Commun. Math. Phys. (2024) 405:272
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-024-05150-z

Communications in
Mathematical Physics

Nonlinear Anderson Localized States at Arbitrary Disorder

Wencai Liu[1], **W.-M. Wang**[2]

[1] Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA.
E-mail: liuwencai1226@gmail.com; wencail@tamu.edu

[2] CNRS and Département de Mathématique, Cergy Paris Université, 95302 Cergy-Pontoise Cedex, France.
E-mail: wei-min.wang@math.cnrs.fr

Received: 8 April 2024 / Accepted: 19 September 2024
Published online: 24 October 2024 –

Abstract: Given an Anderson model $H = -\Delta + V$ in arbitrary dimensions, and assuming the model satisfies localization, we construct quasi-periodic in time (and localized in space) solutions for the nonlinear random Schrödinger equation $i\frac{\partial u}{\partial t} = -\Delta u + Vu + \delta|u|^{2p}u$ for small δ. Our approach combines probabilistic estimates from the Anderson model with the Craig–Wayne–Bourgain method for studying quasi-periodic solutions of nonlinear PDEs.

1. Introduction and the Main Theorem

The method developed in this paper is valid in arbitrary dimensions, we have, however, chosen to focus on the one-dimensional case in the first six sections. The reasons are two fold: firstly in one dimension, Anderson localization for the random schrödinger operator H holds more generally; secondly the main ideas can be illuminated more succinctly. In Sect. 7, however, we will extend the discussion to the arbitrary-dimensional case.

We start with the discrete nonlinear random Schrödinger equation (NLRS) in one dimension:

$$i\frac{\partial u}{\partial t} = -\Delta u + Vu + \delta|u|^{2p}u, \ p \in \mathbb{N}, \tag{1}$$

where Δ is the discrete Laplacian:

$$(\Delta u)(x) = u(x+1) + u(x-1),$$

and $V = \{v_x\}$ is a family of independently identically distributed random variables on $[0, 1]$, with distribution density g. Assume that g is bounded, $g \in L^\infty$. Denote by $\mathbb{P}$ the probability measure, the product measure on $[0, 1]^{\mathbb{Z}}$.

Let

$$H = -\Delta + V, \tag{2}$$

be the random Schrödinger operator. It is well-known, as a consequence of Furstenberg's theorem on positive Lyapunov exponent for products of random SL(2, $\mathbb{R}$) matrices [1], that with probability 1, H has Anderson localization, namely, pure point spectrum with exponentially decaying eigenfunctions [2,3]. (See also [4–8].) In higher dimensions and for large disorder, i.e., replacing V by λV, $\lambda \gg 1$, Anderson localization has been established using multiscale analysis [9], see also [10], or fractional moment method [11].

Assume that H has Anderson localization. Let $\{\phi_j^V\}_{j\in\mathbb{Z}}$ be the (real) eigen-basis of H. Assume that ℓ_j^V satisfies

$$|\phi_j^V(\ell_j^V)| = \max_x |\phi_j^V(x)|.$$

(If the maximum is not unique, one may choose among the maxima arbitrarily.) We call ℓ_j^V, the *localization center*. It suffices to say here that as a consequence of localization, there is an eigenfunction labelling such that when $j_1 < j_2$, $\ell_{j_1}^V \le \ell_{j_2}^V$ (see Sect. 3 and appendix A for details), and that we use this labelling.

So for a given V such that H has Anderson localization, let $j \in \mathbb{Z}$, and denote by ϕ_j^V and μ_j^V the eigenfunctions and corresponding eigenvalues of H. When $\delta = 0$, all solutions to (1) are of the form

$$\sum_{j\in\mathbb{Z}} c_j e^{-i\mu_j^V t}\phi_j^V,$$

with appropriate c_j, which decay to 0, as $j \to \pm\infty$. This paper is concerned with the case $c_j \neq 0$ for finite (but arbitrary) number of j.

We prove the following nonlinear analogue:

Theorem 1.1. *Consider the discrete NLRS in one dimension:*

$$i\frac{\partial u}{\partial t} = -\Delta u + Vu + \delta|u|^{2p}u,\ p \in \mathbb{N}. \tag{3}$$

For any given $\epsilon > 0$, there exists a length $l_\epsilon > 0$. Fix $L \ge l_\epsilon$ and b lattice sites: $\beta_k \in \mathbb{Z}$, $k = 1, 2, \cdots, b$ satisfying $10L \le |\beta_k| \le L^3$ and $|\beta_k - \beta_{k'}| \ge 10L$ for $k \neq k'$, $k, k' = 1, 2, \cdots, b$. Then there exists a subset of potentials $X_\epsilon \subset [0, 1]^{\mathbb{Z}}$ with $\mathbb{P}(X_\epsilon) \ge 1 - \epsilon$ and $\delta_0 > 0$ (depending on g, ϵ and L) such that the following holds: Fix any $V \in X_\epsilon$ and $0 < \delta \le \delta_0$, and consider any b eigenfunctions $\phi_{\alpha_k}^V$ with localization centers near β_k, satisfying $\ell_{\alpha_k}^V \in B_k = \{\ell \in \mathbb{Z} : |\ell - \beta_k| \le L\}$, $k = 1, 2, \cdots, b$. Let $a = (a_1, a_2, \cdots, a_b) \in [1, 2]^b$ and consider the solution to the linear equation:

$$u_0(t, x) = \sum_{k=1}^{b} a_k e^{-i\mu_{\alpha_k}^V t}\phi_{\alpha_k}^V(x).$$

There exists a subset of amplitudes $A_\delta \subset [1, 2]^b$ of measure at least $1 - e^{-|\log\delta|^{1/2}}$, such that for any $a \in A_\delta$, the nonlinear Eq. (3) has a solution $u(t, x)$ satisfying

$$u(t, x) = \sum_{(n,j)\in\mathbb{Z}^b\times\mathbb{Z}} \hat{u}(n, j)e^{in\cdot\omega t}\phi_j^V(x) = \sum_{k=1}^{b} a_k e^{-i\omega_k t}\phi_{\alpha_k}^V(x) + O(\delta^{1/2}),$$

where $\omega = (\omega_1, \omega_2, \cdots, \omega_b) = (\mu_{\alpha_1}^V, \mu_{\alpha_2}^V, \cdots, \mu_{\alpha_b}^V) + O(\delta)$, and $\hat{u}(n, j)$ decay exponentially as $|(n, j)| \to \infty$.

Remark 1. (1) We note that l_ϵ also depends on g and b. For large L and any B_k, $k = 1, 2, \cdots, b$, there are at least $2(1-\epsilon)L$ normalized eigenfunctions $\phi^V_{\alpha_k}$ such that the localization centers $\ell^V_{\alpha_k}$ lie in B_k.
(2) We could replace L^3 with e^{L^κ}, $0 < \kappa < 1$.
(3) We observe that we need to remove a set of measure $e^{-|\log\delta|^{1/2}}$, which is larger than $|\delta|^{1/2}$, the typical measure removed by KAM-type approaches.
(4) The proof of Theorem 1.1 is general. In Sect. 7, we discuss the generalization to arbitrary dimensions by stating the required spectral conditions on H. These conditions are satisfied in arbitrary dimensions at high disorder, when V is replaced by λV with $\lambda \gg 1$.
(5) In Theorem 1.1, we need to exclude potentials with a small probability ϵ and carefully select eigenfunctions with specific localization centers. While these restrictions could potentially be relaxed, they cannot be entirely removed due to the nature of our proof.

1.1. About Theorem 1.1. The linear solution u_0 is localized in space, and quasi-periodic in time, with frequencies the b eigenvalues of the linear random Schrödinger operator. Theorem 1.1 shows that under small nonlinear perturbations, for a large set of amplitudes, there is a solution to the nonlinear equation nearby. This nonlinear solution u remains localized in space and quasi-periodic in time; moreover the frequencies are harmonics of the modulated b Fourier modes of the linear random Schrödinger equation.

The NLRS in (1) can be viewed as an effective equation for a many body system. Theorem 1.1 can then be seen as showing the existence of finite particle localized states. For related physics literature, see e.g., [12,13], and for a review on many body Anderson localization, see [14].

Theorem 1.1 is a KAM-type persistence result. Most of these results relate to perturbations of systems where the solutions are explicitly known. The random Schrödinger equation belongs, however, to an *entirely different* category: its eigenfunctions can only be known *qualitatively*. Nonetheless, Theorem 1.1 shows persistence of time quasi-periodic, localized solutions. Moreover there is an abundance of such solutions. This is the main novelty.

The NLRS is a nonlinear parameter dependent (the random potentials) *difference equation* on $\mathbb{Z}^d$. For KAM results on nonlinear PDE's, such as the parameter dependent nonlinear Schrödinger equation on the torus $\mathbb{T}^d$, see e.g., [15]. Note, however, that by Fourier series, for the latter, the solutions to the linear equation are known explicitly.

Remark 2. Previously the paper [16], see also the recent improvement [17], established existence of quasi-periodic solutions, originating from linear combinations of Dirac δ-functions at large disorder. It perturbs about the diagonal operator λV, $\lambda \gg 1$, in the canonical $\mathbb{Z}^d$ basis. The method is not applicable here. For periodic solutions originating from eigenfunctions at large disorder, see [18,19].

For other nonlinear models, see e.g., [20–23].

Remark 3. The symbol $\hat{u}$ is merely a notation here and does not carry the connotation of being the dual of u.

1.2. Ideas of the proof. One of the main ideas is to use properties of Anderson localization to first identify a subset in V, a "good" set of large measure, and subsequently in the nonlinear analysis, *fix* a V in this set, and work in the eigenfunction basis provided by the random Schrödinger operator. We give the requirements to be good in Sect. 1.3 below. Fixing a potential circumvents the lack of control of the eigenfunctions as the potentials vary. Note moreover that, as mentioned earlier, generally it is not possible to know precisely the eigenfunctions of the random Schrödinger operators, even for large disorder, see [24].

So fix indeed such a good potential V (see Sect. 1.3), and let

$$u_0(t,x) = \sum_{k=1}^{b} a_k e^{-i\mu_{\alpha_k}^V t}\phi_{\alpha_k}^V(x), \tag{4}$$

be a solution to the linear equation, as in Theorem 1.1. As an ansatz, we seek solutions of the form:

$$u(t,x) = \sum_{(n,j)\in\mathbb{Z}^b\times\mathbb{Z}} \hat{u}(n,j)e^{in\cdot\omega t}\phi_j^V(x). \tag{5}$$

Note that $u(t,x)$ of the above form are closed under multiplication and complex conjugation. So we may seek solutions to (1) in this form.

Using (5) in (1) leads to the following nonlinear system of equations on $\mathbb{Z}^b\times\mathbb{Z}$:

$$(n\cdot\omega+\mu_j^V)\hat{u}(n,j)+\delta W_{\hat{u}}(n,j)=0,\ (n,j)\in\mathbb{Z}^b\times\mathbb{Z}, \tag{6}$$

where, to give an idea, when $p=1$,

$$\begin{aligned} W_{\hat{u}}(n,j) = &\sum_{\substack{n_1+n_2-n_3=n\\ n_1,n_2,n_3\in\mathbb{Z}^b}}\ \sum_{j_1,j_2,j_3\in\mathbb{Z}} \hat{u}(n_1,j_1)\hat{u}(n_2,j_2)\overline{\hat{u}(n_3,j_3)}\\ &\left(\sum_{x\in\mathbb{Z}}\phi_j^V(x)\phi_{j_1}^V(x)\phi_{j_2}^V(x)\phi_{j_3}^V(x)\right); \end{aligned} \tag{7}$$

while for general p,

$$\begin{aligned} W_{\hat{u}}(n,j) = &\sum_{\substack{n'+\sum_{m=1}^p(n_m-n'_m)=n\\ n',n_m,n'_m\in\mathbb{Z}^b}}\ \sum_{j',l_m,l'_m\in\mathbb{Z}} \hat{u}(n',j')\prod_{m=1}^{p}\hat{u}(n_m,l_m)\overline{\hat{u}(n'_m,l'_m)}\\ &\left(\sum_{x\in\mathbb{Z}}\phi_j^V(x)\phi_{j'}^V(x)\prod_{m=1}^{p}\phi_{l_m}^V(x)\phi_{l'_m}^V(x)\right). \end{aligned} \tag{8}$$

1.3. The good potentials. The linear solution u_0 solves (1) to order δ. One may write u_0 in the form (5), with $\hat{u}(-e_k,\alpha_k)=a_k$, $k=1,2,\cdots,b$, where e_k is the kth basis vector of $\mathbb{Z}^b$, $\hat{u}(n,j)=0$ otherwise, and $\omega_k=\mu_{\alpha_k}^V$, $k=1,2,\cdots,b$. The vector $W_{\hat{u}}$ in (6) depends on a_k, $k=1,2,\cdots,b$. Generally speaking, one would need parameters to solve the nonlinear equation (6) using a Newton scheme, starting from the approximate solution u_0. Since V is fixed, the a_k's are the parameters in the problem. There is however, a δ factor in front of $W_{\hat{u}}$.

The small $\mathcal{O}(\delta)$ parameters pose difficulties mainly at small scales: $|(n,j)|\ll\delta^{-1}$, when estimating the inverse of the linearized operators. The key new idea is that we can overcome this difficulty if the diagonals of the linear operator in (6) satisfy a *clustering property*. Roughly speaking, this means that if two diagonal elements are "not close", then they are "far apart". (One may think of the integers, which have this property: if two integers are not equal, then they are at least of distance 1.) This then permits localizing about the diagonals in $\mathcal{O}(\delta)$ intervals, which compensates for the small $\mathcal{O}(\delta)$ parameters. The potentials V that lead to clustering properties, in addition to Anderson localization, are *good* potentials.

It should be emphasized that the clustering property is only needed at *small* scales, and *not* large ones. This makes the approach robust, potentially applicable to many problems.

Remark 4. Deterministic clustering was first used by Wang to prove existence of time quasi-periodic solutions to nonlinear partial differential equations on the torus in arbitrary dimensions, see the papers [25] and [26]. In [26], this was established using number theory.

1.4. Anderson localization and clustering property of the diagonals. Given a linear operator A, we say that A has the *clustering property* if there exists $c > 0$, so that the spectrum of A, $\sigma(A)$ is contained in the union of intervals:

$$\sigma(A) \subset \bigcup_i I_i,$$

satisfying dist $(I_i, I_j) > c > 0$, if $i \neq j$.

For the random Schrödinger operator H in (2), we use Anderson localization to establish probabilistic clustering at small scales. So one may set ω to be the frequencies of the u_0 in (4), which are b eigenvalues of the H. The diagonals in (6) then correspond to a family of harmonics, i.e., certain linear combinations of the eigenvalues of H.

The proof of the clustering of these (low lying) harmonics is rather delicate. The Minami estimate [27] on eigenvalue spacing plays a fundamental role. Semi-uniform property of Anderson localization, see [28,29], is essential. Wegner estimate [30] comes into play as well. This is done in sects. 2 and 3, and the conclusion is summarized in Theorem 3.5, which also provides lower bounds on the diagonals. The clustering property permits the analysis to go beyond small perturbative scales, and is one of the main points of the paper.

1.5. Small scale analysis. The clustering property indicates that at *small scales* the spectrum of the diagonal operator has many *gaps*. Using perturbation theory, the linearized operators are invertible in the gaps; while away from the gaps, one may work locally in intervals of size $\mathcal{O}(\delta)$, and consequently parameters of size $\mathcal{O}(\delta)$ (extracted from the nonlinear term) suffice for the analysis. This is the case for the proof of the large deviation theorem when applying the Cartan estimates in Sect. 4, as well as for the semi-algebraic projection in Sect. 6.

1.6. Large scale analysis. Large scale analysis is related to what has been done before in [16], cf. also Chap. 18 [31], which are a priori tailored for parameters of order 1. However, after incorporating the local argument recounted above, it can be adapted and used to prove Theorem 1.1.

1.7. Comparison with other methods. In the nonlinear setting, as mentioned earlier in Remark 2, most other method that we are aware of starts from explicit solutions to the underlying non-perturbed system. This includes the previous results on the existence of nonlinear Anderson localized states at high disorder in [16,17], which start from the linear equation:

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

and *not* the usual random Schrödinger equation. Compared to [16,17], an additional job done in this paper is the fine linear analysis in sects. 2–4. This seems indispensable when perturbing from a system, which is known only qualitatively, and renders the method more robust.

1.8. Organization of the paper. Section 2 establishes the subset of *good* potentials; Sect. 3 makes linear estimates for small scales; Sect. 4 proves a large deviation theorem, to be used for the nonlinear analysis at large scales; Sects. 5 and 6 finally solve (6) and hence (1), using a Lyapunov–Schmidt decomposition and Theorems 3.5 and 4.1. In Sect. 7, we discuss the arbitrary dimension generalization.

2. One Dimensional Random Schrödinger Operators in Finite Volumes

As mentioned earlier in the introduction, we seek solutions in the form (5). Using (5), the NLRS maybe expressed as a nonlinear matrix equation. In order to have a precise form for this matrix equation, namely the W in (7) and (8), we need to analyze the eigenfunctions of the linear random Schrödinger operator on $\ell^2(\mathbb{Z})$. Recall that unlike the exponentials, there is no known exact form for the eigenfunctions of the random Schrödinger operator.

Toward that purpose, we first restrict to some finite volume $\Lambda \subset \mathbb{Z}$ and study the random Schrödinger operator on $\ell^2(\Lambda)$. We then use properties of Anderson localization in Sect. 3, to deduce, in particular, properties of the eigenvalues and the eigenfunctions on $\ell^2(\mathbb{Z})$.

Our main aim is to give lower bounds on the diagonals of the matrix, as well as their spacings. This is to prepare for the invertibility analysis in Sect. 4. These estimates will be probabilistic.

Below we study the one dimensional random Schrödinger operator $H = -\Delta + V$ on $\ell^2(\Lambda)$. For an operator H on $\ell^2(\mathbb{Z}^d)$ and $\Lambda \subset \mathbb{Z}^d$, let $H_\Lambda = R_\Lambda H R_\Lambda$, where R_Λ is the restriction to Λ. For $n = (n_1, n_2, \cdots, n_d) \in \mathbb{Z}^d$, let $|n| = \max_{j\in\{1,2,\cdots,d\}} |n_j|$ denote the ℓ^∞ norm. For $\Lambda \subset \mathbb{Z}$, denote by $\tilde{\mu}_j^\Lambda$, $j \in \Lambda$, eigenvalues of $H_\Lambda = R_\Lambda H R_\Lambda$, with corresponding normalized eigenvectors $\tilde{\phi}_j^\Lambda$.

Remark 5. Note that $\tilde{\mu}_j^\Lambda$ and $\tilde{\phi}_j^\Lambda$ depend on the realization of the potentials in Λ. It is convenient here to label the eigenvalues and normalized eigenvectors by $j \in \Lambda$, instead of $j \in \{1, 2, \cdots, |\Lambda|\}$.

For a ball $B = \{\ell \in \mathbb{Z} : |\ell - \ell_0| \le l\}$ of size l with center ℓ_0, denote by rB, the dilation: $rB = \{\ell \in \mathbb{Z} : |\ell - \ell_0| \le rl\}$. Fix $L > 0$. Let $\tilde{B}_{(k,L)}$ be balls of size L, $k = 1, 2, \cdots, b$. Assume that

$$L \le \operatorname{dist}(\tilde{B}_{(k,L)}, 0) \le L^4,$$

and for any distinct k and k',

$$\mathrm{dist}(\tilde{B}_{(k,L)}, \tilde{B}_{(k',L)}) \geq 6L. \tag{9}$$

Since L will be fixed, we omit the dependence on L of $\tilde{B}_{(k,L)}$ and write simply $\tilde{B}_k$.

Recall that e_k, $k = 1, 2, \cdots, b$, are the canonical basis vectors of $\mathbb{Z}^b$. Let $C_0 > 0$, $\gamma_0 > 0$, $q_0 > 0$ be three fixed constants, which will be determined later. Assume that H_Λ has b eigen-pairs $\tilde{\mu}^\Lambda_{\tilde{\alpha}_k}$ and $\tilde{\phi}^\Lambda_{\tilde{\alpha}_k}$, $k = 1, 2, \cdots, b$ such that for any $k = 1, 2, \cdots, b$, there exists $\tilde{\ell}^\Lambda_{\tilde{\alpha}_k} \in \tilde{B}_k$ such that

$$|\tilde{\phi}^\Lambda_{\tilde{\alpha}_k}(\ell)| \leq C_0(1 + |\tilde{\ell}^\Lambda_{\tilde{\alpha}_k}|)^{q_0} e^{-\gamma_0|\ell - \tilde{\ell}^\Lambda_{\tilde{\alpha}_k}|}, \ell \in \Lambda. \tag{10}$$

Let $\tilde{\omega}^\Lambda_k = \tilde{\mu}^\Lambda_{\tilde{\alpha}_k}$, $k = 1, 2, \cdots, b$ and $\tilde{\omega}^\Lambda = (\tilde{\omega}^\Lambda_1, \tilde{\omega}^\Lambda_2, \cdots, \tilde{\omega}^\Lambda_b) \in \mathbb{R}^b$. When there is no ambiguity, we omit the dependence on Λ. In the following, $\delta > 0$ is sufficiently small.

Remark 6. Condition (10) is motivated by semi-uniform properties of Anderson localization in infinite volume, see Theorem 3.1.

2.1. Estimates on the diagonals. Let $\Lambda_1 = \left[-2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor, 2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor\right]$, where $\lfloor x \rfloor$ is the integer part of x. (The choice of scales is in view of the later nonlinear analysis.) Denote by S^1 the probability event that H_{Λ_1} has eigen-pairs $\tilde{\mu}^{\Lambda_1}_{\tilde{\alpha}_k}$ and $\tilde{\phi}^{\Lambda_1}_{\tilde{\alpha}_k}$, $k = 1, 2, \cdots, b$ satisfying (10) (with $\Lambda = \Lambda_1$), and there exists either

$$(n, j) \in \left[-2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor, 2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor\right]^{b+1} \backslash \{-e_k, \tilde{\alpha}_k\}_{k=1}^b, \tag{11}$$

such that

$$|n \cdot \tilde{\omega}^{\Lambda_1} + \tilde{\mu}^{\Lambda_1}_j| \leq 4\delta^{\frac{1}{8}} \tag{12}$$

or

$$(n, j) \in \left[-2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor, 2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor\right]^{b+1} \backslash \{e_k, \tilde{\alpha}_k\}_{k=1}^b, \tag{13}$$

such that

$$|-n \cdot \tilde{\omega}^{\Lambda_1} + \tilde{\mu}^{\Lambda_1}_j| \leq 4\delta^{\frac{1}{8}} \tag{14}$$

and the normalized eigenvector corresponding to $\tilde{\mu}^{\Lambda_1}_j$ satisfies for some $\tilde{\ell}^{\Lambda_1}_j \in \Lambda_1$,

$$|\tilde{\phi}^{\Lambda_1}_j(\ell)| \leq C_0(1 + |\tilde{\ell}^{\Lambda_1}_j|)^{q_0} e^{-\gamma_0|\ell - \tilde{\ell}^{\Lambda_1}_j|}, \ell \in \Lambda_1. \tag{15}$$

Fix a large constant $q_1 > 0$. Denote by S_N the probability event such that for any $j \in [-N, N]$, $j' \in [-N, N]$ and $j \neq j'$,

$$|\tilde{\mu}^\Lambda_j - \tilde{\mu}^\Lambda_{j'}| \geq \frac{1}{N^{q_1}}, \Lambda = [-N, N]. \tag{16}$$

Denote by $\mathbb{P}$ the measure as before, and $\mathbb{E}$ the expectation.

Theorem 2.1. *For small δ, we have*

$$\mathbb{P}(S^1) \le e^{C|\log\delta|^{\frac{3}{4}}}\delta^{\frac{1}{8}},$$

where C is a large constant independent of δ.

We will use the following three lemmas to prove Theorem 2.1.

Lemma 2.2. (Wegner estimate [30], see also [32]) *Let $\Lambda \subset \mathbb{Z}$. For any $E \in \mathbb{R}$ and $\varepsilon > 0$,*

$$\mathbb{E}(\mathrm{dist}(E, \sigma(H_\Lambda)) \le \varepsilon) \le C|\Lambda|\varepsilon.$$

Lemma 2.3. (Minami estimate [27], see also [32]) *Let $\Lambda \subset \mathbb{Z}$ and $J \subset \mathbb{R}$ be an interval. Then we have*

$$\mathbb{E}([\mathrm{tr}(\mathbf{1}_J(H_\Lambda))]\cdot[\mathrm{tr}(\mathbf{1}_J(H_\Lambda)) - 1]) \le C|\Lambda|^2|J|^2,$$

where $\mathbf{1}_J$ is the characteristic function of the interval J. In particular, we have that

$$\mathbb{P}(\textit{there exist distinct } j, j' \in \Lambda \textit{ such that } |\tilde{\mu}_j^\Lambda - \tilde{\mu}_{j'}^\Lambda| \le \varepsilon) \le C\varepsilon|\Lambda|^2,$$

and hence

$$\mathbb{P}(S_N) \le CN^{-q_1+2}.$$

Lemma 2.4. *Let $\Lambda \subset \mathbb{Z}$. Assume that eigenvalues $\tilde{\mu}_j$, $j \in \Lambda$, of H_Λ are simple and let*

$$d = \min_{j \neq j', j\in\Lambda, j'\in\Lambda} |\tilde{\mu}_j - \tilde{\mu}_{j'}| > 0. \tag{17}$$

Let $(\tilde{\mu}, \tilde{\phi})$ be an eigen-pair of H_Λ. Let $l \in \Lambda$, and $(\tilde{\mu}^s, \tilde{\phi}^s)$ (depending continuously on s) be an eigen-pair of $H_\Lambda + sI_{\{l\}}$ with $|s| \le \frac{d}{10}$ satisfying $\tilde{\mu}^s|_{s=0} = \tilde{\mu}$ and $\tilde{\phi}^s|_{s=0} = \tilde{\phi}$. Then

$$\frac{d\tilde{\mu}^s}{ds} = |\tilde{\phi}(l)|^2 + |s|O\left(\frac{|\Lambda|}{d}\right) + |s|^2 O\left(\frac{|\Lambda|^2}{d^2}\right),$$

and hence

$$\tilde{\mu}^s - \tilde{\mu} = s|\tilde{\phi}(l)|^2 + |s|^2 O\left(\frac{|\Lambda|}{d}\right) + |s|^3 O\left(\frac{|\Lambda|^2}{d^2}\right).$$

Proof. Let $\{(\tilde{\mu}_j, \tilde{\phi}_j),\ j \in \Lambda\}$ be the complete set of eigen-pairs of H_Λ. Without loss of generality, assume that $\tilde{\mu} = \tilde{\mu}_1$ and $\tilde{\phi} = \tilde{\phi}_1$.

By a standard perturbation argument and (17), one has that for any $|s| \le \frac{d}{10}$,

$$|\tilde{\mu}^s - \tilde{\mu}_1| \le \frac{d}{2}$$

and

$$|\tilde{\mu}^s - \tilde{\mu}_{j'}| \ge \frac{d}{2},\ j' \in \Lambda\backslash\{1\}. \tag{18}$$

Let

$$\tilde{\phi}^s = \sum_{j\in\Lambda} c_j^s \tilde{\phi}_j.$$

Then

$$(H_\Lambda + sI_{\{l\}})\tilde{\phi}^s = \tilde{\mu}^s\tilde{\phi}^s = \sum_{j\in\Lambda} \tilde{\mu}^s c_j^s \tilde{\phi}_j \tag{19}$$

and

$$H_\Lambda\tilde{\phi}^s = \sum_{j\in\Lambda} c_j^s \tilde{\mu}_j \tilde{\phi}_j. \tag{20}$$

Since $\|(H_\Lambda + sI_{\{l\}})\tilde{\phi}^s - H_\Lambda\tilde{\phi}^s\| = O(s)$, by (18), (19) and (20), one has that for any $j \in \Lambda\backslash\{1\}$,

$$|c_j^s| \le O\left(\frac{s}{d}\right).$$

This implies that $1 - O\left(\frac{|\Lambda|s^2}{d^2}\right) \le (c_1^s)^2 \le 1$. Therefore, one has that

$$|\tilde{\phi}^s(l)|^2 = |\tilde{\phi}(l)|^2 + |s|O\left(\frac{|\Lambda|}{d}\right) + |s|^2 O\left(\frac{|\Lambda|^2}{d^2}\right). \tag{21}$$

By eigenvalue variations (aka Hellmann-Feynman Theorem) and (21), one has that

$$\begin{aligned}
\frac{d\tilde{\mu}^s}{ds} &= \langle\tilde{\phi}^s, \frac{d(H_\Lambda + sI_{\{l\}})}{ds}\tilde{\phi}^s\rangle \\
&= |\tilde{\phi}^s(l)|^2 \\
&= |\tilde{\phi}(l)|^2 + |s|O\left(\frac{|\Lambda|}{d}\right) + |s|^2 O\left(\frac{|\Lambda|^2}{d^2}\right).
\end{aligned}$$

We conclude that

$$\tilde{\mu}^s - \tilde{\mu} = s|\tilde{\phi}(l)|^2 + |s|^2 O\left(\frac{|\Lambda|}{d}\right) + |s|^3 O\left(\frac{|\Lambda|^2}{d^2}\right).$$

□

For any $n = (n_1, n_2, \cdots, n_b) \in \mathbb{Z}^b$, denote by

$$\operatorname{supp} n = \#\{n_k : n_k \neq 0, k = 1, 2, \cdots, b\},$$

where # denotes the number of elements in a set.

Proof of Theorem 2.1. Let $N = 2\left\lfloor e^{|\log\delta|^{\frac{3}{4}}}\right\rfloor$ and $\tilde{S}_1 = S_N$. By Lemma 2.3 (Minami estimate), it suffices to prove that

$$\mathbb{P}(S^1 \cap \tilde{S}_1) \leq e^{C|\log\delta|^{\frac{3}{4}}}\delta^{\frac{1}{8}}.$$

Since the total number of (n, j) in (11) and (13) is bounded by $(2N+1)^{b+1}$, it suffices to prove that for a fixed (n, j) in the set,

with probability at most $e^{C|\log\delta|^{\frac{3}{4}}}\delta^{\frac{1}{8}}$, either (12) or (14) holds. Without loss of generality, we only consider the probability event $P_{n,j}$ (P for simplicity) that (15) holds, and

$$|n\cdot\tilde{\omega} + \tilde{\mu}_j| \leq 4\delta^{\frac{1}{8}}. \tag{22}$$

Our goal is to show that

$$\mathbb{P}(P_{n,j} \cap \tilde{S}_1) \leq e^{C|\log\delta|^{\frac{3}{4}}}\delta^{\frac{1}{8}}. \tag{23}$$

Case 1: $n = 0$. In this case, Wegner estimate implies (23).

Case 2: supp $n = 1$

Case 2_1: supp $n = 1$ and $n = -e_k, k = 1, 2, \cdots, b$.

Without loss of generality, assume that $n = (-1, 0, \cdots, 0)$. In this case, $n\cdot\tilde{\omega}+\tilde{\mu}_j = -\tilde{\mu}_{\tilde{\alpha}_1} + \tilde{\mu}_j$ and $j \neq \tilde{\alpha}_1$. In this case, it follows from (16).

Case 2_2: supp $n = 1$, $n = e_k, k = 1, 2, \cdots, b$ or $n = re_k, k = 1, 2, \cdots, b$ with $|r| \geq 2$.

Without loss of generality, assume that $n = (n_1, 0, \cdots, 0)$ with $n_1 = 1$ or $|n_1| \geq 2$.

Denote by $\tilde{\phi}_{\tilde{\alpha}_1}$ and $\tilde{\phi}_j$ eigenvectors of eigenvalues $\tilde{\mu}_{\tilde{\alpha}_1} = \tilde{\omega}_1$ and $\tilde{\mu}_j$. For any $l \in 2\tilde{B}_1$, denote by P_l the probability event such that

$$|\tilde{\phi}_{\tilde{\alpha}_1}(l)|^2 \geq \frac{4}{5}|\tilde{\phi}_j(l)|^2 + \frac{1}{L^{10}}, \tag{24}$$

where $L \geq L_0(C_0, q_0, \gamma_0)$ is sufficiently large.

By (10), one has that

$$\sum_{\ell\notin 2\tilde{B}_1} |\tilde{\phi}_{\tilde{\alpha}_1}(\ell)|^2 \leq e^{-\frac{\gamma_0}{2}L},$$

and hence

$$\sum_{\ell\in 2\tilde{B}_1} |\tilde{\phi}_{\tilde{\alpha}_1}(\ell)|^2 \geq 1 - e^{-\frac{\gamma_0}{2}L}. \tag{25}$$

By (24) and (25), we claim that

$$P_{n,j} \subset \bigcup_{l\in 2\tilde{B}_1} P_l. \tag{26}$$

Otherwise,

$$\sum_{\ell\in 2\tilde{B}_1} |\tilde{\phi}_{\tilde{\alpha}_1}(\ell)|^2 < \frac{10}{L^9} + \sum_{\ell\in 2\tilde{B}_1} |\tilde{\phi}_j(l)|^2$$
$$< \frac{4}{5} + \frac{10}{L^9},$$

which contradicts (25).

Take any $l \in 2\tilde{B}_1$ such that (24) holds. Split $[0, 1]$ into N^{8q_1} intervals of size N^{-8q_1} and take any interval $I = [f_1, f_2]$. Define the probability event $P_{l,I}$ to be $P_{l,I} = \{V \in P_l : V_l \in I\}$. Applying Lemma 2.4, one has that if we fix V_ℓ, $\ell \in \Lambda_1\backslash\{l\}$, then for any $f \in I$,

$$\frac{d\tilde{\mu}_{\tilde{\alpha}_1}^{V_l=f}}{df} = |\tilde{\phi}_{\tilde{\alpha}_1}^{V_l=f_1}(l)|^2 + N^{-5q_1}O(1), \tag{27}$$

and

$$\frac{d\tilde{\mu}_j^{V_l=f}}{df} = |\tilde{\phi}_j^{V_l=f_1}(l)|^2 + N^{-5q_1}O(1), \tag{28}$$

where $\tilde{\mu}_i^{V_l=f}$ is the eigenvalue with the potential $V_l = f$ and fixed V_ℓ, $\ell \in \Lambda_1\backslash\{l\}$.

Obviously,

$$\frac{d\tilde{\mu}_j^{V_l=f}}{df} \geq 0. \tag{29}$$

From (24), (27), (28) and (29), one has that

$$\begin{aligned}\left|\frac{d(n\cdot\tilde{\omega}^{V_l=f} + \tilde{\mu}_j^{V_l=f})}{df}\right| &\geq \frac{3}{5}|\tilde{\phi}_{\tilde{\alpha}_1}^{V_l=f_1}(l)|^2 - N^{-5q_1}O(1)\\ &\geq \frac{3}{5L^{10}} - N^{-5q_1}O(1)\\ &\geq \frac{1}{2L^{10}}.\end{aligned} \tag{30}$$

Since $g \in L^\infty[0, 1]$, by (22) and (30), one has that for any $l \in 2\tilde{B}_1$,

$$\mathbb{P}(P_{n,j} \cap P_l \cap \tilde{S}_1) \leq O(1)N^{8q_1}L^{10}\delta^{\frac{1}{8}}. \tag{31}$$

By (26) and (31), one has that

$$\mathbb{P}(P_{n,j} \cap \tilde{S}_1) \leq O(1)N^{8q_1}L^{11}\delta^{\frac{1}{8}}. \tag{32}$$

It implies (23) since δ is sufficiently small (depending on L).

Case 3: supp $n \geq 2$

Let n_i be such that $|n_i| = \max_{i\in\{1,2,\cdots,b\}} |n_i|$.

Case 3_1: $|n_i| \geq 2$

Without loss of generality, assume that $n_i = n_1$. It is easy to see that for any $\ell \in 2\tilde{B}_1$

$$|\tilde{\phi}_{\tilde{\alpha}_k}(\ell)|^2 \leq e^{-\gamma_0 L}, k = 2, 3, \cdots, b. \tag{33}$$

For any $l \in 2\tilde{B}_1$, denote by P_l^1 the probability event that

$$|\tilde{\phi}_{\tilde{\alpha}_1}(\ell)|^2 \geq \frac{4}{5}|\tilde{\phi}_j(\ell)|^2 + L^2\left(\sum_{k=2}^{b}|\tilde{\phi}_{\tilde{\alpha}_k}(\ell)|^2\right) + \frac{1}{L^{10}}. \tag{34}$$

By (25), (33) and (34), one has that (the proof is similar to that of (26))

$$P_{n,j} \subset \bigcup_{l \in 2\tilde{B}_1} P_l^1. \tag{35}$$

Replacing (26) with (35), and following the proof of Case 2_2 (using Lemma 2.4), we still have (23).

Case 3_2: $|n_i| \leq 1$, namely $n_k = 0, \pm 1, k = 1, 2, \cdots, b$.

In this case, clearly, there exist at least two non-zero n_k, $k = 1, 2, \cdots, b$.

Without loss of generality, assume that $n_1 \neq 0$ and $n_2 \neq 0$. In this case, if $\tilde{\ell}_j \geq 2L^4$, by (15), one has that for any $\ell \in \cup_{k=1}^{b} 2\tilde{B}_k$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}. \tag{36}$$

If $\tilde{\ell}_j \leq 2L^4$, by (9) and (15), one has that either for any $\ell \in 2\tilde{B}_1$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}, \tag{37}$$

or for any $\ell \in 2\tilde{B}_2$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}. \tag{38}$$

Without loss of generality assume that (37) holds. Therefore, for any $\ell \in 2\tilde{B}_1$,

$$|\tilde{\phi}_i(\ell)|^2 \leq e^{-\gamma_0 L}, i = \tilde{\alpha}_2, \tilde{\alpha}_3, \cdots, \tilde{\alpha}_b, \text{ and } i = j. \tag{39}$$

For any $l \in 2\tilde{B}_1$, denote by P_l^2 the probability event that

$$|\tilde{\phi}_{\tilde{\alpha}_1}(l)|^2 \geq L^2|\tilde{\phi}_j(l)|^2 + L^2\left(\sum_{k=2}^{b}\tilde{\phi}_{\alpha_k}(l)|^2\right) + \frac{1}{L^{10}}. \tag{40}$$

By (25), (39) and (40), one has that

$$P_{n,j} \subset \bigcup_{l \in 2\tilde{B}_1} P_l^2. \tag{41}$$

Replacing (26) with (41), and following the proof of Case 2_2 or Case 3_1 (also using Lemma 2.4), we have (23). □

2.2. Spacing of the diagonals. For the purpose of nonlinear analysis, it suffices to work with scales $|\log\delta|^s$, $s>1$. So let $\Lambda_2^s=\left[-2\lfloor|\log\delta|^s\rfloor, 2\lfloor|\log\delta|^s\rfloor\right]$. Denote by S_2^s the probability event that there exist eigen-pairs $\tilde{\mu}_{\tilde{\alpha}_k}^{\Lambda_2^s}$ and $\tilde{\phi}_{\tilde{\alpha}_k}^{\Lambda_2^s}$, $k=1,2,\cdots,b$ satisfying (10), and $m\in\mathbb{Z}^b$ with $|m|\le|2\log\delta|^s$, either $|m_k|\ge 2$ for some $k\in\{1,2,\cdots,b\}$ or supp $m\ge 3$, and $j, j'\in\Lambda_2^s$ satisfying

$$|m\cdot\tilde{\omega}^{\Lambda_2^s}+\tilde{\mu}_j^{\Lambda_2^s}-\tilde{\mu}_{j'}^{\Lambda_2^s}|\le 4\delta^{\frac{1}{8}}, \tag{42}$$

and the eigenvectors corresponding to $\tilde{\mu}_j^{\Lambda_2^s}$ and $\tilde{\mu}_{j'}^{\Lambda_2^s}$ satisfy that there exist $\tilde{\ell}_j^{\Lambda_2^s}\in\Lambda_2^s$ and $\tilde{\ell}_{j'}^{\Lambda_2^s}\in\Lambda_2^s$ such that for any $\ell\in\Lambda_2^s$,

$$|\tilde{\phi}_j^{\Lambda_2^s}(\ell)|\le C_0(1+|\tilde{\ell}_j^{\Lambda_2^s}|)^{q_0}e^{-\gamma_0|\ell-\tilde{\ell}_j^{\Lambda_2^s}|},\ |\tilde{\phi}_{j'}^{\Lambda_2^s}(\ell)|\le C_0(1+|\tilde{\ell}_{j'}^{\Lambda_2^s}|)^{q_0}e^{-\gamma_0|\ell-\tilde{\ell}_{j'}^{\Lambda_2^s}|}. \tag{43}$$

We should mention that we allow $j=j'$.

Theorem 2.5. *For small δ, we have*

$$\mathbb{P}(S_2^s)\le(|\log\delta|)^{Cs}\delta^{\frac{1}{8}}.$$

Proof. Let $N=2\lfloor|\log\delta|^s\rfloor$ and $\tilde{S}_2^s=S_N$. By Lemma 2.3 (Minami estimate), it suffices to prove that

$$\mathbb{P}(S_2^s\cap\tilde{S}_2^s)\le(|\log\delta|)^{Cs}\delta^{\frac{1}{8}}.$$

Denote by $P_{m,j,j'}$ the probability event that (10) and (42) hold. Let m_i be such that $|m_i|=\max\{|m_k|, k=1,2,\cdots,b\}$ (If there are several, choose one such m_i). Without loss of generality, assume that $m_i=m_1$. Let us prove the case $|m_1|\ge 2$ first. Without loss of generality, assume that $m_1\ge 0$. It is easy to see that for any $\ell\in 2\tilde{B}_1$,

$$|\tilde{\phi}_{\tilde{\alpha}_k}(\ell)|^2\le e^{-\gamma_0 L},\ k=2,3,\cdots,b. \tag{44}$$

For any $l\in 2\tilde{B}_1$, denote by P_l^3 the probability event that

$$|\tilde{\phi}_{\tilde{\alpha}_1}(l)|^2\ge\frac{4}{5}|\tilde{\phi}_{j'}(l)|^2+L^2(\sum_{k=2}^{b}\tilde{\phi}_{\tilde{\alpha}_k}(l)|^2)+\frac{1}{L^{10}}. \tag{45}$$

By (25), (44) and (45), one has that

$$P_{m,j,j'}\subset\bigcup_{l\in 2\tilde{B}_1}P_l^3. \tag{46}$$

Now the proof follows from Lemma 2.4, which is similar to the proof of Case 2_2 or Case 3_1. Below are the details.

Take any $l\in 2\tilde{B}_1$ such that (45) holds. Split $[0,1]$ into N^{8q_1} intervals of size N^{-8q_1} and take any interval $I=[f_1, f_2]$. Denote the probability event: $P_{l,I}^3=\{V\in P_l: V_l\in I\}$.

Applying Lemma 2.4, one has that if we fix V_ℓ, $\ell \in \Lambda_1\backslash\{l\}$, then for any $f \in I$,

$$\frac{d\tilde{\mu}_{\tilde{\alpha}_1}^{V_l=f}}{df} = |\tilde{\phi}_{\tilde{\alpha}_1}^{V_l=f_1}(l)|^2 + N^{-5q_1}O(1), \tag{47}$$

and

$$\frac{d\tilde{\mu}_{j'}^{V_l=f}}{df} = |\tilde{\phi}_{j'}^{V_l=f_1}(l)|^2 + N^{-5q_1}O(1). \tag{48}$$

From (47), (48) and (29), one has that

$$\begin{aligned}\frac{d(m\cdot\tilde{\omega}^{V_l=f} + \tilde{\mu}_j^{V_l=f} - \tilde{\mu}_{j'}^{V_l=f})}{df} &\geq \frac{3}{4}|\tilde{\phi}_{\tilde{\alpha}_1}^{V_l=f_1}(l)|^2 + N^{-5q_1}O(1)\\ &\geq \frac{3}{4L^{10}} + N^{-5q_1}O(1)\\ &\geq \frac{1}{2L^{10}}.\end{aligned} \tag{49}$$

The proof now follows that of Case 2_2 or Case 3_1 of Theorem 2.1.

Let us proceed to the case supp $m \geq 3$. Without loss of generality, assume that $m_i \neq 0, i = 1, 2, 3$. In this case, if $\tilde{\ell}_j \geq 2L^4$, by (43), one has that for any $\ell \in \cup_{k=1}^b 2\tilde{B}_k$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}. \tag{50}$$

If $\tilde{\ell}_j \leq 2L^4$, by (9) and (43), one has that either for any $\ell \in 2\tilde{B}_1 \cup 2\tilde{B}_2$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}, \tag{51}$$

or for any $\ell \in 2\tilde{B}_2 \cup 2\tilde{B}_3$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}, \tag{52}$$

or for any $\ell \in 2\tilde{B}_1 \cup 2\tilde{B}_3$,

$$|\tilde{\phi}_j(\ell)| \leq e^{-\gamma_0 L}. \tag{53}$$

Clearly, (50)-(53) also hold for j'. Therefore, we have that there exists $i \in \{1, 2, 3\}$, such that for any $m \in \{\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \tilde{\alpha}_b\}\backslash\{\tilde{\alpha}_i\}$ and $m = j, j'$,

$$|\tilde{\phi}_m(\ell)| \leq e^{-\gamma_0 L}, \ell \in 2\tilde{B}_i. \tag{54}$$

The proof subsequently follows that of Case 3_2 of Theorem 2.1.

□

Denote by $\hat{S}_2^s$ the probability event that there exist $\tilde{\omega}^{\Lambda_2^s}$ satisfying (10), $m \in \mathbb{Z}^b$ with $\sum_{k=1}^b m_k \neq 0$, $|m_k| \leq 1, k = 1, 2, \cdots, b$, and $j \in \Lambda_2^s$, $j' \in \Lambda_2^s$ satisfying

$$|m\cdot\tilde{\omega}^{\Lambda_2^s} + \tilde{\mu}_j^{\Lambda_2^s} - \tilde{\mu}_{j'}^{\Lambda_2^s}| \leq 4\delta^{\frac{1}{8}}. \tag{55}$$

Theorem 2.6. *For small δ, we have*

$$\mathbb{P}(\hat{S}_2^s) \leq (|\log\delta|)^{Cs}\delta^{\frac{1}{8}}.$$

Proof. Again, let $N = 2\lfloor|\log\delta|^s\rfloor$ and $\tilde{S}_2^s = S_N$. By Lemma 2.3 (Minami estimate), it suffices to prove that

$$\mathbb{P}(\hat{S}_2^s \cap \tilde{S}_2^s) \leq (|\log\delta|)^{Cs}\delta^{\frac{1}{8}}.$$

Since $\sum_{k=1}^b m_k \neq 0$, by the normalization of of eigenvectors there exists $l \in [-N, N]$, such that

$$|(\sum_{k=1}^b m_k|\tilde{\phi}_{\tilde{\alpha}_k}(l)|^2) + |\tilde{\phi}_j(l)|^2 - |\tilde{\phi}_{j'}(l)|^2| \geq \frac{1}{N^{10}}. \tag{56}$$

We may use Lemma 2.4 to conclude as in the proof of Case 2_2 of Theorem 2.1.

□

Remark 7. We could replace the constant 4 in (12), (14), (42) and (55) with any fixed constant and all the statements in the theorems in this section would still hold.

3. One Dimensional Random Schrödinger Operators

Using Anderson localization properties, we show in this section that the lower bounds on the diagonals and their spacings remain essentially unaltered in the infinite volume, $\mathbb{Z}$. There are now, however, infinite number of eigenfunctions and we need to label them appropriately. Appx. A provides such a labelling scheme. We will not enter the details here, except mentioning that semi-uniform localization properties of the eigenfunctions (stated as Theorem 3.1 below) plays an essential role.

Let

$$H = -\Delta + V,$$

be the random Schrödinger operator on $\ell^2(\mathbb{Z})$ as before. Let $\{\varphi_j^V\}_j$ be the eigen-basis and assume that ι_j^V satisfies

$$|\varphi_j^V(\iota_j^V)| = \max_{x\in\mathbb{Z}} |\varphi_j^V(x)|.$$

(As mentioned earlier, if the maximum is not unique, the choice could be any one of the maxima.) We have

Theorem 3.1. (See e.g., [5] or [33, sect. 1.6]) *There exist some $q > 0$ and $\gamma_1 > 0$ such that, with probability 1,*

$$|\varphi_j^V(\ell)| \leq C_V(1 + |\iota_j^V|)^q e^{-\gamma_1|\ell-\iota_j^V|}, \tag{57}$$

where $\mathbb{E}(C_V) < \infty$.

Remark 8. • Recall that ι_j^V is called the localization center.

• γ_1 and q only depend on the distribution g.

From Theorem 3.1 and a proof similar to that of Theorem 7.1 in [28], one has the following Lemma concerning the localization centers.

Lemma 3.2. *For any ϵ, there exist $\mathcal{V}_\epsilon$ with $\mathbb{P}(\mathcal{V}_\epsilon) > 1-\epsilon$ and a constant l_ϵ such that the following holds. For any $V \in \mathcal{V}_\epsilon$ and $k \in [-L^4, L^4]$ with $L \geq l_\epsilon$,*

$$(1-\epsilon)L \leq \#\{j : \iota_j^V \in [k, k+L]\} \leq (1+\epsilon)L. \tag{58}$$

See appendix B for a proof. □

Basing on (58), one may (re)label the eigenfunctions so that if $j > j'$, then the localization centers of the corresponding eigenfunctions ϕ_j^V and $\phi_{j'}^V$ satisfy $\ell_j^V \geq \ell_{j'}^V$. The construction of such a map is presented in appendix A. Here after relabelling, we use the notations ϕ_j^V instead of φ_j^V and ℓ_j^V instead of ι_j^V. Recall that μ_j^V is the eigenvalue corresponding to eigenfunction ϕ_j^V. When there is no ambiguity, we omit the dependence on V.

Below we summarize properties of the eigenfunction basis in this labelling.

Lemma 3.3. *There exist $q > 0$ and $\gamma > 0$ such that for any ϵ, there exist $\mathcal{V}_\epsilon$ with $\mathbb{P}(\mathcal{V}_\epsilon) > 1-\epsilon$ and constants C_ϵ and ℓ_ϵ such that for any $V \in \mathcal{V}_\epsilon$, the following statements hold:*

- *for any $\ell \in \mathbb{Z}$,*

$$|\phi_j^V(\ell)| \leq C_\epsilon(1+|\ell_j^V|)^q e^{-\gamma|\ell-\ell_j^V|}, \tag{59}$$

- *for any $|\ell_j^V| \geq \ell_\epsilon$,*

$$|\ell_j^V - j| \leq \epsilon|j|, \tag{60}$$

- *for any $k \in [-L^4, L^4]$,*

$$(1-\epsilon)L \leq \#\{j : \ell_j^V \in [k, k+L]\} \leq (1+\epsilon)L. \tag{61}$$

Lemma 3.4. *Let $\Lambda = [-2N, 2N]$ with sufficiently large N. Choose any $V \in \mathcal{V}_\epsilon \cap S_{2N}$. Consider any eigen-pair (μ_j, ϕ_j), $|j| \leq N$, of $H = -\Delta + V$. Then there exists an eigen-pair $(\tilde{\mu}_{\tilde{j}}^\Lambda, \tilde{\phi}_{\tilde{j}}^\Lambda)$, $\tilde{j} \in \Lambda$ such that*

$$|\mu_j - \tilde{\mu}_{\tilde{j}}^\Lambda| \leq e^{-\frac{\gamma}{2}N},$$

and

$$\|\phi_j - \tilde{\phi}_{\tilde{j}}^\Lambda\| \leq e^{-\frac{\gamma}{2}N}.$$

Proof. By (60), one has that $|\ell_j| \leq (1+\epsilon)N$. Then

$$\sum_{|\ell|\geq 2N+1} |\phi_j(\ell)|^2 \leq e^{-\frac{3\gamma}{2}N},$$

and

$$\|H_\Lambda\phi_j - \mu_j\phi_j\| \leq e^{-\frac{3\gamma}{4}N}. \tag{62}$$

Therefore, there exists $\tilde{j} \in \Lambda$ such that

$$|\mu_j - \tilde{\mu}_{\tilde{j}}^{\Lambda}| \le e^{-\frac{3\gamma}{4}N}.$$

Since $V \in S_{2N}$, one has that for any distinct j_1 and j_2 in Λ,

$$|\tilde{\mu}_{j_1}^{\Lambda} - \tilde{\mu}_{j_2}^{\Lambda}| \ge \frac{1}{(2N)^{q_1}},$$

and hence for any $m \neq \tilde{j}$,

$$|\tilde{\mu}_m^{\Lambda} - \mu_j| \ge \frac{1}{2^{q_1+1}N^{q_1}}. \tag{63}$$

Let $\tilde{\phi}_m$ and $\tilde{\mu}_m$ (as usual, for simplicity we have dropped the dependence on Λ from $\tilde{\phi}_m^{\Lambda}$ and $\tilde{\mu}_m^{\Lambda}$, $m \in \Lambda$), be the eigen-pairs of H_Λ. Let

$$I_\Lambda \phi_j = \sum_{m\in\Lambda} c_m \tilde{\phi}_m. \tag{64}$$

From (62) and (64), one has that

$$H_\Lambda \phi_j = \sum_{m\in\Lambda} \tilde{\mu}_m c_m \tilde{\phi}_m, \tag{65}$$

and

$$\| \sum_{m\in\Lambda} \tilde{\mu}_m c_m \tilde{\phi}_m - \sum_{m\in\Lambda} \mu_j c_m \tilde{\phi}_m \| = O(1)e^{-\frac{3\gamma}{4}N}. \tag{66}$$

By (63), (65) and (66), one has that for any $m \neq \tilde{j}$,

$$|c_m| \le O(1)N^{q_1}e^{-\frac{3\gamma}{4}N}.$$

Therefore, $1 - O(1)N^{3q_1}e^{-\frac{3\gamma}{2}N} \le c_{\tilde{j}}^2 \le 1$. We conclude that

$$\|\phi_j - \tilde{\phi}_{\tilde{j}}\| \le \|\phi_j - I_\Lambda\phi_j\| + \|I_\Lambda\phi_j - \tilde{\phi}_{\tilde{j}}\| \le e^{-\frac{\gamma}{2}N}.$$

□

We now state the conclusion:

Theorem 3.5. *For any $\epsilon > 0$, there exists l_ϵ such that the following statements hold. Fix any $L \ge \ell_\epsilon$ and $\beta_k \in \mathbb{Z}$, $k = 1, 2, \cdots, b$ satisfying $10\,L \le |\beta_k| \le L^3$ and $|\beta_k - \beta_{k'}| \ge 10L$, for any distinct $k, k' \in \{1, 2, \cdots, b\}$, there exists a subset of potentials X_ϵ with $\mathbb{P}(X_\epsilon) \ge 1-\epsilon$ and $\delta_0 > 0$ (depending on g, ϵ and L) such that for any $V \in X_\epsilon$ and $0 < \delta \le \delta_0$,*

(1)

$$|\phi_j(\ell)| \le C_\epsilon(1 + |\ell_j|)^q e^{-\gamma|\ell-\ell_j|}, \tag{67}$$

(2) for any $|\ell_j| \geq l_\epsilon$,

$$|\ell_j - j| \leq \epsilon |j|, \tag{68}$$

(3) for large N *(depending on* ϵ*),* $|j|, |j'| \leq N$ *and* $j \neq j'$,

$$|\mu_j - \mu_{j'}| \geq \frac{1}{2^{q_1+1} N^{q_1}}, \tag{69}$$

and

$$|\mu_j| \geq \frac{1}{2N^{q_1}}, \tag{70}$$

(4) for any eigenfunction ϕ_{α_k} *with* $\ell_{\alpha_k} \in B_k = \{\ell \in \mathbb{Z} : |\ell - \beta_k| \leq L\}$, $k = 1, 2, \cdots, b$, *we have that for any* $(n, j) \in [-e^{|\log\delta|^{\frac{3}{4}}}, e^{|\log\delta|^{\frac{3}{4}}}]^{b+1} \backslash \{(-e_k, \alpha_k)\}_{k=1}^{b}$,

$$|n \cdot \omega^{(0)} + \mu_j| \geq 2\delta^{\frac{1}{8}}, \tag{71}$$

where $\omega^{(0)} = (\omega_1^{(0)}, \cdots, \omega_b^{(0)}) = (\mu_{\alpha_1}, \cdots, \mu_{\alpha_b})$, *and for any* $(n, j) \in [-e^{|\log\delta|^{\frac{3}{4}}}, e^{|\log\delta|^{\frac{3}{4}}}]^{b+1} \backslash \{(e_k, \alpha_k)\}_{k=1}^{b}$,

$$|-n \cdot \omega^{(0)} + \mu_j| \geq 2\delta^{\frac{1}{8}}, \tag{72}$$

(5) for any $\theta \in \mathbb{R}$, *there are at most* b *vertices* $(n, j) \in [-|\log\delta|^s, |\log\delta|^s]^{b+1}$, *such that*

$$|(n \cdot \omega^{(0)} + \theta) + \mu_j| \leq \delta^{\frac{1}{8}}, \tag{73}$$

for any $\theta \in \mathbb{R}$, *there are at most* b *vertices* $(n, j) \in [-|\log\delta|^s, |\log\delta|^s]^{b+1}$, *such that*

$$|-(n \cdot \omega^{(0)} + \theta) + \mu_j| \leq \delta^{\frac{1}{8}}. \tag{74}$$

Proof. By Lemmas 2.2 (Wegner estimate), 2.3 (Minami estimate), 3.3, 3.4 and Borel-Cantelli type arugments,, we have (67)-(70).

We apply the Theorems (together with Remark 7) in the previous section with $\tilde{B}_k = B_k, k = 1, 2, \cdots, b$ and $\delta = 2^{-n}, n = 1, 2, \cdots$. Then by Borel-Cantelli type arugments, we have that there exists X_ϵ with $\mathbb{P}(X_\epsilon) \geq 1 - \epsilon$ and $X_\epsilon \cap (\tilde{S}_1 \cup S_2^s \cup \hat{S}_2^s) = \emptyset$ for any small δ.

Equations (71) and (72) follow from Lemma 3.4 with $N = \left\lfloor e^{|\log\delta|^{\frac{3}{4}}} \right\rfloor$.

The proof of (73) and (74) takes more time. Without loss of generality, we only prove (73). Assume that there are $(n^{(m)}, j^{(m)}) \in [-|\log\delta|^s, |\log\delta|^s]^{b+1}, m = 1, 2, \cdots, b+1$, satisfying

$$|(n^{(m)} \cdot \omega^{(0)} + \theta) + \mu_{j^{(m)}}| \leq \delta^{\frac{1}{8}}. \tag{75}$$

By Lemma 3.4, one has that there exist $\tilde{\omega}^{\Lambda_2^s}$ satisfying (10) and $(n^{(m)}, \tilde{j}^{(m)}) \in [-|\log\delta|^s, |\log\delta|^s]^b \times [-2|\log\delta|^s, 2|\log\delta|]^s$, $m = 1, 2, \cdots, b + 1$ such that

$$|(n^{(m)} \cdot \tilde{\omega}^{\Lambda_2^s} + \theta) + \tilde{\mu}_{\tilde{j}^{(m)}}^{\Lambda_2^s}| \leq 2\delta^{\frac{1}{8}}. \tag{76}$$

We can assume that $n^{(m)}$, $m = 1, 2, \cdots, b+1$ are distinct, otherwise Lemma 2.3 (Minami estimate) gives the proof.

Choose any $m_1, m_2 \in \{1, 2, \cdots, b+1\}$ such that (76) holds. When $|n_k^{(m_1)} - n_k^{(m_2)}| \geq 2$ for some $k \in \{1, 2, \cdots, b\}$ or supp $(n^{(m_1)} - n^{(m_2)}) \geq 3$, the proof follows from Theorem 2.5. When $\sum_{k=1}^{b}(n_k^{(m_1)} - n_k^{(m_2)}) \neq 0$ and $|n_k^{(m_1)} - n_k^{(m_2)}| \leq 1, k = 1, 2, \cdots, b$, the proof follows from Theorem 2.6.

So the only exceptional case is when for *all* $m_1, m_2 \in \{1, 2, \cdots, b+1\}$, $n^{(m_1)}$ and $n^{(m_2)}$ satisfy supp $(n^{(m_1)} - n^{(m_2)}) = 2$, $\sum_{k=1}^{b}(n_k^{(m_1)} - n_k^{(m_2)}) = 0$ and $n_k^{(m_1)} - n_k^{(m_2)} = \pm 1, 0, k = 1, 2, \cdots, b$. We will show that this is not possible. Shifting $n^{(m)}$ by $n^{(1)}$, one may assume that $n^{(1)} = (0, 0, \cdots, 0)$.

Consider $b \geq 3$ first. Without loss of generality, assume that $n^{(2)} = (1, -1, 0, \cdots, 0)$. Thus either $n_1^{(m)} = 1$ or $n_2^{(m)} = -1$ for all $m \in \{3, 4, \cdots, b+1\}$. Without loss of generality, assume that $n^{(3)} = (1, 0, -1, 0, 0, \cdots, 0)$. Therefore, for all $m \in \{2, 3, \cdots, b+1\}$, $n_1^{(m)} = 1$. This contradicts with $n^{(m)}$, $m \in \{1, 2, \cdots, b+1\}$ being distinct. The $b = 2$ can be proved in a similar fashion (indeed much easier).

□

4. The Large Deviation Theorem

In this section, we prepare for the nonlinear analysis in Sect. 5, by proving a large deviation theorem for the linearized operators. The assumptions below on $\tilde{H}$ are motivated (and will be shown to be satisfied) by such operators.

Assume that $\tilde{H}$ is an operator on $\ell^2(\mathbb{Z}^{b+1} \times \{0, 1\})$, Töplitz with respect to $n \in \mathbb{Z}^b$. We now write $\mathbb{Z}^{b+1}$ interchangeably with $\mathbb{Z}^b \times \mathbb{Z}$. Assume that there exist functions $h_{r,r'}(n, j, j')$, $r, r' \in \{0, 1\}$, on $\mathbb{Z}^b \times \mathbb{Z} \times \mathbb{Z}$, such that for any $u_r(n, j)$, $r \in \{0, 1\}$ and $(n, j) \in \mathbb{Z}^{b+1}$,

$$(\tilde{H}u)_r(n, j) := \sum_{(n',j') \in \mathbb{Z}^b \times \mathbb{Z},\, r' \in \{0,1\}} h_{r,r'}(n - n', j, j') u_{r'}(n', j'). \tag{77}$$

Assume that there exist $C_1 > 0$ and $c_1 > 0$ such that

$$|h_{r,r'}(n, j, j')| \leq C_1(1 + |n|)^{C_1} e^{-c_1(|n| + |j - j'|) - c_1 \max\{|j|, |j'|\}}. \tag{78}$$

Remark 9. The polynomial component in (78) can be integrated into the exponentially decaying term. However, as polynomial components naturally emerge during the iteration process (refer to Lemma 5.1), we find it more practical to retain these polynomial factors.

Let $D(\theta)$ be a family of operators from $\mathbb{R}$ to Op$[\ell^2(\mathbb{Z}^{b+1} \times \{0, 1\})]$:

$$D(\theta) = \begin{bmatrix} D+ & 0 \\ 0 & D_- \end{bmatrix}, \tag{79}$$

where $D_\pm = \operatorname{diag}(\pm(n \cdot \omega + \theta) + \mu_j)$, $(n, j) \in \mathbb{Z}^{b+1}$. Define $T = T(\theta)\colon \mathbb{R} \to$ Op$[\ell^2(\mathbb{Z}^{b+1} \times \{0, 1\})]$ as

$$T(\theta) = D(\theta) + \delta \tilde{H}. \tag{80}$$

Denote by Q_N an elementary region of size N centered at 0, which is one of the following regions,

$$Q_N = [-N, N]^{b+1}$$

or

$$Q_N = [-N, N]^{b+1} \setminus \{m \in \mathbb{Z}^{b+1} : m_i \varsigma_i 0, 1 \leq i \leq b+1\},$$

where for $i = 1, 2, \cdots, b+1$, $\varsigma_i \in \{<, >, \emptyset\}$ and at least two ς_i are not $\emptyset$. Here $m_i \emptyset 0$ means that no restriction is imposed on m_i.

Denote by $\mathcal{E}_N^0$ the set of all elementary regions of size N centered at 0. Let $\mathcal{E}_N$ be the set of all translates of elementary regions with center at 0, namely,

$$\mathcal{E}_N := \{m + Q_N : m \in \mathbb{Z}^{b+1}, Q_N \in \mathcal{E}_N^0\}.$$

For simplicity, we call elements in $\mathcal{E}_N$ elementary regions. Let $Q_N(j_0) = \{(n, j) \in \mathbb{Z}^b \times \mathbb{Z} : (n, j - j_0) \in Q_N\}$.

The width of a subset $\Lambda \subseteq \mathbb{Z}^{b+1}$, is defined to be the maximum of $M \in \mathbb{N}$ such that for any $m \in \Lambda$, there exists $\hat{M} \in \mathcal{E}_M$ such that

$$m \in \hat{M} \subset \Lambda$$

and

$$\operatorname{dist}(m, \Lambda \backslash \hat{M}) \geq M/2.$$

A generalized elementary region is defined to be a subset $\Lambda \subset \mathbb{Z}^{b+1}$ of the form

$$\Lambda := R \backslash (R + z),$$

where $z \in \mathbb{Z}^{b+1}$ is arbitrary and R is a rectangle,

$$R = \{(m_1, m_2, \cdots, m_{b+1}) \in \mathbb{Z}^{b+1} : |m_1 - m_1'| \leq M_1, \cdots, |m_{b+1} - m_{b+1}'| \leq M_{b+1}\}.$$

For $\Lambda \subset \mathbb{Z}^{b+1}$, we introduce its diameter,

$$\operatorname{diam}(\Lambda) = \sup_{m, m' \in \Lambda} |m - m'|.$$

Denote by $\mathcal{R}_N$ all generalized elementary regions with diameters less than or equal to N. Denote by $\mathcal{R}_N^M$ all generalized elementary regions in $\mathcal{R}_N$ with width larger than or equal to M.

With a slight abuse of notation, we also use $\mathcal{E}_N$, $\mathcal{E}_N^0$, Q_N, $Q_N(j_0)$, $\mathcal{R}_N$ and $\mathcal{R}_N^M$ to denote $\mathcal{E}_N \times \{0, 1\}$, $\mathcal{E}_N^0 \times \{0, 1\}$, $Q_N \times \{0, 1\}$, $Q_N(j_0) \times \{0, 1\}$, $\mathcal{R}_N \times \{0, 1\}$ and $\mathcal{R}_N^M \times \{0, 1\}$ respectively. Similarly for any $\Lambda \subset \mathbb{Z}^{b+1}$, denote by R_Λ the restriction to $\Lambda \times \{0, 1\}$.

We say that T (given by (80)) satisfies the large deviation theorem (LDT) at scale N with parameter $\tilde{c}_1 > 0$ if there exists a subset $\Theta_N \subset \mathbb{R}$ such that

$$\operatorname{Leb}(\Theta_N) \leq e^{-N^{\frac{1}{30}}},$$

and for any $j_0 \in [-2N, 2N]$, $Q_N \in \mathcal{E}_N^0$, and $\theta \notin \Theta_N$,

$$\|(R_{Q_N(j_0)} T(\theta) R_{Q_N(j_0)})^{-1}\| \le e^{N^{\frac{9}{10}}}, \tag{81}$$

and for any (n, j) and (n', j') satisfying $|n - n'| + |j - j'| \ge N^{\frac{1}{2}}$,

$$|(R_{Q_N(j_0)} T(\theta) R_{Q_N(j_0)})^{-1}(n, j; n', j')| \le e^{-\tilde{c}_1(|n-n'|+|j-j'|)}, \ \tilde{c}_1 > 0. \tag{82}$$

Let K_1 be a large constant depending only on b. Let $K = K_1^{100}$, $K_2 = K_1^5$.

Theorem 4.1. *Assume that ω satisfies*

(1) $|\omega_k - \mu_{\alpha_k}| \le C_2\delta$, $k = 1, 2, \cdots, b$;
(2) for any fixed $N \ge (\log \frac{1}{\delta})^K$,
and any $\tilde{N}$ with $(\log \frac{1}{\delta})^K \le \tilde{N} \le N$ and $0 \ne |n| \le 2\tilde{N}$,

$$|n \cdot \omega| \ge e^{-\tilde{N}^{\frac{1}{K_2}}}, \tag{83}$$

and for any $|j| \le 3\tilde{N}$, $|j'| \le 3\tilde{N}$, $|n| \le 2\tilde{N}$ with $(n, j - j') \ne 0$,

$$|n \cdot \omega - \mu_j + \mu_{j'}| \ge e^{-\tilde{N}^{\frac{1}{K_2}}}. \tag{84}$$

Then for small enough δ, the LDT holds at any scale $\tilde{N} \le N$ with a proper parameter $c_N > \frac{1}{2}c_1$.

Before proving the Theorem, let us note that the point-wise estimate (82) is required to hold at *sub-linear* scales in N, instead of (the usual) linear scales, as in e.g., [34,35]. This is in view of applications to the nonlinear analysis in sects. 5 and 6, where T will be given by the linearized operators, which are typically long range quasi-periodic operators in the n-direction. The range is, in turn, determined by the decay rate of the Green's functions. On the other hand, however, the range clearly gives an upper bound on the decay rate itself!

To free the analysis from this conundrum, He-Shi-Shi-Yuan were the first to propose in [36] that (82) hold at *sub-linear* scales and subsequently proved that this prevents the deterioration of the decay rate.[1] Henceforth we adopt this direct, modified approach to LDT. (The reader is welcome to consult appendix C for further technical comments.) We mention, however, Lemma D.1 in appendix D shows that the modified LDT could, in fact, be deduced from the standard one. Therefore proper usage of the latter could avoid the deterioration of constant as well (see appendix D).

Remark 10. The proof of Theorem 4.1 uses ideas from the work of Bourgain, Goldstein and Schlag [34], and combines the more recent, quantitative, work of Liu [35], with that of He-Shi-Shi-Yuan [36].

[1] We thank Ilya Kachkovskiy for a useful discussion concerning this point.

4.1. Preparations. The following perturbation argument will be used repeatedly during the proof.

Lemma 4.2. *Let $S \subset \mathbb{Z}^d$. Assume that A and B are two matrices, with entries $A(m, m')$ and $B(m, m')$, where m and $m' \in S$. Assume that $|B(m, m')| \le \epsilon_2 e^{-c|m-m'|}$, $|A^{-1}(m, m')| \le \epsilon_1^{-1} e^{-c|m-m'|}$ and $\|A^{-1}\| \le \epsilon_1^{-1}$, with ϵ_1, ϵ_2 and $c > 0$. Suppose that $|S|\epsilon_2\epsilon_1^{-1} \le \frac{1}{2}$. Then*

$$\|(A + B)^{-1}\| \le 2\epsilon_1^{-1}, \tag{85}$$

and

$$|(A + B)^{-1}(m, m') - A^{-1}(m, m')| \le 2|S|\epsilon_2\epsilon_1^{-2} e^{-c|m-m'|}. \tag{86}$$

Proof. Let $N = |S|$. Then $\|B\| \le \sqrt{N}\epsilon_2$ and

$$\|BA^{-1}\| \le \frac{1}{2}.$$

By Neumann series expansion, we have

$$(A + B)^{-1} = A^{-1} \sum_{s \ge 0} (-BA^{-1})^s. \tag{87}$$

Thus one has

$$\|(A + B)^{-1}\| \le \|A^{-1}\| \frac{1}{1 - \|BA^{-1}\|} \le 2\epsilon_1^{-1}, \tag{88}$$

and

$$\begin{aligned}|(A + B)^{-1}(m, m') - A^{-1}(m, m')| &\le \|A^{-1}\| \sum_{\substack{s \ge 1\\ k_i \in S}} (\epsilon_2\epsilon_1^{-1})^s e^{-c|m-k_1|-c|k_1-k_2|-\cdots-c|k_s-m'|}\\ &\le \epsilon_1^{-1} e^{-c|m-m'|} \sum_{s \ge 1} (N\epsilon_2\epsilon_1^{-1})^s\\ &\le 2N\epsilon_2\epsilon_1^{-2} e^{-c|m-m'|}.\end{aligned}$$

□

Lemma 4.3. [31, Prop. 14.1] *[35, Lemma 5.1] Let $T(x)$ be a $N \times N$ matrix function of a parameter $x \in [-\tau, \tau]$ satisfying the following conditions:*

(i) $T(x)$ is real analytic in $x \in [-\tau, \tau]$ and has a holomorphic extension to

$$\mathcal{D}_{\tau,\tau_1} = \{z : |\Re z| \le \tau, |\Im z| \le \tau_1\}$$

satisfying

$$\sup_{z \in \mathcal{D}_{\tau,\tau_1}} \|T(z)\| \le B_1, B_1 \ge 1. \tag{89}$$

(ii) For all $x \in [-\tau, \tau]$, there is a subset $\Lambda \subset [1, N]$ with

$$|\Lambda| \le M,$$

and

$$\|(R_{[1,N]\setminus\Lambda} T(x) R_{[1,N]\setminus\Lambda})^{-1}\| \le B_2,\ B_2 \ge 1. \tag{90}$$

(iii)

$$\mathrm{Leb}\{x \in [-\tau, \tau] : \ \|T^{-1}(x)\| \ge B_3\} \le 10^{-3}\tau_1(1+B_1)^{-1}(1+B_2)^{-1}. \tag{91}$$

Let

$$0 < \epsilon \le (1 + B_1 + B_2)^{-10M}. \tag{92}$$

Then

$$\mathrm{Leb}\left\{x \in [-\tau/2, \tau/2] : \ \|T^{-1}(x)\| \ge \epsilon^{-1}\right\} \le C\tau e^{-c\left(\frac{\log \epsilon^{-1}}{M \log(B_1+B_2+B_3)}\right)}, \tag{93}$$

where C and c are absolute constants.

To apply Lemma 4.3, we also need to introduce semi-algebraic sets. A set $\mathcal{S} \subset \mathbb{R}^d$ is called *semi-algebraic* if it is a finite union of sets defined by a finite number of polynomial equalities and inequalities. More precisely, let $\{P_1, \cdots, P_s\} \subset \mathbb{R}[x_1, \cdots, x_d]$ be a family of real polynomials whose degrees are bounded by κ. A (closed) semi-algebraic set $\mathcal{S}$ is given by an expression

$$\mathcal{S} = \bigcup_{l} \bigcap_{\ell \in \mathcal{L}_l} \left\{x \in \mathbb{R}^d : \ P_\ell(x) \varsigma_{l\ell} 0\right\}, \tag{94}$$

where $\mathcal{L}_l \subset \{1, \cdots, s\}$ and $\varsigma_{l\ell} \in \{\ge, \le, =\}$. Then we say that $\mathcal{S}$ has degree at most $s\kappa$. In fact, the degree of $\mathcal{S}$ which is denoted by $\deg(\mathcal{S})$, means the smallest $s\kappa$ over all representations as in (94).

Following are some basic properties of these sets. They are special cases of that in [37], and restated in [31].

Lemma 4.4. [31, Theorem 9.3] *[37, Theorem 1] Let $\mathcal{S} \subset [0, 1]^d$ be a semi-algebraic set of degree B. Then the number of connected components of $\mathcal{S}$ does not exceed $(1+B)^{C(d)}$.*

Lemma 4.5. [31, Proposition 9.2] *Let $\mathcal{S} \subset [0, 1]^{d_1+d_2}$ be a semi-algebraic set of degree B. Let $(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then the projection $\mathrm{proj}_{x_1}(S)$ is a semi-algebraic set of degree at most $(1 + B)^{C(d_1,d_2)}$.*

4.2. Large deviation theorem for small scales. Below we prove Theorem 4.1 for small scales, namely $N \le (\log \frac{1}{\delta})^{10}$.

Proof. In this case, let

$$\tilde{\Theta}_N = \{\theta \in \mathbb{R} : \text{ there exists } (n, j) \in [-N, N]^b \times [-3N, 3N] \text{ such that either}$$
$$|(n \cdot \omega + \theta) + \mu_j| \le 2e^{-N^{\frac{1}{20}}} \text{ or } |(-n \cdot \omega - \theta) + \mu_j| \le 2e^{-N^{\frac{1}{20}}}\}.$$

Clearly, $\mathrm{Leb}(\tilde{\Theta}_N) \le N^{C(b)} e^{-N^{\frac{1}{20}}}$.

When $N \le (\log \frac{1}{\delta})^{10}$, δ is much smaller than $e^{-N^{\frac{1}{20}}}$. Now (81) and (82) follow from standard perturbation arguments (Lemma 4.2).

□

4.3. Large deviation theorem for intermediate scales. In this section, we will prove Theorem 4.1 for intermediate scales, namely $(\log \frac{1}{\delta})^{10} \leq N \leq (\log \frac{1}{\delta})^{K}$.

Proof. Let Θ_N be such that at least one of (81) and (82) does not hold for $\theta \in \Theta_N$. Choose any $N \in [(\log \frac{1}{\delta})^{10}, (\log \frac{1}{\delta})^{K}]$.

Assume that (81) and (82) do not hold for some θ. Then we must have for some $(n, j) \in [-N, N]^b \times [-3N, 3N]$, either $|\theta+n\cdot\omega^{(0)}+\mu_j| \leq \delta^{\frac{3}{8}}$ or $|\theta+n\cdot\omega^{(0)}-\mu_j| \leq \delta^{\frac{3}{8}}$. Otherwise, standard perturbation arguments yield that for any j_0 with $|j_0| \leq 2N$,

$$\|(R_{Q_N(j_0)}TR_{Q_N(j_0)})^{-1}\|| \leq 2\delta^{-\frac{3}{8}} \leq e^{N^{\frac{9}{10}}},$$

and for any (n, j) and (n', j') such that $|n-n'|+|j-j'| \geq \sqrt{N}$,

$$|(R_{Q_N(j_0)}TR_{Q_N(j_0)})^{-1}(n, j; n', j')| \leq e^{-c_1(|n-n'|+|j-j'|)}.$$

Therefore, we can restrict θ to be in $10^{b+1}N^{b+1}$ intervals of size $2\delta^{\frac{3}{8}}$. Denote all the intervals by $\{I_i\}$ and take one of them, I_0, into consideration. Without loss of generality, assume that I_0 comes from the + sector, namely

$$I_0 = \{\theta : |\theta + n_0 \cdot \omega^{(0)} + \mu_{j_0}| \leq \delta^{\frac{3}{8}} \text{ and } (n_0, j_0) \in [-N, N]^b \times [-3N, 3N]\}.$$

For the $-$ sector $I_0 = \{\theta : |\theta + n_0 \cdot \omega^{(0)} - \mu_{j_0}| \leq \delta^{\frac{3}{8}}$ and $(n_0, j_0) \in [-N, N]^b \times [-3N, 3N]\}$, the proof is similar.

Let

$$\mathcal{A}_1^{\theta} = \{(n, j) \in [-N, N]^b \times [-3N, 3N] : |\theta + n \cdot \omega^{(0)} + \mu_j| \leq \delta^{\frac{1}{8}}\},$$

and

$$\mathcal{A}_2^{\theta} = \{(n, j) \in [-N, N]^b \times [-3N, 3N] : |\theta + n \cdot \omega^{(0)} - \mu_j| \leq \delta^{\frac{1}{8}}\}.$$

By (5) of Theorem 3.5, one has that for any θ,

$$\#\mathcal{A}_1^{\theta} \leq b, \#\mathcal{A}_2^{\theta} \leq b. \tag{95}$$

Since the size of I_0 is $2\delta^{\frac{3}{8}}$, we have that there exist $\mathcal{A}_1$ and $\mathcal{A}_2$ independent of $\theta \in I_0$ such that

$$\#\mathcal{A}_1 \leq b, \#\mathcal{A}_2 \leq b, \tag{96}$$

and for any $(n, j) \in [-N, N]^b \times [-3N, 3N]\backslash(\mathcal{A}_1 \cup \mathcal{A}_2)$ and $\theta \in I_0$

$$|\theta + n \cdot \omega^{(0)} \pm \mu_j| \geq \frac{1}{2}\delta^{\frac{1}{8}}. \tag{97}$$

Take any $\tilde{\Lambda} \in R_{N_1}^{\sqrt{N}}$ with $N_1 \in [\sqrt{N}, 6N]$ and $\tilde{\Lambda} \subset [-N, N]^b \times [-3N, 3N]$. By perturbation arguments (Lemma 4.2), we have that for any $\theta \in I_0$,

$$\|(R_{\tilde{\Lambda}\backslash(\mathcal{A}_1\cup\mathcal{A}_2)}T(\theta)R_{\tilde{\Lambda}\backslash(\mathcal{A}_1\cup\mathcal{A}_2)})^{-1}\| \leq 3\delta^{-\frac{1}{8}}. \tag{98}$$

We are going to apply Cartan's estimate, Lemma 4.3. For this reason, let $\tau = \delta^{\frac{3}{8}}$, $\tau_1 = 1$, $\Lambda = \mathcal{A}_1 \cup \mathcal{A}_2$, $M = 2b$, $B_1 = O(1)(\log \frac{1}{\delta})^K$, $B_2 = 3\delta^{-\frac{1}{8}}$, $B_3 = 1$ and $\epsilon = e^{-N_1{}^{3/4}}$. We note that since I_0 has size $\delta^{\frac{3}{8}}$, (91) holds automatically.

Applying Cartan's estimate (Lemma 4.3) in all possible $\tilde{\Lambda} \in R_{N_1}^{\sqrt{N}}$ (in total N^C), there exists a subset $\tilde{\Theta}_{N_1} \subset \mathbb{R}$ such that

$$\mathrm{Leb}(\tilde{\Theta}_{N_1}) \leq e^{-\frac{N_1{}^{\frac{3}{4}}}{|\log\delta|^2}},$$

and for any $\theta \notin \tilde{\Theta}_{N_1}$ and any $\tilde{\Lambda} \in R_{N_1}^{\sqrt{N}}$ with $\tilde{\Lambda} \subset [-N, N]^b \times [-3N, 3N]$,

$$\|(R_{\tilde{\Lambda}} T R_{\tilde{\Lambda}})^{-1}\| \leq e^{N_1^{\frac{3}{4}}}. \tag{99}$$

Let $N_0 = \sqrt{N}$. We call a box $(n_1, j_1) + Q_{N_0} \in \mathcal{E}_{N_0}$, $(n_1, j_1) \in [-N, N]^b \times [-3N, 3N]$ good if

$$\|(R_{(n_1,j_1)+Q_{N_0}} T R_{(n_1,j_1)+Q_{N_0}})^{-1}\| \leq e^{N_0^{\frac{9}{10}}}$$

and for any (n, j) and (n', j') such that $|n - n'| + |j - j'| \geq \sqrt{N_0}$,

$$|(R_{(n_1,j_1)+Q_{N_0}} T R_{(n_1,j_1)+Q_{N_0}})^{-1}(n, j; n', j')| \leq e^{-c_1(|n-n'|+|j-j'|)}.$$

Otherwise, we call $(n_1, j_1) + Q_{N_0} \in [-N, N]^b \times [-3N, 3N]$ bad. By (96), (97) and perturbation arguments (Lemma 4.2), we have that there are at most $2b$ disjoint bad boxes of size $N_0 = N^{1/2}$ contained in $[-N, N]^b \times [-3N, 3N]$.

We have sublinear bound and (99). By [35, Theorem 2.1] and (2) in Remark 11 (see also Appendix A in [36]), for any $\theta \notin \bigcup_{\{I_i\}} \bigcup_{N_1 \in [\sqrt{N}, 6N]} \tilde{\Theta}_{N_1}$, (81) and (82) hold for the scale N. Therefore,

$$\Theta_N \subset \bigcup_{\{I_i\}} \bigcup_{N_1 \in [\sqrt{N}, 6N]} \tilde{\Theta}_{N_1}$$

and hence

$$\mathrm{Leb}(\Theta_N) \leq e^{-N^{\frac{1}{10}}}.$$

□

4.4. Large deviation theorem for large scales. In this section, we are going to prove Theorem 4.1 for large scales, namely $N \geq (\log \frac{1}{\delta})^K$.

Proof. Let $N_2 = N_1^{K_1}$ and $N_3 \in [N_2^{K_1}, N_2^{2K_1}]$. Assume that $N_3 \geq (\log \frac{1}{\delta})^K$ and that the LDT holds at both scales N_1 and N_2 with parameter $\tilde{c}_1$.

We will show that there are at most N_1^C bad disjoint boxes of size N_1 contained in $[-N_3, N_3]^b \times [-3N_3, 3N_3]$. Let $(n_1, l_1) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ be such that $(n_1, l_1) + Q_{N_1}$ is bad for some $Q_{N_1} \in \mathcal{E}_{N_1}^0$.

We first bound the case when $|l_1| \le 2N_1$. By the LDT at scale N_1, there exists a set Θ_{N_1} with $\mathrm{Leb}(\Theta_{N_1}) \le e^{-N_1^{1/30}}$ such that for any $\theta \notin \Theta_{N_1}$ and any $Q_{N_1} \in \mathcal{E}^0_{N_1}$, Q_{N_1} is good. Since the operator is Töplitz with respect to $n \in \mathbb{Z}^b$, one has that for any (n_1, l_1) with $\theta + n_1 \cdot \omega \notin \Theta_{N_1}$, $(n_1, l_1) + Q_{N_1}$ is good for any $Q_{N_1} \in \mathcal{E}^0_{N_1}$. By standard arguments, we can assume that Θ_{N_1} is a semi-algebraic set of degree at most N_1^C, namely, there exist N_1^C intervals I_i of size $e^{-N_1^{1/30}}$, such that $\Theta_{N_1} \subset \cup_i I_i$. The assumption on ω indicates that, for any nonzero n with $|n| \le 2N_3$,

$$|n \cdot \omega| \ge e^{-(2N_3)^{\frac{1}{K_2}}} \ge e^{-N_1^{\frac{1}{30}}}. \tag{100}$$

Therefore, for any $|l_1| \le 2N_1$, there is at most one bad box (n_1, l_1) such that $n_1 \cdot \omega \in I_i$. This leads to at most N_1^C bad boxes in this case.

When $|l_1| \ge 2N_1$, we will show that there are at most *two* disjoint bad boxes of size N_1. First, if a box $(n, j) + Q_{N_1}$ is bad, by (78) and perturbation arguments (Lemma 4.2), we must have that for some $(n_1, l_1) \in (n, j) + Q_{N_1}$, either

$$|\theta + n_1 \cdot \omega + \mu_{l_1}| \le 2e^{-N_1^{9/10}} \tag{101}$$

or

$$|\theta + n_1 \cdot \omega - \mu_{l_1}| \le 2e^{-N_1^{9/10}}. \tag{102}$$

Assume that indeed there are three bad boxes. We have that there are two from D_+, namely (101) (or D_-, namely (102)). Therefore, we have that for two distinct vertices $(n, j) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ and $(n', j') \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$,

$$|m \cdot \omega - \mu_j + \mu_{j'}| \le 4e^{-N_1^{9/10}}, m = n - n'.$$

This contradicts the assumption (84).

Let $\tilde{\Theta}_{N_2} \subset \mathbb{R}$ be such that for some $(n, j) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ such that either $|\theta + n \cdot \omega + \mu_j| \le 2e^{-N_2^{9/10}}$ or $|\theta + n \cdot \omega - \mu_j| \le 2e^{-N_2^{9/10}}$. Since for any $|l_1| \ge 2N_2$, the matrix is essentially diagonal, we have that for any $\theta \notin \tilde{\Theta}_{N_2}$, $(n_1, l_1) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ with $|l_1| \ge 2N_2$ and $Q_{N_2} \in \mathcal{E}^0_{N_2}$, $(n_1, l_1) + Q_{N_2}$ is good. Let $\hat{\Theta}_{N_2} = \{\theta :$ for some $n \in [-N_3, N_3]^b, \theta + n \cdot \omega \in \Theta_{N_2}\}$. Therefore, for any $\theta \notin \hat{\Theta}_{N_2}$, $(n_1, l_1) \in [-N_3, N_3]^b \times [-2N_2, 2N_2]$ and $Q_{N_2} \in \mathcal{E}^0_{N_2}$, $(n_1, l_1) + Q_{N_2}$ is good. Clearly

$$\mathrm{Leb}(\tilde{\Theta}_{N_2} \cap \hat{\Theta}_{N_2}) \le e^{-N_2^{1/31}},$$

and for any $\theta \notin \tilde{\Theta}_{N_2} \cap \hat{\Theta}_{N_2}$, $(n_1, l_1) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ and $Q_{N_2} \in \mathcal{E}^0_{N_2}$, $(n_1, l_1) + Q_{N_2}$ is good.

Applying Lemma 4.3 (see proof of [35, Theorem 2.2] for details), for any N_3, there exists a subset $\tilde{\Theta}_{N_3} \subset \mathbb{R}$ such that

$$\mathrm{Leb}(\tilde{\Theta}_{N_3}) \le e^{-N_3{}^{\frac{1}{4}}},$$

and for any $N \in [N_3^{1/2}, N_3]$, $\tilde{\Lambda} \in R_{6N}^{N_3^{1/2}}$ with $\tilde{\Lambda} \subset [-N_3, N_3]^b \times [-3N_3, 3N_3]$, and for any $\theta \notin \tilde{\Theta}_{N_3}$,

$$\|(R_{\tilde{\Lambda}} T R_{\tilde{\Lambda}})^{-1}\| \le e^{N^{\frac{3}{4}}}. \tag{103}$$

Let $N_0 = {N_3}^{1/2}$. We call a box $(n_1, l_1) + Q_{N_0} \in \mathcal{E}_{N_0}^0$, $(n_1, l_1) \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ good if

$$\|(R_{(n_1,l_1)+Q_{N_0}} T R_{(n_1,l_1)+Q_{N_0}})^{-1}\| \le e^{N_0^{\frac{9}{10}}},$$

and for any (n, j) and (n', j') such that $|n - n'| + |j - j'| \ge \sqrt{N}_0$,

$$|(R_{(n_1,l_1)+Q_{N_0}} T R_{(n_1,l_1)+Q_{N_0}})^{-1}(n, j; n', j')| \le e^{-c_2(|n-n'|+|j-j'|)},$$

where $c_2 = \tilde{c}_1 - N_1^{-\kappa}$ with a proper $\kappa > 0$. Otherwise, we call $(n_1, l_1) + Q_{N_0} \in [-N_3, N_3]^b \times [-3N_3, 3N_3]$ bad. Since there are at most N_1^C bad boxes of size N_1, by resolvent identity, we have that there are at most N_1^C disjoint bad boxes of size $N_0 = N_3^{1/2}$ contained in $[-N_3, N_3]^b \times [-3N_3, 3N_3]$.

We have achieved the sublinear bound and (103). By [35, Theorem 2.1] (see Remark C), we have that the LDT holds for scale N_3 with parameter $\tilde{c}_1 - N_1^{-\kappa}$, where κ is a proper small positive constant. Now the proof follows from standard inductions (e.g. [38, Section 4]). □

We close the section by a remark, to benefit the upcoming nonlinear analysis:

Remark 11. (1) One may choose $c_N = \frac{9}{10}c_1$ for $N \le (\log \frac{1}{\delta})^{10}$ and a sequence γ_N, $N \ge (\log \frac{1}{\delta})^{10} + 1$ (let $k_0 = (\log \frac{1}{\delta})^K + 1$) with $\sum_{j \ge k_0} \gamma_j < \frac{c_1}{10}$ such that $c_N = c_1 - \sum_{j=k_0}^{N} \gamma_j$.

(2) From (83) and (84), in order to have LDT at all scales, we only need to remove a set of measure (with respect to ω) less than $e^{-\frac{1}{2}(\log \frac{1}{\delta})^{\frac{K}{K_2}}} \le e^{-(\log \frac{1}{\delta})^{K_1^{90}}} \ll \delta$.

5. The Nonlinear Analysis

Fix $V \in X_\epsilon$, so that the conclusions of Theorem 3.5 hold and Theorem 4.1 is available. (As before, we omit the superscript V, as it is fixed.) We are now ready to solve the nonlinear matrix Eq. (6) using a Lyapunov-Schmidt decomposition.

5.1. Lyapunov-Schmidt decomposition. To simplify notations, we write u for $\hat{u}$, namely $u(n, j) = \hat{u}(n, j)$. Let v be the complex conjugate of u, more precisely, $v(n, j) = \overline{\hat{u}(-n, j)}$. From (8), W_u is a vector on $\ell^2(\mathbb{Z}^{b+1})$, which is now given by

$$W_u(n, j) = \sum_{\substack{n' + \sum_{m=1}^{p}(n_m + n'_m) = n \\ n', n_m, n'_m \in \mathbb{Z}^b}} \sum_{l_m, l'_m, j' \in \mathbb{Z}} u(n', j') \prod_{m=1}^{p} u(n_m, l_m) v(n'_m, l'_m)$$
$$\left(\sum_{x \in \mathbb{Z}} \phi_j(x) \phi_{j'}(x) \prod_{m=1}^{p} \phi_{l_m}(x) \phi_{l'_m}(x) \right). \tag{104}$$

Let $\widetilde{W}_u$ be a vector on $\ell^2(\mathbb{Z}^{b+1})$, which is given by

$$\widetilde{W}_u(n,j) = \sum_{\substack{n'+\sum_{m=1}^{p}(n_m+n'_m)=n\\ n',n_m,n'_m\in\mathbb{Z}^b}} \sum_{j',l_m,l'_m\in\mathbb{Z}} v(n',j')\prod_{m=1}^{p} v(n_m,l_m)u(n'_m,l'_m)$$
$$\left(\sum_{x\in\mathbb{Z}}\phi_j(x)\phi_{j'}(x)\prod_{m=1}^{p}\phi_{l_m}(x)\phi_{l'_m}(x)\right). \tag{105}$$

We remark that $\widetilde{W}_u$ and W_u are functions of u and v. We only indicate the dependence on u for simplicity and the fact that v is the conjugate of u.

Writing the equation for v as well, leads to the system of nonlinear equations on $\mathbb{Z}^{b+1}\times\{0,1\}$:

$$\begin{aligned}(D_+u)(n,j)+\delta W_u(n,j)&=0,\\ (D_-u)(n,j)+\delta\widetilde{W}_u(n,j)&=0,\end{aligned} \tag{106}$$

where $D_\pm$ are the diagonal matrices with entries

$$D_\pm(n,j) := D_\pm(n,j;n,j) = \pm n\cdot\omega+\mu_j. \tag{107}$$

Define

$$S=\{(-e_k,\alpha_k)\times\{0\},(e_k,\alpha_k)\times\{1\},\ k=1,2,\cdots,b\}, \tag{108}$$

and denote the complement by S^c:

$$S^c=\mathbb{Z}^{b+1}\times\{0,1\}\backslash S. \tag{109}$$

Write D for the diagonal matrix composed of the diagonal blocks $D_\pm$ and write (106) in the form $F(u,v)=0$. Since v is the conjugate of u, we simply write $F(u)=0$ for $F(u,v)=0$. We make a Lyapunov-Schmidt decomposition of (106) into the P-equations:

$$F(u)|_{S^c}=0. \tag{110}$$

and the Q-equations:

$$F(u)|_S=0. \tag{111}$$

5.2. The P-equations. The P-equations are infinite dimensional. They are solved using a Newton scheme, starting from the initial approximation $u^{(0)}=u_0$. The recent paper [39] contains a detailed, step by step account of the resolution of the P-equations. We refer the reader to [39] for more in depth reading. Below we give a brief account of the P-equations in this paper.

Let F' be the linearized operator on $\ell^2(\mathbb{Z}^{b+1})\times\{0,1\}$,

$$F'(u)=D+\delta\mathcal{W}_u,$$

where

$$\mathcal{W}_u=\begin{pmatrix}\frac{\partial W_u}{\partial u} & \frac{\partial W_u}{\partial v}\\ \frac{\partial W_u}{\partial u} & \frac{\partial W_u}{\partial v}\end{pmatrix}.$$

It is easy to see that $\mathcal{W} = \mathcal{W}_u$ is Töplitz with respect to $n \in \mathbb{Z}^b$, namely for any $j \in \mathbb{Z}$, $j' \in \mathbb{Z}$, $k \in \mathbb{Z}^b$, $n \in \mathbb{Z}^b$, $n' \in \mathbb{Z}^b$, $r \in \{0, 1\}$, $r' \in \{0, 1\}$

$$\mathcal{W}_{r,r'}(n, j; n', j') = \mathcal{W}_{r,r'}(n + k, j; n' + k, j').$$

The operator $\mathcal{W}_{r,r'}(n, j; n', j')$ plays the role of $\tilde{H}$ in (77). In the lemma below, we will show that it satisfies (78). This will then enable us to use the large deviation theorem, Theorem 4.1.

Lemma 5.1. *Assume that* $|u(n, j)| \leq e^{-c_1|n|-c_2|j|}$ *and* $0 < c_2 \ll \gamma$. *We have that for any* $\epsilon > 0$,

$$|\mathcal{W}_{r,r'}(n, j; n', j')| \leq C(|n - n'| + 1)^C e^{-c_1|n-n'|-\frac{\gamma}{2}|j-j'|-(2pc_2-\epsilon)\max\{|j|,|j'|\}}, C > 1.$$

Proof. From the definition, one has that

$$|\mathcal{W}_{r,r'}(n, j; n', j')| \leq \sum_{\substack{\sum_{m=1}^p (n_m+n'_m)=n-n' \\ n_m,n'_m\in\mathbb{Z}^b}} \sum_{l_m,l'_m\in\mathbb{Z}} e^{-\sum_{m=1}^p (c_1|n_m|+c_2|l_m|+c_1|n'_m|+c_2|l'_m|)} \tag{112}$$

$$|\sum_{x\in\mathbb{Z}} \phi_j(x)\phi_{j'}(x) \prod_{m=1}^p \phi_{l_m}(x)\phi_{l'_m}(x)| = AB, \tag{113}$$

where

$$A = \sum_{\substack{\sum_{m=1}^p (n_m+n'_m)=n-n' \\ n_m,n'_m\in\mathbb{Z}^b}} e^{-c_1\sum_{m=1}^p (|n_m|+|n'_m|)}$$

and

$$B = \sum_{l_m,l'_m\in\mathbb{Z}} e^{-c_2\sum_{m=1}^p (|l_m|+|l'_m|)} |\sum_{x\in\mathbb{Z}} \phi_j(x)\phi_{j'}(x) \prod_{m=1}^p \phi_{l_m}(x)\phi_{l'_m}(x)|.$$

Direct computation implies that

$$A \leq C(|n - n'| + 1)^C e^{-c_1|n-n'|}, \ C > 1.$$

We are in a position to estimate B.

By (59), one has that

$$B \leq \sum_{l_m,l'_m,x\in\mathbb{Z}} C(1 + |l_m|)^q (1 + |l'_m|)^q (1 + |j|)^q (1 + |j'|)^q$$
$$e^{-\gamma|x-\ell_j|-\gamma|x-\ell_{j'}|-\sum_{m=1}^p (c_2|l_m|+c_2|l'_m|+\gamma|x-\ell_{l_m}|+\gamma|x-\ell_{l'_m}|)}; \tag{114}$$

and by (60),

$$\sum_{l\in\mathbb{Z}} (1 + |l|)^q e^{-c_2|l|-\gamma|x-\ell_l|} \leq \sum_{l\in\mathbb{Z}} (1 + |l|)^q e^{-(c_2-\epsilon)|l|-\gamma|x-l|}$$
$$\leq C(1 + |x|)^C e^{-(c_2-\epsilon)|x|}. \tag{115}$$

From (115), (114) and (60), we therefore have (assume $|j| = \max\{|j|, |j'|\}$) that

$$\begin{aligned}
B &\le C\sum_{x\in\mathbb{Z}}(1+|j|)^C(1+|j'|)^C(1+|x|)^C e^{-(2pc_2-\epsilon)|x|-\gamma|x-\ell_j|-\gamma|x-\ell_{j'}|}\\
&\le C\sum_{x\in\mathbb{Z}}(1+|x|)^C e^{-(2pc_2-\epsilon)|x|-\gamma|x-j|-\gamma|x-j'|+\epsilon|j|+\epsilon|j'|}\\
&\le C\sum_{x\in\mathbb{Z}}(1+|x|)^C e^{-(2pc_2-\epsilon)|x|-\frac{\gamma}{2}|x-j|-\frac{\gamma}{2}|x-j'|+\epsilon|j|+\epsilon|j'|-\frac{\gamma}{2}|j-j'|}\\
&\le C\sum_{x\in\mathbb{Z}}(1+|x|)^C e^{-(2pc_2-\epsilon)|x|-\frac{\gamma}{2}|x-j|+\epsilon|j|-\frac{\gamma}{2}|j-j'|}\\
&\le Ce^{-(2pc_2-\epsilon)|j|-\frac{\gamma}{2}|j-j'|}
\end{aligned}$$

This finishes the proof. □

The operator F' is to be evaluated near $\omega = \omega^{(0)} = (\mu_{\alpha_1}, \mu_{\alpha_2}, \cdots, \mu_{\alpha_b})$, the linear frequency, and $u = u^{(0)}$ and $v = v^{(0)}$. As earlier, we have that $u^{(0)}(-e_k, \alpha_k) = a_k$, $k = 1, 2, \cdots, b$; $u^{(0)}(n, j) = 0$, otherwise.

Recall next the formal Newton scheme:

$$\Delta_{\rm cor}\begin{pmatrix}u\\ v\end{pmatrix} = -[F'_{S^c}(u)]^{-1}F(u)|_{S^c},$$

where the left side denotes the correction to $\begin{pmatrix}u\\ v\end{pmatrix}$, $F'_{S^c}(u)$ is the linearized operator evaluated at (u, v): $F'(u)$, and restricted to S^c:

$$F'_{S^c}(u)(x, y) = F'(u)(x, y),$$

for $x, y \in S^c$; likewise $F(u)|_{S^c}$ is $F(u)$ restricted to S^c:

$$[F(u)|_{S^c}](x) = F(u)(x),$$

for $x \in S^c$.

Since we seek solutions close to $(u^{(0)}, v^{(0)})$, which has compact support in $\mathbb{Z}^{b+1} \times \{0, 1\}$, we adopt a *multiscale* Newton scheme as follows:
At iteration step $(r + 1)$, choose an appropriate scale $N = M^r$ (M is a large constant which will be determined later) and estimate $[F'_N]^{-1}$, where F'_N is F' restricted to

$$[-N, N]^{b+1} \times \{0, 1\}\backslash S \subset \mathbb{Z}^{b+1} \times \{0, 1\},$$

and evaluated at $u^{(r)}$ and $v^{(r)}$: $F'_N = F'_N(u^{(r)})$. Define the $(r + 1)$-th correction to be:

$$\Delta_{\rm cor}\begin{pmatrix}u^{(r+1)}\\ v^{(r+1)}\end{pmatrix} = -[F'_N(u^{(r)})]^{-1}F_N(u^{(r)}),\ N = M^r,$$

where F_N is F restricted to

$$[-N, N]^{b+1} \times \{0, 1\}\backslash S \subset \mathbb{Z}^{b+1} \times \{0, 1\},$$

and

$$\begin{aligned} u^{(r+1)} &= u^{(r)} + \Delta_{\text{cor}} u^{(r+1)}, \\ v^{(r+1)} &= v^{(r)} + \Delta_{\text{cor}} v^{(r+1)}, \end{aligned}$$

for all $r = 0, 1, 2, \cdots$.

At step r, $\mathcal{W}_{u^{(r)}(\omega,a)}$ (depending on $u^{(r)}(\omega, a)$) is a function of ω and a. We write $T_{u^{(r)}}(\theta, \omega, a)$ for the operator $F' = D(\theta) + \delta\mathcal{W}_{u^{(r)}(\omega,a)}$, and $\tilde{T}(\theta, \omega, a)$ the operator $F' = D(\theta) + \delta\mathcal{W}_{u^{(r)}(\omega,a)}$, restricted to $\mathbb{Z}^{b+1} \times \{0, 1\}\backslash S$, where

$$D(\theta) = \begin{bmatrix} \text{diag } (n \cdot \omega + \theta + \mu_j) & 0 \\ 0 & \text{diag } (-n \cdot \omega - \theta + \mu_j) \end{bmatrix}, (n, j) \in \mathbb{Z}^{b+1}. \quad (116)$$

For simplicity, write $\tilde{T}_{u^{(r)}}(\omega, a)$ for $\tilde{T}_{u^{(r)}}(0, \omega, a)$ and $T_{u^{(r)}}(\omega, a)$ for $T_{u^{(r)}}(0, \omega, a)$.

The analysis of the linearized operators F'_N uses Theorem 3.5 for small scales; for large scales, it also uses Theorem 4.1 and semi-algebraic projection to convert estimates in θ into that of ω, and finally a.

5.3. The Q-equations. The Q-equations are $2b$ dimensional, but due to symmetry leading to b equations only. They are used to relate ω with a. Recall that the P-equations also depend on a from the linearized nonlinear term. So both the P and the Q-equations depend on a. Consequently, they are solved together, consecutively, and *not* independently. This is different from that in [39], and is a general feature when the parameters are extracted from the nonlinear term.

To solve the Q-equations, we fix the amplitudes $u(n, j)$ on S, i.e., we fix $u(-e_{\alpha_k}, \alpha_k) = a_k$, $k = 1, 2, \cdots, b$, and the same for the complex conjugate. These b equations are then seen as equations for the frequencies instead, and we have

$$\omega_k = \mu_{\alpha_k} + \delta \frac{W_u(-e_k, \alpha_k)}{a_k}, k = 1, 2, \cdots, b. \quad (117)$$

When $u = u^{(0)}$, let us compute the terms in the Q-equations (117). For $k \in \{1, 2, \cdots, b\}$, we have

$$W_{u^{(0)}}(-e_k, \alpha_k) = \sum_{n' + \sum_{m=1}^{p} n_m - n'_m = -e_k} u^{(0)}(n', l') \prod_{m=1}^{p} u^{(0)}(n_m, l_m) v^{(0)}(n'_m, l'_m)$$
$$\left(\sum_{x \in \mathbb{Z}} \phi_{\alpha_k}(x) \phi_{l'}(x) \prod_{m=1}^{p} \phi_{l_m}(x) \phi_{l'_m}(x) \right). \quad (118)$$

The sum in (118) runs over $l_m \in \mathbb{Z}$, $l'_m \in \mathbb{Z}$, $n'_m \in \mathbb{Z}^b$, $n_m \in \mathbb{Z}^b$, $n' \in \mathbb{Z}^b$, $l' \in \mathbb{Z}$, $m = 1, 2, \cdots, p$.

Since $u^{(0)}$ has support $\{(-e_k, \alpha_k)\}_{k=1}^{b}$, in order to contribute to (118), one has that

$$l_m \in \{\alpha_k\}_{k=1}^{b}, l'_m \in \{\alpha_k\}_{k=1}^{b}, m = 1, 2, \cdots, b \text{ and } l' \in \{\alpha_k\}_{k=1}^{b}. \quad (119)$$

Take $\sum_{x\in\mathbb{Z}}\phi_{\alpha_k}(x)\phi_{l'}(x)\prod_{m=1}^{p}\phi_{l_m}(x)\phi_{l'_m}(x)$ into consideration. Assume $l'=\alpha_k$ and $l_m=l'_m=\alpha_k$, $m=1,2,\cdots,b$. It is easy to see that (similar to the proof of (25)),

$$\sum_{\ell\in\mathbb{Z},|\ell-\ell_{\alpha_k}|\le\frac{L}{2}}|\phi_{\alpha_k}(\ell)|^2\ge 1-e^{-\frac{\gamma}{4}L}. \tag{120}$$

This implies that there exists $\ell\in\mathbb{Z}$ with $|\ell-\ell_{\alpha_k}|\le\frac{L}{2}$ such that

$$|\phi_{\alpha_k}(\ell)|\ge\frac{1}{L^3}. \tag{121}$$

Therefore, in this case,

$$\begin{aligned}\sum_{x\in\mathbb{Z}}\phi_{\alpha_k}(x)\phi_{l'}(x)\prod_{m=1}^{p}\phi_{l_m}(x)\phi_{l'_m}(x)&=\sum_{x\in\mathbb{Z}}|\phi_{\alpha_k}(x)|^{2p+2}\\&\ge\frac{1}{L^{10p}}.\end{aligned} \tag{122}$$

Except for the case $l'=\alpha_k$ and $l_m=l'_m=\alpha_k$, $m=1,2,\cdots,b$, by (119), we have that

$$\Big|\sum_{x\in\mathbb{Z}}\phi_{\alpha_k}(x)\phi_{l'}(x)\prod_{m=1}^{p}\phi_{l_m}(x)\phi_{l'_m}(x)\Big|\le e^{-cL}. \tag{123}$$

Denote by $A_k=\sum_{x\in\mathbb{Z}}|\phi_{\alpha_k}(x)|^{2p+2}$. From (118), (122) and (123) (the leading contribution in the sum of (118) is when $(n',l')=(n'_m,l'_m)=(n_m,l_m)=(-e_k,\alpha_k)$), we have that

$$\omega_k^{(0)}=\mu_{\alpha_k}+\delta(A_k a_k^{2p}+O(1)e^{-cL}),$$

and $\frac{1}{L^{10p}}\le A_k\le 1$.

Denote by $\Omega_0=[\mu_{\alpha_1},\mu_{\alpha_1}+2^{2p+1}\delta]\times[\mu_{\alpha_2},\mu_{\alpha_2}+2^{2p+1}\delta]\times[\mu_{\alpha_b},\mu_{\alpha_b}+2^{2p+1}\delta]\subset\mathbb{R}^b$. Assume that after r steps, we obtain a C^1 function $u^{(r)}(\omega,a)$ on $\Omega_0\times[1,2]^b$. Substituting $u^{(r)}(\omega,a)$ and $v^{(r)}(\omega,a)$ into (117), the implicit function theorem yields,

$$\begin{aligned}\omega&=\omega^{(r+1)}(a),\\a&=a^{(r+1)}(\omega).\end{aligned} \tag{124}$$

Moreover, for some C^1 functions f_k, $k=1,2,\cdots,b$,

$$\omega_k=\mu_{\alpha_k}+\delta(A_k a_k^{2p}+f_k(a_1,a_2,\cdots,a_b)). \tag{125}$$

Denote by Γ_r, the graph of (ω,a) at step r. Denote by P_x the projection onto the x-variable, where $x=a$, ω or (ω,a).

5.4. The induction hypothesis. Let M be a large integer, and denote by $B(0, R)$ the ℓ^∞ ball on $\mathbb{Z}^{b+1}$ centered at the origin with radius R. Set

$$r_0 = \left\lfloor \frac{|\log\delta|^{\frac{3}{4}}}{\log M} \right\rfloor.$$

The proof of the Theorem is an induction. So we first lay down the induction hypothesis, which we prove in Sect. 6. In the following C is a large constant and $c > 0$ is a small constant. Recall that

$$\Omega_0 = [\mu_{\alpha_1}, \mu_{\alpha_1} + 2^{2p+1}\delta] \times [\mu_{\alpha_2}, \mu_{\alpha_2} + 2^{2p+1}\delta] \times \cdots \times [\mu_{\alpha_b}, \mu_{\alpha_b} + 2^{2p+1}\delta] \subset \mathbb{R}^b.$$

For $r \geq 1$, we assume that the following holds:

Hi. $u^{(r)}(\omega, a)$ is a C^1 map on $\Omega_0 \times [1, 2]^b$, and $\operatorname{supp} u^{(r)} \subseteq B(0, M^r)$ ($\operatorname{supp} u^{(0)} \subset B(0, \frac{1}{10}M)$).

Hii. $\|\Delta_{\mathrm{cor}} u^{(r)}\| \leq \delta_r$, $\|\partial\Delta_{\mathrm{cor}} u^{(r)}\| \leq \bar{\delta}_r$, where ∂ denotes ∂_x, x stands for ω_i, a_i, $i = 1, 2, \cdots, b$ and $\|\ \| := \sup_{(\omega,a)} \|\ \|_{\ell^2(\mathbb{Z}^{b+1})}$.

Hiii. $|u^{(r)}(n, j)| \leq Ce^{-c(|n|+|j|)}$.

Hiv. There exists Λ_r, a set of open sets I in (ω, a) of size $M^{-r^{10C}}$ when $r \geq r_0$ (the total number of open sets is therefore bounded above by $M^{r^{10C}}$), such that for any $(\omega, a) \in \bigcup_{I\in\Lambda_r} I$ when $r \geq r_0$ and $(\omega, a) \in \Omega_0 \times [1, 2]^b$ when $1 \leq r \leq r_0 - 1$,

(1) $u^{(r)}(\omega, a)$ is a rational function in (ω, a) of degree at most M^{r^3};

(2)

$$\|F(u^{(r)})\| \leq \kappa_r, \ \|\partial F(u^{(r)})\| \leq \bar{\kappa}_r; \tag{126}$$

(3) For $r \leq r_0 - 1$,

$$\|(R_{[-M^r,M^r]^{b+1}} \tilde{T}_{u^{(r-1)}}(\omega, a) R_{[-M^r,M^r]^{b+1}})^{-1}\| \leq 2\delta^{-\frac{1}{8}}, \tag{127}$$

and

$$|((R_{[-M^r,M^r]^{b+1}} \tilde{T}_{u^{(r-1)}}(\omega, a) R_{[-M^r,M^r]^{b+1}})^{-1} - \tilde{T}_{M^r}^{-1})(n, j; n', j')| \leq \delta^{\frac{1}{2}} e^{-c(|n-n'|+|j-j'|)}, \tag{128}$$

where $\tilde{T}_{M^r}$ is the diagonal component of $R_{[-M^r,M^r]^{b+1}} \tilde{T}_{u^{(r-1)}}(\omega, a) R_{[-M^r,M^r]^{b+1}}$.

(4) For $r \geq r_0$,

$$\|(R_{[-M^r,M^r]^{b+1}} \tilde{T}_{u^{(r-1)}}(\omega, a) R_{[-M^r,M^r]^{b+1}})^{-1}\| \leq M^{r^C}, \tag{129}$$

and for $|n - n'| + |j - j'| > r^C$,

$$|(R_{[-M^r,M^r]^{b+1}} \tilde{T}_{u^{(r-1)}}(\omega, a) R_{[-M^r,M^r]^{b+1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{130}$$

(5) Each $I \in \Lambda_r$ is contained in an open set $I' \in \Lambda_{r-1}$, $r \geq r_0$, and

$$\mathrm{Leb}(P_a(\Gamma_{r-1} \cap (\bigcup_{I' \in \Lambda_{r-1}} I' \backslash \bigcup_{I \in \Lambda_r} I))) \leq e^{-|\log \delta|^{K_1^{90}}} + M^{-\frac{r_0}{2^b}}, r = r_0; \quad (131)$$

and

$$\mathrm{Leb}(P_a(\Gamma_{r-1} \cap (\bigcup_{I' \in \Lambda_{r-1}} I' \backslash \bigcup_{I \in \Lambda_r} I))) \leq M^{-\frac{r}{2^b}}, r \geq r_0 + 1; \quad (132)$$

(6) for $(\omega, a) \in \bigcup_{I \in \Lambda_r} I$ with $r \geq r_0$, ω satisfies the conditions (83), and (84) for $n \neq 0$, in the scales $\tilde{N}$ with any $\tilde{N} \in [(\log \frac{1}{\delta})^K, r^C]$;

Hv. The iteration holds with

$$\delta_r = \delta^{\frac{1}{2}} M^{-(\frac{4}{3})^r}, \ \bar{\delta}_r = \delta^{\frac{1}{8}} M^{-\frac{1}{2}(\frac{4}{3})^r}; \kappa_r = \delta^{\frac{3}{4}} M^{-(\frac{4}{3})^{r+2}}, \ \bar{\kappa}_r = \delta^{\frac{3}{8}} M^{-\frac{1}{2}(\frac{4}{3})^{r+2}}. \quad (133)$$

Remark 12. As usual in multi-scale arguments, the constant c depends on r. From step r to step $r+1$, c becomes slightly smaller (decreases by γ_r' and $\sum_r \gamma_r'$ is small). We may neglect the dependence since it is essentially irrelevant (c has a uniformly (positive) lower bound).

Remark 13. The Lyapunov-Schmidt approach to quasi-periodic solutions was initiated in the paper [40], and greatly generalized by Bourgain starting from the paper [41]. See [42] for a review on this method.

6. Proof of Theorem 1.1

We now prove Theorem 1.1 by using induction. The general scheme of the proof is that, for small scales, we use (71) and (72) of Theorem 3.5 to solve the P-equations; while for larger scales, we use (73) and (74) of Theorem 3.5, Theorem 4.1 and semi-algebraic projection. We refer again to [39] for step by step constructions.

Let us state the projection lemma.

Lemma 6.1. *Let $\mathcal{S} \subset [0,1]^{d_1} \times [0,1]^{d_2} := [0,1]^d$, be a semi-algebraic set of degree B and $\mathrm{meas}_d \mathcal{S} < \eta$, $\log B \ll \log 1/\eta$. Denote by $(x, y) \in [0,1]^{d_1} \times [0,1]^{d_2}$ the product variable. Fix $\epsilon > \eta^{1/d}$. Then there is a decomposition*

$$\mathcal{S} = \mathcal{S}_1 \bigcup \mathcal{S}_2,$$

with $\mathcal{S}_1$ satisfying

$$Leb(\mathrm{Proj}_x \mathcal{S}_1) \leq B^C \epsilon,$$

and $\mathcal{S}_2$ the transversality property

$$Leb(\mathcal{S}_2 \cap L) \leq B^C \epsilon^{-1} \eta^{1/d},$$

for any d_2-dimensional hyperplane L in $[0,1]^{d_1+d_2}$ such that

$$\max_{1 \leq l \leq d_1} |\mathrm{Proj}_L(e_l)| \leq \frac{1}{100}\epsilon,$$

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

The above lemma is the basic tool, underlining the semi-algebraic techniques used in the subject. It is stated as (1.5) in [43], cf., Lemma 9.9 [31] and Proposition 5.1 [34], and relies on the Yomdin-Gromov triangulation theorem. For a complete proof of the latter, see [44]. Together with Theorem 4.1, (73) and (74), it enables us to go beyond the perturbative scales in (71) and (72).

Proof of the induction hypothesis. Assume that the induction holds for all scales up to r. We will prove that it holds for $r+1$. From our construction, it is easy to see that $u^{(r)}(\omega, a)$ is a rational function in (ω, a) of degree at most $M^{(r+1)^3}$ (see p.159 in [31]).

For $r \leq r_0 - 1$, by (71), (72) and standard perturbation arguments (Lemma 4.2), we have that for any $(\omega, a) \in \Omega_0 \times [1,2]^b$, (127) and (128) hold. Clearly, for $r = r_0 - 1$, (127) and (128) imply (129) and (130). We are in a position to treat the case $r \geq r_0$.

For each $I \in \Lambda_r$, split it into intervals with size $M^{-(r+1)^{10C}}$. Denote by $\mathcal{X}$ collections of those intervals I_3 (with size $M^{-(r+1)^{10C}}$) satisfying for any $(\omega, a) \in I_3$, (ω, a) satisfies Hiv, namely (ω, a) satisfies the $r+1$-step of (6) in Hiv and both (129) and (130) hold. Denoting the collection by Λ_{r+1}, we have constructed Λ_{r+1}. Except for (131) and (132), it is now routine that the rest of Hi-v hold for $r+1$, see Chap. 18, IV, (18.36)-(18.41) [31] and Lemmas 5.2 and 5.3 [25]. See appendix E for more details.

We proceed to the proof of the measure estimates (132) and (131). Let $N = M^{r+1}$ and $N_1 = (\log N)^C$ with a large constant C. Let

$$r_1 = \left\lfloor \frac{2 \log N_1}{\log \frac{4}{3}} \right\rfloor + 1, \tag{134}$$

so that $\delta_{r_1} < e^{-N_1^2}$. Consider $T_{u^{(r_1)}}$. Pick one interval $I \in \Lambda_{r_1}$ of size $M^{-r_1^{10C}}$ and let $I_1 = P_\omega(I \cap \Gamma_{r_1})$. For $k > (\log \frac{1}{\delta})^K$, denote by $\mathrm{DC}(k)$ all $\omega \in \Omega_0$ satisfying (83) and (84) for any $(\log \frac{1}{\delta})^K \leq \tilde{N} \leq k$. Denote by DC^{r+1} all $\omega \in \mathrm{DC}(r^C)$ satisfying (83) and (84) with a factor 2 on the right side for any $r^C \leq \tilde{N} \leq (r+1)^C$. Let $I_2 = I_1 \cap \mathrm{DC}^{r+1}$. We remark that DC^{r+1} is slightly smaller than $\mathrm{DC}(N_1) = \mathrm{DC}((r+1)^C)$ since we modify the arithmetic condition from scales r^C to $(r+1)^C$.

From the Q-equation, we have that the size of I_1 is smaller than $C\delta$. Solving the Q-equation at step $r_1 - 1$, one has that $a = a^{(r_1)}(\omega)$, $\omega \in I_2$. Since $\omega \in \Omega_0$, one has that the first assumption of Theorem 4.1 always holds. By Theorem 4.1 (see Remark 15) together with Remark 11, there exists X_{N_1} (depending on ω) such that for any $\theta \notin X_{N_1}$,

$$\|(R_{Q_{N_1}} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) R_{Q_{N_1}})^{-1}\| \leq e^{N_1^{\frac{9}{10}}}, \tag{135}$$

and for any $(n, j) \in \mathbb{Z}^{b+1}$ and $(n', j') \in \mathbb{Z}^{b+1}$ with $|n - n'| + |j - j'| \geq \sqrt{N_1}$,

$$|(R_{Q_{N_1}} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) R_{Q_{N_1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}, \tag{136}$$

and X_{N_1} satisfies

$$\mathrm{Leb}(X_{N_1}) \leq e^{-N_1^{\frac{1}{30}}}. \tag{137}$$

Let $K_{N_1} = \{\pm n \cdot \omega + \mu_j : |n| \leq N_1, |j| \leq 3N_1\}$ and I_{N_1} (depending on ω) be the $N_1^{10b}\delta$ neighbour of K_{N_1}. Assume $\theta \notin I_{N_1}$. Then the diagonal entries D_+, D_- are larger than $N_1^{10b}\delta$. Perturbation argument (Lemma 4.2) leads to, for any $|j_0| \le 2N_1$ and $Q_{N_1} \in \mathcal{E}^0_{N_1}$,

$$\|(R_{Q_{N_1}(j_0)} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) R_{Q_{N_1}(j_0)})^{-1}\| \le \frac{1}{\delta},$$

and for any (n, j) and (n', j') satisfying $|n - n'| + |j - j'| \ge \sqrt{N_1}$,

$$\begin{aligned} |(R_{Q_{N_1}(j_0)} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) Q_{Q_{N_1}(j_0)})^{-1}(n, n')| &\le 4\frac{N_1^{10b}}{\delta} e^{-c(|n-n'|+|j-j'|)} \\ &\le e^{-c(|n-n'|+|j-j'|)}, \end{aligned}$$

where the second inequality holds by the fact that $N_1 = (\log \frac{1}{\delta})^C$. Therefore, (135) and (136) hold. This implies that $X_{N_1} \subset I_{N_1}$. Since Ω_0 has size $C\delta$, we can assume that X_{N_1} is in a union of a collection of intervals of size δ with total number N_1^C (independent of ω). Pick one interval Θ. Let $\mathcal{X}_{N_1}(\omega, \theta) \subset I_2 \times \Theta$ be such that there exists some $Q_{N_1} \in \mathcal{E}_{N_1}$ such that either (135) or (136) is not true. By (137) and Fubini theorem, one has that

$$\mathrm{Leb}(\mathcal{X}_{N_1}) \le C\delta e^{-N_1^{\frac{1}{30}}} \le \delta e^{-N_1^{\frac{1}{31}}}. \tag{138}$$

We can assume that $\mathcal{X}_{N_1} \subset I_2 \times \Theta$ is a semi-algebraic set of degree at most $N_1^C M^{Cr_1^3}$. This can be seen as follows. Let $\tilde{X}_{N_1} \subset \Omega_0 \times [1, 2]^b \times \mathbb{R}$ be such that there exists some $Q_{N_1} \in \mathcal{E}_{N_1}$ such that one of the following is not true:

$$\|(R_{Q_{N_1}} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) R_{Q_{N_1}})^{-1}\| \le e^{N_1^{\frac{9}{10}}}, \tag{139}$$

and for any $(n, j) \in \mathbb{Z}^{b+1}$ and $(n', j') \in \mathbb{Z}^{b+1}$ with $|n - n'| + |j - j'| \ge \sqrt{N_1}$,

$$|(R_{Q_{N_1}} T_{u^{(r_1)}}(\theta, \omega, a^{(r_1)}(\omega)) R_{Q_{N_1}})^{-1}(n, j; n', j')| \le e^{-c(|n-n'|+|j-j'|)}. \tag{140}$$

Therefore, $\mathcal{X}_{N_1} = P_{(\omega,\theta)}(\tilde{X}_{N_1} \cap (\Gamma_{r_1} \times \mathbb{R}))$. Clearly, both $\tilde{X}_{N_1}$ and Γ_{r_1} are semi-algebraic sets of degree at most $N_1^C M^{Cr_1^3}$. Lemma 4.5 implies that $\mathcal{X}_{N_1}$ is a semi-algebraic set of degree at most $N_1^C M^{Cr_1^3}$.

Let

$$\epsilon_l = M^{-\frac{r}{2^{b-l}}}, l = 1, 2, \cdots, b-1, \text{ and } \epsilon_b = 10M^{-r}. \tag{141}$$

Choose any $|j_0| \le 2N_1$. Recall that $T_{u^{(r_1)}}$ is Töplitz with respect to $n \in \mathbb{Z}^d$. Denote by $\epsilon_{b+1} = e^{-N_1^{\frac{1}{40}}}$. Applying Lemma 6.1 in all possible directions (see (3.26) in [43]) and also on all possible open sets and Θ (the total number is bounded by $N_1^C M^{Cr_1^3}$), there exists a set of ω, $I_2^r \subset I_2$ such that

$$\mathrm{Leb}(I_2^r) \le \delta M^{Cr_1^3} N_1^C \left(\sum_{k=2}^{b+1} \left(\prod_{l=1}^{k-1} \epsilon_l^{-1} \right) \epsilon_k \right) \le \delta M^{-\frac{r}{2^{b-1}}} N_1^C M^{Cr_1^3}, \tag{142}$$

and for any $\omega \in I_2 \backslash I_2^r$, one has that for any $(n_0, j_0) \in [-N, N]^b \times [-2N_1, 2N_1]$ with $\max\{|n_0|, |j_0|\} \geq \frac{M^r}{10}$ and $(\omega, a) \in \Gamma_{r_1}$,

$$\|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}\| \leq e^{N_1^{\frac{9}{10}}}, \tag{143}$$

and for any (n, j) and (n', j') satisfying $|n - n'| + |j - j'| \geq \sqrt{N_1}$,

$$|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{144}$$

Let us explain where the factor δ in (142) is from. We apply Lemma 6.1 in $I_1 \times \Theta$, where both I_1 and Θ have sizes $C\delta$. By scaling, we have such a δ factor.

Assume $|j_0| > 2N_1$. In this case, by the standard perturbation theory, we can assume that $R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}}$ is essentially a diagonal matrix. So, we only need to remove ω such that for some $(n, j) \in [-N, N]^{b+1}$, $|n \cdot \omega + \mu_j| \leq 2e^{-N_1^{\frac{9}{10}}}$ or $|n \cdot \omega - \mu_j| \leq 2e^{-N_1^{\frac{9}{10}}}$. This can not happen when $n = 0$ because of (70). Direct compuations imply that there exists a set of ω, $\tilde{I}_2^r \subset I_2$ such that

$$\mathrm{Leb}(\tilde{I}_2^r)) \leq N^{C(b)} e^{-N_1^{\frac{9}{10}}} < \delta M^{-\frac{r}{2^{b-1}}}, \tag{145}$$

and for any $\omega \in I_2 \backslash \tilde{I}_2^r$, one has that for any $(n_0, j_0) \in [-N, N]^{b+1}$ with $\max\{|n_0|, |j_0|\} \geq \frac{M^r}{10}$, $|j_0| \geq 2N_1$ and $(\omega, a) \in \Gamma_{r_1}$,

$$\|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}\| \leq e^{N_1^{\frac{9}{10}}}, \tag{146}$$

and for any (n, j) and (n', j') such that $|n - n'| + |j - j'| \geq \sqrt{N_1}$,

$$|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{147}$$

Therefore, we have that for any $(n_0, j_0) \in [-M^{r+1}, M^{r+1}]^{b+1}$ with $\max\{|n_0|, |j_0|\} \geq \frac{M^r}{10}$, $(\omega, a) \in \Gamma_{r_1}$ and $\omega \in I_2 \backslash (I_2^r \cup \tilde{I}_2^r)$,

$$\|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}\| \leq e^{N_1^{\frac{9}{10}}}, \tag{148}$$

and for any (n, j) and (n', j') such that $|n - n'| + |j - j'| \geq \sqrt{N_1}$,

$$|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r_1)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{149}$$

Since the distance between Γ_{r_1} and Γ_r is less than $C\delta_{r_1} \leq Ce^{-N_1^2}$ and $\|u^{(r_1)} - u^{(r)}\| \leq C\delta_{r_1} \leq Ce^{-N_1^2}$, by perturbation arguments, for any $(n_0, j_0) \in [-M^{r+1}, M^{r+1}]^{b+1}$ with $\max\{|n_0|, |j_0|\} \geq \frac{M^r}{10}$, $(\omega, a) \in \Gamma_r$ and $\omega \in I_2 \backslash (I_2^r \cup \tilde{I}_2^r)$,

$$\|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}\| \leq 2e^{N_1^{\frac{9}{10}}}, \tag{150}$$

and for any (n, j) and (n', j') such that $\max\{|n - n'|, |j - j'|\} \geq \sqrt{N_1}$,

$$|(R_{(n_0,j_0)+Q_{N_1}} T_{u^{(r)}}(\omega, a) R_{(n_0,j_0)+Q_{N_1}})^{-1}(n, j; n', j')| \leq 2e^{-c(|n-n'|+|j-j'|)}. \tag{151}$$

Since $\|\Delta_{\rm cor}u^{(r)}\| \le \delta_r$, by Hiii one has that (see Remark 14)

$$|T_{u^{(r)}}(n, j; n', j') - T_{u^{(r-1)}}(n, j; n', j')| \le M^{-10r^C} e^{-c(|n-n'|+|j-j'|)}. \tag{152}$$

By (129) and (130) at step r, and using perturbation arguments (Lemma 4.2) with (152), one has that for any $(\omega, a) \in (\cup_{I\in\Lambda_r} I)$,

$$\|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r)}}(\omega, a)R_{[-M^r,M^r]^{b+1}})^{-1}\| \le 2M^{r^C}, \tag{153}$$

and for any (n, j) and (n', j') such that $|n-n'|+|j-j'| > r^C$,

$$|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r)}}(\omega, a)R_{[-M^r,M^r]^{b+1}})^{-1}(n, j; n', j')| \le 2e^{-c(|n-n'|+|j-j'|)}. \tag{154}$$

From (150)-(154) and resolvent expansion as in Lemma 3.6 [36], one has that for any $(\omega, a) \in (\cup_{I\in\Lambda_r} I) \cap \Gamma_r$ and $\omega \in I_2\backslash(I_2^r \cup \tilde{I}_2^r)$

$$\|(R_{[-M^{r+1},M^{r+1}]^{b+1}}\tilde{T}_{u^{(r)}}(\omega, a)R_{[-M^{r+1},M^{r+1}]^{b+1}})^{-1}\| \le \frac{1}{2}M^{(r+1)^C}, \tag{155}$$

and for any (n, j) and (n', j') such that $|n-n'|+|j-j'| \ge (r+1)^C$,

$$|(R_{[-M^{r+1},M^{r+1}]^{b+1}}\tilde{T}_{u^{(r)}}(\omega, a)R_{[-M^{r+1},M^{r+1}]^{b+1}})^{-1}(n, j; n', j')| \le \frac{1}{2}e^{-c(|n-n'|+|j-j'|)}. \tag{156}$$

Assume $|(\omega_1, a_1) - (\omega, a)| \le M^{-(r+1)^{10C}}$. Similar to (152) (see Remark 14), one has that

$$|(\tilde{T}_{u^{(r)}}(\omega, a))(n, j; n', j') - (\tilde{T}_{u^{(r)}}(\omega_1, a_1))(n, j; n', j')| \le M^{-10(r+1)^C} e^{-c(|n-n'|+|j-j'|)}. \tag{157}$$

This implies that (155) and (156) remain the same in a $M^{-(r+1)^{10C}}$-neighborhood of (ω, a) (a perturbation argument, Lemma 4.2). Therefore, we have that for any $(\omega, a) \in (\cup_{I\in\Lambda_r} I) \cap \Gamma_r$ and $\omega \in I_2\backslash(I_2^r \cup \tilde{I}_2^r)$, $(\omega, a) \in (\cup_{I\in\Lambda_{r+1}} I)$.

Clearly, one has

$$\mathrm{Leb}(\mathrm{DC}^1((r+1)^C)\backslash \mathrm{DC}(r^C)) \le e^{-r^C}. \tag{158}$$

In order to have eq. (5) in Hiv at the beginning step ($r = r_0 - 1$), we have to remove ω of measure less than $\delta^2 e^{-|\log\delta|^{K_1^{90}}}$, namely,

$$\mathrm{Leb}(\mathrm{DC}((r_0-1)^C)\backslash \mathrm{DC}(|\log\delta|^K)) \le \delta^2 e^{-|\log\delta|^{K_1^{90}}}. \tag{159}$$

By counting all possible intervals I_1 (total number is bounded by $M^{r_1^{10C}}$), and by (142), (145), (158) and (159), one has that for $r \ge r_0$, one has

$$\mathrm{Leb}(P_\omega(\Gamma_r \cap (\bigcup_{I'\in\Lambda_r} I' \backslash \bigcup_{I\in\Lambda_{r+1}} I))) \le \delta M^{r_1^C} M^{-\frac{r}{2^{b-1}}} + e^{-r^C}, \tag{160}$$

and for $r = r_0 - 1$,

$$\mathrm{Leb}(P_\omega(\Gamma_r \cap (\bigcup_{I' \in \Lambda_r} I' \setminus \bigcup_{I \in \Lambda_{r+1}} I))) \le \delta M^{r_1^C} M^{-\frac{r}{2^{b-1}}} + \delta^2 e^{-|\log\delta|^{K_1^{90}}} + e^{-r^C}. \quad (161)$$

This implies that for $r \ge r_0$,

$$\mathrm{Leb}(P_a(\Gamma_r \cap (\bigcup_{I' \in \Lambda_r} I' \setminus \bigcup_{I \in \Lambda_{r+1}} I))) \le M^{-\frac{r}{2^b}}, \quad (162)$$

and for $r = r_0 - 1$,

$$\mathrm{Leb}(P_a(\Gamma_r \cap (\bigcup_{I' \in \Lambda_r} I' \setminus \bigcup_{I \in \Lambda_{r+1}} I))) \le M^{-\frac{r}{2^b}} + e^{-|\log\delta|^{K_1^{90}}}. \quad (163)$$

□

Remark 14. We give more details about the proof of (152) and (157). In order to avoid repetition, we only prove (157). When $|n - n'| + |j - j'| \le (r+1)^{5C} \log M$, one has that

$$\begin{aligned}
|(\tilde{T}_{u^{(r)}}(\omega, a))(n, j; n', j') - (\tilde{T}_{u^{(r)}}(\omega_1, a_1))(n, j; n', j')| \\
&\le C\|u^{(r)}(\omega, a) - u^{(r)}(\omega_1, a_1)\| \\
&\le C M^{-(r+1)^{10C}} \\
&\le M^{-10(r+1)^C} e^{-c(|n-n'|+|j-j'|)}.
\end{aligned}$$

When $|n - n'| + |j - j'| > (r+1)^{5C} \log M$, by Hiii and Lemma 5.1, one has that

$$\begin{aligned}
|(\tilde{T}_{u^{(r)}}(\omega, a))(n, j; n', j') - (\tilde{T}_{u^{(r)}}(\omega_1, a_1))(n, j; n', j')| \\
&\le C|n - n'|^C e^{-c(|n-n'|+|j-j'|)} \\
&\le M^{-10(r+1)^C} e^{-\bar{c}(|n-n'|+|j-j'|)},
\end{aligned}$$

where $\bar{c} = c - (r+1)^{-3C}$.

Remark 15. We need to adjust Theorem 4.1 in the applications to the nonlinear analysis. In Theorem 4.1, the operator T in (80) is fixed; while in the nonlinear analysis, T typically varies with each step. For the scale N, let

$$r(N) = \left\lfloor \frac{2\log N}{\log\frac{4}{3}} \right\rfloor + 1, \quad (164)$$

so that $\delta_{r(N)} < e^{-N^2}$. (Eq. (164) appeared previously as (134).) When we study the LDT at scale N, the operator $T = T_{u^{(r(N))}}$ depends on N. Now the "lossless" inductive estimate (see (1) in appendix C) still holds. This can be seen as follows.

We aim to establish the LDT at scale N_3 from scales N_1 and N_2. Since $|T_{u^{(r(N_1))}} - T_{u^{(r(N_3))}}| \le e^{-N_1^{3/2}}$ and $|T_{u^{(r(N_2))}} - T_{u^{(r(N_3))}}| \le e^{-N_2^{3/2}}$, perturbation arguments allow us to replace $T_{u^{(r(N_1))}}$ and $T_{u^{(r(N_2))}}$ with $T_{u^{(r(N_3))}}$. This means that essentially the operator does not change from scales N_1 and N_2 to the scale N_3.

Proof of Theorem 1.1. Theorem 1.1 follows immediately from the (by now verified) hypothesis (Hi-v). □

7. NLRS in Arbitrary Dimensions

We discuss the discrete nonlinear random Schrödinger equation (NLRS) in arbitrary dimensions:

$$i\frac{\partial u}{\partial t} = -\Delta u + Vu + \delta|u|^{2p}u,\ p \in \mathbb{N}, \tag{165}$$

where

Δ is the discrete Laplacian on $\mathbb{Z}^d$:

$$(\Delta u)(x) = \sum_{x' \in \mathbb{Z}^d, \sum_{j=1}^d |x_j - x'_j| = 1} u(x'),$$

and $V = \{v_x\}, x \in \mathbb{Z}^d$ is a family of independent identically distributed random variables on [0, 1], with distribution density g. Assume that g is bounded, $g \in L^\infty$.

Let

$$H = -\Delta + V, \tag{166}$$

be the random Schrödinger operator. Assume that H has Anderson localization.

Let $\{\varphi_j^V\}_{j\in\mathbb{Z}^d}$ be the (real) eigen-basis of H. Assume that ι_j^V satisfies

$$|\varphi_j^V(\iota_j^V)| = \max_{x\in\mathbb{Z}^d} |\varphi_j^V(x)|,$$

and that

A1. there exist some $q > 0$ and $\gamma_1 > 0$ such that, with probability 1,

$$|\varphi_j^V(\ell)| \leq C_V(1 + |\iota_j^V|)^q e^{-\gamma_1|\ell - \iota_j^V|}, \tag{167}$$

where $\mathbb{E}(C_V) < \infty$.

A2. H satisfies the Wegner estimate: for and $\Lambda \subset \mathbb{Z}^d$, $E \in \mathbb{R}$ and $\varepsilon > 0$,

$$\mathbb{E}(\text{dist}(E, \sigma(H_\Lambda)) \leq \varepsilon) \leq C|\Lambda|\varepsilon.$$

A3. H satisfies Minami estimate. Let $\Lambda \subset \mathbb{Z}^d$ and $J \subset \mathbb{R}$ be an interval. Then we have

$$\mathbb{E}([\text{tr}(\mathbf{1}_J(H_\Lambda))] \cdot [\text{tr}(\mathbf{1}_J(H_\Lambda)) - 1]) \leq C|\Lambda|^2|J|^2,$$

where $\mathbf{1}_J$ is the characteristic function of the interval J.

Remark 16. As mentioned previously, assumptions A1-3 are satisfied in arbitrary dimensions at high disorder, i.e., when V is replaced by λV with $\lambda \gg 1$.

Based on assumption A1, we can relabel the eigenfunctions of $H = -\Delta + V$ similar to Lemma 3.3. Denote by eigenfunctions, ϕ_j^V, and eigenvalues, μ_j^V, $j \in \mathbb{Z}^d$ after the relabelling. It can be readily seen that the proof of Theorem 1.1 only uses that A1-3 hold. Using A1-3, we can prove its generalization to arbitrary dimensions, as stated below. We do not repeat the proof, as it basically follows that of Theorem 1.1 verbatim.

Theorem 7.1. *Consider the discrete NLRS in (arbitrary) dimension $d \geq 1$:*

$$i\frac{\partial u}{\partial t} = -\Delta u + Vu + \delta|u|^{2p}u,\ p \in \mathbb{N}. \tag{168}$$

Assume that $-\Delta + V$ satisfies assumptions A1-A3. Let $a = (a_1, a_2, \cdots, a_b) \in [1, 2]^b$. For any $\epsilon > 0$, there exists l_ϵ such that the following holds. Fix any $L \geq l_\epsilon$ and $\beta_k \in \mathbb{Z}^d$, $k = 1, 2, \cdots, b$ satisfying $10L \leq |\beta_k| \leq L^3$ and $|\beta_k - \beta_{k'}| \geq 10\,L$ for any distinct $k, k' \in \{1, 2, \cdots, b\}$, there exist a subset X_ϵ with $\mathbb{P}(X_\epsilon) \geq 1-\epsilon$ and $\delta_0 > 0$ (depending on g, ϵ and L) such that for any $V \in X_\epsilon$ and $0 < \delta \leq \delta_0$, any b eigenfunctions $\phi^V_{\alpha_k}$ with $\ell^V_{\alpha_k} \in B_k = \{\ell \in \mathbb{Z}^d : |\ell - \beta_k| \leq L\}$, $k = 1, 2, \cdots, b$, there exists a set $A_\delta \subset [1, 2]^b$ of measure at least $1 - e^{-|\log\delta|^{1/2}}$, such that for any $a \in A_\delta$, the nonlinear equation (3) has a solution $u(t, x)$ satisfying

$$u(t, x) = \sum_{(n,j)\in\mathbb{Z}^b\times\mathbb{Z}^d} \hat{u}(n, j)e^{in\cdot\omega t}\phi^V_j(x) = \sum_{k=1}^{b} a_k e^{-i\omega_k t}\phi^V_{\alpha_k}(x) + O(\delta^{1/2}),$$

where $\omega = (\omega_1, \omega_2, \cdots, \omega_b) = (\mu^V_{\alpha_1}, \mu^V_{\alpha_2}, \cdots, \mu^V_{\alpha_b}) + O(\delta)$,
and $\hat{u}(n, j)$ decay exponentially as $|(n, j)| \to \infty$.

Acknowledgments W. Liu was supported by NSF DMS-2246031, DMS-2000345 and DMS-2052572. W.-M. Wang acknowledges support from CY Initiative of Excellence (Grant "Investissements d'Avenir" ANR-16-IDEX-0008).

Data Availability Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Appendix A. Eigenfunction Relabelling Map

For a fixed $V \in \mathcal{V}_\epsilon$, basing on (58), one may relabel the eigenfunctions in a more intrinsic way. Write φ_i and ι_i for φ^V_i and ι^V_i, since V is fixed. The goal is that in the new labelling scheme, if $j > j'$, then the localization centers of the corresponding eigenfunctions ϕ_j and $\phi_{j'}$ satisfy $\ell_{j'} \geq \ell_j$. Below we provide such a relabelling map.
For a given eigenfunction φ_i, we first select a vertex among the set of vertices, on which φ_i achieves its maximum. (This selection could be arbitrary, but it is practical to have a rule.) So for $i \in \mathbb{Z}$, let

$$\mathcal{M}_i = \{x_0 \in \mathbb{Z} : |\varphi_i(x_0)| = \max_{x\in\mathbb{Z}} |\varphi_i(x)|\}.$$

Define $\mathcal{M}^+_i = \mathcal{M}_i \cap \{\{0\} \cup \mathbb{Z}_+\}$. If $\mathcal{M}^+_i \neq \emptyset$, define $\iota_i = \min x_0$, $x_0 \in \mathcal{M}^+_i$; otherwise define $\iota_i = \min -x_0$, $x_0 \in \mathcal{M}_i$.

Define

$$f_1 : \mathbb{Z} \mapsto \mathbb{Z},\ f_1(i) = \iota_i.$$

Let $\mathcal{L}$ be the range of f_1,

$$\mathcal{L} = \mathrm{Ran}(f_1), \mathcal{L} \subseteq \mathbb{Z}.$$

For a given $l \in \mathcal{L}$, let $I_l = \{i | \iota_i = l\}$. Define

$$f_2 : \mathcal{L} \mapsto \bigsqcup_{l \in \mathcal{L}} I_l.$$

From (58): If $|l| < l_\epsilon$, $\#\{\cup_{|l| \le l_\epsilon} I_l\} \le (1+\epsilon) l_\epsilon$; and if $|l| \ge l_\epsilon$,

$$(1-\epsilon) l \le \#\{\cup_{|l| \le l_\epsilon} I_l\} \le (1+\epsilon) l.$$

The upper bound gives that for all l, I_l is finite. So we may define a map

$$f_3 : \bigsqcup_{l \in \mathcal{L}} I_l \mapsto \mathbb{Z},$$

such that if $x \in I_l$ and $y \in I_{l'}$, with $l' > l$, then $f_3(y) > f_3(x)$.
Finally define the map f to be $f = f_3 \circ f_2 \circ f_1$,

$$f : \mathbb{Z} \mapsto \mathbb{Z}.$$

Using the relabelling map f yields our ortho-normal eigen-basis $\{\phi_j\}_{j \in \mathbb{Z}}$.

Remark 17. The above method is direct and could be generalized to arbitrary dimensions. In the large disorder case, one may, alternatively, invoke Hall's Marriage Theorem from graph theory to label the eigenfunctions as in (4.46)–(4.53), Sect. 4 [24]. Note, however, that at high disorder, with large probability, the eigenvectors of the random Schrödinger operator are close to the canonical basis of $\ell^2(\mathbb{Z}^d)$ for arbitrary d.

Appendix B. Proof of Lemma 3.2

Proof. It is similar to the proof of Theorem 7.1 in [28]. Below are the details. By Theorem 3.1 and Chebyshev's inequality, one has that for any ϵ_1, there exists $\mathcal{V}_{\epsilon_1}$ with $\mathbb{P}(\mathcal{V}_{\epsilon_1}) > 1 - \epsilon_1$ such that for any $V \in \mathcal{V}_{\epsilon_1}$,

$$|\varphi_j^V(\ell)| \le C_{\epsilon_1}(1 + |\iota_j^V|)^q e^{-\gamma_1 |\ell - \iota_j^V|}. \tag{169}$$

For simplicity, below we drop the superscript V.
Clearly, we have

$$\sum_{\ell \in \mathbb{Z}} |\varphi_j(\ell)|^2 = 1. \tag{170}$$

and

$$\sum_{j \in \mathbb{Z}} |\varphi_j(\ell)|^2 = 1. \tag{171}$$

Let ϵ be an arbitrarily small constant. Assume L is large enough, depending on ϵ and ϵ_1. If $k \le \iota_j \le k+L$ with $|k| \le L^4$, then

$$\sum_{\ell \le k-\epsilon L} |\varphi_j(\ell)|^2 + \sum_{\ell \ge k+(1+\epsilon)L} |\varphi_j(\ell)|^2 \le \sum_{m \ge \epsilon L} C_{\epsilon_1} L^{8q} e^{-\gamma_1 m} \tag{172}$$

$$\le Ce^{-\epsilon L}, \tag{173}$$

where C depends on ϵ_1 and ϵ. By (170) and (173), one has that for any j with $k \le \iota_j \le k+L$,

$$\sum_{k-\epsilon L \le \ell \le k+(1+\epsilon)L} |\varphi_j(\ell)|^2 \ge 1 - Ce^{-\epsilon L}. \tag{174}$$

By (171) and (174), we have that

$$(1+\epsilon)L \ge \sum_{\substack{k-\epsilon L \le \ell \le k+(1+\epsilon)L \\ j \in \mathbb{Z}}} |\varphi_j(\ell)|^2 \tag{175}$$

$$\ge \sum_{\substack{k-\epsilon L \le \ell \le k+(1+\epsilon)L \\ j \in \mathbb{Z}: k \le \iota_j \le k+L}} |\varphi_j(\ell)|^2 \tag{176}$$

$$\ge (1 - Ce^{-\epsilon L}) \#\{j : k \le \iota_j \le k+L\}. \tag{177}$$

This implies that for any $k \in [-L^4, L^4]$,

$$\#\{j : k \le \iota_j \le k+L\} \le (1+\epsilon)L. \tag{178}$$

For any $\ell \in [k+\epsilon L, k+(1-\epsilon)L]$ with $|k| \le L^4$, by (169), one has

$$\begin{aligned}
\sum_{j \in \mathbb{Z}: \iota_j \notin [k,k+L]} |\varphi_j(\ell)|^2 &\le \sum_{m=0}^{\infty} \sum_{\substack{j \in \mathbb{Z}: mL \le |\iota_j| \le (m+1)L \\ |\iota_j - \ell| \ge \epsilon L}} C(1+|\iota_j|)^{2q} e^{-\gamma_1 |\ell - \iota_j|} \\
&\le \sum_{m \le 10L^4} \sum_{j \in \mathbb{Z}: |\iota_j| \le 20L^5} CL^{10q} e^{-\gamma_1 \epsilon L} + \\
&\quad \sum_{m=10L^4}^{\infty} C(1+mL)^{2q} \#\{j : |\iota_j| \le (m+1)L\} e^{-\frac{1}{2}\gamma_1 mL} \\
&\le Ce^{-\epsilon L} + \sum_{m=1}^{\infty} C(1+mL)^{3q} e^{-\frac{1}{2}\gamma_1 mL} \\
&\le Ce^{-\epsilon L},
\end{aligned} \tag{179}$$

where (179) holds by (178). It implies that

$$\sum_{\substack{j \in \mathbb{Z}: \iota_j \notin [k,k+L] \\ \ell \in [k+\epsilon L, k+(1-\epsilon)L]}} |\varphi_j(\ell)|^2 \le Ce^{-\epsilon L}. \tag{180}$$

By (171) and (180), one has that

$$\begin{aligned}(1-\epsilon)L &= \sum_{\ell\in[k+\epsilon L,k+(1-\epsilon)L],\, j\in\mathbb{Z}} |\varphi_j(\ell)|^2 \\ &\le Ce^{-\epsilon L} + \sum_{\substack{j\in\mathbb{Z}:\iota_j\in[k,k+L]\\ \ell\in[k+\epsilon L,k+(1-\epsilon)L]}} |\varphi_j(\ell)|^2 \\ &\le Ce^{-\epsilon L} + \#\{j : k\le \iota_j \le k+L\}.\end{aligned}$$

This yields that for any k with $|k|\le L^4$,

$$\#\{j : k\le \iota_j\le k+L\}\ge (1-\epsilon)L. \tag{181}$$

Now (58) follows from (178) and (181). □

Appendix C. More about the Proof of Theorem 4.1

(1) For small scales, the proof follows directly from standard perturbation arguments. For large scales, the proof is based on multi-scale analysis: if the LDT holds at scales N_1, $N_2 = N_1^{C_1}$ with parameter $\tilde{c}_1$, we prove the LDT for scales $N_3 \in [N_2^{C_2}, N_2^{C_3}]$, with parameter $\tilde{c}_2 = \tilde{c}_1 - N_1^{-\kappa}$, where C_1, C_2 and C_3 are properly chosen large constants and κ is a small constant. We should mention that here $\tilde{c}_2 = \tilde{c}_1 - N_1^{-\kappa}$ is crucial. Had we used the standard definition of LDT, namely assumed that (82) holds for $|n-n'|+|j-j'|\ge \frac{N}{10}$, we would only have been able to prove $\tilde{c}_2 = \frac{4}{5}\tilde{c}_1 - N_1^{-\kappa}$ for long range quasi-periodic operators [35,38,45] (In appendix D, we show that we could improve $\frac{4}{5}\tilde{c}_1 - N_1^{-\kappa}$ to $\tilde{c}_1 - N_1^{-\kappa}$ by modifying proofs in [35,38,45]).
The above deterioration of constant (from 1 to 4/5)
may cause problems in the Newton iteration when solving the P-equations, since the encountered quasi-periodic operators are typically long range, as observed in [36]. (This problem may, in fact, have already surfaced earlier in Chap. 18 in [31].) Consequently, the definition of LDT was modified in [36] (namely, (82) holds for $|n-n'|+|j-j'|\ge\sqrt{N}$) to overcome this technical difficulty. We follow [36] and use the modified LDT approach in this paper.

(2) Using the modified definition, the main results in [35] hold without deterioration in the constants. In the standard LDT in [35], the deterioration comes from expression (72) in [35]. If we use the modified definition, namely $|n-n_1|\ge\sqrt{M}$ (not $|n-n_1|\ge\frac{M}{10}$), then (72) holds for $c = c_1 - O(M^{-\frac{1}{2}})$ ($\tilde{\sigma}$ in (72) is 1, since the matrices have (at least) exponential off-diagonal decay in this paper).

(3) In Theorem 4.1, we assumed that $h_{r,r'}$ satisfies (78), which has a polynomial prefactor. This differs slightly from the settings in [34–36,38]. However, this factor affects neither the proof of nor the statement in Theorem 4.1.

Appendix D. An Alternative Way

Let A be a Töplitz matrix on $\mathbb{Z}^d$. For a given subset $\Lambda\subset\mathbb{Z}^d$, let R_Λ be the restriction to Λ, $A_\Lambda = R_\Lambda A R_\Lambda$, and $G_\Lambda = (R_\Lambda A R_\Lambda)^{-1}$, assuming that the inverse is well defined.

Lemma D.1. *Let* $\sigma_1, \kappa, \sigma \in (0, 1)$ *and* $\sigma_1 > \kappa > \sigma$. *Assume that* $\operatorname{diam}(\Lambda) \leq 2N + 1$ *and let* $M = N^\kappa$. *Suppose that for any* $n \in \Lambda$, *there exists some* $W = W(n) \in \mathcal{E}_M$ *such that* $n \in W$, $\operatorname{dist}(n, \Lambda\backslash W) \geq \frac{M}{2}$, $W \subset \Lambda$ *and*

$$\|G_W\| \leq e^{M^\sigma}, \tag{182}$$

$$|G_W(n; n')| \leq e^{-c_1|n-n'|} \text{ for } |n - n'| \geq \frac{M}{10}. \tag{183}$$

Then

$$\|G_\Lambda\| \leq e^{N^\sigma}, \tag{184}$$

and

$$|G_\Lambda(n; n')| \leq e^{-\bar{c}|n-n'|} \textit{ for any } |n - n'| \geq N^{\sigma_1},\ 0 < \sigma_1 < 1,$$

where

$$\bar{c} = c_1 - M^{-\kappa_1}, \kappa_1 = \kappa_1(\sigma, \sigma_1, \kappa). \tag{185}$$

Proof. Choose σ_2 with $\sigma < \sigma_2 < \kappa$. Let $\tilde{A}(n; n') = A(n; n')$ for $|n - n'| < N^{\sigma_2}$ and $\tilde{A}(n; n') = 0$ for $|n - n'| \geq N^{\sigma_2}$. Let $\tilde{A}_\Lambda = R_\Lambda \tilde{A} R_\Lambda$ and $\tilde{G}_\Lambda = (R_\Lambda \tilde{A} R_\Lambda)^{-1}$. Clearly, for any $n, n' \in \Lambda$, one has

$$|\tilde{A}(n; n') - A(n; n')| \leq \frac{1}{N^{20d} e^{10N^\sigma}} e^{-\bar{c}|n-n'|}, \tag{186}$$

where $\bar{c} = c - N^{-\kappa_1}$.
From (186), Lemma 4.2, (182) and (183), one has that

$$\|\tilde{G}_W\| \leq 2e^{M^\sigma}, \tag{187}$$

$$|\tilde{G}_W(n; n')| \leq 2e^{-c_1|n-n'|} \text{ for } |n - n'| \geq \frac{M}{10}. \tag{188}$$

Using resolvent expansion type arguments and the fact that the band of $\tilde{A}$ is much smaller than N^κ, one has that

$$\|G_\Lambda\| \leq \tfrac{1}{2} e^{N^\sigma},$$
$$|G_\Lambda(n; n')| \leq \tfrac{1}{2} e^{-\bar{c}|n-n'|} \text{ for any } |n - n'| \geq N^{\sigma_1}.$$

The proof now follows from Lemma 4.2 and (186). □

Let us call the standard LDT (coming from (182) and (183)), LDT, and the modified one (coming from (81) and (82)), mLDT. If LDT holds at all scales, one may use the above Lemma to deduce mLDT: if LDT holds at scale N, then mLDT holds at a slightly larger scale $N' > N$. Therefore mLDT could be seen as a corollary, and the issue of deterioration of constant does not arise.

Appendix E. Bourgain's Induction Estimates

Assume that Hi-v hold at step r, and that Hi-v hold at step $r+1$ except for (133). We show that (133) holds at step $r+1$. This follows from Chap. 18 [31] with minor modifications. Let us state the difference between our setting and that of [31]. Bourgain assumed that

$$\|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r-1)}}R_{[-M^r,M^r]^{b+1}})^{-1}\| \leq M^{r^C}, \tag{189}$$

and that for (n, j) and (n', j') satisfying $|n-n'|+|j-j'| > r^C$,

$$|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r-1)}}R_{[-M^r,M^r]^{b+1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{190}$$

Then he proved that the induction holds if

$$\begin{aligned}
\delta_{r+1} &\geq M^{(r+1)^{10C}}\kappa_r,\\
\bar{\delta}_{r+1} &\geq M^{2(r+1)^{10C}}\bar{\kappa}_r + M^{(r+1)^{10C}}\delta_{r+1},\\
\kappa_{r+1} &\geq e^{-\frac{c}{3}M^{r+1}}\kappa_r + \delta_{r+1}^2,\\
\bar{\kappa}_{r+1} &\geq M^{2(r+1)^{10C}}\kappa_r + e^{-\frac{c}{3}M^{r+1}}\bar{\kappa}_r + \delta_{r+1}\bar{\delta}_{r+1}.
\end{aligned}$$

We assume that for a proper $0 < \nu < 1$,

$$\|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r-1)}}R_{[-M^r,M^r]^{b+1}})^{-1}\| \leq \delta^{-\nu}M^{r^C}, \tag{191}$$

and that for (n, j) and (n', j') satisfying $|n-n'|+|j-j'| > r^C$,

$$|(R_{[-M^r,M^r]^{b+1}}\tilde{T}_{u^{(r-1)}}R_{[-M^r,M^r]^{b+1}})^{-1}(n, j; n', j')| \leq e^{-c(|n-n'|+|j-j'|)}. \tag{192}$$

We remark that (191) and (192) are the combination of (127), (128), (129), and (130) (with $\nu = \frac{1}{8}$).

Following Bourgain's proof[2], we obtain the following new relations

$$\begin{aligned}
\delta_{r+1} &\geq \delta^{-\nu}M^{(r+1)^{10C}}\kappa_r,\\
\bar{\delta}_{r+1} &\geq \delta^{-2\nu}M^{2(r+1)^{10C}}\bar{\kappa}_r + \delta^{-\nu}M^{(r+1)^{10C}}\delta_{r+1},\\
\kappa_{r+1} &\geq \delta^{1-\nu}e^{-\frac{c}{3}M^{r+1}}\kappa_r + \delta_{r+1}^2 + \delta e^{-\frac{c}{2}M^{r+1}},\\
\bar{\kappa}_{r+1} &\geq \delta^{-2\nu}M^{2(r+1)^{10C}}\kappa_r + \delta^{1-\nu}e^{-\frac{c}{3}M^{r+1}}\bar{\kappa}_r + \delta_{r+1}\bar{\delta}_{r+1} + \delta e^{-\frac{c}{2}M^{r+1}},.
\end{aligned}$$

For example, we may take $\nu = \frac{1}{8}$,

$$\delta_r = \delta^{\frac{1}{2}}M^{-(\frac{4}{3})^r},\ \bar{\delta}_r = \delta^{\frac{1}{8}}M^{-\frac{1}{2}(\frac{4}{3})^r},\ \kappa_r = \delta^{\frac{3}{4}}M^{-(\frac{4}{3})^{r+2}},\ \bar{\kappa}_r = \delta^{\frac{3}{8}}M^{-\frac{1}{2}(\frac{4}{3})^{r+2}}.$$

[2] There are, however, two modifications: first, we need to take the factor δ in the nonlinear term in Eq. (1) into account; second, to estimate $\Delta_{r+1}u^{(r+1)}(n, j)$ for $r \leq r_0$, we need to use (128).

References

1. Furstenberg, H.: Non-commuting random products. Trans. Am. Math. Soc. **108**, 377–428 (1963)
2. Goldshied, I., Molchanov, S., Pastur, L.: A pure point spectrum of the stochastic one-dimensional Schrödinger equation. Funct. Anal. Appl. **11**, 1–10 (1977)
3. Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux differences finies aléatoire. Commun. Math. Phys. **78**, 201–246 (1980)
4. Bucaj, V., Damanik, D., Filman, J., Gerbuz, V., VandenBoom, T., Wang, F., Zhang, Z.: Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent. Trans. Am. Math. Soc. **372**, 3619–3667 (2019)
5. Ge, L., Zhao, X.: Exponential dynamical localization in expectation for the one dimensional Anderson model. J. Spectr. Theory. **10**(3), 887–904 (2020)
6. Gorodetski, A., Kleptsyn, V.: Parametric Furstenburg theorem on random products of SL(2, $\mathbb{R}$) matrices. Adv. Math. **378**, 107522 (2021)
7. Jitomirskaya, S., Zhu, X.: Large deviation of the Lyapunov exponent and localization for the 1d Anderson model. Commun. Math. Phys. **370**, 311–324 (2019)
8. Macera, D., Sodin, S.: Anderson localisation for quasi-one-dimensional random operators. Ann. Henri Poincaré **23**(12), 4227–4247 (2022)
9. Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model. Commun. Math. Phys. **88**, 151–184 (1983)
10. von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285–299 (1989)
11. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**, 245–278 (1993)
12. Johansson, M., Kopidakis, G., Aubry, S.: KAM tori in 1D random discrete nonlinear Schrödinger model? EPL **91**, 50001 (2010)
13. Laptyeva, T.V., Ivanchenko, M.V., Flach, S.: Nonlinear waves in heterogeneous media. J. Phys. A Math. Theor. **47**, 493001 (2014)
14. Abdul-Rahman, H., Nachtergaele, B., Sims, R., Stolz, G.: Localization properties of the xy spin chain: A review of mathematical results with an eye toward many-body localization. Ann. der Physik **529**(7), 1600280 (2017)
15. Eliasson, L.H., Kuksin, S.: KAM for the nonlinear Schrödinger equation. Ann. Math. **172**(2), 371–435 (2010)
16. Bourgain, J., Wang, W.-M.: Quasi-periodic solutions of nonlinear random Schrödinger equations. J. Eur. Math. Soc. **10**, 1–45 (2008)
17. Geng, J., Sun, Y., Wang, W.-M.: Quasi-periodic solutions to nonlinear random Schrödinger equations at fixed potential realizations. J. Math. Phys. **64**, 032701 (2023)
18. Albanese, C., Fröhlich, J.: Periodic solutions of some infinite dimensional Hamiltonian systems associated with nonlinear partial differentiation equations. i. Commun. Math. Phys. **116**, 475–502 (1988)
19. Albanese, C., Fröhlich, J., Spencer, T.: Periodic solutions of some infinite dimensional Hamiltonian systems associated with nonlinear partial differentiation equations. ii. Commun. Math. Phys. **119**, 677–699 (1988)
20. Fröhlich, J., Spencer, T., Wayne, C.E.: Localization in disordered, nonlinear dynamical systems. J. Stat. Phys. **42**(3/4), 247–274 (1986)
21. Geng, J., You, J., Zhao, Z.: Localization in one-dimensional quasi-periodic nonlinear systems. Geom. Funct. Anal. **24**(1), 116–158 (2014)
22. Yuan, X.: KAM theorems and open problems for infinite-dimensional Hamiltonian with short range. Sci. China Math. **57**(7), 1479–1486 (2014)
23. Zhao, Z., Geng, J.: Linearly stable quasi-periodic breathers in a class of random Hamiltonian systems. J. Dyn. Diff. Equat. **23**, 961–997 (2011)
24. Elgart, A., Klein, A.: An eigensystem approach to Anderson localization. J. Funct. Anal. **271**, 3465–3512 (2016)
25. Wang, W.-M.: Energy supercritical nonlinear Schrödinger equations: Quasi-periodic solutions. Duke Math. J. **165**(6), 1129–1192 (2016)
26. Wang, W.-M.: Quasi-periodic solutions to a nonlinear Klein-Gordon equation with a decaying nonlinear term. arXiv:1609.00309, (2021)
27. Minami, N.: Local fluctuations of a multidimensional Anderson tight binding model. Commun. Math. Phys. **177**, 709–725 (1996)
28. 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)
29. Germinet, F., Klein, A.: New characterization of the region of complete localization for random Schrödinger operators. J. Stat. Phys. **122**, 73–94 (2006)

30. Wegner, F.: Bounds on the density of states in disordered systems. Z. Phys. B. **44**, 9–15 (1981)
31. Bourgain, J.: Green's function estimates for lattice Schrödinger operators and applications. Annals of Mathematics Studies, vol. 158. Princeton University Press, Princeton, NJ (2005)
32. Aizenman, M., Warzel, S.: Random Operators: Disorder Effects on Quantum Spectra and Dynamics, volume 168 of Graduate Studies in Mathematics. American Mathematical Society, (2015)
33. Germinet, F., Klopp, F.: Spectral statistics for random Schrödinger operators in the localized regime. J. Eur. Math. Soc. (JEMS) **16**(9), 1967–2031 (2014)
34. Bourgain, J., Goldstein, M., Schlag, W.: Anderson localization for Schrödinger equations on $\mathbb{Z}^2$ with quasi-periodic potential. Acta Math. **188**, 41–86 (2002)
35. Liu, W.: Quantitative inductive estimates for Green's functions of non-self-adjoint matrices. Anal. PDE **15**(8), 2061–2108 (2022)
36. He, X., Shi, J., Shi, Y., Yuan, X.: On linear stability of KAM tori via the Craig–Wayne–Bourgain method. SIAM J. Math. Anal. **56**(3), 3605–3645 (2024)
37. Basu, S.: On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets. Discrete Comput. Geom. **22**(1), 1–18 (1999)
38. Jitomirskaya, S., Liu, W., Shi, Y.: Anderson localization for multi-frequency quasi-periodic operators on $\mathbb{Z}^D$. Geom. Funct. Anal. **30**(2), 457–481 (2020)
39. Kachkovskiy, I., Liu, W., Wang, W.-M.: Spacetime quasiperiodic solutions to a nonlinear Schrödinger equation on $\mathbb{Z}$. J. Math. Phys. **65**(1), 34 (2024)
40. Craig, W., Wayne, C.E.: Newton's method and periodic solutions of nonlinear equations. Commun. Pure Appl. Math. **46**, 1409–1498 (1993)
41. Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of $2d$ linear Schrödinger equations. Ann. Math. **148**, 363–439 (1998)
42. Wang, W.-M.: Semi-algebraic sets method in PDE and mathematical physics. J. Math. Phys. **62**, 021506 (2021)
43. Bourgain, J.: Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary. Geom. Funct. Anal. **17**(3), 682–706 (2007)
44. Binyamini, G., Novikov, D.: Complex cellular structures. Ann. Math. **190**(2), 145–248 (2019)
45. Shi, Y.: Spectral theory of the multi-frequency quasi-periodic operator with a Gevrey type perturbation. J. Anal. Math. **148**(1), 305–338 (2022)

Communicated by S. Dyatlov