Forum of Mathematics, Sigma (2026), **Vol. 14**:e88 1–66
doi:10.1017/fms.2026.10241

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Anderson localized states for the quasi-periodic nonlinear Schrödinger equation on $\mathbb{Z}^d$

Yunfeng Shi[1] and Wei-Min Wang[2]

[1]School of Mathematics, Sichuan University, China; E-mail: yunfengshi@scu.edu.cn (Corresponding author).
[2]CNRS and Department De Mathematique, Cergy Paris Universite, France; E-mail: wei-min.wang@math.cnrs.fr

Received: 3 April 2025; **Revised:** 24 March 2026; **Accepted:** 14 May 2026

2020 Mathematical Subject Classification: *Primary* – 37K55; *Secondary* – 37K10, 37K45

Abstract
We establish large sets of Anderson localized states for the quasi-periodic nonlinear Schrödinger equation on $\mathbb{Z}^d$, thus extending Anderson localization from the linear (cf. Bourgain [Geom. Funct. Anal., 17(3):682–706, 2007]) to a nonlinear setting, and from the random (cf. Bourgain-Wang [J. Eur. Math. Soc., 10(1):1–45, 2008]) to a deterministic setting. Among the main ingredients are a new Diophantine estimate of quasi-periodic functions in arbitrary-dimensional phase space, and the application of Bourgain's geometric lemma in [Geom. Funct. Anal., 17(3):682–706, 2007].

1. Introduction and Main Result

Let

$$\begin{aligned}\boldsymbol{\alpha} &= (\alpha_1, \cdots, \alpha_d) \in [0,1]^d,\\ \boldsymbol{\theta} &= (\theta_1, \cdots, \theta_d) \in [0,1]^d,\end{aligned}$$

and consider the nonlinear Schrödinger equation with a quasi-periodic potential on $\mathbb{Z}^d$:

$$iu_t + (\varepsilon\Delta + V(\boldsymbol{\theta} + \boldsymbol{n}\boldsymbol{\alpha})\delta_{\boldsymbol{n},\boldsymbol{n}'})u + \delta|u|^{2p}u = 0, \tag{1.1}$$

where ε and δ are parameters in $[0,1]$, $\boldsymbol{n} \in \mathbb{Z}^d$, $p \in \mathbb{N}$ and $\Delta(\boldsymbol{n}, \boldsymbol{n}') = \delta_{|\boldsymbol{n}-\boldsymbol{n}'|_1,1}$ denotes the adjacency Laplacian with $|\boldsymbol{n}|_1 := \sum_{\ell=1}^{d} |n_\ell|$ and

$$\boldsymbol{n}\boldsymbol{\alpha} = (n_1\alpha_1, \cdots, n_d\alpha_d) \in \mathbb{R}^d;$$

the potential V is a trigonometric polynomial:

$$V(\boldsymbol{\theta}) = \sum_{\boldsymbol{\ell}\in\Gamma_K} v_{\boldsymbol{\ell}} \cos 2\pi(\boldsymbol{\ell}\cdot\boldsymbol{\theta}),\ \boldsymbol{\ell}\cdot\boldsymbol{\theta} = \sum_{s=1}^{d} \ell_s\theta_s,$$

with $\Gamma_K \subset [-K, K]^d \setminus \{\mathbf{0}\},\ K \geq 1$ and

$$\boldsymbol{v} = (v_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\Gamma_K} \in \mathbb{R}^{\#\Gamma_K} \setminus \{\mathbf{0}\},$$

where $\#(\cdot)$ denotes the cardinality of a set. We further assume that the set Γ_K is *maximal* so that:

$$\begin{aligned}&\text{if } \boldsymbol{\ell} \in \Gamma_K, \text{then}, \ell_s \neq 0 \text{ for } \forall\ 1 \le s \le d;\\ &\text{and if } \ \boldsymbol{\ell}, \boldsymbol{\ell}' \in \Gamma_K, \ \text{then } \boldsymbol{\ell} + \boldsymbol{\ell}' \neq \mathbf{0}.\end{aligned} \tag{1.2}$$

We assume that V is nondegenerate: For any fixed $1 \le \ell \le d$ and $(\theta_s)_{s\neq\ell}$, $f(\theta_\ell) := V((\theta_s)_{s\neq\ell}, \theta_\ell)$ is not a constant function in θ_ℓ. Note that the first condition in (1.2) then says that, in fact, each term in the polynomial is nondegenerate.

When $\delta = 0$, it is known from the breakthrough paper [Bou07] (cf. also the recent refinement [JLS20]) that the linear operator

$$\mathcal{H} = \varepsilon\Delta + V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})\delta_{\boldsymbol{n},\boldsymbol{n}'} \text{ on } \mathbb{Z}^d, \tag{1.3}$$

exhibits Anderson localization, namely pure point spectrum with exponentially decaying eigenfunctions, for small ε, on a large set in $(\boldsymbol{\alpha}, \boldsymbol{\theta})$. It follows that when $0 < \varepsilon \ll 1$, the linear Schrödinger equation

$$iu_t + (\varepsilon\Delta + V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})\delta_{\boldsymbol{n},\boldsymbol{n}'})u = 0 \tag{1.4}$$

has only Anderson localized states, that is, roughly speaking, wave packets localized about the origin remain localized for all time.

The present paper addresses the persistence question when $\delta \neq 0$, namely the existence of Anderson-localized-type solutions for the nonlinear Schrödinger equation (1.1) when $0 < \varepsilon, \delta \ll 1$, under appropriate conditions on $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. Since both ε and δ are small, we start from the (decoupled) equation

$$iu_t + V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})\delta_{\boldsymbol{n},\boldsymbol{n}'}u = 0. \tag{1.5}$$

It has solutions of the form

$$u^{(0)}(t, \boldsymbol{n}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} e^{i\mu_{\boldsymbol{n}} t}, \tag{1.6}$$

where

$$\mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})$$

and $a_{\boldsymbol{n}}$ decays rapidly. The solutions (1.6), in general, have infinite number of frequencies and are *almost-periodic* in time.

We study the persistence of the above type of solutions with finite (but arbitrary) number of frequencies in time, that is, the *quasi-periodic* in time solutions. Denoting the frequency by $\boldsymbol{\omega}$, $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_b)$, as an ansatz, we seek solutions to (1.1) in the form of a convergent series:

$$u(t, \boldsymbol{n}) = \sum_{(\boldsymbol{k},\boldsymbol{n})\in\mathbb{Z}^b\times\mathbb{Z}^d} \hat{u}(\boldsymbol{k}, \boldsymbol{n}) e^{i\boldsymbol{k}\cdot\boldsymbol{\omega} t}. \tag{1.7}$$

Note that solutions of the above form are consistent with the nonlinearity in (1.1).

Denote by $\mathrm{meas}(\cdot)$ the Lebesgue measure of a set. Let $|\cdot|$ denote the supremum norm. We have

Theorem 1.1. *Fix (any) distinct $\boldsymbol{n}_1, \cdots, \boldsymbol{n}_b \in \mathbb{Z}^d$ and let*

$$u^{(0)}(t, \boldsymbol{n}) = \sum_{\ell=1}^{b} a_\ell e^{i\omega_\ell^{(0)} t} \delta_{\boldsymbol{n}_\ell,\boldsymbol{n}} \tag{1.8}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

be a solution to (1.5), where $\boldsymbol{a} = (a_\ell)_{\ell=1}^b \in [1,2]^b$ *and*

$$\boldsymbol{\omega}^{(0)} = (\omega_\ell^{(0)})_{\ell=1}^b = (V(\boldsymbol{n}_\ell \boldsymbol{\alpha} + \boldsymbol{\theta}))_{\ell=1}^b.$$

Then for $0 < \varepsilon \le \delta \le \log^{-1}\frac{1}{\varepsilon} \le \delta_0(b,d,K,V,\max_{1\le\ell\le b}|\boldsymbol{n}_\ell|) \ll 1$, *there is a set* $\mathcal{W} \subset [0,1]^{2d}$ *satisfying* $\operatorname{meas}([0,1]^{2d} \setminus \mathcal{W}) \le \log^{-c}\frac{1}{\varepsilon+\delta}(c = c(b,d) > 0)$ *such that, for any* $(\boldsymbol{\alpha},\boldsymbol{\theta}) \in \mathcal{W}$, *there exists a diffeomorphism* $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{a})$ *on* $[0,1]^b$ *and some* $\mathcal{R} \subset [1,2]^b$ *of* $\operatorname{meas}([1,2]^b \setminus \mathcal{R}) \le (\varepsilon+\delta)^c$ *so that the following holds true: If* $\boldsymbol{a} \in \mathcal{R}$ *and* $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{a})$, *then* $|\boldsymbol{\omega} - \boldsymbol{\omega}^{(0)}| \lesssim \delta$ *and*

$$u(t,\boldsymbol{n}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^b} \hat{u}(\boldsymbol{k},\boldsymbol{n}) e^{i\boldsymbol{k}\cdot\boldsymbol{\omega}t}$$

is a solution to (1.1). Moreover, one has

$$\hat{u}(\boldsymbol{e}_\ell,\boldsymbol{n}_\ell) = a_\ell,\ 1 \le \ell \le b,$$
$$\sum_{(\boldsymbol{k},\boldsymbol{n})\notin\mathcal{S}_+} |\hat{u}(\boldsymbol{k},\boldsymbol{n})| e^{|\boldsymbol{k}|+|\boldsymbol{n}|} < \sqrt{\varepsilon+\delta},$$

where $\mathcal{S}_+ = \{(\boldsymbol{e}_\ell,\boldsymbol{n}_\ell)_{\ell=1}^b\}$ *with* $\boldsymbol{e}_\ell$ $(1 \le \ell \le b)$ *the standard basis vectors for* $\mathbb{Z}^b$ *(i.e.,* $\boldsymbol{e}_\ell = (\delta_{\ell,j})_{j=1}^b$*).*

Remark 1.2. Theorem 1.1 readily generalizes to potentials V, which are given by finite Fourier series with both cosine and sine terms.

Remark 1.3. Note that we do not require the linear equation (1.4) to satisfy Anderson localization. However, linear Anderson localization does hold on a large subset of $\mathcal{W}$ (cf. [Bou07, JLS20]).

1.1. Ideas of the proof

The multi-dimensional phase space, namely $\boldsymbol{\theta}$ being a vector instead of a scalar as in, for example, [SW23] poses a major challenge. Diophantine properties of the eigenvalues of the linear Schrödinger operator (1.3) play an essential role in the proof. To $O(\delta)$, this could be replaced by Diophantine properties of the potential $V(\boldsymbol{n}\boldsymbol{\alpha}+\boldsymbol{\theta})$ at different lattice sites $\boldsymbol{n} \in \mathbb{Z}^d$. We need to show that the values of the function at these chosen sites are, in some sense, *linearly independent*. This is, however, a priori, not obvious, since V is a *given function* on $\mathbb{R}^d$. To prove linear independence, we use a generalized Wronskian approach and bound the determinant away from zero by carefully removing some $(\boldsymbol{\alpha},\boldsymbol{\theta})$. This method is applicable to *any* trigonometric polynomials, generalizing the approach in [SW23] for the cosine function on $\mathbb{R}$. It is independent of the main body of the proof and could be of independent interest. Such arguments may be closely related to Diophantine approximations on manifold developed by Kleinbock and Margulis [KM98].

The other main ingredient of the proof is the application of Bourgain's approach in the study of Anderson localization for linear quasi-periodic Schrödinger operators in d-dimensions [Bou07] (e.g., the $\mathcal{H}$ in (1.3)). Bourgain's approach seems to be, so far, the only method available to deal with the d-dimensional phase space, that is, $\boldsymbol{\theta} \in [0,1]^d$, $d \ge 3$. Among other innovations, it uses semi-algebraic geometry arguments (e.g., Yomdin-Gromov algebraic lemma) to control the resonances. The difficulty of d-dimensional phase space manifests as well in our paper, in that for the linear analysis, semi-algebraic geometry already seems indispensable to control the resonances; while in previous works, see [BW08, Wan16, LW24, SW23], this role is filled by Diophantine properties. To control the resonances in time, we also need weak second Melnikov-type estimates. This could be readily obtained, similar to [Bou05a, BW08, LW24, SW23]. In this paper, however, we have to overcome much more serious resonances in space (cf. Figure 1), which seems novel. The proof is then accomplished by doing multi-scale analysis with some new ideas.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

The general problem discussed here is the persistence of quasi-periodic solutions of linear or integrable equations after Hamiltonian perturbation. This subject is closely related to the well-known "KAM theory" of invariant tori in smooth dynamical systems. Results along this line were first obtained by Kuksin [Kuk87] and Wayne [Way90], stimulating considerable progress, see, for example, [KP96, CY00, BG01, Bou05b, EK10, LY10, LY11, GT11, BBP13, PP15, EGK16, BBHM18, BKM18, CLSY18, Yua21, BHM23] to name but a few. The existence of quasi-periodic solutions can also be obtained using the direct Craig-Wayne-Bourgain method [CW93, Bou94, Bou98, Bou05a] (cf. [BW08, BB13, Wan16, HSSY20, Wan21a, Wan21b, LW24, SW23, KLW24] for more recent results), which turns out to be more robust when dealing with multi-dimensional Hamiltonian PDEs (cf. Chapters 18–20 of [Bou05a]).

Finally, we mention the work [Yua02] in which a KAM scheme was first developed to prove the existence of quasi-periodic solutions for some nonlinear discrete equations without external parameters (i.e., regarding the initial states as parameters). Later in [GYZ14], the authors also applied the KAM method to obtain the existence of quasi-periodic solutions for the nonlinear quasi-periodic Schrödinger equation on $\mathbb{Z}$.

1.2. The nonlinear random Schrödinger equation

When the potential V, with $V(\boldsymbol{n}) = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})$, is replaced by a random potential, for example, with $V(\boldsymbol{n})$ a family of independent identically distributed random variables (with the uniform distribution), it is known from [BW08] that the nonlinear random Schrödinger equation

$$iu_t + \varepsilon\Delta u + Vu + \delta|u|^{2p}u = 0 \text{ on } \mathbb{Z}^d$$

has large sets of Anderson localized states for small ε and δ. When $d = 1$, it is shown further in the important work [LW24] that large sets of Anderson localized states persist for all $\varepsilon \neq 0$ (similar results could be proven for the nonlinear wave equation). Recall that the first proof of the Anderson localization for the linear random Schrödinger equation

$$iu_t + \varepsilon\Delta u + Vu = 0 \text{ on } \mathbb{Z}^d$$

for small ε was based on multi-scale-analysis-type Green's function estimates developed in [FS83], see also [AM93].

It is generally believed that the linear random Schrödinger equation and the linear quasi-periodic Schrödinger equation should share common localization features in the perturbative regime, despite the quasi-periodic problems being more delicate and the results more difficult to obtain. The use of semi-algebraic geometry and Cartan's Lemma [BGS02, Bou07, JLS20], for example, certainly renders the quasi-periodic issue more involved than that of the random, and this seems to continue to the nonlinear case, as mentioned earlier. Theorem 1.1 generalizes this circle of ideas to the nonlinear setting by providing a concrete example where the quasi-periodic and the random have indeed similar behavior.

1.3. Structure of the paper

In Section 2, we focus on the linear estimates, in the form of a large deviation theorem. The paper concludes by constructing the quasi-periodic in-time solutions in Section 3. The proof of the Diophantine-type estimates and some important lemmas on resolvent identities are presented in the Appendix.

2. Linear analysis

In this section, we investigate the properties of linearized operators on $\mathbb{Z}^{b+d}$. Of particular importance is the *large deviation theorem (LDT)* for the corresponding Green's functions.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

2.1. *The operator on the lattice*

For a vector $\boldsymbol{x} \in \mathbb{R}^r$, we denote by $|\boldsymbol{x}|$ (resp. $|\boldsymbol{x}|_2$) the supremum norm (resp. the Euclidean norm). For an operator (or a matrix), we denote by $\|\cdot\|$ its operator norm.

We define the lattice:

$$\mathbb{Z}^{b+d}_{\mathrm{pm}} = \mathbb{Z}^{b+d} \times \{+,-\}.$$

Let $\mathcal{S}_+ = \{(\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)_{\ell=1}^b\} \subset \mathbb{Z}^{b+d}$ and we identify $\mathcal{S}_+$ with $\{(\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)_{\ell=1}^b\} \times \{+\} \subset \mathbb{Z}^{b+d}_{\mathrm{pm}}$. Similarly, let $\mathcal{S}_- = \{(-\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)_{\ell=1}^b\} \subset \mathbb{Z}^{b+d}$, which is then identified with $\{(-\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)_{\ell=1}^b\} \times \{-\} \subset \mathbb{Z}^{b+d}_{\mathrm{pm}}$. We let $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_-$ and

$$\mathbb{Z}^{b+d}_{\mathrm{pm},*} = \mathbb{Z}^{b+d}_{\mathrm{pm}} \setminus \mathcal{S}.$$

In the following, we study the operator on $\mathbb{Z}^{b+d}_{\mathrm{pm},*}$ the operator

$$H(\sigma) = D(\sigma) + \varepsilon(\Delta \oplus \Delta) + \delta S,\ \sigma \in \mathbb{R}, \tag{2.1}$$

where

$$D(\sigma) = \mathrm{diag}_{(\boldsymbol{k},\boldsymbol{n}) \in \mathbb{Z}^{b+d}} \begin{pmatrix} -\sigma - \boldsymbol{k}\cdot\boldsymbol{\omega} + \mu_{\boldsymbol{n}} & 0 \\ 0 & \sigma + \boldsymbol{k}\cdot\boldsymbol{\omega} + \mu_{\boldsymbol{n}} \end{pmatrix},\ \mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})$$

and the operator S refers typically to the linearized operator of the nonlinear perturbation. So we may let S satisfy the following properties:

- First

$$S:\ \ell^2(\mathbb{Z}^{b+d}_{\mathrm{pm}}) \to \ell^2(\mathbb{Z}^{b+d}_{\mathrm{pm}})$$

  is a bounded self-adjoint operator.
- For all $\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{k}'' \in \mathbb{Z}^b$, $\boldsymbol{n}, \boldsymbol{n}' \in \mathbb{Z}^d$ and $\xi, \xi' \in \{+,-\}$, we have the Töplitz property in the $\boldsymbol{k}$-variable:

$$S((\boldsymbol{k}'+\boldsymbol{k}, \boldsymbol{n}, \xi); (\boldsymbol{k}''+\boldsymbol{k}, \boldsymbol{n}', \xi')) = S((\boldsymbol{k}', \boldsymbol{n}, \xi); (\boldsymbol{k}'', \boldsymbol{n}', \xi')).$$

- We have for some $C_2 > 0$ and $\gamma \in (\frac{1}{2}, 10)$,

$$|S((\boldsymbol{k}, \boldsymbol{n}, \xi); (\boldsymbol{k}', \boldsymbol{n}', \xi'))| \leq C_2(1 + |\boldsymbol{k} - \boldsymbol{k}'|)^{C_2} e^{-\gamma|\boldsymbol{k}-\boldsymbol{k}'| - \gamma|\boldsymbol{n}|} \delta_{\boldsymbol{n},\boldsymbol{n}'}.$$

The frequency $\boldsymbol{\omega}$ satisfies

$$\boldsymbol{\omega} \in \Omega = \boldsymbol{\omega}^{(0)} + [-C\delta, C\delta]^b, \tag{2.2}$$

where $C > 0$ and $\boldsymbol{\omega}^{(0)} = \boldsymbol{\omega}^{(0)}(\boldsymbol{\alpha}, \boldsymbol{\theta})$ is defined in Theorem 1.1.

2.2. *The (generalized) elementary regions on the lattice*

For some technical reasons, we will introduce the definitions of *elementary regions* and *generalized elementary regions* on $\mathbb{Z}^{b+d}$ originating from [BGS02, Bou07]. We refer to [Liu22] (cf. Section 2 in [Liu22]) for some important clarifications and improvements on those concepts.

Denote by $\Lambda_N(\boldsymbol{x})$ ($\boldsymbol{x} \in \mathbb{Z}^r,\ r = b+d, d$),

$$\Lambda_N(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{Z}^r :\ |\boldsymbol{y} - \boldsymbol{x}| \leq N\},$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

the cube of radius N centered at $\boldsymbol{x}$. In particular, we write $\Lambda_N = \Lambda_N(\mathbf{0})$. For $\Lambda \subset \mathbb{Z}^r$, we define its diameter to be $\operatorname{diam}\Lambda = \sup_{\boldsymbol{x},\boldsymbol{y}\in\Lambda} |\boldsymbol{x}-\boldsymbol{y}|$. For $\boldsymbol{x} \in \mathbb{Z}^r$ and $\boldsymbol{w} \in \mathbb{R}^r_+$, let $R_{\boldsymbol{w}}(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{Z}^r : |y_\ell - x_\ell| \le w_l,\ 1 \le \ell \le r\}$ denote the rectangle.

A *generalized elementary region* is defined to be a set Λ of the form

$$\Lambda = R_{\boldsymbol{w}}(\boldsymbol{x}) \setminus (R_{\boldsymbol{w}}(\boldsymbol{x}) + z),$$

where $z \in \mathbb{Z}^r$ is arbitrary. The size of a generalized elementary region is simply its diameter. The set of all *generalized elementary regions* of size at most M will be denoted by $\mathcal{E}_M$.

An *elementary region* Q_N of size N and centered at $\mathbf{0}$ is one of the following regions

$$Q_N = \Lambda_N \text{ or } Q_N = \Lambda_N \setminus \{\boldsymbol{x} \in \mathbb{Z}^r : x_\ell \square_\ell 0,\ 1 \le \ell \le r\},$$

where $\square_\ell \in \{<, >, \emptyset\}$ and at least two $\square_\ell$'s are not $\emptyset$. Denote by $\mathcal{ER}_{\mathbf{0}}(N)$ the set of all elementary regions of size N and centered at $\mathbf{0}$. Let

$$\mathcal{ER}(N) = \{\boldsymbol{x} + Q_N : \boldsymbol{x} \in \mathbb{Z}^r,\ Q_N \in \mathcal{ER}_{\mathbf{0}}(N)\}.$$

Let $\mathcal{E}_M^L$ denote the set of all M-size generalized elementary regions of width at least $L \le M$, namely, $\Lambda \in \mathcal{E}_M^L$ iff $\Lambda \in \mathcal{E}_M$ and, for any $\boldsymbol{x} \in \Lambda, 0 < L' < L$, there is some $\Lambda' \in \mathcal{ER}(L')$ so that $\boldsymbol{x} \in \Lambda' \subset \Lambda$, $\operatorname{dist}(\boldsymbol{x}, \Lambda \setminus \Lambda') \ge L'/2$, where $\operatorname{dist}(\cdot,\cdot)$ is induced by the supremum norm. So we have $\mathcal{ER}(N) \subset \mathcal{E}_{2N}^N$.

With a slight abuse of notation, we also use $\mathcal{ER}_{\mathbf{0}}(N), \mathcal{ER}(N), \mathcal{E}_N, \mathcal{E}_N^L, \Lambda_N(\boldsymbol{x}),\ \Lambda_N$ to denote $\mathcal{ER}_{\mathbf{0}}(N) \times \{+,-\}, \mathcal{ER}(N) \times \{+,-\}, \mathcal{E}_N \times \{+,-\}, \mathcal{E}_N^L \times \{+,-\}, \Lambda_N(\boldsymbol{x}) \times \{+,-\}, \Lambda_N \times \{+,-\}$. Similarly, for any $\Lambda \subset \mathbb{Z}^r$, denote by R_Λ the restriction to $\Lambda \times \{+,-\}$.

2.3. Green's functions and LDE

We now define Green's functions on subsets of $\mathbb{Z}^{b+d}_{\mathrm{pm},*}$. Let $\Lambda \subset \mathbb{Z}^{b+d}_{\mathrm{pm},*}$ be a nonempty set, R_Λ the restriction operator to Λ, and define the $(\#\Lambda) \times (\#\Lambda)$ matrix:

$$(R_\Lambda H(\sigma) R_\Lambda)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi')) = H(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi')),\ (\boldsymbol{x},\xi),(\boldsymbol{x}',\xi') \in \Lambda, \tag{2.3}$$

where $H(\sigma)$ is defined by (2.1). Define the Green's function to be (if it exists)

$$G_\Lambda(\sigma) = (R_\Lambda H(\sigma) R_\Lambda)^{-1}.$$

Before addressing the LDT for Green's functions, let us first give the definition of large deviation estimates (LDE):

Definition 2.1 (LDE). Let $\rho > 0$, $0 < \gamma' < \gamma$ and $M \in \mathbb{N}$. We say $H(\sigma)$ defined by (2.1) satisfies the (ρ, γ', M)-LDE if there exists a set $\Sigma_M \subset \mathbb{R}$ with

$$\operatorname{meas}(\Sigma_M) \le e^{-M^\rho},$$

so that for $\sigma \notin \Sigma_M$ the following estimates hold: If $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(M)$ satisfies $|\boldsymbol{n}| \le 10M$, then

$$\begin{gathered}\|G_\Lambda(\sigma)\| \le e^{M^{\frac{3}{4}}},\\ |G_\Lambda(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| \le e^{-\gamma'|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'| \ge M^{\frac{8}{9}} \text{ and } \xi,\xi' \in \{\pm\},\end{gathered}$$

where $\boldsymbol{x} = (\boldsymbol{k}', \boldsymbol{n}'), \boldsymbol{x}' = (\boldsymbol{k}'', \boldsymbol{n}'')$.

We will show in this section that the $(\rho, \frac{\gamma}{2}, N)$-LDE hold for some small

$$\rho = \rho(b, d) > 0$$

and all $N \gg 1$, under certain nonresonant conditions on $\boldsymbol{\alpha}, \boldsymbol{\theta}$ and $\boldsymbol{\omega}$. This leads to the LDT, Theorem 2.11, below. Since we view $\boldsymbol{\omega}$ as an $O(\delta)$ perturbation of $\boldsymbol{\omega}^{(0)}$, the properties of $\boldsymbol{\omega}^{(0)}$ become essential in the proof of LDT. So, in the following, we first establish nonresonant properties (called Diophantine estimates) of $\boldsymbol{\omega}^{(0)}$. Then the proof of LDT will follow from a multi-scale analysis scheme in the spirit of [BGS02, Bou05a, Bou07].

2.4. Diophantine estimates

Recall that

$$\begin{aligned} \boldsymbol{\omega}^{(0)} &= \boldsymbol{\omega}^{(0)}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = (V(\boldsymbol{n}_\ell \boldsymbol{\alpha} + \boldsymbol{\theta}))_{\ell=1}^{b}, \\ \mu_{\boldsymbol{n}} &= \mu_{\boldsymbol{n}}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta}), \ \boldsymbol{n} \in \mathbb{Z}^d. \end{aligned}$$

The main purpose of this section is to obtain lower bounds on

$$\begin{aligned} &|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)}| \text{ for } \boldsymbol{k} \in \mathbb{Z}^b \setminus \{\boldsymbol{0}\}, \\ &| \pm \boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}}| \text{ for } (\boldsymbol{k}, \boldsymbol{n}) \in \mathbb{Z}^{b+d} \setminus \mathcal{S}, \\ &|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}} - \mu_{\boldsymbol{n}'}| \text{ for } \boldsymbol{n} \neq \boldsymbol{n}' \in \mathbb{Z}^d, \end{aligned}$$

under certain restrictions on $(\boldsymbol{\alpha}, \boldsymbol{\theta})$. As mentioned in the Introduction, due to the d-dimensional phase space, such Diophantine-type estimates are highly nontrivial. We use a generalized Wronskian approach and work out the details in the Appendix (cf. Theorem A.2 and Corollary A.4). These estimates can be read independently, and are key to the existence of quasi-periodic solutions to the nonlinear equation.

Applying Corollary A.4 then leads to

Theorem 2.2. *There exist some* $0 < c_1 = c_1(b, d, K, V) < \frac{1}{100b}, C_1 = C_1(b, d, K, V) > 1$ *such that, if* $0 < \varepsilon + \delta \leq \delta_0(b, d, K, V, \max_{1 \leq \ell \leq b} |\boldsymbol{n}_\ell|) \ll 1$*, then there is some* $\mathcal{M} \subset [0, 1]^{2d}$ *satisfying*

$$\operatorname{meas}([0, 1]^{2d} \setminus \mathcal{M}) \leq \log^{-1} \frac{1}{\varepsilon + \delta}, \tag{2.4}$$

so that the following properties hold true for $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{M}$ *and* $L_{\varepsilon,\delta} := 100(\varepsilon + \delta)^{-c_1}$.

(1) *For all* $\log \frac{1}{\varepsilon+\delta} \leq L \leq L_{\varepsilon,\delta}$ *and all* $(\boldsymbol{k}, \boldsymbol{n}) \in \Lambda_L \setminus \mathcal{S}$, *we have*

$$\min_{\xi = \pm 1} |\xi \boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}}| > L^{-C_1}. \tag{2.5}$$

(2) *We have for any* $\xi = \pm 1$,

$$\sup_{\sigma \in \mathbb{R}} \# \left\{ (\boldsymbol{k}, \boldsymbol{n}) \in \Lambda_{L_{\varepsilon,\delta}} : \ |\xi(\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)}) + \mu_{\boldsymbol{n}}| < \frac{(\varepsilon + \delta)^{\frac{1}{8b}}}{4} \right\} \leq b. \tag{2.6}$$

Remark 2.3. The estimate in (1) plays an essential role in the initial iteration steps for the nonlinear analysis. The conclusion (2) of this theorem will be only used in the proof of intermediate scales LDE.

Proof. The proof is similar to that in [SW23]. First, the measure estimate (2.4) is a consequence of Corollary A.4. Next, the inequality (2.5) in (1) follows directly from (iii) of Corollary A.4.

So, it suffices to establish (2.6) of (2). We prove it by the contradiction. Without loss of generality, we consider the case $\xi = +1$. Assume that there are $b+1$ distinct $\{(\boldsymbol{k}_\ell, \boldsymbol{n}'_\ell)\}_{1\le\ell\le b+1}$ satisfying

$$|\sigma + \boldsymbol{k}_\ell \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}'_\ell}| < \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{4}. \tag{2.7}$$

We claim that all $\boldsymbol{k}_\ell$ ($1 \le \ell \le b+1$) are distinct. Actually, if there are $\boldsymbol{k}_{\ell_1} = \boldsymbol{k}_{\ell_2}$ for some $1 \le \ell_1 \ne \ell_2 \le b+1$, then it must be that $\boldsymbol{n}'_{\ell_1} \ne \boldsymbol{n}'_{\ell_2}$. As a result, we obtain using (2.7) that

$$\begin{aligned}|\mu_{\boldsymbol{n}'_{\ell_1}} - \mu_{\boldsymbol{n}'_{\ell_2}}| &\le |\sigma + \boldsymbol{k}_{\ell_1} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}'_{\ell_1}}| \\ &\quad + |\sigma + \boldsymbol{k}_{\ell_2} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}'_{\ell_2}}| \\ &\le \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{2},\end{aligned}$$

which contradicts (i) of Corollary A.4. Next, we claim that $\{\boldsymbol{n}'_\ell\}_{1\le\ell\le b+1} \subset \{\boldsymbol{n}_\ell\}_{1\le\ell\le b}$. In fact, if $\boldsymbol{n}'_{\ell_1} \notin \{\boldsymbol{n}_\ell\}_{1\le\ell\le b}$ for some $1 \le \ell_1 \le b+1$, we can assume that there is $\ell_2 \ne \ell_1$ so that $\boldsymbol{n}'_{\ell_2} \ne \boldsymbol{n}'_{\ell_1}$. Otherwise, we may have

$$|(\boldsymbol{k}_{\ell_1} - \boldsymbol{k}_{\ell_2}) \cdot \boldsymbol{\omega}^{(0)}| \le \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{2},$$

which contradicts (ii) of Corollary A.4, since $\boldsymbol{k}_{\ell_1} \ne \boldsymbol{k}_{\ell_2}$ as claimed above. So we obtain for some $\ell_2 \ne \ell_1$ and $1 \le \ell_2 \le b+1$ with $\boldsymbol{n}'_{\ell_2} \ne \boldsymbol{n}'_{\ell_1}$ that

$$|(\boldsymbol{k}_{\ell_1} - \boldsymbol{k}_{\ell_2}) \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}'_{\ell_1}} - \mu_{\boldsymbol{n}'_{\ell_2}}| \le \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{2},$$

which contradicts (iv) of Corollary A.4. So it suffices to assume $\{\boldsymbol{n}'_\ell\}_{1\le\ell\le b+1} \subset \{\boldsymbol{n}_\ell\}_{1\le\ell\le b}$. However, in this case, we have by the pigeonhole principle that there are $1 \le \ell_1 \ne \ell_2 \le b+1$ so that $\boldsymbol{n}'_{\ell_1} = \boldsymbol{n}'_{\ell_2}$. Then we get

$$|(\boldsymbol{k}_{\ell_1} - \boldsymbol{k}_{\ell_2}) \cdot \boldsymbol{\omega}^{(0)}| \le \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{2},$$

which contradicts (ii) of Corollary A.4 as shown above. This proves (2.6). □

2.5. *Bourgain's geometric lemma*

To proceed with the analysis, we will also need Anderson localization-type properties for the linear quasi-periodic Schrödinger operator $\mathcal{H}$ in (1.3), namely,

$$\mathcal{H} = \mathcal{H}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \varepsilon\Delta + V(\boldsymbol{\theta} + \boldsymbol{n}\boldsymbol{\alpha})\delta_{\boldsymbol{n},\boldsymbol{n}'}.$$

This was proven by Bourgain in the remarkable work [Bou07], where he established Anderson localization for general quasi-periodic Schrödinger operators on $\mathbb{Z}^d$ with nondegenerate analytic potentials on $\mathbb{T}^d$ (identified with $[0,1]^d$). This result was later significantly extended by Jitomirskaya-Liu-Shi [JLS20] (cf. also [Shi22, Liu22]) to potentials on $\mathbb{T}^b$ for $b \ge d$.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

In the present paper, due to the d-dimensional phase space, we need to apply Bourgain's arguments. Using short-range property of S in the $\boldsymbol{n}$-variable, Bourgain's geometric lemma is made available to handle LDE for Green's functions at large scales. *This is a main new aspect not present in [SW23]*. The general scheme is similar to that in Chapter 19 of [Bou05a], where Bourgain made a great breakthrough and first proved the existence of quasi-periodic solutions for the nonlinear Schrödinger equation on arbitrarily dimensional torus (a KAM theorem was later established by Eliasson-Kuksin in an important paper [EK10]). Among others, Bourgain [Bou05a] employed heavily the *separation property* of eigenvalues of the Laplace operator on the higher dimensional torus, while the *separation property* is missing in the present setting (indeed, the spectrum of $\mathcal{H}$ may contain an interval).

We start with introducing the definition of semi-algebraic sets.

Definition 2.4. A set $X \subset \mathbb{R}^r$ 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 $\{P_1, \cdots, P_k\} \subset \mathbb{R}[x_1, \cdots, x_r]$ be a family of real polynomials whose degrees are bounded by p. A (closed) semi-algebraic set X is given by an expression

$$X = \bigcup_{s} \bigcap_{\ell \in \mathcal{K}_s} \{\boldsymbol{x} \in \mathbb{R}^r : \ P_\ell(\boldsymbol{x}) \varsigma_{s\ell} 0\}, \tag{2.8}$$

where $\mathcal{K}_s \subset \{1, \cdots, k\}$ and $\varsigma_{s\ell} \in \{\geq, \leq, =\}$. Then we say that X has degree at most kp. In fact, the degree of X which is denoted by $\deg X$, means the smallest kp over all representations as in (2.8).

Denote by $\mathcal{ER}_{\mathbb{Z}^d}(M)$ (resp. $\mathcal{ER}_{\mathbb{Z}^d,\boldsymbol{0}}(M)$) the set of all M-size elementary regions on $\mathbb{Z}^d$ (resp. centered at $\boldsymbol{0}$). We have

Lemma 2.5 (cf. the Claim (page 694) in [Bou07] and Theorem 2.7 in [JLS20]). *There are small constants $0 < \kappa_1 < \kappa_2 < 1$ depending only on d so that the following holds true. For $0 < \varepsilon \leq \varepsilon_0(V, d) \ll 1$ and $N \geq N_1 := 10^{-\frac{32}{\kappa_1^2}} \log^{\frac{4}{\kappa_1}}(\log \frac{1}{\varepsilon})$, there is a semi-algebraic set $\mathcal{A}_N \subset [0,1]^d$ satisfying*

$$\deg \mathcal{A}_N \leq N^{4d}, \ \operatorname{meas}\Big(\bigcap_{N \geq N_1} \mathcal{A}_N\Big) = 1 - O\Big(\log^{-c} \frac{1}{\varepsilon}\Big), \ c = c(d) > 0,$$

so that the following properties hold true for $\boldsymbol{\alpha} \in \mathcal{A}_N$.

(1) *For all $E \in \mathbb{R}$ and all $\boldsymbol{\theta} \in [0,1]^d$, there is $L \in [N^{\kappa_1}, N^{\kappa_2}]$ so that, for all $Q \in \mathcal{ER}_{\mathbb{Z}^d}(L_1)$ satisfying $L_1 \sim (\log N)^{\frac{4}{\kappa_1}}$ and*

$$Q \subset [-L, L]^d \setminus [-L^{\frac{1}{10d}}, L^{\frac{1}{10d}}]^d,$$

one has

$$\|\mathcal{H}_Q^{-1}(E; \boldsymbol{\theta})\| \leq e^{\sqrt{L_1}},$$
$$|\mathcal{H}_Q^{-1}(E; \boldsymbol{\theta})(\boldsymbol{n}; \boldsymbol{n}')| \leq e^{-\frac{1}{2}|\log \varepsilon| \cdot |\boldsymbol{n} - \boldsymbol{n}'|} \text{ for } |\boldsymbol{n} - \boldsymbol{n}'| \geq L_1^{\frac{8}{9}},$$

where

$$\begin{aligned}\mathcal{H}_Q(E; \boldsymbol{\theta}) &= R_Q\big(\varepsilon\Delta + V(\boldsymbol{\theta} + \boldsymbol{n}\boldsymbol{\alpha})\delta_{\boldsymbol{n},\boldsymbol{n}'}\big)R_Q - E \\ &:= \mathcal{L}_Q(\boldsymbol{\theta}) - E.\end{aligned}$$

(2) *There is some $\Theta_N = \Theta_N(\boldsymbol{\alpha}) \subset [0,1]^d$ satisfying*

$$\operatorname{meas}([0,1]^d \setminus \Theta_N) \leq e^{-N^{\frac{\kappa_1}{4}}}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

so that, if $\boldsymbol{\theta} \in \Theta_N$, then one has for all $\Lambda \in \mathcal{ER}_{\mathbb{Z}^d,\mathbf{0}}(N)$,

$$\|\mathcal{H}_\Lambda^{-1}(\boldsymbol{\alpha},\boldsymbol{\theta})\| \le e^{\sqrt{N}},$$
$$|\mathcal{H}_\Lambda^{-1}(\boldsymbol{\alpha},\boldsymbol{\theta})(\boldsymbol{n};\boldsymbol{n}')| \le e^{-\frac{1}{2}|\log\varepsilon|\cdot|\boldsymbol{n}-\boldsymbol{n}'|} \text{ for } |\boldsymbol{n}-\boldsymbol{n}'| \ge N^{\frac{8}{9}}.$$

Denote

$$\mathcal{A} = \bigcap_{N\ge N_1}\mathcal{A}_N,\ \tilde{\Theta}_N(\boldsymbol{\alpha}) = \bigcap_{|\boldsymbol{n}|\le e^{N^{\frac{\kappa_1}{5}}}}\{\boldsymbol{\theta} : \ \boldsymbol{\theta}+\boldsymbol{n}\boldsymbol{\alpha}\in\Theta_N\},$$
$$\mathcal{W}' = \left\{(\boldsymbol{\alpha},\boldsymbol{\theta})\in[0,1]^{2d} : \ \boldsymbol{\alpha}\in\mathcal{A},\ \boldsymbol{\theta}\in\bigcap_{N\ge N_1}\tilde{\Theta}_N(\boldsymbol{\alpha})\right\}. \tag{2.9}$$

Then

$$\mathrm{meas}(\mathcal{W}') \ge 1 - \log^{-c}\frac{1}{\varepsilon}. \tag{2.10}$$

Remark 2.6. In this lemma, the constants κ_1, κ_2 correspond to c_3, c_4 of [JLS20], respectively. Originally, the off-diagonal exponential decay distance (cf., e.g., decay estimates in (2) of Lemma 2.5) in [Bou07, JLS20] is $|\boldsymbol{n}-\boldsymbol{n}'| \ge \frac{N}{10}$. However, it can be improved to the present sublinear scale $N^{\frac{8}{9}}$. This issue is important for the nonlinear analysis, and can be resolved with little effort, cf. Remark B.7 and Remark B.4 in the Appendix for more details.

Remark 2.7. The proof of this lemma is based on analysis of semi-algebraic sets (cf. [Bou07, JLS20]), and applies to general nondegenerate analytic potentials.

Remark 2.8. We mention that the set $\mathcal{W}$ on which Theorem 1.1 holds satisfies

$$\mathcal{W} = \mathcal{W}' \cap \mathcal{M}, \tag{2.11}$$

where $\mathcal{M}$ is the set on which the Diophantine estimates in Theorem 2.2 hold, and $\mathcal{W}'$ is defined by (2.9). So, we have

$$\mathrm{meas}(\mathcal{W}) \ge 1 - \log^{-c}\frac{1}{\varepsilon+\delta},\ c = c(d) > 0$$

as shown in Theorem 1.1. For $(\boldsymbol{\alpha},\boldsymbol{\theta}) \in \mathcal{W}$, the conclusion in (2) plays an essential role in the nonlinear analysis.

Proof. It suffices to establish the measure bounds. The proof is a small modification of that of Theorem 4.1 in [JLS20] (cf. pages 475–476). Indeed, the only difference is that we will apply Theorem 3.7 in [JLS20] starting from $N_1 = 10^{-\frac{32}{\kappa_1^2}}\log^{\frac{4}{\kappa_1}}(\log\frac{1}{\varepsilon})$ rather than $\log\log\frac{1}{\varepsilon}$ as in [JLS20]. This change makes sense since the large deviation estimates in [JLS20] hold for the initial scales of $N_0(V,d) \le N \le N_2 := \frac{1}{10}\log^2\frac{1}{\varepsilon}$ (via the Neumann series argument), and for $f(x) = e^{x^{\frac{\kappa_1}{4}}}$,

$$[N_1, (f(N_1))^{\frac{8}{\kappa_1}}] \subset [N_0, N_2].$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

This modification then leads to the following measure estimate:

$$\begin{aligned}\mathrm{meas}(\bigcap_{N\geq N_1}\mathcal{A}_N) &\geq 1-\sum_{N\geq N_1} f(N)^{-\kappa_1}\\ &\geq 1-\sum_{N\geq N_1} e^{-\kappa_1 N^{\frac{\kappa_1}{4}}}\\ &\geq 1-C(d)e^{-\frac{\kappa_1}{2}N_1^{\frac{\kappa_1}{4}}}\\ &\geq 1-C(d)\log^{-\frac{\kappa_1}{2}\times 10^{-\frac{8}{\kappa_1}}}\frac{1}{\varepsilon}.\end{aligned}$$

For the proof of (2.10), it follows directly from applying Fubini's theorem and

$$\mathrm{meas}(\bigcap_{N\geq N_1}\tilde{\Theta}_N(\boldsymbol{\alpha}))\geq 1-C(d)\log^{-\frac{1}{5}\times 10^{-\frac{8}{\kappa_1}}}\frac{1}{\varepsilon}.$$

□

2.6. *Large deviation theorem*

In this section, we establish that the LDE (cf. Definition 2.1) hold for *all* sufficiently large scales by a multi-scale analysis. From (2.2), however, the variation of the frequency, $\boldsymbol{\omega}$ is only $O(\delta)$ and not $O(1)$, hence as we will see shortly, the proof of LDT will be accomplished in three steps instead of the usual two:

- The first step deals with the small scales, for which we establish LDE for all $\boldsymbol{\omega}\in\Omega$ and scales

$$N_0\leq N\leq 10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta},$$

  for some $\rho>0$. In this step, we only use the Neumann series argument, which leads to the above scales (cf. Lemma 2.15).
- Next, we establish intermediate scales LDE, that is, scales in the range

$$10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta}\leq N\leq(\varepsilon+\delta)^{-c_1}$$

  for all $\boldsymbol{\omega}\in\Omega$, where c_1 is defined in Theorem 2.2. Note that Theorem 2.2 is available in the above interval of scales. The proof of such intermediate scales LDE is based on preparation type theorem together with clustering properties of the spectrum of the diagonal matrix $D(0)$ (cf. (2.6), Theorem 2.2).

  Remark 2.9. The intermediate scales are needed here because for a *fixed* $(\boldsymbol{\alpha},\boldsymbol{\theta})$, hence *fixed* $\boldsymbol{\omega}^{(0)}$, $\boldsymbol{\omega}$ only varies in an interval of size $O(\delta)$, see (2.2), as mentioned above. Consequently, there is a gap in the range of scales before Bourgain-type analysis becomes applicable. Fortunately, the Diophantine estimates for $(\boldsymbol{\alpha},\boldsymbol{\theta})\in\mathcal{M}$ can control the resonances directly as in (2.6), leading to LDE for the range of scales in the gap, which we term the *intermediate scales*.

- For the large scales (i.e., $N>(\varepsilon+\delta)^{-c_1}$) LDE, we will apply matrix-valued Cartan's lemma and semi-algebraic sets theory of [BGS02, Bou07]. Of particular importance is Bourgain's geometric lemma (cf. Lemma 2.5). This step requires both the Diophantine condition and the so-called weak second Melnikov's condition on $\boldsymbol{\omega}$.

  To be coherent, since the variation in $\boldsymbol{\omega}$ is only $O(\delta)$, the excised measure in Ω needs to be less than δ^b. So we make a particular choice of the exponent and impose the following Diophantine condition:

$$\begin{aligned}&\mathrm{DC}_{\omega}(N)\\&=\left\{\boldsymbol{\omega}\in\Omega:\ |\boldsymbol{k}\cdot\boldsymbol{\omega}|\geq e^{-N^{\rho^4}}\ \text{for}\ \forall\ 0<|\boldsymbol{k}|\leq 100N^2\right\}.\end{aligned}\tag{2.12}$$

Then we have for $\varepsilon \leq \delta$,

$$\operatorname{meas}\left(\Omega\setminus\left(\bigcap_{N>(\varepsilon+\delta)^{-c_1}}\mathrm{DC}_{\omega}(N)\right)\right)\ll\delta^b.\tag{2.13}$$

The weak second Melnikov's condition is imposed in order to restrict the time $\boldsymbol{k}$-direction where resonances could possibly occur. This greatly reduces the number of resonances, and paves the way for the application of Lemma 2.25.

Remark 2.10. The condition on the size of ε relative to δ stems again from the modulation in $\boldsymbol{\omega}$ being $O(\delta)$, while the excised measure depends on both ε and δ, as we will see in the proofs. This also ensures, in addition, that the amplitude-frequency map: $\boldsymbol{a}\to\boldsymbol{\omega}(\boldsymbol{a})$ is nondegenerate in the nonlinear analysis, so that we may indeed vary $\boldsymbol{\omega}$ in the linear analysis.

Before stating the LDT, let us summarize various parameters and the relation between ε, δ.

- Throughout this paper, we first fix constants

$$\kappa_1=\kappa_1(d),\kappa_2=\kappa_2(d)\in(0,1)$$

given in Bourgain's geometric lemma (cf. Lemma 2.5).
- The most important parameter $\rho>0$ is then fixed so that it depends only on b, d and satisfies

$$\rho=c_2\kappa_1\leq\frac{\kappa_1}{10^4(b+d)^2},\tag{2.14}$$

where $c_2=c_2(b,d)$ is given by Lemma 2.22.
- The parameters

$$0<c_1=c_1(b,d,K,V)<\frac{1}{100b},\ C_1=C_1(b,d,K,V)>1$$

are fixed, which appear in the Diophantine estimates (cf. Theorem 2.2). Also, the parameter c_1 is the power-law exponent of the upper bound of intermediate scales.
- Fix

$$\kappa=\frac{1}{2}\min\left\{\vartheta,\ \frac{1}{10}\right\}\leq\frac{1}{20},$$

where $\vartheta>0$ is an absolute constant appearing in the coupling lemma of [Liu22] (cf. Lemma B.3). Then κ (cf. (2.35)) describes the changes of off-diagonal exponential decay rates of Green's functions along the multi-scale iterations.
- For convenience, we assume the decay rate of S satisfies $\gamma\in(\frac{1}{2},10)$. It turns out that the (ρ,γ_∞,N)-LDE hold true for *all* $N\geq\log^3\frac{1}{\varepsilon+\delta}$ and for some $\gamma_\infty\in[\frac{\gamma}{2},\gamma]$. The constant $C_2>0$ is the power-law index in S.
- With these parameters having been fixed, we aim to find

$$\delta_0=\delta_0(b,d,K,V,\max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|,\kappa_1,\kappa_2,\rho,c_1,C_2)>0,$$

so that, if $0<\varepsilon+\delta\leq\delta_0$, then LDT can be established. Since we will use the multi-scale analysis method to prove LDT, various conditions on δ_0 are required, and it is unclear a priori if such restrictions are meaningful. Fortunately, we will divide the LDT proof into 3 steps, and at each step, we can find a desired δ_ℓ ($\ell = 1, 2, 3$). Finally, it suffices to take $\delta_0 = \min\{\delta_1, \delta_2, \delta_3\}$.

◦ As we will see later, we further assume $\delta \le \log^{-1}\frac{1}{\varepsilon}$ when proving large scales LDE.

Let $\rho, \kappa_1, \kappa_2, c_1, \gamma, C_2$ be fixed as above. We have

Theorem 2.11 (LDT). *There is some*

$$\delta_0 = \delta_0(b, d, K, V, \max_{1\le\ell\le b}|\boldsymbol{n}_\ell|, \kappa_1, \kappa_2, \rho, c_1, C_2) > 0$$

so that the following holds true for $0 < \varepsilon + \delta \le \delta_0$.

(1) *For* $\log^3\frac{1}{\varepsilon+\delta} < N \le 10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta}$, *the* (ρ, γ_N, N)*-LDE hold true for all* $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0, 1]^{2d}$ *(and thus for all* $\boldsymbol{\omega} \in \Omega$*) with* $\gamma_N \equiv \gamma_0 = \gamma - \log^{-2}\frac{1}{\varepsilon+\delta} \in [\frac{\gamma}{2}, \gamma]$.

(2) *For* $10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta} \le N \le (\varepsilon+\delta)^{-c_1}$, *the* (ρ, γ_N, N)*-LDE hold true for* $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{W}$ *(cf. (2.11)) and all* $\boldsymbol{\omega} \in \Omega$ *(depending on* $(\boldsymbol{\alpha}, \boldsymbol{\theta})$*) with* $\gamma_N = \gamma_0 - N^{-\vartheta} \in [\frac{\gamma}{2}, \gamma]$.

(3) *Fix* $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{W}$ *(cf. (2.11)) and let* $\varepsilon \le \delta \le \log^{-1}\frac{1}{\varepsilon}$. *Let* $N > (\varepsilon+\delta)^{-c_1}$, $N_1 \sim N^{\rho^2}$ *and* $N_2 \sim N^{2\rho}$. *Then there are some* $\widetilde{\Omega}_N \subset \Omega$ *(independent of S) satisfying* $\mathrm{meas}(\Omega \setminus \widetilde{\Omega}_N) \le e^{-\frac{1}{5}N^{\frac{3\kappa_1\rho^2}{4}}}$ *and some* $\gamma_N \in [\frac{\gamma}{2}, \gamma]$ *so that the* (ρ, γ_N, N)*-LDE hold true for* $\boldsymbol{\omega} \in \Omega_N := \widetilde{\Omega}_N \cap \mathrm{DC}_{\boldsymbol{\omega}}(N) \cap \Omega_{N_2}$ *with* $\mathrm{DC}_{\boldsymbol{\omega}}(N)$ *given by (2.12) and*

$$\widetilde{\Omega}_N = \bigcap_{1\le\ell,\ell'\le N^{4(b+d)},\,\xi=\pm1,\,\xi'=\pm1,\,0<|\boldsymbol{k}|\le 2N^2} \widetilde{\Omega}_{\boldsymbol{k},\ell,\ell',\xi,\xi'},$$

where

$$\widetilde{\Omega}_{\boldsymbol{k},\ell,\ell',\xi,\xi'} := \left\{\boldsymbol{\omega} \in \Omega :\ |\boldsymbol{k}\cdot\boldsymbol{\omega} + \xi\lambda_\ell - \xi'\lambda_{\ell'}| > 10e^{-\frac{1}{4}N^{\frac{3\kappa_1\rho^2}{4}}}\right\},$$

$\gamma_N \ge \gamma_{N_1} - N^{-\kappa}$ *and* $\lambda_\ell = \lambda_\ell(\boldsymbol{\alpha}, \boldsymbol{\theta})$ *(* $1 \le \ell \le N^{4(b+d)}$*) are all eigenvalues of the operators* $\mathcal{H}_Q(\boldsymbol{\alpha}, \boldsymbol{\theta})$ *for all* $Q \subset [-100N^2, 100N^2]^d$ *satisfying*

$$Q \in \bigcup_{\tilde{N}\in[\frac{1}{4}N^{\kappa_1\rho^2}, N^{\kappa_2\rho^2}]} \mathcal{ER}_{\mathbb{Z}^d}(\tilde{N}).$$

Remark 2.12. Note that the sets Ω_N are *independent of S*. This will be important for later nonlinear applications. Moreover, we have $\Omega_N \subset \Omega_{N-1} \subset \cdots \subset \Omega_{N_0} = \Omega$ and Ω_N is a semi-algebraic set of $\deg\Omega_N \le N^{10(b+d)}$ and $\mathrm{meas}(\Omega_{N^{2\rho}} \setminus \Omega_N) \le e^{-\frac{1}{2}N^{\rho^4}}$.

Remark 2.13. One may observe that this LDT is for a fixed operator H; while the linearized operator is varying in a Newton scheme. This issue, however, can be remedied by using that the Green's function estimates in LDE (at scale N) are stable under perturbations of order e^{-N^2} using Lemma B.1. The sup-exponentially decaying error (cf. (3.14)) in the Newton scheme can ensure this. More precisely, let

$$\begin{aligned}
&\gamma \ge \gamma_1 \ge \gamma_2 \ge \cdots \ge \frac{\gamma}{2},\\
&\gamma \ge \gamma_1' \ge \gamma_2' \ge \cdots \ge \frac{\gamma}{2},\\
&\gamma_\ell' \le \min\{\gamma_{[\ell^{\rho^2}]+1}, \gamma'_{[\ell^{\rho^2}]+1}\} - \ell^{-\kappa}.
\end{aligned}$$

We can replace H in LDT with a sequence of operators $H^{(\ell)}(\sigma) = D(\sigma) + \delta S_\ell$ ($\ell \geq 1$) satisfying (we hide the dependence on $\xi, \xi' \in \{\pm\}$)

$$|S_\ell((\boldsymbol{k}, \boldsymbol{n}); (\boldsymbol{k}', \boldsymbol{n}'))| \leq C_2(1 + |\boldsymbol{k} - \boldsymbol{k}'|)^{C_2} e^{-\gamma_\ell |\boldsymbol{k}-\boldsymbol{k}'| - \gamma_\ell |\boldsymbol{n}|} \delta_{\boldsymbol{n},\boldsymbol{n}'},$$
$$\|H^{(\ell+1)} - H^{(\ell)}\| \leq e^{-\ell^{\frac{2}{\rho^2}}}.$$

It is remarkable that in our LDT, Ω_N does not depend on S. So with this modification, the conclusion of LDT becomes for $\boldsymbol{\omega} \in \Omega_N$, $H^{(\ell)}$ has (ρ, γ'_N, N)-LDE for all $\ell \geq N \geq \log^3 \frac{1}{\varepsilon+\delta}$ (Σ_N depending only on S_ℓ, $1 \leq \ell \leq N$).

Indeed, assume we have established the LDE at scale $N_1 = N^{\rho^2}$ for all $H^{(\ell)}$ with $\ell \geq N_1$. Then it suffices to establish (ρ, γ'_N, N)-LDE for $H^{(N)}$, since for all $\ell > N$,

$$\|H^{(\ell)} - H^{(N)}\| \leq \sum_{\ell \geq N} e^{-\ell^{\frac{2}{\rho^2}}} < e^{-\frac{N^{\frac{2}{\rho^2}}}{2}} \ll e^{-100N},$$

and (ρ, γ'_N, N)-LDE for $H^{(\ell)}$ can be obtained via the perturbation argument (cf. Lemma B.1). Next, we fix $S_\ell \equiv S_{N_1}$ for $\ell \in [N_1, N]$. Similar to the proof of Lemma 2.20, we can show using the modified resolvent identities Lemmas B.3, B.5, B.6 (e.g., taking into account the power-law factor in the estimate of S_ℓ) that $H^{(N_1)}$ satisfies (ρ, γ''_N, N)-LDE with

$$\gamma''_N \geq \min\{\gamma_{N_1}, \gamma'_{N_1}\} - N^{-\kappa} \geq \gamma'_N.$$

Finally, using again the perturbation argument Lemma B.1 and

$$\|H^{(N_1)} - H^{(N)}\| \leq e^{-\frac{N_1^{\frac{2}{\rho^2}}}{2}} \leq e^{-\frac{N^2}{5}} \ll e^{-100N},$$

we get that $H^{(N)}$ satisfies (ρ, γ'_N, N)-LDE by the assumptions on $\gamma_\ell, \gamma'_\ell$.

Remark 2.14. In the large scales case, we assume $\delta \in [\varepsilon, \log^{-1} \frac{1}{\varepsilon}]$. The condition $\delta \geq \varepsilon$ is necessary since (2.13) as mentioned above. The restriction $\delta \leq \log^{-1} \frac{1}{\varepsilon}$, however, is due to the fact that we need to use Bourgain's geometric lemma in the form of Lemma 2.5 to handle the large $|\boldsymbol{n}|$ regime estimates. In fact, if we assume that $\mathcal{W}$ (cf. (2.11)) satisfies $\mathrm{meas}([0,1]^{2d} \setminus \mathcal{W}) = o(1)$ (rather than the quantitative version $\mathrm{meas}([0,1]^{2d} \setminus \mathcal{W}) = O(\log^{-c} \frac{1}{\varepsilon+\delta})$ in the present work), we can remove the condition $\delta \leq \log^{-1} \frac{1}{\varepsilon}$ (cf. Remark 2.24 for details).

This LDT will be proved in detail in the following three sections.

2.6.1. The small scales LDE

This section is devoted to the proof of LDE for small scales, that is,

$$\log^3 \frac{1}{\varepsilon+\delta} \leq N \leq 10^{-\frac{1}{\rho}} \log^{\frac{1}{\rho}} \frac{1}{\varepsilon+\delta}.$$

In this case, we only use the standard Neumann series argument. As a result, we can establish LDE for all $\boldsymbol{\omega} \in \Omega$. We have

Lemma 2.15. *Let*

$$\Sigma_N = \bigcup_{\Lambda \in (\boldsymbol{0},\boldsymbol{n}) + \mathcal{E}\mathcal{R}_{\boldsymbol{0}}(N),\ |\boldsymbol{n}| \leq 10N} \left\{\sigma \in \mathbb{R} :\ \min_{(\boldsymbol{k},\boldsymbol{n},\xi) \in \Lambda} |\xi(\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega}) + \mu_{\boldsymbol{n}}| \leq e^{-2N^\rho}\right\}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Let $0 < \rho < \frac{1}{10^4}$. *Then there is some* $\delta_1 = \delta_1(b, d, C_2, \rho) > 0$ *so that if* $0 < \varepsilon + \delta \leq \delta_1$, *for*

$$\log^3 \frac{1}{\varepsilon + \delta} \leq N \leq 10^{-\frac{1}{\rho}} \log^{\frac{1}{\rho}} \frac{1}{\varepsilon + \delta},$$

we have

$$\operatorname{meas}(\Sigma_N) \leq e^{-N^\rho}. \tag{2.15}$$

Moreover, if $\sigma \notin \Sigma_N$, *we have for all* $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N)$ *with* $|\boldsymbol{n}| \leq 10N$,

$$\|G_\Lambda(\sigma)\| \leq e^{N^{\frac{3}{4}}}, \tag{2.16}$$

$$|G_\Lambda(\sigma)((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \leq e^{-\gamma_0 |\boldsymbol{x} - \boldsymbol{x}'|} \text{ for } |\boldsymbol{x} - \boldsymbol{x}'| \geq N^{\frac{8}{9}}, \tag{2.17}$$

where

$$\gamma_0 = \gamma - \log^{-2} \frac{1}{\varepsilon + \delta} \in [\frac{\gamma}{2}, \gamma]. \tag{2.18}$$

Remark 2.16. In this lemma, we make no restrictions on $\boldsymbol{\omega}$.

Proof. We first establish the measure bound (2.15). We have

$$\operatorname{meas}(\Sigma_N) \leq C(b, d) N^{C(b,d)} e^{-2N^\rho} \leq e^{-N^\rho},$$

where in the last inequality we use $0 < \varepsilon + \delta \leq \delta_1^{(1)}(b, d, \rho) \ll 1$ to ensure $N \geq \log^3 \frac{1}{\varepsilon+\delta} \geq N_0(b, d, \rho) \gg 1$.

Next, we let $\sigma \notin \Sigma_N$ and fix any $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N)$ with $|\boldsymbol{n}| \leq 10N$. Let $A = D_\Lambda(\sigma), B = \delta S_\Lambda$. We have

$$|A^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \leq e^{N^\rho} \delta_{\boldsymbol{x}, \boldsymbol{x}'}, \ |B(\boldsymbol{x}, \boldsymbol{x}')| \leq C_2 \delta (1 + |\boldsymbol{x} - \boldsymbol{x}'|)^{C_2} e^{-\gamma |\boldsymbol{x} - \boldsymbol{x}'|}.$$

Note that

$$\begin{aligned} C N^{2(b+d)+C_2} \delta e^{N^\rho} &\leq C(b, d, C_2) N^{2(b+d)+C_2} (\varepsilon + \delta)^{\frac{9}{10}} \\ &\leq C(b, d, C_2) 10^{-\frac{2(b+d)+C_2}{\rho}} (\varepsilon + \delta)^{\frac{9}{10}} \log^{-\frac{2(b+d)+C_2}{\rho}} \frac{1}{\varepsilon + \delta} \\ &\leq \frac{1}{2} \end{aligned}$$

assuming $0 < \varepsilon + \delta \leq \delta_1^{(2)}(b, d, C_2, \rho) \ll 1$. So we can apply Lemma B.1 with $A = D_\Lambda(\sigma), B = \delta S_\Lambda, \epsilon_2 = C_2 \delta, \epsilon_1 = e^{-N^\rho}, M = 0, X = \Lambda$ to obtain (we hide the dependence on ξ, ξ')

$$\begin{aligned} \|G_\Lambda(\sigma)\| &\leq 2 e^{N^\rho} < e^{N^{\frac{3}{4}}} \ (\text{since } 0 < \rho < \frac{1}{10^4}), \\ |G_\Lambda(\sigma)(\boldsymbol{x}, \boldsymbol{x}')| &\leq e^{N^\rho} e^{-\gamma |\boldsymbol{x} - \boldsymbol{x}'|}. \end{aligned}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Then for $|\boldsymbol{x}-\boldsymbol{x}'| \geq N^{\frac{8}{9}}$,

$$\begin{aligned}|G_\Lambda(\sigma)(\boldsymbol{x},\boldsymbol{x}')| &\leq e^{-|\boldsymbol{x}-\boldsymbol{x}'|(\gamma-\frac{N^\rho}{|\boldsymbol{x}-\boldsymbol{x}'|})}\\ &\leq e^{-|\boldsymbol{x}-\boldsymbol{x}'|(\gamma-N^{-\frac{7}{8}})} \ (\text{since } 0<\rho<\frac{1}{10^4})\\ &:= e^{-\gamma'|\boldsymbol{x}-\boldsymbol{x}'|},\end{aligned}$$

which implies

$$\gamma' > \gamma_0 = \gamma - \log^{-2}\frac{1}{\varepsilon+\delta}.$$

Then both (2.16) and (2.17) hold true if we assume that

$$0<\varepsilon+\delta\leq\delta_1=\min\{\delta_1^{(1)}(b,d,C_2,\rho),\delta_1^{(2)}(b,d,C_2,\rho)\},$$

and $\sigma\notin\Sigma_N$. □

2.6.2. The intermediate scales LDE

This section and the next deal with intermediate-scale LDE. We will show that the LDE hold true for all $\boldsymbol{\omega}\in\Omega$ and scales in the range

$$10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta}\leq N\leq(\varepsilon+\delta)^{-c_1},$$

where c_1 is given by Theorem 2.2. In this case, the analysis in Section 2.4 becomes essential.

Lemma 2.17. *Let $0<\rho<\frac{1}{10^4}$. There is some $\delta_2=\delta_2(b,d,K,V,\rho,c_1,C_2)>0$ so that the following holds for $0<\varepsilon+\delta\leq\delta_2$. Assume $(\boldsymbol{\alpha},\boldsymbol{\theta})\in\mathcal{M}$ (cf. Theorem 2.2) and fix*

$$10^{-\frac{1}{\rho}}\log^{\frac{1}{\rho}}\frac{1}{\varepsilon+\delta}\leq N\leq(\varepsilon+\delta)^{-c_1}.$$

Then for each $\boldsymbol{\omega}\in\Omega$, there exists a set $\Sigma_N\subset\mathbb{R}$ with

$$\operatorname{meas}(\Sigma_N)\leq(\varepsilon+\delta)^{-\frac{9}{10}}e^{-\frac{1}{2b}N^{\frac{3}{8}}}<e^{-N^\rho},$$

so that if $\sigma\notin\Sigma_N$, then for all $\Lambda\in(\boldsymbol{0},\boldsymbol{n})+\mathcal{ER}_{\boldsymbol{0}}(N)$ with $|\boldsymbol{n}|\leq 10N$,

$$\|G_\Lambda(\sigma)\|\leq e^{N^{\frac{3}{4}}},$$
$$|G_\Lambda(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))|\leq e^{-\gamma_N|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'|\geq N^{\frac{8}{9}},$$

where

$$\gamma_N=\gamma_0-N^{-\vartheta}\in[\frac{\gamma}{2},\gamma],$$

with γ_0 given by (2.18) and $\vartheta>0$ by Lemma B.3.

Remark 2.18. Note that this lemma also holds for all $\boldsymbol{\omega}\in\Omega$.

Proof. The proof is based on the clustering property (cf. (2) of Theorem 2.2), Schur's complement argument, Rouché's theorem and the iteration of resolvent identities. So we first choose $0<\varepsilon+\delta\leq c(b,d,K,V,\max\limits_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|)\ll 1$ so that Theorem 2.2 holds true. Then the proof can be decomposed into two steps.

Step 1. We define

$$\Sigma_N = \bigcup_{\sqrt{N} \leq N' \leq N,\ \Lambda \in (\mathbf{0},\boldsymbol{n}) + \mathcal{E}\mathcal{R}_{\mathbf{0}}(N'),\ |\boldsymbol{n}| \leq 10N'} \Sigma_\Lambda, \tag{2.19}$$

where

$$\Sigma_\Lambda = \left\{\sigma \in \mathbb{R} : \ \|G_\Lambda(\sigma)\| \geq e^{(\operatorname{diam}\Lambda)^{\frac{3}{4}}}\right\}.$$

We first show that

$$\Sigma_\Lambda \subset \bigcup_{\xi = \pm 1, \boldsymbol{i} = (\boldsymbol{k},\boldsymbol{n}) \in \Lambda} I_{\xi,\boldsymbol{i}}, \tag{2.20}$$

where

$$I_{\xi,\boldsymbol{i}} = \left\{\sigma \in \mathbb{R} : \ |\xi(\sigma + \boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)}) + \mu_{\boldsymbol{n}}| \leq \frac{1}{100}(\varepsilon+\delta)^{\frac{1}{7b}}\right\}.$$

In fact, if $\sigma \notin \bigcup_{(\xi,\boldsymbol{i}) \in \Lambda} I_{\xi,\boldsymbol{i}}$, then for each $\boldsymbol{\omega} \in \Omega$,

$$\begin{aligned}
&\inf_{(\xi,\boldsymbol{i}) \in \Lambda} |\xi(\sigma + \boldsymbol{k}\cdot\boldsymbol{\omega}) + \mu_{\boldsymbol{n}}| \\
&\geq \inf_{(\xi,\boldsymbol{i}) \in \Lambda} |\xi(\sigma + \boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)}) + \mu_{\boldsymbol{n}}| - \sup_{\boldsymbol{k} \in \Lambda} |\boldsymbol{k}\cdot(\boldsymbol{\omega} - \boldsymbol{\omega}^{(0)})| \\
&\geq \frac{1}{200}(\varepsilon+\delta)^{\frac{1}{7b}}.
\end{aligned}$$

Since $N \geq 10^{-\frac{1}{\rho}} \log^{\frac{1}{\rho}} \frac{1}{\varepsilon+\delta}$, the Neumann series argument implies

$$\|G_\Lambda(\sigma)\| \leq C(\varepsilon+\delta)^{-\frac{2}{7b}} < e^{(\operatorname{diam}\Lambda)^{\frac{3}{4}}}$$

when $\sigma \notin \bigcup_{(\xi,\boldsymbol{i}) \in \Lambda} I_{\xi,\boldsymbol{i}}$. This proves (2.20) and

$$\Sigma_\Lambda = \bigcup_{(\xi,\boldsymbol{i}) \in \Lambda} (\Sigma_\Lambda \cap I_{\xi,\boldsymbol{i}}).$$

In the following, we estimate $\operatorname{meas}(\Sigma_\Lambda \cap I_{\xi,\boldsymbol{i}})$. For this purpose, we fix any $\xi^* \in \{+,-\}$, $\boldsymbol{i}^* = (\boldsymbol{k}^*, \boldsymbol{n}^*)$ and denote

$$I^* = I_{\xi^*,\boldsymbol{i}^*} = \left[\sigma^* - \frac{1}{100}(\varepsilon+\delta)^{\frac{1}{7b}}, \sigma^* + \frac{1}{100}(\varepsilon+\delta)^{\frac{1}{7b}}\right],$$

where

$$\sigma^* = -\boldsymbol{k}^*\cdot\boldsymbol{\omega}^{(0)} - \xi^*\mu_{\boldsymbol{n}^*}. \tag{2.21}$$

We will study $G_\Lambda(\sigma)$ for $\sigma \in I^*$. From (2) of Theorem 2.2, we have for every $\sigma^* \in \mathbb{R}$, there exists $B_* = B_*^+ \cup B_*^- \subset \Lambda$ with $\# B_*^\pm \leq b$ so that

$$\inf_{(\xi,\boldsymbol{i}) \in \Lambda \setminus B_*} |\xi(\sigma^* + \boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)}) + \mu_{\boldsymbol{n}}| \geq \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{4}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Similar to the proof of (2.20), if (we denote $\mathbb{D}_r(z) := \{w \in \mathbb{C} : |w - z| \le r\}$) $z \in \mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{8b}}/10}(\sigma^*)$, then

$$\inf_{(\xi,\boldsymbol{i})\in\Lambda\setminus B_*} |\xi(z + \boldsymbol{k}\cdot\omega) + \mu_{\boldsymbol{n}}| \ge \frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{8}.$$

So we have an analytic extension of $H_\Lambda(z)$ to $z \in \mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{8b}}/10}(\sigma^*)$ and

$$\sup_{(\boldsymbol{x},\xi),(\boldsymbol{x}',\xi')\in\Lambda,\ z\in\mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{8b}}/10}(\sigma^*)} |D_\Lambda(z)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| \le C(b)(\varepsilon+\delta)^{-c_1}.$$

More importantly, if we denote $\Lambda^c = \Lambda \setminus B_*$, then by the Neumann series argument, we have for all $z \in \mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{8b}}/10}(\sigma^*)$,

$$\|G_{\Lambda^c}(z)\| \le (\varepsilon+\delta)^{-\frac{1}{4b}}.$$

In the following, we estimate $G_\Lambda(z)$ using the Schur complement reduction. We define

$$S(z) = R_{B_*}H(z)R_{B_*} - R_{B_*}H(z)R_{\Lambda^c}G_{\Lambda^c}(z)R_{\Lambda^c}H(z)R_{B_*}.$$

Obviously, we have

$$R_{B_*}H(z)R_{B_*} = R_{B_*}D(z)R_{B_*} + O(\varepsilon+\delta) \tag{2.22}$$

and

$$R_{B_*}H(z)R_{\Lambda^c}G_{\Lambda^c}(z)R_{\Lambda^c}H(z)R_{B_*} = O((\varepsilon+\delta)^{2-\frac{1}{4b}}). \tag{2.23}$$

Now we consider $\det S(z)$. Combining (2.22) and (2.23) yields

$$\begin{aligned}
&\det S(z)\\
&= \prod_{\boldsymbol{i}=(\boldsymbol{k},\boldsymbol{n})\in B_*^+} (z + \boldsymbol{k}\cdot\omega + \mu_{\boldsymbol{n}}) \prod_{\boldsymbol{i}=(\boldsymbol{k},\boldsymbol{n})\in B_*^-} (-z - \boldsymbol{k}\cdot\omega + \mu_{\boldsymbol{n}})\\
&\quad + O((\varepsilon+\delta)^{3/4})\\
&= (-1)^{\#B_*^-}\prod_{\ell=1}^{\#B_*}(z - \sigma_\ell) + O((\varepsilon+\delta)^{3/4}),
\end{aligned}$$

where $\sigma_\ell \in \{-\boldsymbol{k}\cdot\omega - \mu_{\boldsymbol{n}}\}_{(\boldsymbol{k},\boldsymbol{n})\in B_*^+} \cup \{-\boldsymbol{k}\cdot\omega + \mu_{\boldsymbol{n}}\}_{(\boldsymbol{k},\boldsymbol{n})\in B_*^-}$. Recalling the definition of B_*, we must have

$$|z - \sigma_\ell| \le \frac{4}{5}(\varepsilon+\delta)^{\frac{1}{8b}}.$$

From (2.21), there exists at least one σ_ℓ so that $|\sigma_\ell - \sigma^*| = O((\varepsilon+\delta)^{1-c_1})$. At this stage, we need a useful result.

Lemma 2.19. *Let* $\varphi = \left(\frac{14}{15}\right)^{\frac{1}{5b}} < 1$. *There is* $0 \le \ell_* \le 5b - 1$ *so that*

$$(\mathbb{A}_{\ell_*+1} \setminus \mathbb{A}_{\ell_*}) \cap \{\sigma_\ell\}_{1\le\ell\le\#B_*} = \emptyset,$$

where for $0 \le \ell \le 5b$,

$$\mathbb{A}_\ell = \mathbb{D}_{(\varepsilon+\delta)^{\frac{\varphi^\ell}{7b}}}(\sigma^*).$$

Proof of Lemma 2.19. It suffices to note that

$$\mathbb{A}_{5b} \setminus \mathbb{A}_0 = \bigcup_{\ell=0}^{5b-1} (\mathbb{A}_{\ell+1} \setminus \mathbb{A}_\ell).$$

The proof then follows from the pigeonhole principle since $\#B_* \leq 2b$. □

Let $r_* = (\varepsilon+\delta)^{\frac{\varphi^{\ell_*}}{7b}}$, where φ and ℓ_* are defined in Lemma 2.19. From Lemma 2.19, we can restrict our considerations on $\mathbb{A}_*$ with

$$\mathbb{A}_* = \mathbb{D}_{\frac{r_*+r_*^{\varphi}}{2}}(\sigma^*).$$

Then

$$\mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{7b}}}(\sigma^*) \subset \mathbb{A}_{\ell_*} \subset \mathbb{A}_* \subset \mathbb{D}_{(\varepsilon+\delta)^{\frac{1}{7.5b}}}(\sigma^*).$$

Let

$$\mathcal{K} = \{1 \leq \ell \leq \#B_* : \ \sigma_\ell \in \mathbb{D}_{10r_*}(\sigma^*)\}.$$

We have $\mathcal{K} \neq \emptyset$ and for $\ell \notin \mathcal{K}$,

$$\sigma_\ell \notin \mathbb{A}_*, \ \operatorname{dist}(\sigma_\ell, \partial\mathbb{D}_{20r_*}(\sigma^*)) \geq \operatorname{dist}(\sigma_\ell, \partial\mathbb{A}_*) \geq \frac{r_*^{\varphi}}{4} \gg (\varepsilon+\delta)^{\frac{1}{7b}}.$$

As a result, we can write for $z \in \mathbb{A}_*$,

$$\frac{\det S(z)}{\prod_{\ell\notin\mathcal{K}}(z-\sigma_\ell)} = \pm\prod_{\ell\in\mathcal{K}}(z-\sigma_\ell) + O((\varepsilon+\delta)^{\frac{1}{3}}). \tag{2.24}$$

It holds that

$$\prod_{\ell\in\mathcal{K}} |z-\sigma_\ell| \big|_{\partial\mathbb{D}_{20r_*}(\sigma^*)} \geq (\varepsilon+\delta)^{\frac{\#\mathcal{K}}{7b}} \geq (\varepsilon+\delta)^{2/7} \gg (\varepsilon+\delta)^{\frac{1}{3}}.$$

Applying Rouché's theorem shows that the function $\frac{\det S(z)}{\prod_{\ell\notin\mathcal{K}}(z-\sigma_\ell)}$ has exactly $\#\mathcal{K}$ zeros (denoted by $z_\ell, \ \ell \in \mathcal{K}$) in $\mathbb{D}_{20r_*}(\sigma^*)$. From $\sigma_\ell \in \mathbb{D}_{10r_*}(\sigma^*)$, we also get

$$|\sigma_\ell - z_\ell| \leq 30r_*. \tag{2.25}$$

From similar analysis, we can obtain that (2.24) also has exactly $\#\mathcal{K}$ zeros in $\mathbb{A}_*$, which implies that $\{z_\ell\}_{\ell\in\mathcal{K}}$ are all zeros of (2.24) in $\mathbb{A}_*$. So on $\mathbb{A}_*$, we have

$$\frac{\det S(z)}{\prod_{\ell\notin\mathcal{K}}(z-\sigma_\ell)} = g(z)\prod_{\ell\in\mathcal{K}}(z-z_\ell), \tag{2.26}$$

where g is analytic on $\mathbb{A}_*$ with $\inf_{z\in\mathbb{A}_*}|g(z)| > 0$. Using (2.25) and (2.26), we have

$$g\big|_{\partial\mathbb{A}_*} = 1 + O((\varepsilon+\delta)^{\frac{1}{21}}).$$

Using the maximum principle then leads to

$$g\big|_{\mathbb{A}_*} = 1 + O((\varepsilon+\delta)^{\frac{1}{21}}).$$

We have established on $\mathbb{A}_*$,

$$\det S(z) = g(z) \prod_{\ell \notin \mathcal{K}} (z - \sigma_\ell) \prod_{\ell \in \mathcal{K}} (z - z_\ell).$$

In particular, one has for all $z \in \mathbb{A}_*$,

$$\begin{aligned} |\det S(z)| &\geq \frac{4}{5} (\varepsilon + \delta)^{\frac{\# B_* - \# \mathcal{K}}{7b}} \prod_{\ell \in \mathcal{K}} |z - z_\ell| \\ &\geq (\varepsilon + \delta)^{\frac{2}{7}} \prod_{\ell \in \mathcal{K}} |z - z_\ell| \end{aligned}$$

assuming $0 < \varepsilon + \delta \leq c(b) \ll 1$. Thus combining Hadamard's inequality and Cramer's rule implies for $z \in \mathbb{A}_*$,

$$\begin{aligned} \|S^{-1}(z)\| &\leq C(b)(\varepsilon + \delta)^{-2bc_1} (\varepsilon + \delta)^{-\frac{2}{7}} \prod_{\ell \in \mathcal{K}} |z - z_\ell|^{-1} \\ &\leq (\varepsilon + \delta)^{-\frac{1}{3}} \prod_{\ell \in \mathcal{K}} |z - z_\ell|^{-1} \ (\text{since } c_1 < \frac{1}{100b}). \end{aligned}$$

Applying the Schur complement argument shows that for $z \in \mathbb{A}_*$,

$$\begin{aligned} \|G_\Lambda(z)\| &\leq 4(1 + \|G_{\Lambda^c}(z)\|)^2 (1 + \|S^{-1}(z)\|) \\ &\leq (\varepsilon + \delta)^{-\frac{7}{8}} \prod_{\ell \in \mathcal{K}} |z - z_\ell|^{-1}. \end{aligned}$$

Recall that $I^* \subset \mathbb{A}_*$. So for $\sigma \in I^*$, we have

$$\|G_\Lambda(\sigma)\| \leq (\varepsilon + \delta)^{-\frac{7}{8}} \prod_{\ell \in \mathcal{K}} |\sigma - z_\ell|^{-1}$$

and

$$\Sigma_\Lambda \cap I^* \subset \bigcup_{\ell \in \mathcal{K}} \left\{ \sigma \in I^* : \ |\sigma - z_\ell| \leq (\varepsilon + \delta)^{-\frac{7}{8}} e^{-\frac{1}{2b} (\operatorname{diam} \Lambda)^{\frac{3}{4}}} \right\}.$$

This implies

$$\operatorname{meas}(\Sigma_\Lambda \cap I^*) \leq C(\varepsilon + \delta)^{-\frac{7}{8}} e^{-\frac{1}{2b} (\operatorname{diam} \Lambda)^{\frac{3}{4}}},$$

and subsequently

$$\operatorname{meas}(\Sigma_\Lambda) \leq (\varepsilon + \delta)^{-\frac{8}{9}} e^{-\frac{1}{2b} (\operatorname{diam} \Lambda)^{\frac{3}{4}}},$$

which combined with (2.19) shows

$$\operatorname{meas}(\Sigma_N) \leq (\varepsilon + \delta)^{-\frac{9}{10}} e^{-\frac{1}{2b} N^{\frac{3}{8}}}.$$

Step 2. In this step, we prove the off-diagonal exponential decay of $G_\Lambda(\sigma)((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))$ assuming $\sigma \notin \Sigma_N$ and $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N)$ with $|\boldsymbol{n}| \leq 10N$. By Theorem 2.2, we know that there exists $B \subset \Lambda$

so that

$$\inf_{(\xi,\boldsymbol{k},\boldsymbol{n})\in\Lambda\setminus B}|\xi(\sigma+\boldsymbol{k}\cdot\boldsymbol{\omega})+\mu_{\boldsymbol{n}}|\geq\frac{(\varepsilon+\delta)^{\frac{1}{8b}}}{8}.$$

Using the Neumann series argument (cf. Lemma B.1) yields that for any $\Lambda'\subset\Lambda$ with $\Lambda'\cap B=\emptyset$,

$$\|G_{\Lambda'}(\sigma)\|\leq 50(\varepsilon+\delta)^{-\frac{1}{4b}}, \tag{2.27}$$

$$|G_{\Lambda'}(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))|\leq 50(\varepsilon+\delta)^{-\frac{1}{4b}}e^{-\gamma|\boldsymbol{x}-\boldsymbol{x}'|}\text{ for }\boldsymbol{x}\neq\boldsymbol{x}'. \tag{2.28}$$

We say that $\mathcal{ER}(\sqrt{N})\ni Q,\ Q\subset\Lambda$ is $\sqrt{N}$-regular if both (2.27) and (2.28) hold with $\Lambda'=Q$. Otherwise, Q is called $\sqrt{N}$-singular. Let $\mathcal{F}$ be any family of pairwise disjoint $\sqrt{N}$-singular sets contained in Λ. We obtain

$$\#\mathcal{F}\leq 2b.$$

We have established the sublinear bound. Then applying the coupling lemma, that is, Lemma B.3 of [Liu22], shows

$$|G_{\Lambda}(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))|\leq e^{-\gamma_N|\boldsymbol{x}-\boldsymbol{x}'|}\text{ for }|\boldsymbol{x}-\boldsymbol{x}'|>N^{\frac{8}{9}},$$

where

$$\begin{aligned}\gamma_N&\geq\gamma-N^{-\frac{8}{9}}\cdot\log\frac{1}{\varepsilon+\delta}-N^{-\vartheta}\\&\geq\gamma-\log^{-\frac{8}{9\rho}+1}\frac{1}{\varepsilon+\delta}-N^{-\vartheta}\\&\geq\gamma_0-N^{-\vartheta}\ (\text{since }0<\rho<\frac{1}{10^4}).\end{aligned}$$

This proves the off-diagonal exponential decay estimates.

We have completed the proof of Lemma 2.17. □

2.6.3. The large scales LDE

In this section, we will prove the LDE for

$$(\varepsilon+\delta)^{-c_1}<N<\infty.$$

To perform the multi-scale scheme, we need two small scales $N_1<N_2$ with

$$N_2=N_1^{\frac{2}{\rho}},\ N=N_1^{\frac{1}{\rho^2}}.$$

We assume the LDE hold for $\boldsymbol{\omega}\in\Omega_{N_2}\subset\Omega_{N_1}$.

Then we have

Lemma 2.20. *There are some* $0<c_2=c_2(b,d)\leq\frac{1}{10^4(b+d)^2}$ *and some*

$$\delta_3=\delta_3(b,d,K,V,\max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|,\kappa_1,\kappa_2,\rho,c_1,C_2)>0$$

so that the following holds true for $0<\varepsilon+\delta\leq\delta_3$, $\varepsilon\leq\delta\leq\log^{-1}\frac{1}{\varepsilon}$ *and* $0<\rho\leq c_2\kappa_1$ *with* κ_1 *given by Lemma 2.5. Let* $(\boldsymbol{\alpha},\boldsymbol{\theta})\in\mathcal{W}$ *(cf. (2.11)) and* N *satisfy*

$$(\varepsilon+\delta)^{-c_1}<N<\infty.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Then there is a semi-algebraic set $\widetilde{\Omega}_N \subset \Omega$ *(cf. Lemma 2.26) with* $\mathrm{meas}(\Omega \setminus \widetilde{\Omega}_N) \leq e^{-\frac{1}{5}N^{\frac{3\kappa_1\rho^2}{4}}}$ *and* $\deg \widetilde{\Omega}_N \leq N^{5(b+d)}$ *so that, for* $\omega \in \mathrm{DC}_\omega(N) \cap \widetilde{\Omega}_N \cap \Omega_{N_2}$ *(with* $\mathrm{DC}_\omega(N)$ *given by (2.12)), there is some* $\Sigma_N \subset \mathbb{R}$ *with* $\mathrm{meas}(\Sigma_N) \leq e^{-N^\rho}$*, so that, if* $\sigma \notin \Sigma_N$*, then for all* $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N)$ *with* $|\boldsymbol{n}| \leq 10N$*,*

$$\|G_\Lambda(\sigma)\| \leq e^{N^{\frac{3}{4}}},$$
$$|G_\Lambda(\sigma)((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \leq e^{-\gamma_N |\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x} - \boldsymbol{x}'| \geq N^{\frac{8}{9}},$$

where $[\frac{\gamma}{2}, \gamma] \ni \gamma_N \geq \gamma_{N_1} - N^{-\kappa}$ *(cf. (2.35)).*

We begin with a useful definition.

Definition 2.21 (σ-good). For $\Lambda \in \mathcal{ER}(L)$, we say Λ is σ-*good* if

$$\|G_\Lambda(\sigma)\| \leq e^{L^{\frac{3}{4}}},$$
$$|G_\Lambda(\sigma)(\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \leq e^{-\gamma_L |\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x} - \boldsymbol{x}'| \geq L^{\frac{8}{9}},$$

where $\gamma_L \in [\frac{\gamma}{2}, \gamma]$, and σ-*bad*, if one of the inequalities is violated.

The proof of Lemma 2.20 is based on matrix-valued Cartan's lemma (Lemma 2.25) and iterations of resolvent identities. To apply Cartan's lemma, one needs to establish the sublinear bound on σ-bad $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$ contained in Λ_N: If $|\boldsymbol{n}| \leq 10N_1$, the sublinear bound can be obtained via the Töplitz property of H in the $\boldsymbol{k}$-direction together with the semi-algebraic sets theory (cf. Lemma 2.22); for $|\boldsymbol{n}| > 10N_1$, we will use the short-range property of S and Bourgain's geometric lemma (Lemma 2.5) to prove the sublinear bound. Once the sublinear bound is derived, one can use the Cartan's lemma and resolvent identities to finish the proof of large scales LDE.

We need the following ingredients to prove Lemma 2.20.

Lemma 2.22. *There is some* $0 < c_2 = c_2(b, d) \leq \frac{1}{10^4(b+d)^2}$ *so that the following holds true. Assume that LDE hold true at scale* N_1 *with*

$$N_1 \geq N_0(b, d, \rho, K, V).$$

If $0 < \rho \leq c_2$*, then for all* $\sigma \in \mathbb{R}$*, the number of* σ*-bad* $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$ *satisfying* $|\boldsymbol{k}| \leq N^2$ *and* $|\boldsymbol{n}| \leq 10N_1$ *is at most* $N^{\frac{1}{10^3(b+d)^2}}$.

Proof. First, we define $\widetilde{\Sigma}_{N_1}$ to be the set of $\sigma \in \mathbb{R}$ so that, there is $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$ satifying $|\boldsymbol{n}| \leq 10N_1$, which is σ-bad. We then obtain using LDE at scale N_1,

$$\mathrm{meas}(\widetilde{\Sigma}_{N_1}) \leq C(b, d) N_1^{C(b,d)} e^{-N_1^\rho} < e^{-\frac{1}{2}N_1^\rho}$$

for $N_1 \geq N_0(b, d, \rho)$. Using the Hilbert-Schmidt norm and Cramer's rule, we can identify $\widetilde{\Sigma}_{N_1}$ with a semi-algebraic set of $\deg \widetilde{\Sigma}_{N_1} \leq N_1^{C(b,d)}$. Then using Basu-Pollack-Roy Theorem [BPR96] on Betti numbers of a semi-algebraic set (cf. also Proposition 9.2 of [Bou05a]), we have a decomposition of

$$\widetilde{\Sigma}_{N_1} \subset \bigcup_{1 \leq \ell \leq N_1^{C(b,d)}} I_\ell,$$

where each I_ℓ is an interval of length $|I_\ell| \leq e^{-\frac{1}{2}N_1^\rho}$. Now assume that $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$ and $\Lambda' \in (\boldsymbol{k}', \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$, which are all σ-bad for some $|\boldsymbol{n}| \leq 10N_1, \boldsymbol{k} \neq \boldsymbol{k}'$ and $|\boldsymbol{k}|, |\boldsymbol{k}'| \leq N^2$. Then from the Töplitz property of $H(\sigma)$ in the $\boldsymbol{k}$-direction, we obtain

$$\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega},\ \sigma + \boldsymbol{k}' \cdot \boldsymbol{\omega} \in \widetilde{\Sigma}_{N_1}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

As a result, by $\boldsymbol{k} \neq \boldsymbol{k}'$ with $|\boldsymbol{k}|, |\boldsymbol{k}'| \leq N^2$, $\boldsymbol{\omega} \in \mathrm{DC}_{\omega}(N)$ (cf. (2.12)) and $N \sim N_1^{\frac{1}{\rho^2}}$, we have for $N_1 \geq N_0(b, d, \rho, K, V)$,

$$e^{-\frac{1}{2}N_1^{\rho}} = e^{-\frac{1}{2}N^{\rho^3}} < e^{-N^{\rho^4}} \leq |(\boldsymbol{k} - \boldsymbol{k}') \cdot \boldsymbol{\omega}|,$$

which implies $\boldsymbol{k}$ and $\boldsymbol{k}'$ cannot stay in the same interval I_ℓ for $1 \leq \ell \leq N_1^{C(b,d)}$. Thus we have shown that the number of σ-bad $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(N_1)$ with $|\boldsymbol{n}| \leq 10N_1$ and $|\boldsymbol{k}| \leq N^2$ is at most

$$N_1^{C(b,d)} \leq N^{\rho^2 \cdot C(b,d)} \leq N^{\frac{1}{10^3(b+d)^2}},$$

assuming $0 < \rho \leq c_2(b, d) \leq \frac{1}{10^4(b+d)^2}$. This proves Lemma 2.22. □

Next, we will deal with $|\boldsymbol{n}| > 10N_1$ by combining the short-range property of S and Bourgain's geometric lemma (cf. Lemma 2.5). For $\Lambda \subset \mathbb{Z}^{b+d}$, denote by $\Pi_b\Lambda$ (resp. $\Pi_d\Lambda$) the projection of Λ on $\mathbb{Z}^b$ (resp. $\mathbb{Z}^d$). For each $\boldsymbol{k} \in \Pi_b\Lambda$, define

$$\Lambda(\boldsymbol{k}) = \{\boldsymbol{n} \in \mathbb{Z}^d : \ (\boldsymbol{k}, \boldsymbol{n}) \in \Lambda\}.$$

We have

Lemma 2.23. *Assume that $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{W}'$ (cf. Lemma 2.5) and $L \geq 10^{-\frac{32}{\kappa_1^2}} \log^{\frac{4}{\kappa_1}}(\log \frac{1}{\varepsilon})$. We have*

(1) *Fix $\Lambda \in (\boldsymbol{k}', \boldsymbol{n}') + \mathcal{ER}_{\boldsymbol{0}}(L)$ with $|\boldsymbol{n}'| > 10L$. There is a collection of $\{\lambda_\ell = \lambda_\ell(\boldsymbol{\alpha}, \boldsymbol{\theta})\}_{1 \leq \ell \leq L^{b+2d}}$ depending only on $\Lambda, \boldsymbol{\alpha}, \boldsymbol{\theta}, V$ so that, if*

$$\min_{\boldsymbol{k} \in \Pi_b\Lambda, 1 \leq \ell \leq L^{b+2d}, \xi = \pm 1} |\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega} + \xi\lambda_\ell| > e^{-\frac{1}{4}L^{\frac{3\kappa_1}{4}}},$$

then Λ is σ-good with $\gamma_L = \gamma$, where κ_1 is given in Lemma 2.5.

(2) *Let $L \leq N \leq e^{\log^2 L}$. There exists some $\tilde{\Sigma}_L \subset \mathbb{R}$ with $\mathrm{meas}(\tilde{\Sigma}_L) \leq e^{-\frac{1}{5}L^{\frac{3\kappa_1}{4}}}$ so that, if $\sigma \notin \tilde{\Sigma}_L$, then each $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(L)$ satisfying $|(\boldsymbol{k}, \boldsymbol{n})| \leq 100N^2$ and $|\boldsymbol{n}| > 10L$ is σ-good with $\gamma_L = \gamma$.*

Proof. (1) Let $\boldsymbol{k} \in \Pi_b\Lambda$. *It is important that either $\Lambda(\boldsymbol{k}) \in \mathcal{ER}_{\mathbb{Z}^d}(L)$, or $\Lambda(\boldsymbol{k})$ is a rectangle of width at least L (cf. Lemma B.8).* In each of the two cases, we have $\Lambda(\boldsymbol{k}) \in \mathcal{E}_{2L}^L$.

For $Q \subset \Lambda(\boldsymbol{k})$, write

$$\begin{aligned} A_{\boldsymbol{k},Q} = A_{\boldsymbol{k},Q}(\sigma) &= R_Q(D(\sigma) + \varepsilon(\Delta \oplus \Delta))R_Q \\ &= (-\sigma - \boldsymbol{k} \cdot \boldsymbol{\omega} + \mathcal{L}_Q(\boldsymbol{\theta})) \oplus (\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega} + \mathcal{L}_Q(\boldsymbol{\theta})), \end{aligned}$$

where $\mathcal{L}_Q(\boldsymbol{\theta}) := \mathcal{H}_Q(\boldsymbol{\alpha}, \boldsymbol{\theta})$ is defined in Lemma 2.5. Then the set of all eigenvalues of $A_{\boldsymbol{k},Q}$ is given by

$$\{\pm(\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega}) + \lambda_{\ell,Q}(\boldsymbol{\alpha}, \boldsymbol{\theta})\}_{\ell=1}^{\#Q},$$

where $\{\lambda_\ell := \lambda_{\ell,Q}(\boldsymbol{\alpha}, \boldsymbol{\theta})\}_{\ell=1}^{\#Q}$ denotes the spectrum of $\mathcal{L}_Q(\boldsymbol{\theta})$. So we have

$$\|A_{\boldsymbol{k},Q}^{-1}\| \leq \max_{\xi = \pm 1, 1 \leq \ell \leq \#Q} |\xi(\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega}) + \lambda_\ell|^{-1}.$$

We now can use Lemma 2.5 and resolvent identities. Since $\boldsymbol{\alpha} \in \mathcal{A} \subset \mathcal{A}_L$, applying (1) of Lemma 2.5 with $E_{\boldsymbol{k}} = \sigma + \boldsymbol{k} \cdot \boldsymbol{\omega}$ similar to [JLS20] (cf. pages 471–472, the proof of Theorem 3.7) gives the following:

For each $\boldsymbol{m} \in \Lambda(\boldsymbol{k})$, there exist $\frac{1}{4}L^{\kappa_1} \le \tilde{L} \le L^{\kappa_2}$, $Q_{\boldsymbol{m}} \in \mathcal{ER}_{\mathbb{Z}^d}(\tilde{L})$ and $\tilde{Q}_{\boldsymbol{m}}$, such that

$$Q_{\boldsymbol{m}} \subset \boldsymbol{m} + [-\tilde{L}, \tilde{L}]^d, \ \boldsymbol{m} + [-\tilde{L}^{\frac{1}{10d}}, \tilde{L}^{\frac{1}{10d}}]^d \subset \tilde{Q}_{\boldsymbol{m}},$$
$$\boldsymbol{m} \in Q_{\boldsymbol{m}} \subset \Lambda(\boldsymbol{k}), \mathrm{dist}(\boldsymbol{m}, \Lambda(\boldsymbol{k}) \setminus Q_{\boldsymbol{m}}) \ge \frac{\tilde{L}}{2}, \ \mathrm{diam}\, \tilde{Q}_{\boldsymbol{m}} \le 4\tilde{L}^{\frac{1}{10d}}.$$

Also for any $\boldsymbol{n} \in Q_{\boldsymbol{m}} \setminus \tilde{Q}_{\boldsymbol{m}}$, there exist $L_1 \sim (\log L)^{\frac{4}{\kappa_1}}$ and some $\mathcal{ER}_{\mathbb{Z}^d}(L_1) \ni W \subset Q_{\boldsymbol{m}} \setminus \tilde{Q}_{\boldsymbol{m}}$ such that

$$\boldsymbol{n} \in W, \ \mathrm{dist}(\boldsymbol{n}, Q_{\boldsymbol{m}} \setminus \tilde{Q}_{\boldsymbol{m}} \setminus W) \ge \frac{L_1}{2},$$

and

$$\|\mathcal{H}_W^{-1}(E_{\boldsymbol{k}}; \boldsymbol{\theta})\| \le e^{\sqrt{L_1}},$$
$$|\mathcal{H}_W^{-1}(E_{\boldsymbol{k}}; \boldsymbol{\theta})(\boldsymbol{n}; \boldsymbol{n}')| \le e^{-\frac{1}{2}|\log \varepsilon| \cdot |\boldsymbol{n} - \boldsymbol{n}'|} \text{ for } |\boldsymbol{n} - \boldsymbol{n}'| \ge L_1^{\frac{8}{9}}.$$

Next, assume that (set $Q = Q_{\boldsymbol{m}}$)

$$\min_{\boldsymbol{m} \in \Lambda(\boldsymbol{k}), \xi = \pm 1, 1 \le \ell \le \# Q_{\boldsymbol{m}}} |\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega} + \xi \lambda_\ell| > e^{-\frac{1}{4}\tilde{L}^{\frac{3}{4}}}.$$

Then applying Lemma B.6 of [JLS20] yields for

$$A_{\boldsymbol{k}} := A_{\boldsymbol{k}}(\sigma) = A_{\boldsymbol{k}, \Lambda(\boldsymbol{k})}(\sigma),$$

the following estimates

$$\|A_{\boldsymbol{k}}^{-1}\| \le e^{\frac{1}{2}L^{\frac{3}{4}}},$$
$$|A_{\boldsymbol{k}}^{-1}(\boldsymbol{n}; \boldsymbol{n}')| \le e^{-\frac{1}{4}|\log \varepsilon| \cdot |\boldsymbol{n} - \boldsymbol{n}'|} \text{ for } |\boldsymbol{n} - \boldsymbol{n}'| \ge L^{\frac{8}{9}}.$$

Note that

$$\tilde{G}_\Lambda(\sigma) = \bigoplus_{\boldsymbol{k} \in \Pi_b \Lambda} A_{\boldsymbol{k}}^{-1},$$

where

$$\tilde{G}_\Lambda(\sigma) = (R_\Lambda(D(\sigma) + \varepsilon(\Delta \oplus \Delta))R_\Lambda)^{-1}.$$

By taking into account all $\boldsymbol{k} \in \Pi_b \Lambda$, we obtain that if

$$\min_{\boldsymbol{k} \in \Pi_b \Lambda, \boldsymbol{m} \in \Lambda(\boldsymbol{k}), \xi = \pm 1, 1 \le \ell \le \# Q_{\boldsymbol{m}}} |\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega} + \xi \lambda_\ell| > e^{-\frac{1}{4}L^{\frac{3\kappa_1}{4}}}, \tag{2.29}$$

then

$$\|\tilde{G}_\Lambda(\sigma)\| \le e^{\frac{1}{2}L^{\frac{3}{4}}},$$
$$|\tilde{G}_\Lambda(\sigma)((\boldsymbol{k}, \boldsymbol{n}, \xi); (\boldsymbol{k}', \boldsymbol{n}', \xi'))| \le \delta_{\boldsymbol{k}, \boldsymbol{k}'} e^{-\frac{1}{4}|\log \varepsilon| \cdot |\boldsymbol{n} - \boldsymbol{n}'|} \text{ for } |\boldsymbol{n} - \boldsymbol{n}'| \ge L^{\frac{8}{9}}.$$

Note that the total number of λ_ℓ in (2.29) is at most

$$C(b, d)L^b L^d L^{\kappa_2} L^{\kappa_2 d} \ll L^{b+2d}.$$

Finally, since $|\boldsymbol{n}| > 10L$, we know that S has the decay estimate

$$|S((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| \le C_2(1+|\boldsymbol{x}-\boldsymbol{x}'|)^{C_2}e^{-6\gamma L}e^{-2\gamma|\boldsymbol{x}-\boldsymbol{x}'|}.$$

Then using the perturbation Lemma B.1 implies that Λ is σ-good assuming (2.29) holds true. In fact, we can apply Lemma B.1 (we hide the dependence on ξ,ξ') with $B=\delta S_\Lambda(\boldsymbol{x},\boldsymbol{x}')$, $A^{-1}(\boldsymbol{x},\boldsymbol{x}')=\tilde{G}_\Lambda(\boldsymbol{x},\boldsymbol{x}')$ and $c=2\gamma,\epsilon_1=e^{-\frac{1}{2}L^{\frac{3}{4}}},M=L^{\frac{8}{9}},\epsilon_2=C\delta e^{-6\gamma L}$. Since $\gamma\in(\frac{1}{2},10)$, we obtain for $L\ge N_0(b,d,C_2)\gg 1$,

$$C(b,d,C_2)L^{2(b+d)+C_2}\delta e^{-6\gamma L}e^{2\gamma L^{\frac{8}{9}}}e^{\frac{1}{2}L^{\frac{3}{4}}} \le \frac{1}{2}.$$

This verifies all assumptions of Lemma B.1. So using Lemma B.1, we get

$$\|G_\Lambda(\sigma)\| \le 2e^{\frac{1}{2}L^{\frac{3}{4}}} < e^{L^{\frac{3}{4}}},$$

and for $|\boldsymbol{x}-\boldsymbol{x}'|\ge L^{\frac{8}{9}}$ (since $\frac{1}{4}|\log\varepsilon|>2\gamma>1$),

$$\begin{aligned}|G_\Lambda(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| &\le |\tilde{G}_\Lambda(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))|+e^{\frac{1}{2}L^{\frac{3}{4}}}e^{-2\gamma|\boldsymbol{x}-\boldsymbol{x}'|}\\ &\le e^{L^{\frac{3}{4}}-L^{\frac{8}{9}}}e^{-\gamma|\boldsymbol{x}-\boldsymbol{x}'|}\\ &\le e^{-\gamma|\boldsymbol{x}-\boldsymbol{x}'|}.\end{aligned}$$

This proves (1) of Lemma 2.23.

Remark 2.24.

- Let $L\le N\le e^{(\log L)^2}$. We would also like to remark that if we take into account all these $\Lambda\in(\boldsymbol{k},\boldsymbol{n})+\mathcal{ER}_{\boldsymbol{0}}(L)$ with $|(\boldsymbol{k},\boldsymbol{n})|\le 100N^2$ and $|\boldsymbol{n}|>10L$ (the total number of these Λ is at most $C(b,d)N^{2(b+d)}$), then we get a sequence of real numbers $\{\lambda_\ell(\boldsymbol{\alpha},\boldsymbol{\theta})\}_{1\le\ell\le N^{4(b+d)}}$. As a result, there is a set $\tilde{\Sigma}_L\subset\mathbb{R}$ satisfying

$$\operatorname{meas}(\tilde{\Sigma}_L)\le e^{C(b,d)\log^2 L}e^{-\frac{1}{4}L^{\frac{3}{4}\kappa_1}} \le e^{-\frac{1}{5}L^{\frac{3}{4}\kappa_1}}$$

  so that, for $\sigma\notin\tilde{\Sigma}_L$, all $\Lambda\in(\boldsymbol{k},\boldsymbol{n})+\mathcal{ER}_{\boldsymbol{0}}(L)$ satisfying $|(\boldsymbol{k},\boldsymbol{n})|\le 100N^2$ and $|\boldsymbol{n}|>10L$ are σ-good.
- In the application, we typically choose $L=N_1\ge 10^{-\frac{32}{\kappa_1^2}}\log^{\frac{4}{\kappa_1}}(\log\frac{1}{\varepsilon})$. It then follows from $N_1\sim N^{\rho^2}\ge(\varepsilon+\delta)^{-c_1\rho^2}$ that a restriction like $\delta\le\log^{-1}\frac{1}{\varepsilon}$ is needed.

(2) The conclusion is just that in Remark 2.24.
This completes the proof of Lemma 2.23. □

By combining Lemma 2.22 and Lemma 2.23, we are ready to prove LDE at the scale N (indeed all scales in $[N,N^2]$). We first recall an important lemma.

Lemma 2.25 (Matrix-valued Cartan's lemma, cf. Proposition 14.1 in [Bou05a]). *Let $T(\sigma)$ be a self-adjoint $N\times N$ matrix-valued function of a parameter $\sigma\in[-\beta,\beta]$ satisfying the following conditions:*

(i) *$T(\sigma)$ is real analytic in σ and has a holomorphic extension to*

$$\mathbb{D}=\{z\in\mathbb{C}:\ |\Re z|\le\beta,\ |\Im z|\le\beta_1\}$$

satisfying $\sup_{z\in\mathbb{D}}\|T(z)\|\le K_1,\ K_1\ge 1.$

(ii) *For each $\sigma \in [-\beta, \beta]$, there is a subset $Y \subset [1, N]$ with $\#Y \le M$ such that*

$$\|(R_{[1,N]\setminus Y} T(\sigma) R_{[1,N]\setminus Y})^{-1}\| \le K_2, \ K_2 \ge 1.$$

(iii) *Assume*

$$\operatorname{meas}\left(\{\sigma \in [-\beta, \beta] : \ \|T^{-1}(\sigma)\| \ge K_3\}\right) \le 10^{-3}\beta_1 (1+K_1)^{-1}(1+K_2)^{-1}.$$

Let $0 < \epsilon \le (1 + K_1 + K_2)^{-10M}$. Then we have

$$\operatorname{meas}\left(\left\{\sigma \in [-\beta/2, \beta/2] : \ \|T^{-1}(\sigma)\| \ge \frac{1}{\epsilon}\right\}\right) \le C e^{\frac{-c \log \frac{1}{\epsilon}}{M \log(M+K_1+K_2+K_3)}}, \tag{2.30}$$

where $C, c > 0$ are some absolute constants.

Now, we can prove Lemma 2.20.

Proof of Lemma 2.20. Recalling

$$N_2 = N_1^{\frac{2}{\rho}}, \ N = N_1^{\frac{1}{\rho^2}},$$

we assume that for $L = N_1, N_2$ and $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(L)$ with $|\boldsymbol{n}| \le 10L$, $G_\Lambda(\sigma)$ satisfies the LDE.

Without loss of generality, we only establish LDE for the Green's function on $\Lambda_{\mathrm{pm},N} = \{(\boldsymbol{x}, \xi) \in \mathbb{Z}^{b+d}_{\mathrm{pm},*} : \ |\boldsymbol{x}| \le N\} = (\Lambda_N \times \{+,-\}) \cap \mathbb{Z}^{b+d}_{\mathrm{pm},*}$, the other cases (i.e., $\Lambda \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_3)$ for $|\boldsymbol{n}| \le 10N_3$, $N_3 \in [N, N^2]$) can be dealt with similarly. We define three intermediate scales between $[N_2, N]$:

$$L_1 = N^{\frac{1}{100(b+d)^2}}, \ L_2 = N^{\frac{1}{50(b+d)^2}}, \ L_3 = N^{\frac{1}{10(b+d)}}. \tag{2.31}$$

Note that from $0 < \rho \le c_2(b, d) \le \frac{1}{10^4(b+d)^2}$ (cf. Lemma 2.22), we have

$$\log N_1 \ll \log N_2 \ll \log L_1 = \frac{1}{100(b+d)^2} \log N.$$

We first cover the $\Lambda_{\mathrm{pm},N}$ with two overlapped subregions, $\Lambda_{\mathrm{pm},N} = \mathcal{R}_1 \cup \mathcal{R}_2$ satisfying

$$\mathcal{R}_1 = \{(\boldsymbol{k}, \boldsymbol{n}, \xi) \in \Lambda_{\mathrm{pm},N} : \ |\boldsymbol{n}| \le L_2\},$$
$$\mathcal{R}_2 = \{(\boldsymbol{k}, \boldsymbol{n}, \xi) \in \Lambda_{\mathrm{pm},N} : \ |\boldsymbol{n}| \ge L_1\}.$$

We will apply matrix-valued Cartan's lemma (cf. Lemma 2.25) and Bourgain's geometric lemma [Bou07] (cf. Lemma 2.5) to obtain the sub-exponential growth estimates on Green's functions. Then we apply the coupling lemmas of [BGS02, JLS20, Liu22] (cf., e.g., Lemmas B.3, B.5, B.6) based on the iteration of resolvent identities to prove the off-diagonal exponential decay of $G_{\Lambda_{\mathrm{pm},N}}(\sigma)$ if desired σ are removed.

Before we enter the details, let us first explain why we need to divide $\Lambda_{\mathrm{pm},N}$ into the above two regions. This is quite different from that in [SW23]. Based on Lemma 2.23, we can impose the so-called weak second Melnikov's condition on $\boldsymbol{\omega}$ so that all σ-bad (cf. Definition 2.21) $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1) \subset \Lambda_{\mathrm{pm},N}$ can stay only in a **narrow strip region** (cf. Figure 1), namely, there is some $\boldsymbol{k}_* \in [-N, N]^b$ so that

$$|\boldsymbol{k} - \boldsymbol{k}_*| \le 10N_1 \le N^{0+}.$$

This gives the desired control on σ-bad Λ in the $\boldsymbol{k}$-direction. To establish the sublinear bound, it remains to handle the $\boldsymbol{n}$-direction. In [SW23], the $\boldsymbol{n}$-direction estimates can be obtained directly using the fact that $\theta \in \mathbb{R}$, Diophantine $\boldsymbol{\omega}$ and LDT for the corresponding quasi-periodic Schrödinger operators. Such arguments inevitably fail for $\boldsymbol{\theta} \in \mathbb{R}^d$: There is simply no one-dimensional interval decomposition of d-dimensional semi-algebraic sets, and more essentially, the Diophantine property of $\boldsymbol{\omega}$ cannot ensure the sublinear estimate in the $\boldsymbol{n}$-direction. This motivates us to restrict $|\boldsymbol{n}| \le L_2$ (i.e., the region $\mathcal{R}_1$). We can use the Cartan's lemma and resolvent identities to deal with $\mathcal{R}_1$. For $\mathcal{R}_2$, we take advantage of Bourgain's geometric lemma and the short-range property of H using Lemma 2.23. Finally, it suffices to cover $\Lambda_{\mathrm{pm},N}$ with **four types** of regions of sizes $N_1, L_1^{\frac{9}{10}}, L_1, L_3$ (will be specified below) contained in $\mathcal{R}_1, \mathcal{R}_2$, and apply the resolvent identities again. See Figure 1.

We first deal with $\mathcal{R}_2$, and then $\mathcal{R}_1$.

Analysis in $\mathcal{R}_2$

In this region, we can apply directly the conclusion of Lemma 2.23. In this case, we make no restriction on $\boldsymbol{\omega}$. We let

$$L_0 = L_1^{\frac{9}{10}} = N^{\frac{9}{1000(b+d)^2}}.$$

For any $\Lambda \subset \mathcal{R}_2$ satisfying $\Lambda = (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}(L_0)$, we have $|\boldsymbol{n}| > L_1 > 10L_0$. Then applying Lemma 2.23 with $L = L_0$ and since $0 < \rho < c_2(b,d)\kappa_1 \le \frac{\kappa_1}{10^4(b+d)^2}$, we get a set $\tilde{\Sigma}_{N,1} \subset \mathbb{R}$ satisfying

$$\mathrm{meas}(\tilde{\Sigma}_{N,1}) \le e^{-\frac{1}{5}L_0^{\frac{3\kappa_1}{4}}} \le e^{-\frac{1}{5}N^{\frac{27\kappa_1}{4000(b+d)^2}}} < e^{-N^{3\rho}},$$

so that, if $\sigma \notin \tilde{\Sigma}_{N,1}$, all $\Lambda \in \mathcal{ER}(L_0)$ contained in $\mathcal{R}_2$ are σ-good (cf. Definition 2.21).

We remark that in this case, we do not use Cartan's lemma.

Analysis in $\mathcal{R}_1$

In this region, we need to make further restriction on $\boldsymbol{\omega}$ called the weak second Melnikov's condition. More precisely, we have

Lemma 2.26. *There is a sequence of real numbers $\{\lambda_\ell = \lambda_\ell(\boldsymbol{\alpha}, \boldsymbol{\theta})\}_{1 \le \ell \le N^{4(b+d)}}$ depending only on $V, \boldsymbol{\alpha}, \boldsymbol{\theta}$ so that the following holds true. If $\boldsymbol{\omega} \in \widetilde{\Omega}_N$ with*

$$\widetilde{\Omega}_N := \bigcap_{1 \le \ell, \ell' \le N^{4(b+d)}, \xi = \pm 1, \xi' = \pm 1, 0 < |\boldsymbol{k}| \le 2N^2} \left\{ \boldsymbol{\omega} \in \Omega : \ |\boldsymbol{k} \cdot \boldsymbol{\omega} + \xi \lambda_\ell - \xi' \lambda_{\ell'}| > 10 e^{-\frac{1}{4}N^{\frac{3\kappa_1\rho^2}{4}}} \right\}, \tag{2.32}$$

then for any $\sigma \in \mathbb{R}$, there is some $\boldsymbol{k}_* \in [-N^2, N^2]^b$ so that, all σ-bad $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(N_1)$ satisfying $|\boldsymbol{k}| \le N^2$ and $10N_1 \le |\boldsymbol{n}| \le 100N^2$ must obey

$$\Pi_b \Lambda \subset [\boldsymbol{k}_* - 10N_1, \boldsymbol{k}_* + 10N_1]^b. \tag{2.33}$$

In particular, we have

$$\mathrm{meas}(\Omega \setminus \widetilde{\Omega}_N) \le e^{-\frac{1}{5}N^{\frac{3\kappa_1\rho^2}{4}}}. \tag{2.34}$$

Remark 2.27. Indeed, in this lemma, $\lambda_\ell(\boldsymbol{\alpha}, \boldsymbol{\theta})$ ($1 \le \ell \le N^{4(b+d)}$) are eigenvalues of the operator $\mathcal{H}_Q(\boldsymbol{\alpha}, \boldsymbol{\theta})$ for all $Q \subset [-100N^2, 100N^2]^d$ satisfying $Q \in \mathcal{ER}_{\mathbb{Z}^d}(\tilde{N}_1)$ for some $\tilde{N}_1 \in [\frac{1}{4}N_1^{\kappa_1}, N_1^{\kappa_2}]$.

Proof of Lemma 2.26. The proof follows from Lemma 2.23. Applying (1) of Lemma 2.23 (cf. also Remark 2.24) with $L = N_1$ and taking into account of all $\Lambda \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(N_1)$ with $|(\boldsymbol{k}, \boldsymbol{n})| \le 100N^2$ and $|\boldsymbol{n}| > 10N_1$ yield a sequence $\{\lambda_\ell\}_{\le \ell \le N^{4(b+d)}} \subset \mathbb{R}$ and each λ_ℓ depending only on $\boldsymbol{\alpha}, \boldsymbol{\theta}$ (but not on

$\sigma, \boldsymbol{\omega}$), so that if $\Lambda = (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\mathbf{0}}(N_1)$ satisfies $|(\boldsymbol{k}, \boldsymbol{n})| \le 100N^2$, $|\boldsymbol{n}| > 10N_1$ and

$$\min_{\boldsymbol{k} \in \Pi_b \Lambda, \xi = \pm 1, 1 \le \ell \le N^{4(b+d)}} |\sigma + \boldsymbol{k} \cdot \boldsymbol{\omega} + \xi \lambda_\ell| > e^{-\frac{1}{4} N_1^{\frac{3\kappa_1}{4}}} = e^{-\frac{1}{4} N^{\frac{3\kappa_1 \rho^2}{4}}},$$

then Λ is σ-good. So recalling (2.32), we have

$$\operatorname{meas}(\Omega \setminus \widetilde{\Omega}_N) \le N^{C(b,d)} e^{-\frac{1}{4} N^{\frac{3\kappa_1 \rho^2}{4}}} \le e^{-\frac{1}{5} N^{\frac{3\kappa_1 \rho^2}{4}}},$$

which implies (2.34).

In the following, we assume $\boldsymbol{\omega} \in \widetilde{\Omega}_N$.

Suppose now Λ' with $\Lambda' \in (\boldsymbol{k}', \boldsymbol{n}') + \mathcal{ER}_{\mathbf{0}}(N_1)$ satisfying $|\boldsymbol{k}'| \le N^2$ and $10N_1 < |\boldsymbol{n}'| \le 100N^2$ is σ-bad. Then there are some $\boldsymbol{k}_* \in \Pi_b \Lambda' \subset [-N^2, N^2]^b$, $\xi_* = \pm 1$ and some $\ell_* \in [1, N^{4(b+d)}]$ so that

$$|\sigma + \boldsymbol{k}_* \cdot \boldsymbol{\omega} + \xi_* \lambda_{\ell_*}| \le e^{-\frac{1}{4} N_1^{\frac{3\kappa_1}{4}}}.$$

Now let Λ'' (satisfying $\Lambda'' \in (\boldsymbol{k}'', \boldsymbol{n}'') + \mathcal{ER}_{\mathbf{0}}(N_1)$ with $|\boldsymbol{k}''| \le N^2$ and $10N_1 < |\boldsymbol{n}''| \le 100N^2$) be another σ-bad region. From $\boldsymbol{\omega} \in \widetilde{\Omega}_N$, we must have

$$\boldsymbol{k}_* \in \Pi_b \Lambda''.$$

Otherwise, there must be some $\boldsymbol{k}''' \in \Pi_b \Lambda'', \xi' \in \{\pm 1\}, \ell'$ so that, $\boldsymbol{k}_* \ne \boldsymbol{k}'''$ and

$$|\sigma + \boldsymbol{k}''' \cdot \boldsymbol{\omega} + \xi' \lambda_{\ell'}| \le e^{-\frac{1}{4} N_1^{\frac{3\kappa_1}{4}}},$$

which contradicts $\boldsymbol{\omega} \in \widetilde{\Omega}_N$. We have established that all σ-bad elementary N_1-regions Λ with centers $(\boldsymbol{k}, \boldsymbol{n})$ satisfying $|\boldsymbol{k}| \le N^2$ and $10N_1 < |\boldsymbol{n}| \le 100N^2$ must satisfy (2.33). □

In the following, we always assume $\boldsymbol{\omega} \in \widetilde{\Omega}_N$. Under this condition, we first give a geometric construction of the a region $\Lambda^\dagger$, on which the worst resonances appear. It turns out that this region has the **annulus structure** similar to [Bou07, JLS20].

Lemma 2.28. *Let $\omega \in \widetilde{\Omega}_N$ and $\sigma \in \mathbb{R}$. Let $\boldsymbol{k}_* \in [-N, N]^b$ (if it exists) be defined by Lemma 2.26. Then there are some $\Lambda_1^\dagger \subset \Lambda^\dagger \subset \mathcal{R}_1$ satisfying* $\operatorname{diam} \Lambda_1^\dagger \le 4L_1, \boldsymbol{k}_* \in \Lambda_1^\dagger$ *and*

$$\Lambda^\dagger = ((\boldsymbol{k}_* + [-L_3 + O(L_1), L_3 + O(L_1)]^b) \times [-L_2, L_2]^d) \cap \mathcal{R}_1,$$

so that the following holds true (recalling $L_1 \ll L_2 \ll L_3$ given by (2.31)). For any $\boldsymbol{x} \in \Lambda^\dagger \setminus \Lambda_1^\dagger$, there are some $L \in \{N_1, L_1, L_0\}$ and some $\mathcal{ER}(L) \ni W \subset \Lambda^\dagger \setminus \Lambda_1^\dagger$ so that

$$\boldsymbol{x} \in W, \ \operatorname{dist}(\boldsymbol{x}, \Lambda^\dagger \setminus \Lambda_1^\dagger \setminus W) \ge \frac{L}{2}.$$

Proof. The proof is similar to that of Theorem 3.7 of [JLS20] (cf. pages 471–472). We initially set $Q^\dagger = (\boldsymbol{k}_* + [-L_3, L_3]^b) \times [-L_2, L_2]^d$ and $Q_1 = (\boldsymbol{k}_* + [-L_1, L_1]^b) \times [-L_1, L_1]^d$. Then we slightly change $Q^\dagger, Q_1^\dagger$ to $\Lambda^\dagger, \Lambda_1^\dagger$ with the diameter modulations of length $2L_1$ if the boundary of $Q^\dagger, Q_1^\dagger$ approaches the boundary of $\mathcal{R}_1$ with the distance of $0 < L \le 2L_1$.

Next, let $\boldsymbol{x} = (\boldsymbol{k}, \boldsymbol{n}) \in \Lambda^\dagger \setminus \Lambda_1^\dagger$. For $|\boldsymbol{n}| \le 10N_1$, we use L_1 size regions to cover $\boldsymbol{x}$; for $10N_1 < |\boldsymbol{n}| \le L_1$, we use the N_1 size regions to cover $\boldsymbol{x}$; for $L_1 < |\boldsymbol{n}| \le L_2$, we use the L_0 size regions to cover $\boldsymbol{x}$. □

Next, we show that all $Q \in (\boldsymbol{k}, \mathbf{0}) + \mathcal{ER}_{\mathbf{0}}(L_1)$ satisfying $|\boldsymbol{k}| \le N$ and $|\boldsymbol{k} - \boldsymbol{k}_*| > 10N_1$ are σ-good for "most" σ, by using Cartan's lemma and coupling lemma of [Liu22]. More precisely, we have

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Lemma 2.29. *Let $\omega \in \widetilde{\Omega}_N \cap \Omega_{N_2}$. There is some $\tilde{\Sigma}_{N,2} \subset \mathbb{R}$ satisfying*

$$\mathrm{meas}(\tilde{\Sigma}_{N,2}) \leq e^{-N^{3\rho}},$$

so that for $\sigma \notin \tilde{\Sigma}_{N,2}$, the following holds true. If $Q \subset \mathcal{R}_1$ satisfies

$$Q \in \mathcal{Q} := \bigcup_{|\boldsymbol{k}-\boldsymbol{k}_*|>10N_1, |\boldsymbol{k}|\leq N} \{(\boldsymbol{k},\boldsymbol{0}) + Q : \ Q \in \mathcal{ER}_{\boldsymbol{0}}(L_1)\},$$

then Q is σ-good with $\gamma_{L_1} = \gamma_{N_1} - N_1^{-\vartheta}$, where $\vartheta > 0$ is an absolute constant defined in Lemma B.3.

Proof of Lemma 2.29. We choose any $Q \in \mathcal{ER}(L_1)$ with $Q \in \mathcal{Q}$. Then we take any $Q' \in \mathcal{R}_{L_*}^{\sqrt{L_1}}$ satisfying $Q' \subset Q$ and $\sqrt{L_1} \leq L_* \leq 2L_1$. From Lemma 2.22 and $\omega \in \widetilde{\Omega}_N$, we have for any $\sigma \in \mathbb{R}$,

$$\#\{\Lambda \in \mathcal{ER}(N_1) : \ \Lambda \text{ is } \sigma\text{−bad}, \ \Lambda \subset Q'\} \leq N^{\frac{1}{10^3(b+d)^2}}.$$

Denote by $Y_1(\sigma) := \bigcup_{1\leq \ell \leq k} Q_{\boldsymbol{x}_\ell}, k \leq N^{\frac{1}{10^3(b+d)^2}}$ all those σ-bad N_1 size elementary regions. We claim that, there is some $Y = Y(\sigma) \subset Q'$ satisfying $\#Y \leq C(b,d)N_1^{b+d}N^{\frac{1}{10^3(b+d)^2}}$, so that if $\boldsymbol{x} \in Q' \setminus Y$, then there is some σ-good $W \in \mathcal{ER}(N_1), W \subset Q' \setminus Y$ satisfying

$$\boldsymbol{x} \in W, \ \mathrm{dist}(\boldsymbol{x}, Q' \setminus Y \setminus W) \geq \frac{N_1}{2}.$$

Indeed, we pave Q' with Λ_α for $\Lambda_\alpha = Q' \cap \Lambda_{10N_1}(\boldsymbol{x}), \boldsymbol{x} \in 10N_1\mathbb{Z}^{b+d}$. Then we can enlarge Λ_α to $\tilde{\Lambda}_\alpha \subset Q'$ so that $\tilde{\Lambda}_\alpha$ has width at least N_1 and diameter at most $C(b,d)N_1$, and $\bigcup_\alpha \tilde{\Lambda}_\alpha$ remains a tiling of Q'. It suffices to take

$$Y(\sigma) = \bigcup_{\alpha: \ \tilde{\Lambda}_\alpha \cap Y_1(\sigma) \neq \emptyset} \tilde{\Lambda}_\alpha.$$

Since each $Q_{\boldsymbol{x}_\ell}$ can only intersect with at most $C(b,d)$ many $\tilde{\Lambda}_\alpha$, we have

$$\#Y(\sigma) \leq C(b,d)N_1^{b+d}N^{\frac{1}{10^3(b+d)^2}}.$$

We are ready to apply Cartan's lemma (cf. Lemma 2.25). From Lemma B.1, any σ-good N_1 region remains essentially σ'-good if $|\sigma' - \sigma| \leq e^{-10\gamma N_1}$. So on an interval I of length at most $e^{-10\gamma N_1}$, we can fix Y to associate with the midpoint of I. In addition, we can assume $|\sigma| \leq C(b,d)N$ since if $|\sigma| > CN$, then $\|D_{Q'}^{-1}(\sigma)\| \leq 1$. In this case, Q' becomes σ-good via the Neumann series argument. We apply Lemma 2.25 with

$$T(\sigma) = H_{Q'}(\sigma), \ \beta = \beta_1 = e^{-10\gamma N_1}.$$

It remains to verify the assumptions of Lemma 2.25. First, one has by the definition of the operator $H(\sigma)$,

$$K_1 = CN, C = C(V,b,d,C_2) > 0.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

We let Y be defined in the above claim. Then using sublinear bound conclusion and Lemma B.5, we have

$$M = \#Y \le C(b,d)N^{\rho^2(b+d)}N^{\frac{1}{10^3(b+d)^2}} < N^{\frac{1}{999(b+d)^2}} \le L_*^{\frac{1}{4}}$$
$$(\text{since } \rho \le c_2 \le \frac{1}{10^4(b+d)^2},\ L_* \ge N^{\frac{1}{200(b+d)^2}}),$$
$$K_2 = e^{2N_1^{\frac{3}{4}}}.$$

To apply the matrix-valued Cartan's lemma, we need the scale

$$N_2 = N_1^{\frac{2}{\rho}}.$$

Recall that the LDE hold at scale N_2. Applying the resolvent identity Lemma B.5 and using (2) of Lemma 2.23 with $L = N_2$ yield

$$\|T^{-1}(\sigma)\| \le e^{2N_2^{\frac{3}{4}}} = K_3$$

for σ away from a set of measure at most (since $0 < \rho < c_2\kappa_1$)

$$CN^{C(b,d)}e^{-N_2^{\rho}} + e^{-\frac{1}{5}N_2^{\frac{3\kappa_1}{4}}} \le e^{-N_2^{\rho}/2}.$$

It follows from $N_1 = N_2^{\frac{\rho}{2}}$ that

$$\begin{aligned} 10^{-3}\beta_1(1+K_1)^{-1}(1+K_2)^{-1} &\ge e^{-10\gamma N_1}N^{-2}e^{-2N_1^{\frac{3}{4}}} \\ &> e^{-N_2^{\rho}/2}. \end{aligned}$$

This verifies (iii) of Lemma 2.25. If $\epsilon = e^{-L_*^{\frac{1}{2}}}$, then one has $\epsilon < (1+K_1+K_2)^{-10M}$. Cover $[-CN, CN]$ with disjoint intervals of length $e^{-10\gamma N_1}$. Define

$$\tilde{\Sigma}_{Q'} = \{\sigma \in \mathbb{R} : \ \|G_{Q'}(\sigma)\| \ge e^{L_*^{\frac{3}{4}}}\}.$$

Then by (2.30) of Lemma 2.25, one obtains

$$\operatorname{meas}(\tilde{\Sigma}_{Q'}) \le CN^{C}e^{10\gamma N_1}e^{-\frac{cL_*^{\frac{3}{4}}}{CN_1N_2L_*^{\frac{1}{4}}\log N}} \le e^{-L_*^{\frac{1}{3}}} \ll e^{-N^{3\rho}}.$$

Note that the number of $Q' \subset Q$ with $Q' \in \mathcal{R}_{L_*}^{\sqrt{L_1}}$, $\sqrt{L_1} \le L_* \le 2L_1$ is at most $L^{C(b,d)}$ and the number of $Q \in \mathcal{Q}$ is at most $N^{(b,d)}$. We take the union over all these $Q' \subset Q, Q \in \mathcal{Q}$ leading to the desired $\tilde{\Sigma}_{N,2} \subset \mathbb{R}$ with (since $0 < \rho < c_2 \le \frac{1}{10^4(b+d)^2}$)

$$\operatorname{meas}(\tilde{\Sigma}_{N,2}) \le N^{C(b,d)}e^{-L_1^{\frac{1}{6}}} < e^{-N^{\frac{1}{600(b+d)^2}}} \ll e^{-N^{3\rho}}.$$

To finish the proof, it suffices to prove the off-diagonal exponential decay of $G_Q(\sigma)$ for $Q \in \mathcal{Q}$ and $\sigma \notin \tilde{\Sigma}_{N,2}$. This follows from using the coupling lemma of [Liu22]. Indeed, fix $Q \in \mathcal{Q}$ (then $Q \in \mathcal{ER}(L_1)$). Let $\mathcal{F}$ be any family of pairwise disjoint $\sqrt{L_1}$ size elementary regions in Q. Then using Lemma 2.22 and $\omega \in \widetilde{\Omega}_N$, we get

$$\#\{\Lambda \in \mathcal{F} : \ \Lambda \text{ is } \sigma\text{−bad}\} \le N^{\frac{1}{10^3(b+d)^2}} \le L_1^{\frac{1}{9}}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

So applying Lemma B.3 leads to the off-diagonal exponential decay of $G_Q(\sigma)$ with the decay rate

$$\gamma_{L_1} = \gamma_{N_1} - N_1^{-\vartheta}. \qquad \square$$

We then deal with $G_{\Lambda^\dagger}(\sigma)$ with $\Lambda^\dagger$ given by Lemma 2.28. We have

Lemma 2.30. *Let $\omega \in \widetilde{\Omega}_N \cap \Omega_{N_2}$. Then there is some $\widetilde{\Sigma}_{N,3} \subset \mathbb{R}$ satisfying*

$$\mathrm{meas}(\tilde{\Sigma}_{N,3}) < e^{-N^{3\rho}}$$

so that, if $\sigma \notin \bigcup_{\ell=1,2,3} \tilde{\Sigma}_{N,\ell}$, then $G_{\Lambda^\dagger}(\sigma)$ is σ-good in the sense of Definition 2.21 (replacing Λ with $\Lambda^\dagger$) with $\gamma_{L_2} = \gamma_{L_1} - N_1^{-\frac{1}{10}}$.

Proof of Lemma 2.30. The proof is similar to that of Lemma 2.29, but is more in the spirit of [JLS20] when dealing with off-diagonal exponential decay estimate, as $\Lambda^\dagger$ has the special annulus structure.

To get the sub-exponential growth estimate of $\|G_{\Lambda^\dagger}(\sigma)\|$, the key point is to apply Cartan's lemma only involving N_1, N_2 sizes regions similar to the proof of Lemma 2.29. This means precisely that we do not choose $Y = \Lambda_1^\dagger$ when applying Lemma 2.25.

From $\omega \in \widetilde{\Omega}_N$, the definition of $\Lambda^\dagger$ (cf. Lemma 2.28) and Lemma 2.22, we have for any $\sigma \in \mathbb{R}$,

$$\begin{aligned}
&\#\{\Lambda \in \mathcal{ER}(N_1) : \ \Lambda \text{ is } \sigma\text{−bad}, \ \Lambda \subset \Lambda^\dagger\} \\
&\le N^{\frac{1}{10^3(b+d)^2}} + C(b,d) N_1^b L_2^d \\
&\le N^{\frac{1}{10^3(b+d)^2}} + C(b,d) N^{\rho^2 b + \frac{1}{50(b+d)}} \\
&\le 2N^{\rho^2 b + \frac{1}{50(b+d)}}.
\end{aligned}$$

Similar to the proof of Lemma 2.29, we can find some $Y = Y(\sigma) \subset \Lambda^\dagger$ satisfying $\#Y \le C(b,d) N^{2\rho^2(b+d)} N^{\frac{1}{50(b+d)}} < N^{\frac{1}{49(b+d)}}$ so that, if $\boldsymbol{x} \in \Lambda^\dagger \backslash Y$, then there is some σ-good $W \in \mathcal{ER}(N_1)$, $W \subset \Lambda^\dagger \setminus Y$ satisfying

$$\boldsymbol{x} \in W, \ \mathrm{dist}(\boldsymbol{x}, \Lambda^\dagger \setminus Y \setminus W) \ge \frac{N_1}{2}.$$

Let $\omega \in \widetilde{\Omega}_N$. Again, both the site $\boldsymbol{k}_* \in [-N,N]^b$ (given in Lemma 2.26) and $Y(\sigma)$ depend on σ. From Lemma B.1, any σ-good N_1 size region remains essentially σ'-good if $|\sigma' - \sigma| \le e^{-10\gamma N_1}$. So on an interval I of length at most $e^{-10\gamma N_1}$, we can choose $Y, \boldsymbol{k}_*$ to associate with the midpoint of I. Also, we can assume $|\sigma| \le C(b,d)N$ since if $|\sigma| > CN$, then $\|D_{\Lambda^\dagger}^{-1}(\sigma)\| \le 1$, and $\Lambda^\dagger$ becomes σ-good via the Neumann series argument.

We now apply Lemma 2.25 with

$$T(\sigma) = H_{\Lambda^\dagger}(\sigma), \ \beta = \beta_1 = e^{-10\gamma N_1}.$$

We have $K_1 = CN$. Using the sublinear bound conclusion and the resolvent identity Lemma B.5, we have

$$M = \#Y \le N^{\frac{1}{49(b+d)}} < L_3^{\frac{1}{4}}, \ K_2 = e^{2N_1^{\frac{3}{4}}}.$$

We again use the scale $N_2 = N_1^{\frac{2}{\rho}}$, at which the LDE hold. Similar to the proof of Lemma 2.29, we can verify all assumptions of Lemma 2.25. Define

$$\tilde{\Sigma}_{N,3} = \{\sigma \in \mathbb{R} : \ \|G_{\Lambda^\dagger}(\sigma)\| \ge e^{L_3^{\frac{3}{4}}}\}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Then by (2.30) of Lemma 2.25, one obtains

$$\mathrm{meas}(\tilde{\Sigma}_{N,3}) \leq CN^C e^{10\gamma N_1} e^{-\frac{cL_3^{\frac{3}{4}}}{CN_1N_2L_3^{\frac{1}{4}}\log N}} \leq e^{-L_3^{\frac{1}{3}}} \ll e^{-N^{3\rho}}.$$

Finally, it suffices to prove the off-diagonal exponential decay of $G_{\Lambda^\dagger}(\sigma)$. This follows directly from the resolvent identity of [JLS20], that is, Lemma B.6. In fact, by the analysis in $\mathcal{R}_2$, Lemma 2.29 and Lemma 2.28, all elementary regions contained in $\Lambda^\dagger \backslash \Lambda_1^\dagger$ with sizes N_1, L_0, L_1 are σ-good. Then applying Lemma B.6 with $\Lambda = \Lambda^\dagger, \Lambda_1 = \Lambda_1^\dagger$ (since diam $\Lambda^\dagger \sim L_3 > 100^{2(b+d)} L_1^{2(b+d)} \geq (\mathrm{diam}\, \Lambda_1^\dagger)^{2(b+d)}$) leads to the desired off-diagonal exponential decay estimate of $G_{\Lambda^\dagger}(\sigma)$ with

$$\gamma_{L_2} = \gamma_{L_1} - N_1^{-\frac{1}{10}} = \gamma_{N_1} - N_1^{-\vartheta} - N_1^{\frac{1}{10}}.$$

□

Application of the resolvent identity again

Finally, we define $\Sigma_N = \bigcup_{\ell=1,2,3} \tilde{\Sigma}_{N,\ell}$ with $\tilde{\Sigma}_{N,1}, \tilde{\Sigma}_{N,2}, \tilde{\Sigma}_{N,3}$ given by the analysis in $\mathcal{R}_2$, Lemma 2.29, Lemma 2.30, respectively. Then

$$\mathrm{meas}(\Sigma_N) \leq 3e^{-2N^{3\rho}} \ll e^{-N^{2\rho}},$$

and for $\sigma \notin \Sigma_N$, we can cover $\Lambda_{\mathrm{pm},N}$ with σ-good regions of sizes N_1, L_0, L_1, L_3. More precisely, we have (cf. Figure 1)

- For $\boldsymbol{x} \in \mathcal{R}_2$, cover $\boldsymbol{x}$ with $\Lambda \in \mathcal{ER}(L_0)$ with $L_0 \sim N^{\frac{9}{1000(b+d)^2}}$.
- For $\boldsymbol{x} \in \Lambda_1^\dagger$, cover $\boldsymbol{x}$ with $\Lambda^\dagger$.
- For $\boldsymbol{x} \in \mathcal{R}_1 \setminus \Lambda_1^\dagger$, cover $\boldsymbol{x}$ with regions in

$$\bigcup_{L=N_1,L_0,L_1} \mathcal{ER}(L).$$

As a result, for $\sigma \notin \Sigma_N$, we can apply Lemma B.5 and Lemma B.6 to show that $\Lambda_{\mathrm{pm},N}$ is σ-good. Indeed, we first apply Lemma B.5 with

$$\Lambda = \Lambda_{\mathrm{pm},N}, W \in \bigcup_{L=N_1,L_0,L_1} (\mathcal{ER}(L) \cup \{\Lambda^\dagger\})$$

to obtain

$$\|G_{\Lambda_{\mathrm{pm},N}}(\sigma)\| \leq e^{N^{\frac{3}{4}}}.$$

Next, for the off-diagonal exponential decay, we use Lemma B.6 with $\Lambda = \Lambda_{\mathrm{pm},N}$, $\Lambda_1 = \emptyset$, $M_0 = N_1$ to obtain (recalling the decay rate estimates in Lemmas 2.29, 2.30) for $|\boldsymbol{x} - \boldsymbol{x}'| > N^{\frac{8}{9}}$,

$$\begin{aligned} |G_{\Lambda_{\mathrm{pm},N}}(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| &\leq e^{-(\gamma_{L_2} - N_1^{-\frac{1}{10}})|\boldsymbol{x}-\boldsymbol{x}'|} \\ &\leq e^{-(\gamma_{N_1} - N_1^{-\vartheta} - 2N_1^{-\frac{1}{10}})|\boldsymbol{x}-\boldsymbol{x}'|} \\ &= e^{-\gamma_N|\boldsymbol{x}-\boldsymbol{x}'|}, \end{aligned}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

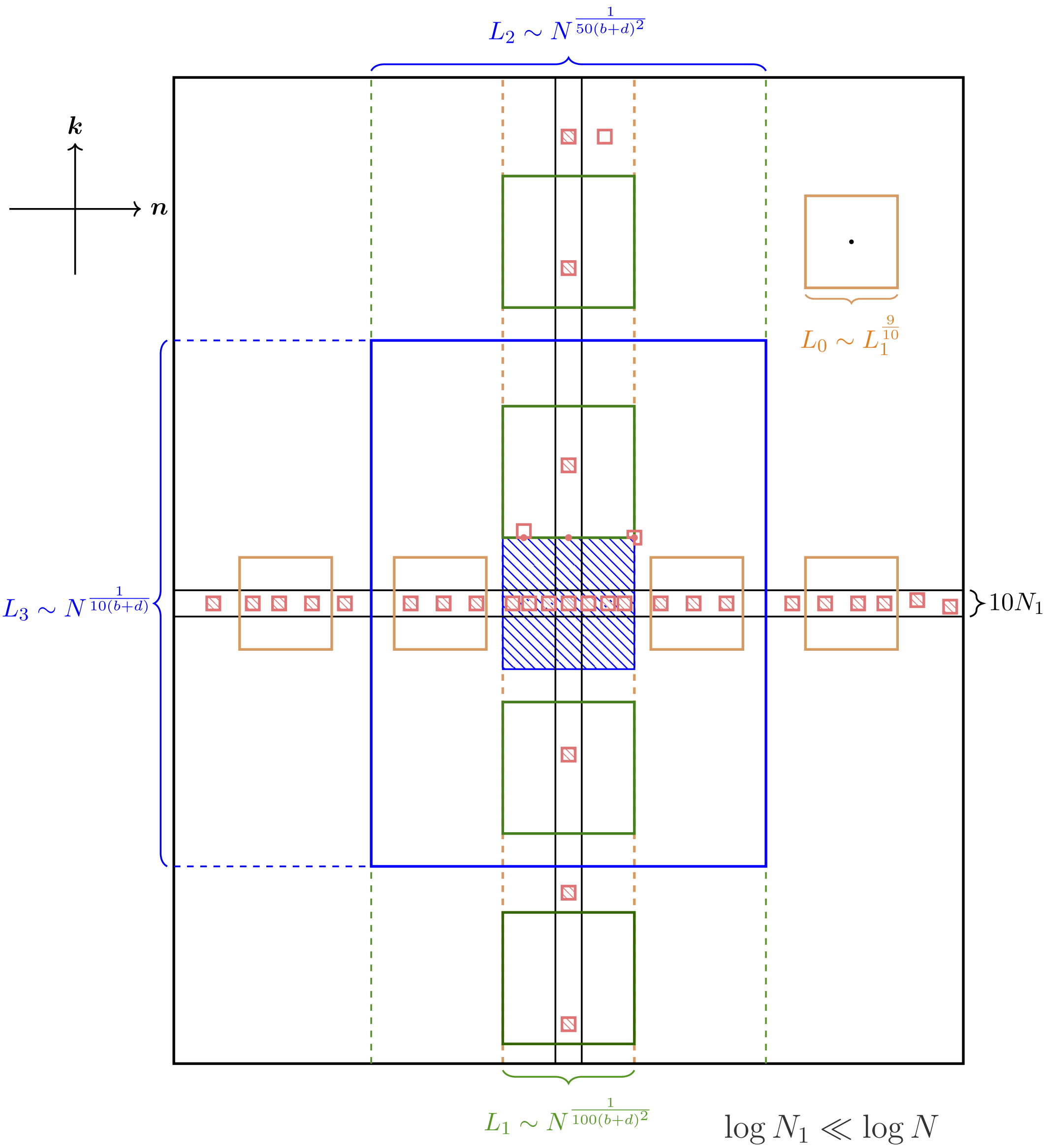

Figure 1. *The distribution of resonant blocks.*

where

$$\gamma_N \geq \gamma_{N_1} - N^{-\kappa}, \ \kappa = \frac{1}{2}\min\{\vartheta, \frac{1}{10}\} \leq \frac{1}{20}. \tag{2.35}$$

Since $N \sim N_1^{\frac{1}{\rho^2}}$, we choose ℓ_* so that $N^{\rho^{2\ell_*}} := N_* \sim \log^{\frac{1}{\kappa}} \frac{1}{\varepsilon+\delta}$. It follows from applying Lemma 2.15 (i.e., $\gamma_{N^{\rho^{2\ell_*}}} = \gamma_0 = \gamma - \log^{-2} \frac{1}{\varepsilon+\delta}$) and iterating (2.35) that

$$\begin{aligned}\gamma_N &\geq \gamma_{N_*} - \sum_{\ell=1}^{\ell_*} N_*^{-\frac{1}{\rho^{2\ell}}\kappa} \\ &\geq \gamma_{N^{\rho^{2\ell_*}}} - N_*^{-\frac{\kappa}{\rho^2}} - \sum_{\ell\geq 2} N_*^{-\frac{1}{\rho^{2\ell}}\kappa}\end{aligned}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

$$\begin{aligned} &\geq \gamma - \log^{-2} \frac{1}{\varepsilon+\delta} - 2N_*^{-\frac{\kappa}{\rho^2}} \\ &\geq \gamma - 3\log^{-2} \frac{1}{\varepsilon+\delta} \geq \frac{\gamma}{2} \end{aligned} \tag{2.36}$$

provided $0 < \varepsilon + \delta \leq c(\rho, c_1, \kappa)$.

The above arguments hold indeed for all scales in $[N, N^2]$. Then we can propagate the LDE from small scales interval $[N_1, N_2]$ to large scales one $[N, N^2]$ similar to the proof of Theorem 4.1 in [JLS20]. This completes the proof of Lemma 2.20. □

Proof of Theorem 2.11. The proof is just a combination of Lemmas 2.15, 2.17 and 2.20 by taking

$$0 < \varepsilon + \delta \leq \delta_0 = \min\{\delta_1, \delta_2, \delta_3\}.$$

□

3. Nonlinear analysis

In this section, we focus on nonlinear analysis. We work directly with the nonlinear equation (1.1). Recall that

$$\mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta}),\ \boldsymbol{\omega}^{(0)} = \left(\omega_\ell^{(0)} = \mu_{\boldsymbol{n}_\ell}\right)_{\ell=1}^{b}.$$

We aim to construct solutions to (1.1) of the form

$$u(t, \boldsymbol{n}) = \sum_{(\boldsymbol{k},\boldsymbol{n})\in\mathbb{Z}^b\times\mathbb{Z}^d} \hat{u}(\boldsymbol{k}, \boldsymbol{n}) e^{i\boldsymbol{k}\cdot\boldsymbol{\omega}t}.$$

Using the ansatz, we get the nonlinear lattice equation

$$\operatorname{diag}_{(\boldsymbol{k},\boldsymbol{n})}(-\boldsymbol{k}\cdot\boldsymbol{\omega} + \mu_{\boldsymbol{n}})\hat{u} + \varepsilon\Delta\hat{u} + \delta(\hat{u} * \hat{v})^{*p} * \hat{u} = 0, \tag{3.1}$$

where $\hat{v}(\boldsymbol{k}, \boldsymbol{n}) = \overline{\hat{u}(-\boldsymbol{k}, \boldsymbol{n})}$, the convolution is only in the $\boldsymbol{k}$-variable:

$$\hat{u} * \hat{v}(\boldsymbol{k}, \boldsymbol{n}) = \sum_{\boldsymbol{k}_1+\boldsymbol{k}_2=\boldsymbol{k}} \hat{u}(\boldsymbol{k}_1, \boldsymbol{n})\hat{v}(\boldsymbol{k}_2, \boldsymbol{n})$$

and

$$(\hat{u} * \hat{v})^{*p} = \underbrace{(\hat{u} * \hat{v}) * \cdots * (\hat{u} * \hat{v})}_{p \text{ times}}.$$

Note that Δ acts only on the $\boldsymbol{n}$-variable. To solve for $\hat{u}, \hat{v}$, we also need a conjugate equation of (3.1):

$$\operatorname{diag}_{(\boldsymbol{k},\boldsymbol{n})}(\boldsymbol{k}\cdot\boldsymbol{\omega} + \mu_{\boldsymbol{n}})\hat{v} + \varepsilon\Delta\hat{v} + \delta(\hat{u} * \hat{v})^{*p} * \hat{v} = 0.$$

As a result, we obtain the following vector-valued nonlinear lattice equation:

$$F(\vec{u}) = 0,\ \vec{u} = \begin{pmatrix}\hat{u}\\ \hat{v}\end{pmatrix}, \tag{3.2}$$

where

$$F(\vec{u}) = H_0\vec{u} + \delta P(\vec{u})$$

with

$$H_0 = \mathrm{diag}_{(\boldsymbol{k},\boldsymbol{n})\in\mathbb{Z}^{b+d}}\begin{pmatrix} -\boldsymbol{k}\cdot\boldsymbol{\omega}+\mu_{\boldsymbol{n}}+\varepsilon\Delta & 0 \\ 0 & \boldsymbol{k}\cdot\boldsymbol{\omega}+\mu_{\boldsymbol{n}}+\varepsilon\Delta \end{pmatrix}$$
$$:= D(0)+\varepsilon(\Delta\oplus\Delta) \tag{3.3}$$

and

$$P(\vec{u}) = \begin{pmatrix} (\hat{u}*\hat{v})^{*p}*\hat{u} \\ (\hat{u}*\hat{v})^{*p}*\hat{v} \end{pmatrix}.$$

The linearized operator of P at $\vec{u}$ is then given by

$$T_{\vec{u}} = \begin{pmatrix} (p+1)(\hat{u}*\hat{v})^{*p}* & p(\hat{u}*\hat{v})^{*(p-1)}*\hat{u}*\hat{u}* \\ p(\hat{u}*\hat{v})^{*(p-1)}*\hat{v}*\hat{v}* & (p+1)(\hat{u}*\hat{v})^{*p}* \end{pmatrix}. \tag{3.4}$$

3.1. The Lyapunov-Schmidt decomposition

We look for quasi-periodic solutions to (1.1) near

$$u^{(0)} = u^{(0)}(t,\boldsymbol{n}) = \sum_{\ell=1}^{b} a_\ell e^{i\omega_\ell^{(0)}t}\delta_{\boldsymbol{n}_\ell,\boldsymbol{n}},$$

using the Lyapunov-Schmidt decomposition. The Fourier support of $u^{(0)}$ is $\mathcal{S}_+ = \{(\boldsymbol{e}_\ell,\boldsymbol{n}_\ell)_{\ell=1}^b\} \subset \mathbb{Z}^{b+d}$ (we identify $\mathcal{S}_+$ with $\{(\boldsymbol{e}_\ell,\boldsymbol{n}_\ell)_{\ell=1}^b\}\times\{+\} \subset \mathbb{Z}^{b+d}_{\mathrm{pm}}$). Similarly, let $\mathcal{S}_- = \{(-\boldsymbol{e}_\ell,\boldsymbol{n}_\ell)_{\ell=1}^b\} \subset \mathbb{Z}^{b+d}$ be the Fourier support of $v^{(0)} := \overline{u^{(0)}}$, which is then identified with $\{(-\boldsymbol{e}_\ell,\boldsymbol{n}_\ell)_{\ell=1}^b\}\times\{-\} \subset \mathbb{Z}^{b+d}_{\mathrm{pm}}$. Let $\mathcal{S} = \mathcal{S}_+\cup\mathcal{S}_-$ and write $\mathbb{Z}^{b+d}_{\mathrm{pm},*} = \mathbb{Z}^{b+d}_{\mathrm{pm}}\setminus\mathcal{S} := \mathcal{S}^c$.

We divide the equation (3.2) into the P-equations

$$F(\vec{u})\big|_{\mathcal{S}^c} = 0, \tag{3.5}$$

and the Q-equations

$$F(\vec{u})\big|_{\mathcal{S}} = 0. \tag{3.6}$$

We iteratively solve the Q-equations and then the P-equations.

3.2. The Q-equations and extraction of parameters

The Q-equations are $2b$-dimensional.

On $\mathcal{S}$, $\vec{u}$ is held fixed, and the Q-equations (3.6) are viewed instead as equations for $\boldsymbol{\omega}$, and will be solved using the implicit function theorem. The Q-equations are used to relate the frequencies $\boldsymbol{\omega}$ to the amplitudes $\boldsymbol{a}$, permitting amplitude-frequency modulation to solve the P-equations.

Note that

$$u^{(0)}(t,\boldsymbol{n}) = \sum_{(\boldsymbol{k},\boldsymbol{n})\in\mathcal{S}_+} \hat{u}^{(0)}(\boldsymbol{k},\boldsymbol{n})e^{i\boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)}t}, \tag{3.7}$$

where

$$\hat{u}^{(0)}(\boldsymbol{e}_\ell,\boldsymbol{n}_\ell) = \hat{v}^{(0)}(-\boldsymbol{e}_\ell,\boldsymbol{n}_\ell) = a_\ell \text{ for } 1\le\ell\le b.$$

The solutions $\vec{u} = (\hat{u}, \hat{v})^t$ (with $(\cdot)^t$ denoting the transpose) on $\mathcal{S}$ are held *fixed*: $\vec{u} \equiv (\hat{u}^{(0)}, \hat{v}^{(0)})^t = \vec{u}^{(0)}$ on $\mathcal{S}$, the Q-equations are used instead to solve for the frequencies. Due to symmetry, the Q-equations for $\hat{v}$ are the same as those for $\hat{u}$. So we only need to consider those for $\hat{u}$. Solving the equations at the initial step

$$F(\vec{u}^{(0)})\big|_{\mathcal{S}_+} = 0$$

leads to, for $1 \le \ell \le b$, the initial modulated frequencies

$$\omega_\ell = \omega_\ell^{(1)} = \omega_\ell^{(0)} + \delta a_\ell^{2p} \tag{3.8}$$

satisfying

$$\det\left(\frac{\partial \boldsymbol{\omega}^{(1)}}{\partial \boldsymbol{a}}\right) = \delta^b \prod_{\ell=1}^{b} (2pa_\ell^{2p-1}).$$

Along the way, we will show that $\boldsymbol{\omega}^{(r)} = \boldsymbol{\omega}^{(0)} + O(\delta)$, for all $r = 0, 1, 2, \cdots$. So $\boldsymbol{\omega}^{(r)} \in \Omega$, the frequency set in Theorem 2.11, hence Theorem 2.11 is at our disposal to solve the P-equations. To construct approximate solutions to the P-equations, it is convenient to view $\boldsymbol{\omega} \in \Omega$ as an independent parameter, and work in the $(\boldsymbol{\omega}, \boldsymbol{a})$-variable, $(\boldsymbol{\omega}, \boldsymbol{a}) \in \Omega \times [1, 2]^b := (\boldsymbol{\omega}^{(0)} + [\delta, 2^{2p}\delta]^b) \times [1, 2]^b$. The approximate solutions to the nonlinear matrix equation (3.1) will then be obtained by taking into account the solutions to the Q-equations as well, which restricts $(\boldsymbol{\omega}, \boldsymbol{a})$ to the b-dimensional hypersurface $\boldsymbol{\omega} = \boldsymbol{\omega}^{(r)}(\boldsymbol{a})$, at the r-th iteration.

3.3. The P-equations

The P-equations are infinite-dimensional.

We use (3.5) to solve for $\vec{u}|_{\mathcal{S}^c}$. Solving the P-equations requires analyzing the invertibility of the linearized operator $H = H_0 + \delta T_{\vec{u}}$ restricted to $\mathbb{Z}^{b+d}_{\mathrm{pm},*}$, which we do in the previous section. Then the resolution of the P-equations combines the Newton scheme and the linear analysis in Section 2. The large deviation estimates in σ in Theorem 2.11 will be converted into estimates in the amplitudes $\boldsymbol{a}$ by using a semi-algebraic projection lemma (cf. Lemma 3.7 below), and amplitude-frequency modulation, $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{a})$. This is an established scheme, which has recently been extensively elaborated on with detailed proofs in sects. IV-VI, [KLW24] and [LW24]. They will serve as the basic reference point for the present section.

Recall first the formal Newton scheme. Let $\vec{u}$ be an approximate solution to the P-equations (3.5). For the next approximation $\vec{u}'$, write

$$\vec{u}' = \vec{u} + \Delta_{\mathrm{cor}}\vec{u}'.$$

Then $\Delta_{\mathrm{cor}}\vec{u}'$ is set to be

$$\Delta_{\mathrm{cor}}\vec{u}' = -(H(\vec{u}))^{-1} F(\vec{u}),$$

where the linearized operator $H(\vec{u}) = H_0 + \delta T_{\vec{u}}(\boldsymbol{\omega}, \boldsymbol{a})$ (cf. (3.3) and (3.4)). This is, however, only *indicative* since we assume that $H(\vec{u})$ is invertible. Due to the small divisor difficulties in the inversion process, we need to regularize the linearized operator $H(\vec{u})$ and do a *multi-scale* Newton iteration instead.

Toward that purpose, let M be a large integer and consider the geometric sequence of scales M^ℓ, $\ell = 1, 2, \cdots$. Denote by $\vec{u}^{(r)}$, the r-th approximate solution to the P-equations (3.5). For the $(r+1)$-th

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

approximation, write

$$\vec{u}^{(r+1)} = \vec{u}^{(r)} + \Delta_{\rm cor}\vec{u}^{(r+1)}.$$

We define (if it exists)

$$\Delta_{\rm cor}\vec{u}^{(r+1)} = -(R_N H(\vec{u}^{(r)}) R_N)^{-1} F(\vec{u}^{(r)}),$$

where $N = M^{r+1}$ and R_N is the restriction operator to the cube $\Lambda_N \cap \mathbb{Z}^{b+d}_{{\rm pm},*}$. This leads to estimate the inverse (i.e., the Green's function),

$$G_N(\vec{u}) := \left(R_N H(\vec{u}^{(r)}) R_N\right)^{-1}, \ r = 0, 1, 2, \cdots .$$

The following estimates are crucial for the convergence of the approximate solutions with exponential decay:

$$\|G_N(\vec{u}^{(r)})\| \le e^{(\log N)^C},$$
$$|G_N(\vec{u}^{(r)})((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \le e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x} - \boldsymbol{x}'| \ge (\log N)^C,$$

for $\boldsymbol{x} = (\boldsymbol{k}, \boldsymbol{n})$ and some $c, C > 0$, see sect. V, [KLW24].

The LDT, Theorem 2.11, in the previous section plays an essential role in the above Green's function estimates. The operator $T_{\vec{u}}$ plays the role of S in (2.1) (cf. Remark 2.13 for details). The following lemma implies that $T_{\vec{u}}$ satisfies the decay assumption of S (cf. also Remark 2.13). This will then enable us to use Theorem 2.11.

Lemma 3.1.

(1) *For all $\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{k}'' \in \mathbb{Z}^b$, $\boldsymbol{n}, \boldsymbol{n}' \in \mathbb{Z}^d$ and $\xi, \xi' \in \{+, -\}$, we have the Toeplitz property in the $\boldsymbol{k}$-variable:*

$$T_{\vec{u}}((\boldsymbol{k}' + \boldsymbol{k}, \boldsymbol{n}, \xi); (\boldsymbol{k}'' + \boldsymbol{k}, \boldsymbol{n}', \xi')) = T_{\vec{u}}((\boldsymbol{k}', \boldsymbol{n}, \xi); (\boldsymbol{k}'', \boldsymbol{n}', \xi')).$$

(2) *Assume that $|\vec{u}(\boldsymbol{k}, \boldsymbol{n})| \le e^{-c(|\boldsymbol{k}|+|\boldsymbol{n}|)}, c > 0$. Then we have for some $C = C(b, p) > 0$,*

$$|T_{\vec{u}}((\boldsymbol{k}, \boldsymbol{n}, \xi); (\boldsymbol{k}', \boldsymbol{n}', \xi'))| \le C(1 + |\boldsymbol{k} - \boldsymbol{k}'|)^C e^{-c|\boldsymbol{k}-\boldsymbol{k}'|-c|\boldsymbol{n}|}\delta_{\boldsymbol{n},\boldsymbol{n}'}.$$

Proof. The proof is similar to that of Lemma 5.1 in [LW24] and Proposition 4.1 in [LSZ25] (our case is easier since the convolution here involves only the $\boldsymbol{k}$-variable). We omit the details. □

3.4. The induction hypothesis and the proof

Let M be a large integer and recall that $\Lambda_R \subset \mathbb{Z}^{b+d}$ denotes the ℓ^∞-norm-induced cube of radius R and centered at the origin.

In the following, $C > 0$ is a large constant and $c > 0$ is a small one.

We begin with the analysis of initial induction steps, which relies on the Neumann series argument and Diophantine estimates (2.5) of Theorem 2.2.

3.4.1. The initial $2r_\star - r_0$ steps

We choose $1 < r_0 < r_\star$ such that

$$M^{r_0} \le \log \frac{1}{\varepsilon + \delta} \le M^{r_0+1}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

and

$$M^{2r_\star} \le (\varepsilon+\delta)^{-c_1} < M^{2r_\star+1}, \tag{3.9}$$

where $c_1 > 0$ is given in Theorem 2.11.

Let $\vec{u}^{(0)}(\boldsymbol{\omega}, \boldsymbol{a}) = \vec{u}^{(0)}(\boldsymbol{\omega}^{(0)}, \boldsymbol{a})$ with $\boldsymbol{\omega} = \boldsymbol{\omega}^{(1)}$ given by (3.8).

We start by constructing $\vec{u}^{(1)}_{\mathrm{in}}(\boldsymbol{\omega}, \boldsymbol{a})$. Obviously,

$$F(\vec{u}^{(0)}) = O(\varepsilon+\delta)$$

and (the support) supp $F(\vec{u}^{(0)}) \subset \Lambda_{M_0}$ for some $M_0 = M_0(p, \mathcal{S}) > 0$. Let $L_1 = M^{r_0+1} > M_0$. It suffices to estimate

$$H_{L_1}^{-1} = (R_{\Lambda_{L_1}}(D(0) + \varepsilon(\Delta \oplus \Delta) + \delta T_{\vec{u}^{(0)}})R_{\Lambda_{L_1}})^{-1}.$$

It follows from (2.5) of Theorem 2.2, $\boldsymbol{\omega} = \boldsymbol{\omega}^{(0)} + O(\delta)$ and the Neumann series argument (cf. Lemma B.1) that

$$\begin{aligned} \|H_{L_1}^{-1}\| &\le 2L_1^{C_1}, \\ |H_{L_1}^{-1}((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| &\le L_1^{C_1} e^{-\frac{\log\frac{1}{\varepsilon+\delta}}{2}|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } \boldsymbol{x} \neq \boldsymbol{x}' \in \Lambda_{L_1}, \end{aligned}$$

since $L_1 = M^{r_0+1} \ll (\varepsilon+\delta)^{-c_1}$. Then we have

$$\Delta_{\mathrm{cor}}\vec{u}^{(1)}_{\mathrm{in}} = -H_{L_1}^{-1}F(\vec{u}^{(0)}),$$

and consequently

$$\vec{u}^{(1)}_{\mathrm{in}} = \vec{u}^{(0)} + \Delta_{\mathrm{cor}}\vec{u}^{(1)}_{\mathrm{in}} = \vec{u}^{(0)} - H_{L_1}^{-1}F(\vec{u}^{(0)}).$$

Substituting $\vec{u}^{(1)}_{\mathrm{in}}$ into the Q-equations using

$$\hat{u}^{(1)}_{\mathrm{in}}\big|_{\mathcal{S}+} = \hat{v}^{(1)}_{\mathrm{in}}\big|_{\mathcal{S}-} = \boldsymbol{a},$$

we obtain

$$\begin{aligned} \omega_\ell = \omega_\ell^{(0)} &+ \varepsilon\left(\frac{1}{a_\ell}(\Delta\hat{u}^{(1)}_{\mathrm{in}})(\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)\right)(\boldsymbol{\omega}, \boldsymbol{a}) \\ &+ \delta\left(\frac{1}{a_\ell}(\hat{u}^{(1)}_{\mathrm{in}} * \hat{v}^{(1)}_{\mathrm{in}})^{*p} * \hat{u}^{(1)}_{\mathrm{in}}(\boldsymbol{e}_\ell, \boldsymbol{n}_\ell)\right)(\boldsymbol{\omega}, \boldsymbol{a}),\ \ell = 1, 2, \cdots, b. \end{aligned}$$

Since the second and the third terms are smooth in $\boldsymbol{\omega}$ and $\boldsymbol{a}$, using the implicit function theorem yields $\boldsymbol{\omega}^{(2)}_{\mathrm{in}} = \boldsymbol{\omega}^{(2)}_{\mathrm{in}}(\boldsymbol{a}) = (\omega^{(2)}_{\mathrm{in},\ell}(\boldsymbol{a}))_{\ell=1}^b$, written in the form:

$$\omega^{(2)}_{\mathrm{in},\ell}(\boldsymbol{a}) = \omega_\ell^{(0)} + (\varepsilon+\delta)\varphi^{(2)}_{\mathrm{in},\ell}(\boldsymbol{a}),\ \ell = 1, 2, \cdots, b,$$

with a smooth function $\varphi^{(2)}_{\mathrm{in},\ell}$ (comparing with (3.8)).

The above constructions can be inductively performed for $2r_\star - r_0$ steps, with $r_\star$ satisfying (3.9). So we have obtained $\vec{u}^{(r)}_{\mathrm{in}} = \vec{u}^{(r)}_{\mathrm{in}}(\boldsymbol{\omega}, \boldsymbol{a})$ ($1 \le r \le 2r_\star - r_0$) for all $(\boldsymbol{\omega}, \boldsymbol{a}) \in \Omega \times [1, 2]^b$.

3.4.2. The Inductive Theorem

Recall that

$$\Omega = \boldsymbol{\omega}^{(0)} + [\delta, 2^{2p}\delta]^b \subset \mathbb{R}^b.$$

We first state the induction hypothesis, which can be verified if $0 < \varepsilon \le \delta \le \log^{-1}\frac{1}{\varepsilon} \le \delta_0 \ll 1$ and $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{W}$ (cf. (2.11)). Given its significance, we refer to it as the induction theorem.

Theorem 3.2. *For $r \ge r_\star$, we have*

(Hi) $\operatorname{supp} \vec{u}^{(r)} \subset \Lambda_{M^{\tilde{r}}}$, $\tilde{r} = r + r_\star$.

(Hii) $\|\Delta_{\mathrm{cor}}\vec{u}^{(r)}\| < \delta_r$, $\|\partial\Delta_{\mathrm{cor}}\vec{u}^{(r)}\| < \bar{\delta}_r$, *where ∂ refers to derivation in $\boldsymbol{\omega}$ or $\boldsymbol{a}$,*

$$\vec{u}^{(r)} = \vec{u}^{(r-1)} + \Delta_{\mathrm{cor}}\vec{u}^{(r)},$$

and $\|\cdot\| = \sup_{\boldsymbol{\omega},\boldsymbol{a}} \|\cdot\|_{\ell^2(\mathbb{Z}^{b+d}_{\mathrm{pm}})}$.

Remark 3.3. The size of $\delta_r, \bar{\delta}_r$ will satisfy $\log\log\frac{1}{\delta_r+\bar{\delta}_r} \sim r$.

(Hiii) $|\vec{u}^{(r)}(\boldsymbol{k}, \boldsymbol{n})| \le e^{-c(|\boldsymbol{k}|+|\boldsymbol{n}|)}$ *for some $c > 0$.*

Remark 3.4. The constant $c > 0$ will decrease slightly along the iterations but remain bounded away from 0. This will become clear in the proof, cf. sect. V, G. Step 4, [KLW24]. We also remark that $\vec{u}^{(r)}$ can be defined as a C^1 function on the entire parameter space $(\boldsymbol{\omega}, \boldsymbol{a}) \in \Omega \times [1, 2]^b$ by using an extension argument, cf., [Bou98, BW08] and sect. V, L. Step 9, [KLW24]. So one can apply the implicit function theorem to solve the Q-equations leading to

$$\omega_\ell^{(r)}(\boldsymbol{a}) = \omega_\ell^{(0)} + (\varepsilon+\delta)\varphi_\ell^{(r)}(\boldsymbol{a}) \ (1 \le \ell \le b), \tag{3.10}$$

where $\|\partial\boldsymbol{\varphi}^{(r)}\| = \sup_{1\le\ell\le b}\|\partial\varphi_\ell^{(r)}\| \lesssim 1$ and $\boldsymbol{\varphi}^{(r)} = (\varphi_\ell^{(r)})_{\ell=1}^b$. By (**Hii**), we have

$$|\boldsymbol{\varphi}^{(r)} - \boldsymbol{\varphi}^{(r-1)}| \lesssim (\varepsilon+\delta)\|\vec{u}^{(r)} - \vec{u}^{(r-1)}\| \lesssim (\varepsilon+\delta)\delta_r.$$

Denote by Γ_r the graph of $\boldsymbol{\omega}^{(r)} = \boldsymbol{\omega}^{(r)}(\boldsymbol{a})$. We have $\|\Gamma_r - \Gamma_{r-1}\| \lesssim (\varepsilon+\delta)\delta_r$. Recall that $\boldsymbol{\omega}$ is given by (3.8) and $\boldsymbol{\varphi}^{(0)} = \mathbf{0}$. Thus we have a diffeomorphism from $\boldsymbol{\omega}^{(r)} \in \Omega$ to $\boldsymbol{a} \in [1, 2]^b$.

(Hiv) *There is a collection $\mathcal{I}_r$ of intervals $I \subset \Omega \times [1, 2]^b$ of size $M^{-\tilde{r}^{10C}}$, so that*

(a) *On each $I \in \mathcal{I}_r$, both $\hat{u}^{(r)}(\boldsymbol{\omega}, \boldsymbol{a})$ and $\hat{v}^{(r)}(\boldsymbol{\omega}, \boldsymbol{a})$ are given by rational functions in $(\boldsymbol{\omega}, \boldsymbol{a})$ of degree at most $M^{\tilde{r}^3}$.*

(b) *For $(\boldsymbol{\omega}, \boldsymbol{a}) \in \bigcup_{I\in\mathcal{I}_r} I$,*

$$\|F(\vec{u}^{(r)})\| \le \kappa_r, \ \|\partial F(\vec{u}^{(r)})\| \le \bar{\kappa}_r,$$

where ∂ refers to derivation in $\boldsymbol{\omega}$ or $\boldsymbol{a}$, and $\log\log\frac{1}{\kappa_r+\bar{\kappa}_r} \sim r$.

(c) *For $(\boldsymbol{\omega}, \boldsymbol{a}) \in \bigcup_{I\in\mathcal{I}_r} I$ and $H = H(\vec{u}^{(r-1)})$, one has*

$$\|H_{M^{\tilde{r}}}^{-1}\| \le M^{\tilde{r}^C}, \tag{3.11}$$

$$|H_{M^{\tilde{r}}}^{-1}((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \le e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'| > \tilde{r}^C, \tag{3.12}$$

where $\boldsymbol{x} = (\boldsymbol{k}, \boldsymbol{n}), \boldsymbol{x}' = (\boldsymbol{k}', \boldsymbol{n}')$ and $H_{M^{\tilde{r}}}$ refers to the restriction of H to $\Lambda_{M^{\tilde{r}}}$.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

(d) *Each $I \in \mathcal{I}_r$ is contained in some $I' \in \mathcal{I}_{r-1}$ and*

$$\operatorname{meas}\left(\Pi_{\boldsymbol{a}}(\Gamma_{r-1} \cap (\bigcup_{I' \in \mathcal{I}_{r-1}} I' \setminus \bigcup_{I \in \mathcal{I}_r} I))\right) \leq M^{-\frac{\tilde{r}}{C(b)}},$$

where $C(b) > 0$ depends only on b, and $\Pi_{\boldsymbol{a}}$ denotes the projection of the set on the $\boldsymbol{a}$-variable.

(Hv) *The following precise relations hold true*

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

3.5. *Proof of the Inductive Theorem*

From the analysis in Section 3.4.1, we have constructed $\vec{u}_{\text{in}}^{(r)}$ for $1 \leq r \leq 2r_\star - r_0$. In particular, we obtain that the Inductive Theorem holds for $r = r_\star$ as we can set $\vec{u}^{(r_\star)} = \vec{u}_{\text{in}}^{(2r_\star - r_0)}$.

Assume now the Inductive Theorem holds for $r > r_\star$. We will prove it for $\ell = r + 1$.

We first check the degree bound on $\vec{u}^{(r+1)}$ (i.e., (**Hiv, a**) for $r + 1$). In fact, the Newton scheme gives

$$\Delta_{\text{cor}} \vec{u}^{(r+1)} = -H_N^{-1}(\vec{u}^{(r)}) F(\vec{u}^{(r)}), \ N = M^{\tilde{r}+1}, \ \tilde{r} = r + r_\star, \tag{3.13}$$

and by (**Hiv, a**), $\vec{u}^{(r)}(\omega, \boldsymbol{a})$ is a rational function of degree at most $M^{\tilde{r}^3}$ (in $(\omega, \boldsymbol{a})$). By the convolution structure of $T_{\vec{u}^{(r)}}$, we have

$$\deg F(\vec{u}^{(r)}) \leq (2p + 1) M^{\tilde{r}^3}, \ \deg H_N(\vec{u}^{(r)}) \leq 2p M^{\tilde{r}^3}.$$

Then by Cramer's rule, we have

$$H_N^{-1}(\vec{u}^{(r)}) = \frac{A}{\det H_N(\vec{u}^{(r)})},$$

where A is the adjacent matrix of $H_N(\vec{u}^{(r)})$. Hence,

$$\deg H_N^{-1}(\vec{u}^{(r)}) \leq (4p + 2) M^{\tilde{r}+1} M^{\tilde{r}^3}.$$

Thus, using (3.13) shows

$$\deg \Delta_{\text{cor}} \vec{u}^{(r+1)} \leq (4p + 2) M^{\tilde{r}+1} M^{\tilde{r}^3} + (2p + 1) M^{\tilde{r}^3} \leq M^{(\tilde{r}+1)^3}.$$

This proves (**Hiv, a**) for $r + 1$.

The other inductive assumptions can be verified in the usual way, cf. [LW24] and sect. V, [KLW24], for example, the verification of (**Hiv, b**) can be found in I. Step 6 of [KLW24]; the verification of the constant relations (**Hv**) can be found in Appendix E of [LW24].

It remains to establish (**Hiv, c**) and (**Hiv, d**) for $r + 1$. For this purpose, we set (recalling $\tilde{r} = r + r_\star$)

$$N = M^{\tilde{r}+1}$$

and a smaller scale,

$$N_1 = (\log N)^{\frac{4}{\rho^4}},$$

where $0 < \rho \ll 1$ is given in Theorem 2.11.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

To establish (**Hiv, c**), we first make approximations on $T_{\vec{u}^{(\ell)}}$ for different $\vec{u}^{(\ell)}$. For the set

$$W = \Big(([-N, N]^{b+d} \setminus [-N/2, N/2]^{b+d}) \times \{+, -\} \Big) \cap \mathbb{Z}^{b+d}_{\mathrm{pm},*},$$

we use N_1 size regions to do estimates, in order to lower the degree of the associated semi-algebraic set, which will be essential for the upcoming semi-algebraic projection arguments. We set

$$r_1 = 2\left[\frac{\log N_1}{\log \frac{4}{3}}\right] + 1.$$

Then

$$\|T_{\vec{u}^{(r)}} - T_{\vec{u}^{(r_1)}}\| \lesssim \delta_{r_1} \lesssim \sqrt{\varepsilon + \delta}\, e^{-(\frac{4}{3})^{r_1}} \ll e^{-N_1^2}. \tag{3.14}$$

So we can estimate $H_W^{-1}(\vec{u}^{(r)})$ using $H_Q^{-1}(\vec{u}^{(r_1)})$ for $Q \subset W$. More precisely, we want to show that the following estimates hold for every $Q \subset W$ with $Q \in \mathcal{ER}(N_1)$:

$$\|H_Q^{-1}(\vec{u}^{(r_1)})\| \le e^{N_1^{\frac{3}{4}}}, \tag{3.15}$$

$$|H_Q^{-1}(\vec{u}^{(r_1)})((\boldsymbol{x}, \xi); (\boldsymbol{x}', \xi'))| \le e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x} - \boldsymbol{x}'| > N_1^{\frac{8}{9}}. \tag{3.16}$$

This requires additional restrictions on $(\boldsymbol{\omega}, \boldsymbol{a})$. For this, we divide into the following cases:

Case 1. Assume $Q \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_0(N_1)$ with $|\boldsymbol{n}| > 10N_1$ and $Q \subset W$. Denote by $\mathcal{C}_1$ the set of all these Q. To estimate $H_Q^{-1}(\vec{u}^{(r_1)})$, we can impose as in the proof of Lemma 2.23 (with $\sigma = 0$) the following condition

$$\min_{\boldsymbol{k} \in \Pi_b Q, \xi = \pm 1, 1 \le \ell \le N_1^{b+2d}} |\boldsymbol{k} \cdot \boldsymbol{\omega} + \xi \lambda_\ell| > e^{-\frac{1}{4} N_1^{\frac{3\kappa_1}{4}}},$$

where λ_ℓ is given by Lemma 2.23. For $\boldsymbol{k} \neq \boldsymbol{0}$, we can remove $\boldsymbol{\omega}$ directly (i.e., a set in $\boldsymbol{\omega}$ of measure at most $N^C e^{-\frac{1}{4} N_1^{\frac{3\kappa_1}{4}}}$) to control $A_{\boldsymbol{k}}^{-1}(0)$ (cf. the proof of Lemma 2.23 for this notation). For the case $\boldsymbol{k} = \boldsymbol{0}$ of which we cannot remove $\boldsymbol{\omega}$, we employ Green's function estimates in (2) of Lemma 2.5. Indeed, similar to the proof of Lemma 2.23 (1), the estimate of $A_{\boldsymbol{0}}^{-1}(0) = \mathcal{L}_{Q(\boldsymbol{0})}^{-1}(\boldsymbol{\theta}) \oplus \mathcal{L}_{Q(\boldsymbol{0})}^{-1}(\boldsymbol{\theta})$ follows directly from (2) of Lemma 2.5 since we assume $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{W}$ (cf. (2.11)). So by taking into account all $\boldsymbol{k} \in \Pi_b Q$, all $Q \in \mathcal{C}_1$ and by Fubini's theorem, we can find $I_1 \subset \Omega \times [1, 2]^b$ with (recalling $\rho \le c_2 \kappa_1 \le \frac{\kappa_1}{10^4 (b+d)^2}$)

$$\mathrm{meas}(I_1) \le N^C e^{-\frac{1}{4} (\log N)^{\frac{3\kappa_1}{\rho^4}}} \le e^{-(\log N)^{10}} \ll (\varepsilon + \delta)^{b+1} M^{-\tilde{r}},$$

so that, for all $(\boldsymbol{\omega}, \boldsymbol{a}) \notin I_1$ and all $Q \in \mathcal{C}_1$, $H_Q^{-1}(\vec{u}^{(r_1)})$ satisfies (3.15) and (3.16).

Case 2. Assume $Q \in (\boldsymbol{k}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(N_1)$ with $|\boldsymbol{n}| \le 10N_1$ and $Q \subset W$. Denote by $\mathcal{C}_2$ the set of all these Q. In this case, it must be that $|\boldsymbol{k}| \ge N/2$. We will use the projection lemma and LDT to remove $\boldsymbol{\omega}$. For any $Q \in \mathcal{C}_2$, we can write $Q = (\boldsymbol{k}, \boldsymbol{n}) + Q_0$ with $Q_0 \in (\boldsymbol{0}, \boldsymbol{n}) + \mathcal{ER}_{\boldsymbol{0}}(N_1)$ for $|\boldsymbol{n}| \le 10N_1$, and $N/2 \le |\boldsymbol{k}| \le N$. This motivates us to consider

$$H(\sigma) = H(\vec{u}^{(r_1)}; \sigma) = D(\sigma) + \varepsilon(\Delta \oplus \Delta) + \delta T_{\vec{u}^{(r_1)}},$$

which has been investigated in Section 2. Recall that $G_Q(\sigma)$ denotes the Green's function of

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

$H(\sigma)$ restricted to Q. Then the Toeplitz property in the $\boldsymbol{k}$-direction of $H(\sigma)$ implies

$$H_Q^{-1}(\vec{u}^{(r_1)}) = G_Q(0) = G_{Q_0}(\boldsymbol{k}\cdot\boldsymbol{\omega}).$$

Note that Ω_{N_1} given by Theorem 2.11 is a semi-algebraic set, *independent of* $\boldsymbol{a}$, and of degree at most $N_1^{10(b+d)}$. Without loss of generality, we may assume $\Pi_\omega(I\cap\Gamma_{r_1})\subset\Omega_{N_1}$ for each $I\in\mathcal{I}_{r_1}$. For otherwise, we can replace $\Pi_\omega(I\cap\Gamma_{r_1})$ with $(\Pi_\omega(I\cap\Gamma_{r_1}))\cap(\Omega_{N_1'}\setminus\Omega_{N_1})$, where $N_1'=(\log M^{\tilde{r}})^{\frac{4}{\rho^4}}\ll N_1$ since by Remark 2.12,

$$\operatorname{meas}(\Omega_{N_1'}\setminus\Omega_{N_1})\le e^{-\frac{1}{2}N_1'^{\rho^4}}\le e^{-\frac{1}{2}(\log N)^4}\ll(\varepsilon+\delta)^{b+1}N^{-10}.$$

Then we apply the LDT at scale N_1 on each $I\in\mathcal{I}_{r_1}$ (cf. Remark 2.13) to obtain that for all $Q\in(\boldsymbol{0},\boldsymbol{n})+\mathcal{ER}_{\boldsymbol{0}}(N_1)$ with $|\boldsymbol{n}|\le 10N_1$, and for σ outside a set of measure at most $e^{-N_1^{\rho}}$, the following estimates hold true:

$$\|G_Q(\sigma)\|\le e^{N_1^{\frac{3}{4}}}, \tag{3.17}$$

$$|G_Q(\sigma)((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))|\le e^{-c|\boldsymbol{x}-\boldsymbol{x}'|}\text{ for }|\boldsymbol{x}-\boldsymbol{x}'|>N_1^{\frac{8}{9}}. \tag{3.18}$$

For the usage of projection lemma in the present setting (i.e., $\operatorname{meas}(\Omega)\sim\delta^b$), we need to make more precise descriptions of the admitted σ first. We say σ is Q-bad if either (3.17) or (3.18) fails.

Lemma 3.5. *Fix $Q\in(\boldsymbol{0},\boldsymbol{n})+\mathcal{ER}_{\boldsymbol{0}}(N_1)$ with $|\boldsymbol{n}|\le 10N_1$. Denote by Σ_Q the set of Q-bad $\sigma\in\mathbb{R}$. Then*

$$\Sigma_Q\subset\bigcup_{1\le\ell\le N_1^C}J_\ell,$$

where each J_ℓ is an interval of length $\sim(\varepsilon+\delta)$.

Proof. Note that $N=M^{\tilde{r}+1}\ge M^{2r_\star+1}>(\varepsilon+\delta)^{-c_1}$ since $r>r_\star$. So we have

$$N_1=(\log N)^{\frac{4}{\rho^4}}>c_1\log^{\frac{4}{\rho^4}}\frac{1}{\varepsilon+\delta}>\log^{\frac{2}{\rho^4}}\frac{1}{\varepsilon+\delta},$$

which implies

$$(\varepsilon+\delta)^{-1}<e^{N_1^{\rho^4/2}}. \tag{3.19}$$

So we can define for each $\boldsymbol{x}=(\boldsymbol{k},\boldsymbol{n})\in Q$ and $\xi=\pm1$ the set

$$J_{\boldsymbol{x},\xi}=\{\sigma\in\mathbb{R}:\ |\sigma+\boldsymbol{k}\cdot\boldsymbol{\omega}+\xi\mu_{\boldsymbol{n}}|\le C(\varepsilon+\delta)\},$$

where $C>1$ depends only on $\|\Delta\|,\|T_{\vec{u}^{(r_1)}}\|$. From (3.19) and the Neumann series argument (cf. Lemma B.1), we have

$$\Sigma_Q\subset\bigcup_{\boldsymbol{x}\in Q,\ \xi=\pm1}J_{\boldsymbol{x},\xi}.$$

This proves Lemma 3.5. ◻

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Now fix $I_0 \in \mathcal{I}_{r_1}$. Solving the Q-equations at $r = r_1$ leads to the graph Γ_{r_1}. Then $\tilde{I} = \Pi_{\boldsymbol{\omega}}(\Gamma_{r_1} \cap I_0)$ is an interval of size at most $\varepsilon+\delta$. For $\boldsymbol{\omega} \in \tilde{I}$, let $\tilde{\Sigma}$ be the set of $\sigma \in \mathbb{R}$ so that either (3.17) or (3.18) fails for some $Q \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{E}\mathcal{R}_{\mathbf{0}}(N_1)$ with $|\boldsymbol{n}| \leq 10N_1$. Then by Lemma 3.5, $\tilde{\Sigma}$ can be covered by N_1^C intervals of size $\sim (\varepsilon+\delta)$. Pick J to be one of such intervals and consider $\mathcal{K} := \tilde{I} \times (\tilde{\Sigma} \cap J) \subset \tilde{I} \times J \subset \mathbb{R}^b \times \mathbb{R}$. We will show that $\mathcal{K}$ is a semi-algebraic set of degree $\deg \mathcal{K} \leq N_1^C M^{C\tilde{r}_1^3}$ and measure $\operatorname{meas}(\mathcal{K}) \leq C(\delta+\varepsilon)^b e^{-N_1^\rho}$. This measure bound follows directly from the LDT at scale N_1 and Fubini's theorem. For the semi-algebraic description of $\mathcal{K}$, we need the following lemma.

Lemma 3.6 (Tarski-Seidenberg Principle, cf., for example, [Bou07]). *Denote by $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{d_1+d_2}$ the product variable. If $X \subset \mathbb{R}^{d_1+d_2}$ is semi-algebraic of degree B, then its projections $\Pi_{\boldsymbol{x}} X \subset \mathbb{R}^{d_1}$ and $\Pi_{\boldsymbol{y}} X \subset \mathbb{R}^{d_2}$ are semi-algebraic of degree at most B^C, where $C = C(d_1, d_2) > 0$.*

We define $X \subset I_0 \times J$ to be the set of all $(\boldsymbol{\omega}, \boldsymbol{a}, \sigma)$ so that either (3.17) or (3.18) fails for some $Q \in (\mathbf{0}, \boldsymbol{n}) + \mathcal{E}\mathcal{R}_{\mathbf{0}}(N_1)$ with $|\boldsymbol{n}| \leq 10N_1$. Similar to the proof of Lemma 2.22, we can regard X as a semi-algebraic set of degree N_1^C. From (**Hiv**) and solving the Q-equations for $r = r_1$, we know that $\Gamma_{r_1} \cap I_0$ is given by an algebraic equation in $\boldsymbol{\omega}, \boldsymbol{a}$ (cf. (3.10)) of degree $M^{C\tilde{r}_1^3}$. Note also that

$$\mathcal{K} = \Pi_{\boldsymbol{\omega}, \sigma}\big(X \cap ((\Gamma_{r_1} \cap I_0) \times \mathbb{R})\big),$$

which together with Lemma 3.6 implies $\deg \mathcal{K} \leq N_1^C M^{C\tilde{r}_1^3}$.

Our next aim is to estimate

$$\operatorname{meas}\left(\bigcup_{N/2 \leq |\boldsymbol{k}| \leq N} \{\boldsymbol{\omega} : \ (\boldsymbol{\omega}, \boldsymbol{k}\cdot\boldsymbol{\omega}) \in \mathcal{K}\}\right). \tag{3.20}$$

At this stage, we need an important projection lemma.

Lemma 3.7 (cf., for example, [Bou07]). *Let $Y \subset [0,1]^{d=d_1+d_2}$ be a semi-algebraic set of degree $\deg Y = B$ and $\operatorname{meas}(Y) \leq \eta$, where $\log B \ll \log \frac{1}{\eta}$. Denote by $(\boldsymbol{x}, \boldsymbol{y}) \in [0,1]^{d_1} \times [0,1]^{d_2}$ the product variable. Suppose $\eta^{\frac{1}{d}} \leq \epsilon$. Then there is a decomposition of Y as*

$$Y = Y_1 \cup Y_2$$

with the following properties: The projection of Y_1 on $[0,1]^{d_1}$ has small measure

$$\operatorname{meas}(\Pi_{\boldsymbol{x}} Y_1) \leq B^{C(d)}\epsilon,$$

and Y_2 has the transversality property

$$\operatorname{meas}(\mathbb{H} \cap Y_2) \leq B^{C(d)}\epsilon^{-1}\eta^{\frac{1}{d}},$$

where $\mathbb{H}$ is any d_2-dimensional hyperplane in $\mathbb{R}^d$, s.t., $\max\limits_{1 \leq j \leq d_1} |\Pi_{\mathbb{H}}(\boldsymbol{e}_j)| < \epsilon$ (we denote by $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{d_1}$ the $\boldsymbol{x}$-coordinate standard basis).

Remark 3.8. This lemma permits scaling. It is based on the Yomdin-Gromov triangulation theorem [Gro87]; for a complete proof, see [BN19].

Taking $\eta = C(\varepsilon+\delta)^b e^{-N_1^\rho}$, $\epsilon = 10/N$, we obtain

$$C(\varepsilon+\delta)^b e^{-\frac{1}{b+1}N_1^\rho} \leq C(\varepsilon+\delta)^b e^{-\frac{1}{b+1}(\log N)^4} \ll \epsilon.$$

Using Lemma 3.7, we have a decomposition

$$\mathcal{K} = \mathcal{K}_1 \cup \mathcal{K}_2,$$

where meas $(\Pi_\omega \mathcal{K}_1) \le C(\varepsilon+\delta)^{b+1} N_1^C M^{C\tilde{r}_1^3} M^{-\tilde{r}} \le (\varepsilon+\delta)^{b+1} M^{-\frac{\tilde{r}}{2}}$ and the factor $(\varepsilon+\delta)^{b+1}$ comes from scaling when applying Lemma 3.7. Note however that $|\boldsymbol{k}| \ge N/2$ cannot ensure $\min_{1\le \ell \le b} |k_\ell| \ge N/2$. So the hyperplane $\{(\boldsymbol{\omega}, \boldsymbol{k}\cdot\boldsymbol{\omega})\}$ does not satisfy the steepness condition. To address this issue, we will need to take into account all possible directions and apply Lemma 3.7 for b times, as done by Bourgain (cf. (3.26) of [Bou07] and also (142) of [LW24]). This then leads to an upper bound $C(\varepsilon+\delta)^{b+1} M^{-\frac{\tilde{r}+1}{C(b)}}$ on (3.20), where $C(b) > 0$ only depends on b.

Taking into account all J, we have shown the existence of $\tilde{I}_1 \subset \tilde{I}$ with $\mathrm{meas}_b(\tilde{I}_1) \le C(\varepsilon+\delta)^{b+1} N_1^C M^{-\frac{\tilde{r}+1}{C(b)}} \le C(\varepsilon+\delta)^{b+1} M^{-\frac{\tilde{r}+1}{C(b)}}$ so that for $\boldsymbol{\omega} \in \tilde{I} \setminus \tilde{I}_1$, the estimates (3.15) and (3.16) hold for all $Q \in \mathcal{C}_2$. Let I_0 range over $\mathcal{I}_{r_1}$. The total measure removed from $\Pi_\omega \Gamma_{r_1}$ is at most $C(\varepsilon+\delta)^{b+1} M^{\tilde{r}_1^C} M^{-\frac{\tilde{r}+1}{C(b)}} \le (\varepsilon+\delta)^{b+1} M^{-\frac{\tilde{r}+1}{C(b)}}$. Since (3.15)–(3.16) allow $O(e^{-N_1^2})$ perturbation of $(\boldsymbol{\omega}, \boldsymbol{a})$ (again using Lemma B.1) and $\|\Gamma_r - \Gamma_{r_1}\| \lesssim \delta_{r_1} \ll e^{-N_1^2}$, we obtain a subset $\Gamma_r' \subset \Gamma_r$ with $\mathrm{meas}(\Pi_\omega \Gamma_r') \le (\varepsilon+\delta)^{b+1} M^{-\frac{\tilde{r}+1}{C(b)}}$ so that (3.15) and (3.16) hold on

$$\bigcup_{I \in \mathcal{I}_{r_1}} (I \cap (\Gamma_r \setminus \Gamma_r')),$$

and hence on

$$\bigcup_{I \in \mathcal{I}_{r}} (I \cap (\Gamma_r \setminus \Gamma_r')).$$

Next, by the induction hypothesis at step r, we have that (3.11) and (3.12) hold for r. Also, by (**Hii**) at step r, if $|\boldsymbol{x}-\boldsymbol{x}'| \le (\tilde{r}+1)^{5C} \log M$, then (we hide the dependence on $\xi, \xi', \boldsymbol{\omega}, \boldsymbol{a}$)

$$\begin{aligned}
&|T_{\vec{u}^{(r)}}(\boldsymbol{x},\boldsymbol{x}') - T_{\vec{u}^{(r-1)}}(\boldsymbol{x},\boldsymbol{x}')| \\
&\le C\|\Delta_{\mathrm{cor}} \vec{u}^{(r)}\| \\
&\le C\delta_r \le CM^{-(\tilde{r}+1)^{10C}} \\
&\le M^{-10(\tilde{r}+1)^C} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|}.
\end{aligned}$$

If $|\boldsymbol{x}-\boldsymbol{x}'| > (\tilde{r}+1)^{5C} \log M$, by (**Hiii**) at step r and Lemma 3.1, we have

$$\begin{aligned}
&|T_{\vec{u}^{(r)}}(\boldsymbol{x},\boldsymbol{x}') - T_{\vec{u}^{(r-1)}}(\boldsymbol{x},\boldsymbol{x}')| \\
&\le C(1+|\boldsymbol{x}-\boldsymbol{x}'|)^C e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \\
&\le M^{-10(\tilde{r}+1)^C} e^{-(c-(\tilde{r}+1)^{-3C})|\boldsymbol{x}-\boldsymbol{x}'|}.
\end{aligned}$$

The above estimates imply that for all $\boldsymbol{x}, \boldsymbol{x}'$,

$$|T_{\vec{u}^{(r)}}(\boldsymbol{x},\boldsymbol{x}') - T_{\vec{u}^{(r-1)}}(\boldsymbol{x},\boldsymbol{x}')| \le M^{-10(\tilde{r}+1)^C} e^{-(c-(\tilde{r}+1)^{-3C})|\boldsymbol{x}-\boldsymbol{x}'|}.$$

Hence, combining (3.11)–(3.12) (at step r) and Lemma B.1, it follows that for any $(\boldsymbol{\omega}, \boldsymbol{a}) \in \cup_{I\in\mathcal{I}_r} I$, we can replace $\vec{u}^{(r-1)}$ in (3.11)–(3.12) with $\vec{u}^{(r)}$. That is to say,

$$\|H^{-1}_{M^{\tilde{r}}}(\vec{u}^{(r)})\| \le 2M^{\tilde{r}^C}, \tag{3.21}$$

and for $|\boldsymbol{x}-\boldsymbol{x}'| > \tilde{r}^C$,

$$|H^{-1}_{M^{\tilde{r}}}(\vec{u}^{(r)})(\boldsymbol{x},\boldsymbol{x}')| \leq e^{-(c-(\tilde{r}+1)^{-3C})|\boldsymbol{x}-\boldsymbol{x}'|}. \tag{3.22}$$

Assume $|(\boldsymbol{\omega},\boldsymbol{a})-(\boldsymbol{\omega}',\boldsymbol{a}')| \leq M^{-(\tilde{r}+1)^{10C}}$. Similar to the above proof, we have for all $\boldsymbol{x},\boldsymbol{x}'$,

$$\begin{aligned}&|H(\vec{u}^{(r)};(\boldsymbol{\omega},\boldsymbol{a}))(\boldsymbol{x},\boldsymbol{x}')-H(\vec{u}^{(r)};(\boldsymbol{\omega}',\boldsymbol{a}'))(\boldsymbol{x},\boldsymbol{x}')|\\ \leq &M^{-10(\tilde{r}+1)^C}e^{-(c-(\tilde{r}+1)^{-3C})|\boldsymbol{x}-\boldsymbol{x}'|}.\end{aligned}$$

Using Lemma B.1 again, we have that the estimates (3.21)–(3.22) and (3.15)–(3.16) (replacing $\vec{u}^{(r_1)}$ with $\vec{u}^{(r)}$) allow $O(M^{-(\tilde{r}+1)^{10C}})$ perturbation of $(\boldsymbol{\omega},\boldsymbol{a})$. Combining conclusions in the above **Case 1**–**Case 2** and the resolvent identity of [HSSY20] (cf. Lemma 3.6) gives a collection $\mathcal{I}_{r+1}$ of intervals of size $M^{-(\tilde{r}+1)^{10C}}$ so that for $I \in \mathcal{I}_{r+1}$,

$$\begin{aligned}\|H_N^{-1}(\vec{u}^{(r)})\| &\leq M^{(\tilde{r}+1)^C},\\ |H_N^{-1}(\vec{u}^{(r)})((\boldsymbol{x},\xi);(\boldsymbol{x}',\xi'))| &\leq e^{-c_{r+1}|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'|>(\tilde{r}+1)^C,\end{aligned}$$

which concludes (**Hiv, c**) at scale $r+1$. We remark that in the above off-diagonal exponential decay estimates, $c_{r+1} \geq c_r - (\log M)^C(\tilde{r}+1)^{-C}$, which implies $\inf_r c_r > 0$. This explains why we need the sublinear-distance off-diagonal decay in LDT.

Finally, we have

$$\begin{aligned}&\operatorname{meas}\left(\Pi_{\boldsymbol{\omega}}(\Gamma_r \cap (\bigcup_{I'\in\mathcal{I}_r} I' \setminus \bigcup_{I\in\mathcal{I}_{r+1}} I))\right)\\ \leq\ &(\varepsilon+\delta)^{b+1}M^{-\frac{\tilde{r}+1}{C(b)}} + e^{-\frac{1}{2}N_1^{\rho^4}} < (\varepsilon+\delta)^{b+1}M^{-\frac{\tilde{r}+1}{C(b)}}.\end{aligned}$$

Recalling that $\boldsymbol{\omega} \longrightarrow \boldsymbol{a}$ is a C^1 diffeomorphism and $\det(\frac{\partial\boldsymbol{\omega}}{\partial\boldsymbol{a}}) \sim \delta^b$, we obtain since $\varepsilon \leq \delta$ that

$$\operatorname{meas}\left(\Pi_{\boldsymbol{a}}(\Gamma_r \cap (\bigcup_{I'\in\mathcal{I}_r} I' \setminus \bigcup_{I\in\mathcal{I}_{r+1}} I))\right) \leq M^{-\frac{\tilde{r}+1}{C(b)}},$$

which yields (**Hiv, d**) at step $r+1$. □

Proof of Theorem 1.1. The proof of Theorem 1.1 is a direct corollary of the Inductive Theorem. We refer to [BW08, LW24] and Sections V and VI of [KLW24] for further details. □

Appendix A. Diophantine estimates

Below we provide Diophantine estimates when $V(\boldsymbol{\theta})$ is an (arbitrary) trigonometric polynomial. These estimates hold on a large set in $(\boldsymbol{\alpha},\boldsymbol{\theta})$. The main idea is to utilize an appropriate generalized Wronskian approach. To our knowledge, these estimates did not appear to be known in the literature before. Given the relation between analytic functions (hence trigonometric polynomials) and generalized Wronskians, this approach also seems natural in hindsight.

To our knowledge, using transversality properties to make Diophantine approximations on manifolds was first developed by Pyartli [Pya69]. This method was later introduced into the study of KAM theory [XYQ97, Eli02, Bam03]. It turns out that this idea also plays a key role in the Craig-Wayne-Bourgain type argument used in the present paper. In addition, we can handle Diophantine estimates on general quasi-periodic polynomials on $[0,1]^d$. The main difference between [XYQ97, Eli02, Bam03] and the present work is that we have no *prior* transversality estimate, which requires much additional effort.

Recall that

$$V(\boldsymbol{\theta}) = \sum_{\boldsymbol{\ell} \in \Gamma_K} v_{\boldsymbol{\ell}} \cos 2\pi(\boldsymbol{\ell} \cdot \boldsymbol{\theta}) \text{ with } \boldsymbol{\theta} \in [0,1]^d,$$

where $\Gamma_K \subset [-K,K]^d \setminus \{\mathbf{0}\}$ ($K \in \mathbb{N}$) is *maximal* so that

$$\begin{aligned} &\text{if } \boldsymbol{\ell} \in \Gamma_K, \text{then}, \ell_s \neq 0, \forall\, 1 \leq s \leq d; \\ &\text{and if } \ \boldsymbol{\ell}, \boldsymbol{\ell}' \in \Gamma_K, \ \text{ then } \boldsymbol{\ell} + \boldsymbol{\ell}' \neq \mathbf{0}, \end{aligned} \tag{A.1}$$

and

$$\boldsymbol{v} = (v_{\boldsymbol{\ell}})_{\boldsymbol{\ell} \in \Gamma_K} \in \mathbb{R}^{\#\Gamma_K} \setminus \{\mathbf{0}\}.$$

We also assume $V(\boldsymbol{\theta})$ is *nondegenerate*.

To handle $\mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})$, we need the following Łojasiewicz-type lemma.

Lemma A.1 (cf., for example, Lemma 4.2 in [JLS20]). *There exist some constants $C_V > 0, 0 < c_V < 1$ depending only on V so that, for any $\eta > 0$, one has*

$$\text{meas}\left(\left\{\boldsymbol{\theta} \in [0,1]^d : |V(\boldsymbol{\theta})| \leq \eta\right\}\right) \leq C_V \eta^{c_V}.$$

Next, we fix

$$\boldsymbol{n}_1', \cdots, \boldsymbol{n}_D' \in \mathbb{Z}^d$$

to be D ($D \geq 2$) distinct lattice points. We are mainly concerned with Diophantine estimates for

$$\boldsymbol{\omega}' = (\omega_1', \cdots, \omega_D'), \ \omega_s' = V(\boldsymbol{\theta} + \boldsymbol{n}_s'\boldsymbol{\alpha}) \ (1 \leq s \leq D).$$

Throughout this Appendix, all constants $C, c > 0$ are assumed to (depend only on d, K, D, V) be independent of $\boldsymbol{n}_s'$ ($1 \leq s \leq D$).

Our main result on the Diophantine estimates is

Theorem A.2. *Let $0 < \eta < 1$. We have for $R = R_D = D \cdot (\#\Gamma_K) = D\frac{(2K+1)^d - 1}{2}$,*

$$\text{meas}(\text{DC}_{\text{nls}}(\eta)) \geq 1 - C \cdot (\max_{1 \leq s \leq D} |\boldsymbol{n}_s'|)^{5dR^2} \eta^{\frac{1}{10dR^2}},$$

where $C = C(d, K, D, V) > 0$ and

$$\begin{aligned} &\text{DC}_{\text{nls}}(\eta) \\ &= \left\{(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0,1]^{2d} : |\boldsymbol{k}' \cdot \boldsymbol{\omega}'| \geq \frac{\eta}{|\boldsymbol{k}'|^{4R^2}} \text{ for } \forall\, \boldsymbol{k}' \in \mathbb{Z}^D \setminus \{\mathbf{0}\}\right\}. \end{aligned}$$

We also have for any $\boldsymbol{k}' \in \mathbb{Z}^D \setminus \{\mathbf{0}\}$,

$$\text{meas}(\{(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0,1]^{2d} : |\boldsymbol{k}' \cdot \boldsymbol{\omega}'| \leq \eta\}) \leq C \cdot (\max_{1 \leq s \leq D} |\boldsymbol{n}_s'|)^{5dR^2} \eta^{\frac{1}{10dR^2}}.$$

Remark A.3. In contrast, if $\boldsymbol{\omega}'$ were a free parameter varying in $[0,1]^D$, we would have

$$\text{meas}\left(\left\{\boldsymbol{\omega}' \in [0,1]^D : |\boldsymbol{k}' \cdot \boldsymbol{\omega}'| \geq \frac{\eta}{|\boldsymbol{k}'|^{4R^2}} \text{ for } \forall\, \boldsymbol{k}' \in \mathbb{Z}^D \setminus \{\mathbf{0}\}\right\}\right) = 1 - O(\eta).$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

So the absence of independence of the coordinates of $\boldsymbol{\omega}'$ may lead to the corresponding measure bound (from $1 - O(\eta)$ to) $1 - O(\eta^{\frac{1}{10dR^2}})$.

Recall that we have fixed in Theorem 1.1,

$$\boldsymbol{n}_1, \cdots, \boldsymbol{n}_b \in \mathbb{Z}^d,\ \boldsymbol{\omega}^{(0)} = (V(\boldsymbol{n}_1\boldsymbol{\alpha} + \boldsymbol{\theta}), \cdots, V(\boldsymbol{n}_b\boldsymbol{\alpha} + \boldsymbol{\theta})),$$

and for arbitrary $\boldsymbol{n}$, we have $\mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})$.

As a consequence of Lemma A.1 and Theorem A.2, we obtain

Corollary A.4. *There exist some* $0 < c_1 = c_1(b, d, K, V) < \frac{1}{100b}, C_1 = C_1(b, d, K, V) > 1$ *such that, if* $0 < \varepsilon + \delta \leq \delta_0(b, d, K, V, \max_{1 \leq \ell \leq b} |\boldsymbol{n}_\ell|) \ll 1$*, then there is some* $\mathcal{M} \subset [0, 1]^{2d}$ *satisfying*

$$\text{meas}([0, 1]^{2d} \setminus \mathcal{M}) \leq \log^{-1} \frac{1}{\varepsilon + \delta}$$

so that the following properties hold true for $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \mathcal{M}$ *and* $L_{\varepsilon,\delta} = 100(\varepsilon + \delta)^{-c_1}$.

(i). *For any* $\boldsymbol{n} \neq \boldsymbol{n}'$ *satisfying* $|\boldsymbol{n}|, |\boldsymbol{n}'| \leq L_{\varepsilon,\delta}$,

$$|\mu_{\boldsymbol{n}} - \mu_{\boldsymbol{n}'}| \geq (\varepsilon + \delta)^{\frac{1}{8b}}.$$

(ii). *For any* $\boldsymbol{k} \in \mathbb{Z}^b$ *satisfying* $0 < |\boldsymbol{k}| \leq 2L_{\varepsilon,\delta}$,

$$|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)}| \geq (\varepsilon + \delta)^{\frac{1}{8b}}.$$

(iii). *For any* $\log \frac{1}{\varepsilon+\delta} \leq L \leq L_{\varepsilon,\delta}$ *and all* $(\boldsymbol{k}, \boldsymbol{n}) \in \Lambda_L \setminus \mathcal{S}$,

$$\min_{\xi = \pm 1} |\xi \boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}}| \geq L^{-C_1}.$$

(iv). *For any* $(\boldsymbol{k}, \boldsymbol{n}, \boldsymbol{n}') \in \mathbb{Z}^b \times \mathbb{Z}^d \times \mathbb{Z}^d$ *satisfying* $|\boldsymbol{k}| \leq 2L_{\varepsilon,\delta}, |(\boldsymbol{n}, \boldsymbol{n}')| \leq L_{\varepsilon,\delta}$ *and* $\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}} - \mu_{\boldsymbol{n}'} \not\equiv 0$,

$$|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}} - \mu_{\boldsymbol{n}'}| \geq (\varepsilon + \delta)^{\frac{1}{8b}}.$$

Proof. It suffices to apply Lemma A.1 (i.e., $D = 1$) and Theorem A.2 with $\eta = (\varepsilon + \delta)^{\frac{1}{8b}}, L^{-C_1}$ and $D = 2, b + 1, b + 2$. Without loss of generality, we only deal with (iii) and (iv) since (i) and (ii) can be handled similarly.

(iii). Denote by $\mathcal{M}_{\text{iii}}$ the set of $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0, 1]^{2d}$ so that the conclusion of (iii) holds true.

For the case of $\boldsymbol{k} = \boldsymbol{0}$, applying Lemma A.1 yields for all $\boldsymbol{\alpha} \in [0, 1]^d$,

$$\begin{aligned}
&\text{meas}\left(\bigcup_{\log \frac{1}{\varepsilon+\delta} \leq L \leq L_{\varepsilon,\delta}} \bigcup_{|\boldsymbol{n}| \leq L} \{\boldsymbol{\theta} \in [0, 1]^d : |\mu_{\boldsymbol{n}} = V(\boldsymbol{n}\boldsymbol{\alpha} + \boldsymbol{\theta})| \leq L^{-C_1}\} \right) \\
&\leq \sum_{\log \frac{1}{\varepsilon+\delta} \leq L \leq L_{\varepsilon,\delta}} \sum_{|\boldsymbol{n}| \leq L} C_V L^{-C_1 c_V} \\
&\leq \sum_{\log \frac{1}{\varepsilon+\delta} \leq L \leq L_{\varepsilon,\delta}} C(V, d) L^{-C_1 c_V + d} \\
&\leq \log^{-3} \frac{1}{\varepsilon + \delta} \text{ (choosing } C_1 \text{ so that } C_1 c_V - d > 10),
\end{aligned}$$

where in the last inequality we assume $0 < \varepsilon + \delta \leq \delta_0(V, d) \ll 1$.

For the case of $\boldsymbol{k} \neq \boldsymbol{0}$ and $\boldsymbol{n} \notin \{\boldsymbol{n}_1, \cdots, \boldsymbol{n}_b\}$, we apply Theorem A.2 with $D = b+1, \eta = L^{-C_1}$ and $\boldsymbol{k}' = (\pm\boldsymbol{k}, 1)$ to get for

$$\mathcal{M}_{\mathrm{iii}}^{(1)}(L) := \bigcup_{|(\boldsymbol{k},\boldsymbol{n})|\leq L, \boldsymbol{n}\notin\{\boldsymbol{n}_1,\cdots,\boldsymbol{n}_b\}} \left\{(\boldsymbol{\alpha},\boldsymbol{\theta}) \in [0,1]^{2d} : \min_{\xi=\pm 1} |\xi\boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}}| \leq L^{-C_1}\right\},$$

the following

$$\begin{aligned}
&\mathrm{meas}\left(\bigcup_{\log\frac{1}{\varepsilon+\delta}\leq L\leq L_{\varepsilon,\delta}} \mathcal{M}_{\mathrm{iii}}^{(1)}(L)\right) \\
&\leq \sum_{\log\frac{1}{\varepsilon+\delta}\leq L\leq L_{\varepsilon,\delta}} C(b,d,K,V) L^{2d} L^{5dR_{b+1}^2} L^{-\frac{C_1}{10dR_{b+1}^2}} \text{ (assume } L > \max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|) \\
&\leq \sum_{\log\frac{1}{\varepsilon+\delta}\leq L\leq L_{\varepsilon,\delta}} C(b,d,K,V) L^{-10} \text{ (choosing } C_1 \geq C(b,d,K) \gg 1) \\
&\leq \log^{-3}\frac{1}{\varepsilon+\delta},
\end{aligned}$$

where in the last inequality we assume $0 < \varepsilon+\delta \leq \delta_0(b,d,K,V, \max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|) \ll 1$. The case of $\boldsymbol{k} \neq \boldsymbol{0}$ and $\boldsymbol{n} \in \{\boldsymbol{n}_1, \cdots, \boldsymbol{n}_b\}$ can be handled similarly and we omit the details. Since there are only 3 cases to deal with, $C_1 = C_1(b,d,K,V) > 0$ can be well-defined.

By combining all the above estimates, we obtain the desired measure bound on $\mathcal{M}_{\mathrm{iii}}$, namely,

$$\mathrm{meas}([0,1]^{2d} \setminus \mathcal{M}_{\mathrm{iii}}) \leq \log^{-2}\frac{1}{\varepsilon+\delta}.$$

(iv). If $x, y \in \mathbb{R}$, we denote $x \wedge y = \min\{x,y\}$ and $x \vee y = \max\{x,y\}$. Denote by $\mathcal{M}_{\mathrm{iv}}$ the set of $(\boldsymbol{\alpha},\boldsymbol{\theta}) \in [0,1]^{2d}$ so that the conclusion of (iv) holds true. First we consider the case of $(\boldsymbol{k},\boldsymbol{n},\boldsymbol{n}') \in \mathcal{T}_1$ with $\mathcal{T}_1$ the set of all $(\boldsymbol{k},\boldsymbol{n},\boldsymbol{n}')$ satisfying: $0 < |\boldsymbol{k}| \leq 2L_{\varepsilon,\delta}, |(\boldsymbol{n},\boldsymbol{n}')| \leq L_{\varepsilon,\delta}, \boldsymbol{n} \neq \boldsymbol{n}'$, both $\boldsymbol{n}$ and $\boldsymbol{n}' \notin \{\boldsymbol{n}_1, \cdots, \boldsymbol{n}_b\}$. We let

$$\mathcal{M}_{\mathrm{iv}}^{(1)} := \bigcup_{(\boldsymbol{k},\boldsymbol{n},\boldsymbol{n}')\in\mathcal{T}_1} \left\{(\boldsymbol{\alpha},\boldsymbol{\theta}) \in [0,1]^{2d} : \min_{\xi=\pm 1} |\boldsymbol{k}\cdot\boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}} - \mu_{\boldsymbol{n}'}| \leq (\varepsilon+\delta)^{\frac{1}{8b}}\right\}.$$

So we can apply Theorem A.2 with $D = b+2, \eta = (\varepsilon+\delta)^{\frac{1}{8b}}$ and $\boldsymbol{k}' = (\boldsymbol{k}, 1, -1)$ to get

$$\begin{aligned}
\mathrm{meas}\left(\mathcal{M}_{\mathrm{iv}}^{(1)}\right) &\leq \sum_{(\boldsymbol{k},\boldsymbol{n},\boldsymbol{n}')\in\mathcal{T}_1} C(b,d,K,V)((\max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|) \vee |\boldsymbol{n}| \vee |\boldsymbol{n}'|)^{5dR_{b+2}^2} (\varepsilon+\delta)^{\frac{1}{80bdR_{b+2}^2}} \\
&\leq C(b,d,K,V) L_{\varepsilon,\delta}^{b+2d+5dR_{b+1}^2} (\varepsilon+\delta)^{\frac{1}{80bdR_{b+2}^2}} \text{ (assume } L_{\varepsilon,\delta} > \max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|) \\
&\leq C(b,d,K,V)(\varepsilon+\delta)^{-c_1(b+2d+5dR_{b+1}^2)+\frac{1}{80bdR_{b+2}^2}} \\
&\leq (\varepsilon+\delta)^{\frac{1}{100bdR_{b+2}^2}} \\
&\quad \text{(choosing } 0 < c_1 \leq \frac{1}{(b+2d+5dR_{b+1}^2)720bdR_{b+2}^2}),
\end{aligned}$$

where in the last inequality we assume $0 < \varepsilon+\delta \leq \delta_0(b,d,K,V, \max_{1\leq\ell\leq b}|\boldsymbol{n}_\ell|) \ll 1$.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Next, we consider the case of $(\boldsymbol{k}, \boldsymbol{n}, \boldsymbol{n}')$ with $\boldsymbol{k} \cdot \boldsymbol{\omega}^{(0)} + \mu_{\boldsymbol{n}} \equiv 0$. In this case, applying Lemma A.1 yields for all $\boldsymbol{\alpha} \in [0,1]^d$,

$$\begin{aligned}
&\operatorname{meas}\left(\bigcup_{|\boldsymbol{n}'| \le L_{\varepsilon,\delta}} \left\{\boldsymbol{\theta} \in [0,1]^d : \ |\mu_{\boldsymbol{n}'} = V(\boldsymbol{n}'\boldsymbol{\alpha} + \boldsymbol{\theta})| \le (\varepsilon+\delta)^{\frac{1}{8b}}\right\}\right) \\
&\le \sum_{|\boldsymbol{n}'| \le L_{\varepsilon,\delta}} C_V (\varepsilon+\delta)^{\frac{c_V}{8b}} \\
&\le C(V,d)(\varepsilon+\delta)^{-c_1 d + \frac{c_V}{8b}} \\
&\le (\varepsilon+\delta)^{\frac{c_V}{10b}} \ (\text{choosing } 0 < c_1 \le \frac{c_V}{72db}),
\end{aligned}$$

where in the last inequality we assume $0 < \varepsilon + \delta \le \delta_0(b,d,V) \ll 1$. Other cases are easier to handle, and we omit the details. By taking into account all the above estimates, we have

$$\operatorname{meas}([0,1]^{2d} \setminus \mathcal{M}_{\mathrm{iv}}) \le (\varepsilon+\delta)^{\frac{c_V}{1000bdR_{b+2}^2}}.$$

Note that since there are only finitely many (say, at most 10) cases to deal with, $c_1 = c_1(b,d,K,V) > 0$ can be well-defined. □

A.1. Some useful lemmas

In this section, we will introduce some useful lemmas.

Lemma A.5 [KM98]. *Let $I \subset \mathbb{R}$ be an interval of finite length (i.e., $0 < |I| < \infty$) and $k \ge 1$. If $f \in C^k(I;\mathbb{R})$ satisfies*

$$\inf_{x \in I} |\frac{d^k}{dx^k} f(x)| \ge A > 0, \tag{A.2}$$

then for all $\varepsilon > 0$,

$$\operatorname{meas}(\{x \in I : \ |f(x)| \le \varepsilon\}) \le \zeta_k (\frac{\varepsilon}{A})^{\frac{1}{k}}, \tag{A.3}$$

where $\zeta_k = k(k+1)((k+1)!)^{\frac{1}{k}}$.

Proof. The following proof is taken from Kleinbock-Margulis [KM98].

From (A.2) and Rolle's theorem, the equation $\frac{d}{dx} f(x) = 0$ has at most k zeros on I. This then implies that the set $\{x \in I : \ |f(x)| \le \varepsilon\}$ consists of at most $k+1$ intervals. Denote by I_1 one of those subintervals having maximal length. So we get

$$\operatorname{meas}(\{x \in I : \ |f(x)| \le \varepsilon\}) \le (k+1)|I_1|.$$

To prove (A.3), it remains to estimate $|I_1| > 0$. We divide I_1 into k equal parts by points $x_1, \cdots, x_{k+1}$. Let $P(x)$ be the Lagrange polynomial of f on points $x_1, \cdots, x_{k+1}$, namely,

$$P(x) = \sum_{i=1}^{k+1} f(x_i) \frac{\prod_{1 \le j \ne i \le k+1} (x - x_i)}{\prod_{1 \le j \ne i \le k+1} (x_j - x_i)}.$$

By applying Rolle's theorem k times and since $f(x_i) = P(x_i)$ ($1 \le i \le k+1$), we can find $x_0 \in I_1$ so that

$$\frac{d^k}{dx^k} f(x_0) = \frac{d^k}{dx^k} P(x_0) = k! \sum_{i=1}^{k+1} \frac{f(x_i)}{\prod_{1 \le j \ne i \le k+1}(x_j - x_i)},$$

which together with (A.2) yields

$$A \le k! \sum_{i=1}^{k+1} \frac{\varepsilon}{|I_1|^k k^{-k}}.$$

This implies

$$|I_1| \le k(k+1)!)^{\frac{1}{k}} (\frac{\varepsilon}{A})^{\frac{1}{k}}.$$

We have finished the proof. □

In the following, we turn to the analysis of multi-variable functions. We first introduce the notations.

- For any $\boldsymbol{x} = (x_1, \cdots, x_d)$ and any $d \ge 1$, let

$$|\boldsymbol{x}|_2 = \sqrt{\sum_{i=1}^{d} |x_i|^2},\ |\boldsymbol{x}|_1 = \sum_{i=1}^{d} |x_i|.$$

- Let $\boldsymbol{a} = (a_1, \cdots, a_d), \boldsymbol{b} = (b_1, \cdots, b_d)$. Define the d-dimensional interval to be

$$I = I_{\boldsymbol{a},\boldsymbol{b}} = \prod_{i=1}^{d} [a_i, b_i] \subset \mathbb{R}^d. \tag{A.4}$$

$$|\boldsymbol{a} \vee \boldsymbol{b}| = \sum_{i=1}^{d} \max(|a_i|, |b_i|).$$

- For $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_d) \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$, define

$$d_{\boldsymbol{\beta}} := \sum_{i=1}^{d} \beta_i \partial_i,\ \partial_i := \frac{\partial}{\partial x_i}.$$

  We also write for $\ell \ge 1$,

$$d_{\boldsymbol{\beta}}^{\ell} = d_{\boldsymbol{\beta}} \cdots d_{\boldsymbol{\beta}}\ (\ell - \text{times}).$$

  Denote

$$\nabla = (\partial_1, \cdots, \partial_d).$$

- For a multi-index $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_d) \in \mathbb{N}^d$, denote

$$|\boldsymbol{\gamma}| = \sum_{i=1}^{d} \gamma_i,\ \partial^{\gamma} = \partial_1^{\gamma_1} \cdots \partial_d^{\gamma_d}.$$

We have

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Lemma A.6. *Fix $k \in \mathbb{N}$. Let $I = I_{\boldsymbol{a},\boldsymbol{b}} \subset \mathbb{R}^d$ be defined by (A.4). Assume that the function $f \in C^{k+1}(I;\mathbb{R})$ satisfies for some $A > 0$,*

$$\inf_{\boldsymbol{x}\in I} \sup_{1\le \ell\le k} \left| d_{\boldsymbol{\beta}}^{\ell} f(\boldsymbol{x})\right| \geq A. \tag{A.5}$$

Let

$$\|f\|_{k+1} = \sup_{\boldsymbol{x}\in I} \sup_{1\le|\boldsymbol{\gamma}|\le k+1} |\partial^{\boldsymbol{\gamma}} f(\boldsymbol{x})| < \infty.$$

Then for $0 < \varepsilon < 1$,

$$\begin{aligned} &\operatorname{meas}(\{\boldsymbol{x} \in I : \ |f(\boldsymbol{x})| \le \varepsilon\}) \\ &\le C(\boldsymbol{\beta}, k, d)(\|f\|_{k+1}|\boldsymbol{b} - \boldsymbol{a}|_2 + 1)^d |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} A^{-d} (A \wedge 1)^{-1} \varepsilon^{\frac{1}{k}}, \end{aligned} \tag{A.6}$$

where $C = C(\boldsymbol{\beta}, k, d) > 0$ depends only on $\boldsymbol{\beta}, k, d$ (but not on f).

Remark A.7. If f is a polynomial on $\mathbb{R}^d$, [CW01] proved some strong distributional inequalities, which may directly imply (A.6). However, in the present case, we have a weak transversality condition (A.5). So we would like to give an elementary proof applying for more general functions (cf. [LSZ25] for an application), which also produces an upper bound in the estimate (A.6) with the explicit dependence on $\|f\|_{k+1}$.

Proof. The proof can be divided into two steps:

Step 1. Assume that for some $1 \le k_1 \le k$ and interval $I_1 = \prod_{i=1}^d [h_i, m_i]$,

$$\inf_{\boldsymbol{x}\in I_1} \left| d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x})\right| \geq A. \tag{A.7}$$

We will combine Lemma A.5 and Fubini's theorem to deal with the measure estimate. Let $\boldsymbol{e}_1 = \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|_2}$ and choose a normalized orthogonal basis $\boldsymbol{e}_2, \cdots, \boldsymbol{e}_d \in \mathbb{R}^d$ of

$$\boldsymbol{\beta}^{\perp} = \{\boldsymbol{x} \in \mathbb{R}^d : \ \boldsymbol{x} \cdot \boldsymbol{\beta} = 0\}.$$

We can then define the coordinate transform:

$$\boldsymbol{x} = \boldsymbol{\varphi}(s, \boldsymbol{y}), (s, \boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^{d-1},$$

via

$$\boldsymbol{\varphi}(s, \boldsymbol{y}) = s\boldsymbol{e}_1 + \sum_{i=2}^{d} y_i \boldsymbol{e}_i.$$

So the Jacobian satisfies

$$|\det \frac{\partial \boldsymbol{x}}{\partial(s, \boldsymbol{y})}| = 1.$$

Define

$$S_\varepsilon = \{\boldsymbol{x} \in I_1 : \ |f(\boldsymbol{x})| \le \varepsilon\},$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

and denote by $\chi_{(\cdot)}$ the indicator function. We have

$$\begin{aligned}\mathrm{meas}(S_\varepsilon) &= \int_{I_1} \chi_{S_\varepsilon}(\boldsymbol{x})d\boldsymbol{x} \\ &= \int_{\boldsymbol{\varphi}^{-1}(I_1)} \chi_{S_\varepsilon}(\boldsymbol{\varphi}(s,\boldsymbol{y}))dsd\boldsymbol{y}.\end{aligned} \tag{A.8}$$

We proceed to control $\boldsymbol{\varphi}^{-1}(I_1)$. From $\boldsymbol{x} = s\boldsymbol{e}_1 + \sum_{i=2}^d y_i\boldsymbol{e}_i \in I_1$, we obtain

$$s = \boldsymbol{x}\cdot\boldsymbol{e}_1 \in [-|\boldsymbol{h}\vee\boldsymbol{m}|, |\boldsymbol{h}\vee\boldsymbol{m}|].$$

Similarly, we also have $y_i \in [-|\boldsymbol{h}\vee\boldsymbol{m}|, |\boldsymbol{h}\vee\boldsymbol{m}|]$ for $2\le i\le d$, which yields $\boldsymbol{\varphi}^{-1}(I_1)\subset[-|\boldsymbol{h}\vee\boldsymbol{m}|, |\boldsymbol{h}\vee\boldsymbol{m}|]^d$. We define

$$\Pi_1\boldsymbol{\varphi}^{-1}(I_1) = \{\boldsymbol{y}\in\mathbb{R}^{d-1} :\ \exists\, s\in\mathbb{R}\ s.t.,\ (s,\boldsymbol{y})\in\boldsymbol{\varphi}^{-1}(I_1)\}$$

and, define for $\boldsymbol{y}\in\Pi_1\boldsymbol{\varphi}^{-1}(I_1)$, the interval $I_{\boldsymbol{y}} = \{s\in\mathbb{R} :\ (s,\boldsymbol{y})\in\boldsymbol{\varphi}^{-1}(I_1)\}$ of finite length. Applying Fubini's theorem together with (A.8) implies that

$$\begin{aligned}\mathrm{meas}(S_\varepsilon) &= \int_{\Pi_1\boldsymbol{\varphi}^{-1}(I_1)} d\boldsymbol{y}\int_{I_{\boldsymbol{y}}} \chi_{S_\varepsilon}(\boldsymbol{\varphi}(s,\boldsymbol{y}))ds \\ &\le \mathrm{meas}_{d-1}(\Pi_1\boldsymbol{\varphi}^{-1}(I_1)) \sup_{\boldsymbol{y}\in\Pi_1\boldsymbol{\varphi}^{-1}(I_1)}\int_{I_{\boldsymbol{y}}} \chi_{S_\varepsilon}(\boldsymbol{\varphi}(s,\boldsymbol{y}))ds \\ &\le 2^{d-1}|\boldsymbol{h}\vee\boldsymbol{m}|^{d-1} \sup_{\boldsymbol{y}\in\Pi_1\boldsymbol{\varphi}^{-1}(I_1)}\int_{I_{\boldsymbol{y}}} \chi_{S_\varepsilon}(\boldsymbol{\varphi}(s,\boldsymbol{y}))ds \\ &\quad (\text{since } \Pi_1\boldsymbol{\varphi}^{-1}(I_1)\subset[-|\boldsymbol{h}\vee\boldsymbol{m}|, |\boldsymbol{h}\vee\boldsymbol{m}|]^{d-1}).\end{aligned}$$

For fixed $\boldsymbol{y}\in\Pi_1\boldsymbol{\varphi}^{-1}(I_1)$, we get

$$\begin{aligned}&\int_{I_{\boldsymbol{y}}} \chi_{S_\varepsilon}(\boldsymbol{\varphi}(s,\boldsymbol{y}))ds \\ &= \mathrm{meas}(\{s\in I_{\boldsymbol{y}} :\ |f(s\boldsymbol{e}_1 + \sum_{i=2}^d y_i\boldsymbol{e}_i)|\le\varepsilon\}).\end{aligned}$$

Denote $g(s) = f(s\boldsymbol{e}_1 + \sum_{i=2}^d y_i\boldsymbol{e}_i)$. From (A.7), we have for all $s\in I_{\boldsymbol{y}}$,

$$|\frac{d^{k_1}}{ds^{k_1}}g(s)| = |\frac{1}{|\boldsymbol{\beta}|_2^{k_1}}d_{\boldsymbol{\beta}}^{k_1}f(s\boldsymbol{e}_1 + \sum_{i=2}^d y_i\boldsymbol{e}_i)| \ge \frac{A}{|\boldsymbol{\beta}|_2^{k_1}}.$$

So applying Lemma A.5 yields

$$\mathrm{meas}(S_\varepsilon) \le \zeta_{k_1}2^{d-1}|\boldsymbol{h}\vee\boldsymbol{m}|^{d-1}|\boldsymbol{\beta}|_2(\frac{\varepsilon}{A})^{\frac{1}{k_1}}. \tag{A.9}$$

Step 2. We deal with the general case in this step. So we first divide each one-dimensional interval $[a_i,b_i]$ into N (will be specified below) equal subintervals I_{i,j_i} ($1\le j_i\le N$). This then induces a decomposition of I, namely,

$$I = \bigcup_{1\le j_1,\cdots,j_d\le N}\prod_{i=1}^d I_{i,j_i} = \bigcup_{\boldsymbol{J}\in([1,N]\cap\mathbb{N})^d} I_{\boldsymbol{J}}.$$

We will apply the argument proved in **Step 1** on each $I_{\boldsymbol{J}}$. For this purpose, we fix $I_{\boldsymbol{J}}$. For any $\boldsymbol{x}, \boldsymbol{y} \in I_{\boldsymbol{J}}$, we have

$$\begin{aligned} |\boldsymbol{x} - \boldsymbol{y}|_2 &\leq \sqrt{\sum_{i=1}^{d} \frac{|b_i - a_i|^2}{N^2}} \\ &\leq \frac{|\boldsymbol{b} - \boldsymbol{a}|_2}{N}. \end{aligned}$$

Fixing any $\boldsymbol{x}_0 \in I_{\boldsymbol{J}}$, we get from the assumption (A.5) that there exists $1 \leq k_1 \leq k$ with

$$|d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x}_0)| \geq A > 0.$$

As a result, for any $\boldsymbol{x} \in I_{\boldsymbol{J}}$, we obtain

$$\begin{aligned} |d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x})| &\geq |d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x}_0)| - |d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x}) - d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x}_0)| \\ &\geq A - \sup_{\boldsymbol{x} \in I_{\boldsymbol{J}}} |\nabla d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x})|_2 \cdot |\boldsymbol{x} - \boldsymbol{x}_0|_2 \\ &\geq A - \frac{\|f\|_{k_1+1} L_{\boldsymbol{\beta},d} |\boldsymbol{b} - \boldsymbol{a}|_2}{N} \\ &\geq A - \frac{\|f\|_{k+1} L_{\boldsymbol{\beta},d} |\boldsymbol{b} - \boldsymbol{a}|_2}{N}, \end{aligned} \tag{A.10}$$

where

$$L_{\boldsymbol{\beta},d} = \sqrt{d} \sum_{1 \leq l_1, \cdots, l_d \leq d} |\beta_{l_1} \cdots \beta_{l_d}|.$$

In fact, $L_{\boldsymbol{\beta},d}$ is derived from

$$\begin{aligned} |\nabla d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x})|_2 &= \sqrt{\sum_{j=1}^{d} \Big| \sum_{1 \leq l_1, \cdots, l_d \leq d} \beta_{l_1} \cdots \beta_{l_d} \partial_{j l_1 \cdots l_d}^{k_1+1} f(\boldsymbol{x})\Big|^2} \\ &\leq \|f\|_{k_1+1} \sqrt{\sum_{j=1}^{d} \Big(\sum_{1 \leq l_1, \cdots, l_d \leq d} |\beta_{l_1} \cdots \beta_{l_d}|\Big)^2} \\ &:= \|f\|_{k_1+1} L_{\boldsymbol{\beta},d}. \end{aligned}$$

From (A.10), we can set

$$N = [2\|f\|_{k+1} L_{\boldsymbol{\beta},d} |\boldsymbol{b} - \boldsymbol{a}|_2 A^{-1}] + 1$$

so that

$$\inf_{\boldsymbol{x} \in I_{\boldsymbol{J}}} |d_{\boldsymbol{\beta}}^{k_1} f(\boldsymbol{x})| \geq A/2 > 0,$$

where $[x]$ denotes the integer part of $x \in \mathbb{R}$. Thus applying (A.9) yields (since $I_J \subset I, 0 < \varepsilon < 1$)

$$\begin{aligned} \operatorname{meas}(\{\boldsymbol{x} \in I_{\boldsymbol{J}} : \ |f(\boldsymbol{x})| \leq \varepsilon\}) &\leq \zeta_{k_1} 2^{d-1} |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} 2^{\frac{1}{k_1}} |\boldsymbol{\beta}|_2 \Big(\frac{\varepsilon}{A}\Big)^{\frac{1}{k_1}} \\ &\leq C(k,d) |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} |\boldsymbol{\beta}|_2 (A \wedge 1)^{-1} \varepsilon^{\frac{1}{k}}. \end{aligned}$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Collecting all $I_{\boldsymbol{J}}$ leads to,

$$\begin{aligned}
&\operatorname{meas}(\{\boldsymbol{x} \in I: \ |f(\boldsymbol{x})| \leq \varepsilon\}) \\
&\leq C(k,d) N^d |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} |\boldsymbol{\beta}|_2 (A \wedge 1)^{-1} \varepsilon^{\frac{1}{k}} \\
&\leq C(\boldsymbol{\beta},k,d)(\|f\|_{k+1}|\boldsymbol{b}-\boldsymbol{a}|_2 A^{-1}+1)^d |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} (A \wedge 1)^{-1} \varepsilon^{\frac{1}{k}} \\
&\leq C(\boldsymbol{\beta},k,d)(\|f\|_{k+1}|\boldsymbol{b}-\boldsymbol{a}|_2+1)^d |\boldsymbol{a} \vee \boldsymbol{b}|^{d-1} A^{-d} (A \wedge 1)^{-1} \varepsilon^{\frac{1}{k}}.
\end{aligned}$$

We have proven (A.6). □

Finally, we introduce a key lemma (cf. Proposition in Appendix B, [BGG85] for a more precise form) on determinant estimates.

Lemma A.8. *Let $\boldsymbol{v}^{(1)}, \cdots, \boldsymbol{v}^{(r)} \in \mathbb{R}^r$ be r linearly independent vectors with $|\boldsymbol{v}^{(\ell)}|_1 \leq M$ for $1 \leq \ell \leq r$. Then for any $\boldsymbol{w} \in \mathbb{R}^r$, we have*

$$\max_{1 \leq \ell \leq r} |\boldsymbol{w} \cdot \boldsymbol{v}^{(\ell)}| \geq r^{-3/2} M^{1-r} |\boldsymbol{w}|_2 \cdot |\det\left[\boldsymbol{v}^{(\ell)}\right]_{1 \leq \ell \leq r}|.$$

Proof. For completeness, we give a proof based on Cramer's rule and the Hadamard's inequality. We assume $\boldsymbol{v}^{(\ell)}$ is the row vector in $\mathbb{R}^r$ ($1 \leq \ell \leq r$). Denote by $X = [\boldsymbol{v}^{(1)}, \cdots, \boldsymbol{v}^{(r)}]$ the $r \times r$ matrix and consider the equation

$$X\boldsymbol{w} = \boldsymbol{\xi}, \ \boldsymbol{\xi} = (\xi_1, \cdots, \xi_r) \text{ with } \xi_\ell = \boldsymbol{w} \cdot \boldsymbol{v}^{(\ell)}.$$

By the Cramer's rule, we obtain

$$\boldsymbol{w} = X^{-1}\boldsymbol{\xi} = \frac{X^*}{\det X}\boldsymbol{\xi},$$

where $X^* = [\tilde{v}_{ij}]_{1 \leq i,j \leq r}$ denotes the adjugate matrix of X. As a result, we get

$$\begin{aligned}
|\boldsymbol{w}|_2 \cdot |\det X| &= |X^*\boldsymbol{\xi}|_2 \\
&= \sqrt{\sum_{i=1}^r |\sum_{j=1}^r \tilde{v}_{ij}\xi_j|^2} \\
&\leq \max_{1 \leq j \leq r} |\xi_j| \sqrt{\sum_{i=1}^r (\sum_{j=1}^r |\tilde{v}_{ij}|)^2} \\
&\leq r^{3/2} \max_{1 \leq j \leq r} |\xi_j| \max_{1 \leq i,j \leq r} |\tilde{v}_{ij}|.
\end{aligned}$$

It remains to bound $\max_{1 \leq i,j \leq r} |\tilde{v}_{ij}|$, which will be completed by using the Hadamard's inequality. In fact, we have

$$\begin{aligned}
|\tilde{v}_{ij}| &\leq \prod_{1 \leq \ell \neq i \leq r} \sqrt{\sum_{1 \leq m \neq j \leq r} |v_m^{(\ell)}|^2} \\
&\leq \prod_{1 \leq \ell \neq i \leq r} \sum_{1 \leq m \neq j \leq r} |v_m^{(\ell)}| \\
&\leq M^{r-1} \ (\text{since } \max_{1 \leq \ell \leq r} |\boldsymbol{v}^{(\ell)}|_1 \leq M).
\end{aligned}$$

This finishes the proof. □

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

A.2. *Proof of Theorem A.2*

In this section, we prove Theorem A.2.

Proof of Theorem A.2. Recall that

$$\begin{aligned}\boldsymbol{k}'\cdot\boldsymbol{\omega}' &= \sum_{s=1}^{D} k_s' V(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha})\\ &= \sum_{s=1}^{D}\sum_{\boldsymbol{\ell}\in\Gamma_K} k_s' v_{\boldsymbol{\ell}} \cos 2\pi\boldsymbol{\ell}\cdot(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha})\\ &= (\boldsymbol{k}'\otimes\boldsymbol{v})\cdot\boldsymbol{V},\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{k}'\otimes\boldsymbol{v} &= (k_s' v_{\boldsymbol{\ell}})_{1\le s\le D,\boldsymbol{\ell}\in\Gamma_K},\\ \boldsymbol{V} &= (V_{(s,\boldsymbol{\ell})})_{1\le s\le D,\boldsymbol{\ell}\in\Gamma_K} := (\cos 2\pi\boldsymbol{\ell}\cdot(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha}))_{1\le s\le D,\boldsymbol{\ell}\in\Gamma_K}.\end{aligned}$$

Denote $\boldsymbol{\xi}=(\boldsymbol{\beta},\boldsymbol{q})$ for $\boldsymbol{\beta},\boldsymbol{q}\in[0,1]^d$. Write also

$$d_{\boldsymbol{\xi}} = \sum_{i=1}^{d}\Big(\beta_i\frac{\partial}{\partial\alpha_i}+q_i\frac{\partial}{\partial\theta_i}\Big)$$

and accordingly $d_{\boldsymbol{\xi}}^{\ell}$ for $\ell\ge 1$.

A key observation is that for all $j\ge 1$ and $\boldsymbol{\ell}\in\Gamma_K$,

$$\begin{aligned}&d_{\boldsymbol{\xi}}^{2j}\cos 2\pi\boldsymbol{\ell}\cdot(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha})\\ &=(-(2\pi(\boldsymbol{\ell}\boldsymbol{n}_s')\cdot\boldsymbol{\beta}+2\pi\boldsymbol{q}\cdot\boldsymbol{\ell})^2)^j\cos 2\pi\boldsymbol{\ell}\cdot(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha}).\end{aligned}$$

This motivates us to consider the Wronskian

$$W=[d_{\boldsymbol{\xi}}^{2j}V_{(s,\boldsymbol{\ell})}]_{s,\boldsymbol{\ell}}^{1\le j\le R}\ \text{with}\ R=D\cdot(\#\Gamma_K),$$

which is a $R\times R$ real matrix. Direct computations show

$$\begin{aligned}&|\det W|\\ &=\prod_{s,\boldsymbol{\ell}}|\cos 2\pi\boldsymbol{\ell}\cdot(\boldsymbol{\theta}+\boldsymbol{n}_s'\boldsymbol{\alpha})|\times\prod_{s,\boldsymbol{\ell}}(2\pi)^2((\boldsymbol{\ell}\boldsymbol{n}_s')\cdot\boldsymbol{\beta}+\boldsymbol{q}\cdot\boldsymbol{\ell})^2\times|\det W_1|,\end{aligned}\tag{A.11}$$

where W_1 is a Vandermonde matrix with

$$|\det W_1|=\sqrt{\prod_{(s,\boldsymbol{\ell})\ne(s',\boldsymbol{\ell}')}(2\pi)^2|((\boldsymbol{\ell}\boldsymbol{n}_s')\cdot\boldsymbol{\beta}+\boldsymbol{q}\cdot\boldsymbol{\ell})^2-((\boldsymbol{\ell}'\boldsymbol{n}_{s'}')\cdot\boldsymbol{\beta}+\boldsymbol{q}\cdot\boldsymbol{\ell}')^2|}.\tag{A.12}$$

Then the proof can be divided into three steps:

Step 1. Prior conditions on $(\boldsymbol{\alpha},\boldsymbol{\theta})$

We aim to get a lower bound for (A.11), which requires the prior restrictions on $\boldsymbol{\alpha},\boldsymbol{\theta}$. Let $\|x\|_{\mathbb{T}/2}=\operatorname{dist}(x,\mathbb{Z}/2)$. From

$$4\|x-\frac{1}{4}\|_{\mathbb{T}/2}\le|\cos 2\pi x|\le 2\pi\|x-\frac{1}{4}\|_{\mathbb{T}/2},$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

we can define for $\delta > 0$ the resonant set (of $\boldsymbol{\alpha}, \boldsymbol{\theta}$)

$$S_\delta = \bigcup_{1\le s\le D, \boldsymbol{\ell}\in\Gamma_K} \left\{(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0,1]^{2d} : \|\boldsymbol{\ell}\cdot(\boldsymbol{\theta} + \boldsymbol{n}'_s\boldsymbol{\alpha}) - \frac{1}{4}\|_{\mathbb{T}/2} \le \delta\right\}. \tag{A.13}$$

Denote $L = \max_{1\le s\le D} K|\boldsymbol{n}'_s|$. We have

Lemma A.9. *Let* $0 < \delta < 1$. *Then* S_δ *can be covered by* $CL^{2d-1}\delta^{-2d+1}$ *intervals of side length* $\le \delta$, *where* $C = C(d, R) > 0$.

Proof. It suffices to cover for each $\boldsymbol{n}'_s, \boldsymbol{\ell}$ the set

$$S_{\boldsymbol{n}'_s,\boldsymbol{\ell}} = \{(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in [0,1]^{2d} : \|\boldsymbol{\ell}\cdot(\boldsymbol{\theta} + \boldsymbol{n}'_s\boldsymbol{\alpha}) - \frac{1}{4}\|_{\mathbb{T}/2} \le \delta\}$$

with $C(d)L^{2d-1}\delta^{-2d+1}$ intervals of side length $\le \delta$ since the total number of all $\boldsymbol{n}'_s, \boldsymbol{\ell}$ is at most R. Since $|\boldsymbol{\ell}| > 0$, we may assume $\ell_1 \ne 0$ without loss of generality. For any $\boldsymbol{x} = (x_1, \cdots, x_d)$, write $\boldsymbol{x} = (x_1, \boldsymbol{x}_1^c)$. So we can divide $[0,1]^d \times [0,1]^{d-1} \ni (\boldsymbol{\alpha}, \boldsymbol{\theta}_1^c)$ into $C_1 = [100L\delta^{-1}]^{2d-1}$ cubes $\boldsymbol{C}_i$ ($1 \le i \le C_1$) with disjoint exteriors of the same side length $\le \frac{\delta}{L}$. Denote by $(\boldsymbol{\alpha}_i, \boldsymbol{\theta}^c(i)) \in [0,1]^d \times [0,1]^{d-1}$ ($1 \le i \le C_1$) the centers of those cubes. We further define

$$\begin{aligned} &I_{\boldsymbol{n}'_s,\boldsymbol{\ell}} \\ &= \bigcup_{1\le i\le C_1} \{(\boldsymbol{\alpha}, \boldsymbol{\theta}_1^c, \theta_1) \in \boldsymbol{C}_i \times [0,1] : \|\ell_1\theta_1 + \boldsymbol{\ell}_1^c\cdot\boldsymbol{\theta}^c(i) + (\boldsymbol{n}'_s\boldsymbol{\ell})\cdot\boldsymbol{\alpha}_i - \frac{1}{4}\|_{\mathbb{T}/2} \le 2\delta\} \\ &:= \bigcup_{1\le i\le C_1} J_i. \end{aligned}$$

We first claim that $S_{\boldsymbol{n}'_s,\boldsymbol{\ell}} \subset I_{\boldsymbol{n}'_s,\boldsymbol{\ell}}$. Let $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in S_{\boldsymbol{n}'_s,\boldsymbol{\ell}}$. Then by the definition (of $S_{\boldsymbol{n}'_s,\boldsymbol{\ell}}$), we have

$$\|\ell_1\theta_1 + \boldsymbol{\ell}_1^c\cdot\boldsymbol{\theta}_1^c + (\boldsymbol{\ell}\boldsymbol{n}'_s)\cdot\boldsymbol{\alpha} - \frac{1}{4}\|_{\mathbb{T}/2} \le \delta.$$

Since $\boldsymbol{C}_i$ ($1 \le i \le C_1$) covers $[0,1]^{2d-1}$, there is $1 \le i_0 \le C_1$ so that $(\boldsymbol{\alpha}, \boldsymbol{\theta}_1^c) \in \boldsymbol{C}_{i_0}$, namely, $|\boldsymbol{\alpha} - \boldsymbol{\alpha}_{i_0}| + |\boldsymbol{\theta}_1^c - \boldsymbol{\theta}^c(i_0)| \le \frac{\delta}{2L}$. This implies that

$$\begin{aligned} &\|\ell_1\theta_1 + \boldsymbol{\ell}_1^c\cdot\boldsymbol{\theta}^c(i_0) + (\boldsymbol{n}'_s\boldsymbol{\ell})\cdot\boldsymbol{\alpha}_{i_0} - \frac{1}{4}\|_{\mathbb{T}/2} \\ &\le \|\boldsymbol{\ell}\cdot(\boldsymbol{\theta} + \boldsymbol{n}'_s\boldsymbol{\alpha}) - \frac{1}{4}\|_{\mathbb{T}/2} + |\boldsymbol{\ell}_1^c\cdot(\boldsymbol{\theta}^c(i_0) - \boldsymbol{\theta}_1^c)| \\ &\quad + |(\boldsymbol{\ell}\boldsymbol{n}'_s)\cdot(\boldsymbol{\alpha} - \boldsymbol{\alpha}_{i_0})| \\ &\le 2\delta. \end{aligned}$$

The claim is proved. So it remains to cover each J_i with intervals of side length $\le \delta$. This can be accomplished by the fact that the set $\{\theta_1 \in [0,1] : \|\ell_1\theta_1 + x\|_{\mathbb{T}/2} \le 2\delta\}$ ($x \in \mathbb{R}$) consists of $r \sim |\ell_1|$ one-dimensional intervals of length $2\delta/|\ell_1|$.

We have finished the proof. □

Given $0 < \delta < 1/10$, we divide $[0,1]$ into $N = [\delta^{-2}] + 1$ intervals of equal length $\frac{1}{N} \sim \delta^2$. This then induces a partition of $[0,1]^{2d}$ into cubes of equal side length $1/N$. We denote by $\Lambda_{1/N}(\boldsymbol{x}_i)$ ($1 \le i \le N^{2d}$) all these cubes. We have

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Lemma A.10. *Let* $0 < \delta < 1/10$ *and let* S_δ *be defined by (A.13). Then we have for* $N = [\delta^{-2}] + 1$,

$$\#\{1 \le i \le N^{2d} : \Lambda_{1/N}(\boldsymbol{x}_i) \cap S_\delta \ne \emptyset\} \le C(d, R) L^{2d-1} \delta^{-4d+1}. \tag{A.14}$$

In particular, we have

$$\operatorname{meas}\left(\bigcup_{1 \le i \le N^{2d} :\ \Lambda_{1/N}(\boldsymbol{x}_i) \cap S_\delta \ne \emptyset} \Lambda_{1/N}(\boldsymbol{x}_i) \right) \le C(d, R) L^{2d-1} \delta. \tag{A.15}$$

Proof. Let $\boldsymbol{C}_i$ be given as in the proof of Lemma A.9. Then $S_\delta \subset \cup_{1 \le i \le C(d,R) L^{2d-1} \delta^{-2d+1}} \boldsymbol{C}_i$. So it suffices to bound

$$L_1 = \#\{1 \le i \le N^{2d} : (\cup_{1 \le j \le C L^{2d-1} \delta^{-2d+1}} \boldsymbol{C}_j) \cap \Lambda_{1/N}(\boldsymbol{x}_i) \ne \emptyset\}.$$

Note that if $\Lambda_{1/N}(\boldsymbol{x}_i) \cap \boldsymbol{C}_j \ne \emptyset$, then

$$\Lambda_{1/N}(\boldsymbol{x}_i) \subset \boldsymbol{C}'_j,$$

where $\boldsymbol{C}'_j$ is also an interval containing $\boldsymbol{C}_j$ and satisfying

$$\operatorname{dist}(\partial \boldsymbol{C}_j, \partial \boldsymbol{C}'_j) \le 10/N,$$

with ∂ the boundary. Regarding the disjointness (of the exterior) of $\Lambda_{1/N}(\boldsymbol{x}_i)$ for different i, we must have

$$L_1 \le C L^{2d-1} \delta^{-2d+1} (\delta + 10/N)^{2d} \delta^{-4d} \le C L^{2d-1} \delta^{-4d+1}.$$

This implies the bound (A.14). Then the measure bound follows directly from (A.15). □

From this lemma, we get immediately that

Lemma A.11. *Let* $0 < \delta < 1/10$ *and define* $N = [\delta^{-2}] + 1$. *Then there is a collection of cubes* $\Lambda_{1/N}(\boldsymbol{x}_i) \subset [0, 1]^{d+d}$ *(* $1 \le i \le N_1$*) with disjoint exteriors so that*

$$\operatorname{meas}\left([0, 1]^{2d} \setminus \cup_{1 \le i \le N_1} \Lambda_{1/N}(\boldsymbol{x}_i) \right) \le C L^{2d-1} \delta,$$

and if $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \cup_{1 \le i \le N_1} \Lambda_{1/N}(\boldsymbol{x}_i)$, *then*

$$\prod_{1 \le s \le D, \boldsymbol{\ell} \in \Gamma_K} |\cos 2\pi \boldsymbol{\ell} \cdot (\boldsymbol{\theta} + \boldsymbol{n}'_s \alpha)| \ge c \delta^R, \ c = c(d, R) > 0.$$

Step 2. Conditions on $(\boldsymbol{\beta}, \boldsymbol{q})$

Next we impose conditions on $(\boldsymbol{\beta}, \boldsymbol{q})$.

Lemma A.12. *Let* $\epsilon \in (0, 1)$ *and* $\boldsymbol{B} \subset [0, 1]^d \times [0, 1]^d$ *be the set of* $\boldsymbol{\xi} = (\boldsymbol{\beta}, \boldsymbol{q})$ *satisfying*

$$\begin{aligned}
&\min_{s, \boldsymbol{\ell}} ((\boldsymbol{\ell} \boldsymbol{n}'_s) \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot \boldsymbol{\ell})^2 \ge \epsilon > 0, \\
&\min_{(s, \boldsymbol{\ell}) \ne (s', \boldsymbol{\ell}')} |(\boldsymbol{\ell} \boldsymbol{n}'_s - \boldsymbol{\ell}' \boldsymbol{n}'_{s'}) \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}')| \ge \epsilon > 0, \\
&\min_{(s, \boldsymbol{\ell}) \ne (s', \boldsymbol{\ell}')} |(\boldsymbol{\ell} \boldsymbol{n}'_s + \boldsymbol{\ell}' \boldsymbol{n}'_{s'}) \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot (\boldsymbol{\ell} + \boldsymbol{\ell}')| \ge \epsilon > 0.
\end{aligned}$$

Then

$$\operatorname{meas}\left([0, 1]^{2d} \setminus \boldsymbol{B} \right) \le 3 R^2 \sqrt{\epsilon}.$$

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Moreover, for $\boldsymbol{\xi} \in \boldsymbol{B}$, *we have*

$$\begin{aligned}
&\prod_{s,\boldsymbol{\ell}} ((\boldsymbol{\ell} \boldsymbol{n}_s') \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot \boldsymbol{\ell})^2 \times |\det W_1| \\
&= \prod_{s,\boldsymbol{\ell}} ((\boldsymbol{\ell} \boldsymbol{n}_s') \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot \boldsymbol{\ell})^2 \\
&\quad \times \prod_{(s,\boldsymbol{\ell}) \neq (s',\boldsymbol{\ell}')} |((\boldsymbol{\ell} \boldsymbol{n}_s') \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot \boldsymbol{\ell})^2 - ((\boldsymbol{\ell}' \boldsymbol{n}_{s'}') \cdot \boldsymbol{\beta} + \boldsymbol{q} \cdot \boldsymbol{\ell}')^2| \\
&\geq \epsilon^{R^3}.
\end{aligned}$$

Proof. The proof is based on Fubini's theorem, (A.1) and the assumption on $\boldsymbol{n}_s'$ ($1 \leq s \leq D$). Actually, in all cases we have

$$\begin{aligned}
(\boldsymbol{\ell} \boldsymbol{n}_s', \boldsymbol{\ell}) &\neq \boldsymbol{0}, \\
(\boldsymbol{\ell} \boldsymbol{n}_s' - \boldsymbol{\ell}' \boldsymbol{n}_{s'}', \boldsymbol{\ell} - \boldsymbol{\ell}') &\neq \boldsymbol{0}, \\
(\boldsymbol{\ell} \boldsymbol{n}_s + \boldsymbol{\ell}' \boldsymbol{n}_{s'}', \boldsymbol{\ell} + \boldsymbol{\ell}') &\neq \boldsymbol{0},
\end{aligned}$$

which combined with $s \leq D$ and $\boldsymbol{\ell} \in \Gamma_K$ leads to the excision of $(\boldsymbol{\beta}, \boldsymbol{q})$ in a set of measure $O(\sqrt{\epsilon})$. □

Step 3. Usage of the Lemma A.8

Now combining Lemma A.11 and Lemma A.12 (we set $\epsilon = 10^{-2} R^{-4}$) leads to

Lemma A.13. *Let* $0 < \delta < 1/10$ *and define* $N = [\delta^{-2}] + 1$. *Then there is a collection of cubes* $\Lambda_{1/N}(\boldsymbol{x}_i) \subset [0,1]^{2d}$ *(* $1 \leq i \leq N_1$ *for some* $N_1 > 1$*) with disjoint exteriors, and a set* $\boldsymbol{B} \subset [0,1]^{2d}$ *so that*

$$\begin{aligned}
\mathrm{meas}\left([0,1]^{2d} \setminus \cup_{1 \leq i \leq N_1} \Lambda_{1/N}(\boldsymbol{x}_i)\right) &\leq C L^{2d-1} \delta, \\
\mathrm{meas}([0,1]^{2d} \setminus \boldsymbol{B}) &< 1/10.
\end{aligned}$$

Moreover, if $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \cup_{1 \leq i \leq N_1} \Lambda_{1/N}(\boldsymbol{x}_i)$ *and* $\boldsymbol{\xi} \in \boldsymbol{B}$, *then*

$$|\det W| \geq c \delta^R, \ c = c(d, R) > 0.$$

We are ready to conclude the proof of Theorem A.2. Fix $\boldsymbol{\xi} \in \boldsymbol{B}$ and a $1/N \sim \delta^2$-cube $\Lambda_i = \Lambda_{1/N}(\boldsymbol{x}_i) \subset [0,1]^{2d}$ in Lemma A.13. We will deal with $(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \Lambda_i$. Denote for $\boldsymbol{k}' \in \mathbb{Z}^D \setminus \{\boldsymbol{0}\}$,

$$f(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \boldsymbol{k}' \cdot \boldsymbol{\omega}' = (\boldsymbol{k}' \otimes \boldsymbol{v}) \cdot \boldsymbol{V}.$$

Direct computations show that for $\boldsymbol{\xi} = (\boldsymbol{\beta}, \boldsymbol{q})$,

$$\begin{aligned}
&\sup_{(\boldsymbol{\alpha},\boldsymbol{\theta}) \in [0,1]^{2d}} \sup_{1 \leq j \leq R} |\partial_{\boldsymbol{\xi}}^{2j} \boldsymbol{V}|_1 \\
&= \sup_{(\boldsymbol{\alpha},\boldsymbol{\theta}) \in [0,1]^{2d}} \sup_{1 \leq j \leq R} \sum_{1 \leq s \leq D, \boldsymbol{\ell} \in \Gamma_K} |(2\pi (\boldsymbol{\ell} \boldsymbol{n}_s') \cdot \boldsymbol{\beta} + 2\pi \boldsymbol{q} \cdot \boldsymbol{\ell})^{2j} \cos 2\pi \boldsymbol{\ell} \cdot (\boldsymbol{\theta} + \boldsymbol{n}_s' \boldsymbol{\alpha})| \\
&\leq C R (\sup_s |\boldsymbol{n}_s'| + dK)^{2R} := M.
\end{aligned}$$

We have by Lemma A.13 and Lemma A.8 that

$$\inf_{(\boldsymbol{\alpha},\boldsymbol{\theta}) \in \Lambda_i} \max_{1 \leq j \leq R} |d_{\boldsymbol{\xi}}^{2j} f(\boldsymbol{\alpha}, \boldsymbol{\theta})| \geq c \delta^R M^{1-R} |\boldsymbol{k}'| \cdot |\boldsymbol{v}|, \tag{A.16}$$

where $c = c(d, R, V) > 0$.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

It suffices to apply Lemma A.6 by letting $k = 2R$. The transversality condition has been verified by (A.16). For $f(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \boldsymbol{k}' \cdot \boldsymbol{\omega}'$, we have

$$\|f\|_{2R+1} = \sup_{|\gamma| \le 2R+1, (\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \Lambda_i} |\partial_\xi^\gamma f(\boldsymbol{\alpha}, \boldsymbol{\theta})| \le C(d, R, V) \cdot (\sup_{1 \le s \le D} |\boldsymbol{n}'_s|)^{2R+1} |\boldsymbol{k}'|.$$

So using Lemma A.6 by setting $\varepsilon = \frac{\eta}{|\boldsymbol{k}'|^{4R^2}}$ leads to

$$\begin{aligned}
&\operatorname{meas}\left(\left\{(\boldsymbol{\alpha}, \boldsymbol{\theta}) \in \Lambda_i : |\boldsymbol{k}' \cdot \boldsymbol{\omega}'| \le \frac{\eta}{|\boldsymbol{k}'|^{4R^2}}\right\}\right) \\
&\le C(d, R, V)(\max_{1 \le s \le D} |\boldsymbol{n}'_s|)^{3Rd} M^{2Rd} \delta^{-R(d+1)} \eta^{\frac{1}{2R}} |\boldsymbol{k}'|^{-2R} \\
&\le C(d, R, V)(\max_{1 \le s \le D} |\boldsymbol{n}'_s|)^{5dR^2} \frac{\delta^{3dR}}{|\boldsymbol{k}'|^{2D}} \ (\text{choose } \delta = \frac{1}{10} \eta^{\frac{1}{10dR^2}}).
\end{aligned}$$

Together with Lemma A.13 this yields

$$\begin{aligned}
&\operatorname{meas}([0,1]^{2d} \setminus \mathrm{DC}_{\mathrm{nls}}(\eta)) \\
&\le C(d, R, V) L^{2d-1} \delta + C(d, R, V)(\max_{1 \le s \le D} |\boldsymbol{n}'_s|)^{5dR^2} \delta^{-4d} \sum_{\boldsymbol{k}' \in \mathbb{Z}^D \setminus \{\boldsymbol{0}\}} \frac{\delta^{3dR}}{|\boldsymbol{k}'|^{2D}} \\
&\le C(d, K, D, V)(\max_{1 \le s \le D} |\boldsymbol{n}'_s|)^{5dR^2} \delta \ (\text{since } R \ge 2 \text{ and } L = K \max_{1 \le s \le D} |\boldsymbol{n}'_s|) \\
&\le C(d, K, D, V)(\max_{1 \le s \le D} |\boldsymbol{n}'_s|)^{5dR^2} \eta^{\frac{1}{10dR^2}}.
\end{aligned}$$

This proves Theorem A.2. □

Appendix B. Some important lemmas

Let $M_2(\mathbb{C})$ denote the Banach algebra of all (2×2)-complex matrices equipped with the standard operator norm. Since the index $\xi \in \{+, -\}$, we can identify the matrix from $\mathbb{Z}^{b+d}_{\mathrm{pm}}$ to $\mathbb{C}$ with that from $\mathbb{Z}^{b+d}$ to $M_2(\mathbb{C})$. So, in the following analysis, we only focus on matrices with their entries indexes $\boldsymbol{x} = (\boldsymbol{k}, \boldsymbol{n}) \in \mathbb{Z}^{b+d}$.

We first introduce an important perturbation lemma.

Lemma B.1 (cf. Lemma 4.2 of [LW24] and Lemma A.1 of [Shi22]). *Let $X \subset \mathbb{Z}^{b+d}$ with $|X| = \#X$. Assume that A and B are two matrices with entries $A(\boldsymbol{x}, \boldsymbol{x}')$ and $B(\boldsymbol{x}, \boldsymbol{x}')$, where $\boldsymbol{x}, \boldsymbol{x}' \in X$. Assume that for some positive constants $\epsilon_1, \epsilon_2 \in (0,1)$, $C_2 \ge 0$, $c > 0$ and $0 \le M < \operatorname{diam} X$: (1) for all $\boldsymbol{x}$ and $\boldsymbol{x}'$, $|B(\boldsymbol{x}, \boldsymbol{x}')| \le \epsilon_2 (1 + |\boldsymbol{x} - \boldsymbol{x}'|)^{C_2} e^{-c|\boldsymbol{x} - \boldsymbol{x}'|}$; (2) for $|\boldsymbol{x} - \boldsymbol{x}'| > M$, $|A^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \le e^{-c|\boldsymbol{x} - \boldsymbol{x}'|}$. Suppose that $\|A^{-1}\| \le \epsilon_1^{-1}$ and*

$$|X|^2 e^{2cM} (1 + \operatorname{diam} X)^{C_2} \cdot \epsilon_2 \epsilon_1^{-1} \sum_{\boldsymbol{x} \in \mathbb{Z}^{b+d}} e^{-c|\boldsymbol{x}|} \le \frac{1}{2}.$$

Then

$$\|(A+B)^{-1}\| \le 2\epsilon_1^{-1},$$
$$|(A+B)^{-1}(\boldsymbol{x}, \boldsymbol{x}') - A^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \le \epsilon_1^{-1} e^{-c|\boldsymbol{x} - \boldsymbol{x}'|}.$$

Proof. The proof is based on the Neumann series argument similar to [Shi22, LW24]. For completeness, we give a proof.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

First, we note that $|A^{-1}(\boldsymbol{x},\boldsymbol{x}')| \le \epsilon_1^{-1} e^{cM} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|}$ for all $\boldsymbol{x},\boldsymbol{x}' \in X$. For $\boldsymbol{x}^0 = \boldsymbol{x}, \boldsymbol{x}^s = \boldsymbol{x}' \in X$ and $s \ge 1$, we can obtain by direct computations that

$$(BA^{-1})^s(\boldsymbol{x},\boldsymbol{x}') = \sum_{\boldsymbol{x}^1,\cdots,\boldsymbol{x}^{s-1},\boldsymbol{y}^1,\cdots,\boldsymbol{y}^s \in X} \prod_{t=1}^{s} B(\boldsymbol{x}^{t-1},\boldsymbol{y}^t) A^{-1}(\boldsymbol{y}^t,\boldsymbol{x}^t).$$

Thus for $s \ge 1$, one has

$$\begin{aligned}
&|(BA^{-1})^s(\boldsymbol{x},\boldsymbol{x}')| \\
&\le \sum_{\boldsymbol{x}^1,\cdots,\boldsymbol{x}^{s-1},\boldsymbol{y}^1,\cdots,\boldsymbol{y}^s \in X} e^{scM}(1+\operatorname{diam} X)^{sC_2} \cdot \epsilon_2^s \epsilon_1^{-s} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \\
&\le |X|^{2s-1} e^{scM}(1+\operatorname{diam} X)^{sC_2} \cdot \epsilon_2^s \epsilon_1^{-s} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|},
\end{aligned}$$

and for $\boldsymbol{x} \ne \boldsymbol{x}'$,

$$\begin{aligned}
&\left|\left(\sum_{s\ge1}(-BA^{-1})^s\right)(\boldsymbol{x},\boldsymbol{x}')\right| \\
&\le \sum_{s\ge1} |X|^{2s-1} e^{scM}(1+\operatorname{diam} X)^{sC_2} \cdot \epsilon_2^s \epsilon_1^{-s} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \\
&= |X| e^{cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2 \epsilon_1^{-1} \cdot \frac{1}{1-|X|^2 e^{cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2\epsilon_1^{-1}} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \\
&\le 2|X| e^{cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2 \epsilon_1^{-1} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|}.
\end{aligned}$$

Hence, from the Schur's test, we have

$$\|BA^{-1}\| \le |X| e^{cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2\epsilon_1^{-1} \cdot \sum_{\boldsymbol{x} \in \mathbb{Z}^{b+d}} e^{-c|\boldsymbol{x}|} \le \frac{1}{2}.$$

Using the Neumann series argument, we get

$$(A+B)^{-1} = A^{-1} \sum_{s\ge0} (-BA^{-1})^s.$$

Thus

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

and for $\forall\, \boldsymbol{x},\boldsymbol{x}' \in X$,

$$\begin{aligned}
&|(A+B)^{-1}(\boldsymbol{x},\boldsymbol{x}') - A^{-1}(\boldsymbol{x},\boldsymbol{x}')| \\
&\le \sum_{\boldsymbol{y}\in X} |A^{-1}(\boldsymbol{x},\boldsymbol{y})| \cdot \left|\left(\sum_{s\ge1}(-BA^{-1})^s\right)(\boldsymbol{y},\boldsymbol{x}')\right| \\
&\le \sum_{\boldsymbol{y}\in X} \epsilon_1^{-1} e^{cM} e^{-c|\boldsymbol{x}-\boldsymbol{y}|} \cdot 2|X| e^{cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2\epsilon_1^{-1} e^{-c|\boldsymbol{y}-\boldsymbol{x}'|} \\
&\le 2|X|^2 e^{2cM}(1+\operatorname{diam} X)^{C_2} \cdot \epsilon_2 \epsilon_1^{-2} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|} \\
&\le \epsilon_1^{-1} e^{-c|\boldsymbol{x}-\boldsymbol{x}'|}.
\end{aligned}$$

We complete the proof. □

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

In the following, let A be a matrix on $\mathbb{Z}^{b+d}$ satisfying

$$|A(\boldsymbol{x},\boldsymbol{x}')| \leq C_2(1+|\boldsymbol{x}-\boldsymbol{x}'|)^{C_2}e^{-b_1|\boldsymbol{x}-\boldsymbol{x}'|},\ C_2, b_1 > 0.$$

We introduce two types of coupling lemmas by iterating resolvent identities, which are repeatedly used to derive off-diagonal exponential decay estimates for Green's functions.

Remark B.2. In (linear) spectral problems (cf. [Bou07, JLS20]), one typically deals with operators such as

$$|A(\boldsymbol{x},\boldsymbol{x}')| \leq C_2 e^{-b_1|\boldsymbol{x}-\boldsymbol{x}'|}$$

without the polynomial factor $(|\boldsymbol{x}-\boldsymbol{x}'|+1)^{C_2}$. However, the extra polynomial factor does not significantly affect the proofs in [JLS20] (cf. Lemma 3.2 and Theorem 3.3) and [Liu22] (cf. Theorem 2.1), as one can replace, at each step, the estimate

$$\sum_{\substack{\boldsymbol{x}_1\in W\\ \boldsymbol{x}_2\in\Lambda\setminus W}} e^{-b_1|\boldsymbol{x}-\boldsymbol{x}'|}|G_\Lambda(\boldsymbol{x}_2,\boldsymbol{x}')| \leq (2N+1)^{2d}\sup_{\boldsymbol{x}_2\in\Lambda\setminus W}|G_\Lambda(\boldsymbol{x}_2,\boldsymbol{x}')|$$

with

$$\sum_{\substack{\boldsymbol{x}_1\in W\\ \boldsymbol{x}_2\in\Lambda\setminus W}} (|\boldsymbol{x}-\boldsymbol{x}'|+1)^{C_2}e^{-b_1|\boldsymbol{x}-\boldsymbol{x}'|}|G_\Lambda(\boldsymbol{x}_2,\boldsymbol{x}')| \leq (2N+1)^{2d+C_2}\sup_{\boldsymbol{x}_2\in\Lambda\setminus W}|G_\Lambda(\boldsymbol{x}_2,\boldsymbol{x}')|.$$

We say that an elementary region $\Lambda\in\mathcal{ER}(L)$ is in the class $\mathbb{G}$ (good) if ($G_\Lambda=(R_\Lambda AR_\Lambda)^{-1}$)

$$|G_\Lambda(\boldsymbol{x},\boldsymbol{x}')| \leq e^{-b_2|\boldsymbol{x}-\boldsymbol{x}'|}\text{ for }|\boldsymbol{x}-\boldsymbol{x}'|\geq L^{\frac{8}{9}}, \tag{B.1}$$

where $\frac{1}{2}b_1 < b_2 \leq b_1$.

We have a revised version of the coupling lemma of [Liu22] (cf. also [BGS02]).

Lemma B.3 (cf. Theorem 2.1, [Liu22]). *Let $\tilde\Lambda_0\in\mathcal{ER}(N)$ be an elementary region with the property that for all $\Lambda\subset\tilde\Lambda_0, \Lambda\in\mathcal{R}_L^{\sqrt{N}}$ with $\sqrt{N}\leq L\leq N$, the Green's function $(R_\Lambda AR_\Lambda)^{-1}$ satisfies*

$$\|(R_\Lambda AR_\Lambda)^{-1}\| \leq e^{L^{\frac{3}{4}}}.$$

Assume that for any family $\mathcal{F}$ of pairwise disjoint elementary regions of size $M=[N^{\xi}]$ contained in $\tilde\Lambda_0$,

$$\#\{\Lambda\in\mathcal{F}:\ \Lambda\text{ is not in class }\mathbb{G}\} \leq N^{\frac{1}{4}}.$$

Then for large N (depending on C_2, b_1), we have

$$|(R_{\tilde\Lambda_0}AR_{\tilde\Lambda_0})^{-1}(\boldsymbol{x},\boldsymbol{x}')| \leq e^{-(b_2-N^{-\vartheta})|\boldsymbol{x}-\boldsymbol{x}'|}\text{ for }|\boldsymbol{x}-\boldsymbol{x}'|\geq N^{\frac{8}{9}}, \tag{B.2}$$

where $\vartheta=\vartheta\in(0,1)$ is an absolute constant.

Remark B.4. There is an important difference between the original form of Lemma B.3 in [Liu22] and the present one. In [Liu22], (B.1)–(B.2) hold under the condition of $|\boldsymbol{x}-\boldsymbol{x}'|\geq\frac{N}{10}$, $0<b_2\leq\frac{4}{5}b_1$. However, the relation $0<b_2\leq\frac{4}{5}b_1$ leads to the deterioration of decay rates in the Newton iteration when solving the P-equations. To address this issue, we impose a stronger restriction $|\boldsymbol{x}-\boldsymbol{x}'|>N^{\frac{8}{9}}$ (this idea was first introduced in [HSSY20]), which ensures that $0<b_2\leq b_1$. So Lemma B.3 is proved via the argument of [Liu22] combined with the argument of [HSSY20] (we can take $b=\frac{3}{4},\tau=\frac{1}{2},\theta=\frac{8}{9}$ in Lemma 4.2 of [HSSY20]). We omit the details here.

We also need the following lemmas of [JLS20] (with again minor modifications).

Lemma B.5 (cf. Lemma 3.2, [JLS20]). *Let* $M_0 \geq (\log N)^{\frac{4}{3}}$, $\frac{b_1}{2} < b_2 \leq b_1$ *and* $M_1 \leq N$. *Suppose that* $\Lambda \subset \mathbb{Z}^{b+d}$ *is connected and* $\operatorname{diam}(\Lambda) \leq 2N+1$. *Suppose that for any* $\boldsymbol{y} \in \Lambda$, *there exists some* $W = W(\boldsymbol{y}) \in \mathcal{E}_M$ *with* $M_0 \leq M \leq M_1$ *such that* $\boldsymbol{y} \in W \subset \Lambda, \operatorname{dist}(\boldsymbol{y}, \Lambda \setminus W) \geq \frac{M}{2}$ *and*

$$\|(R_W A R_W)^{-1}\| \leq e^{M^{\frac{3}{4}}},$$

$$|(R_W A R_W)^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \leq e^{-b_2|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'| \geq M^{\frac{8}{9}}.$$

Assume further that N is large enough such that

$$\sup_{M_0 \leq M \leq M_1} e^{M^{\frac{3}{4}}}(2M+1)^{b+d+C_2} e^{b_2\sqrt{N}} \sum_{j=0}^{\infty} (M+2j+1)^{b+d} e^{-b_2(j+M/2)} \leq \frac{1}{2}.$$

Then we have

$$\|(R_\Lambda A R_\Lambda)^{-1}\| \leq 4(2M_1+1)^{b+d} e^{M_1^{\frac{3}{4}}}.$$

Lemma B.6 (cf. Theorem 3.3, [JLS20]). *Assume* $\Lambda \subset \mathbb{Z}^{b+d}$ *is connected and* $\operatorname{diam}\Lambda \leq 2N+1$. *Assume* $\operatorname{diam}\Lambda_1 \leq N^{\frac{1}{2(b+d)}}$. *Let* $M_0 \geq (\log N)^{\frac{4}{3}}$, $\frac{b_1}{2} < b_2 \leq b_1$. *Suppose that for any* $\boldsymbol{y} \in \Lambda$, *there exists some* $W = W(\boldsymbol{y}) \in \mathcal{ER}(M)$ *with* $M_0 \leq M \leq N^{\frac{3}{4}}$ *such that* $\boldsymbol{y} \in W \subset \Lambda, \operatorname{dist}(\boldsymbol{y}, \Lambda \setminus \Lambda_1 \setminus W) \geq \frac{M}{2}$ *and*

$$\|(R_W A R_W)^{-1}\| \leq e^{M^{\frac{3}{4}}},$$
$$|(R_W A R_W)^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \leq e^{-b_2|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'| \geq M^{\frac{8}{9}}.$$

Suppose that

$$\|R_\Lambda A R_\Lambda\| \leq e^{N^{\frac{3}{4}}}.$$

Then

$$|(R_\Lambda A R_\Lambda)^{-1}(\boldsymbol{x}, \boldsymbol{x}')| \leq e^{-\left(b_2 - \frac{C(b,d)}{M_0^{\frac{1}{9}}}\right)|\boldsymbol{x}-\boldsymbol{x}'|} \text{ for } |\boldsymbol{x}-\boldsymbol{x}'| \geq M^{\frac{8}{9}}.$$

Remark B.7. A similar issue in Lemma B.3 also appears in Lemma B.5 and Lemma B.6. Again, we can improve the original estimates of [JLS20] from $|\boldsymbol{x}-\boldsymbol{x}'| \geq \frac{N}{10}$ to $|\boldsymbol{x}-\boldsymbol{x}'| \geq N^{\frac{8}{9}}$.

The following lemma deals with the projection property of elementary regions.

Lemma B.8. *Let* $\Lambda \in \mathcal{ER}_{\mathbb{Z}^{b+d}}(N)$ *and denote by* $\Pi_b : \mathbb{Z}^{b+d} \to \mathbb{Z}^b$ *the canonical projection map. Then for any* $\boldsymbol{k} \in \Pi_b\Lambda$, *the section*

$$\Lambda(\boldsymbol{k}) = \{\boldsymbol{n} \in \mathbb{Z}^d : (\boldsymbol{k}, \boldsymbol{n}) \in \Lambda\}$$

either belongs to $\mathcal{ER}_{\mathbb{Z}^d}(N)$ *or is a rectangle of width at least N.*

Proof. By translation invariance, it suffices to consider the regions centered at the origin, that is, $\Lambda \in \mathcal{ER}_{\boldsymbol{0},\mathbb{Z}^{b+d}}(N)$. We proceed by analyzing the two possible forms of Λ.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

Case 1: $\Lambda = \Lambda_{N,\mathbb{Z}^{b+d}} := [-N,N]^{b+d} \cap \mathbb{Z}^{b+d}$.

In this case, the section at any $\boldsymbol{k} \in \Pi_b\Lambda$ is simply $\Lambda(\boldsymbol{k}) = [-N,N]^d \cap \mathbb{Z}^d$, which belongs to $\mathcal{ER}_{\mathbb{Z}^d}(N)$.

Case 2: $\Lambda = \Lambda^{\boldsymbol{\iota}}_{N,\mathbb{Z}^{b+d}}$ for some constraint vector $\boldsymbol{\iota} = (\iota_i)_{1\le i\le b+d} \in \{>,<,\emptyset\}^{b+d}$ with at least two nonempty components.

By definition, Λ can be expressed as:

$$\begin{aligned}\Lambda &= \Lambda_{N,\mathbb{Z}^{b+d}} \setminus \{\boldsymbol{x} = (x_i)_{1\le i\le b+d} : \ x_i \ \iota_i \ 0, \ \forall\, 1 \le i \le b+d\}\\ &= \Lambda_{N,\mathbb{Z}^{b+d}} \setminus \left(\bigcap_{1\le i\le b+d} \{\boldsymbol{x} : \ x_i \ \iota_i \ 0\}\right).\end{aligned} \tag{B.3}$$

Using De Morgan's laws, we can rewrite (B.3) as:

$$\Lambda = \Lambda_{N,\mathbb{Z}^{b+d}} \bigcap \left(\bigcup_{1\le i\le b+d} \{\boldsymbol{x} : \ x_i \ \iota_i \ 0\}^c\right), \tag{B.4}$$

where the complement sets are given by

$$\left\{\boldsymbol{x} \,\middle|\, x_i \ \iota_i \ 0\right\}^c = \begin{cases} \emptyset, & \text{if } \iota_i = \emptyset,\\ \{\boldsymbol{x} : \ x_i \ \overline{\iota_i} \ 0\}, & \text{otherwise,}\end{cases}$$

with the convention $\overline{>} := <$ and $\overline{<} := >$.

Note that the sectioning operation $A \mapsto A(\boldsymbol{k}) = \{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in A\}$ has the following properties. Given $A, B \in \mathbb{Z}^{b+d}$:

- For $\boldsymbol{k} \in \prod_b(A\cup B)$, one has $(A\cup B)(\boldsymbol{k}) = A(\boldsymbol{k}) \cup B(\boldsymbol{k})$ since
$$\{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in (A\cup B)\} = \{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in A\} \cup \{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in B\}.$$
- For $\boldsymbol{k} \in \prod_b(A\cap B)$, one has $(A\cap B)(\boldsymbol{k}) = A(\boldsymbol{k}) \cap B(\boldsymbol{k})$ since
$$\{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in (A\cap B)\} = \{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in A\} \cap \{\boldsymbol{n} : \ (\boldsymbol{k},\boldsymbol{n}) \in B\}.$$

Therefore, for a fixed $\boldsymbol{k} \in \Pi_b\Lambda$, taking the section of (B.4) yields:

$$\Lambda(\boldsymbol{k}) = \Lambda_{N,\mathbb{Z}^d} \bigcap \left(\bigcup_{1\le i\le b+d} \{\boldsymbol{x} : \ x_i \ \iota_i \ 0\}^c(\boldsymbol{k})\right). \tag{B.5}$$

We now analyze the sections of the complements for $1 \le i \le b+d$. If $\iota_i = \emptyset$, the section is trivially empty. Thus it is sufficient to consider the case $\iota_i \neq \emptyset$. We decouple the Cartesian product $\boldsymbol{x} \in \mathbb{Z}^{b+d}$ as $(\boldsymbol{k},\boldsymbol{n}) \in \mathbb{Z}^b \times \mathbb{Z}^d$:

$$\left\{\boldsymbol{x} \in \mathbb{Z}^{b+d} : \ x_i \ \overline{\iota_i} \ 0\right\}(\boldsymbol{k}) = \begin{cases} \mathbb{Z}^d, & 1 \le i \le b \ \text{ and } \ k_i \ \overline{\iota_i} \ 0,\\ \emptyset, & 1 \le i \le b \ \text{ and } \ k_i \ \iota_i \ 0,\\ \left\{\boldsymbol{n} \in \mathbb{Z}^d : \ n_{i-b} \ \overline{\iota_i} \ 0\right\}, & b+1 \le i \le b+d.\end{cases} \tag{B.6}$$

Combining (B.5) and (B.6), we observe that if there exists any $1 \le i \le b$ such that $k_i \ \overline{\iota_i} \ 0$ ($\iota_i \neq \emptyset$), the union in (B.5) evaluates to $\mathbb{Z}^d$. In this scenario, $\Lambda(\boldsymbol{k}) = \Lambda_{N,\mathbb{Z}^d}$, which belongs to $\mathcal{ER}_{\mathbb{Z}^d}(N)$.

Otherwise, the first b constraints do not contribute to the union, and the geometric shape is entirely determined by the remaining d spatial constraints. Thus, there exists a projected constraint vector

$\boldsymbol{\iota}' \in \{>, <, \emptyset\}^d$ such that

$$\bigcup_{1\leq i\leq b+d} \{\boldsymbol{x} : \ x_i \ \iota_i \ 0\}^c(\boldsymbol{k}) = \bigcup_{1\leq j\leq d} \left\{\boldsymbol{n}\in\mathbb{Z}^d : \ n_j \ \iota_j' \ 0\right\}^c.$$

Substituting this back into (B.5) and applying De Morgan's laws once more, we obtain

$$\Lambda(\boldsymbol{k}) = \Lambda_{N,\mathbb{Z}^d} \setminus \left(\bigcap_{1\leq j\leq d} \left\{\boldsymbol{n}\in\mathbb{Z}^d : \ n_j \ \iota_j' \ 0\right\}\right).$$

Let $|\boldsymbol{\iota}'|$ denote the number of nonempty components (i.e., $\iota_j' \neq \emptyset$) in $\boldsymbol{\iota}'$. We classify $\Lambda(\boldsymbol{k})$ as follows:

- If $|\boldsymbol{\iota}'| = 0$, then $\Lambda(\boldsymbol{k}) = \Lambda_{N,\mathbb{Z}^d}$, which belongs to $\mathcal{ER}_{\mathbb{Z}^d}(N)$.
- If $|\boldsymbol{\iota}'| = 1$, $\Lambda(\boldsymbol{k})$ is the full box $\Lambda_{N,\mathbb{Z}^d}$ minus a single half-space. This leaves a rectangle where $(d-1)$ sides have length $2N$ and one side has length N. Hence, it is a rectangle of width at least N.
- If $|\boldsymbol{\iota}'| \geq 2$, by the definition, $\Lambda(\boldsymbol{k}) = \Lambda^{\boldsymbol{\iota}'}_{N,\mathbb{Z}^d}$, which implies $\Lambda(\boldsymbol{k}) \in \mathcal{ER}_{\mathbb{Z}^d}(N)$.

This completes the proof for the origin-centered case. □

Acknowledgments. The authors would like to thank the handling editor and the anonymous reviewers for valuable suggestions.

Competing interests. The authors have no competing interests to declare.

Financial support. Y. Shi was supported by the National Key R&D Program (2021YFA1001600) and the NSFC (12522110). W.-M. Wang acknowledges support from the CY Initiative of Excellence, "Investissements d'Avenir" Grant No. ANR-16-IDEX-0008.

References

[AM93] M. Aizenman and S. Molchanov, 'Localization at large disorder and at extreme energies: an elementary derivation, *Comm. Math. Phys.* **157**(2) (1993), 245–278.

[Bam03] D. Bambusi, 'Birkhoff normal form for some nonlinear PDEs', *Comm. Math. Phys.* **234**(2) (2003), 253–285.

[BB13] M. Berti and P. Bolle, 'Quasi-periodic solutions with Sobolev regularity of NLS on T^d with a multiplicative potential', *J. Eur. Math. Soc.* **15**(1) (2013), 229–286.

[BBHM18] P. Baldi, M. Berti, E. Haus, and R. Montalto, 'Time quasi-periodic gravity water waves in finite depth', *Invent. Math.* **214**(2) (2018), 739–911.

[BBP13] M. Berti, L. Biasco, and M. Procesi, 'KAM theory for the Hamiltonian derivative wave equation', *Ann. Sci. Éc. Norm. Supér. (4)*, **46**(2) (2013), 301–373.

[BG01] D. Bambusi and S. Graffi, 'Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods', *Comm. Math. Phys.* **219**(2) (2001), 465–480.

[BGG85] G. Benettin, L. Galgani and A. Giorgilli, 'A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems', *Celestial Mech.* **37**(1) (1985), 1–25.

[BGS02] J. Bourgain, M. Goldstein and W. Schlag, 'Anderson localization for Schrödinger operators on $\mathbf{Z}^2$ with quasi-periodic potential', *Acta Math.* **188**(1) (2002), 41–86.

[BHM23] M. Berti, Z. Hassainia and N. Masmoudi, 'Time quasi-periodic vortex patches of Euler equation in the plane', *Invent. Math.* **233**(3) (2023), 1279–1391.

[BKM18] M. Berti, T. Kappeler and R. Montalto, 'Large KAM tori for perturbations of the defocusing NLS equation', *Astérisque* **403** (2018), viii+148.

[BN19] G. Binyamini and D. Novikov, 'Complex cellular structures', *Ann. of Math. (2)* **190**(1) (2019), 145–248.

[Bou94] J. Bourgain, 'Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE', *Internat. Math. Res. Notices* **11** (1994), 475–497.

[Bou98] J. Bourgain, 'Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations', *Ann. of Math. (2)* **148**(2) (1998), 363–439.

[Bou05a] J. Bourgain, *Green's Function Estimates for Lattice Schrödinger operators and Applications*, volume 158 of *Annals of Mathematics Studies* (Princeton University Press, Princeton, NJ, 2005).

[Bou05b] J. Bourgain, 'On invariant tori of full dimension for 1D periodic NLS', *J. Funct. Anal.* **229**(1) (2005), 62–94.

[Bou07] J. Bourgain, 'Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary', *Geom. Funct. Anal.* **17**(3) (2007), 682–706.
[BPR96] S. Basu, R. Pollack and M.-F. Roy, 'On the combinatorial and algebraic complexity of quantifier elimination', *J. ACM* **43**(6) (1996), 1002–1045.
[BW08] J. Bourgain and W.-M. Wang, 'Quasi-periodic solutions of nonlinear random Schrödinger equations', *J. Eur. Math. Soc.* **10**(1) (2008), 1–45.
[CLSY18] H. Cong, J. Liu, Y. Shi and X. Yuan, 'The stability of full dimensional KAM tori for nonlinear Schrödinger equation', *J. Differential Equations* **264**(7) (2018), 4504–4563.
[CW93] W. Craig and C. E. Wayne, 'Newton's method and periodic solutions of nonlinear wave equations', *Comm. Pure Appl. Math.* **46**(11) (1993), 1409–1498.
[CW01] A. Carbery and J. Wright, 'Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$', *Math. Res. Lett.* **8**(3) (2001), 233–248.
[CY00] L. Chierchia and J. You, 'KAM tori for 1D nonlinear wave equations with periodic boundary conditions', *Comm. Math. Phys.* **211**(2) (2000), 497–525.
[EGK16] L. H. Eliasson, B. Grébert and S. B. Kuksin, 'KAM for the nonlinear beam equation', *Geom. Funct. Anal.* **26**(6) (2016), 1588–1715.
[EK10] L. H. Eliasson and S. B. Kuksin, 'KAM for the nonlinear Schrödinger equation', *Ann. of Math. (2)* **172**(1) (2010), 371–435.
[Eli02] L. H. Eliasson, 'Perturbations of linear quasi-periodic system', in *Dynamical Systems and Small Divisors (Cetraro, 1998)*, volume 1784 of *Lecture Notes in Math.* (Springer, Berlin, 2002), 1–60.
[FS83] J. Fröhlich and T. Spencer, 'Absence of diffusion in the Anderson tight binding model for large disorder or low energy', *Comm. Math. Phys.* **88**(2) (1983), 151–184.
[Gro87] M. Gromov, 'Entropy, homology and semialgebraic geometry', *Astérisque* **145**(146) (1987), 225–240.
[GT11] B. Grébert and L , Thomann. 'KAM for the quantum harmonic oscillator', *Comm. Math. Phys.* **307**(2) (2011), 383–427.
[GYZ14] J. Geng, J. You and Z. Zhao, 'Localization in one-dimensional quasi-periodic nonlinear systems', *Geom. Funct. Anal.* **24**(1) (2014), 116–158.
[HSSY20] X. He, J. Shi, Y. Shi and X. Yuan, 'On linear stability of KAM tori via the Craig-Wayne-Bourgain method', 2020, arXiv:2003.01487.
[JLS20] S. Jitomirskaya, W. Liu and Y. Shi, 'Anderson localization for multi-frequency quasi-periodic operators on $\mathbb{Z}^d$', *Geom. Funct. Anal.* **30**(2) (2020), 457–481.
[KLW24] I. Kachkovskiy, W. Liu and W.-M. Wang, 'Spacetime quasiperiodic solutions to a nonlinear Schrödinger equation on $\mathbb{Z}$', *J. Math. Phys.* **65**(1) (2024), Paper No. 011502, 34.
[KM98] D. Y. Kleinbock and G. A. Margulis, 'Flows on homogeneous spaces and Diophantine approximation on manifolds', *Ann. of Math. (2)* **148**(1) (1998), 339–360.
[KP96] S. Kuksin and J. Pöschel, 'Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation', *Ann. of Math. (2)* **143**(1) (1996), 149–179.
[Kuk87] S. B. Kuksin, 'Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum', *Funktsional. Anal. i Prilozhen.* **21**(3) (1987), 22–37, 95.
[Liu22] W. Liu, 'Quantitative inductive estimates for Green's functions of non-self-adjoint matrices', *Anal. PDE* **15**(8) (2022), 2061–2108.
[LSZ25] S. Liu, Y. Shi and Z. Zhang, 'Anderson localized states for the nonlinear Maryland model on $\mathbb{Z}^d$, 2025, arXiv:2502.16397.
[LW24] W. Liu and W.-M. Wang, 'Nonlinear Anderson localized states at arbitrary disorder', *Comm. Math. Phys.* **405**(11) (2024), Paper No. 272, 48.
[LY10] J. Liu and X. Yuan, 'Spectrum for quantum Duffing oscillator and small-divisor equation with large-variable coefficient', *Comm. Pure Appl. Math.* **63**(9) (2010), 1145–1172.
[LY11] J. Liu and X. Yuan, 'A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations', *Comm. Math. Phys.* **307**(3) (2011), 629–673.
[PP15] C. Procesi and M. Procesi, 'A KAM algorithm for the resonant non-linear Schrödinger equation', *Adv. Math.* **272** (2015), 399–470.
[Pya69] A. S. Pyartli, 'Diophantine approximations on submanifolds of Euclidean space', *Functional analysis and Its Applications* **3**(4) (1969), 303–306.
[Shi22] Y. Shi, 'Spectral theory of the multi-frequency quasi-periodic operator with a Gevrey type perturbation', *J. Anal. Math.* **148**(1) (2022), 305–338.
[SW23] Y. Shi and W.-M. Wang, 'Anderson localized states for the quasi-periodic nonlinear wave equation on $\mathbb{Z}^d$', 2023, arXiv:2306.00513.
[Wan16] W.-M. Wang, 'Energy supercritical nonlinear Schrödinger equations: quasiperiodic solutions', *Duke Math. J.* **165**(6) (2016), 1129–1192.
[Wan21a] W.-M. Wang, 'Quasi-periodic solutions to a nonlinear Klein-Gordon equation with a decaying nonlinear term', 2021, arXiv:1609.00309.

https://doi.org/10.1017/fms.2026.10241 Published online by Cambridge University Press

[Wan21b] W.-M. Wang, 'Semi-algebraic sets method in PDE and mathematical physics', *J. Math. Phys.* **62**(2) (2021), Paper No. 021506, 12.

[Way90] C. E. Wayne, 'Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory', *Comm. Math. Phys.* **127**(3) (1990), 479–528.

[XYQ97] J. Xu, J. You and Q. Qiu, 'Invariant tori for nearly integrable Hamiltonian systems with degeneracy', *Math. Z.* **226**(3) (1997), 375–387.

[Yua02] X. Yuan, 'Construction of quasi-periodic breathers via KAM technique', *Comm. Math. Phys.* **226**(1) (2002), 61–100.

[Yua21] X. Yuan, 'KAM theorem with normal frequencies of finite limit points for some shallow water equations', *Comm. Pure Appl. Math.* **74**(6) (2021), 1193–1281.