \log L_1$. We work in an interval $\Theta = \{\theta' \mid |\theta' - \theta| < e^{-N_0}\}$. Using (4.40) for the Λ_2 's at scale L_2 and in Λ (Lemma B), we obtain an upper bound on $\|[T_N^{\theta'}]^{-1}\|$ except for a set of θ' of measure smaller than $e^{-\mathcal{O}(L_2^2)}$. So there exists $y \in \Theta$ such that we have *both* an upper bound on $\|[A(\theta')]\|$ at $\theta' = y$, hence a lower bound on the smallest eigenvalue of $A(\theta)$, and an a priori upper bound on $\|A(\theta')\|$, which comes from the boundedness of $T_N^{\theta'}$ and the bound on $[T_{\Lambda_*}^{\theta'}]^{-1}$ (see (4.45)).
- Transferring the estimates on $\|A(\theta')\|, \|[A(\theta')]\|$ into estimates on $\log |\det A(\theta')|$, which is subharmonic, and using either a Cartan-type theorem (see Sect. 11.2 in [Le]) or proceeding as in the proof of Lemma 4.4 of [BGS] or Chap. XIV of [B5], we obtain the lemma by covering the interval $I = (-\mathcal{O}(N_1^{\sigma C_1}), \mathcal{O}(N_1^{\sigma C_1}))$ with intervals of size e^{-N_0} . (Recall (4.39) and that for all $\theta \notin I$, $T_N^{\theta'}$ is automatically invertible.) \square

Iteration lemma

To obtain exponential decay of T_N^{-1} , we need

Lemma 4.6. *Suppose $M, N \in \mathbb{N}^+$ are such that for some $0 < \tau < 1$,*

$$N^\tau \leq M \leq 2N^\tau. \quad (4.46)$$

Let $\Lambda_0 \in \mathcal{ER}(N)$ be an elementary region with the property that for all $\Lambda \subset \Lambda_0$ such that $\Lambda \in \mathcal{ER}(L)$ with $M \leq L \leq N$,

$$\|[T_L^\theta]^{-1}\| \leq e^{L^\beta} \quad (0 < \beta < 1). \quad (4.47)$$

We say that $\Lambda \in \mathcal{ER}(L)$ with $\Lambda \subset \Lambda_0$ is good if in addition to (4.47),

$$\|[T_L^\theta]^{-1}(k, k')\| \leq e^{-\gamma|k-k'|} \quad (4.48)$$

for all $k, k' \in \Lambda$ with $|k - k'| > L/10$. Otherwise Λ is called bad. Assume that for any family \mathcal{F} of pairwise disjoint bad M' -regions in Λ_0 with $M + 1 \leq M' \leq 2M + 1$,

$$\#\mathcal{F} \leq N^\beta \quad (0 < \beta < 1). \quad (4.49)$$

Then

$$|[T_N^\theta]^{-1}(k, k')| \leq e^{-\gamma'|k-k'|} \quad (4.50)$$

for all $k, k' \in \Lambda_0$ with $|k - k'| > N/10$, and $\gamma' = \gamma - N^{-\delta'}$ ($\delta' > 0$), provided $N \gg N_0(\beta, \tau, \gamma)$.

The proof of the above lemma is written out in detail in [BGS]. The only difference is that instead of being tridiagonal, T_L^θ has exponentially decaying off-diagonal elements by (H3). So $\gamma = \gamma(\alpha)$. The proof goes through. So we do not repeat it here. (See also [BW].) The gist is as follows.

The exponential decay estimate at scale N in (4.50) is obtained from the exponential decay estimate in (4.48) at smaller scales M' by using (4.15), the norm estimate in (4.47) and the resolvent identity. To implement this, we use a sequence of scales $M_{j+1} = M_j^{C'}$, with $M_0 = M$ and C' such that $C'\beta < 1$ and $C'\tau \leq 1$. For each elementary region $\Lambda(M_{j+1})$ at scale M_{j+1} and for each $k \in \Lambda(M_{j+1})$, we exhaust $\Lambda(M_{j+1})$ by an increasing sequence of annuli centered at k of width $2M_j$, or more precisely the intersections of this sequence with $\Lambda(M_{j+1})$. Roughly speaking, an annulus is good if it does not intersect a bad cube of the previous scale M_j .

In each of the connected components of the complement of the bad annuli, we apply the resolvent identity using the estimate in (4.48) for elementary regions of size M_j . In the bad annuli, we use (4.47). From (4.49), the number of bad annuli is at most sublinear in M_{j+1} . Using a multiscale induction argument to reach the scale N , we obtain the exponential decay in (4.50) when $|k - k'| > N/10$; δ' is determined from (4.47, 4.49), $\delta' \simeq \tau(1 - \beta C')$. \square

Let $\mathcal{B}_{\beta, \gamma}(N, i)$, $i \in \mathbb{Z}^{d+v}$, be a set such that on the complement, (4.9) hold. When $i = 0$, we write $\mathcal{B}_{\beta, \gamma}(N, 0) = \mathcal{B}_{\beta, \gamma}(N)$. As before let

$$\begin{aligned} X_{N, i} \text{ and } \Omega_{N, i} \text{ be the probability frequency subsets on which and} \\ \text{on the complement of which } \text{mes } \mathcal{B}_{\beta, \gamma}(N, i) \leq e^{-N^\sigma}. \end{aligned} \quad (4.51)$$

Combining Lemmas 4.5, 4.6 with (4.23–4.25) and Lemmas 4.3 and 4.4, we obtain

Lemma 4.7. *Assume that for any $\tilde{N}_0 \leq N_0 \leq \tilde{N}_0^C$, $\max(1/\sigma, 6(d+v)) \ll C < \tilde{N}_0^{\sigma/2}$ ($\tilde{N}_0(\epsilon, \delta) \gg 1$), and any $\Lambda(N_0) \in \mathcal{ER}(N_0, i)$, $i \in \mathbb{Z}^{d+v}$,*

$$\text{mes } \mathcal{B}_{\beta, \gamma}(N_0, i) \leq e^{-N_0^\sigma} \quad (0 < \sigma < 1). \quad (4.52)$$

Let $[\tilde{N}_0^C, \tilde{N}_0^{C^2\sigma}]$ be the next interval of scales. For any $N \in [\tilde{N}_0^C, \tilde{N}_0^{C^2\sigma}]$, write $N = N_0^C$ with $N_0 \in [\tilde{N}_0, \tilde{N}_0^C]$. Let $X'_N, \tilde{X}_{N^{1/C}}$ be the sets defined in (4.23, 4.24) and Lemma 4.3,

and Ω'_N the set defined in Lemma 4.3, satisfying

$$\begin{aligned} \text{mes } X'_N &\geq 1 - \frac{\mathcal{O}(1)}{N^{\frac{2}{c}(p'-d(C+1)-1)}} \quad (p' > d(C^2 + 1) + 1), \\ \text{mes } \tilde{X}_{N^{1/c}} &\geq 1 - \frac{\mathcal{O}(1)}{N^{\frac{2}{c^2}(p'-d(C^2+1)-1)}} \quad (p' > d(C^2 + 1) + 1), \\ \text{mes } \Omega'_N &\leq e^{-N^{\beta/2C}}, \end{aligned} \quad (4.53)$$

in view of (4.25, 4.26).

Let $\bar{X}_N, \bar{\Omega}_N$ be the sets such that on \bar{X}_N and $DC_{A,c}(2N) \setminus \bar{\Omega}_N$, (4.52) holds for all $i \in [-\tilde{N}_0^{C^2\sigma}, \tilde{N}_0^{C^2\sigma}]^{d+\nu}$ and all $N_0 \in [\tilde{N}_0, \tilde{N}_0^C]$. If

$$\begin{aligned} x &\in X'_N \cap \tilde{X}_{N^{1/c}} \cap \bar{X}_N =: X_N, \\ \text{mes } X_N &\geq 1 - 1/N^{p''} \text{ with } p'' = p''(p', C, d + \nu) > 1, \\ &p' \text{ large enough,} \\ \omega &\in DC_{A,c}(2N) \setminus \{\Omega'_N \cup \bar{\Omega}_N\} =: DC_{A,c}(2N) \setminus \Omega_N, \\ \text{mes } \Omega_N &\leq e^{-N^{\beta/3C^2}}, \\ \Omega_N &\text{ is semi-algebraic of degree less than } N^{C^2(d+\nu)}, \end{aligned} \quad (4.54)$$

then

$$\text{mes } \mathcal{B}_{\beta, \gamma'}(N) \leq e^{-N^\sigma} \quad (0 < \sigma < 1), \quad (4.55)$$

where $\gamma' = \gamma - N^{-\delta'}$ ($\delta' > 0$).

Proof. Applying Lemma 4.5 using (4.52), we obtain the large deviation estimate on $\| [T_N^\theta]^{-1} \|$. Choosing $0 < \tau < 1/C\sigma$, for x, ω in the sets defined in (4.54), (4.49) is satisfied. (Here we need the definition that an N^τ -region is bad if it intersects a bad N_0 -region. Otherwise it is good. On the good N^τ -region, (4.48) is obtained by using the resolvent expansion (Lemma B) and (4.9) for N_0 -regions.) Hence Lemma 4.6 is available and we obtain (4.55). The estimates on X_N, Ω_N follow from (4.51, 4.53) and the constructions in Lemmas 4.3–4.5. \square

Proof of Lemma 4.1. Assume \bar{N}_0 (to be determined below) is such that Lemma 4.7 is available. For the scales $\bar{N}_0 \leq N \leq \bar{N}_0^C$, we use Neumann series in ϵ and δ à la Lemma 4.2 and its proof. For the scales $N \geq \bar{N}_0^C$, we use Lemma 4.7.

From Lemma 4.2, we need

$$\bar{N}_0 \geq M_0(d + \nu, \sigma, \beta) \quad (4.56)$$

and

$$\epsilon_0 = \delta_0 = \frac{1}{2} e^{-2\bar{N}_0^{C\beta}}. \quad (4.57)$$

From Lemma 4.7 and the choice of σ , the expansion factor C needs to satisfy

$$\max(1/\sigma, 6(d+v)) \ll C < \bar{N}_0^{\sigma/2}. \quad (4.58)$$

From Theorem A and (4.19), \bar{N}_0 needs further to satisfy

$$\bar{N}_0 > \max(Q, M_0), \quad (4.59)$$

where $M_0 = M_0(d+v, \beta, \sigma)$ as in Lemma 4.2. Fix \bar{N}_0 satisfying (4.59), and C satisfying (4.58). Then (4.57) determines ϵ_0, δ_0 .

For the scales $\bar{N}_0 \leq N \leq \bar{N}_0^C$, with \bar{N}_0 satisfying (4.59), the estimates

$$\begin{aligned} \|[T_N^\theta]^{-1}\| &< e^{N^\beta} \quad (0 < \beta < 1), \\ |[T_N^\theta]^{-1}(\ell, \ell')| &< e^{-\alpha|\ell-\ell'|} \quad (\alpha \text{ as in (H3)}), \\ |\ell - \ell'| &> N/10, \end{aligned}$$

for $0 < \epsilon \leq \epsilon_0, 0 < \delta \leq \delta_0$ are obtained using Neumann series by shifting in θ only. So

$$\text{mes } X_{N,i} = 1, \quad \text{mes } \Omega_{N,i} = 0 \quad (\bar{N}_0 \leq N \leq \bar{N}_0^C),$$

where $X_{N,i}, \Omega_{N,i}$ are as defined in (4.51), and (4.52) holds.

Let

$$X := \bigcap_N \bigcap_{i \in [-3NC, 3NC]^{d+v}} X_{N,i} = \bigcap_N \bigcap_{j \in [-3NC, 3NC]^d} X_{N,j}, \quad (4.60)$$

where the second equality follows from the remark after Lemma 4.5, j being the \mathbb{Z}^d coordinate of i ; and

$$\Omega := \bigcup_N \bigcup_{i \in [-3NC, 3NC]^{d+v}} \Omega_{N,i}, \quad (4.61)$$

where $X_{N,j}, \Omega_{N,i}$ are as defined in (4.51).

On X and $DC_{A,c} \setminus \Omega$, Lemma 4.7 is available with the initial $\gamma = \alpha$ from Lemma 4.2 for iteration to all scales. Estimating the measure of X and Ω using (4.53, 4.54, 4.60, 4.61), using the measure estimates on the bad set in θ from Lemmas 4.2 and 4.7, we obtain the assertion of Lemma 4.1 by taking $p' \simeq \mathcal{O}_d(1)C^3$. \square

5. Invertibility of $T(y_i)$, Q -equations and determination of ω

Fix $x \in X \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^v$, defined in (4.12), which generates a corresponding set Ω as in Lemma 4.1. The main work of this section is to convert the measure estimates in θ for fixed $\omega \in \mathbb{R}^v \setminus \Omega$ and fixed $\mathcal{V} \in \mathbb{R}^v$ in Lemma 4.1 into measure estimates in $\omega = \omega(\mathcal{V}) \in \mathbb{R}^v$ and extend them to $(\omega, \mathcal{V}) \in \mathbb{R}^{2v}$ in the neighborhood of $\omega = \omega(\mathcal{V})$, while keeping θ fixed: $\theta = 0$ and addressing the original family of linearized operators $T_N(y_i)$ defined in (4.2–4.5), where y_i is the i^{th} approximant for the P -equations (4.1). This is possible because T_N^θ is only a function of $n \cdot \omega + \theta$, $T_N^\theta = T'(n \cdot \omega + \theta)$ (cf. (4.2–4.4, 4.7)).

Since θ is now fixed at 0, before making the conversion, we need to make a further restriction in X in order that the spectrum of the various restricted random Schrödinger operators stay away from 0. This is needed when $n = 0$ and we cannot vary ω to have invertibility of the linearized operators (cf. (4.3)).

So we modify the definition of X'_N in (4.23, 4.24) to include the condition

$$\text{dist}(\sigma(\bar{\Lambda}(N')), 0) > e^{-N_0^{\beta'}} \quad (0 < \beta' < \beta) \quad (5.0)$$

for all $\bar{\Lambda}(N') \cap [-N, N]^d \neq \emptyset$ ($N_0 \leq N' \leq 2N_0$), i.e., we require (4.30) to hold also when $\lambda = \mu_j$, eigenvalues of restricted random Schrödinger operators and not just differences of pairs of eigenvalues. In view of the Wegner estimate (A7), this leaves the measure of the set X'_N in (4.25) essentially unchanged.

This generates the restricted probability set $\tilde{X} \subset X \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^v$, on which Lemma 4.1 holds. Rename \tilde{X} as X and let Ω be its corresponding frequency set.

To make the conversion, we need to supplement the measure estimates in (4.14) by the fact that the bad set $\mathcal{B}(N)$ defined from (4.9) has a semi-algebraic description in terms of $(\omega, \mathcal{V}, \theta)$, enabling us to use the decomposition Lemma 9.9 of [B5]. Once we have the necessary estimates on $[T_N(y_i)]^{-1}$ after removing a small set of $\omega_i = \omega_i(\mathcal{V})$ (we put the suffix i here to stress that it is the i^{th} approximation), we construct the next approximant y_{i+1} according to (3.9), which in turn is used to construct $\omega_{i+1} = \omega_{i+1}(\mathcal{V})$.

In this section, we primarily address the invertibility of $T_N(y_i)$, $N = M^{i+1}$. Since the estimate on $T_N(y_i)$ and the construction of y_{i+1} are interconnected, it is good at this point to lay down the complete induction hypothesis. (In Section 4, we used the first part of the induction hypothesis (H1, 3) and the first inequality of (H2) to derive Lemma 4.1.) The first few approximations are obtained by using direct perturbation series in ϵ, δ (see (4.9), together with the text and the remark afterwards). So $0 < \epsilon \leq \epsilon_0 \ll 1$, $0 < \delta \leq \delta_0 \ll 1$, and we have $\alpha = \mathcal{O}(1)|\log(\epsilon + \delta)|$.

On the *entire* (ω, \mathcal{V}) parameter space, we assume:

$$(H1) \quad \text{supp } y_i \subset [-M^i, M^i]^{d+v} \quad (i \geq 1),$$

$$(H2) \quad \|\Delta_i y\| < \delta_i, \quad \|\partial \Delta_i y\| < \bar{\delta}_i,$$

where $\|\cdot\|$ stands for $\sup_{\omega, \mathcal{V}} \|\cdot\|_{\ell^2(\mathbb{Z}^{d+v})}$ (recall that we identify y with \hat{y} ; ∂ refers to derivation in ω or \mathcal{V}). $\delta_i, \bar{\delta}_i$ will be shown to satisfy

$$\delta_i < \sqrt{\epsilon + \delta} M^{-(4/3)^i}, \quad \bar{\delta}_i < \sqrt{\epsilon + \delta} M^{-(4/3)^i/2}.$$

$$(H3) \quad |y_i(k)| < e^{-\alpha|k|} \text{ for some } \alpha > 0$$

(the constant in front of the exponential in (H3) is 1, because $|a_\ell| \ll 1$, $\ell = 1, \dots, v$, see (1.9)).

From (H2), y is a C^1 function of (ω, \mathcal{V}) . Application of the implicit function theorem to the Q -equations in (2.12) with

$$y = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = y_i \quad (5.1)$$

yields

$$\omega_i = \mathcal{V} + (\varepsilon + \delta)\varphi_i(\mathcal{V}) \quad \text{with } \|\partial\varphi_i\| < C, \quad (5.2)$$

whose graph we denote by Γ_i . Recall that a priori, y_i are only defined on certain intervals in (ω, \mathcal{V}) space. It is in order to use the implicit function theorem that we extend y_i to the entire (ω, \mathcal{V}) space, using the estimates on ∂y_i in (H2). (H2) and (2.12) imply, moreover, that

$$\|\varphi_i - \varphi_{i-1}\| \leq \mathcal{O}(1)\|y_i - y_{i-1}\| < \mathcal{O}(1)\delta_i, \quad (5.3)$$

which in turn implies that

$$\Gamma_i \text{ is an } (\varepsilon + \delta)\delta_i\text{-approximation of } \Gamma_{i-1}. \quad (5.4)$$

At each stage i , define

$$M_0 = \mathcal{O}(1)(i+1)^{C/2}(\log M)^{C/2} \quad (5.5)$$

for some $C > 0$, and $\Omega_{M_0, k}$ as in (4.51). Unlike (H1–3), the following hypothesis is only assumed to hold on *certain* intervals in \mathbb{R}^{2v} , the (ω, \mathcal{V}) parameter space.

(H4) There is a collection Σ_i of intervals I in \mathbb{R}^{2v} of size M^{-iC} , with the same C as in (5.5), such that

(0) $I \cap \Gamma_i \subset DC_{A, c}^{(i)} \setminus \Omega_i$, where

$$\begin{aligned} DC_{A, c}^{(i)} &:= DC_{A, c}(2N), \quad N = M^i, \\ \Omega_i &:= \bigcup_{k \in \mathbb{Z}^d \cap [-2M_0, 2M_0]^d} \Omega_{M_0, k}, \quad i > i_0 > 0 \end{aligned}$$

(see the remark after (iv) concerning i_0).

- (i) On $I \in \Sigma_i$, $y_i(\omega, \mathcal{V})$ is a rational function of (ω, \mathcal{V}) of degree at most $M^q i^3$ ($q \in \mathbb{N}^+$).
- (ii) For $(\omega, \mathcal{V}) \in \bigcup_{I \in \Sigma_i} I$,

$$\|F(y_i)\| < \kappa_i, \quad \|\partial F(y_i)\| < \bar{\kappa}_i. \quad (5.6)$$

In (6.20), $\kappa_i, \bar{\kappa}_i$ will be shown to satisfy

$$\kappa_i < \sqrt{\varepsilon + \delta} M^{-(4/3)^{i+2}}, \quad \bar{\kappa}_i < \sqrt{\varepsilon + \delta} M^{-(4/3)^{i+2}/2}. \quad (5.7)$$

- (iii) For $(\omega, \mathcal{V}) \in \bigcup_{I \in \Sigma_i} I$,

$$T_N = T_N(y_{i-1}), \quad N = M^i,$$

satisfies

$$\|T_N^{-1}\| \leq M^{iC}, \quad |T_N^{-1}(k, k')| \leq e^{-\alpha|k-k'|} \quad \text{for } |k - k'| > i^C.$$

(iv) Each $I \in \Sigma_i$ is contained in an interval $I' \in \Sigma_{i-1}$ and

$$\text{mes}_v \left(\Gamma_i \cap \left(\bigcup_{I \in \Sigma_{i-1}} I \setminus \bigcup_{I \in \Sigma_i} I' \right) \right) < M^{-i/5}.$$

Remark. (H4.0) is only needed for $i > i_0$ to ensure the availability of Lemma 4.1. Up to stage i_0 , we use direct ϵ, δ perturbation series, where the Diophantine property of ω is not required (cf. Lemma 4.2).

Unlike the related estimates on $\Delta_i y, \partial \Delta_i y$ in (H2), (5.6, 5.7) *cannot* be extended to the entire (ω, \mathcal{V}) space. This is because, as mentioned earlier, outside the intervals in Σ_i , y_i are no longer close to solutions to the P -equations (2.16).

Invertibility of $T_N(y_i)$, $N = M^{i+1}$

Assume (H1–4) hold at stage i . To construct y_{i+1} , we need to control

$$[T_N(y_i)]^{-1}, \quad N = M^{i+1},$$

with a further restriction on the (ω, \mathcal{V}) parameter set. This will give us (H4.iii) at stage $i + 1$.

We accomplish this by covering $[-M^{i+1}, M^{i+1}]^{d+v}$ with $[-M^i, M^i]^{d+v}$ and intervals $[-M_0, M_0]^{d+v} + k, M_0$ as in (5.5), $k \in \mathbb{Z}^{d+v}, M^i/2 < |k| < M^{i+1}$, and using the resolvent identity. We first estimate $[T_{M^i}(y_i)]^{-1}$. Fix $(\omega, \mathcal{V}) \in \bigcup_{I \in \Sigma_i} I$. (H4.iii) at stage i gives

$$\begin{aligned} \|[T_{M^i}(y_{i-1})]^{-1}\| &\leq M^{iC}, \\ |[T_{M^i}(y_{i-1})]^{-1}(k, k')| &\leq e^{-\alpha|k-k'|} \quad (|k-k'| > i^C). \end{aligned} \quad (5.8)$$

We write

$$T_{M^i}(y_i) = T_{M^i}(y_{i-1}) + [T_{M^i}(y_i) - T_{M^i}(y_{i-1})] =: A + B. \quad (5.9)$$

From the first inequality in (5.8),

$$\|A^{-1}\| \leq M^{iC}. \quad (5.10)$$

The first inequality of (H2) at stage i gives

$$\|B\| < \mathcal{O}(1)M^{-(4/3)^i}. \quad (5.11)$$

So

$$\|[T_{M^i}(y_i)]^{-1}\| \leq 2M^{iC} \quad \text{for } i > C^2. \quad (5.12)$$

To obtain a pointwise estimate on $[T_{M^i}(y_i)]^{-1}$, we use (5.9) and resolvent series. A^{-1} has off-diagonal decay from (5.8), and B has off-diagonal decay from (H3) at stage i . Iterating the resolvent series and using (5.12), we obtain

$$|[T_{M^i}(y_i)]^{-1}(k, k')| \leq e^{-\alpha'|k-k'|} \quad (|k-k'| > i^C) \quad (5.13)$$

with $\alpha' = \alpha - M^{-i\delta'} > \alpha/2$ ($\delta' > 0$), uniformly in i .

We now study $|[T(y_i)]^{-1}|$ on the M_0 intervals, M_0 as in (5.5). We distinguish two types of M_0 intervals J in $[-M^{i+1}, M^{i+1}]^{d+v}$:

- (a) $J \cap [-M_0, M_0]^d \times [-M^{i+1}, M^{i+1}]^v = \emptyset$,
 (b) $J \cap [-M_0, M_0]^d \times [-M^{i+1}, M^{i+1}]^v \neq \emptyset$.

For type (a), we use direct perturbation in view of (H3). For type (b), we use a more delicate construction. We write $M_0 = M^{\tilde{i}_0}$. (M_0 is chosen in order that the total degree of the semi-algebraic set describing the bad set in $\omega, \mathcal{V}, \theta$ is not too large.)

The M_0 intervals are at the $[\tilde{i}_0]^{\text{th}}$ scale. On M_0 intervals of type (b), we use Lemma 4.1, which is for $T_{M_0}(y_{i_0})$, $i_0 \simeq \frac{\log M}{\log b} \tilde{i}_0 < i$ similar to (4.9). Using a decomposition lemma ([B5, Lemma 9.9] restated here as Lemma 5.3) to make appropriate incisions in the (ω, \mathcal{V}) parameter space, applying (H2) between y_{i_0} and y_i and combining with estimates on type (a) intervals, we obtain

$$\begin{aligned} \|(R_J T(y_i) R_J)^{-1}\| &< e^{M_0^\beta} \quad (0 < \beta < 1), \\ |(R_J T(y_i) R_J)^{-1}(k, k')| &< e^{-\alpha''|k-k'|}, \quad k, k' \in J, |k-k'| > M_0/10, \end{aligned} \quad (5.14)$$

where $\alpha'' = \alpha - M^{-i\delta''}$ ($\delta'' > 0$), for all

$$J = [-M_0, M_0]^{d+v} + k, \quad \frac{1}{2}M^i < |k| < M^{i+1}. \quad (5.15)$$

This is the content of Lemma 5.2. We delay its precise statement and proof momentarily. We first prove

Lemma 5.1. *Assume (5.12–5.15) hold and M_0 is as in (5.5). Then*

$$\|[T_{M^{i+1}}(y_i)]^{-1}\| < \mathcal{O}(1)M^{(i+1)C}, \quad (5.16)$$

$$|[T_{M^{i+1}}(y_i)]^{-1}(k, k')| < e^{-\tilde{\alpha}|k-k'|} \quad \text{for } |k-k'| > (i+1)^C, \quad (5.17)$$

with $\tilde{\alpha} = \alpha - M^{-(i+1)\tilde{\delta}}$, $\tilde{\delta} > 0$.

Proof. (5.16, 5.17) are exercises in the resolvent identity or equivalently using Lemma B in the appendix. We first prove (5.16). Then (5.17) follows by using (5.16) and another application of the resolvent identity. Let

$$B_i = [-M^i, M^i]^{d+v}, \quad B_{i+1} = [-M^{i+1}, M^{i+1}]^{d+v}.$$

For any $k, \ell \in B_{i+1}$, we have (assume $T_{B_{i+1}}^{-1}$ is defined)

$$T_{B_{i+1}}^{-1}(k, \ell) = T_{W(k)}^{-1}(k, \ell) + \sum_{\substack{k' \in \partial_* W(k) \\ |k''-k'|=1 \\ k'' \in B_{i+1} \setminus W(k)}} T_{W(k)}^{-1}(k, k') T_{B_{i+1}}^{-1}(k'', \ell); \quad (5.18)$$

here for $|k| \leq \frac{1}{2}M^i$, $W(k) = B_i$, while for $|k| > \frac{1}{2}M^i$, $W(k)$ is a size M_0 interval. It is easy to see that for every k' , there exists $W(k)$ such that $\text{dist}(k', \partial_* W(k)) \geq M_0$, where $\partial_* W(k)$ is the interior boundary of $W(k)$, relative to B_{i+1} . Summing over $\ell \in B_{i+1}$ yields

$$\begin{aligned} \sup_{k \in B_{i+1}} \sum_{\ell \in B_{i+1}} |T_{B_{i+1}}^{-1}(k, \ell)| &\leq \sup_{k \in B_{i+1}} \sum_{\ell \in W(k)} \|T_{W(k)}^{-1}\| \\ &+ \sup_{k \in B_{i+1}} \sum_{k' \in \partial_* W(k)} |T_{W(k)}^{-1}(k, k')| \sup_{k'' \in B_{i+1}} \sum_{\ell \in B_{i+1}} |T_{B_{i+1}}^{-1}(k'', \ell)|. \end{aligned} \quad (5.19)$$

Using (5.12–5.15), we have

$$\begin{aligned} \sup_{k \in B_{i+1}} \sum_{\ell \in B_{i+1}} |T_{B_{i+1}}^{-1}(k, \ell)| &\leq 2M^{iC} \cdot (2M^i + 1)^{d+v} \\ &+ \mathcal{O}(1)e^{-cM_0} M_0^{d+v-1} \sup_{k'' \in B_{i+1}} \sum_{\ell \in B_{i+1}} |T_{B_{i+1}}^{-1}(k'', \ell)|, \end{aligned} \quad (5.20)$$

since $M^i \gg M_0$. So

$$\|T_{B_{i+1}}^{-1}\| \leq M^{(i+1)C},$$

which is (5.16).

To obtain (5.17), we retrace our steps back to (5.18) and restrict to k, ℓ such that $|k - \ell| > (i + 1)^C \gg M_0$, in view of (5.5). Iterating (5.18) along the path from k to ℓ using B_i, J and using (5.16) for the last factor, we obtain (5.17). \square

(5.16, 5.17) will be the conclusion of (H4.iii) at stage $i + 1$ once we specify the new set of intervals Σ_{i+1} , on which they hold. As alluded to earlier, Σ_{i+1} will be determined from Σ_i and the new restriction on (ω, \mathcal{V}) in order that (5.14, 5.15) hold.

Determination of Σ_{i+1}

Lemma 5.2. *Assume (H1–4) at stage i . There exists $\tilde{\Gamma}_i \subset \Gamma_i$ with $\text{mes}_\nu \tilde{\Gamma}_i < M^{-i/4}$ such that (5.14, 5.15) hold on*

$$\bigcup_{I \in \Delta_i} (I \cap (\Gamma_i \setminus \tilde{\Gamma}_i)). \quad (5.21)$$

The proof of (5.21) relies on the measure estimates in Lemma 4.1 and semi-algebraic description of the bad set. We need the following decomposition lemma, which is proven in [B5, Lemma 9.9].

Lemma 5.3. *Let $\mathcal{S} \subset [0, 1]^{2n}$ be a semi-algebraic set of degree B and $\text{mes}_{2n} \mathcal{S} < \eta$, $\log B \ll \log(1/\eta)$. Denote by $(x, y) \in [0, 1]^n \times [0, 1]^n$ the product variable. Fix $\epsilon > \eta^{1/2n}$. Then there is a decomposition $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ with \mathcal{S}_1 satisfying*

$$|\text{Proj}_x \mathcal{S}_1| < B^K \epsilon \quad (K > 0) \quad (5.22)$$

and \mathcal{S}_2 satisfying the transversality property

$$\text{mes}_n(\mathcal{S}_2 \cap L) < B^K \epsilon^{-1} \eta^{1/2n} \quad (K > 0), \quad (5.23)$$

for any n -dimensional hyperplane L such that

$$\max_{1 \leq j \leq n} |\text{Proj}_L(e_j)| < \frac{1}{100} \epsilon, \quad (5.24)$$

where e_j are the basis vectors for the x -coordinates.

Proof of Lemma 5.2. Assume (5.14, 5.15) hold with $T(y_{i_0})$ replacing $T(y_i)$, where as before

$$i_0 \simeq \frac{\log M}{\log b} \tilde{i}_0 \sim \frac{\log M_0}{\log b} < i, \quad (5.25)$$

$$\begin{aligned} \|[R_J T(y_{i_0}) R_J]^{-1}\| &< e^{M_0^\beta} \quad (0 < \beta < 1), \\ \|[R_J T(y_{i_0}) R_J]^{-1}(k, k')\| &< e^{-\alpha|k-k'|}, \quad k, k' \in J, |k-k'| > M_0/10, \end{aligned} \quad (5.26)$$

for all

$$J = [-M_0, M_0]^{d+v} + k, \quad \frac{1}{2}M^i < |k| < M^{i+1}. \quad (5.27)$$

Recall that we are at stage i , so (H2) is satisfied. Hence

$$\|T(y_i) - T(y_{i_0})\| < \mathcal{O}(1)\delta_{i_0} < e^{-\alpha M_0} \quad (5.28)$$

by (5.25). This in turn implies that (5.14, 5.15) hold. So we only need to prove (5.26). ((5.28) is in fact a reason for the choice of M_0 in (5.5).)

Fix $I \in \Sigma_{i_0}$. For type (a) intervals, using (H3) for y_{i_0} , the off-diagonal elements in the n direction of S have exponential decay and $\|S\| \leq \mathcal{O}(\delta e^{-\alpha M_0})$. We use A.L. for the random Schrödinger operator $\epsilon \Delta_j + V_j$ to obtain (5.26) as follows.

To obtain the first estimate in (5.26), we make direct incisions in the frequency space. Let J be a type (a) interval. We require

$$|\pm n \cdot \omega + \epsilon \Delta_j + V_j| = |\pm n \cdot \omega + \mu_j| \geq 2e^{-M_0^\beta}$$

for all $(n, j) \in J$, where μ_j are the eigenvalues of $\epsilon \Delta_j + V_j$ restricted to the projection of J onto \mathbb{Z}^d .

When $n \neq 0$, this amounts to taking away a set in Ω of measure $\leq \mathcal{O}(1)e^{-M_0^\beta} \cdot M_0^{d+v}$. When $n = 0$, this is satisfied for $x \in X$ in view of (5.0). So

$$\|[R_J T(y_{i_0}) R_J]^{-1}\| < e^{M_0^\beta},$$

by using the exponential estimates on S . This gives the first estimate in (5.26).

To obtain the second estimate in (5.26), we use Anderson localization, i.e., on the probability set X defined in (4.12), for all E (here $E = n \cdot \omega$) there exists at most one pairwise disjoint bad elementary region in \mathbb{Z}^d of size $M_0^{1/C}$ in the projection of J onto \mathbb{Z}^d . A resolvent series in the j direction coupled with a resolvent series in the n direction using the above two estimates and the decay property of S gives the second estimate in (5.26) (cf. proof of Lemma 4.3).

Hence there is a set $\bar{\Gamma} \subset \Gamma_{i_0} \cap I$ with

$$\begin{aligned} \text{mes}_v \bar{\Gamma} &< \mathcal{O}(1)e^{-M_0^\beta} M^{(i+1)(d+v)} M_0^{d+v} \quad (0 < \beta < 1) \\ &< M^{-i} \end{aligned} \quad (5.29)$$

such that outside $\bar{\Gamma}$, (5.26) hold for all $J \subset [-M^{i+1}, M^{i+1}]^{d+v}$ satisfying $J \cap [-M_0, M_0]^d \times [-M^{i+1}, M^{i+1}]^v = \emptyset$.

To prove (5.26) for type (b) intervals J , we use Lemma 4.1 at scale $\tilde{i}_0 = [\log M_0 / \log M]$ and the decomposition Lemma 5.3. We illustrate this on the interval $J = [-M_0, M_0]^{d+\nu}$. We consider the set

$$\mathcal{S} = \{(\omega, \mathcal{V}, \theta) \in I \times \mathbb{R} \mid (4.9) \text{ fail for } T_{[-M_0, M_0]^{d+\nu}}^\theta(y_{i_0}), \\ \text{i.e., with } N \text{ replaced by } M_0, \text{ and } y_i \text{ replaced by } y_{i_0}\}, \quad (5.30)$$

where $I \in \Sigma_{i_0}$ is the same fixed interval as earlier. (Recall that $x \in X \subset \mathbb{R}^{\mathbb{Z}^d} \setminus \mathbb{R}^\nu$ is fixed.) Let $T_{M_0}^\theta(y_{i_0})$ denote $T_{[-M_0, M_0]^{d+\nu}}^\theta(y_{i_0})$. Each matrix element of $T_{M_0}^\theta(y_{i_0})$ is a rational function of ω, \mathcal{V} of degree at most $2(p+1)M^q i_0^3$ ($q \in \mathbb{N}^+$) and linear in θ (see (H4.i), (4.2–4.4, 4.7)). As before, the condition in (5.30) can be expressed in terms of determinants and hence polynomials in the matrix elements of $T_{M_0}^\theta(y_{i_0})$. This implies that \mathcal{S} is semi-algebraic of total degree at most $M_0^{C(d+\nu)} M^q i_0^3$ in $\mathbb{R}^{2\nu+1}$.

We now localize to $(\omega, \mathcal{V}) \in \Gamma_{i_0}$. We consider the set

$$\mathcal{S}' := \mathcal{S} \cap \{(I \cap \Gamma_{i_0} \cap \{DC_{A,c}(2M_0) \setminus \Omega_{M_0}\}) \times \mathbb{R}\} \subset \mathbb{R}^{\nu+1}, \quad (5.31)$$

where Ω_{M_0} is as in (4.54) of Lemma 4.7 with N replaced by M_0 , and Γ_{i_0} is determined by the Q -equations (2.12), which are polynomial in (ω, \mathcal{V}) of degree at most $2(p+1)M^q i_0^3$ ($q \in \mathbb{N}^+$). From Lemma 4.7, (4.54), $DC_{A,c}(2M_0) \setminus \Omega_{M_0}$ is determined by $M_0^{C^2(d+\nu)} \sim M^{C^2(d+\nu)i_0}$ monomials of degree 1. So \mathcal{S}' is semi-algebraic of total degree at most $M^{(q+1)i_0^3}$ in $\mathbb{R}^{\nu+1}$.

By Lemma 4.1, each section $\mathcal{S}'(\omega, \mathcal{V}) = \mathcal{S}'(\omega(\mathcal{V}), \mathcal{V})$ is of measure at most $e^{-M_0^\sigma}$ in θ ($0 < \sigma < 1/2$). So

$$\text{mes}_{\nu+1} \mathcal{S}' \leq e^{-M_0^\sigma} \quad (0 < \sigma < 1/2). \quad (5.32)$$

Our aim is to estimate, for $k \in \mathbb{Z}^\nu$, $\frac{1}{2}M^i < |k| \leq M^{i+1}$,

$$\text{mes}_\nu \{\mathcal{V} \mid (\mathcal{V}, \mathcal{V} + (\epsilon + \delta)\varphi_{i_0}(\mathcal{V}), k \cdot (\mathcal{V} + (\epsilon + \delta)\varphi_{i_0}(\mathcal{V}))) \in \mathcal{S}'\}. \quad (5.33)$$

Since Γ_{i_0} is a C^1 function, $(I \cap \Gamma_{i_0}) \times \mathbb{R}$ may be identified with an interval in $\mathbb{R}^{\nu+1}$, say $[0, 1]^\nu \times \mathbb{R}$, \mathcal{S}' defined in (5.31) is a subset of $(I \cap \Gamma_{i_0}) \times \mathbb{R}$, and therefore can be identified with a subset in $[0, 1]^\nu \times \mathbb{R}$. For the purpose of application of Lemma 5.3, we identify $[0, 1]^\nu \times \mathbb{R}$ with $[0, 1]^\nu \times \{0\} \times \mathbb{R} \subset [0, 1]^\nu \times [0, 1]^{\nu-1} \times \mathbb{R}$ and \mathcal{S}' with $\mathcal{S}' \times \{0\} \subset \mathcal{S}' \times [0, 1]^{\nu-1}$. Since T is restricted to the interval $[-M_0, M_0]^{d+\nu}$, we may further restrict the interval in $[0, 1]^\nu \times \mathbb{R}$ to be

$$[0, 1]^\nu \times [-\mathcal{O}(\sqrt{d+\nu})M_0, \mathcal{O}(\sqrt{d+\nu})M_0] \\ \simeq [0, 1]^{2\nu-1} \times [-\mathcal{O}(\sqrt{d+\nu})M_0, \mathcal{O}(\sqrt{d+\nu})M_0]. \quad (5.34)$$

We decompose $[-\mathcal{O}(\sqrt{d+\nu})M_0, \mathcal{O}(\sqrt{d+\nu})M_0]$ into intervals of length 1 and identify each of them with $[0, 1]$. Applying the decomposition Lemma 5.3 to each of these intervals and taking the union, we obtain a subset $\Gamma' \subset \Gamma_{i_0} \cap I$ with

$$\text{mes}_\nu \Gamma' < (M^{(q+1)i_0^3})^C M^{-i} \cdot \mathcal{O}(\sqrt{d+\nu})M_0 < M^{-i/2} \quad (5.35)$$

(recall \mathcal{S}' is of degree $B \leq M^{q_0^3}$ ($q \in \mathbb{N}^+$) and we take $\epsilon = 200M^{-i}$ in (5.22)) such that for all $k \in \mathbb{Z}^\nu$, $|k| > \frac{1}{2}M^i$,

$$\begin{aligned} & \text{mes}_\nu \{ \mathcal{V} \mid (\mathcal{V}, \mathcal{V} + (\epsilon + \delta)\varphi_{i_0}(\mathcal{V}), k \cdot (\mathcal{V} + (\epsilon + \delta)\varphi_{i_0}(\mathcal{V})) \\ & \in \mathcal{S} \cap \{ (\Gamma_{i_0} \setminus \Gamma') \cap I \cap \{ DC_{A,c}(2M_0) \setminus \Omega_{M_0} \} \times \mathbb{R} \} < e^{-M_0^{\sigma/2}} \quad (0 < \sigma < 1/2), \end{aligned} \quad (5.36)$$

where M_0 is as in (5.5).

The same estimates as in (5.35, 5.36) hold when $T_{[-M_0, M_0]^{d+\nu}}$ is replaced by $T_{[-M_0, M_0]^{d+\nu} + \ell}$, $\ell \in \mathbb{Z}^d \cap [-2M_0, 2M_0]^d$. Therefore, there is a set $\Gamma'' \subset \Gamma_{i_0} \cap I$ with

$$\text{mes}_\nu \Gamma'' < \mathcal{O}(1)M^{-i/2} + \mathcal{O}(1)M^{(i+1)(d+\nu)}e^{-M_0^{\sigma/2}} < M^{-i/3} \quad (5.37)$$

(in view of the choice of M_0 in (5.5) and $C\sigma \gg 1$ from (4.58)) such that outside Γ'' , (4.9) hold for all intervals J of the form $[-M_0, M_0]^{d+\nu} + \ell$, $\ell \in \mathbb{Z}^d \cap [-2M_0, 2M_0]^d$, $y_{i'} = y_{i_0}$ and $\theta = k \cdot \omega$, $k \in \mathbb{Z}^\nu$, $\frac{1}{2}M^i < |k| < M^{i+1}$. (The condition $\cap \{ DC_{A,c}(2M_0) \setminus \Omega_{M_0} \}$ in (5.36) does not require additional incisions in the frequency space, as (H4.0) holds starting at stage $i_0 + 1$.)

Combined with (5.29) and the previous perturbation argument of replacing y_i by y_{i_0} in (5.28), this implies that there exists Γ''' with $\text{mes}_\nu \Gamma''' < M^{-i/3}$ such that outside Γ''' , (5.14, 5.15) hold for all k with $\frac{1}{2}M^i < |k| < M^{i+1}$, and a fixed $I \in \Sigma_{i_0}$.

Letting I range over Σ_{i_0} (there can be at most $\mathcal{O}(1)M_0^C$ such intervals, as the (ω, \mathcal{V}) parameter space can be restricted to, say $[0, 1]^{2\nu}$), the total measure removed from Γ_{i_0} is at most $M_0^C \cdot M^{-i/3} < M^{-i/4}$. Since Γ_{i_0} and Γ_i are at distance $< \delta_{i_0} < e^{-\alpha M_0}$ from (5.3), we obtain a subset $\tilde{\Gamma}_i \subset \Gamma_i$ with $\text{mes}_\nu \tilde{\Gamma}_i < M^{-i/4}$ such that (5.14, 5.15) hold for all k with $M^i/2 < |k| < M^{i+1}$ and on

$$\bigcup_{I \in \Sigma_{i_0}} (I \cap (\Gamma_i \setminus \tilde{\Gamma}_i)) \quad (5.38)$$

and hence on

$$\bigcup_{I \in \Sigma_i} (I \cap (\Gamma_i \setminus \tilde{\Gamma}_i)) \quad (5.39)$$

by (H4.iv). This proves the lemma. \square

Lemma 5.1 then gives that on the set in (5.39), (5.16, 5.17) hold. Clearly by perturbation, (5.16, 5.17) remain valid on an $M^{-(i+1)C}$ neighborhood of (5.39) (since $M^{-(i+1)C} \ll e^{-\alpha M_0} \sim e^{-\alpha(i+1)C/2}$ by the choice of M_0 in (5.5)), which in turn generates a collection Λ_{i+1} of intervals in $\mathbb{R}^{2\nu}$ of size $M^{-(i+1)C}$ such that for $(\omega, \mathcal{V}) \in I \in \Lambda_{i+1}$, (5.16, 5.17) hold. So (H4.iii) holds at stage $i + 1$ with $\tilde{\alpha} = \alpha - M^{-(i+1)\tilde{\delta}}$ ($\tilde{\delta} > 0$) replacing α . Moreover we have

$$\text{mes}_\nu \left(\bigcup_{I \in \Lambda_i} (I \cap \Gamma_i) \setminus \bigcup_{I' \in \Lambda_{i+1}} (I' \cap \Gamma_i) \right) \leq \text{mes}_\nu \tilde{\Gamma}_i < M^{-i/4}, \quad (5.40)$$

which will imply (H4.iv) at stage $i + 1$, once we construct y_{i+1} and hence Γ_{i+1} using (H4.iii) at stage $i + 1$.

6. Construction of y_{i+1} and completion of the assemblage

Construction of y_{i+1}

Let $N = M^{i+1}$, and for $(\omega, \mathcal{V}) \in \bigcup_{I \in \Lambda_{i+1}} I$, define

$$\Delta_{i+1}y = y_{i+1} - y_i := -[T_N(y_i)]^{-1}F(y_i) \quad (6.1)$$

(previously (3.9)). In view of (3.1, 2.16, H4.i), this implies that $\Delta_{i+1}y$ is a rational function of (ω, \mathcal{V}) of degree at most

$$\mathcal{O}(1)N^{d+\nu}M^{qi^3} + (2p+1)M^{qi^3} < M^{q(i+1)^3} \quad (q \in \mathbb{N}^+). \quad (6.2)$$

(Recall $p < M$.) So (H4.i) holds at stage $i+1$. (5.16, H4.ii) give

$$\|\Delta_{i+1}y\| < M^{(i+1)C} \kappa_i = \delta_{i+1} \quad (6.3)$$

and

$$\begin{aligned} \|\partial(\Delta_{i+1}y)\| &< \|\partial T_N^{-1}\| \|F(y_i)\| + \|T_N^{-1}\| \|\partial F(y_i)\| \\ &< \|T_N^{-1}\|^2 \|y_i\|_{C^1} \kappa_i + \|T_N^{-1}\| \bar{\kappa}_i \\ &< M^{2(i+1)C} \bar{\kappa}_i = \bar{\delta}_{i+1}, \end{aligned} \quad (6.4)$$

where we also used (H2). Next we obtain a pointwise estimate on $\Delta_{i+1}y$. From (6.1),

$$|\Delta_{i+1}y(k)| \leq \sum_{|k'| \leq N} |T_N^{-1}(k, k')| |F(y_i)(k')|. \quad (6.5)$$

(2.16) gives

$$\begin{aligned} |F(y_i)(k')| &\leq \mathcal{O}(1) \sum_{k_1 + \dots + k_{2p+1} = k'} |y_i(k_1)| \cdots |y_i(k_{2p+1})| \\ &\leq (CM)^{CM} |k'|^{(d+\nu)M} e^{-\alpha|k'|} \end{aligned} \quad (6.6)$$

(since $p < M$). Substituting (5.17, 6.6) into (6.5), we then obtain

$$\begin{aligned} |\Delta_{i+1}y(k)| &\leq (CM)^{CM} \left\{ \sum_{|k-k'| < i^C} M^{i^C} |k'|^{(d+\nu)M} e^{-\alpha|k'|} \right. \\ &\quad \left. + \sum_{|k-k'| \geq i^C} |k'|^{(d+\nu)M} e^{-\tilde{\alpha}(|k'|+|k-k'|)} \right\} \\ &< C' M^{2i^C} |k|^{(d+\nu)M} e^{-\tilde{\alpha}|k|}. \end{aligned} \quad (6.7)$$

Using (6.3) for k such that $\log |k| \lesssim i$ and (6.7) otherwise, we obtain

$$|y_{i+1}(k)| < e^{(\log |k|)^{C'} - \tilde{\alpha}|k|} \leq e^{-\tilde{\alpha}|k|} \quad (6.8)$$

with $\bar{\alpha} = \alpha - M^{-(i+1)\bar{\delta}}$ for some $\bar{\delta} > 0$ independent of i , where we used the estimate on $\tilde{\alpha}$ just above (5.40). This shows that (H3) is essentially preserved at stage $i + 1$. Here we used the fact that $\alpha = \mathcal{O}(1)|\log(\epsilon + \delta)|$ and $0 < \epsilon, \delta \ll 1$.

Since the intervals in Λ_{i+1} are of size $M^{-(i+1)^C}$, we may extend $\Delta_{i+1}y$ to the entire (ω, \mathcal{V}) parameter space as follows. For any $I \in \Lambda_{i+1}$, let $\tilde{I} \subset I$ be such that $\text{dist}(\Lambda_{i+1}^c, \tilde{I}) \sim \frac{1}{3}M^{-(i+1)^C}$. Set $\Delta y'_{i+1} = \Delta y_{i+1}$ on I , $\Delta y'_{i+1} = 0$ on Λ_{i+1}^c . Define a C^1 function

$$\Xi_{i+1} = \begin{cases} 1 & \text{on } \tilde{I}, \\ 0 & \text{on } \Lambda_{i+1}^c. \end{cases} \quad (6.9)$$

Set

$$\Delta \tilde{y}_{i+1} = \Xi_{i+1} \Delta y'_{i+1}. \quad (6.10)$$

Then $\Delta \tilde{y}_{i+1}$ is defined on the whole (ω, \mathcal{V}) parameter space and satisfies

$$\|\partial \Delta \tilde{y}_{i+1}\| < 3M^{(i+1)^C} \delta_{i+1} + M^{2(i+1)^C} \bar{\kappa}_i = \bar{\delta}_{i+1}, \quad (6.11)$$

where the second contribution comes from (6.4). Renaming $\Delta \tilde{y}_i$ as Δy_i and letting $y_{i+1} = y_i + \Delta y_i$, we have thus shown that (H1–3) remain valid at stage $i + 1$ with $\bar{\alpha}$ replacing α .

From y_{i+1} , the \mathcal{Q} -equations (2.12) define Γ_{i+1} at most at a distance $\delta_{i+1} \simeq M^{-b^{i+1}} \ll M^{-i/4}$ from Γ_i . Clearly (5.40) implies

$$\text{mes}_\nu \left(\Gamma_{i+1} \cap \left(\bigcup_{I \in \Lambda_i} I \setminus \bigcup_{I' \in \Lambda_{i+1}} I' \right) \right) < M^{-i/4} < M^{-(i+1)/5}, \quad (6.12)$$

which is (H4.iv) at stage $i + 1$.

It remains to verify the properties of $F(y_{i+1})$ in (H4.ii), stage $i + 1$. From the Taylor series in (3.10),

$$F(y_{i+1}) = -(T - T_N)[T_N(y_i)]^{-1} F(y_i) + \mathcal{O}(1)\|\Delta_{i+1}y\|^2, \quad N = M^{i+1}. \quad (6.13)$$

By construction and (H1), we have $\text{supp } y_i \subset [-M^i, M^i]^{d+\nu}$; therefore (2.16) gives

$$\begin{aligned} \text{supp } F(y_i) &\subset [-(2p+1)M^i, (2p+1)M^i]^{d+\nu} \\ &\subset [-M^{i+1}/10, M^{i+1}/10]^{d+\nu} = [-N/10, N/10]^{d+\nu}. \end{aligned} \quad (6.14)$$

So

$$\begin{aligned} F(y_{i+1}) &= [R_{\mathbb{Z}^{d+\nu} \setminus B(0, N)} T T_N^{-1} R_{B(0, N/10)}] F(y_i) + \mathcal{O}(1)\|\Delta_{i+1}y\|^2, \\ \|F(y_{i+1})\| &\leq \|R_{\mathbb{Z}^{d+\nu} \setminus B(0, N)} T T_N^{-1} R_{B(0, N/10)}\| \|F(y_i)\| + \mathcal{O}(1)\|\Delta_{i+1}y\|^2, \end{aligned} \quad (6.15)$$

where $B(0, N) = [-N, N]^{d+\nu}$, and $R_{B(0, N)}$, $R_{B(0, N/10)}$ are the characteristic functions. Thus

$$\|F(y_{i+1})\| \leq e^{-\alpha N/3} \kappa_i + \mathcal{O}(1)\delta_{i+1}^2 = \kappa_{i+1}, \quad (6.16)$$

where we used (H2,3,4.ii).

Similarly,

$$\begin{aligned}
\mathcal{O}(1)\|\partial F(y_{i+1})\| &\leq \|T_N^{-1}\| \|\partial T\| \|F(y_i)\| + \|\partial T_N^{-1}\| \|F(y_i)\| \|T\| \\
&\quad + \|\mathcal{R}_{\mathbb{Z}^{d+v} \setminus B(0,N)} T T_N^{-1} \mathcal{R}_{B(0,N/10)}\| \|\partial F(y_i)\| \\
&\quad + \|\Delta_{i+1} y\| \|\partial \Delta_{i+1} y\| \\
&< M^{2(i+1)^C} \kappa_i + e^{-\alpha M^{i+1}/3} \bar{\kappa}_i + \delta_{i+1} \bar{\delta}_{i+1},
\end{aligned} \tag{6.17}$$

and we may take

$$\bar{\kappa}_{i+1} = \mathcal{O}(1)(M^{2(i+1)^C} \kappa_i + e^{-\alpha M^{i+1}/3} \bar{\kappa}_i + \delta_{i+1} \bar{\delta}_{i+1}). \tag{6.18}$$

Summarizing (6.3, 6.4, 6.16–6.18), we have

$$\begin{cases} \delta_{i+1} = M^{(i+1)^C} \kappa_i, \\ \bar{\delta}_{i+1} = M^{2(i+1)^C} \bar{\kappa}_i, \\ \kappa_{i+1} = e^{-\alpha M^{i+1}/3} \kappa_i + \mathcal{O}(1) \delta_{i+1}^2, \\ \bar{\kappa}_{i+1} = \mathcal{O}(1)(M^{2(i+1)^C} \kappa_i + e^{-\alpha M^{i+1}/3} \bar{\kappa}_i + \delta_{i+1} \bar{\delta}_{i+1}). \end{cases} \tag{6.19}$$

We start from $\kappa_0, \bar{\kappa}_0 = \mathcal{O}(1)(\epsilon + \delta)$. For $\epsilon + \delta$ small enough, (6.19) is satisfied for $i \geq 1$ if

$$\begin{cases} \delta_i < \sqrt{\epsilon + \delta} M^{-(4/3)^i}, & \kappa_i < \sqrt{\epsilon + \delta} M^{-(4/3)^{i+2}}, \\ \bar{\delta}_i < \sqrt{\epsilon + \delta} M^{-(4/3)^i/2}, & \bar{\kappa}_i < \sqrt{\epsilon + \delta} M^{-(4/3)^{i+2}/2}. \end{cases} \tag{6.20}$$

(H4.0) and initial input for the induction

To ensure (H4.0) at stage $i + 1$, we make further incisions. (This is in order that Lemma 4.1 remains at our disposal at a later stage.) On Γ_{i+1} , we need to eliminate ω such that

$$\begin{cases} |n \cdot \omega_{i+1} + \lambda_{jj'}| \leq e^{-M_0^{\beta'}} & (n \sim \mathcal{O}(1)M_0), \\ \|n \cdot \omega_{i+1}\|_{\mathbb{T}} \leq c/|n|^A & (0 < |n| < 2M^{i+1}), \end{cases} \tag{6.21}$$

where M_0 is as in (5.5, H4.0), $\beta' = \beta/\mathcal{O}(1)$, $\mathcal{O}(1)$ is the same expansion factor as in Lemma 4.6 (denoted C there), $\lambda_{jj'} = \mu_j - \mu_{j'}$, $\mu_j, \mu_{j'}$ are the eigenvalues of the random Schrödinger operator $\epsilon \Delta_j + V_j$ restricted to the myriad elementary regions of size $M_0^{1/\mathcal{O}(1)}$ (the same expansion factor $\mathcal{O}(1)$) in $[-3M_0, 3M_0]^d$ (see Lemmas 4.1, 4.2, 4.6, proof of Lemma 4.2, (4.8, 4.9), the remark after (4.9, 4.26, 4.54) and the definition of Ω_i in (H4.0)). There are at most $\mathcal{O}(1)M_0^{C'd}$ ($C' > 0$) such differences of eigenvalues.

In view of (5.2) at stage $i + 1$, the first equation in (6.21) removes a set $\bar{\Gamma}_{i+1} \subset \Gamma_{i+1}$ with

$$\begin{aligned} \text{mes}_v \bar{\Gamma}_{i+1} &\leq e^{-M_0^{\beta''}} \quad (0 < \beta'' < \beta') \\ &\sim e^{-[(i+1)^{C/2} (\log M)^{C/2}]^{\beta''}} \ll M^{-(i+1)/5}, \end{aligned} \tag{6.22}$$

on using (5.5) and choosing

$$C > \frac{\mathcal{O}(1) \log M}{\beta},$$

which is always possible.

Since

$$\|\omega_{i+1} - \omega_i\| \leq \delta_i = M^{-(4/3)^i} \ll 1/M^{iA} \quad (6.23)$$

from (5.2, 5.3, 6.20), we only need to remove ω_{i+1} such that

$$\|n \cdot \omega_{i+1}\|_{\mathbb{T}} \leq c/|n|^A$$

for $M^i \leq |n| \leq M^{i+1}$, which removes a set $\tilde{\Gamma}_{i+1} \subset \Gamma_{i+1}$ with

$$\text{mes}_v \tilde{\Gamma}_{i+1} \leq \mathcal{O}(1)/M^{iA} \ll M^{-(i+1)/5}. \quad (6.24)$$

Rename α as α_i , and $\tilde{\alpha}$ as α_{i+1} . From (6.8), $\alpha_{i+1} = \alpha_i - M^{-(i+1)\delta} > \alpha/2$ uniformly in i . Combining (6.22, 6.24), we have (H4.0) at stage $i+1$ and (6.12) is preserved. We have thus made a complete induction step from stage i to $i+1$.

7. Proof of the Theorem

The ‘‘proof of the Theorem’’ is now just a matter of juxtaposing Sections 4, 5, 6 and recalling the sequence of events. We recount the spine of the argument.

We use the modified Newton scheme (3.9) to construct approximate solutions:

$$\Delta_{i+1}y = y_{i+1} - y_i = -[T_N(y_i)]^{-1}F(y_i), \quad N = M^{i+1}, \quad (7.1)$$

where T_N is T restricted on $[-N, N]^{d+v}$, and T and F are as in (3.3–3.5, 3.1). Assume we have obtained the first i approximations y_1, \dots, y_i on a set of intervals $\mathbb{R}^{2v} \supset \Lambda_1 \supset \dots \supset \Lambda_i$. To obtain y_{i+1} , we need to control $[T_N(y_i)]^{-1}$ with a further restriction on the new set of intervals Λ_{i+1} in (ω, \mathcal{V}) space. This is accomplished as follows.

To estimate $T_N(y_i)$, we cover $[-M^{i+1}, M^{i+1}]^{d+v}$ with the interval $[-M^i, M^i]^{d+v} = I$ and smaller intervals $J = [-M_0, M_0]^{d+v} + k$, $M^i/2 < |k| < M^{i+1}$, $M_0 \sim (\log N)^{C/2}$ as in (5.5). T_J^{-1} is ‘‘good’’ on Λ_i by using perturbation theory. The J intervals are divided into two types as in Section 5, according to their distances to the \mathbb{Z}^d axis (see equations (a) and (b) between (5.13, 5.14)). T_J^{-1} of type (a) is easily obtained by using the manifest exponential decay properties of y_i and a direct incision in the frequency space. The main task is to control T_J^{-1} of type (b), which leads to further incisions in the frequency space, hence to the new set of intervals Λ_{i+1} .

Since $|J| \ll |I|$, we may consider $T_J(y_{i_0})$ instead of $T_J(y_i)$ for some $i_0 \ll i$ as in (5.25). We add a parameter θ to $T_J(y_{i_0})$ and estimate the measure of the set of θ on the complement of which $[T_J^{(\theta)}]^{-1}$ is ‘‘good’’. This is Lemma 4.1. We then use the decomposition Lemma 5.3 to transfer the estimate in θ into estimates in ω , giving rise to the new set of intervals Λ_{i+1} .

On Λ_{i+1} we construct y_{i+1} according to (7.1). Using the Q -equations (2.12), we obtain Γ_{i+1} . The first i_0 approximations are constructed by using direct ϵ , δ series, in order that Lemma 4.7 and hence Lemma 4.1 are available: $i_0 \simeq \frac{1}{\beta} \log |\log(\epsilon + \delta)|$ from (4.57) and the third expression in (4.9) after setting $N = \tilde{N}_0^C$ and determining i , hence i_0 . (6.20) gives the rate of convergence of this Newton scheme and hence the Theorem. \square

Appendix: Localization results for random Schrödinger operators

A random Schrödinger operator is the operator

$$H = \epsilon \Delta + V \quad \text{on } \ell^2(\mathbb{Z}^d),$$

where $\epsilon > 0$ is a parameter, $\Delta(i, j) = 1$ if $|i - j| = 1$ and zero otherwise, and $V = \{v_i\}_{i \in \mathbb{Z}^d}$ is a family of independent identically distributed (i.i.d.) random variables with common probability distribution g . The spectrum of H is given by

$$\sigma(H) = \sigma(\epsilon \Delta) + \sigma(V) = [-2\epsilon d, 2\epsilon d] + \text{supp } g, \quad \text{a.s.}$$

We summarize below the known results on Anderson localization, which are relevant for the present construction (cf. [DJLS1,2, vDK, GB, GK, Mi, Si]). This is an expanded and more complete version of the appendix in [BW].

For any $L \in \mathbb{N}$, let $\Lambda_L(i)$ denote any elementary region in \mathbb{Z}^d with diameter $2L$, center $i \in \mathbb{Z}^d$ as defined in (4.10, 4.11) with \mathbb{Z}^d replacing $\mathbb{Z}^{d+\nu}$. Let $H_{\Lambda_L(i)}$ be H restricted to $\Lambda_L(i)$. Let $m > 0$ and $E \in \mathbb{R}$. The set $\Lambda_L(i)$ is (m, E) -regular (for a fixed V) if $E \notin \sigma(H_{\Lambda_L(i)})$ and

$$|G_{\Lambda_L(i)}(E; j, j')| \leq e^{-m|j-j'|} \quad (\text{A1})$$

for all $j, j' \in \Lambda_L(i)$ with $|j - j'| > L/4$. The following theorem is an immediate corollary of the corresponding theorem in [vDK] pertaining to cubes, by covering elementary regions with cubes and then applying the resolvent equation (cf. Lemma B).

Theorem A. *Let $I \subset \mathbb{R}$ be a bounded interval. Suppose that for some $L_0 > 0$,*

$$\text{Prob}\{\text{for any } E \in I \text{ either } \Lambda_{L_0}(i) \text{ or } \Lambda_{L_0}(j) \text{ is } (m_0, E)\text{-regular}\} \geq 1 - 1/L_0^{2p'} \quad (\text{A2})$$

for some $p' > d$ and $m_0 > 0$, and any $i, j \in \mathbb{Z}^d$ with $|i - j| > 2L_0$, and

$$\text{Prob}\{\text{dist}(E, \sigma(H_{\Lambda_L(0)})) < e^{-L^\beta}\} \leq 1/L^{q'} \quad (\text{A3})$$

for some β with $0 < \beta < 1$ and q with

$$q' > 4p' + 6d \quad (\text{A4})$$

and all E with

$$\text{dist}(E, I) \leq \frac{1}{2}e^{-L^\beta},$$

and all $L \geq L_0$. Then there exists α , $1 < \alpha < 2$, such that if we set $L_{k+1} = [L_k^\alpha] + 1$, $k = 0, 1, \dots$, and pick m , $0 < m < m_0$, there is $Q < \infty$ such that if $L_0 > Q$, then for any $k = 0, 1, \dots$,

$$\text{Prob}\{\text{for any } E \in I \text{ either } \Lambda_{L_k(i)} \text{ or } \Lambda_{L_k(j)} \text{ is } (m, E)\text{-regular}\} \geq 1 - 1/L_k^{2p'} \quad (\text{A5})$$

for any $i, j \in \mathbb{Z}^d$ with $|i - j| > 2L_k$.

Remark. On the same probability subspace,

$$\text{dist}(\sigma(H_{\Lambda_{L_k}(i)}), \sigma(H_{\Lambda_{L_k}(j)})) > e^{-L_k^\beta}, \quad \beta > 0, \quad (\text{A6})$$

if $|i - j| > 2L_k$. This is part of the proof of Theorem A.

Let $S \subset \mathbb{Z}^d$ be an (arbitrary) finite set. Let H_S be H restricted to S . If the probability distribution is absolutely continuous with a bounded density \tilde{g} , we have the following Wegner lemma:

$$\text{Prob}\{\text{dist}(E, \sigma(H_S)) \leq \kappa\} \leq C\kappa|S| \|\tilde{g}\|_\infty, \quad C, \kappa > 0. \quad (\text{A7})$$

(A2) is satisfied if ϵ is sufficiently small. (A3, 5) are provided by (A7) if $\|\tilde{g}\|_\infty < \infty$. More precisely, if we fix $0 < \beta < 1$, and choose q' and hence L_0 sufficiently large, then there exist ϵ sufficiently small such that (A2, 3) are satisfied. We note from (A4) that the larger the q' , the larger the p' could be. In view of (A5), q' can be chosen large if L_0 is large. So p' can always be large enough by choosing ϵ small enough for the construction in this paper (cf. proof of Lemma 4.1).

Theorem A implies that for $0 < \epsilon \ll 1$ and $\|\tilde{g}\|_\infty < \infty$, $\sigma(H)$ has pure point spectrum almost surely. The pure point spectrum is dense. However, it is simple [Si]. Let ψ_n ($n \in \mathbb{Z}^d$) be the n^{th} eigenfunction of H . Then

$$|\psi_n(j)| \leq C_{n,\omega} e^{-m'|j|} \quad (0 < m' < m).$$

Further improvements of technology (see [A, DJLS1,2, GB, GK]) give in fact

$$|\psi_n(j)| \leq C_\omega P_\omega(j_{n,\omega}) e^{-m'|j - j_{n,\omega}|}, \quad (\text{A8})$$

where the centers $j_{n,\omega}$ satisfy $|j_{n,\omega}| \gtrsim n^{1/d}$, and P_ω is a polynomial which only depends on ω .

A resolvent estimate

Lemma B. Suppose $\Lambda \subset \mathbb{Z}^{d+v}$ is an arbitrary set with the following property: for every $x \in \Lambda$, there is a subset $W(x) \subset \Lambda$ with $x \in W(x)$, $\text{diam } W(x) \leq N$ and such that Green's function $G_{W(x)}(E)$ satisfies, for certain $t, N, A > 0$,

$$\|G_{W(x)}(E)\| < A, \quad (\text{B1})$$

$$|G_{W(x)}(E; x, y)| < e^{-tN} \quad \text{for all } y \in \partial_* W(x). \quad (\text{B2})$$

Here $\partial_* W(x)$ is the interior boundary of $W(x)$ relative to Λ given by

$$\partial_* W(x) = \{y' \in W(x) \mid \exists z \in \Lambda \setminus W(x), |z - y'| = 1\}. \quad (\text{B3})$$

Then

$$\|G_\Lambda(E)\| < 2N^2 A$$

provided $4N^2 e^{-tN} \leq 1/2$.

See [BGS], where it is stated as Lemma 2.2, for a proof using the resolvent equation. See also the proof of Lemma 5.1 in Section 5 of the present paper for an essentially identical exercise in the resolvent equation. \square

Acknowledgments. W.-M. Wang thanks the Institute for Advanced Study for hospitality, where part of this work was done.

Research of J. Bourgain was partially supported by NSF grant DMS-03-22370.

Research of W.-M. Wang was partially supported by NSF grant DMS-97-29992 and DMS 0503563.

References

- [A] Aizenman, M.: Localization at weak disorder: some elementary bounds. *Rev. Math. Phys.* **6**, 1163–1182 (1994) Zbl 0843.47039 MR 1301371
- [AFHS] Aizenman, M., Friedrich, R., Hundertmark, D., Shankar, S.: Constructive fractional-moment criteria for localization in random operators. *Phys. A* **279**, 369–377 (2000) MR 1797147
- [AM] Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. *Comm. Math. Phys.* **157**, 245–278 (1993) Zbl 0782.60044 MR 1244867
- [AF] Albanese, C., Fröhlich, J.: Periodic solutions of some infinite-dimensional Hamiltonian systems associated with non-linear partial difference equations I. *Comm. Math. Phys.* **116**, 475–502 (1988) Zbl 0696.35185 MR 0937771
- [AFS] Albanese, C., Fröhlich, J., Spencer, T.: Periodic solutions of some infinite-dimensional Hamiltonian systems associated with non-linear partial difference equations II. *Comm. Math. Phys.* **119**, 677–699 (1988) Zbl 0699.35242 MR 0973022
- [An] Anderson, P.: Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492 (1958)
- [Ba] Basu, S.: On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets. *Discrete Comput. Geom.* **22**, 1–18 (1999) Zbl 0973.14033 MR 1692627
- [B1] Bourgain, J.: Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE. *Int. Math. Res. Not.* **1994**, no. 11, 475–497 Zbl 0817.35102 MR 1316975
- [B2] Bourgain, J.: Construction of approximate and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.* **6**, 201–230 (1996) Zbl 0872.35007 MR 1384610
- [B3] Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math.* **148**, 363–439 (1998) Zbl 0928.35161 MR 1668547

- [B4] Bourgain, J.: Nonlinear Schrödinger equations. In: *Hyperbolic Equations and Frequency Interactions*, IAS/Park City Math. Ser. 5, 3–157 (1999) Zbl 0952.35127 MR 1662829
- [B5] Bourgain, J.: Green's Function Estimates for Lattice Schrödinger Operators and Applications. *Ann. of Math. Stud.* 158, Princeton Univ. Press (2005) Zbl pre02133396 MR 2100420
- [B6] Bourgain, J.: On invariant tori of full dimension for 1D periodic NLS. *J. Funct. Anal.* **229**, 62–94 (2005) Zbl 1088.35061 MR 2180074
- [BGS] Bourgain, J., Goldstein, M., Schlag, W.: Anderson localization for Schrödinger operators on \mathbb{Z}^2 with quasi-periodic potential. *Acta Math.* **188**, 41–86 (2002) Zbl 1022.47023 MR 1947458
- [BW] Bourgain, J., Wang, W.-M.: Anderson localization for time quasi-periodic random Schrödinger and wave equations. *Comm. Math. Phys.* **3**, 429–466 (2004) MR 2076917
- [CP] Chierchia, L., Peretti, P.: Maximal almost-periodic solutions for Lagrangian equations on infinite dimensional tori. In: *Seminar on Dynamical Systems*, S. Kuksin et al. (eds.), Birkhäuser, Basel, 203–212 (1994) Zbl 0796.34028 MR 1279399
- [CW1] Craig, W., Wayne, C. E.: Newton's method and periodic solutions of nonlinear equations. *Comm. Pure Appl. Math.* **46**, 1409–1498 (1993) Zbl 0794.35104 MR 1239318
- [CW2] Craig, W., Wayne, C. E.: Periodic solutions of nonlinear Schrödinger equations and the Nash–Moser method. In: *Hamiltonian Mechanics*, Nato Adv. Sci. Inst. Ser. B Phys. 331, Plenum, NY, 103–122 (1994) MR 1316671
- [CFKS] Cycon, H. L., Froese, R. G., Kirsch, W., Simon, B.: *Schrödinger Operators, with Application to Quantum Mechanics and Global Geometry*. Springer (1987) Zbl 0619.47005 MR 0883643
- [DJLS1] Del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: What is localization. *Phys. Rev. Lett.* **75**, 117–119 (1995)
- [DJLS2] Del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: Operators with singular continuous spectrum IV: Hausdorff dimensions, rank one perturbations, and localization. *J. Anal. Math.* **69**, 153–200 (1996) Zbl 0908.47002 MR 1428099
- [vDK] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. *Comm. Math. Phys.* **124**, 285–299 (1989) Zbl 0698.60051 MR 1012868
- [E] Eliasson, L. H.: Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **15**, 115–147 (1988) Zbl 0685.58024 MR 1001032
- [EK] Eliasson, L. H., Kuksin, S. E.: KAM for the nonlinear Schrödinger equation. Preprint (2006)
- [FMSS] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in Anderson tight binding model. *Comm. Math. Phys.* **101**, 21–46 (1985) Zbl 0573.60096 MR 0814541
- [FS] Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy. *Comm. Math. Phys.* **88**, 151–184 (1983) Zbl 0519.60066 MR 0696803
- [FSW] Fröhlich, J., Spencer, T., Wayne, C. E.: Localization in disordered, nonlinear dynamical systems. *J. Statist. Phys.* **42**, 247–274 (1986) Zbl 0629.60105 MR 0833019
- [GB] Germinet, F., de Bièvre, S.: Dynamical localization for discrete and continuous random Schrödinger operators. *Comm. Math. Phys.* **194**, 323–341 (1998) Zbl 0911.60099 MR 1627657
- [GK] Germinet, F., Klein, A.: Bootstrap multiscale analysis and localization in random media. *Comm. Math. Phys.* **222**, 415–448 (2001) Zbl 0982.82030 MR 1859605

- [GMP] Gol'dsheid, Ya., Molchanov, S., Pastur, L.: Pure point spectrum of stochastic one dimensional Schrödinger operators. *Funct. Anal. Appl.* **11**, no. 1, 1–8 (1977) Zbl 0368.34015 MR 0470515
- [KP] Kuksin, S., Pöschel, J.: Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math.* **143**, 149–179 (1996) Zbl 0847.35130 MR 1370761
- [Le] Levin, Ya. B.: *Lectures on Entire Functions*. Transl. Math. Monogr. 150, Amer. Math. Soc., Providence, RI (1996) Zbl 0856.30001 MR 1400006
- [LZ] Levitan, B. M., Zhikov, V. V.: *Almost-Periodic Functions and Differential Equations*. Cambridge Univ. Press (1982) Zbl 0499.43005 MR 0690064
- [LL] Lieb, E. H., de Llano, M.: Solitons and the delta function fermion gas in the Hartree–Fock theory. *J. Math. Phys.* **19**, 860 (1978)
- [Mi] Minami, N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Comm. Math. Phys.* **177**, 709–725 (1996) Zbl 0851.60100 MR 1385082
- [O] Overhauser, A. W.: Structure of nuclear matter. *Phys. Rev. Lett.* **4**, 415–418 (1960) Zbl 0091.23005
- [PF] Pastur, L., Figotin, A.: *Spectra of Random and Almost Periodic Operators*. Springer (1992) Zbl 0752.47002 MR 1223779
- [Pö1] Pöschel, J.: Small divisors with spatial structure in infinite dimensional Hamiltonian systems. *Comm. Math. Phys.* **127**, 351–393 (1990) Zbl 0702.58065 MR 1037110
- [Pö2] Pöschel, J.: On the construction of almost periodic solutions for a nonlinear Schrödinger equation. *Ergodic Theory Dynam. Systems* **22**, 1537–1549 (2002) Zbl 1020.37044 MR 1934149
- [Sh] Shastry, B. S.: Exact solutions of a nonlinear eigenvalue problem in one dimension. *Phys. Rev. Lett.* **50**, 633–636 (1983) MR 0700922
- [Si] Simon, B.: Cyclic vectors in the Anderson model. *Rev. Math. Phys.* **6**, 1183–1185 (1994) Zbl 0841.60081 MR 1301372
- [W] Wang, W.-M.: Pure point spectrum of the Floquet Hamiltonian for the quantum harmonic oscillator under time quasi-periodic perturbations. Submitted (2006)
- [Wy] Wayne, C. E.: Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Comm. Pure Appl. Math.* **127**, 479–528 (1990) Zbl 0708.35087 MR 1040892