

ONSET OF THE WAVE TURBULENCE DESCRIPTION OF THE LONGTIME BEHAVIOR OF THE NONLINEAR SCHRÖDINGER EQUATION

T. BUCKMASTER, P. GERMAIN, Z. HANI, J. SHATAH

ABSTRACT. Consider the cubic nonlinear Schrödinger equation set on a d -dimensional torus, with data whose Fourier coefficients have phases which are uniformly distributed and independent. We show that, on average, the evolution of the moduli of the Fourier coefficients is governed by the so-called *wave kinetic equation*, predicted in wave turbulence theory, on a nontrivial timescale.

We dedicate this manuscript to the memory of Jean Bourgain (1954 – 2018).

CONTENTS

1. Introduction	1
2. The general result	6
3. Formal derivation of the kinetic equation	7
4. Feynman trees: bounding the terms in the expansion	10
5. Deterministic local well-posedness	20
6. Improved integrability through randomization	24
7. Proof of the main theorem	25
8. Number theoretic results	27
References	45

1. INTRODUCTION

1.1. The Kinetic Equation. The central theme in the theory of non-equilibrium statistical physics of interacting particles is the derivation of a kinetic equation that describes the distribution of particles in phase space. The main example here is Boltzmann's kinetic theory: rather than looking at the individual trajectories of N -point particles following N -body Newtonian dynamics, Boltzmann derived a kinetic equation that described the effective dynamics of the distribution function in a certain large-particle limit (so-called the Boltzmann-Grad limit).

A parallel kinetic theory for waves, being as fundamental as particles, was proposed by physicists in the past century. Much like the Boltzmann theory, the aim is to understand the effective behavior and energy-dynamics of systems where many waves interact nonlinearly according to time-reversible dispersive or wave equations. The theory predicts that the macroscopic behavior of such nonlinear wave systems is described by a *wave kinetic equation* that gives the average distribution of energy among the available wave numbers (frequencies). Of course, the shape of this kinetic equation depends directly on the particular dispersive system/PDE that describes the reversible microscopic dynamics.

The aim of this work is to start the rigorous investigation of such passage from a reversible nonlinear dispersive PDE to an irreversible kinetic equation that describes its effective dynamics. For this, we consider the cubic nonlinear Schrödinger equations on a generic torus of size L (with periodic

boundary conditions) and with a parameter $\lambda > 0$ quantifying the importance of nonlinear effects (or equivalently via scaling, the size of the initial datum):

$$\begin{cases} i\partial_t u - \Delta_\beta u = -\lambda^2 |u|^2 u, & x \in \mathbb{T}_L^d = [0, L]^d, \\ u(0, x) = u_0(x). \end{cases} \quad (\text{NLS})$$

The spatial dimension is $d \geq 3$. Here, and throughout the paper, we denote

$$\Delta_\beta := \frac{1}{2\pi} \sum_{i=1}^d \beta_i \partial_i^2,$$

where $\beta := (\beta_1, \dots, \beta_d) \in [1, 2]^d$, and we denote $\mathbb{Z}_L^d := \frac{1}{L} \mathbb{Z}^d$, the dual space to $\mathbb{T}_L^d$.

Typically in this theory, the initial data are randomly distributed in an appropriate fashion. For us, we consider random initial data of the form

$$u_0(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \sqrt{\phi(k)} e^{2\pi i[k \cdot x + \vartheta_k(\omega)]}, \quad (1.1)$$

for some nice (smooth and localized) deterministic function $\phi : \mathbb{R}^d \rightarrow [0, \infty)$. The phases $\vartheta_k(\omega)$ are independent random variables, uniformly distributed on $[0, 1]$. Notice that the normalization of the Fourier transform is chosen so that

$$\|u_0\|_{L^2} \sim 1.$$

Filtering by the linear group and expanding in Fourier series, we write

$$u(t, x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} a_k(t) e^{2\pi i[k \cdot x + tQ(k)]}, \quad \text{where} \quad Q(k) := \sum_{i=1}^d \beta_i (k_i)^2. \quad (1.2)$$

The main conjecture of wave turbulence theory is that as $L \rightarrow \infty$ (large box limit) and $\frac{\lambda^2}{L^d} \rightarrow 0$ (weakly nonlinear limit), the quantity

$$\rho_k^L(t) = \mathbb{E}|a_k(t)|^2$$

converges to a solution of a kinetic equation. More precisely, it is conjectured that, as $L \rightarrow \infty$, $t \rightarrow \infty$ and $\frac{\lambda^2}{L^d} \rightarrow 0$, then $\rho_k^L(t) \sim \rho(t, k)$, where $\rho : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ satisfies the wave kinetic equation

$$\begin{cases} \partial_t \rho = \frac{1}{\tau} \mathcal{T}(\rho) = \frac{1}{\tau} \int_{(\mathbb{R}^d)^3} \delta(\Sigma) \delta(\Omega) \prod_{i=0}^3 \rho(k_i) \left[\sum_{i=0}^3 \frac{(-1)^i}{\rho(k_i)} \right] \prod_{i=0}^3 dk_i, \\ \rho(0, k) = \phi(k). \end{cases} \quad (\text{WKE})$$

where $\tau \sim \left(\frac{L^d}{\lambda^2}\right)^2$, we introduced the convention $k_0 = k$ and the notation

$$\begin{cases} \Sigma = \Sigma(k, k_1, \dots, k_3) = \sum_{i=0}^3 (-1)^i k_i \\ \Omega = \Omega(k, k_1, \dots, k_3) = \sum_{i=0}^3 (-1)^i Q(k_i), \end{cases}$$

and finally $\delta(\Sigma)\delta(\Omega)$ is to be understood in the sense of distributions: $\delta(\Sigma)$ is just the convolution integral over $k_1 - k_2 + k_3 = k$, whereas $\delta(\Omega = 0) := \lim_{\epsilon \rightarrow 0} \int \varphi\left(\frac{\Omega}{\epsilon}\right) dk_1 dk_2 dk_3$ for some $\varphi \in C_c^\infty(\mathbb{R})$ with $\int \varphi = 1$. Note that this is absolutely continuous to the surface measure through the formula $\delta(\Omega) = \frac{1}{|\nabla \Omega|} d\mu_\Omega$, with $d\mu_\Omega$ being the surface measures on $\{\Omega = 0\}$.

1.2. Background. In the physics literature, the wave kinetic equation (WKE) was first derived by Peierls [33] in his investigations of solid state physics; it was discovered again by Hasselmann [23, 24] in his work on the energy spectrum of water waves. The subject was revived and systematically investigated by Zakharov and his collaborators [38], particularly after the discovery of special power-type stationary solutions for the kinetic equation that serve as analogs of the Kolmogorov spectra of hydrodynamic turbulence. These so-called *Kolmogorov-Zakharov spectra* predict steady states of the corresponding microscopic system (possibly with forcing and dissipation at well-separated extreme scales), where the energy *cascades* at a constant flux through the (intermediate) frequency scales. Nowadays, wave turbulence is a vibrant area of research in nonlinear wave theory with important practical applications in several areas including oceanography and plasma physics, to mention a few. We refer to [31, 32] for recent reviews.

The analysis of (WKE) is full of very interesting questions, see [16, 22, 34] for recent developments, but we will focus here on the problem of its rigorous derivation. Several partial or heuristic derivations have been put forward for (WKE), or the closely related quantum Boltzmann equations [1, 2, 3, 13, 10, 17, 28, 30, 36]. However, to the best of our knowledge, there is no rigorous mathematical statement on the derivation of (WKE) from random data. The closest attempt in this direction is due to Lukkarinen and Spohn [29], who studied the large box limit for the discrete nonlinear Schrödinger equation at statistical equilibrium (corresponding to a stationary solution to (WKE)).

In preparation for such a study, one can first try to understand the large box and weakly nonlinear limit of (NLS) without assuming any randomness in the data. In the case where (NLS) is set on a rational torus, it is possible to extract a governing equation by retaining only exact resonances [18, 21, 20, 6]. The limiting equation is then Hamiltonian and dictates the behavior of the microscopic system (NLS on $\mathbb{T}_L^d$) on the timescales L^2/λ^2 (up to a log loss for $d = 2$) and for sufficiently small λ . It is worth mentioning that such a result is not possible if the equation is set on generic tori, since most of the exact resonances are destroyed there.

Finally, we point out that there are very few instances where the derivation of kinetic equations has been done rigorously. The fundamental result of Lanford [27], later clarified in [19], deals with the N -body Newtonian dynamics, from which emerges, in the Grad limit, the Boltzmann equation. This can be understood as a classical analog of the rigorous derivation on (WKE). Another instance of such success was the case of random linear Schrödinger operators (Anderson's model) [12, 14, 15, 35]. This can be understood as a linear analog of the problem of rigorously deriving (WKE).

1.3. The difficulties of the problem. There are several difficulties in proving the validity of (WKE) which we now enumerate:

- (a) The textbook derivation of the wave kinetic equation is done under the assumption that the independence of the data propagates for all time. This assumption cannot be verified for any nonlinear model. A way around this difficulty is to Taylor expand the profile a_k in terms of the initial data. Such an expansion can be represented by Feynman trees, and permits us to utilize the statistical independence of the data in computing the expected value of $|a_k|^2$. Moreover one needs to control the errors in such an expansion to derive the kinetic equation (WKE). These calculations are presented in Sections 4 and 5.
- (b) The wave kinetic equation induces an $O(1)$ change on its initial configuration at a timescale of τ . Thus we need to establish that for solutions of (NLS), the expansion mentioned above converge up to time τ . This requires a local existence result on a timescale which is several orders of magnitude longer than what is known. This shortcoming is a main reason why our argument cannot reach the kinetic timescale τ , and we have to contend with a derivation

over timescales where the kinetic equation only affects a relatively small change on the initial distribution, and as such coincides (up to negligible errors) with its first time-iterate.

Therefore, a pressing issue is to increase the length of the time interval $[0, T]$, over which the Taylor expansion gives a good representation of solutions to the nonlinear problem. For deterministic data, the best known results that give effective bounds in terms of L come from our previous work [6] which gives a description of the solution up to times $\sim L^2/\lambda^2$ (up to a $\log L$ loss for $d = 2$) and for $\lambda \ll 1$. Such timescale would be very short for our purposes.

To increase T , we have to rely on the randomness of the initial data. Roughly speaking, for a random field that is normalized to 1 in $L^2(\mathbb{T}_L^d)$, its L^∞ norm can be heuristically bounded on average by $L^{-d/2}$. Therefore, regarding the nonlinearity $\lambda^2|u|^2u$ as a nonlinear potential Vu with $V = \lambda^2|u|^2$ and $\|V\|_{L^\infty} \lesssim \lambda^2 L^{-d}$, one would hope that this should get a convergent expansion on an interval $[0, T]$ provided that $T\lambda^2 L^{-d} \ll 1$, which amounts to $T \leq \sqrt{\tau}$. This is the target in this manuscript.

The heuristic presented above can be implemented by relying on Khinchine-type improvements to the Strichartz norms of a linear solution $e^{it\Delta_\beta}u_0$ with random initial data u_0 . Similar improvements have been used to lower the regularity threshold for well-posedness of nonlinear dispersive PDE. Here, the aim is to prolong the existence time and improve the Taylor approximation. The randomness gives us better control on the size of the linear solution over the interval $[0, T]$, while an improved *deterministic* Strichartz estimate for $\|e^{it\Delta_\beta}\psi\|_{L^p([0, T] \times \mathbb{T}^d)}$ with $\psi \in L^2(\mathbb{T}^d)$, allows us to maintain the random improvement for the nonlinear problem. The genericity of the (β_i) is crucial (as was first observed in [11]), and allows us to go beyond the limiting $T^{1/p}$ growth that occurs on the rational torus. Unfortunately, the available estimates here (including those in [11]) are not optimal for some ranges of the parameters λ and L , which is why, in $d = 3$, our result in Theorem 1.1 below falls short of the timescale $\sqrt{\tau} \sim L^3/\lambda$.

- (c) To derive the kinetic equation in the large box limit, using the expansion for $\rho_k^L(t) = \mathbb{E}|a_k(t)|^2$, one has to prove equidistribution theorems for the quasi-resonances over a very fine scale, i.e., T^{-1} . Since T could be $\gg L^2$, such scales are much finer than the any equidistribution scale on the rational torus. Again, here the genericity of the (β_i) is crucial. For this we use and extend a recent result of Bourgain on pair correlation for irrational quadratic forms [5].

1.4. The main result. Precise statements of our results in arbitrary dimensions $d \geq 3$ will be given in Section 2. Those statements depend on several parameters coming from equidistribution of lattice points and Strichartz estimates. For the purposes of this introductory section, we present a less general theorem without the explicit appearance of these parameters.

Theorem 1.1. *Consider the cubic (NLS) on the three-dimensional torus $\mathbb{T}_L^3$. Assume that the initial data are chosen randomly as in (1.1). There exists $\delta > 0$ such that the following holds for L sufficiently large and $L^{-A} \leq \lambda \leq L^B$ (for positive A and B):*

$$\mathbb{E}|a_k(t)|^2 = \phi(k) + \frac{t}{\tau}\mathcal{T}(\phi)(k) + O_{\ell^\infty}\left(L^{-\delta}\frac{t}{\tau}\right), \quad L^\delta \leq t \leq T, \quad (1.3)$$

where $\tau = \frac{1}{2}\left(\frac{L^3}{\lambda^2}\right)^2$ and $T \sim \frac{L^{3-\gamma}}{\lambda^2}$, for some $0 < \gamma < 1$ stated explicitly in Theorem 2.2.

We note that the right-hand side of (1.3) is nothing but the first time-iterate of the wave kinetic equation (WKE) with initial data ϕ (cf. (1.1)) which coincides (up to the error term in (1.3)) with the exact solution of the (WKE) over long times scales, but shorter than the kinetic timescale τ .

The proof this theorem can be split into three components:

- (1) Section 4: Feynman tree representation. In this section we derive the Taylor expansion of the nonlinear solution in terms of the initial data. Roughly speaking, we write the Fourier modes of the nonlinear solution $a_k(t)$ (see (1.2)) as follows:

$$a_k(t) = \sum_{n=0}^N \mathcal{J}_n(t, k)(\mathbf{a}^{(0)}) + R_{N+1}(t, k)(\mathbf{a}^{(t)}),$$

where $\mathcal{J}_n$ are sums of monomials of degree $2n + 1$ in the initial data $\mathbf{a}^{(0)}$, and R_N is the remainder which depends on the nonlinear solution $\mathbf{a}^{(t)}$. Each term of $\mathcal{J}_n$ can be represented by a Feynman tree which makes the calculations of $\mathbb{E}(\mathcal{J}_n \mathcal{J}_{n'})$ more transparent. Such terms appear in the expansion of $\mathbb{E}|a_k|^2$. The estimates in this section rely on essentially sharp bounds on quasi-resonant sums of the form

$$\sum_{\vec{k} \in \mathbb{Z}_L^{rd}} \mathbf{1}(|\vec{k}| \lesssim 1) \mathbf{1}(|\mathcal{Q}(k)| \sim 2^{-A}) \lesssim 2^{-A} L^{rd} \quad (1.4)$$

where $\mathbf{1}(S)$ denotes the characteristic function of a set S and $\mathcal{Q}$ is an irrational quadratic form. Since A will be taken large $2^A \sim T \gg L^2$, such estimates belong to the realm of number theory and will be a consequence the third component of this work.

The bounds we obtain for such interaction are good up to times of order $\sqrt{\tau}$ which is sufficient given the restrictions on the time interval of convergence imposed by the second component below.

- (2) Section 5: Construction of solutions. In this section we construct solutions on a time interval $[0, T]$ via a contraction mapping argument. To maximize T while maintaining a contraction, we rely on the Khinchine improvement to the space-time Strichartz bounds, as well as the long-time Strichartz estimates on generic irrational tori proved in [11]. It is here that our estimates are very far from optimal, since there is no proof to the conjectured optimal Strichartz estimates.
- (3) Section 8: Equidistribution of irrational quadratic forms. The purpose of this section is two-fold. The first is proving bounds on quasi-resonant sums like those in (1.4) for the largest possible T , and the second is to extract the exact asymptotic, with effective error bounds, of the leading part of the sum. It is this leading part that converges to the kinetic equation collision kernel as $L \rightarrow \infty$.

Here we remark, that if $\mathcal{Q}$ is a rational form, then the largest A for which one could hope for an estimate like (1.4) is $2^A \sim L^2$ which reflects the fact that a rational quadratic form cannot be equidistributed at scales smaller than L^{-2} (at the level of NLS, it would yield a time interval restriction of $T \lesssim L^2$ for the rational torus). However, for generic irrational quadratic forms, $\mathcal{Q}$ is actually equidistributed at much finer scales than L^{-2} . Here, we adapt a recent work of Bourgain [5] which will allow us to reach equidistribution scales essentially up to L^{-d} .

1.5. Notations. In addition to the notation introduced earlier for $\mathbb{T}_L^d = [0, L]^d$ and $\mathbb{Z}_L^d = \frac{1}{L}\mathbb{Z}^d$, we use standard notations. A function f on $\mathbb{T}_L^d$ and its Fourier transform $\hat{f}$ on $\mathbb{Z}_L^d$ are related by

$$f(x) = \frac{1}{L^d} \sum_{\mathbb{Z}_L^d} \hat{f}_k e^{2\pi i k \cdot x} \quad \text{and} \quad \hat{f}_k = \int_{\mathbb{T}_L^d} e^{-2\pi i k \cdot x} f(x) dx.$$

Parseval's theorem becomes

$$\|f\|_{L^2(\mathbb{T}_L^d)}^2 = \|\hat{f}\|_{\ell^2(\mathbb{Z}_L^d)}^2 = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} |\hat{f}_k|^2.$$

We adopt the following definition for weighted ℓ^p spaces: if $p \geq 1$, $s \in \mathbb{R}$, and $b \in \ell^p$,

$$\|b\|_{\ell^{p,s}(\mathbb{Z}_L^d)} = \left[\frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} (\langle k \rangle^s |b_k|)^p \right]^{1/p}.$$

Sobolev spaces $H^s(\mathbb{T}^d)$ are then defined naturally by

$$\|f\|_{H^s(\mathbb{T}^d)} = \|\langle k \rangle^s \hat{f}\|_{\ell^{2,s}(\mathbb{Z}_L^d)}.$$

For functions defined on $\mathbb{R}^d$, we adopt the normalization

$$f(x) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot x} \hat{f}(\xi) d\xi \quad \text{and} \quad \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i k \cdot x} f(x) dx.$$

We denote by C any constant whose value does not depend on λ or L . The notation $A \lesssim B$ means that there exists a constant C such that $A \leq CB$. We also write $A \lesssim L^{r^+} B$, if for any $\epsilon > 0$ there exists C_ϵ such that $A \leq C_\epsilon L^{r^+ + \epsilon} B$. Similarly $A \gtrsim L^{r^-} B$, if for any $\epsilon > 0$ there exists C_ϵ such that $A \geq C_\epsilon L^{r^- - \epsilon} B$. Finally we use the notation $u = O_X(B)$ to mean $\|u\|_X \lesssim B$.

We would like to thank Peter Sarnak for pointing us to unpublished work by Bourgain [5]. This reference helped us improve an earlier version of our work. We also would like to thank Peter and Simon Myerson for many helpful and illuminating discussions.

2. THE GENERAL RESULT

We start by writing the equations for the interaction representation $(a_k(t))_{k \in \mathbb{Z}_L^d}$, given in (1.2):

$$\begin{cases} i\dot{a}_k = -\left(\frac{\lambda}{L^d}\right)^2 \sum_{\substack{(k_1, \dots, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} a_{k_1} \overline{a_{k_2}} a_{k_3} e^{-2\pi i t \Omega(k, k_1, k_2, k_3)} \\ a_k(0) = a_k^0 = \sqrt{\phi(k)} e^{i\vartheta_k(\omega)}, \end{cases} \quad (2.1)$$

where we recall $\Omega(k, k_1, k_2, k_3) = Q(k) - Q(k_1) + Q(k_2) - Q(k_3)$, and $\vartheta_k(\omega)$ are i.i.d. random variables that are uniformly distributed in $[0, 2\pi]$. Our results depend on two parameters: the equidistribution parameter ν , and a Strichartz parameter θ_p , which we now explain.

2.1. The Equidistribution parameter ν . The interaction frequency $\Omega(k, k_1, k_2, k_3)$ above is an irrational quadratic form. Such quadratic forms can be equidistributed at scales that are much smaller than the finest scale $\sim L^{-2}$ of rational forms.

We will denote by ν the largest real number such that for all $k \in \mathbb{Z}_L^d$, $|k| \leq 1$, and $\epsilon > 0$, there exists $\delta > 0$ such that, for $|a|, |b| < 1$ with $b - a \geq L^{-\nu^-}$,

$$\sum_{\substack{a \leq \Omega(k, k_1, k_2, k_3) \leq b \\ |k_1|, |k_2|, |k_3| \leq 1 \\ k - k_1 + k_2 - k_3 = 0}} 1 = (1 + O(L^{-\delta})) L^{2d} \int_{|k_1|, |k_2|, |k_3| \leq 1} \mathbb{1}_{a \leq \Omega(k, k_1, k_2, k_3) \leq b} \delta(k - k_1 + k_2 - k_3) dk_1 dk_2 dk_3.$$

Proposition 2.1. *With the above definition for ν , we have*

- (i) *If $\beta_i = 1$ for all $i \in \{1, \dots, d\}$, $\nu = 2$.*
- (ii) *If the β_i are generic, $\nu = d$.*

Proof. The first assertion is classical, e.g., see [6]. The second assertion is proved in Section 8. $\square$

2.2. The Strichartz parameter θ_d . Our proof relies on long-time Strichartz estimates, which are used to maintain linear bounds for the nonlinear problem. The genericity of the β 's gives crucial improvements from the rational case. The improved estimates for generic β 's were proved in [11],

$$\|e^{it\Delta_\beta} P_N \psi\|_{L^p_{t,x}([0,T] \times \mathbb{T}^d)} \lesssim N^{\frac{d}{2} - \frac{d+2}{p}} \left(1 + \frac{T}{N^{\gamma(d,p)}}\right)^{1/p} \|\psi\|_{L^2(\mathbb{T}^d)}$$

for some $0 \leq \gamma(d,p) \leq d-2$. The N^γ term can be thought of as the time it takes for a focused wave with localized wave number $\leq N$, to focus again. For the rational torus $\gamma = 0$.

Here we only need to use the $L^4_{t,x}([0,T] \times \mathbb{T}^d_L)$ norm, and therefore we introduce a parameter θ_d to record how the constant in the $L^4_{t,x}([0,T] \times \mathbb{T}^d_L)$ estimates depends on L . By scaling, the result in [11] translates into,

$$\|e^{it\Delta_\beta} P_{k \leq 1} \psi\|_{L^4_{t,x}([0,T] \times \mathbb{T}^d_L)} \lesssim L^{0+} \left(1 + \frac{T}{L^{\theta_d}}\right)^{1/4} \|\psi\|_{L^2(\mathbb{T}^d_L)} \quad (2.2)$$

$$\text{where } \theta_d := \begin{cases} \frac{4}{13} + 2, & d = 3 \\ \frac{(d-2)^2}{2(d-1)} + 2, & d \geq 4. \end{cases}$$

2.3. The approximation theorem. With these parameters defined, we state the approximation theorem for the cubic NLS in dimension $d \geq 3$ and generic β 's.

Theorem 2.2. *Assume the β 's are generic and $d \geq 3$. Let $\phi_0 : \mathbb{R}^d \rightarrow \mathbb{R}^+$, a rapidly decaying smooth function. Suppose that $a_k(0) = \sqrt{\phi(k)} e^{i\vartheta_k(\omega)}$ where $\vartheta_k(\omega)$ are i.i.d. random variables uniformly distributed in $[0, 2\pi]$. For every ϵ_0 , a sufficiently small constant, and $L > L_*(\epsilon_0)$ sufficiently large, the following holds:*

There exists a set $E_{\epsilon_0, L}$ of measure $\mathbb{P}(E_{\epsilon_0, L}) \geq 1 - e^{-L^{\epsilon_0}}$ such that: if $\omega \in E_{\epsilon_0, L}$, then for any $L > L_(\epsilon_0)$, the solution $a_k(t)$ of (NLS) exists in $C_t H^s([0, T] \times \mathbb{T}^d_L)$ for*

$$T \sim \begin{cases} \lambda^{-2} L^{\frac{d+\theta_d}{2} - 4\epsilon_0} & \text{if } L^{-\frac{d+\theta_d}{4}} \lesssim \lambda \lesssim L^{\frac{d-\theta_d}{4} - 2\epsilon_0}, \\ \lambda^{-4} L^{d-8\epsilon_0} & \text{if } \lambda \geq L^{\frac{d-\theta_d}{4} - 2\epsilon_0}. \end{cases}$$

Moreover,

$$\mathbb{E} \left[|a_k(t)|^2 \mathbf{1}_{E_{\epsilon_0, L}} \right] = \phi(k) + \frac{t}{\tau} \mathcal{T}_3(\phi)(k) + O_{\ell^\infty} \left(L^{-\epsilon_0} \frac{t}{\tau} \right), \quad L^{\epsilon_0} \leq t \leq T, \text{ and } \tau = \frac{L^{2d}}{2\lambda^4}.$$

For $d = 3, 4$, the solutions exist globally in time [4, 26], and one has the same estimate without multiplying with $\mathbf{1}_{E_{\epsilon_0}}$ inside the expectation.

Here we note that the error could be controlled in a much stronger norm than ℓ^∞ , and that other randomizations of the data are possible (complex Gaussians for instance) without any significant changes in the proof.

3. FORMAL DERIVATION OF THE KINETIC EQUATION

In this section, we present the formal derivation of the kinetic equation, whose basic steps we shall follow in the proof. The starting point is equation (2.1) integrated in time,

$$a_k(t) = a_k^0 + \frac{i\lambda^2}{L^{2d}} \int_0^t \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} a_{k_1} \overline{a_{k_2}} a_{k_3} e^{-2\pi i s \Omega(k, k_1, k_2, k_3)} ds \quad (3.1)$$

The derivation of the kinetic equation proceeds as follows:

Step 1: expanding in the data. Noting the symmetry in (3.1) in the variables k_1 and k_3 , we have upon integrating by parts twice, and substituting (2.1) for $\dot{a}_k$,

$$a_k(t) = a_k^0 \quad (3.2a)$$

$$+ \frac{\lambda^2}{L^{2d}} \sum_{k - k_1 + k_2 - k_3 = 0} a_{k_1}^0 \overline{a_{k_2}^0} a_{k_3}^0 \frac{1 - e^{-2\pi i t \Omega(k, k_1, k_2, k_3)}}{2\pi \Omega(k, k_1, k_2, k_3)} \quad (3.2b)$$

$$+ 2 \frac{\lambda^4}{L^{4d}} \sum_{\substack{k - k_1 + k_2 - k_3 = 0 \\ k_1 - k_4 + k_5 - k_6 = 0}} a_{k_4}^0 \overline{a_{k_5}^0} a_{k_6}^0 \overline{a_{k_2}^0} a_{k_3}^0 \frac{1}{2\pi \Omega(k, k_1, k_2, k_3)} \left[\frac{e^{-2\pi i t \Omega(k, k_4, k_5, k_6, k_2, k_3)} - 1}{2\pi \Omega(k, k_4, k_5, k_6, k_2, k_3)} - \frac{e^{-2\pi i t \Omega(k_1, k_4, k_5, k_6)} - 1}{2\pi \Omega(k_1, k_4, k_5, k_6)} \right] \quad (3.2c)$$

$$+ \frac{\lambda^4}{L^{4d}} \sum_{\substack{k - k_1 + k_2 - k_3 = 0 \\ k_2 - k_4 + k_5 - k_6 = 0}} a_{k_1}^0 \overline{a_{k_4}^0} a_{k_5}^0 \overline{a_{k_6}^0} a_{k_3}^0 \frac{1}{2\pi \Omega(k, k_1, k_2, k_3)} \left[\frac{e^{-2\pi i t \Omega(k, k_1, k_4, k_5, k_6, k_3)} - 1}{2\pi \Omega(k, k_1, k_4, k_5, k_6, k_3)} - \frac{e^{-2\pi i t \Omega(k_2, k_4, k_5, k_6)} - 1}{2\pi \Omega(k_2, k_4, k_5, k_6)} \right] \quad (3.2d)$$

$$+ \{\text{higher order terms}\}. \quad (3.2e)$$

where we denoted $\Omega(k, k_1, k_2, k_3, k_4, k_5) = Q(k) + \sum_{i=1}^5 (-1)^i Q(k_i)$; we also used the convention that, if $a = 0$, $\frac{e^{2\pi i t a} - 1}{2\pi a} = it$, while, if $a = b = 0$, $\frac{1}{2\pi a} \left(\frac{e^{2\pi i t(a+b)} - 1}{2\pi(a+b)} - \frac{e^{2\pi i t a} - 1}{2\pi a} \right) = -\frac{1}{2} t^2$.

Step 2: parity pairing. We now compute $\mathbb{E}|a_k|^2$, where the expectation $\mathbb{E}$ is understood with respect to the random phases (random parameter ω). The key observation is,

$$\mathbb{E}(a_{k_1}^0 \dots a_{k_s}^0 \overline{a_{\ell_1}^0} \dots \overline{a_{\ell_s}^0}) = \begin{cases} \phi_{k_1} \dots \phi_{k_s} & \text{if there exists a } \gamma \text{ such that } k_{\gamma(i)} = \ell_i \\ 0 & \text{otherwise.} \end{cases}$$

(for $k \in \mathbb{Z}_L^d$, we write $\phi_k = \phi(k)$). Computing $\mathbb{E}(|a_k|^2)$ with the help of the above formula, we see that, there are no terms of order λ^2 . There are two kinds of terms of order λ^4 obtained as follows: either by pairing the term of order λ^2 , namely (3.2b), with its conjugate, or by pairing one of the terms of order λ^4 , (3.2c) or (3.2d), with the term of order 1, namely a_k^0 . Overall, this leads to

$$\mathbb{E}|a_k|^2(t) = \phi_k + \frac{2\lambda^4}{L^{4d}} \sum_{k - k_1 + k_2 - k_3 = 0} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(t\pi \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 + \{\text{higher order terms}\} + \{\text{degenerate cases}\},$$

where degenerate cases occur for instance if k, k_1, k_2, k_3 are not distinct¹. The details of the computation are as follows:

¹Degenerate cases, like higher order terms, have smaller order of magnitude, on the timescales we consider as will be illustrated in Section 4.

(a) Consider first $\mathbb{E}|(3.2b)|^2 = \mathbb{E}(3.2b)\overline{(3.2b)}$, and denote k_1, k_2, k_3 the indices in (3.2b) and k'_1, k'_2, k'_3 the indices in $\overline{(3.2b)}$. There are two possibilities:

- $\{k_1, k_3\} = \{k'_1, k'_3\}$, in which case $k_2 = k'_2$, and $\Omega(k, k_1, k_2, k_3) = \Omega(k, k'_1, k'_2, k'_3)$.
- $(k_2 = k_1 \text{ or } k_3)$ and $(k'_2 = k'_1 \text{ or } k'_3)$, in which case $\Omega(k, k_1, k_2, k_3) = \Omega(k, k'_1, k'_2, k'_3) = 0$.

Overall, we find, neglecting degenerate cases (which occur if k, k_1, k_2, k_3 are not distinct),

$$\mathbb{E}|(3.2b)|^2 = \frac{2\lambda^4}{L^{4d}} \sum_{k-k_1+k_2-k_3=0} \phi_{k_1} \phi_{k_2} \phi_{k_3} \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 + \frac{4\lambda^4}{L^{4d}} t^2 \sum_{k_1, k_3} \phi_k \phi_{k_1} \phi_{k_2}.$$

(b) Consider next the pairing of a_k^0 with (3.2c), which contributes $2\mathbb{E}\Re \left[(3.2c) \overline{a_k^0} \right]$. The possible pairings are

- $\{k, k_2\} = \{k_4, k_6\}$, implying $k_3 = k_5$, and leading to $\Omega(k_1, k_4, k_5, k_6) = -\Omega(k, k_1, k_2, k_3)$, and $\Omega(k, k_4, k_5, k_6, k_2, k_1) = 0$.
- $(k_3 = k_2 \text{ or } k)$ and $(k_5 = k_4 \text{ or } k_6)$ in which case $\Omega(k, k_1, k_2, k_3) = \Omega(k_1, k_4, k_5, k_6) = 0$.

This gives, neglecting degenerate cases,

$$\begin{aligned} 2\mathbb{E}\Re \left[\overline{a_k^0} (3.2c) \right] &= \frac{8\lambda^4}{L^{4d}} \times \\ &\sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_2} \phi_{k_3} \Re \left[\frac{e^{-2\pi i t \Omega(k, k_1, k_2, k_3)} - 1}{4\pi^2 \Omega(k, k_1, k_2, k_3)^2} \right] - \frac{8\lambda^4}{L^{4d}} t^2 \sum_{k_1, k_3} \phi_k \phi_{k_2} \phi_{k_3} \\ &= -\frac{2\lambda^4}{L^{4d}} \sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_2} \phi_k \left[\frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 - \frac{8\lambda^4}{L^{4d}} t^2 \sum_{k_1, k_3} \phi_k \phi_{k_2} \phi_{k_3}, \end{aligned}$$

where we used in the last line the symmetry between the variables k_1 and k_3 , as well as the identity $\Re(e^{iy} - 1) = -2|\sin(y/2)|^2$, for $y \in \mathbb{R}$.

(c) Finally, the pairing of a_k^0 with (3.2d) can be discussed similarly, to yield

$$2\mathbb{E}\Re \left[\overline{a_k^0} (3.2d) \right] = \frac{2\lambda^4}{L^{4d}} \sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_3} \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 + \frac{4\lambda^4}{L^{4d}} t^2 \sum_{k_1, k_3} \phi_k \phi_{k_2} \phi_{k_3},$$

Summing the above expressions for $\mathbb{E}|(3.2b)|^2$, $2\mathbb{E}\Re \left[\overline{a_k^0} (3.2c) \right]$ and $2\mathbb{E}\Re \left[\overline{a_k^0} (3.2d) \right]$ gives the desired result.

Step 3: the big box limit $L \rightarrow \infty$. Assuming that $\Omega(k, k_1, k_2, k_3)$ is equidistributed on a scale

$$L^{-\nu} \ll \frac{1}{t}, \quad (3.3)$$

we see that, as $L \rightarrow \infty$,

$$\begin{aligned} &\sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 \sim \\ &L^{2d} \int \delta(\Sigma) \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 dk_1 dk_2 dk_3. \end{aligned}$$

Step 4: the large time limit $t \rightarrow \infty$ Observe that² $\int \frac{(\sin x)^2}{x^2} dx = \pi^2$, so that, in the sense of distributions,

$$\left| \frac{\sin(\pi t \Omega)}{\pi \Omega} \right|^2 \sim t \delta(\Omega) \quad \text{as } t \rightarrow \infty.$$

Therefore, as $t \rightarrow \infty$,

$$\begin{aligned} & \sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 \\ & \sim t L^{2d} \int \delta(\Sigma) \delta(\Omega) \phi(k) \phi(k_1) \phi(k_2) \phi(k_3) \left[\frac{1}{\phi(k)} - \frac{1}{\phi(k_1)} + \frac{1}{\phi(k_2)} - \frac{1}{\phi(k_3)} \right] dk_1 dk_2 dk_3 \\ & = t L^{2d} \mathcal{T}(\phi, \phi, \phi). \end{aligned}$$

Conclusion: relevant timescales for the problem. Overall, we find, assuming that the above limits are justified

$$\mathbb{E}|a_k|^2(t) = \phi_k + 2 \frac{\lambda^4}{L^{2d}} t \mathcal{T}(\phi, \phi, \phi) + \{\text{lower order terms}\}. \quad (3.4)$$

This suggests that the actual timescale of the problem is

$$\tau = \frac{L^{2d}}{2\lambda^4},$$

and that, setting $s = \frac{t}{\tau}$, the governing equation should read

$$\partial_s \phi = \mathcal{T}(\phi, \phi, \phi) \quad (3.5)$$

In which regime is this approximation expected? Let T be the timescale over which we consider the equation.

- In order for (3.4) to hold, the condition (3.3) has to hold, and the limits $L \rightarrow \infty$ and $T \rightarrow \infty$ have to be taken: one needs

$$T \ll L^\nu, \quad L \gg 1, \quad \text{and} \quad T \gg 1.$$

- In order for the nonlinear evolution of (3.5) to affect an $O(\kappa)$ change on the initial data, the two conditions above should be satisfied; in addition T should be of the order of $\kappa\tau$ (equivalently $s \sim \kappa$). Thus we find the conditions

$$1 \ll T \approx \kappa\tau \ll L^\nu \quad \text{and} \quad \kappa^{\frac{1}{4}} L^{d/2} \gg \lambda \gg \kappa^{\frac{1}{4}} L^{d/2-\nu/4}.$$

4. FEYNMAN TREES: BOUNDING THE TERMS IN THE EXPANSION

Since we are considering the problem with rapidly decaying ϕ , then the rapid decay of ϕ yields all the bounds one needs for wave numbers $|k| \geq L^{0+}$, thus we might as well consider ϕ to be compactly supported.

²This follows from Plancherel's theorem, and the fact that the Fourier transform of $\frac{1}{\pi} \frac{\sin x}{x}$ is the characteristic function of $[-\frac{1}{2\pi}, \frac{1}{2\pi}]$.

4.1. Expansion of the solution in the data. We follow mostly the notations in Lukkarinen-Spohn [29], Section 3 (see also [9]).

The iterates of ϕ , considered in the previous section, can be represented through trees (at least up to lower order error terms). To explain these trees, let us start with the equation satisfied by the amplitude of the wave number a_k

$$a_k(t) = a_k^0 + \frac{i\lambda^2}{L^{2d}} \int_0^t \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} a_{k_1} \bar{a}_{k_2} a_{k_3} e^{-2\pi i s \Omega(k, k_1, k_2, k_3)} ds,$$

$$a_k(t) = a_k^0 + \frac{i\lambda^2}{L^{2d}} \int_0^t \mathcal{P}_3(a)(s) e^{-2\pi i s \Omega} ds.$$

where the subscript in $\mathcal{P}_3$ is to indicate that it is a monomial of degree 3, and where we suppressed the k dependence for convenience. The expansion can be obtained by integrating by parts on the oscillating factor $e^{-2\pi i s \Omega}$. Thus the first integration by parts gives the cubic expansion,

$$a_k(t) = a_k^0 + \frac{i\lambda^2}{L^{2d}} \mathcal{P}_3(a)(0) F_0^t + \frac{i\lambda^2}{L^{2d}} \int_0^t \dot{\mathcal{P}}_3(a)(s) F_s^t ds, \quad F_s^t := \int_s^t e^{-2\pi i \tau \Omega} d\tau.$$

Using the equation for a , we see that $\dot{\mathcal{P}}_3(a)$ consists of three monomials of degree 5, and if we denote one of them by $\mathcal{P}_5$, then the integral term consists of three integrals of the type,

$$\left(\frac{i\lambda^2}{L^{2d}} \right)^2 \int_0^t \mathcal{P}_5(a)(s) e^{-2\pi i s \Omega} F_s^t ds.$$

Another integration by parts gives the quintic expansion, which consist of three terms of the form

$$\left(\frac{i\lambda^2}{L^{2d}} \right)^2 \mathcal{P}_5(a)(0) G_0^t + \left(\frac{i\lambda^2}{L^{2d}} \right)^2 \int_0^t \dot{\mathcal{P}}_5(a)(s) G_s^t ds, \quad G_s^t = \int_s^t e^{-2\pi i \tau \Omega} F_\tau^t d\tau.$$

Consequently, to compute the expansion to order N we need to integrate by parts N times on the oscillating exponentials, giving the expansion,

$$a_k(t) = \sum_{n=0}^N \mathcal{J}_n(t, k)(\mathbf{a}^{(0)}) + R_{N+1}(t, k)(\mathbf{a}^{(t)}), \quad (4.1)$$

where $\mathcal{J}_n = \sum_{\ell} \mathcal{J}_{n, \ell}$, and each $\mathcal{J}_{n, \ell}$ is a monomial of degree $2n + 1$ generated by the n^{th} integration by parts. The index ℓ is a vector whose entries keep track of the history of how the monomial $\mathcal{J}_{n, \ell}$ was generated. R_{N+1} is the remaining time integral.

Each $\mathcal{J}_{n, \ell}$ can be represented by a tree similar to Figure 1 below. which we now explain.

The trees will be constructed in reverse order of their usage. Therefore the labeling of the wave numbers will be done backwards: $n - j$, $0 \leq j \leq n$.

The tree corresponding to $\mathcal{J}_{n, \ell}$, is given as follows.

- There are $n + 1$ levels in the tree, the bottom level is the 0^{th} level. Descending from the top to the bottom, each level is generated from the previous level by an integration by parts step. Thus level j represents the terms present after $n - j$ integration by parts.

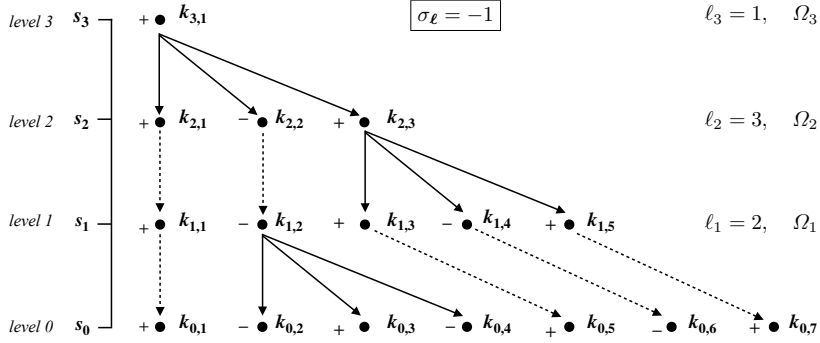


FIGURE 1. tree of depth 3.

- $k_{j,m}$ denote the wave numbers present in level j , and therefore $1 \leq m \leq 2(n-j)+1$.
- $k_{j,m}$ has a parity σ_m due to complex conjugation. For m odd or even, $\sigma_m = +1$ or $\sigma_m = -1$ respectively.

$$a_{k_{j,m},\sigma_m} = \begin{cases} a_{k_{j,m}} & \text{if } \sigma_m = +1 \\ \overline{a_{k_{j,m}}} & \text{if } \sigma_m = -1 \end{cases}.$$

- For each level j , we associate a number ℓ_j , which signals out the wave number k_{j,ℓ_j} which has 3 branches. This is the wave number of the a (or $\bar{a}$) that was differentiated by the j^{th} integration by parts. The index vector ℓ , keeps track of the integration by parts history in the tree for $\mathcal{J}_{n,\ell}$. The entries ℓ_j , $1 \leq j \leq n$, are given by

$$\ell = (\ell_1, \dots, \ell_n) \in \{1, \dots, 2n-1\} \times \{1, \dots, 2n-3\} \times \dots \times \{1, 2, 3\} \times \{1\}.$$

- The tree has a signature $\sigma_\ell = \prod_{j=1}^n (-1)^{\ell_j+1}$.
- *Transition rules.* To go from level j to level $j-1$, the wave numbers are related as follows

$$\begin{cases} k_{j,m} = k_{j-1,m} & \text{for } m < \ell_j \\ k_{j,m} = k_{j-1,m+2} & \text{for } \ell_j < m \\ k_{j,\ell_j} = k_{j-1,\ell_j} - k_{j-1,\ell_j+1} + k_{j-1,\ell_j+2} \end{cases} \quad (4.2)$$

Note that for any j , $\sum_{m=1}^{2(n-j)+1} (-1)^{m+1} k_{j,m} = k_{n,1} = k$. The wave numbers at level 0, i.e., those present in $\mathcal{J}_{n,\ell}$, are labeled

$$\mathbf{k} = (k_{0,1}, \dots, k_{0,2n+1}) \in (\mathbb{Z}_L^d)^{2n+1}.$$

- At each level j , the derivative of the element with wave number k_{j,ℓ_j} (due to the integration by parts), generates a oscillatory term with frequency

$$\Omega_j(\mathbf{k}) = (-1)^{\ell_j+1} (Q(k_{j,\ell_j}) - Q(k_{j-1,\ell_j}) + Q(k_{j-1,\ell_j+1}) - Q(k_{j-1,\ell_j+2})).$$

- We introduce variables $\mathbf{s} = (s_0, \dots, s_n) \in \mathbb{R}_+^{n+1}$; $t_j(\mathbf{s}) = \sum_{k=0}^{j-1} s_k$, $1 \leq j \leq n$. This choice of variables can be explained as follows. Repeated integration by parts generates terms of the form

$$\int_0^t g_0(s_0) \int_{s_0}^t g_1(s_1) \dots \int_{s_{n-2}}^t g_{n-1}(s_{n-1}) = \int_0^t g_0(s_0) \int_0^{t-s_0} g_1(s_0 + s_1) \dots \int_0^{t-s_0-\dots-s_{n-2}} g_{n-1}(s_0 + \dots + s_{n-1})$$

which can be written as

$$\int_{\mathbb{R}_+^{n+1}} g_0(s_0) g_1(s_0 + s_1) \dots g_{n-1}(s_0 + \dots + s_{n-1}) \delta(t - \sum_{l=0}^n s_l)$$

With this notation at hand,

$$\begin{aligned} \mathcal{J}_0 &= a_k^0, \quad \mathcal{J}_1 = \mathcal{J}_{1,1} = (3.2b), \quad \mathcal{J}_2 = \mathcal{J}_{2,(1,1)} + \mathcal{J}_{2,(2,1)} + \mathcal{J}_{2,(3,1)}, \\ \mathcal{J}_{2,(2,1)} &= (3.2d), \quad \mathcal{J}_{2,(1,1)} = \mathcal{J}_{2,(3,1)} = \frac{1}{2}(3.2c), \end{aligned}$$

and Figure 1 represents $\mathcal{J}_{3,(2,3,1)}$. The general formula for $\mathcal{J}_{n,\ell}$ is given by

$$\mathcal{J}_{n,\ell}(t, \mathbf{k}) = \left(\frac{i\lambda^2}{L^{2d}} \right)^n \sigma_\ell \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1}} \delta_{k_{n,1}}^k \prod_{j=1}^{2n+1} a_{k_{0,j}, \sigma_j}^0 \int_{\mathbb{R}_+^{n+1}} \prod_{m=1}^n e^{-2\pi i t_m(s) \Omega_m(\mathbf{k})} \delta\left(t - \sum_0^n s_i\right) d\mathbf{s} \quad (4.3)$$

Here and throughout the manuscript we write

$$\delta_j^k = \begin{cases} 1, & k = j, \\ 0, & k \neq j, \end{cases}$$

while $\delta(\cdot)$ is the Dirac delta.

Finally, we write $R_n(t, k)(\mathbf{a}) = \sum_{\ell} \int_0^t R_{n,\ell}(t, s_0; k)(\mathbf{a}^{(s_0)}) ds_0$, where

$$\begin{aligned} R_{n,\ell}(t, s_0; k)(\mathbf{b}) &= \left(\frac{i\lambda^2}{L^{2d}} \right)^n \sigma_\ell \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1}} \delta_{k_{n,1}}^k \prod_{j=1}^{2n+1} b_{k_{0,j}, \sigma_j} \int_{\mathbb{R}_+^n} \prod_{j=1}^n e^{-2\pi i t_j(s) \Omega_j(\mathbf{k})} \times \\ &\quad \delta\left(t - s_0 - \sum_1^n s_i\right) d\mathbf{s}. \end{aligned} \quad (4.4)$$

4.2. Bound on the correlation. Our aim is to prove the following proposition.

Proposition 4.1. *If $t < L^{d-\epsilon_0}$, then*

$$\left| \sum_{n+n'=S} \sum_{\ell, \ell'} \mathbb{E}(\mathcal{J}_{n,\ell}(t, k) \overline{\mathcal{J}_{n',\ell'}(t, k)}) \right| \lesssim_S (\log t)^2 \left(\frac{t}{\sqrt{\tau}} \right)^S \frac{1}{t}. \quad (4.5)$$

Remark 4.2. *The trivial estimate would be that*

$$\left| \sum_{n+n'=S} \sum_{\ell, \ell'} \mathbb{E}(\mathcal{J}_{n,\ell}(t, k) \overline{\mathcal{J}_{n',\ell'}(t, k)}) \right| \lesssim \left(\frac{t}{\sqrt{\tau}} \right)^S.$$

Indeed, $\mathcal{J}_{n,\ell}\mathcal{J}_{n',\ell'}$ comes with a prefactor $\left(\frac{\lambda^2}{L^{2d}}\right)^{n+n'}$; the size of the domains where the time integration takes place is $O(t^{n+n'})$; and the summation over $\mathbf{k}$ and $\mathbf{k}'$ is over $2d(n+n'+1)$ dimensions, half of which are canceled by the pairing (see below), out of which d further dimensions are canceled by the requirement that $k_{n,1} = k$. Overall, this gives a bound $\left(\frac{\lambda^2}{L^{2d}}\right)^{n+n'} \times t^{n+n'} \times L^{d(n+n')} = \left(\frac{t}{\sqrt{\tau}}\right)^{n+n'}$.

Therefore, the above proposition essentially allows a gain of $\frac{1}{t}$ over the trivial bound. This gain of $\frac{1}{t}$ comes from cancelations in the “non degenerate interactions” as will be exhibited by equation (4.13).

Before we start the proof of Proposition 4.1, we shall classify the transitions (4.2) as degenerate if

$$k_{j,\ell_j} \in \{k_{j-1,\ell_j}, k_{j-1,\ell_j+2}\},$$

i.e., if the parallelogram with vertices $(k_{j,\ell_j}, k_{j-1,\ell_j-1}, k_{j-1,\ell_j}, k_{j-1,\ell_j+2})$ degenerates into a line. In this case $\Omega_{\ell_j}(\mathbf{k}) = 0$. When *all* transitions in a tree that represents $\mathcal{J}_{n,\ell}$ are degenerate we denote the term by $D_{n,\ell}(t, k)$, and if one transition is non degenerate we denote it by $\tilde{\mathcal{J}}_{n,\ell}(t, k)$, that is

$$\mathcal{J}_{n,\ell}(t, k) = \tilde{\mathcal{J}}_{n,\ell}(t, k) + D_{n,\ell}(t, k) \quad (4.6)$$

$$\begin{aligned} D_{n,\ell}(t, k) &= \left(\frac{i\lambda^2}{L^{2d}}\right)^n \sigma_\ell \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1}} \delta_{k_{n,1}}^k (1 - \Delta(\mathbf{k})) \prod_{j=1}^{2n+1} a_{k_{0,j},\sigma_j}^0 \int_{\mathbb{R}_+^{n+1}} \delta\left(t - \sum_0^n s_i\right) d\mathbf{s} \\ &= 2^n \frac{t^n}{n!} \left(\frac{i\lambda^2}{L^{2d}}\right)^n \sigma_\ell a_k^0 \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^n} \prod_{j=1}^n |a_{k_j}^0|^2, \end{aligned} \quad (4.7)$$

$$\tilde{\mathcal{J}}_{n,\ell}(t, k) = \left(\frac{i\lambda^2}{L^{2d}}\right)^n \sigma_\ell \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1}} \delta_{k_{n,1}}^k \Delta(\mathbf{k}) \prod_{j=1}^{2n+1} a_{k_{0,j},\sigma_j}^0 \int_{\mathbb{R}_+^{n+1}} \prod_{m=1}^n e^{-2\pi i t_m(s) \Omega_m(\mathbf{k})} \delta\left(t - \sum_0^n s_i\right) d\mathbf{s}, \quad (4.8)$$

where $\Delta(\mathbf{k}) = 1 - \prod_{j=1}^n \delta_{\{k_{j-1,\ell_j+1}, k_{j-1,\ell_j+1+\sigma_j,\ell_j}\}}^{k_{j,\ell_j}}$. Note that $\Delta(\mathbf{k}) = 1$ whenever $\tilde{\mathcal{J}}_{n,\ell}(t, k) \neq 0$.

4.3. Cancellation of degenerate interactions. As can be seen from a simple computation in the formula for $D_{n,\ell}$, the contribution of each $\mathbb{E}(D_{n,\ell}(t, k) \bar{D}_{n',\ell'}(t, k))$ to the sum in (4.5) is of size $\sim \left(\frac{t}{\sqrt{\tau}}\right)^S$, which is too large. Luckily, all those terms cancel out as shows the lemma below.

Note that this cancellation between graph expectations is essentially due to the invariance of the expectation $\mathbb{E}|a_k|^2$ under Wick renormalization, which is a classical trick in the analysis of the nonlinear Schrödinger equation that eliminates all degenerate interactions. However, working at the level of graph expectations might be applicable in more general contexts.

Lemma 4.3. *For any $S \geq 2$*

$$\sum_{n+n'=S} \sum_{\ell,\ell'} \mathbb{E}(D_{n,\ell}(t, k) \bar{D}_{n',\ell'}(t, k)) = 0.$$

Proof. First we note that since each level in the tree has parity equal to 1, then

$$\sum_{\ell} \sigma_\ell = \prod_{j=1}^n (\text{parity of line } j) = (1)^n = 1.$$

Hence by equation (4.7)

$$\sum_{\ell} D_{n,\ell}(t, k) = 2^n \frac{t^n}{n!} \left(\frac{i\lambda^2}{L^{2d}} \right)^n \left(\sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^n} \prod_{j=1}^n |a_{k_j}^0|^2 \right) a_k^0.$$

Thus we obtain

$$\sum_{n+n'=S} \sum_{\ell, \ell'} \mathbb{E}(D_{n,\ell}(t, k) \overline{D_{n',\ell'}(t, k)}) = 2^S t^S \left(\frac{\lambda^2}{L^{2d}} \right)^S \left(\sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^S} \prod_{j=1}^S |a_{k_j}^0|^2 \right) |a_k^0|^2 \left(\sum_{n+n'=S} \frac{i^{n-n'}}{n!n'!} \right).$$

The result will follow once we show that

$$\sum_{n+n'=S} \frac{i^{n-n'}}{n!n'!} = 0.$$

This follows by parametrizing the above sum as $\{(n, n') = (S - j, j) : j = 0, \dots, S\}$, which gives

$$\sum_{n+n'=S} \frac{i^{n-n'}}{n!n'!} = i^S \sum_{j=0}^S \frac{(-1)^j}{(S-j)!j!} = \frac{i^S}{S!} \sum_{j=0}^S \frac{(-1)^j S!}{(S-j)!j!} = \frac{i^S}{S!} (1+x)^S \Big|_{x=-1} = 0.$$

□

4.4. Estimate on non-degenerate interactions. Proposition 4.1 now follows from the following lemma:

Lemma 4.4. *Suppose $G_{n',\ell'}(t, k) \in \{\tilde{\mathcal{J}}_{n',\ell'}(t, k), D_{n',\ell'}(t, k)\}$, then for $0 < t < L^{d-\epsilon_0}$,*

$$\left| \mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t, k) \overline{G_{n',\ell'}(t, k)}) \right| \lesssim_n (\log t)^2 \left(\frac{t}{\sqrt{\tau}} \right)^{n+n'} \frac{1}{t}.$$

Proof. We will only consider the case of $G_{n',\ell'}(t, k) = \tilde{\mathcal{J}}_{n',\ell'}(t, k)$, since the case $G_{n',\ell'}(t, k) = D_{n',\ell'}(t, k)$ is easier to bound. Using the identity

$$\delta \left(t - \sum_0^n s_j \right) = \frac{1}{2\pi} \int e^{-i\alpha(t - \sum_{j=0}^n s_j)} d\alpha,$$

we can write for any $(e_0, \dots, e_n) \in \mathbb{R}^{n+1}$ and $\eta > 0$

$$\int_{\mathbb{R}_+^{n+1}} \prod_{j=0}^n e^{-is_j e_j} \delta \left(t - \sum_0^n s_i \right) ds = \frac{e^{\eta t}}{2\pi} \int_{\mathbb{R}} e^{-i\alpha t} \prod_{j=0}^n \frac{i}{\alpha - e_j + i\eta} d\alpha. \quad (4.9)$$

Thus by choosing $\eta = \frac{1}{t}$, we have

$$\begin{aligned} \tilde{\mathcal{J}}_{n,\ell}(t, k) &= \frac{ie}{2\pi} \left(-\frac{\lambda^2}{L^{2d}} \right)^n \sigma_{\ell} \sum_{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1}} \delta_{k_{n,1}}^k \Delta(\mathbf{k}) \prod_{j=1}^{2n+1} a_{k_{0,j}, \sigma_j}^0 \times \\ &\quad \int \frac{e^{-i\alpha t} d\alpha}{(\alpha - \Omega_1 - \Omega_2 - \dots - \Omega_n + \frac{i}{t}) \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})}. \end{aligned}$$

Here we employed the notation $\Omega_j = 2\pi\Omega_j(\mathbf{k})$.

To bound $\mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t,k)\overline{\tilde{\mathcal{J}}_{n',\ell'}(t,k)})$, we will simplify the notation by setting $k_{0,2n+1+j} = k'_{0,j}$, which also preserves the parity convention. Consequently, we have $2n + 2n' + 2$ wave numbers present in the expression for $\mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t,k)\overline{\tilde{\mathcal{J}}_{n',\ell'}(t,k)})$.

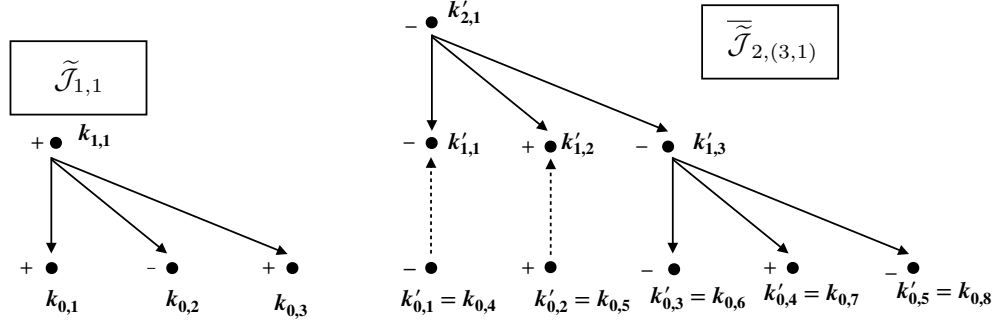


FIGURE 2. Relabeling trees.

Next, since the phases are i.i.d. with mean 0, then only specific pairing of the wave numbers contribute nonzero terms, namely the pairing should be between terms with the same wave number and opposite parity. For this reason we introduce $\mathcal{P} = \mathcal{P}(n, n', \sigma, \sigma')$ a class of pairings indices and parities, as illustrated in Figure 3

$$\mathcal{P} \ni \psi : \{1, \dots, 2n + 2n' + 2\} \rightarrow \{1, \dots, 2n + 2n' + 2\} \Leftrightarrow \psi(j) = l \Rightarrow \psi(l) = j, \quad \text{and} \quad \sigma_{\psi(j)} = -\sigma_j$$

Furthermore, we define the pairing of wave numbers induced by ψ , $\Gamma_\psi(\mathbf{k}, \mathbf{k}') = \prod_{j=1}^{2n+2n'+2} \delta_{k_{0,\psi(j)}}^{k_{0,j}}$. By the independence of the phases $\vartheta_{k_{0,j}}(\omega)$, we have,

$$\left| \mathbb{E}_\omega \left[\prod_{j=1}^{2n+1} e^{i\sigma_{0,j}\vartheta_{k_{0,j}}(\omega)} \prod_{j'=1}^{2n'+1} e^{-i\sigma_{0,j'}\vartheta_{k_{0,j'}}(\omega)} \right] \right| \lesssim \sum_{\psi \in \mathcal{P}} \Gamma_\psi(\mathbf{k}, \mathbf{k}'),$$

Hence we obtain

$$\begin{aligned} \left| \mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t,k)\overline{\tilde{\mathcal{J}}_{n',\ell'}(t,k)}) \right| &\lesssim \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum_{\psi \in \mathcal{P}} \sum_{\substack{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1} \\ \mathbf{k}' \in (\mathbb{Z}_L^d)^{2n'+1}}} \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \times \\ &\quad \left| \int \frac{e^{-i\alpha t} d\alpha}{(\alpha - \Omega_1 - \dots - \Omega_n + \frac{i}{t}) \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})} \times \right. \\ &\quad \left. \int \frac{e^{i\alpha' t} d\alpha'}{(\alpha' - \Omega'_1 - \dots - \Omega'_{n'} + \frac{i}{t}) \dots (\alpha' - \Omega'_{n'} + \frac{i}{t})(\alpha' + \frac{i}{t})} \right|. \end{aligned}$$

where $\mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') = \delta_k^{k_{n,1}} \delta_{k'_{n',1}}^k \Delta(\mathbf{k}) \Delta(\mathbf{k}') \Gamma_\psi(\mathbf{k}, \mathbf{k}') \prod_{j=1}^{2n+1} \sqrt{\phi(k_{0,j})} \prod_{j'=1}^{2n'+1} \sqrt{\phi(k'_{0,j'})}$.

By Hölder's inequality, for any $m \geq 1$ and $b_1, \dots, b_{n'+1} \in \mathbb{R}$,

$$\int_{\mathbb{R}} \frac{d\alpha'}{|\alpha' - b_1 + \frac{i}{t}| \dots |\alpha' - b_{m+1} + \frac{i}{t}|} \lesssim t^m, \quad (4.10)$$

and applying this bound to the α' integral yields

$$\left| \mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t,k) \overline{\tilde{\mathcal{J}}_{n',\ell'}(t,k)}) \right| \lesssim t^{n'} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum_{\psi \in \mathcal{P}} \sum_{\substack{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1} \\ \mathbf{k}' \in (\mathbb{Z}_L^d)^{2n'+1}}} \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \times \left| \int \frac{d\alpha}{(\alpha - \sum_{l=1}^n \Omega_l + \frac{i}{t}) \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})} \right|.$$

Let $p = p(\mathbf{k})$ be the smallest integer such that $k_{p+1,\ell_p} \notin \{k_{p,\ell_p+1}, k_{p,\ell_p+1+\sigma_{p+1,\ell_p}}\}$, i.e., in the tree for $\tilde{\mathcal{J}}_{n,\ell}$ the transition from level $p+1$ to level p is not degenerate. Note that $0 \leq p \leq n-1$, and

$$\left| \mathbb{E}(\tilde{\mathcal{J}}_{n,\ell}(t,k) \overline{\tilde{\mathcal{J}}_{n',\ell'}(t,k)}) \right| \lesssim t^{n'} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum_{\psi \in \mathcal{P}} \sum_{\substack{\mathbf{k} \in (\mathbb{Z}_L^d)^{2n+1} \\ \mathbf{k}' \in (\mathbb{Z}_L^d)^{2n'+1}}} \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \times \left| \int \frac{d\alpha}{(\alpha - \sum_{l=p+1}^n \Omega_l + \frac{i}{t})^{p+1} \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})} \right| =: \sum_{\psi} \sum_p \mathfrak{I}_{p,\psi}.$$

We now set

$$\begin{aligned} I_1 &= \ell_p, & I_2 &= \ell_p + 1, & I_3 &= \ell_p + 2, & \mathbf{k}_p &= (k_{p,I_1}, k_{p,I_2}, k_{p,I_3}), \\ J_1 &= \psi(I_1), & J_2 &= \psi(I_2), & J_3 &= \psi(I_3). \end{aligned}$$

Note that, by definition of p ,

$$\{I_1, I_2, I_3\} \cap \{J_1, J_2, J_3\} = \emptyset.$$

The figure below illustrate all the introduced notations and parings for the product of two non degenerate terms.

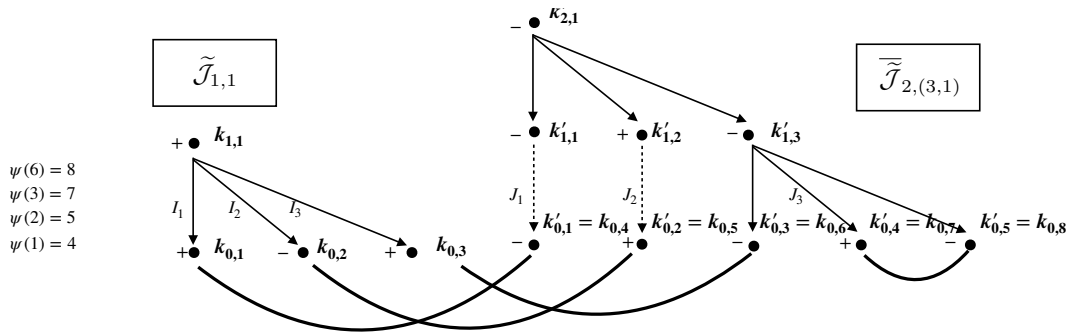


FIGURE 3. Pairing trees.

We distinguish three cases depending on the values of the numbers J_i .

Case 1: $J_1, J_2, J_3 \geq 2n+2$. For a fixed p and ψ we sum over all wave numbers in $\mathfrak{I}_{p,\psi}$ that yield degenerate transitions, i.e., wave numbers generated in rows $0 \leq l \leq p-1$. This contributes L^{dp} to

the bound,

$$\mathfrak{I}_{p,\psi} \lesssim t^{n'} L^{dp} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \sum_{\mathbf{k}_p} \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \int \frac{d\alpha}{|(\alpha - \sum_{l=p+1}^n \Omega_l + \frac{i}{t})^{p+1} \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})|},$$

where $\sum^*$ stands for the sum over $k_{p,j}$, where $1 \leq j \leq 2(n-p)+1$ and $j \notin \{I_1, I_2, I_3\}$, and $k'_{0,j'}$ for $1 \leq j' \leq 2n'+1$ with $j' \notin \{J_1, J_2, J_3\}$.

The contribution of the above integral is acceptable as long as the denominator is $O(\langle \alpha \rangle^{-2})$. Therefore, it suffices to prove the desired bound when the domain of integration reduces to $\alpha \in [-R, R]$, for some $R > 0$, since the resonance moduli Ω_i are bounded. Furthermore by bounding the integrand by $\frac{t^{n-1}}{|\alpha - \Omega_{p+1} - \dots - \Omega_n + \frac{i}{t}| |\alpha + \frac{i}{t}|}$, matters reduce to estimating

$$t^{n'+n-1} L^{dp} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \sum_{\mathbf{k}_p} \int_{-R}^R \frac{d\alpha}{|\alpha - \sum_{l=p+1}^n \Omega_l + \frac{i}{t}| |\alpha + \frac{i}{t}|}. \quad (4.11)$$

By the identity $k_{p+1,I_1} - k_{p,I_2} = \sigma_{p,I_1}(k_{p,I_3} - k_{p,I_1})$, this can also be written

$$t^{n'+n-1} L^{dp} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \int_{-R}^R \left[\sum_{\mathbf{k}_p} \delta_{k_{p,I_2} + \sigma_{p,I_1}(k_{p,I_3} - k_{p,I_1})}^{k_{p+1,I_1}} \frac{1}{|\alpha - \sum_{l=p+1}^n \Omega_l + \frac{i}{t}|} \right] \frac{d\alpha}{|\alpha + \frac{i}{t}|},$$

and since $\sum_{j=1}^{2(n-p)+1} \sigma_{p,j} k_{p,j} = k_{n,1} = k$, we note that

$$\begin{aligned} \Omega_{p+1} + \dots + \Omega_n &= Q(k) - \sum_{j=1}^{2(n-p)+1} \sigma_{p,j} Q(k_{p,j}) = \\ &= -\sigma_{p,I_1} Q(k_{p,I_1}) - \sigma_{p,I_3} Q(k_{p,I_3}) - \sigma_{p,I_2} Q(k_{p+1,I_1} - \sigma_{p,I_1}(k_{p,I_3} - k_{p,I_1})) + C \end{aligned} \quad (4.12)$$

where C depends only on k and the variables $k_{p,j}$ with $j \notin \{I_1, I_2, I_3\}$.

By setting $P = k_{p,I_3}$ and $R = k_{p,I_1}$, for $t < L^\nu$ we bound

$$\sum_{\substack{P, R \in \mathbb{Z}_L^d \\ |P|, |R| \leq 1}} \frac{1}{| -Q(P) + Q(R) - Q(N + P - R) + C + \frac{i}{t} |},$$

using the equidistribution result in Section 8. If $| -Q(P) + Q(R) - Q(N + P - R) + C | \leq t^{-1}$, we have by Corollary 8.5

$$\sum_{\substack{P, R \in \mathbb{Z}_L^d \\ |P|, |R| \leq 1}} \frac{1}{| -Q(P) + Q(R) - Q(N + P - R) + C + \frac{i}{t} |} \lesssim t \left(L^{2d} \frac{1}{t} + L^d \right).$$

Whereas for $| -Q(P) + Q(R) - Q(N + P - R) + C | \geq t^{-1}$, we bound

$$\begin{aligned} \sum_{\substack{P, R \in \mathbb{Z}_L^d \\ |P|, |R| \leq 1}} \frac{1}{| -Q(P) + Q(R) - Q(N + P - R) + C + \frac{i}{t} |} &\lesssim \\ \sum_{\frac{1}{t} < 2^n \lesssim 1} 2^{-n} \sum_{| -Q(P) + Q(R) - Q(N + P - R) + C | \sim 2^n} 1 &\lesssim L^{2d} \sum_{\frac{1}{t} < 2^n \lesssim 1} 1 \lesssim L^{2d} \log t. \end{aligned} \quad (4.13)$$

Therefore, we can bound (4.11) by

$$t^{n'+n-1} L^{dp} \left(L^{2d} \log t + t L^d \right) \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \Gamma_\psi(\mathbf{k}, \mathbf{k}') \int_{-R}^R \frac{d\alpha}{|\alpha + \frac{i}{t}|},$$

The sum $\sum^*$ is over $2(n + n' - p - 2)$ variables; however, because of the pairing Γ_ψ , half of them drop out, so that the remaining sum is $\lesssim L^{d(n+n'-p-2)}$. And since $\int_{-R}^R \frac{d\alpha}{|\alpha + \frac{i}{t}|} \lesssim \log t$, the above expression can be bounded by,

$$t^{n'+n-1} \left((\log t)^2 L^{d(n+n')} + t (\log t) L^{d(n+n'-1)} \right) \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'},$$

which gives the stated bound.

Case 2: only two of J_1, J_2, J_3 are $\geq 2n + 2$. Suppose for instance that $J_2 \leq 2n + 1$. Then, there exists $I_4 \leq 2(n - p) + 1$ such that $\psi(I_4) = J_4 \geq 2n + 2$ (such an index exists because there is an odd number of elements in the set of elements in $\{1, \dots, 2(n - p) + 1\} \setminus \{I_1, I_2, I_3, J_2\}$, so they cannot be paired together completely). One can then follow the above argument replacing I_2 by I_4 .

Case 3: two of J_1, J_2, J_3 are $\leq 2n + 1$ Assume for instance that $J_1, J_3 \leq 2n + 1$. Proceeding as in Case 1, it suffices to bound

$$t^{n'} L^{dp} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \sum_{k_{p,I_1}, k_{p,I_3}} \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \int \frac{d\alpha}{|(\alpha - \sum_{l=p+1}^n \Omega_l + \frac{i}{t})^{p+1} \dots (\alpha - \Omega_n + \frac{i}{t})(\alpha + \frac{i}{t})|},$$

where $\sum^*$ is the sum over $k_{p,j}$, with $j \in \{1, \dots, 2(n - p) + 1\} \setminus \{I_1, I_3, J_1, J_3\}$, and over $k_{0,j'}$, with $j' \in \{1, \dots, 2n' + 1\}$.

A crucial observation is that, since $\sum_{j=1}^{2(n-p)+1} \sigma_{p,j} k_{p,j} = k_{n,1} = k$, the wave numbers k_{p,I_1} and k_{p,I_3} do not contribute to this sum since the pairing $k_{p,I_1} = k_{p,J_1}$ and $k_{p,I_3} = k_{p,J_3}$, causes them to cancel one another. Furthermore, $0 \leq p \leq n - 2$ since $J_1, J_3 \leq 2n + 1$, and therefore we bound the integrand by $\frac{t^{n-1}}{|\alpha - \Omega_{p+2} - \dots - \Omega_n + \frac{i}{t}| |\alpha + \frac{i}{t}|}$. Overall, we can bound the above by

$$L^{dp} t^{n+n'-1} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'} \sum^* \mathcal{A}_\psi(\mathbf{k}, \mathbf{k}') \int_{-R}^R \left(\sum_{k_{p,I_1}, k_{p,I_3}} \frac{1}{|\alpha - \Omega_{p+2} - \dots - \Omega_n + \frac{i}{t}|} \right) \frac{d\alpha}{|\alpha + \frac{i}{t}|}.$$

From equation (4.12), we conclude

$$\sum_{l=p+2}^n \Omega_l = -\sigma_{p,I_1} Q(k_{p,I_1}) - \sigma_{p,I_3} Q(k_{p,I_3}) - \sigma_{p,I_2} Q(k_{p+1,I_1} - \sigma_{p,I_1}(k_{p,I_3} - k_{p,I_1})) + C$$

where C only depends on the variables in $\sum^*$. Applying (4.13) enables us to bound the inner sum by $L^{2d} \log t$, and the α integral by $\log t$. Finally, the number of variables in $\sum^*$ is $2(n + n' - p - 1)$. By pairing them there are only $n + n' - p - 1$, and fixing $k_{n,1} = k$ brings their number down to $n + n' - p - 2$. Thus $\sum^*$ will contribute $\lesssim L^{d(n+n'-p-2)}$. Overall, we obtain the bound

$$\lesssim (\log t)^2 t^{n'+n-1} L^{d(n+n')} \left(\frac{\lambda^2}{L^{2d}} \right)^{n+n'},$$

which is the desired estimate. $\square$

5. DETERMINISTIC LOCAL WELL-POSEDNESS

Local or long time existence of smooth solutions is usually carried out by using Strichartz estimates to bound solutions. The known Strichartz estimates for our problem (2.2) are not sufficient to allow us to prove existence of solutions for a long time interval where the wave kinetic equation (WKE) emerges. However, if the data is assumed to be random, then one has improved estimates due to Khinchin's inequality [7]. Based on this, we first present a local well-posedness theorem provided the data satisfies a certain estimate. In Section 6, we show that such an improved estimate occurs with high probability.

Moreover, to use the results from Sections 4 and 8, we will restrict discussion to the case $T < L^{d-\epsilon_0}$.

5.1. Strichartz estimate. Recall equation (2.2), which can be written as,

$$\|e^{it\Delta_\beta} P_1 \psi\|_{L^4_{t,x}([0,T] \times \mathbb{T}_L^d)} \leq S_{d,\epsilon} \|\psi\|_{L^2(\mathbb{T}_L^d)}, \quad S_{d,\epsilon} := C_{d,\epsilon} L^\epsilon \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4}$$

Moreover if we denote the characteristic function of the unit cube centered at $j \in \mathbb{Z}^d$ by $\mathbf{1}_{B_j}$, and define

$$\widehat{\psi_{B_j}}(k) = \mathbf{1}_{B_j}(k) \widehat{\psi}(k), \quad \text{and therefore } \psi_{B_0} = P_1 \psi.$$

Then, using the Galilean invariance $|e^{-it\Delta_\beta} \psi_{B_j}(x)| = |[e^{-it\Delta_\beta} (e^{2\pi i j x} \psi)_{B_0}](x - 2tj)|$, we have

$$\|e^{it\Delta_\beta} \psi_{B_j}\|_{L^4_{t,x}([0,T] \times \mathbb{T}_L^d)} \leq S_{d,\epsilon} \|\psi\|_{L^2(\mathbb{T}_L^d)}. \quad (5.1)$$

Converting this estimate to its dual, and applying the Christ-Kiselev inequality, one gets

$$\left\| \int_0^T e^{-is\Delta} F_{B_j}(s) ds \right\|_{L^2(\mathbb{T}_L^d)} \leq S_{d,\epsilon} \|F\|_{L^{4/3}_{t,x}([0,T] \times \mathbb{T}_L^d)} \quad (5.2)$$

$$\left\| \int_0^t e^{i(t-s)\Delta} F_{B_j}(s) ds \right\|_{L^4_{t,x}([0,T] \times \mathbb{T}_L^d)} \leq S_{d,\epsilon}^2 \|F\|_{L^{4/3}_{t,x}([0,T] \times \mathbb{T}_L^d)} \quad (5.3)$$

for an appropriate choice of $C_{d,\epsilon}$ used in the definition of $S_{d,\epsilon}$.

5.2. A priori bound in Z_T^s and energy. Let Z_T^s denote the function space defined by the norm,

$$\|u\|_{Z_T^s} = \left(\sum_{j \in \mathbb{Z}^d} \langle j \rangle^{2s} \|u_{B_j}\|_{L^4_{t,x}([0,T] \times \mathbb{T}_L^d)}^2 \right)^{1/2}, \quad (5.4)$$

then the Z_T^s norm of the nonlinearity is bounded.

Lemma 5.1. Fix $s > \frac{d}{2}$. For every $\epsilon_0 > 0$, and an appropriate choice of C_{d,ϵ_0} , we have

$$\left\| \int_0^t e^{i(t-s)\Delta_\beta} |u(s)|^2 u(s) ds \right\|_{Z_T^s} \leq S_*^2 \|u\|_{Z_T^s}^3, \quad S_* := S_{d,\epsilon_0} = C_{d,\epsilon_0} L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4}. \quad (5.5)$$

Proof. Consider $v \in L_{t,x}^{\frac{4}{3}}([0, T] \times \mathbb{T}_L^d)$, and let $\tilde{v}(s, x) = \int_s^T e^{i(s-s')\Delta_\beta} v(s') ds'$, then

$$\begin{aligned}
& \left\| \int_0^t e^{i(t-s)\Delta_\beta} P_{B_j} |u(s)|^2 u(s) ds \right\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d)} \\
&= \sup_{\|v\|_{L_{t,x}^{4/3}}=1} \int_0^T \int_0^t \int_{\mathbb{T}_L^d} \left[e^{i(t-s)\Delta_\beta} P_{B_j} |u(s)|^2 u(s) \right] \overline{v(t, x)} dx, ds dt \\
&= \sup_{\|v\|_{L_{t,x}^{4/3}}=1} \int_0^T \int_0^t \int_{\mathbb{T}_L^d} |u(s)|^2 u(s) \overline{e^{i(t-s)\Delta_\beta} v_{B_j}(t, x)} dx ds dt \\
&= \sup_{\|v\|_{L_{t,x}^{4/3}}=1} \int_0^T \int_{\mathbb{T}_L^d} |u(s)|^2 u(s) \overline{\tilde{v}_{B_j}(s, x)} ds dx.
\end{aligned}$$

Using equation (5.3), we have for every $\epsilon_0 > 0$,

$$\begin{aligned}
& \int_0^T \int_{\mathbb{T}_L^d} |u(s)|^2 u(s) \overline{\tilde{v}_{B_j}(s, x)} ds dx = \sum_{j_1-j_2+j_3-j=O(1)} \int_0^T \int_{\mathbb{T}_L^d} u_{B_{j_1}} \overline{u_{B_{j_2}}} u_{B_{j_3}} \overline{\tilde{v}_{B_j}} ds dx \\
& \lesssim \sum_{\substack{j_1-j_2+j_3-j=O(1) \\ j_1, j_2, j_3 \in \mathbb{Z}^d}} \prod_{k=1}^3 \|u_{B_{j_k}}\|_{L_{t,x}^4} \|\tilde{v}_{B_j}\|_{L_{t,x}^4} \lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/2} \sum_{\substack{j_1-j_2+j_3-j=O(1) \\ j_1, j_2, j_3 \in \mathbb{Z}^d}} \prod_{k=1}^3 \|u_{B_{j_k}}\|_{L_{t,x}^4},
\end{aligned}$$

and therefore

$$\left\| \int_0^t e^{i(t-s)\Delta_\beta} P_{B_j} |u(s)|^2 u(s) ds \right\|_{L_{t,x}^4([0,T] \times \mathbb{T}^d)} \lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/2} \sum_{\substack{j_1-j_2+j_3-j=O(1) \\ j_1, j_2, j_3 \in \mathbb{Z}^d}} \prod_{k=1}^3 \|u_{B_{j_k}}\|_{L_{t,x}^4}.$$

Consequently, for $s > d/2$, we have

$$\begin{aligned}
& \left(\sum_{j \in \mathbb{Z}^d} \langle j \rangle^{2s} \left\| \int_0^t e^{i(t-s)\Delta_\beta} P_{B_j} |u(s)|^2 u(s) ds \right\|_{L_{t,x}^4}^2 \right)^{1/2} \\
& \lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/2} \left(\sum_{j \in \mathbb{Z}^d} \langle j_1 \rangle^{2s} \|u_{B_{j_1}}\|_{L_{t,x}^4}^2 \right)^{1/2} \left(\sum_{\ell \in \mathbb{Z}^d} \|u_{B_\ell}\|_{L_{t,x}^4} \right)^2 \lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/2} \|u\|_{Z_T^s}^3.
\end{aligned}$$

proving equation (5.5). $\square$

Lemma 5.2 (A priori energy estimates).

$$\left\| \int_0^t e^{i(t-s)\Delta_\beta} |u|^2 u ds \right\|_{L_t^\infty H_x^s} \leq S_* \|u\|_{Z_T^s}^3. \quad (5.6)$$

Proof. By duality, we have

$$\begin{aligned} \left\| \int_0^t e^{i(t-s)\Delta_\beta} |u|^2 u \, ds \right\|_{L_t^\infty H_x^s} &\leq \sup_{\substack{\|\psi\|_{L_x^2}=1 \\ 0 \leq t \leq T}} \int_0^t \int_{\mathbb{T}_L^d} |u|^2 u \, e^{is\Delta_\beta} \langle \nabla \rangle^s \bar{\psi} \, dx \, ds \\ &= \sup_{\substack{\|\psi\|_{L_x^2}=1 \\ 0 \leq t \leq T}} \sum_{j_1-j_3+j_3-j_4=O(1)} \int_0^t \int_{\mathbb{T}_L^d} u_{B_{j_1}} \overline{u_{B_{j_2}}} u_{B_{j_3}} \, e^{is\Delta_\beta} \langle \nabla \rangle^s \overline{\psi_{B_{j_4}}} \, dx \, ds. \end{aligned}$$

Applying the Strichartz estimate (5.1) yields

$$\begin{aligned} &\sum_{j_1-j_3+j_3-j_4=O(1)} \left| \int_0^t \int_{\mathbb{T}_L^d} u_{B_{j_1}} \overline{u_{B_{j_2}}} u_{B_{j_3}} \, e^{is\Delta_\beta} \nabla^s \psi_{B_{j_4}} \, dx \, ds \right| \\ &\lesssim \sum_{j_1-j_2+j_3-j_4=O(1)} \langle j_4 \rangle^s \prod_{k=1}^3 \|u_{B_{j_k}}\|_{L_{t,x}^4} \|e^{is\Delta_\beta} \psi_{B_{j_4}}\|_{L_{t,x}^4} \\ &\lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4} \sum_{j_1-j_2+j_3-j_4=O(1)} \langle \max(|j_1|, |j_2|, |j_3|) \rangle^s \prod_{k=1}^3 \|u_{B_{j_k}}\|_{L_{t,x}^4} \|\psi_{B_{j_4}}\|_{L_x^2} \\ &\lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4} \left(\sum_j \|\psi_{B_j}\|_{L^2}^2 \right)^{1/2} \left(\sum_j \langle j \rangle^{2s} \|u_{B_j}\|_{L^4}^2 \right)^{1/2} \left(\sum_j \|u_{B_j}\|_{L^4} \right)^2 \\ &\lesssim L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4} \|u\|_{Z_T^s}^3 \|\psi\|_{L_x^2}. \end{aligned}$$

This establishes the stated bound. $\square$

5.3. Existence theorem. Local well-posedness for (NLS) will be established in the space Z_T^s , with data f of size at most $\mathcal{J}$,

$$\mathcal{J} := L^{\epsilon_0} (TL^{-d})^{\frac{1}{4}} \geq \|e^{it\Delta_\beta} f\|_{Z_T^s}. \quad (5.7)$$

This seemingly strange normalization is actually well adapted to the problem we are considering. Indeed, consider for simplicity initial data f supported on Fourier frequencies $\lesssim 1$, whose L^2 norm is of size L^{ϵ_0} , and with random Fourier coefficients of uncorrelated phases. Then we expect $e^{it\Delta_\beta} f$ to be evenly spread over $\mathbb{T}_L^d$. By conservation of the L^2 norm, this corresponds to $\|e^{it\Delta_\beta} f\|_{Z_T^s} \sim \mathcal{J}$.

Theorem 5.3. *Let $f \in Z_T^s$ with $\mathcal{J}$ and S_* defined in equations (5.7) and (5.5) respectively, then*

$$\begin{cases} i\partial_t u - \Delta_\beta u = -\lambda^2 |u|^2 u \\ u(0, x) = f(x) \end{cases}$$

is locally well-posed in Z_T^s , provided

$$R \stackrel{\text{def}}{=} 12(\lambda S_* \mathcal{J})^2 \leq \frac{1}{2}. \quad (5.8)$$

The solution $u \in Z_T^s$, satisfies $\|u\|_{Z_T^s} \leq 2\mathcal{J}$. Moreover

$$\|u\|_{L_t^\infty H_x^s([0,T] \times \mathbb{T}_L^d)} \leq \|f\|_{H_x^s} + C\lambda^2 S_* \mathcal{J}^3 = \|f\|_{H_x^s} + C \frac{R}{S_*} \mathcal{J} \leq \|f\|_{H_x^s} + CR. \quad (5.9)$$

Remark 5.4. *The time scale T over which the solution can be constructed would be equal to $\sqrt{\tau}$, up to subpolynomial losses in L , if the long-time Strichartz estimate conjectured in [11] for $p = 4$ could be established. Since it is currently not known to be true, the result stated above gives a shorter time scale, with a more complicated numerology.*

Proof. This theorem is proved by using a contraction mapping argument, to find a fixed point of the map,

$$\Phi(u) = e^{it\Delta_\beta} f + i\lambda^2 \int_0^t e^{i(t-s)\Delta_\beta} |u|^2 u(s) ds,$$

in $\{u \in Z_T^s \mid \|u\|_{Z_T^s} \leq 2\mathcal{J}\}$. Consequently $u = \lim_{N \rightarrow \infty} \Phi^N(0)$, where Φ^N stands for the N -th iterate of Φ :

$$\Phi^0(0) = e^{it\Delta_\beta} f, \quad \Phi^{N+1}(0) = e^{it\Delta_\beta} f + i\lambda^2 \int_0^t e^{i(t-s)\Delta_\beta} |\Phi^N(0)|^2 \Phi^N(0) ds.$$

To check that Φ is a contraction on $B_{Z_T^s}(0, 2\mathcal{J})$, note that by equation (5.5),

$$\|\Phi(u) - \Phi^0(0)\|_{Z_T^s} = \left\| \lambda^2 \int_0^t e^{i(t-s)\Delta_\beta} |u|^2 u(s) ds \right\|_{Z_T^s} \leq \lambda^2 S_*^2 \|u\|_{Z_T^s}^3 \leq 8\lambda^2 S_*^2 \mathcal{J}^3 \leq R\mathcal{J} \leq \frac{1}{2}\mathcal{J}.$$

and thus Φ maps $B_{Z_T^s}(0, 2\mathcal{J})$ into itself. Again, by equation (5.5),

$$\|\Phi(u) - \Phi(v)\|_{Z_T^s} \leq 3\lambda^2 S_*^2 (2\mathcal{J})^2 \|u - v\|_{Z_T^s} \leq R\|u - v\|_{Z_T^s} \leq \frac{1}{2}\|u - v\|_{Z_T^s}.$$

Therefore Φ is a contraction on $\{u \in Z_T^s \mid \|u\|_{Z_T^s} \leq 2\mathcal{J}\}$, and the H^s estimate follows from the a priori energy bound. $\square$

Besides the established bounds on u , we need to investigate the rate of convergence of $\Phi^N(u) \rightarrow u$.

Corollary 5.5. *Under the conditions of Theorem 5.3, there holds*

$$\|u - \Phi^N(0)\|_{L^\infty H^s} \leq C \frac{R^N}{S_*} \mathcal{J} \leq CR^N$$

Proof. Since Φ is a contraction with modulus R , then

$$\|\Phi^j(0) - \Phi^{j-1}(0)\|_{Z_T^s} \leq R^{j-1} \|\Phi^0(0)\|_{Z_T^s}.$$

Moreover the energy estimate (5.6) gives

$$\|\Phi^{j+1}(0) - \Phi^j(0)\|_{L_T^\infty H^s} \leq C \frac{R}{S_*} \|\Phi^j(0) - \Phi^{j-1}(0)\|_{Z_T^s}.$$

Consequently by writing $u - \Phi^N(0) = \sum_{j=N}^{\infty} \Phi^{j+1}(0) - \Phi^j(0)$, we bound

$$\begin{aligned} \|u - \Phi^N(0)\|_{L_T^\infty H^s} &\leq \sum_{j=N}^{\infty} \|\Phi^{j+1}(0) - \Phi^j(0)\|_{L_T^\infty H^s} \leq C \frac{R}{S_*} \sum_{j=N}^{\infty} \|\Phi^j(0) - \Phi^{j-1}(0)\|_{Z_T^s} \\ &\leq C \frac{R}{S_*} \sum_{j=N}^{\infty} R^{j-1} \|\Phi^0(0)\|_{Z_T^s} \leq C \frac{R^N}{S_*} \|\Phi^0(0)\|_{Z_T^s}. \end{aligned}$$

□

Next we establish an energy bound for the Feynman trees,

$$U_{n,\ell}(t) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} J_{n,\ell}(t, k) e^{2\pi i(k \cdot x + tQ(k))}.$$

Corollary 5.6. *Under the conditions of Theorem 5.3,*

$$\|U_{n,\ell}\|_{L_T^\infty H^s} \leq C \frac{R^N}{S_*} \mathcal{J} \leq CR^N$$

Proof. Since $U_{n,\ell}$ is the linear propagator of $\mathcal{J}_{n,\ell}$ in physical space, then they can be represented by the following iterative procedure: Set $v_0^m = e^{2\pi i t \Delta_\beta} u_0$ for $0 \leq m \leq 2n+1$ and for any $1 \leq j \leq n$ we define v_m^j for $0 \leq m \leq 2(n-j)+1$ as $v_j^m = v_{j-1}^m$ if $m < \ell_j$, and $v_j^m = v_{j-1}^{m+2}$ if $m > \ell_j$, where we set

$$v^{\ell_j} = i\lambda^2 \int_0^t e^{i(t-s)\Delta_\beta} v_{j-1}^{\ell_j} \overline{v_{j-1}^{\ell_j+1}} v_{j-1}^{\ell_j+2} ds.$$

Hence we have $U_{n,\ell} = v_n^1$.

Using the energy estimate (5.6), we bound

$$\|v_n^1\|_{L_T^\infty H^s} \leq \lambda^2 S_* \|v_{n-1}^1\|_{Z_T^s} \|v_{n-1}^2\|_{Z_T^s} \|v_{n-1}^3\|_{Z_T^s}.$$

We can then descend down the tree by estimating $v_{n-j}^{\ell_{n-j}}$ using the Z^s estimate (5.5). This leads to the stated bound. □

6. IMPROVED INTEGRABILITY THROUGH RANDOMIZATION

Recall that

$$u_0 = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \sqrt{\phi(k)} e^{2\pi i k \cdot x} e^{2\pi i \vartheta_k(\omega)},$$

where the $\vartheta_k(\omega)$ are independent random variables, uniformly distributed on $[0, 2\pi]$.

For any t, s, ω , we have

$$\|e^{it\Delta_\beta} u_0\|_{H^s} = \left[\frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \langle k \rangle^{2s} \phi(k) \right]^{1/2}.$$

In other words, the randomization of the angles of the Fourier coefficients does not have any effect on L^2 based norms. This is not the case for Lebesgue indices larger than 2.

Theorem 6.1. *Assume that $|\phi(k)| \lesssim \langle k \rangle^{-s}$, with $s > \frac{d}{2}$. Then*

$$(i) \quad \mathbb{E} \|e^{it\Delta_\beta} u_0\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d)}^4 \lesssim \frac{T}{L^d} \|u_0\|_{L_x^2}^4$$

(ii) *(large deviation estimate)*

$$\mathbb{P} \left[\|e^{it\Delta_\beta} u_0\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d)}^4 > \lambda \right] \lesssim \exp \left(-c \left(\frac{\lambda}{T^{1/4} L^{-d/4}} \right)^2 \right)$$

Proof. (i) The proof is more or less standard. See [7] for instance.

(ii) We follow the argument in [7]. By Minkowski's inequality (for $p \geq 4$) and Khinchin's inequality,

$$\begin{aligned} \|e^{it\Delta_\beta} u_0\|_{L_\omega^p(\Omega, L_{t,x}^4([0,T] \times \mathbb{T}_L^d))} &\lesssim \|e^{it\Delta_\beta} u_0\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d, L_\omega^p(\Omega))} \\ &\lesssim \frac{\sqrt{p}}{L^d} \left\| \left(\sum \phi(k) \right)^{1/2} \right\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d)} \lesssim \sqrt{p} T^{1/4} L^{-d/4}. \end{aligned}$$

By Chebyshev's inequality,

$$\mathbb{P} \left[\|e^{it\Delta_\beta} u_0\|_{L_{t,x}^p([0,T] \times \mathbb{T}_L^d)} > \lambda \right] \lesssim \lambda^{-p} (C_0 \sqrt{p} T^{1/4} L^{-d/4})^p.$$

The desired inequality is then obvious if $\lambda < 2eC_0 T^{1/4} L^{-d/4}$; if not, it follows upon choosing $p = \left(\frac{\lambda}{C_0 T^{1/4} L^{-d/4} e} \right)^2$. $\square$

As a consequence, we deduce the following proposition.

Proposition 6.2. *Let $\epsilon_0 > 0$, $\alpha > s + \frac{d}{2}$, and assume that $|\phi(k)| \lesssim \langle k \rangle^{-2\alpha}$. Then, for two constant $C, c > 0$,*

$$\mathbb{P} \left[\|e^{it\Delta_\beta} u_0\|_{Z^s} < T^{1/4} L^{\epsilon_0 - d/4} \right] > 1 - C e^{-cL^{\epsilon_0}}.$$

Proof. Applying Theorem 6.1 to $(u_0)_{B_j}$,

$$\mathbb{P} \left[\|e^{it\Delta_\beta} u_0\|_{L_{t,x}^4([0,T] \times \mathbb{T}_L^d)} > \langle j \rangle^{-\alpha} T^{1/4} L^{\frac{\epsilon_0}{2} - \frac{d}{4}} \right] \lesssim \exp(-c\langle j \rangle^{2\alpha} L^{\epsilon_0}).$$

Therefore, for L sufficiently large,

$$\begin{aligned} \mathbb{P} \left[\|e^{it\Delta_\beta} u_0\|_{Z^s} < T^{1/4} L^{\epsilon_0 - d/4} \right] &> 1 - \sum_j \mathbb{P} \left[\|e^{it\Delta_\beta} (u_0)_{B_j}\|_{Z^s} > T^{1/4} L^{\frac{\epsilon_0}{2} - d/4} \langle j \rangle^{-\alpha} \right] \\ &> 1 - C \sum_j \exp(-c\langle j \rangle^{2\alpha} L^{\epsilon_0}) \\ &> 1 - C e^{-cL^{\epsilon_0}}. \end{aligned}$$

$\square$

7. PROOF OF THE MAIN THEOREM

Fix $\epsilon_0 > 0$ sufficiently small, and recall that $T \leq L^d$, with

$$S_* = C_{4,\epsilon_0} L^{\epsilon_0} \left\langle \frac{T}{L^{\theta_d}} \right\rangle^{1/4}, \quad \mathcal{J} = L^{\epsilon_0} (T L^{-d})^{\frac{1}{4}}, \quad \text{and } R \stackrel{\text{def}}{=} 12(\lambda S_* \mathcal{J})^2.$$

- 1) *Excluding exceptional data.* Let $E_{\epsilon_0, L}$ be the event $\{\|e^{it\Delta_\beta} u_0\|_{Z^s} \leq \mathcal{J}\}$, and $F_{\epsilon_0, L}$ its contrary: $\{\|e^{it\Delta_\beta} u_0\|_{Z^s} > \mathcal{J}\}$. By Proposition 6.2,

$$\mathbb{P}(F_{\epsilon_0, L}) \lesssim e^{-cL^{\epsilon_0}}.$$

This is the set appearing in the statement of Theorem 2.2. By conservation of mass

$$\mathbb{E}(|a_k(t)|^2) = \mathbb{E}(|a_k(t)|^2 \mid E_{\epsilon_0, L}) + O_{\ell^\infty}(e^{-cL^{\epsilon_0}} L^d).$$

- 2) *Iterative resolution.* To ensure that $R \leq \frac{1}{2}$ we restrict the range of the parameters λ, T relative to L . There are two regimes depending on the Strichartz constant S_* and the number theory restriction $t \leq L^{d-\epsilon_0}$ (see Remark 8.2).

- $L^{\theta_d} \lesssim T \lesssim L^d$. The condition $R \leq \frac{1}{2}$ translates into $T \sim \lambda^{-2} L^{\frac{d+\theta_d}{2}-4\epsilon_0}$. Therefore we restrict λ to

$$L^{\frac{-d+\theta_d-8\epsilon_0}{4}} \lesssim \lambda \lesssim L^{\frac{d-\theta_d-8\epsilon_0}{4}}.$$

For this range of parameters, the energy inequality (5.9) implies $\|u\|_{L_t^\infty H_x^s([0,T] \times \mathbb{T}_L^d)} \lesssim 1$.

- $T \lesssim L^{\theta_d}$. In this case the condition on R restricts $T \sim \min(L^{\theta_d}, \lambda^{-4} L^{d-8\epsilon_0})$, and therefore

$$L^{\frac{d-\theta_d-8\epsilon_0}{4}} \lesssim \lambda.$$

Here the energy inequality also implies $\|u\|_{L_t^\infty H_x^s([0,T] \times \mathbb{T}_L^d)} \lesssim 1$.

Note that for these ranges of parameters $T \leq L^{-2\delta} \sqrt{\tau}$, where δ is that of Theorem 8.1.

With these restrictions on the range of the parameters we proceed by writing $u = \Phi^N(0) + u - \Phi^N(0)$. Note that since $\Phi^N(0)$ is a polynomial of degree 3^N , we write

$$u = \sum_{n=0}^N U_{n,\ell} + \sum_{(n,\ell) \in S^N} U_{n,\ell} + u - \Phi^N(0),$$

where S^N includes all the terms in $\Phi^N(0)$ of degree greater than N .

By Corollary 5.5 and Proposition 5.6, this implies that

$$u = \sum_{n=1}^N \sum_{\ell} U_{n,\ell} + O_{L_t^\infty H_x^s}(R^N)$$

where the constant depends on N . In terms of Fourier variables this can be written as,

$$|a_k(t)|^2 = \left| \sum_{n=1}^N \sum_{\ell} J_{n,\ell} \right|^2 + O_{\ell_L^{1,2s}}(R^N) = \left| \sum_{n=1}^N \sum_{\ell} J_{n,\ell} \right|^2 + O_{\ell^\infty}(L^d R^N).$$

3) *Pairing.* By Proposition 4.1,

$$\begin{aligned} \left| \sum_{n=1}^N \sum_{\ell} J_{n,\ell} \right|^2 &= \mathbb{E} \left[|J_1(k)|^2 + J_0(k) \overline{J_2(k)} + \overline{J_0(k)} J_2(k) \right] + O\left(\frac{t}{\tau} \frac{t \log t}{\sqrt{\tau}}\right) \\ &= \phi_k + \frac{2\lambda^4}{L^{4d}} \sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \times \\ &\quad \left| \frac{\sin(t\pi\Omega(k, k_1, k_2, k_3))}{\pi\Omega(k, k_1, k_2, k_3)} \right|^2 + O_{\ell^\infty}\left(\frac{t}{\tau} \frac{t \log t}{\sqrt{\tau}}\right) \end{aligned}$$

4) *Large box limit* $L \rightarrow \infty$. By the equidistribution theorem 8.1, we have for $t < L^{d-\epsilon}$

$$\begin{aligned} \frac{2\lambda^4}{L^{4d}} \sum_{k-k_1+k_2-k_3=0} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] &\left| \frac{\sin(t\pi\Omega(k, k_1, k_2, k_3))}{\pi\Omega(k, k_1, k_2, k_3)} \right|^2 \\ &= \frac{2\lambda^4}{L^{2d}} \int \delta(\Sigma) \phi(k) \phi(k_1) \phi(k_2) \phi(k_3) \left[\frac{1}{\phi(k)} - \frac{1}{\phi(k_1)} + \frac{1}{\phi(k_2)} - \frac{1}{\phi(k_3)} \right] \times \\ &\quad \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi\Omega(k, k_1, k_2, k_3)} \right|^2 dk_1 dk_2 dk_3 + O_{\ell^\infty}\left(\frac{t}{\tau} L^{-\delta}\right). \end{aligned}$$

5) *Large time limit* $t \sim T \rightarrow \infty$. Since for a smooth function f ,

$$\int \left| \frac{\sin(\pi t x)}{x} \right|^2 f(x) dx = \pi^2 t f(0) + O(1),$$

then, with $\tau = \frac{L^{2d}}{2\lambda^4}$, we have

$$\begin{aligned} \frac{2\lambda^4}{L^{2d}} \int \delta(\Sigma) \phi(k) \phi(k_1) \phi(k_2) \phi(k_3) \left[\frac{1}{\phi(k)} - \frac{1}{\phi(k_1)} + \frac{1}{\phi(k_2)} - \frac{1}{\phi(k_3)} \right] \times \\ \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 dk_1 dk_2 dk_3 = \frac{t}{\tau} \mathcal{T}(\phi, \phi, \phi) + O\left(\frac{1}{\tau}\right). \end{aligned}$$

Consequently, for ϵ_0 sufficiently small and $t \leq T \leq L^{d-\epsilon_0}$, we choose $L \geq L_1(\epsilon_0)$ to bound the error term in Step 1 by $\frac{t}{\tau} L^{-\epsilon_0}$. Also, since $R \leq \frac{1}{2}$ then by picking N large enough we can bound the error in Step 2 by $O(\frac{t}{\tau} L^{-\epsilon_0})$. Similarly, since $t \log t \leq L^{-\delta} \sqrt{\tau}$, then the error for Steps 3, 4, and 5, are of order $O_{\ell^\infty}(\frac{t}{\tau} L^{-\delta})$, and this concludes the proof of Theorem 2.2.

8. NUMBER THEORETIC RESULTS

Our aim in this section is to prove the asymptotic formula for the following Riemann sum,

Theorem 8.1. *Given $\phi \in \mathcal{S}(\mathbb{R}^d)$ and $\epsilon > 0$, there exists a $\delta > 0$ such that if $0 < t \leq L^{d-\epsilon}$, then*

$$\begin{aligned} \sum_{\substack{k_i \in \mathbb{Z}_L^d \\ k-k_1+k_2-k_3=0}} \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 = \\ L^{2d} \int \delta(\Sigma) \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right] \left| \frac{\sin(\pi t \Omega(k, k_1, k_2, k_3))}{\pi \Omega(k, k_1, k_2, k_3)} \right|^2 dk_1 dk_2 dk_3 \\ + O\left(t L^{2d-\delta}\right) + O\left(L^d\right), \end{aligned}$$

where we recall $\Sigma(k, k_1, k_2, k_3) = k - k_1 + k_2 - k_3$.

The difficulty in proving this theorem is that Ω can be very small, while the stated time interval for the validity of the asymptotic formula is very large. In fact if we restrict ourselves to a timescale which is not too long, then the asymptotic formula is straight forward as will be demonstrated in Proposition 8.10. However to prove this theorem as stated we need to generalize a result of Bourgain on pair correlations of generic quadratic forms [5].

Bourgain considered a positive definite diagonal form,

$$Q(n) = \sum_{i=1}^d \beta_i n_i^2, \quad n = (n_1, \dots, n_d), \quad Q(p, q) := Q(p) - Q(q), \quad (8.1)$$

for generic $\beta = (\beta_1, \dots, \beta_d) \in [1, 2]^d$, and proved that for $d = 3$ the lattice points in the region,

$$R_{\mathbb{Z}} \stackrel{\text{def}}{=} \{(p, q) \in \mathbb{Z}^{2d} \cap [0, L]^{2d} \mid Q(p, q) \in [a, b], p \neq q\},$$

are equidistributed at a scale of $\frac{1}{L^\rho}$, for $0 < \rho < d - 1$. Specifically, he proved,

$$\sum_{R_{\mathbb{Z}}} 1 = L^{2(d-1)}(b-a) \mathcal{H}^{2d-1} \left(\{(x, y) \in [-1, 1]^{2d} \mid Q(x, y) = 0\} \right) + O\left(L^{d-2-\delta}(b-a)\right),$$

provided $|a|, |b| < O(1)$ and $L^{-\rho} < b - a < 1$. Here $\mathcal{H}^{2d-1}$ is the $2d - 1$ Hausdorff measure.

Our quadratic form Ω , restricted to Σ , can be transformed to $Q(p, q)$, given in (8.1), as follows. Rescale time $\mu := tL^{-2}$, let $K_i = Lk_i \in \mathbb{Z}$, and denote by

$$g(x) = \left(\frac{\sin(\pi x)}{\pi x} \right)^2, \quad W_0 \left(\frac{K}{L}, \frac{K_1}{L}, \frac{K_2}{L}, \frac{K_3}{L} \right) = \phi_k \phi_{k_1} \phi_{k_2} \phi_{k_3} \left[\frac{1}{\phi_k} - \frac{1}{\phi_{k_1}} + \frac{1}{\phi_{k_2}} - \frac{1}{\phi_{k_3}} \right].$$

Then the sum can be expressed as

$$t^2 \sum_{\substack{K, K_1, K_2, K_3 \in \mathbb{Z}^d \\ K - K_1 + K_2 - K_3 = 0}} W_0 \left(\frac{K}{L}, \frac{K_1}{L}, \frac{K_2}{L}, \frac{K_3}{L} \right) g(\mu \Omega(K, K_1, K_2, K_3)).$$

By defining

$$u' = K_1 - K \in \mathbb{Z}^d, \quad u'' = K_3 - K \in \mathbb{Z}^d, \quad \text{and } u = (u', u'') \in \mathbb{Z}^{2d}$$

then

$$\Omega(K, K_1, K_2, K_3) = Q_0(u)$$

where

$$Q_0(u) := -2\beta_1 u'_1 u''_1 - 2\beta_2 u'_2 u''_2 - \cdots - 2\beta_d u'_d u''_d. \quad (8.2)$$

Hence the sum can be expressed as

$$t^2 \sum_{(u'_i, u''_i) \in \mathbb{Z}^2} W_0 \left(\frac{K}{L}, \frac{u' + K}{L}, \frac{u' + u'' + K}{L}, \frac{u'' + K}{L} \right) g(\mu Q_0(u)). \quad (8.3)$$

The quadratic form Q_0 can be diagonalized by making the change of coordinates

$$p_i = u'_i + u''_i, \quad q_i = u'_i - u''_i$$

where p_i and q_i are either both even or both odd, i.e.

$$\sum_{u_i \in \mathbb{Z}^2} = \sum_{p_i, q_i \in 2\mathbb{Z}} + \sum_{p_i, q_i \in (2\mathbb{Z}+1)} = \sum_{p_i, q_i \in \mathbb{Z}} - \sum_{p_i \in 2\mathbb{Z}, q_i \in \mathbb{Z}} - \sum_{p_i \in \mathbb{Z}, q_i \in 2\mathbb{Z}} + 2 \sum_{p_i, q_i \in 2\mathbb{Z}}.$$

Consequently, the sum (8.3), can be written as four different sums of the form,

$$t^2 \sum_{(p, q) \in \mathbb{Z}^{2d}} W \left(\frac{p}{L}, \frac{q}{L} \right) g(\mu Q(p, q)), \quad (8.4)$$

where $Q(p, q)$ is given by³ (8.1), and where we suppressed the dependence of W on k for convenience.

Remark 8.2. Note that we do not exclude the points when $p_i^2 = q_i^2$ for all $i \in [1, \dots, n]$, as Bourgain did. These points contribute $O(L^d)$ to the sum and will be considered as lower order terms. They also explain the $O(L^d)$ term in Theorem 8.1.

It is this fact that prevents us from using the full strength of our equidistribution result which holds for $\mu = tL^{-2} \leq L^{d-1-\epsilon}$, and we use the result for $t \leq L^{d-\epsilon}$. This ensures that $O(L^d)$ term is an error in the asymptotic formula.

To prove the asymptotic formula given in Theorem 8.1, with $0 < \mu = tL^{-2} \leq L^{d-1-\epsilon}$, we proceed as follows: 1) identify which part of the sum contributes the leading order term and which part contributes error terms; 2) prove equidistribution of lattice points on a coarse scale; 3) present Bourgain's theorem on equidistribution on a fine scale; and finally 4) prove Theorem 8.1.

³There are factors of 2 missing due to sums over even terms. However, this has no impact since β is generic.

8.1. Identifying main terms vs error terms. To identify the leading order term in the equidistribution formula, we first obtain upper bounds on lattice sums that are optimal up to sub-polynomial factor.

For generic $\beta = (\beta_1, \dots, \beta_d) \in [1, 2]^d$, a good upper bound for the linear form $\beta \cdot n \in [a, b]$, where $n = \mathbb{Z}^d$ is a consequence of the pigeonhole principle:

Lemma 8.3. *The linear form $\beta \cdot n \in [a, b]$ satisfies the following bound*

$$\#\{n \in \mathbb{Z}^d \cap [-M, M]^d \mid a \leq \beta \cdot n \leq b\} = \sum_{\substack{a \leq \beta \cdot n \leq b \\ |n| \leq M}} 1 \lesssim M^{(d-1)^+} (b-a) + 1 \quad (8.5)$$

Proof. Since $\beta = (\beta_1, \dots, \beta_d)$ are generic, then for $0 < |n| \leq M$ (see for example [8], Chapter VII)

$$|\beta \cdot n| \gtrsim \frac{1}{M^{(d-1)^+}}.$$

For arbitrary $n^{(1)} \neq n^{(2)} \in \mathbb{Z}^d$ satisfying $a \leq \beta \cdot n^{(i)} \leq b$ and $0 < |n^{(i)}| \leq M$,

$$\frac{1}{M^{(d-1)^+}} \lesssim \left| \beta \cdot (n^{(1)} - n^{(2)}) \right| \leq b - a.$$

By the pigeonhole principle we obtain (8.5). $\square$

An upper bound on the cardinality of the set,

$$R_{\mathbb{Z}} \stackrel{\text{def}}{=} \{(p, q) \in \mathbb{Z}^{2d} \cap [0, L]^{2d} \mid Q(p, q) \in [a, b], p \neq q\},$$

can be obtained by bounding the number of lattice points in subsets of the form,

$$R_{\mathbb{Z}\ell} = \{(p, q) \in \mathbb{Z}^{2d} \cap [0, L]^{2d} \mid Q(p, q) \in [a, b], p_i \neq q_i, 1 \leq i \leq \ell, \text{ and } p_i = q_i, \ell + 1 \leq i \leq d\},$$

using Lemma 8.3, and by using the divisor bound $d(k) \lesssim_\epsilon k^\epsilon$.

Lemma 8.4. *For $\ell = 1, \dots, d$ the cardinality of $R_{\mathbb{Z}\ell}$ satisfies the bound*

$$\#R_{\mathbb{Z}\ell} = \sum_{R_{\mathbb{Z}\ell}} 1 \lesssim L^{(d+\ell-2)^+} (b-a) + L^{(d-\ell)^+} \quad (8.6)$$

Proof. Define $k_i = (p_i - q_i)(p_i + q_i)$, for $1 \leq i \leq \ell$. Since $p_i = q_i$, for $\ell + 1 \leq i \leq d$, we conclude

$$\#R_{\mathbb{Z}\ell} \lesssim L^{d-\ell} \sum_{\substack{a \leq \sum_{i=1}^{\ell} \beta_i k_i \leq b \\ 0 < |k| \lesssim L^2}} \left(\sum_{(p_i - q_i)(p_i + q_i) = k_i} 1 \right)$$

By the divisor bound

$$\sum_{(p_i - q_i)(p_i + q_i) = k_i} 1 \lesssim L^{0^+},$$

and by (8.5), with $M = L^2$, we obtain

$$\#R_{\mathbb{Z}\ell} \lesssim L^{(d-\ell)^+} \left(L^{2(\ell-1)^+} (b-a) + 1 \right),$$

and (8.6) follows. $\square$

Corollary 8.5. *The number of elements in $R_{\mathbb{Z}}$, can be bounded by*

$$\#R_{\mathbb{Z}} \lesssim L^{2(d-1)^+}(b-a) + L^{(d-1)^+} \quad (8.7)$$

Moreover, if we further assume $|a|, |b| \leq 1$, then we have the improved bound

$$\#R_{\mathbb{Z}} \lesssim L^{2(d-1)^+}(b-a) + L^{(d-2)^+}, \quad (8.8)$$

Proof. It suffices to apply the Lemma 8.4, and to observe that $\ell \in \{1, \dots, d\}$ since $p = q$ is excluded. (8.8) follows from noting that if $|a|, |b| \leq 1$, then $R_{\mathbb{Z}1}$ is empty. $\square$

Remark 8.6. *Note, that in terms of the first estimate (8.7), the second term may be treated as an error as long as $b-a \geq L^{-(d-1)+\epsilon_0}$ for some $\epsilon_0 > 0$. Analogously, the second term of (8.8) may be treated as an error assuming $b-a \geq L^{-d+\epsilon_0}$.*

Following this remark on identifying the leading order term, we can now identify subsets of $R_{\mathbb{Z}}$ that contribute error terms only. The first such subsets are when $|p_i - q_i| \lesssim L^{1-\delta}$ for some fixed $\delta > 0$ and some i that we may without loss of generality assume to be 1.

Lemma 8.7. *For $|a|, |b| \leq 1$, the number of elements in $R_{\mathbb{Z}}$ satisfying $|p_1 - q_1| \lesssim L^{1-\delta}$ satisfy the following bound*

$$\#R_{\mathbb{Z}} \cap \{(p, q) \in \mathbb{Z}^{2d} \mid |p_1 - q_1| \lesssim L^{1-\delta}\} \lesssim L^{2(d-1)^+-\delta}(b-a) + L^{(d-1)^+}.$$

Proof. If $p_i = q_i$ for at least one i , then by Corollary 8.5 with d replaced by $d-1$, we have

$$\#R_{\mathbb{Z}} \cap \{(p, q) \in \mathbb{Z}^{2d} \mid p_i = q_i\} \lesssim L \left(L^{2(d-2)^+}(b-a) + L^{(d-3)^+} \right),$$

which is lower order. Moreover, if $p_i \neq q_i$ for all i , and $|p_1 - q_1| \lesssim L^{1-\delta}$, then the sum over $2 \leq i \leq d$ can be bounded by $L^{2(d-2)^+}(b-a) + L^{0+}$, using Lemma 8.4, while the sum over p_1 and q_1 can be by $L^{2-\delta}$. This gives a bound of $L^{2-\delta} \left(L^{2(d-2)^+}(b-a) + L^{0+} \right)$, which is lower order if $d \geq 3$. $\square$

Next we show that if one p_i or q_i is less than $L^{1-\delta}$, where we may again assume $i = 1$, then the contribution to the number of elements in $R_{\mathbb{Z}}$ is lower order.

Lemma 8.8. *For $|a|, |b| \leq 1$, we have the following estimate*

$$\#R_{\mathbb{Z}} \cap \{(p, q) \in \mathbb{Z}^{2d} \mid |p_1| \lesssim L^{1-\delta}\} \lesssim L^{2(d-1)^+-\delta}(b-a) + L^{(d-1)^+}$$

Proof. If both $|p_1| \lesssim L^{1-\delta}$ and $|q_1| \lesssim L^{1-\delta}$ or $p_i = q_i$ for at least one i , then by Lemma 8.7 we have the stated bound. Otherwise, the sum over $2 \leq i \leq d$ contributes $L^{2(d-2)^+}(b-a) + L^{0+}$, while the sum over p_1 and q_1 contributes $L^{2-\delta}$. $\square$

From Lemma 8.7 and Lemma 8.8, we have

Corollary 8.9. *Setting*

$$R_{\mathbb{Z}\delta} = R_{\mathbb{Z}} \setminus \bigcup_{i=1}^d \left(\{(p, q) \in \mathbb{Z}^{2d} \mid \text{where, } |p_i|, |q_i|, \text{ or } |p_i - q_i| \lesssim L^{1-\delta}, \text{ for at least one } i\} \right).$$

Then, for $|a|, |b| \leq 1$, we have the following cardinality bound on the set difference $R_{\mathbb{Z}} \setminus R_{\mathbb{Z}\delta}$

$$\#R_{\mathbb{Z}} \setminus R_{\mathbb{Z}\delta} \lesssim L^{2(d-1)^+-\delta}(b-a) + L^{(d-1)^+}$$

8.2. Asymptotic formula on a coarse scale. These upper bounds, in particular Corollary 8.5 allow us to present a simple proof of the asymptotic formula for $\#R_{\mathbb{Z}}$ on a coarser scale, e.g. $b - a = L^{\frac{4}{3}}$. Note that this is still better than the trivial Riemann sum scale of $b - a = O(L^2)$.

Proposition 8.10. *Fix $\delta > 0$ sufficiently small, then if $L^{1+4\delta} \leq b - a \leq L^{2-\delta}$, we have the asymptotic formula*

$$\begin{aligned} \# \left\{ (p, q) \in \mathbb{Z}^d \cap [0, L]^{2d} \mid Q(p, q) \in [a, b] \right\} &= L^{2(d-1)}(b - a) \iint_{\mathbb{R}^{2d}} \mathbf{1}_{[0,1]^{2d}}(x, y) \delta(Q(x, y)) dx dy \\ &+ O\left(L^{2(d-1)-\delta}(b - a)\right). \end{aligned}$$

Proof. First we will smooth the characteristic functions by extending the region to a slightly bigger region with a controlled error term. This is done as follows. Let $w_L \in C_c^\infty([-L^\delta, L + L^\delta])$ be a bump function satisfying $w_L(x) = 1$ for $x \in [0, L]$ and

$$\|w_L\|_{C^N} \lesssim L^{-N\delta}.$$

Then by setting $W_L(x, y) = \prod_{i=1}^d w_L(Lx_i)w_L(Ly_i)$, we have,

$$\sum_{p, q \in \mathbb{Z}^d} W_L\left(\frac{p}{L}, \frac{q}{L}\right) - \mathbf{1}_{[0, L]^{2d}}(p, q) = O\left(L^{2d-1+\delta}\right).$$

Moreover, if we denote by $h_L \in C_c^\infty([a - L^{1+2\delta}, b + L^{1+2\delta}])$ a bump function $h_L(x) = 1$ for $x \in [a, b]$ and

$$\|h_L\|_{C^N} \lesssim L^{-N(1+2\delta)}.$$

then by Corollary 8.5, we have

$$\begin{aligned} \sum_{p, q \in \mathbb{Z}^d} W_L\left(\frac{p}{L}, \frac{q}{L}\right) h_L(Q(p, q)) - \mathbf{1}_{[0, L]^{2d}}(p, q) \mathbf{1}_{[a, b]}(Q(p, q)) &= \\ O(L^{2d-1+\delta}) + O(L^{(2d-1+2\delta)^+}) &= O\left(L^{2(d-1)-\delta}(b - a)\right). \end{aligned}$$

assuming that $b - a \geq L^{1+4\delta}$. Thus, it is sufficient to obtain the asymptotic formula for

$$\mathcal{S} := \sum_{p, q \in \mathbb{Z}^d} W_L\left(\frac{p}{L}, \frac{q}{L}\right) h_L(Q(p, q)).$$

Using Fourier transform, we express $\mathcal{S}$ as

$$\mathcal{S} = \int_{-\infty}^{\infty} \widehat{h}_L(s) \sum_{p, q} W_L\left(\frac{p}{L}, \frac{q}{L}\right) e(Q(p, q)s) ds := \int_{-\infty}^{\infty} \widehat{h}_L(s) S(s) ds \quad (8.9)$$

Applying Poisson summation we may rewrite $S(s)$ as

$$S(s) = \sum_{\ell} \int W_L\left(\frac{x}{L}, \frac{y}{L}\right) e(Q(x, y)s - m \cdot x - n \cdot y) dx dy \quad (8.10)$$

$$= L^{2d} \sum_{\ell} \int W_L(z) e(L^2 Q(z)s - L\ell \cdot z) dz \quad (8.11)$$

where $z = (x, y)$, and $\ell = (m, n)$.

The term $\ell = 0$ contributes the asymptotic formula

$$L^{2d} \int W_L(z) h_L(L^2 Q(z)) dz = L^{2(d-1)}(b-a) \int_{\mathbb{R}^{2d}} \mathbb{1}_{[0,1]^{2d}}(z) \delta(Q(z)) dz + O\left(L^{2(d-1)-\delta}(b-a)\right)$$

where we used $(b-a) < L^{2-\delta}$ in replacing $h_L(L^2 Q)$ by $\delta_{\text{dirac}}(Q)$. So it remains to show that the sum for $\ell \neq 0$ can be treated as error. First we estimate the sum for $s \leq \frac{1}{L^{1+\delta}}$. In this case we write $\Phi(z, \ell, s) = L^2 Q(z)s - L\ell \cdot z$, and note that since $|s| \leq \frac{1}{L^{1+\delta}}$ and $|z| \lesssim 1$, then $|\nabla_z \Phi(z, \ell, s)| \geq \frac{L|\ell|}{2}$, where

$$\nabla_z \Phi(z, \ell, s) = L^2 \nabla Q(z)s - L\ell, \quad (8.12)$$

and thus upon integrating (8.11) by parts, we obtain

$$S(s) = \sum_{\ell \neq 0} L^{2d} \int \nabla_j \left(\frac{W_L(z)}{2\pi i \nabla_j \Phi(z, \ell, s)} \right) e(\Phi(z, \ell, s)) dz. \quad (8.13)$$

Since each derivative of W_L contributes $L^{1-\delta}$, then each integration by parts contributes a factor of $\frac{1}{L^\delta |\ell|}$. Applying a sufficient number of integrations by parts, and using the fact that $|\widehat{h}_L(s)| \lesssim b-a$, we may ensure that the contribution for $\ell \neq 0$ and $|s| \leq \frac{1}{L^{1+\delta}}$ is arbitrarily small.

For $|s| \geq \frac{1}{L^{1+\delta}}$ we note that

$$|\widehat{h}_L(s)| \lesssim (b-a) \frac{1}{(L^{1+2\delta}|s|)^N},$$

for all N , and thus this term can be treated as an error. This concludes the stated result. $\square$

8.3. Bourgain's Theorem. Now we present Bourgain's proof of equidistribution.

Theorem 8.11. *Fix $\epsilon > 0$, then for $\delta > 0$ sufficiently small the following statement is true: Suppose $I_j, J_j \subset [0, L]$, $j = 1, \dots, d$ for $d \geq 3$ are intervals with length satisfying*

$$L^{1-\delta} \leq |I_j|, |J_j| \leq L \quad (8.14)$$

Then for a, b satisfying $|a|, |b| \leq 1$ and $L^{-d+1+\epsilon} < b-a < L^{-\epsilon}$ we have

$$\sum_{\substack{a \leq Q(p,q) \leq b \\ p_j \in I_j, q_j \in J_j \\ p \neq q}} 1 = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy + O(L^{2(d-1-d\delta)}(b-a)). \quad (8.15)$$

In order to prove Theorem 8.11, we first make a series of reductions.

Step 1: Restrict to dyadic lengths and discrete intervals (a, b) . We first show that it is sufficient to assume dyadic lengths $L = 2^{N_1}$ for $N_1 \in \mathbb{N}$ and that $(a, b) = (N_2 L^{-d+1+\epsilon}, (N_2 + 1) L^{-d+1+\epsilon})$, for $N_2 \in \mathbb{Z}$ such that $|N_2| \leq 2L^{d-1-\epsilon}$. The restriction to dyadic lengths $L = 2^{N_1}$ is valid since it only has potential effect of modifying the implicit constants in the theorem. Now suppose (8.15) is satisfied for all such L and (a, b) as described above and suppose we are given another interval (a', b') such that a', b' satisfies $|a'|, |b'| \leq 1$ and $L^{-d+1+2\epsilon} < b' - a' < L^{-\epsilon}$. Then, by assuming δ is sufficiently small (depending on ϵ), and summing over intervals of the form $(N_2 L^{-d+1+\epsilon}, (N_2 + 1) L^{-d+1+\epsilon})$ we obtain

$$\sum_{\substack{a' \leq Q(p,q) \leq b' \\ p_j \in I_j, q_j \in J_j \\ p \neq q}} 1 = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbb{1}_{a' \leq Q(x,y) \leq b'} dx dy + O(L^{2(d-1)-\epsilon}(b' - a')).$$

Thus, by again taking δ smaller if needed, we obtain Theorem 8.11 with ε replaced by 2ε , i.e. up to a relabeling of ε , we obtain Theorem 8.11.

Step 2: Ignore intervals that contribute lower order sums. Set $\tilde{\delta} = 4d\delta$, then by Corollary 8.9 we have for $\tilde{\delta}$ sufficiently small,

$$\sum_{R_{\mathbb{Z}}} 1 = \sum_{R_{\mathbb{Z}\tilde{\delta}}} 1 + O\left(L^{2(d-1)^+ - \tilde{\delta}}(b-a)\right) + O(L^{(d-2)^+}) = \sum_{R_{\mathbb{Z}\tilde{\delta}}} 1 + O\left(L^{2(d-1)^+ - \tilde{\delta}}(b-a)\right) \quad (8.16)$$

where we have used the restriction of $a-b$ and assumed δ to be sufficiently small compared to ε .

Thus we restrict our attention to the case where

- (a) $\forall p_i \in E_i$, and $\forall q_i \in F_i$, we have $|p_i| > L^{1-\tilde{\delta}}, |q_i| > L^{1-\tilde{\delta}}$,
- (b) $\text{distance}(E_i, F_i) > L^{1-\tilde{\delta}}$.

With this reduction at hand, we divide each interval into at most $L^{3\tilde{\delta}}$ intervals, $E_i = \cup_{\alpha} I_i^{\alpha}$ and $F_i = \cup_{\alpha} J_i^{\alpha}$ each satisfying

- (c) $\frac{1}{2}L^{1-3\tilde{\delta}} \leq |I_i^{\alpha}|, |J_i^{\alpha}| \leq L^{1-3\tilde{\delta}}$,

and prove that for intervals I_i^{α} and J_i^{α} , satisfying Conditions (a), (b), and (c) we have

$$\sum_{\substack{a \leq Q(p,q) \leq b \\ p_j \in I_j^{\alpha}, q_j \in J_j^{\alpha} \\ p \neq q}} 1 = \int_{I_1^{\alpha} \times \dots \times I_d^{\alpha}} \int_{J_1^{\alpha} \times \dots \times J_d^{\alpha}} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy + O(L^{2(d-1) - (3d+1)\tilde{\delta}}(b-a)). \quad (8.17)$$

Summing in α and using (8.16) we have

$$\begin{aligned} \sum_{\substack{a \leq Q(p,q) \leq b \\ p_j \in I_j, q_j \in J_j \\ p \neq q}} 1 &= \sum_{\alpha} \left(\int_{I_1^{\alpha} \times \dots \times I_d^{\alpha}} \int_{J_1^{\alpha} \times \dots \times J_d^{\alpha}} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy + O(L^{2(d-1) - (3d+1)\tilde{\delta}}(b-a)) \right) \\ &\quad + O\left(L^{2(d-1)^+ - 4d\delta}(b-a)\right) \\ &= \sum_{\alpha} \int_{I_1^{\alpha} \times \dots \times I_d^{\alpha}} \int_{J_1^{\alpha} \times \dots \times J_d^{\alpha}} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy + O\left(L^{2(d-1)^+ - \tilde{\delta}}(b-a)\right). \end{aligned}$$

Using that $\tilde{\delta} = 4d\delta$ and

$$\begin{aligned} &\left| \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy - \sum_{\alpha} \int_{I_1^{\alpha} \times \dots \times I_d^{\alpha}} \int_{J_1^{\alpha} \times \dots \times J_d^{\alpha}} \mathbb{1}_{a \leq Q(x,y) \leq b} dx dy \right| \\ &\lesssim L^{2(d-1)^+ - \tilde{\delta}}(b-a) \end{aligned}$$

we conclude (8.15).

Summarizing, if by abuse of notation, we drop the index α and replace $\tilde{\delta}$ with δ , we have reduced the proof of Theorem 8.11 to proving the following proposition.

Proposition 8.12. *Fix $\epsilon > 0$, then for $\delta > 0$ sufficiently small the following statement is true: Suppose $I_j, J_j \subset [-L, L]$, $j = 1, \dots, d$ for $d \geq 3$ are intervals satisfying*

- (1) $\forall p_i \in I_i$, and $\forall q_i \in J_i$, we have $|p_i| > L^{1-\delta}, |q_i| > L^{1-\delta}$.

$$(2) \text{ distance}(I_i, J_i) > L^{1-\delta}.$$

$$(3) \frac{1}{2}L^{1-3\delta} \leq |I_i|, |J_i| \leq L^{1-3\delta}$$

Then for a, b satisfying $|a|, |b| \leq 1$ and $L^{-d+1+\epsilon} < b-a < L^{-\epsilon}$ we have

$$\sum_{\substack{a \leq Q(p,q) \leq b \\ p_j \in I_j, q_j \in J_j \\ p \neq q}} 1 = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbf{1}_{a \leq Q(x,y) \leq b} dx dy + O(L^{2(d-1)-(3d+1)\delta}(b-a)). \quad (8.18)$$

Let us now suppose I_j and J_j satisfy the hypothesis of Proposition 8.12.

Step 3: Transform the region of summation. The sum can be written as,

$$\sum_{p_i \in I_j, q_j \in J_j} \mathbf{1}_{[a,b]}(Q(p,q)) = \sum_{p_i \in I_j, q_j \in J_j} \mathbf{1}_{[0, \frac{b-a}{2}]} \left(Q(p,q) - \frac{a+b}{2} \right) \quad (8.19)$$

By writing $I_d = [u - \Delta u, u + \Delta u]$, and $J_d = [v - \Delta v, v + \Delta v]$, and utilizing the fact that $|u-v| > L^{1-\delta}$, we express the region $R_{\mathbb{Z}}$ as,

$$\left| \frac{\sum_{j=1}^{d-1} \beta_j(p_j^2 - q_j^2) - \frac{b+a}{2}}{\beta_d(p_d^2 - q_d^2)} + 1 \right| \leq \frac{b-a}{2\beta_d |p_d^2 - q_d^2|},$$

$$\leq \frac{b-a}{2\beta_d |u^2 - v^2|} + O\left(\frac{(b-a)L^{-\delta}}{|u^2 - v^2|}\right)$$

since

$$|p_d^2 - q_d^2 - u^2 + v^2| \lesssim L(\Delta u + \Delta v) \lesssim L^{2-3\delta} \quad \text{and} \quad |u^2 - v^2| \geq L^{2-2\delta}.$$

Setting $\xi = \frac{b+a}{2}$ and $\eta = \frac{b-a}{2}$, then by taking logarithms and Taylor expanding $\ln(x)$ around $x = 1$ we obtain

$$\left| \ln \left(\sum_{j=1}^{d-1} \beta_j(p_j^2 - q_j^2) - \xi \right) - \ln(p_d^2 - q_d^2) - \ln \beta_d \right| \leq \frac{\eta}{\beta_d |u^2 - v^2|} + O\left(\frac{\eta L^{-\delta}}{|u^2 - v^2|}\right), \quad (8.20)$$

here we assumed, without loss of generality, $\sum_{j=1}^{d-1} \beta_j(p_j^2 - q_j^2) - \xi > 0$ and $p_d^2 - q_d^2 > 0$.

Step 4: Replace the sum with an analogous sum.

Instead of considering the sum over the region $R_{\mathbb{Z}}$, we will consider the sum over the region $S_{\mathbb{Z}}$, defined as

$$S_{\mathbb{Z}} = \left\{ (p, q) \in \prod_{j=1}^d I_j \times \prod_{k=1}^d J_k : \left| \ln \left(\sum_{j=1}^{d-1} \beta_j(p_j^2 - q_j^2) - \xi \right) - \ln(p_d^2 - q_d^2) - \ln \beta_d \right| \leq \frac{\eta}{\beta_d |u^2 - v^2|} \right\} \quad (8.21)$$

In order to make this reduction, we need a bound on cardinality of (p, q) satisfying

$$\left| \ln \left(\sum_{j=1}^{d-1} \beta_j(p_j^2 - q_j^2) - \xi \right) - \ln(p_d^2 - q_d^2) - \ln \beta_d \right| = \frac{\eta}{\beta_d |u^2 - v^2|} + O\left(\frac{\eta L^{-\delta}}{|u^2 - v^2|}\right),$$

Such a bound would follow as a consequence of a version of a weaker version of Proposition 8.12 with the asymptotic formula (8.18) replaced with a sharp upper bound, i.e.,

Proposition 8.13. *Fix $\epsilon > 0$, then for $\delta > 0$ sufficiently small the following statement is true: Suppose I_j and J_j satisfy the hypothesis of Proposition 8.12, then for a, b satisfying $|a|, |b| \leq 1$ and $L^{-d+1+\epsilon} < b - a < L^{-\epsilon}$ we have*

$$\sum_{\substack{a \leq Q(p,q) \leq b \\ p_j \in I_j, q_j \in J_j \\ p \neq q}} 1 = O(L^{2(d-1)-3d\delta}(b-a)). \quad (8.22)$$

We note that for Proposition 8.12 compared with Proposition 8.13 we may require a stricter smallness criteria on δ relative to the choice of ϵ . With this in mind, applying Proposition 8.13, the difference in summing in p and q satisfying (8.20) and computing the cardinality of $S_{\mathbb{Z}}$ is of order $O(L^{2(d-1)-(3d+1)\delta}(b-a))$ and hence can be treated as an error. We remark that such arguments will be used later to bound analogous error terms.

By the arguments above, the sum in Proposition 8.13 may be estimated from above by the cardinality of $S_{\mathbb{Z}}$ with η replaced by 2η in the set's definition. Hence up to a factor of 2 in the definition of $S_{\mathbb{Z}}$, to prove both Proposition 8.13 and Proposition 8.12, it suffices to obtain an asymptotic formula for $S_{\mathbb{Z}}$.

If we set

$$F(p, q) = \ln \left(\sum_{j=1}^{d-1} \beta_j (p_j^2 - q_j^2) - \xi \right) - \ln (p_d^2 - q_d^2) - \ln \beta_d, \quad A = \frac{\eta}{|u^2 - v^2|},$$

then we can rewrite the cardinality of $S_{\mathbb{Z}}$ as

$$\sum_{S_{\mathbb{Z}}} 1 = \sum_{(p_j, q_j) \in I_i \times J_j} \mathbf{1}_{[-A, A]}(F(p, q)).$$

For a technical reason (as will be seen in Step 7), we replace $\mathbf{1}_{[-A, A]}$ by a smooth approximation. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth, non-negative, symmetric Friedrich mollifier, that is monotonically decreasing on $\mathbb{R}^+$. Setting $\phi_\epsilon(x) = \epsilon^{-1} \phi(\frac{x}{\epsilon})$. Then, we have

$$\begin{aligned} \sum_{S_{\mathbb{Z}}} 1 &= \sum_{(p_j, q_j) \in I_i \times J_j} (\mathbf{1}_{[-A, A]} * \phi_{L^{-100d}})(F(p, q)) + \sum_{(p_j, q_j) \in I_i \times J_j} (\mathbf{1}_{[-A, A]} - \mathbf{1}_{[-A, A]} * \phi_{L^{-100d}})(F(p, q)) \\ &= I + II. \end{aligned} \quad (8.23)$$

In an analogous argument to showing that the cardinality of $R_{\mathbb{Z}}$ can well approximated by the cardinality of $R_{\mathbb{Z}}$, we may show that sum II can be estimated up to an acceptable error.

Step 5: Expressing the sum using Fourier Transform. The number $\#S_{\mathbb{Z}}$ can be expressed using the Fourier transform as follows. Let

$$F(p, q) = \ln \left(\sum_{j=1}^{d-1} \beta_j (p_j^2 - q_j^2) - \xi \right) - \ln (p_d^2 - q_d^2) - \ln \beta_d, \quad A = \frac{\eta}{|u^2 - v^2|},$$

and write

$$\begin{aligned} I &= \sum_{(p_j, q_j) \in I_i \times J_j} (\mathbf{1}_{[-A, A]} * \phi_{L^{-100d}})(F(p, q)) = \sum_{(p_j, q_j) \in I_i \times J_j} \int e^{iF(p, q)t} (\widehat{\mathbf{1}_{[-A, A]} * \phi_{L^{-100d}}})(t) dt \\ &= \int S_1(t) \overline{S_2(t)} e^{-it \ln \beta_d} \widehat{\mathbf{1}_{[-A, A]}} \widehat{\phi_{L^{-100d}}} dt, \end{aligned}$$

where,

$$S_1(t) = \sum_{\substack{p_i \in I_i, q_i \in J_i \\ i=1, \dots, d-1}} \left(\sum_{j=1}^{d-1} \beta_j (p_j^2 - q_j^2) + \xi \right)^{it} \quad (8.24)$$

$$S_2(t) = \sum_{p_d \in I_d, q_d \in J_d} (p_d^2 - q_d^2)^{it}. \quad (8.25)$$

Step 6: A scaling argument. As mentioned earlier, if A is large compared to L^{-1} , then comparing the sum over $S_{\mathbb{Z}}$ and the area of S is relatively simple. For this reason we split our sum by scaling with a factor $\frac{A}{A_0}$, where $A_0 = \frac{L^{4/3}}{|u^2 - v^2|}$, i.e., split the integral into two terms,

$$\begin{aligned} I &= \frac{A}{A_0} \int S_1(t) \overline{S_2(t)} e^{-it \ln \beta_d} \mathbb{1}_{\widehat{[-A_0, A_0]}} \widehat{\phi_{L^{-100d}}} dt \\ &\quad + \int S_1(t) \overline{S_2(t)} e^{-it \ln \beta_d} \left(\mathbb{1}_{\widehat{[-A, A]}} - \frac{A}{A_0} \mathbb{1}_{\widehat{[-A_0, A_0]}} \right) \widehat{\phi_{L^{-100d}}} dt = III + IV. \end{aligned}$$

Ignoring the factor $\widehat{\phi_{L^{-100d}}}$, the first integral is counting p, q such that

$$\left| \ln \left(\sum_{j=1}^{d-1} \beta_j (p_j^2 - q_j^2) - \xi \right) - \ln (p_d^2 - q_d^2) - \ln \beta_d \right| \leq A_0 \beta_d^{-1}.$$

As in Step 4, the factor $\widehat{\phi_{L^{-100d}}}$ can be ignored, up to a suitable contributing error. Then, one is reduced to counting

$$\left| \sum_{j=1}^d \beta_j (p_j^2 - q_j^2) - \xi \right| \leq L^{\frac{4}{3}} + O(L^{\frac{4}{3}-\delta}).$$

Again, applying a similar upper/lower bounding argument to that used in Step 4 with the use of Proposition 8.13 replaced by the use of Proposition 8.10, we obtain

$$III = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbb{1}_{a \leq Q(x, y) \leq b} dx dy + O(L^{2(d-1)-(3d+1)\delta} (b-a)).$$

For the purpose of proving Proposition 8.13, one simply observes that the first term is of order $O(L^{2(d-1)-3d\delta} (b-a))$. Thus in order to complete the proof of Proposition 8.12, Proposition 8.13, and by implication Theorem 8.11, it suffices to estimate IV .

Step 7: Replace S_2 with a sum involving smooth cut-offs

We now replace the sum S_2 with a sum involving smooth cut-offs. This is a preparatory step, that will be needed for Step 10, in order to apply an argument involving the Mellin transform and Riemann zeta function estimates.

We rewrite S_2 in terms of the coordinates $m = p_d - q_d$, $n = p_d + q_d$ and the set

$$K := \{p_d - q_d \mid (p_d, q_d) \in I_d \times J_d\}.$$

Then S_2 becomes

$$\begin{aligned} S_2 &= \sum_{p_d, q_d} \mathbb{1}_{I_d}(p_d) \mathbb{1}_{J_d}(q_d) (p_d^2 - q_d^2)^{it} \\ &= \sum_{m, n} \mathbb{1}_{[-1, 1]} \left(\frac{m + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{m - n - 2v}{2\Delta v} \right) m^{it} n^{it} \\ &= \sum_{m \in K} m^{it} \sum_n \mathbb{1}_{[-1, 1]} \left(\frac{m + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{m - n - 2v}{2\Delta v} \right) n^{it}. \end{aligned}$$

Without loss of generality, we may assume $\Delta u \leq \Delta v$. Let us cover K by disjoint intervals M_j of length $L^{1-100d\delta}$ and define w_j to be the center of M_j . It is not difficult to show that that may be achieved such that $\#\{M_j\} \lesssim L^{100d\delta}$ we have the following bound on the set difference

$$\# \left(\bigcup_k M_j \right) \setminus K \lesssim L^{1-100d\delta}.$$

Thus we have

$$\left| S_2 - \sum_j \sum_{m \in M_j} m^{it} \sum_n \mathbb{1}_{[-1, 1]} \left(\frac{m + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{m - n - 2v}{2\Delta v} \right) n^{it} \right| \lesssim L^{2-100d\delta}.$$

Using that M_j is of length $L^{1-100d\delta}$, we may also replace m with the midpoints w_j in order to obtain the estimate

$$\left| \sum_n \left(\mathbb{1}_{[-1, 1]} \left(\frac{m + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{m - n - 2v}{2\Delta v} \right) - \mathbb{1}_{[-1, 1]} \left(\frac{w_j + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{w_j - n - 2v}{2\Delta v} \right) \right) \right| \lesssim L^{1-100d\delta},$$

and hence

$$S_2 = \sum_j \sum_{m \in M_j} m^{it} \sum_n \mathbb{1}_{[-1, 1]} \left(\frac{w_j + n - 2u}{2\Delta u} \right) \mathbb{1}_{[-1, 1]} \left(\frac{w_j - n - 2v}{2\Delta v} \right) n^{it} + O(L^{2-100d\delta}).$$

Again, up to an allowable error we may also replace the sharp cut-off cutoff functions with a smooth cut-off $\psi \equiv 1$ on $[-1 + L^{-100d\delta}, 1 - L^{-100d\delta}]$ and supported on the interval $[-1, 1]$, i.e.

$$S_2 = \sum_j \sum_{m \in M_j} m^{it} \underbrace{\sum_n \psi \left(\frac{w_j + n - 2u}{2\Delta u} \right) \psi \left(\frac{w_j - n - 2v}{2\Delta v} \right) n^{it}}_{S_j} + O(L^{2-100d\delta}). \quad (8.26)$$

Finally, the sum in m can be replaced by a sum involving a smooth cut-off, up to an allowable error

$$S_2 = \underbrace{\sum_j \sum_m \psi \left(\frac{w_j - m}{L^{1-100d\delta}} \right) m^{it} S_j}_{\tilde{S}_2} + O(L^{2-100d\delta}). \quad (8.27)$$

We now decompose IV as

$$\begin{aligned} IV &= \int S_1(t) \overline{\tilde{S}_2(t)} e^{-it \ln \beta_d} \left(\widehat{\mathbb{1}_{[-A, A]}} - \frac{A}{A_0} \widehat{\mathbb{1}_{[-A_0, A_0]}} \right) \widehat{\phi_{L^{-100d}}} dt \\ &\quad + \int S_1(t) (\overline{S_2(t)} - \overline{\tilde{S}_2(t)}) e^{-it \ln \beta_d} \left(\widehat{\mathbb{1}_{[-A, A]}} - \frac{A}{A_0} \widehat{\mathbb{1}_{[-A_0, A_0]}} \right) \widehat{\phi_{L^{-100d}}} dt = V + VI \end{aligned}$$

By (8.27) we have

$$|VI| \lesssim L^{2d-100d\delta} \int \left| \left(\widehat{\mathbb{1}_{[-A,A]}} - \frac{A}{A_0} \widehat{\mathbb{1}_{[-A_0,A_0]}} \right) \widehat{\phi_{L^{-100d}}} \right| dt.$$

Observe that

$$\left| \widehat{\mathbb{1}_{[-A,A]}}(t) - \frac{A}{A_0} \widehat{\mathbb{1}_{[-A_0,A_0]}}(t) \right| = A \left| \frac{\sin(At)}{At} - \frac{\sin(A_0 t)}{A_0 t} \right| \lesssim \min \left(AA_0^2 |t|^2, \frac{A}{1+A|t|} \right). \quad (8.28)$$

and for any N we have

$$\left| \widehat{\phi_{L^{-100d}}} \right| \lesssim \frac{1}{(1+L^{-200d}t^2)^N}. \quad (8.29)$$

Thus using that $A \lesssim (b-a)L^{-2+2\delta}$, we have

$$|VI| \lesssim (b-a)L^{2(d-1)-50d\delta},$$

which is an acceptable error.

Step 8: V is an error. Now consider V , we aim to show that

$$|V| \lesssim L^{2(d-1)-3d\delta}\eta \quad (8.30)$$

for a set of (β_2, β_d) of full measure, independent of our choice of length $L = 2^{N_1}$ and interval $(a, b) = (N_2 L^{-d+1+\varepsilon}, (N_2+1)L^{-d+1+\varepsilon})$. By Chebyshev's inequality, it suffices to show

$$\|V\|_{L^2_{\beta_2, \beta_d}} \lesssim L^{2(d-1)-(3d+1)\delta} \eta^{\frac{3}{2}}.$$

To see this, define

$$\Omega_{L, N_2} = \{\beta \in [1, 2]^d \mid |V| > L^{2(d-1)-3d\delta}\eta\}.$$

By Chebyshev's inequality we have

$$|\Omega_{L, N_2}| \lesssim \frac{1}{L^{4(d-1)-6d\delta}\eta^2} \|V\|_{L^2_{\beta_2, \beta_d}}^2 \lesssim L^{-2\delta}\eta.$$

Recall that $\eta = L^{-d+1+\varepsilon}$, then, since

$$\bigcap_{N_1 \geq M, |N_2| \leq 2L^{d-1-\varepsilon}} |\Omega_{2^{N_1}, N_2}| \geq 1 - C \sum_{j=N}^{\infty} 2^{-2j\delta} = 1 - C \frac{4^{\delta(1-N)}}{4^\delta - 1} \rightarrow 1 \quad \text{as } M \rightarrow \infty.$$

we obtain (8.30) for a set of (β_2, β_d) of full measure, where the implicit constant depends on (β_2, β_d) .

Applying (8.28) and (8.29) we have

$$\left| \left(\widehat{\mathbb{1}_{[-A,A]}} - \frac{A}{A_0} \widehat{\mathbb{1}_{[-A_0,A_0]}} \right) \widehat{\phi_{L^{-100d}}} \right| \lesssim \min \left(AA_0^2 |t|^2, \frac{A}{1+A|t|} \right).$$

Averaging in β_2 and β_d , and using Plancherel's theorem for the integral in β_d , we have from the bounds $A = \eta L^{-2+2\delta}$ and $A_0 = L^{-\frac{2}{3}+2\delta}$

$$\begin{aligned} \|V\|_{L^2_{\beta_2, \beta_d}}^2 &\lesssim A^2 \left(A_0^4 \int_{|t| \leq L^{\frac{1}{100}}} t^4 \|S_1\|_{L^2_{\beta_2}}^2 |\tilde{S}_2|^2 dt + \int_{|t| \geq L^{\frac{1}{100}}} \frac{1}{1+A^2 t^2} \|S_1\|_{L^2_{\beta_2}}^2 |\tilde{S}_2|^2 dt \right) \\ &\lesssim \eta^2 L^{\frac{-20}{3}+12\delta} \int_{|t| \leq L^{\frac{1}{100}}} t^4 \|S_1\|_{L^2_{\beta_2}}^2 |\tilde{S}_2|^2 dt + \eta^2 L^{-4+4\delta} \int_{|t| \geq L^{\frac{1}{100}}} \frac{1}{1+\eta^2 L^{-4} t^2} \|S_1\|_{L^2_{\beta_2}}^2 |\tilde{S}_2|^2 dt \\ &\lesssim \underbrace{\eta^2 L^{\frac{-8}{3}+6\delta} \int_{|t| \leq L^{\frac{1}{100}}} t^4 \|S_1\|_{L^2_{\beta_2}}^2 dt}_{VII} + \underbrace{\eta^2 L^{-4+4\delta} \int_{|t| \geq L^{\frac{1}{100}}} \frac{1}{1+\eta^2 L^{-4} t^2} \|S_1\|_{L^2_{\beta_2}}^2 |\tilde{S}_2|^2 dt}_{VIII} \end{aligned}$$

where we have used the trivial bound $\#\tilde{S}_2 \leq \#I_d \#J_d \leq L^{2-6\delta}$.

Step 9: Bounding VII. To bound $\|S_1\|_{L^2_{\beta_2}}$, we rewrite

$$\begin{aligned} |S_1(t)|^2 &= \sum_{\substack{p_i \in I_i, q_i \in J_i \\ i=1, \dots, d-1}} \sum_{\substack{r_j \in I_j, s_j \in J_j \\ j=1, \dots, d-1}} \left(\sum_{j=1}^{d-1} \beta_j (p_j^2 - q_j^2) - \xi \right)^{it} \left(\sum_{j=1}^{d-1} \beta_j (r_j^2 - s_j^2) - \xi \right)^{-it} \\ &= \sum_{\substack{p_i \in I_i, q_i \in J_i \\ i=1, \dots, d-1}} \sum_{\substack{r_j \in I_j, s_j \in J_j \\ j=1, \dots, d-1}} (p_1^2 - q_1^2 + \beta_2(p_2^2 - q_2^2) + \psi_1)^{it} (r_1^2 - s_1^2 + \beta_2(r_2^2 - s_2^2) + \psi_2)^{-it} \\ &= \sum_{\substack{p_i \in I_i, q_i \in J_i \\ i=1, \dots, d-1}} \sum_{\substack{r_j \in I_j, s_j \in J_j \\ j=1, \dots, d-1}} e^{it((\ln(p_1^2 - q_1^2 + \beta_2(p_2^2 - q_2^2) + \psi_1) - \ln(r_1^2 - s_1^2 + \beta_2(r_2^2 - s_2^2) + \psi_2)))} \end{aligned}$$

where

$$\psi_1 := \sum_{j=3}^{d-1} \beta_j (p_j^2 - q_j^2) - \xi \quad \text{and} \quad \psi_2 = \sum_{j=3}^{d-1} \beta_j (r_j^2 - s_j^2) - \xi$$

for $d > 3$ or $\psi_1 = \psi_2 = \xi$ for the case $d = 3$. Setting

$$\phi := \ln(p_1^2 - q_1^2 + \beta_2(p_2^2 - q_2^2) + \psi_1) - \ln(r_1^2 - s_1^2 + \beta_2(r_2^2 - s_2^2) + \psi_2)$$

we have

$$\begin{aligned} |\partial_{\beta_2} \phi| &= \left| \frac{p_2^2 - q_2^2}{p_1^2 - q_1^2 + \beta_2(p_2^2 - q_2^2) + \psi_1} - \frac{r_2^2 - s_2^2}{r_1^2 - s_1^2 + \beta_2(r_2^2 - s_2^2) + \psi_2} \right| \\ &\geq \left| \frac{(p_2^2 - q_2^2)(r_1^2 - s_1^2 + \psi_2) - (r_2^2 - s_2^2)(p_1^2 - q_1^2 + \psi_1)}{L^4} \right|, \end{aligned}$$

then for $t \leq L^4$, and by taking the sup over indices $3 \leq i \leq d-1$, we have

$$\int |S_1(t)|^2 d\beta_2 \lesssim \sup_{\psi_1, \psi_2} L^{2(d-3)} \sum_{\substack{p_i \in I_i, q_i \in J_i \\ r_i \in I_i, s_i \in J_i \\ i=1, 2}} \left(1 + |t| \inf_{\beta_2} |\partial_{\beta_2} \Psi| \right)^{-1}.$$

Here we are using the trivial bound for the case $1 \geq |t| \inf_{\beta_2} |\partial_{\beta_2} \Psi|$, otherwise we use Van der Corput's Lemma (see for example [37] Chapter 8, Proposition 2). For the former case, to apply the proposition, we split the integral into regions for which $\partial_{\beta_2} \Phi$ is monotonic in β_2 .

Set $(p_i - q_i)(p_i + q_i) = w_i$ and $(r_i - s_i)(r_i + s_i) = z_i$, and sum over fixed w_i and z_i using the divisor bound $d(k) \lesssim_\epsilon |k|^\epsilon$, we obtain

$$\int |S_1(t)|^2 d\beta_2 \lesssim \sup_{\psi_1, \psi_2} L^{2(d-3)+} \sum_{L^{2-2\delta} \leq |w_i|, |z_i| \leq L^2} \left(1 + \frac{|t|}{L^4} |w_2(z_1 + \psi_2) - z_2(w_1 + \psi_1)| \right)^{-1}$$

The above sum can be rearranged by summing first over the set,

$$\mathcal{A}_\psi(k, w_2, z_2) = \{L^{2-2\delta} \leq |w_1|, |z_1| \leq L^2 \mid |w_2(z_1 + \psi_2) - z_2(w_1 + \psi_1)| = k\},$$

and then over (k, w_2, z_2) to obtain,

$$\begin{aligned} \int |S_1(t)|^2 d\beta_2 &\lesssim \sup_{\psi_1, \psi_2} L^{2(d-3)+} \sum_{\substack{0 \leq k \lesssim L^2 \\ L^{2-2\delta} \leq |w_2|, |z_2| \leq L^2}} \# \mathcal{A}_\psi(k, w_2, z_2) \frac{L^4}{L^4 + |t|k} \\ &\lesssim \sup_{\psi_1, \psi_2} L^{2(d-3)+} \sum_{L^{2-2\delta} \leq |w_2|, |z_2| \leq L^2} \max_k \# \mathcal{A}_\psi(k, w_2, z_2) \left(1 + \frac{L^{4+}}{|t|}\right) \end{aligned}$$

Now we estimate $\# \mathcal{A}_\psi(k, w_2, z_2)$ for a fixed (k, w_2, z_2) . Assume $\mathcal{A}_\psi(k, w_2, z_2) \neq \emptyset$, then there exists w_0 and z_0 , such that, $L^{2-2\delta} \leq |w_0 - \psi_2| \leq L^2$ and $L^{2-2\delta} \leq |z_0 - \psi_1| \leq L^2$ and

$$[|w_2(z_0) - z_2(w_0)|] = k.$$

Thus

$$\# \mathcal{A}_\psi \lesssim \#\{w_2 \tilde{z}_1 = z_2 \tilde{w}_1 \mid |\tilde{w}_1 - w_0|, |\tilde{z}_1 - z_0| \leq L^2\} = \#\{w_1 = \frac{w_2 \tilde{z}_1}{z_2} \mid |\tilde{w}_1 - w_0|, |\tilde{z}_1 - z_0| \leq L^2\}$$

Since $\tilde{w}_1 \in \mathbb{Z}$ then $\# \mathcal{A}_\psi \lesssim 1 + \frac{L^2 \gcd(w_2, z_2)}{z_2}$, and consequently

$$\begin{aligned} \int |S_1(t)|^2 d\beta_2 &\lesssim L^{2(d-3)+} \left(1 + \frac{L^{4+}}{|t|}\right) \sum_{L^{2-2\delta} \leq |w_2|, |z_2| \leq L^2} \left(1 + \frac{L^2 \gcd(w_2, z_2)}{z_2}\right) \\ &\lesssim L^{2(d-3)+} \left(1 + \frac{L^{4+}}{|t|}\right) \left(\sum_{L^{2-2\delta} \leq |w_2|, |z_2| \leq L^2} 1 + \sum_{\substack{L^{2-2\delta} \leq |w_2|, |z_2| \leq L^2 \\ \gcd(w_2, z_2) \neq 1}} L^{2+} \right) \\ &\lesssim L^{2(d-1)+\delta} \left(1 + \frac{L^4}{|t|}\right). \end{aligned}$$

Hence, applying this bound to *VII* yields

$$VII \lesssim \eta^2 L^{\frac{-8}{3}+6\delta} \int_{|t| \leq L^{\frac{1}{100}}} |t|^3 L^{2(d+1+\delta)} dt \lesssim \eta^2 L^{4(d-1)} L^{-2d+\frac{10}{3}+7\delta+\frac{1}{25}} \lesssim \eta^3 L^{4(d-1)}.$$

where we used that $\eta = L^{-d+1+\varepsilon}$, δ is sufficiently small and $d \geq 3$.

Step 10: Bounding *VIII*. Now consider *VIII*, we have

$$\begin{aligned} VIII &\lesssim \eta^2 L^{2(d-3)+5\delta} \int_{|t| \geq L^{\frac{1}{100}}} \frac{|t| + L^4}{1 + \eta^2 L^{-4} t^2} \frac{1}{|t|} |\tilde{S}_2|^2 dt \\ &\lesssim \eta^2 L^{2(d-3)+5\delta} \sup_{\kappa \geq L^{\frac{1}{100}}} \frac{\kappa^{1+} + L^4 \kappa^{0+}}{1 + \eta^2 L^{-4} \kappa^2} \int_{|t| \geq L^{\frac{1}{100}}} \frac{1}{|t|^{1+}} |\tilde{S}_2|^2 dt \\ &\lesssim \eta^2 L^{2(d-3)+5\delta} \left(\frac{L^{2+}}{\eta} + L^{4+} \right) \left(\sup_k L^{-\frac{1}{100} 2^{-k}} \int_{L^{\frac{1}{100} 2^k}}^{L^{\frac{1}{100} 2^{k+1}}} |\tilde{S}_2|^2 dt \right). \end{aligned} \quad (8.31)$$

We proceed to estimate $|S_j|$, defined in (8.26). Defining

$$\chi(z) = \psi \left(\frac{w_j - 2u}{2\Delta u} + z \right) \psi \left(\frac{w_j - 2v}{2\Delta v} - \frac{z\Delta u}{\Delta v} \right)$$

and letting $\hat{\chi}$ denote Mellin transform of χ , then

$$S_j^2 = \left(\frac{1}{2\pi i} \int_{\Re s=2} \hat{\chi}(s) (2\Delta u)^s \zeta(s-it) ds \right)^2.$$

Shifting the contour to $\Re s = \frac{1}{2}$ we pick up the residue

$$\hat{\chi}(1+it)(2\Delta u)^{1+it},$$

which for $|t| \geq L^{\frac{1}{3}}$ is order $O(L^{-N})$ for any N due to the decay of $\hat{\chi}$ and that Δu and Δv are of comparable size. Then using $\Delta u \sim L^{1-3\delta}$

$$\begin{aligned} |S_j|^2 &= \left| \frac{1}{2\pi} \int_{\Re s=\frac{1}{2}} \hat{\chi}(s) (2\Delta u)^s \zeta(s-it) ds \right|^2 + O(1) \\ &\lesssim L^{1-3\delta} \left| \int_{\Re s=\frac{1}{2}} \hat{\chi}(s) \zeta(s-it) ds \right|^2 + O(1) = L^{1-3\delta} \left| \hat{\chi} \left(\frac{1}{2} + i \cdot \right) * \zeta \left(\frac{1}{2} - i \cdot \right) (t) \right|^2 + O(1). \end{aligned}$$

Again, using the rapid decay of $\hat{\psi}$, we have

$$|S_j|^2 \lesssim L^{1-3\delta} \left| \left(\mathbf{1}_{[-L^{100d\delta}, L^{100d\delta}]}(\cdot) \hat{\chi} \left(\frac{1}{2} + \cdot \right) \right) * \zeta \left(\frac{1}{2} - i(\cdot) \right) (t) \right|^2 + O(1).$$

We now utilize following classical L^4 bound of the zeta function in the critical strip [25]

$$\frac{1}{T} \int_0^T \left| \zeta \left(\frac{1}{2} - it \right) \right|^4 dt \lesssim T^{0+}.$$

Using the above bound yields

$$\begin{aligned} &\left\| \left(\mathbf{1}_{[-L^{100d\delta}, L^{100d\delta}]}(\cdot) \hat{\chi} \left(\frac{1}{2} + i \cdot \right) \right) * \zeta \left(\frac{1}{2} - i \cdot \right) \right\|_{L^4([L^{\frac{1}{100}} 2^k, L^{\frac{1}{100}} 2^{k+1}])}^4 \\ &\lesssim \|\hat{\chi}\|_{L^\infty} \left\| \zeta \left(\frac{1}{2} + i(\cdot) \right) \right\|_{L^4([L^{\frac{1}{100}} 2^k - L^{100d\delta}, L^{\frac{1}{100}} 2^{k+1} + L^{100d\delta}])}^4 \\ &\lesssim L^{\frac{1}{100} + \delta} 2^k \end{aligned}$$

Thus we obtain

$$\|S_j\|_{L^4([L^{\frac{1}{100}} 2^k, L^{\frac{1}{100}} 2^{k+1}])}^4 \lesssim L^{2+\frac{1}{100}-5\delta} 2^k.$$

An analogous argument also yields

$$\left\| \sum_m \psi \left(\frac{w_j - m}{L^{1-100d\delta}} \right) m^{it} \right\|_{L^4([L^{\frac{1}{100}} 2^k, L^{\frac{1}{100}} 2^{k+1}])}^4 \lesssim L^{2+\frac{1}{100}-5\delta} 2^k.$$

Using the decomposition (8.26) and the bound $\#\{M_j\} \lesssim L^{100d\delta}$, we have

$$\|\tilde{S}_2\|_{L^2([L^{\frac{1}{100}} 2^k, L^{\frac{1}{100}} 2^{k+1}])}^2 \lesssim L^{2+\frac{1}{100}+100\delta} 2^k.$$

Thus, combining the above estimate on S_2 with (8.31), we obtain

$$\begin{aligned} VIII &\lesssim \eta^2 L^{2(d-3)+5\delta} \left(\frac{L^{2+}}{\eta} + L^{4+} \right) \left(\sup_k L^{-\frac{1}{100}} 2^{-k} L^{2+\frac{1}{100}+100\delta} 2^k \right) \\ &\lesssim \eta^2 L^{2(d-2)+(5+100d)\delta} \left(\frac{L^{2+}}{\eta} + L^{4+} \right) \\ &\lesssim \eta^2 L^{4(d-1)} L^{-2d+2+200d\delta} \eta^{-1} \end{aligned}$$

where we used that $\eta = L^{-d+1+\varepsilon} > L^2$ since $d \geq 3$. Thus assuming δ to be sufficiently small, then

$$L^{-2d+2+200d\delta} \leq L^{2d+2+\varepsilon-3d\delta} = \eta^2 L^{-3d\delta},$$

and hence

$$VIII \leq L^{2(d-1)-3d\delta} \eta,$$

as desired.

8.4. Proof of Theorem 8.1. . First we note that the sum in Theorem 8.1 can be simplified as follows,

- 1) Ignore all pairs (p, q) such that $|p_j| = |q_j|$ for each j . The sum of such pairs such that $|p|, |q| \leq L^{1+\delta}$ is of order $O(t^2 L^{d(1+\delta)})$ and hence contributes to an admissible error, where here we used the restriction $t \leq L^{d-\epsilon}$.
- 2) We restrict the sum to the positive sector $p, q \in \mathbb{Z}_+^d \cap [0, L^{1+\delta}]$ for $p \neq q$. Here we are using that the subset of (p, q) such that $p_j = 0$ or $q_j = 0$ for some j is an admissible error. This follows as a consequence of Lemma 8.8. To rigorously carry out such an estimate, one must split the contributions when $|Q(p, q)| \leq \mu^{-1}$ and $|Q(p, q)| > \mu^{-1}$. Assuming without loss of generality that $p_1 = 0$, then splitting up the later part dyadically in the size of $|Q(p, q)|$ and using $|g(x)| \lesssim \frac{1}{|x|^2}$ one obtains the estimate

$$t^2 \sum_{\substack{(p,q) \in \mathbb{Z}_+^{2d} \\ p_1=0, p \neq q}} \left| W\left(\frac{p}{L}, \frac{q}{L}\right) g(\mu Q(p, q)) \right| \lesssim t^2 \left(\frac{L^{2(d-1)+-2\delta}}{\mu} + L^{(d-1)+} \right) \lesssim t L^{2d-\delta},$$

where W was defined in (8.4).

With all these reductions in mind, proving Theorem 8.1 will follow as a consequence of the following theorem.

Theorem 8.14 (Equidistribution). *Fix $\epsilon > 0$ and let $\delta > 0$ be sufficiently small. Then for generic $\beta \in [1, 2]^d$, we have that for any function $W \in \mathcal{S}(\mathbb{R}^d)$, the following holds,*

$$\sum_{\substack{(p,q) \in \mathbb{Z}_+^{2d} \\ p \neq q}} W\left(\frac{p}{L}, \frac{q}{L}\right) g(\mu Q(p, q)) = L^{2d} \int_{\mathbb{R}_+^{2d}} W(x, y) g(L^2 \mu Q(x, y)) dx dy + O\left(\frac{L^{2(d-1)-\delta}}{\mu}\right)$$

where $1 < \mu \leq L^{d-1-\epsilon}$.

We remark that the above theorem is actually stronger than required: in view of the restriction on t in the hypothesis of Theorem 8.1, we need only consider μ within the range $0 < \mu \leq L^{d-2-\epsilon}$.

Before we prove Theorem 8.14, we will need a couple of auxiliary lemmas. The following lemma is helpful in bounding errors to the asymptotic formula.

Lemma 8.15. *Let $\epsilon > 0$. Given a generic quadratic form $Q(p, q)$ as defined in (8.1), we have the following estimate*

$$\sum_{\substack{(p,q) \in \mathbb{Z}^{2d} \cap [0, L]^{2d} \\ p \neq q, |Q(p, q)| \geq a}} \frac{1}{Q(p, q)^2} \lesssim \frac{L^{(2d-2)+}}{a}. \quad (8.32)$$

for $a \geq L^{-d+\epsilon}$

Proof. We begin by dyadically subdividing the interval $[a, CL^2]$, for some large C , we define

$$R_{\mathbb{Z}(m)} \stackrel{\text{def}}{=} \{(p, q) \in \mathbb{Z}^{2d} \cap [0, L]^{2d} \mid |Q(p, q)| \in [2^m, 2^{m+1}], p \neq q\},$$

$$m_{\min} = \lfloor \log_2 a \rfloor, \quad \text{and} \quad m_{\max} = \lceil \log_2 CL^2 \rceil.$$

Applying Lemma 8.4 yields

$$\begin{aligned} \sum_{\substack{(p,q) \in \mathbb{Z}^{2d} \cap [0,L]^{2d} \\ p \neq q, |Q(p,q)| \geq a}} \frac{1}{Q(p,q)^2} &\lesssim \sum_{m=m_{\min}}^{m_{\max}} 2^{-2m} \#R_{\mathbb{Z}(m)} \\ &\lesssim \sum_{m=m_{\min}}^{m_{\max}} 2^{-2m} L^{(2d-2)+2m} \\ &\lesssim \sum_{m=m_{\min}}^{m_{\max}} \frac{L^{(2d-2)+}}{a}. \end{aligned}$$

□

The following lemma will be useful localizing the sum in Theorem 8.14.

Lemma 8.16. *Fix $\epsilon > 0$, then for $\delta > 0$ sufficiently small the following statement is true: Suppose $I_j, J_j \subset [0, L]$ for $j = 1, \dots, n$ are intervals with length satisfying*

$$L^{1-\delta} \leq |I_j|, |J_j| \quad (8.33)$$

and define

$$S_{(I,J)} \stackrel{\text{def}}{=} \{(p, q) \in \mathbb{Z}^{2d} \mid p_j \in I_j, q_j \in J_j, p \neq q\}.$$

Then for μ satisfying $L^\epsilon \leq \mu \leq L^{d-\epsilon}$ we have

$$\sum_{(p,q) \in S_{(I,J)}} g(\mu Q(p, q)) = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} g(\mu Q(x, y)) dx dy + O\left(\frac{L^{2(d-1-d\delta)}}{\mu}\right). \quad (8.34)$$

Proof. First note that by Lemma 8.15

$$\begin{aligned} \left| \sum_{(p,q) \in S_{(I,J)}} g(\mu Q(p, q)) - \sum_{\substack{(p,q) \in S_{(I,J)} \\ |Q(p,q)| \leq \mu^{-1} L^{4d\delta}}} g(\mu Q(p, q)) \right| &\lesssim \frac{L^{2(d-1)+(1-4d)\delta}}{\mu} \\ &\lesssim \frac{L^{2(d-1-d\delta)}}{\mu}. \end{aligned}$$

Define the sum $A(y)$ and the integral $\tilde{A}$ as follows

$$A(y) = \sum_{\substack{(p,q) \in S_{(I,J)} \\ |Q(p,q)| \leq y}} 1 \quad \text{and} \quad \tilde{A}(y) = \int_{I_1 \times \dots \times I_d} \int_{J_1 \times \dots \times J_d} \mathbf{1}_{[-y,y]}(Q(u, v)) du dv$$

Then in the sense of distributions

$$\begin{aligned}
\sum_{\substack{(p,q) \in S_{(I,J)} \\ |Q(p,q)| \leq \mu^{-1} L^{4d\delta}}} g(\mu Q(p,q)) &= \int_0^{\mu^{-1} L^{4d\delta}} g(\mu y) A'(y) dy \\
&= -\mu \int_0^{\mu^{-1} L^{4d\delta}} g'(\mu y) A(y) dy + g(L^{4d\delta}) A(\mu^{-1} L^{4d\delta}) \\
&= -\mu \int_0^{\mu^{-1} L^{4d\delta}} g'(\mu y) A(y) dy + O\left(\frac{L^{2(d-2)-4d\delta}}{\mu}\right)
\end{aligned}$$

where in the last inequality we applied Lemma 8.4 and the bound $L^\epsilon \leq \mu \leq L^{d-\epsilon}$. Writing $A = \tilde{A} + (A - \tilde{A})$, we have

$$\sum_{\substack{(p,q) \in S_{(I,J)} \\ |Q(p,q)| \leq \mu^{-1} L^{4d\delta}}} g(\mu Q(p,q)) = -\mu \int_0^{\mu^{-1} L^{4d\delta}} g'(\mu y) \tilde{A}(y) dy + \mu \int_0^{\mu^{-1} L^{4d\delta}} g'(\mu y) (A(y) - \tilde{A}(y)) dy + O\left(\frac{L^{2(d-2)-4d\delta}}{\mu}\right)$$

By Theorem 8.11 (by choosing δ smaller than the δ used in the theorem) it follows that assuming $y \geq L^{-d+\epsilon}$ then

$$|A(y) - \tilde{A}(y)| \lesssim L^{2(d-1)-10d\delta} y.$$

For $y \leq L^{-d+\epsilon}$ by the trivial bound

$$|A(y) - \tilde{A}(y)| \lesssim A(y) + \tilde{A}(y) \lesssim L^{d-2+\epsilon+\delta}.$$

Using the trivial bound $g'(z) \lesssim 1$ we have

$$\begin{aligned}
&\mu \int_0^{\mu^{-1} L^{4d\delta}} |g'(\mu y) (A(y) - \tilde{A}(y))| dy \\
&\lesssim \mu \int_0^{L^{-d+\epsilon}} |g'(\mu y) (A(y) - \tilde{A}(y))| dy + \int_{L^{-d+\epsilon}}^{\mu^{-1} L^{4d\delta}} |g'(\mu y) (A(y) - \tilde{A}(y))| dy \\
&\lesssim \mu L^{-2+2\epsilon+\delta} + \mu L^{2(d-1)-10d\delta} \int_{L^{-d+\epsilon}}^{\mu^{-1} L^{4d\delta}} y dy \\
&\lesssim \mu^{-1} L^{2(d-1)-\epsilon+\delta} + \mu^{-1} L^{2(d-2)-2d\delta}
\end{aligned}$$

where in the last inequality we used $\mu \leq L^{d-\epsilon}$. Choosing δ sufficiently small in relation to ϵ , this constitutes an allowable error. The proof concludes by noting that by integration by parts

$$-\mu \int_0^{\mu^{-1} L^{4d\delta}} g'(\mu y) \tilde{A}(y) dy = \iint_{\substack{I_1 \times \dots \times I_d \\ J_1 \times \dots \times J_d}} \mathbb{1}_{[-\mu^{-1} L^{4d\delta}, \mu^{-1} L^{4d\delta}]}(Q(x, y)) \left(g(\mu Q(x, y)) + g(L^{4d\delta}) \right) dx dy$$

□

Proof of Theorem 8.14. We first note that by symmetry, it is sufficient to restrict ourselves to the positive sector $p, q \in \mathbb{Z}_+^d$. Note that Lemma 8.8 implies the subset of (p, q) such that $p_j = 0$ or

$q_j = 0$ may be treated as an admissible error. Thus, it suffices to show

$$\begin{aligned} \sum_{\substack{(p,q) \in \mathbb{Z}_+^{2d} \\ p \neq q}} W\left(\frac{p}{L}, \frac{q}{L}\right) g(\mu Q(p, q)) &= L^{2d} \int_{\mathbb{R}_+^{2d}} W(x, y) g(L^2 \mu Q(x, y)) dx dy + O\left(\frac{L^{2(d-1)-\delta}}{\mu}\right) \\ &= \frac{L^{2(d-1)}}{\mu} \int_{\mathbb{R}_+^{2d}} W(x, y) \delta(Q(x, y)) dx dy + O\left(\frac{L^{2(d-1)-\delta}}{\mu}\right) \end{aligned}$$

Divide $[0, L^\delta]^d \times [0, L^\delta]^d$ into products of cubes $M_j, N_k \subset \mathbb{R}_+^d$ of length $L^{-10d\delta}$. Define $W_{j,k}$ to be the average of W over $M_j \times N_k$:

$$W_{j,k} := \int_{M_j} \int_{N_k} W(x, y) dx dy$$

Note that if $(x, y) \in M_j \times N_k$ then from the smoothness of W

$$|W(x, y) - W_{j,k}| \lesssim L^{-10d\delta}.$$

Hence using Lemma 8.15

$$\begin{aligned} \left| \sum_{\substack{(p,q) \in \mathbb{Z}_+^{2d} \\ p \neq q}} W\left(\frac{p}{L}, \frac{q}{L}\right) g(\mu Q(p, q)) - \sum_{j,k} \sum_{\substack{p \in LM_j, q \in LN_k \\ p \neq q}} W_{j,k} g(\mu Q(p, q)) \right| &\lesssim \frac{L^{2(d-1)(1+\delta)+\delta-10d\delta}}{\mu} \\ &\lesssim \frac{L^{2(d-1)-\delta}}{\mu} \end{aligned}$$

Applying Lemma 8.16 (taking δ to be sufficiently small) we obtain

$$\begin{aligned} \sum_{j,k} \sum_{\substack{p \in LM_j, q \in LN_k \\ p \neq q}} W_{j,k} g(\mu Q(p, q)) &= \sum_{j,k} \int_{LM_j} \int_{LN_k} W_{j,k} g(\mu Q(x, y)) dx dy + O\left(\frac{L^{2(d-1)-\delta}}{\mu}\right) \\ &= L^{2d} \int_{\mathbb{R}_+^{2d}} W(p, q) g(L^2 \mu Q(x, y)) dx dy + O\left(\frac{L^{2(d-1)-\delta}}{\mu}\right). \end{aligned}$$

□

Acknowledgments. TB was supported by the National Science Foundation grant DMS-1820764. PG was supported by the National Science Foundation grant DMS-1301380. JS was supported by the National Science Foundation grant DMS-1363013. ZH was supported by National Science Foundation grant DMS-1852749, NSF CAREER award DMS-1654692, and a Sloan Fellowship.

REFERENCES

- [1] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti. Some considerations on the derivation of the nonlinear quantum Boltzmann equation. *J. Statist. Phys.*, 116(1-4):381–410, 2004.
- [2] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti. Some considerations on the derivation of the nonlinear quantum Boltzmann equation. II. The low density regime. *J. Stat. Phys.*, 124(2-4):951–996, 2006.
- [3] D. Benedetto, F. Castella, R. Esposito, and M. Pulvirenti. From the N -body Schrödinger equation to the quantum Boltzmann equation: a term-by-term convergence result in the weak coupling regime. *Comm. Math. Phys.*, 277(1):1–44, 2008.

- [4] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geometric and Functional Analysis*, 3(2):107–156, 1993. bibtext: Bourgain.
- [5] J. Bourgain. On pair correlation for generic diagonal forms. *arXiv e-prints*, June 2016.
- [6] T. Buckmaster, P. Germain, Z. Hani, and J. Shatah. Effective dynamics of the nonlinear Schrödinger equation on large domains. *Comm. Pure Appl. Math.*, 71(7):1407–1460, 2018.
- [7] N. Burq and N. Tzvetkov. Random data Cauchy theory for supercritical wave equations. I. Local theory. *Invent. Math.*, 173(3):449–475, 2008.
- [8] J. W. S. Cassels. *An introduction to Diophantine approximation*. Hafner Publishing Co., New York, 1972. Facsimile reprint of the 1957 edition, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45.
- [9] M. Christ. Power series solution of a nonlinear Schrödinger equation. In *Mathematical aspects of nonlinear dispersive equations*, volume 163 of *Ann. of Math. Stud.*, pages 131–155. Princeton Univ. Press, Princeton, NJ, 2007.
- [10] A. de Suzzoni and N. Tzvetkov. On the propagation of weakly nonlinear random dispersive waves. *Arch. Ration. Mech. Anal.*, 212(3):849–874, 2014.
- [11] Y. Deng, P. Germain, and L. Guth. Strichartz estimates for the Schrödinger equation on irrational tori. *J. Funct. Anal.*, 273(9):2846–2869, 2017.
- [12] L. Erdős. Lecture notes on quantum brownian motion. In *Quantum Theory from Small to Large Scales*, pages 3–98. Oxford University Press, may 2012.
- [13] L. Erdős, M. Salmhofer, and H. Yau. On the quantum Boltzmann equation. *J. Statist. Phys.*, 116(1-4):367–380, 2004.
- [14] L. Erdős, M. Salmhofer, and H. Yau. Quantum diffusion of the random Schrödinger evolution in the scaling limit. II. The recollision diagrams. *Comm. Math. Phys.*, 271(1):1–53, 2007.
- [15] L. Erdős, M. Salmhofer, and H. Yau. Quantum diffusion of the random Schrödinger evolution in the scaling limit. *Acta Math.*, 200(2):211–277, 2008.
- [16] M. Escobedo and J. J. L. Velázquez. On the theory of weak turbulence for the nonlinear Schrödinger equation. *Mem. Amer. Math. Soc.*, 238(1124):v+107, 2015.
- [17] E. Faou. Linearized wave turbulence convergence results for three-wave systems. *arXiv e-prints*, May 2018.
- [18] E. Faou, P. Germain, and H. Hani. The weakly nonlinear large-box limit of the 2D cubic nonlinear Schrödinger equation. *Journal of the American Mathematical Society*, 2015.
- [19] I. Gallagher, L. Saint-Raymond, and B. Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.
- [20] P. Germain, H. Hani, and L. Thomann. On the continuous resonant equation for NLS, II: Statistical study. *Anal. PDE*, 8(7):1733–1756, 2015.
- [21] P. Germain, H. Hani, and L. Thomann. On the continuous resonant equation for NLS. I. Deterministic analysis. *J. Math. Pures Appl. (9)*, 105(1):131–163, 2016.
- [22] P. Germain, A. D. Ionescu, and M.-B. Tran. Optimal local well-posedness theory for the kinetic wave equation. *arXiv e-prints*, Nov. 2017.
- [23] K. Hasselmann. On the non-linear energy transfer in a gravity-wave spectrum Part 1. General theory. *Journal of Fluid Mechanics*, 12(04):481–500, Apr. 1962.
- [24] K. Hasselmann. On the non-linear energy transfer in a gravity wave spectrum. II. Conservation theorems; wave-particle analogy; irreversibility. *J. Fluid Mech.*, 15:273–281, 1963.
- [25] D. R. Heath-Brown. The fourth power moment of the Riemann zeta function. *Proc. London Math. Soc. (3)*, 38(3):385–422, 1979.
- [26] A. D. Ionescu and B. Pausader. The energy-critical defocusing NLS on $\mathbb{T}^3$. *Duke Math. J.*, 161(8):1581–1612, 2012.
- [27] O. E. Lanford, III. *Time evolution of large classical systems*. Springer, Berlin, 1975.
- [28] J. Lukkarinen and H. Spohn. Not to normal order—notes on the kinetic limit for weakly interacting quantum fluids. *Journal of Statistical Physics*, 134(5):1133–1172, Mar 2009.
- [29] J. Lukkarinen and H. Spohn. Weakly nonlinear Schrödinger equation with random initial data. *Invent. Math.*, 183(1):79–188, 2011.
- [30] A. J. Majda, D. W. McLaughlin, and E. G. Tabak. A one-dimensional model for dispersive wave turbulence. *J. Nonlinear Sci.*, 7(1):9–44, 1997.
- [31] S. Nazarenko. *Wave turbulence*, volume 825 of *Lecture Notes in Physics*. Springer, Heidelberg, 2011.
- [32] A. C. Newell and B. Rumpf. Wave turbulence: a story far from over. In *Advances in wave turbulence*, volume 83 of *World Sci. Ser. Nonlinear Sci. Ser. A Monogr. Treatises*, pages 1–51. World Sci. Publ., Hackensack, NJ, 2013.
- [33] R. Peierls. Zur kinetischen theorie der wärmeleitung in kristallen. *Annalen der Physik*, 395(8):1055–1101, 1929.
- [34] A. Soffer and M.-B. Tran. On the energy cascade of acoustic wave turbulence: Beyond Kolmogorov-Zakharov solutions. *arXiv e-prints*, Nov. 2018.

- [35] H. Spohn. Derivation of the transport equation for electrons moving through random impurities. *J. Statist. Phys.*, 17(6):385–412, 1977.
- [36] H. Spohn. The phonon Boltzmann equation, properties and link to weakly anharmonic lattice dynamics. *J. Stat. Phys.*, 124(2-4):1041–1104, 2006.
- [37] E. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [38] V. Zakharov, V. L’vov, and G. Falkovich. *Kolmogorov spectra of turbulence I: Wave turbulence*. Springer Science & Business Media, 2012.