

Invariant Gibbs measures and global strong solutions for nonlinear Schrödinger equations in dimension two

By YU DENG, ANDREA R. NAHMOD, and HAITIAN YUE

Dedicated to the memory of Professor Jean Bourgain

Abstract

We consider the defocusing nonlinear Schrödinger equation on $\mathbb{T}^2$ with Wick ordered power nonlinearity, and prove almost sure global well-posedness with respect to the associated Gibbs measure. The heart of the matter is the uniqueness of the solution as limit of solutions to canonically truncated systems, and the invariance of the Gibbs measure under the global dynamics follows as a consequence. The proof relies on the novel idea of *random averaging operators*.

1. Introduction

In this paper we study the (defocusing) Wick ordered nonlinear Schrödinger equation on the torus $\mathbb{T}^2 = (\mathbb{R}/2\pi\mathbb{Z})^2$,

$$(1.1) \quad \begin{cases} (i\partial_t + \Delta)u = W^{2r+1}(u), \\ u(0) = u_{\text{in}}, \end{cases}$$

where r is a given positive integer, and W^{2r+1} is the Wick ordered power nonlinearity of degree $2r + 1$, which will be defined below. We prove that, almost surely with respect to the associated Gibbs measure, [equation \(1.1\)](#) has

Keywords: nonlinear Schrödinger equation, invariance of Gibbs measures, probabilistic well-posedness, propagation of randomness

AMS Classification: Primary: 35Q55, 35R60, 60B20, 60H25, 60H30, 82B10; Secondary: 37E20.

Y. D. is funded in part by NSF-DMS-2246908 and Sloan Fellowship. A.N. is funded in part by NSF DMS-2052740, DMS-2101381 and the Simons Foundation Collaboration Grant on Wave Turbulence (Nahmod's Award ID 651469). H.Y. was partially supported by the Shanghai Technology Innovation Action Plan (No.22JC1402400), and by a Chinese overseas high-level young talents program (2022).

© 2024 Department of Mathematics, Princeton University.

a global *strong* solution, which is the unique limit of solutions to the canonical finite dimensional truncations. This solution map keeps the Gibbs measure invariant.

For $r = 1$ (cubic nonlinearity), this was proved by Bourgain [11]; the results for $r \geq 2$ are new. We remark that in [69], Oh and Thomann constructed almost sure global weak solutions to (1.1) with respect to the Gibbs measure, such that at any time the law of these random solutions is again given by the Gibbs measure. The main point of the current paper is the almost sure *uniqueness* of the solution with respect to the Gibbs measure.

1.1. *Setup and the main theorem.* In this section we setup the problem and state our main theorem. For a review of the background and previous results, see Section 1.2.

1.1.1. *Wick ordering and Gibbs measure.* We will fix a probability space $(\Omega, \mathcal{B}, \mathbb{P})$, and a set of independent complex Gaussian random variables $\{g_k\}_{k \in \mathbb{Z}^2}$ defined on Ω that are normalized, i.e., $\mathbb{E}g_k = 0$ and $\mathbb{E}|g_k|^2 = 1$, such that the law of g_k is rotationally symmetric.

Let $\mathcal{V} = \mathcal{S}'(\mathbb{T}^2)$ be the space of distributions on $\mathbb{T}^2$. We define the $\mathcal{V}$ -valued random variable

$$(1.2) \quad f = f(\omega) : \omega \mapsto \sum_{k \in \mathbb{Z}^2} \frac{g_k(\omega)}{\langle k \rangle} e^{ik \cdot x}, \quad \omega \in \Omega.$$

Let $d\rho$ be the Wiener measure on $\mathcal{V}$, defined for Borel sets $E \subset \mathcal{V}$ by

$$(1.3) \quad \rho(E) = \mathbb{P}(f^{-1}(E)),$$

so $d\rho$ is the law of the random variable f . This measure $d\rho$ is a countably additive Gaussian measure supported in $\bigcap_{\varepsilon > 0} H^{-\varepsilon}(\mathbb{T}^2)$, which we henceforth denote by $H^{0-}(\mathbb{T}^2)$ (similarly $H^{s-}(\mathbb{T}^2) = \bigcap_{\varepsilon > 0} H^{s-\varepsilon}(\mathbb{T}^2)$), but not in $L^2(\mathbb{T}^2)$ (see, e.g., Bogachev [9]). Define the spectral truncation Π_N by

$$(1.4) \quad \mathcal{F}_x \Pi_N u(k) := \mathbf{1}_{\langle k \rangle \leq N} \cdot \mathcal{F}_x u(k),$$

where $\mathcal{F}_x$ is the space Fourier transform, $\langle k \rangle := \sqrt{|k|^2 + 1}$ and $\mathbf{1}_P$ denotes the indicator function of a set or property P , and define the expectation of the truncated mass,

$$(1.5) \quad \sigma_N := \frac{1}{(2\pi)^2} \mathbb{E} \|\Pi_N f(\omega)\|_{L^2}^2 = \sum_{\langle k \rangle \leq N} \frac{1}{\langle k \rangle^2} \sim \log N.$$

For each N and each $p \geq 0$, define the Wick ordered powers,

$$(1.6) \quad \begin{aligned} W_N^{2p}(u) &= \sum_{j=0}^p (-1)^{p-j} \binom{p}{j} \frac{\sigma_N^{p-j} p!}{j!} |u|^{2j}, \\ W_N^{2p+1}(u) &= \sum_{j=0}^p (-1)^{p-j} \binom{p+1}{p-j} \frac{\sigma_N^{p-j} p!}{j!} |u|^{2j} u, \end{aligned}$$

and the canonical finite dimensional truncations for (1.1),

$$(1.7) \quad \begin{cases} (i\partial_t + \Delta)u_N = \Pi_N W_N^{2r+1}(u_N), \\ u_N(0) = \Pi_N u_{\text{in}}. \end{cases}$$

The following proposition ensures the convergence of the right-hand side of (1.7) as $N \rightarrow \infty$ and provides the definition of $W^{2r+1}(u)$ in (1.1).

PROPOSITION 1.1. *Let n be a nonnegative integer. Then almost surely in u with respect to the Wiener measure $d\rho$, the limit*

$$\lim_{N \rightarrow \infty} W_N^n(\Pi_N u) = \lim_{N \rightarrow \infty} \Pi_N W_N^n(\Pi_N u)$$

exists in $H^{0-}(\mathbb{T}^2)$. We will denote this limit by $W^n(u)$.

For each N , we also define the truncated potential energy

$$(1.8) \quad V_N[u] := \frac{1}{r+1} \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} W_N^{2r+2}(\Pi_N u) \, dx.$$

By Proposition 1.1, the limit quantity

$$(1.9) \quad V[u] = \lim_{N \rightarrow \infty} V_N[u] = \frac{1}{r+1} \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} W^{2r+2}(u) \, dx$$

is defined $d\rho$ -almost surely in u . We can verify that (1.7) is a finite dimensional Hamiltonian system with Hamiltonian

$$(1.10) \quad \mathcal{H}_N[u] := \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} |\nabla u|^2 \, dx + V_N[u].$$

This $\mathcal{H}_N[u]$, as well as the mass $\mathcal{M}[u] := \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} |u|^2 \, dx$, is conserved under the flow (1.7).

PROPOSITION 1.2. *Define the measure $d\mu$ by*

$$(1.11) \quad d\mu = Z^{-1} e^{-V[u]} d\rho, \quad Z := \int_{\mathcal{V}} e^{-V[u]} d\rho(u).$$

Then $d\mu$ is mutually absolutely continuous with $d\rho$, and the Radon-Nikodym derivative $Z^{-1} e^{-V[u]}$ belongs to $L^q(d\rho)$ for any $1 \leq q < \infty$. We call this $d\mu$ the Gibbs measure for (1.1).

Propositions 1.1 and 1.2 stem from seminal works of Nelson [66], [67] and of Glimm-Jaffe [36] in the context of quantum field theory (see also Simon [73] and Da Prato-Tubaro [25]). As stated, a proof of these propositions can be found in [69].

1.1.2. *The main theorem.* We can now state our main theorem.

THEOREM 1.3. *There exists a Borel set $\Sigma \subset \mathcal{V}$ with $\mu(\mathcal{V} \setminus \Sigma) = 0$, such that $W^{2r+1}(u) \in H^{0-}(\mathbb{T}^2)$ is well defined for $u \in \Sigma$. Furthermore,*

- (1) *For each $u_{\text{in}} \in \Sigma$ and each $t \in \mathbb{R}$, the solution $u_N(t)$ to (1.7) converges to a unique limit*

$$(1.12) \quad \lim_{N \rightarrow \infty} u_N(t) = u(t)$$

in $H^{0-}(\mathbb{T}^2)$, and $u(t) \in \Sigma$ for each $t \in \mathbb{R}$. This $u(t)$ solves (1.1) in the distributional sense.

- (2) *The limit $u(t)$ in (1.12) defines, for each $t \in \mathbb{R}$, a map from Σ to itself: $u(t) =: \Phi_t u_{\text{in}}$. These maps then satisfy the usual group properties and keep the Gibbs measure $d\mu$ invariant; namely,*

$$(1.13) \quad \mu(E) = \mu(\Phi_t(E))$$

for any Borel set $E \subset \Sigma$.

Remark 1.4. In proving Theorem 1.3, we will replace $H^{0-}(\mathbb{T}^2)$ by $H^{-\varepsilon}(\mathbb{T}^2)$ where $0 < \varepsilon \ll 1$, and throughout the proof we will fix this ε . (We then take a countable intersection in ε .)

Remark 1.5. The Wick ordering (1.6) is needed in order to have a meaningful solution theory, due to the infinite mass in the support $H^{0-}(\mathbb{T}^2)$ of the Gibbs measure $d\mu$; see, e.g., [45].

Remark 1.6. (1) Since the Gibbs measure $d\mu$ is mutually absolutely continuous with the Gaussian measure $d\rho$, part (1) of Theorem 1.3 can also be viewed as an almost sure global well-posedness result with random initial data $u_{\text{in}} = f(\omega)$ as in (1.2).

(2) The truncation frequency N in (1.7) can be any positive number. For simplicity, in the proof below we will assume that $N \in 2^{\mathbb{Z}}$. The general case follows from placing N between $N'/2$ and N' , where $N' \in 2^{\mathbb{Z}}$, and analyzing the difference $u_{N'} - u_N$ in the same way as $u_{N'} - u_{N'/2}$.

(3) The solution $u(t)$ defined in Theorem 1.3 is unique in the following sense. Suppose we replace the truncation in (1.7) by any other canonical truncation; say we replace Π_N in (1.4) by the multiplier with symbol $k \mapsto \varphi(\langle k \rangle/N)$ in both (1.5) and the second line of (1.7), where $\varphi(z)$ is either $\mathbf{1}_{|z| \leq 1}$ or a smooth function supported in $|z| \leq 1$ and $\varphi(0) = 1$. We may also keep the

Π_N in the first line of (1.7) unchanged, or replaced it by 1. Then in any case Theorem 1.3 remains true; moreover, the set Σ and the limit $u(t)$ obtained *do not depend on φ* . The proof is basically the same, with only minor adjustments in a few places. We will not pursue these matters here.

Remark 1.7. (1) Theorem 1.3 is part of the program of constructing invariant Gibbs measures and studying their dynamics for the (renormalized) defocusing nonlinear Schrödinger equation (1.1); see Section 1.2.1 below. With Bourgain [10] completing all cases with $d = 1$ and Theorem 1.3 completing all cases with $d = 2$ after Bourgain [11], the remaining cases that are expected to be solvable¹ are the invariance of Gibbs measure for $(d, r) = (3, 1)$, and invariance of white noise for $(d, r) = (1, 1)$. We expect both to be strictly harder than Theorem 1.3 as they are *critical in the probabilistic scaling*; see Section 1.3.1.

(2) The corresponding problem of constructing invariant Gibbs measures for nonlinear *wave* equations is in general much easier than the Schrödinger problem, due to the derivative gain in Duhamel's formula. Indeed, this has been completely solved, due to the results of Friedlander [33] (see Zhidkov [84] under more restrictive assumptions on the nonlinearity) in dimension $d = 1$, Bourgain [13] in dimension $d = 2$ (see also Oh-Thomann [70]), and the recent result of the authors with Bringmann [16] for $(d, r) = (3, 1)$ proved after the current paper was submitted.

1.2. *A review of previous results.* We start by reviewing previous results and methods on PDEs in the probabilistic setting. As the literature is now extensive, we will put emphasis on the works most relevant to the current paper.

1.2.1. *Invariant measures.* Since the pioneering works of Lebowitz-Rose-Speer [58] and Bourgain [10, 11], there have been numerous results regarding invariant measures for nonlinear dispersive equations. In general, for any Hamiltonian dispersive equation, we may construct the associated Gibbs measure

$$(1.14) \quad d\mu \sim e^{-\beta H} \prod_x dx,$$

where $\beta > 0$ is a fixed parameter (which we may set to be 1) and H is the Hamiltonian. The definition (1.14) is only formal; in some cases it can be

¹Gibbs measures for (1.1) are available only for $d = 1$ (both focusing and defocusing), $d = 2$ (defocusing only), and $(d, r) = (3, 1)$ (defocusing only). This is related to the existence/nonexistence of ϕ_d^{2r+2} theories; see, for example, Aizenman [1], Fröhlich [35], Brydges-Slade [17], and Aizenman-Duminil Copin [2]. The white noise measure is always formally invariant under (1.1), but is compatible with the dynamics only when $(d, r) = (1, 1)$.

justified by using the Gaussian measure as a reference measure and writing $d\mu$ as a weighted Wiener measure. For example, the Hamiltonian for (1.1) is

$$H = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} (|\nabla u|^2 + \frac{1}{r+1} W^{2r+2}(u)) dx$$

and the Gibbs measure

$$d\mu \sim \underbrace{\exp \left[\frac{-1}{r+1} \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} W^{2r+2}(u) dx \right]}_{\text{weight}} \cdot \underbrace{\exp \left[\frac{-1}{(2\pi)^2} \int_{\mathbb{T}^2} |\nabla u|^2 dx \right] \prod_{x \in \mathbb{T}^2} dx}_{\text{Gaussian measure}}$$

can be rigorously defined as a weighted Wiener measure,² as in [Proposition 1.2](#). Defining such Gibbs-type measures and studying their properties under various dynamics is a major problem in constructive quantum field theory.

The Gibbs measure $d\mu$ for a given dispersive equation is *formally* invariant due to a “formal Liouville’s Theorem” and the conservation of Hamiltonian. It is of great interest to establish this invariance rigorously, as this would be the first step in studying the global dynamics from the statistical ensemble point of view. In [\[10\]](#), [\[11\]](#), Bourgain developed a systematic way of showing the invariance of $d\mu$ from the invariance of finite dimensional Gibbs measures, provided we have local well-posedness or almost sure local well-posedness with respect to $d\mu$.

Therefore, justifying the invariance of $d\mu$ (and other similar formally invariant measures) basically reduces to proving almost sure local well-posedness on the support of $d\mu$. As this support is very rough in high dimensions (namely $H^{1-\frac{d}{2}-}$ for (1.1) in dimension d), most known results are limited to one dimension, or require strong symmetry. For the Schrödinger equation (1.1) on the torus $\mathbb{T}^d$, Bourgain [\[10\]](#) solved the case $d = 1$, and he extended this to $d = 2$ and $r = 1$ in [\[11\]](#). These are the only results known for the nonlinear Schrödinger equation prior to the current paper. More is known for wave equations as noted in [Remark 1.7\(2\)](#) above; see [\[33\]](#), [\[84\]](#), [\[13\]](#), [\[70\]](#), [\[16\]](#).

Apart from the standard Schrödinger and wave models on tori, there are many results, again mostly in one dimension or under radial symmetry, where the invariance of corresponding Gibbs measures (or of associated weighted Wiener measures) are justified for various dispersive models on various background manifolds (see, e.g., [\[79\]](#), [\[68\]](#), [\[80\]](#), [\[78\]](#), [\[62\]](#), [\[64\]](#), [\[75\]](#), [\[26\]](#),

²Strictly speaking the measure defined in [Proposition 1.2](#) involves an additional weight which is an exponential of the L^2 mass. As the mass is also conserved, this does not affect any invariance properties.

[27], [31], [72], [77], [76] and references therein). We also mention the compactness method of Alveberio and Cruzeiro [3] that explores the tightness of the sequence of finite dimensional measures and applies the theorems of Prokhorov and Skorokhod to obtain existence of weak solutions (see, e.g., [18], [63], [69], [81]). These are less related to the current paper, and we will not elaborate further.

1.2.2. Probabilistic well-posedness theory. It has long been known that PDEs with *randomness* generally behave better in terms of local well-posedness (i.e., probabilistic well-posedness holds below the deterministic well-posedness threshold). Progress has been made in two parallel directions: random initial data problems and stochastically forced problems.

The first results along this line are due to the seminal works by Bourgain [11], [12] in the random data setting and later to Da Prato-Debussche [23], [24] in the stochastic setting. The idea in both cases is to make a linear-nonlinear decomposition and observe the effect of *probabilistic smoothing*. For example, in [11], equation (1.1) with $r = 1$ on $\mathbb{T}^2$ is studied with random initial data in $H^{-\varepsilon}$ for some $0 < \varepsilon \ll 1$, in which (1.1) is deterministically ill-posed. However with randomness we may construct solutions to (1.1) that have the form $u = e^{it\Delta}u(0) + v$, where $u(0) \in H^{-\varepsilon}$ is the random initial data, and v belongs to some *positive* Sobolev space in which (1.1) is well posed. In other words, the solution contains a rough random part $u_{\text{lin}} := e^{it\Delta}u(0)$ and a smooth remainder v . The point here is that, even though u_{lin} is rough, it has an explicit random structure that allows us to control the nonlinear interactions between u_{lin} and u_{lin} , and between u_{lin} and v , in a more regular space.

Until recently the method of Bourgain, as well as its higher order variants that include some nonlinear interactions of u_{lin} with itself into the rough random part,³ has been the dominant strategy of exploiting randomness in the local well-posedness theory for dispersive and wave equations with random data. After Bourgain's pioneering work, there has been substantial success using this method. (For a sample of works, we refer the readers to [11], [19], [22], [26], [14], [65], [82], [15], [55], [8], [32], [54] and references therein.) However this method by itself has its limitations and does not lead to optimal results in most cases.

A few years ago, a series of important works emerged, which revolutionized the study of local well-posedness for stochastically forced PDEs, in fact reaching the optimal exponents in the parabolic case. These include the *theory*

³This usually results in a finite or infinite tree expansion.

of *regularity structures* of Hairer [46], [48], [47], [49] and the *para-controlled calculus* of Gubinelli-Imkeller-Perkowski [37], [38]. A third method based on Wilsonian renormalization group analysis was independently proposed by Kupiainen in [57].

The theory of regularity structures is based on the local-in-space properties of solutions at fine scales (so it is particularly suitable for parabolic equations); it builds a general theory of distributions that includes the profiles coming from the noise, and allows for multiplications and analysis of the nonlinearity. Since its success with the KPZ equation [46] and the Φ_3^4 model [48], this theory has been further developed by Hairer and collaborators and is now powerful enough to solve a wide range of problems that are subcritical according to a suitable parabolic scaling. We will not get into the details, but we refer the reader to [34], [49], [50], [51], [52], [21], [61], [60] and references therein for nice expositions of these ideas.

The theory of para-controlled calculus, which is in spirit the point of departure of the present paper, takes a different approach and is based on the following idea. In the approach of Bourgain and of Da Prato-Debussche mentioned above, some nonlinear interactions between u_{lin} and v may not have enough regularity despite v being more regular than u_{lin} . However, a key observation is that the only bad terms here are the high-low interactions where the high frequencies come from u_{lin} and the low frequencies come from v , and such terms can be *para-controlled* by the high-frequency inputs (which are nonlinear interactions of u_{lin} with itself). Here f being para-controlled by g simply means that f equals the high-low paraproduct of g with some other function h , up to a smoother remainder. With such structure, these para-controlled terms can be shown to have similar randomness structures as the nonlinear interactions of u_{lin} with itself, and thus can be handled similarly as in Bourgain's or Da Prato-Debussche's approach, leaving an even smoother remainder. The para-controlled calculus also has a higher order variant; see [5], [6]. We refer the reader to [20], [41], [42], [43], [44], [60], [71], [5], [4], [6] and references therein for expositions of these ideas and some other recent developments on this method.

Finally, we would like to mention two recent results of Gubinelli-Koch-Oh [39] and Bringmann [15]. In [39] the authors applied a version of para-controlled calculus to the stochastic wave equation setting, and they obtained almost sure local well-posedness for a quadratic wave equation with additive white noise on $\mathbb{T}^3$. This relied on several new ingredients, including the analysis of a random operator (which is different from and unrelated to the random averaging operator introduced in the current paper).

In [15] the author studied the nonlinear wave equation with quadratic derivative nonlinearity on $\mathbb{R}^3$ and improved the known well-posedness threshold with random initial data, again by analyzing high-low interactions. This work introduced an improved para-controlled scheme that accounts for all the bad terms in the absence of linear smoothing effects, and also made the very important observation that the high-frequency and low-frequency parts can in fact be made *independent*. See Section 1.3.2 for more detailed discussions.

Remark 1.8. We note from previous discussions that, in the same dimension and for the same nonlinearity, the probabilistic improvement (defined as the difference between exponents of the deterministic H^s well-posedness threshold and the obtained probabilistic H^s well-posedness threshold) is always much smaller for *Schrödinger equations* compared to wave and heat equations.

There are two reasons for this. First, heat equations are compatible with Hölder spaces C^s , which scale much higher than H^s , but a function with independent Gaussian Fourier coefficients that belongs to H^s will automatically belong to C^{s-} due to Khintchine's inequality. This allows for a scaling at a higher regularity, and hence we will be in a better situation when studying heat equations. Such an advantage cannot be exploited for Schrödinger and wave equations, since C^s spaces are not compatible even with their linear evolution and cannot be used in any well-posedness theory.

Second, the Duhamel evolution for the heat equation gains two derivatives, and for the wave equation gains one.⁴ This allows for room to apply Sobolev embedding, and also reduces the task of controlling the nonlinearity to the task of making sense of products, which is still hard but at least more manageable. In comparison, the Duhamel evolution for Schrödinger has no smoothing effect, and it can be challenging to close the estimates even when the relevant products are well defined.

1.3. *Difficulties and the strategy.* We now turn to the proof of Theorem 1.3. This proof consists of two parts: (a) proving almost sure local well-posedness for (1.1) on the support of the Gibbs measure, and (b) applying formal invariance to extend local solutions to global ones. Since part (b) is essentially an adaptation of Bourgain's classical proof [11], we shall focus on the local theory in part (a). For exposition simplicity, we will also replace $W^{2r+1}(u)$ by the pure power $|u|^{2r}u$ in the discussion below.

⁴Note that the *derivative* wave equation studied in [15] behaves similarly to Schrödinger as the one derivative gain is cancelled by the derivative nonlinearity. Therefore the probabilistic improvement obtained in [15] is also a tiny amount compared to other results for wave equations.

The obvious difficulty here is that the Gibbs measure $d\mu$ is supported in H^{0-} , while the (deterministic) scaling threshold, below which (1.1) is ill-posed, is $s_c = 1 - \frac{1}{r} \rightarrow 1$ as $r \rightarrow \infty$. In the language of Remark 1.8, we need to obtain a probabilistic improvement ≈ 1 . Therefore, it is important to understand exactly how randomness allows us to beat scaling. This is contained in the notion of *probabilistic scaling*, which we discuss below.

1.3.1. *The probabilistic scaling.* Consider the nonlinear Schrödinger equation

$$(1.15) \quad (i\partial_t + \Delta)u = |u|^{2r}u, \quad u(0) = u_{\text{in}}$$

on $\mathbb{T}^d$. The scaling critical threshold for (1.15) is

$$(1.16) \quad s_c = \frac{d}{2} - \frac{1}{r},$$

and (1.15) is expected to be locally well-posed in H^s only if $s \geq s_c$. This can be demonstrated in multiple ways, but the one most relevant to us is as follows. Suppose the initial data u_{in} has Fourier transform $\mathcal{F}_x u_{\text{in}}(k)$ supported in $|k| \sim N$ with $|\mathcal{F}_x u_{\text{in}}(k)| \sim N^{-\alpha}$ with $\alpha = s + \frac{d}{2}$; then $\|u_{\text{in}}\|_{H^s} \sim 1$. If local well-posedness holds, then we should expect that the second iteration (say at time $t = 1$),

$$u^{(1)} := \int_0^1 e^{i(1-t')\Delta} (|e^{it'\Delta} u_{\text{in}}|^{2r} e^{it'\Delta} u_{\text{in}}) dt',$$

satisfies $\|u^{(1)}\|_{H^s} \lesssim 1$. By performing Fourier expansions, we essentially get

$$(1.17) \quad \mathcal{F}_x u^{(1)}(k) \sim \sum_{\substack{k_1 - \dots + k_{2r+1} = k \\ |k_j| \lesssim N}} \frac{1}{\langle \Sigma \rangle} \prod_{j=1}^{2r+1} \widehat{u_{\text{in}}}(k_j), \quad \Sigma = |k|^2 - |k_1|^2 + \dots - |k_{2r+1}|^2,$$

where complex conjugates are omitted. In the worst scenario this gives, up to logarithmic factors,

$$|\mathcal{F}_x u^{(1)}(k)| \lesssim N^{-(2r+1)\alpha} \sup_{m \in \mathbb{Z}} \#S_m,$$

where $S_m = \{(k_1, \dots, k_{2r+1}) : k_1 - \dots + k_{2r+1} = k, |k_j| \lesssim N, \Sigma = m\}$. By dimension counting, we expect that $\#S_m \lesssim N^{2rd-2}$, so in order for $\|u^{(1)}\|_{H^s} \lesssim 1$ we need $-(2r+1)\alpha + 2rd - 2 \leq -\alpha$, or equivalently $s \geq s_c$.

Now, in the *random data* setting, suppose the Fourier coefficients of initial data $\{\widehat{u_{\text{in}}}(k)\}$ are *independent Gaussians* of size $N^{-\alpha}$. The sum (1.17) will then be a sum of products of independent Gaussian random variables, which is reminiscent of the classical Central Limit Theorem. Recall that in the latter we have a sum of M independent random objects of unit size, and under certain

general conditions, this sum scales only like $\sqrt{M}$ as opposed to M if without randomness. In the same way, we would expect essentially a “square root gain” here, that is,

$$|\mathcal{F}_x u^{(1)}(k)| \lesssim N^{-(2r+1)\alpha} \sup_{m \in \mathbb{Z}} (\#S_m)^{\frac{1}{2}} \lesssim N^{-(2r+1)\alpha + rd - 1},$$

so in order for $\|u^{(1)}\|_{H^s} \lesssim 1$, it suffices to have $-(2r+1)\alpha + rd - 1 \leq -\alpha$, or equivalently

$$(1.18) \quad s \geq s_p := -\frac{1}{2r}.$$

Note that s_p is independent of the dimension and that we always have $s_p \leq s_c$. We will call this s_p the critical threshold for *probabilistic scaling*.⁵ We refer the reader to [30] for further discussions about the probabilistic scaling paradigm.

Remark 1.9. With the above heuristics, it is natural to expect that (1.1) will be almost surely locally well-posed with random initial data in $H^{s-}(\mathbb{T}^d)$; that is,

$$u_{\text{in}} = \sum_{k \in \mathbb{Z}^d} \frac{g_k}{\langle k \rangle^\alpha} e^{ik \cdot x}, \quad \alpha = s + \frac{d}{2},$$

in any *probabilistically subcritical* space with $s > s_p$.

Indeed, after the current paper was submitted, the authors completed the paper [29], which establishes this result. The proof in [29] is based on the *theory of random tensors*, which is the natural extension of the main technique of the current paper, i.e., the method of *random averaging operators* introduced in Section 1.3.3 below.

1.3.2. Discussions of earlier methods. With the scaling heuristics in Section 1.3.1, note that the support of the Gibbs measure $d\mu$, which is H^{0-} , is *above* the probabilistic critical space H^{s_p-} , where $s_p = -\frac{1}{2r}$. Therefore it is reasonable to believe that almost sure local well-posedness of (1.1) should hold in the support of $d\mu$. However, the justification of such heuristics is far from trivial, due to the intrinsic difficulties associated with the Schrödinger equation explained in Remark 1.8.

To motivate the method of random averaging operators introduced in this paper, we first discuss the possibility of applying existing methods⁶ to the

⁵This is associated with Gaussian random variables, but by the Central Limit Theorem, the scaling is the same for more general types of random variables.

⁶We will not discuss the regularity structure theory [48], as it relies on local expansions in physical space and is thus not compatible with Schrödinger or wave equations.

setting of (1.1), the ideas behind these other methods, and why these ideas do not work here. Below we set $u_{\text{in}} = f(\omega)$ in (1.1).

Method 1: Bourgain-Da Prato-Debussche. We start with the classical idea of Bourgain [11] and Da Prato-Debussche [23], [24]. As discussed in Section 1.2.2, this idea is based on the following observation:

- Observation 1: Nonlinear components of random data solutions always have higher regularity than linear ones.

Now, suppose we apply this idea to (1.1), which leads to the ansatz $u(t) = e^{it\Delta}f(\omega) + w$, where w belongs to $C_t^0 H_x^s$, or more precisely $X^{s, \frac{1}{2}+}$ (see Section 2.3 for relevant definitions) for some positive s . In particular, this w will contain components of form

$$(1.19) \quad u^{(1)}(t) = I(|e^{it\Delta}f(\omega)|^{2r} e^{it\Delta}f(\omega)), \quad IF(t) := \int_0^t e^{i(t-t')\Delta} F(t') dt'.$$

However, it is shown in [11] that even when $r = 1$ (and obviously also for larger r), the $u^{(1)}$ defined in (1.19) belongs to $X^{s, \frac{1}{2}+}$ *only for* $s < \frac{1}{2}$. As the space $X^{\frac{1}{2}-, \frac{1}{2}+}$ is still supercritical with respect to deterministic scaling for $d = 2$ and $r \geq 2$, there will be no hope of solving (1.1) using the above ansatz. We may perform higher order Picard iterations, but it turns out that regardless of the order, there is always some contribution in the remainder that has regularity $X^{\frac{1}{2}-, \frac{1}{2}+}$, and the problem persists.

Method 2: Para-controlled calculus. Next we may try the idea of para-controlled calculus of Gubinelli-Imkeller-Perkowski [37] (and Gubinelli-Koch-Oh [39] for wave). This is based on the following observation, in addition to Observation 1 above:

- Observation 2: In probabilistically subcritical settings, the bad regularity of the nonlinear (and subsequent non-explicit) terms only come from high-low interactions.

Now, applying this idea to (1.1) would lead to the ansatz

$$(1.20) \quad u = u_{\text{lin}} + X + Y,$$

where

$$(1.21) \quad u_{\text{lin}} = e^{it\Delta}f(\omega), \quad X = \sum_N I(P_N u_{\text{lin}} \cdot |P_{\ll N} u|^{2r}),$$

I is as in (1.19), and P_N are the standard Littlewood-Paley projections. The term X para-controlled by u_{lin} will be constructed in some less regular space (say $X^{\frac{1}{2}-, \frac{1}{2}+}$), which allows the remainder Y to be constructed in a more regular space (say $X^{1-, \frac{1}{2}+}$).

However, since the Duhamel operator I gains no derivative, the term Y will contain contributions of form

$$(1.22) \quad Z := \sum_N I(P_N X \cdot |P_{\ll N} u|^{2r}),$$

which in fact also has regularity $X^{\frac{1}{2}-, \frac{1}{2}+}$. This shows that the ansatz (1.20) is still not sufficient for (1.1).

Method 3: Bringmann's method. Bringmann's paper [15] introduced a refinement to the para-controlled ansatz in [37], [39] designed for derivative wave equations (where the Duhamel evolution gains no derivative), based on the following observation in addition to Observations 1 and 2 above:

- Observation 3: The subsequent expansion terms of the form Z that are similar to (1.22) can be packed in a para-linearized solution, and moreover the high and low frequency inputs are independent.

More precisely, [15] suggests modification of the ansatz to

$$(1.23) \quad u = u_{\text{lin}} + X + Y,$$

where X is the solution to the para-linearized equation

$$(1.24) \quad X = \sum_N I(P_N(u_{\text{lin}} + X) \cdot |P_{\ll N^\alpha} u|^{2r}),$$

for some $\alpha \in (0, 1)$, which is constructed in the less regular space $X^{\frac{1}{2}-, \frac{1}{2}+}$, and the remainder Y is constructed in the more regular space $X^{1-, \frac{1}{2}+}$. Moreover, the low frequency input $P_{\ll N^\alpha} u$ is essentially *independent* with $P_N u_{\text{lin}}$, which is a crucial observation first made in [15]. This allows for an inductive construction, and also allows the use of more powerful probabilistic tools such as multilinear Wiener chaos estimates.

However, the above ansatz is still not sufficient to control both X and Y in the desired regularities, due to the following reason. Since Y contains high-high interactions, where by “high” we mean frequencies $\gtrsim N^\alpha$, it is easy to see that α has to be chosen close to 1 for large r , in order for Y to be bounded in $X^{1-, \frac{1}{2}+}$; in fact a calculation similar to those in Section 1.3.1 shows that $\alpha \geq 1 - 2/r$. But when α is close to 1, the expression of X will involve terms

$$\sum_N I(P_N u_{\text{lin}} \cdot |P_{\ll N^\alpha} X|^{2r}).$$

To control this contribution, even with independence between $P_N u_{\text{lin}}$ and $P_{\ll N^\alpha} X$, we would still need to bound the power $|P_{\ll N^\alpha} X|^{2r}$ in a suitable space; however in the ansatz of [15] we only know that X is a function in

$X^{\frac{1}{2}-, \frac{1}{2}+}$, which does not imply any useful control for $|P_{\ll N^\alpha} X|^{2r}$ due to super-criticality of this space. Of course we may also exploit [equation \(1.24\)](#) satisfied by X , but each iteration of this equation produces higher powers of X , which again cannot be bounded using the ansatz of [\[15\]](#).

Note that when $r = 2$, it might be possible to carry out the scheme of [\[15\]](#) with a small α by doing some refined analysis (which is by no means immediate in view of all the difficulties for the Schrödinger equation in [Remark 1.8](#)). Our approach, which is described below, allows instead for a *unified* treatment for all values of r by synthesizing the main underlying ideas and crucially capturing the true randomness structure of the solution.

1.3.3. Random averaging operators. It is now clear that, to solve [\(1.1\)](#), it is necessary to exploit all three observations in [Section 1.3.2](#) (in particular, it is very important to exploit the independence between high and low frequency inputs as first done by Bringmann in [\[15\]](#)), but this is still not enough. In fact, the missing piece is the following observation, in addition to Observations 1–3 in [Section 1.3.2](#):

- Observation 4: The input $P_{\ll N^\alpha} u$ in [\(1.24\)](#) is *not* an arbitrary function in $X^{\frac{1}{2}-, \frac{1}{2}+}$; it has its own randomness structure, which must be captured in order for the ansatz to be complete.

This randomness property of $P_{\ll N^\alpha} u$ cannot be captured by any norm of the *para-controlled term* X defined in [\(1.24\)](#). Indeed, to see and exploit this randomness property, we need to perform a **shift of paradigm**, by turning the focus from the para-controlled term X to the **linear operator** defined by

$$(1.25) \quad \mathcal{Q} : y \rightarrow z, \quad \text{where} \quad z = \sum_N I(P_N(y + z) \cdot |P_{\ll N^\alpha} u|^{2r});$$

note that by this definition we have $X = \mathcal{Q}(u_{\text{lin}})$. In practice it is more convenient to consider the simpler operator

$$(1.26) \quad \mathcal{P} : y \mapsto \sum_N I(P_N y \cdot |P_{\ll N^\alpha} u|^{2r}),$$

which obeys the same estimates; in fact from [\(1.25\)](#) and [\(1.26\)](#) it is easy to see that $\mathcal{Q} = \mathcal{P}(1 - \mathcal{P})^{-1}$. Now we shall *extract all the randomness properties of $P_{\ll N^\alpha} u$, as well as properties of the multilinear expression $|\cdot|^{2r}$ when applied to these random objects, and turn them into two particular norm bounds for the operator $\mathcal{P}$ defined in [\(1.26\)](#) (or $\mathcal{Q}$ in [\(1.25\)](#))*.

This is the key idea behind our method of **random averaging operators**. The norms we choose are the operator norm $\|\cdot\|_{\text{OP}}$ and the Hilbert-Schmidt norm $\|\cdot\|_{\text{HS}}$ with $\mathcal{P}$ (or $\mathcal{Q}$) viewed as a linear operator from the space

$X^{s, \frac{1}{2}+}$ to itself; this does not depend on s so we may in fact choose $s = 0$. When the input y in (1.26) is a linear Schrödinger flow, we get an operator from H^s to $X^{s, \frac{1}{2}+}$, and we will also measure the corresponding operator and Hilbert-Schmidt norms. Suppose the maximum frequency of $P_{\ll N^\alpha} u$ in (1.26) is $L \ll N^\alpha$. Then roughly speaking, these norm bounds will look like

$$(1.27) \quad \|\mathcal{P}\|_{\text{OP}} \lesssim L^{-\delta_0}, \quad \|\mathcal{P}\|_{\text{HS}} \lesssim N^{\frac{1}{2}+\delta_1} L^{-\frac{1}{2}},$$

where $\delta_1 \ll \delta_0 \ll 1$. Note that these bounds are obviously false for arbitrary functions $P_{\ll N^\alpha} u \in X^{\frac{1}{2}-, \frac{1}{2}+}$, so they indeed encode the (implicit) randomness structure of $P_{\ll N^\alpha} u$, which is reflected via the multilinear expression $|\cdot|^{2r}$.

We also remark that the choices of the operator and Hilbert-Schmidt norms and the bounds in (1.27) are natural in the following sense: the operator norm bound in (1.27) is what guarantees the solvability of the para-linearized equation (1.24), and the Hilbert-Schmidt norm bound in (1.27) is what guarantees that the term X defined in (1.24) belongs to $X^{\frac{1}{2}-, \frac{1}{2}+}$. Moreover, the estimate (1.27) is preserved under multiplication (which is important in the proof), due to the simple inequalities

$$\begin{aligned} \|\mathcal{P}_1 \mathcal{P}_2\|_{\text{OP}} &\leq \|\mathcal{P}_1\|_{\text{OP}} \|\mathcal{P}_2\|_{\text{OP}}, \\ \|\mathcal{P}_1 \mathcal{P}_2\|_{\text{HS}} &\leq \min(\|\mathcal{P}_1\|_{\text{HS}} \|\mathcal{P}_2\|_{\text{OP}}, \|\mathcal{P}_1\|_{\text{OP}} \|\mathcal{P}_2\|_{\text{HS}}). \end{aligned}$$

See Proposition 3.2 for details.

These operators $\mathcal{P}$ and $\mathcal{Q}$, which will be of central importance in our proof, depend on the object $P_{\ll N^\alpha} u$ that has an implicit randomness structure. Moreover, in the Fourier variables, this operator can be viewed as a weighted average taken over smaller scales $L \ll N^\alpha$. We thus call it a *random averaging operator*, which explains the name of our method.

1.3.4. The full ansatz. We can now describe the full ansatz of the solution u to (1.1), with random averaging operators. On the surface we consider the same framework as in [15], namely $u = u_{\text{lin}} + X + Y$, where X is defined as in (1.24) and $Y \in X^{1-, \frac{1}{2}+}$ is a smooth remainder. However, in our ansatz, instead of merely estimating X in $X^{\frac{1}{2}-, \frac{1}{2}+}$, we exploit the fact that X equals a random averaging operator applied to u_{lin} , where this random averaging operator satisfies a bound of form (1.27). To be precise, choosing $\alpha < 1$ close to 1, we further write that

$$(1.28) \quad X = \mathcal{Q}(u_{\text{lin}}),$$

so we have the ansatz

$$(1.29) \quad u = u_{\text{lin}} + \mathcal{Q}(u_{\text{lin}}) + Y.$$

Here the random averaging operator $\mathcal{Q}$ is $\mathcal{P}(1 - \mathcal{P})^{-1}$, and $\mathcal{P}$ has the form

$$\mathcal{P} = \sum_N \sum_{L \ll N^\alpha} \mathcal{P}_{NL},$$

where $\mathcal{P}_{NL}$ has the form (1.26) with the maximum frequency of $P_{\ll N^\alpha} u$ in (1.26) being L . This $\mathcal{P}_{NL}$ is a Borel function of $(g_k(\omega))_{\langle k \rangle \leq L}$, which is independent with $e^{it\Delta} P_N f(\omega)$, and satisfies (1.27); the operator $\mathcal{Q}$ has a similar decomposition into $\mathcal{Q}_{NL}$ that satisfies the same bounds (1.27). See Section 3.1 for the precise formulas.

With the ansatz (1.29), the proof of local well-posedness then proceeds by inducting on frequencies to show (1.27) and to bound Y in $X^{1-, \frac{1}{2}+}$. In fact, suppose these are true for components of frequency $\ll N^\alpha$. Then the bounds (1.27) and high-low frequency independence imply that the part of $P_{\ll N^\alpha} u$ involving the random averaging operators really behaves like a linear Schrödinger flow, so in $\mathcal{P}_{NL}$ (see (1.26)), we can effectively assume that $P_{\ll N^\alpha} u$ is replaced by either a linear flow or a smooth function in $X^{1-, \frac{1}{2}+}$. Hence (1.27) follows from large deviation estimates for multilinear Gaussians and a T^*T (random matrix) argument like the one of Bourgain [11], and the estimate for Y follows from standard contraction mapping arguments. See Sections 3 and 5 for details.

1.4. *Further discussions.* The notion of probabilistic scaling introduced in Section 1.3.1 is a general philosophy and is not specific to (1.1); in fact it can be extended to more general situations. These include, but are not limited to, the following ones.

1.4.1. *Wave equations.* For the wave equation (say with a power non-linearity as in (1.15)), we can apply the same heuristics as in Section 1.3.1. However, due to the gain of one derivative in the Duhamel formula, instead of (1.17) we essentially have that

$$(1.30) \quad \mathcal{F}_x u^{(1)}(k) \sim \frac{1}{\langle k \rangle} \sum_{\substack{k_1 - \dots - k_{2r+1} = k \\ |k_j| \lesssim N}} \frac{1}{\langle \Sigma \rangle} \prod_{j=1}^{2r+1} \widehat{u_{\text{in}}}(k_j), \quad \Sigma = |k| - |k_1| + \dots - |k_{2r+1}|.$$

Now assume $|k| \sim N$. Then compared to (1.17) we gain an extra factor N^{-1} due to the antiderivative, while in the dimension counting argument we gain one less power of N as Σ is now linear instead of quadratic. In the deterministic setting this leads to the same scaling condition as the Schrödinger equation, but in the probabilistic setting this trade-off leads to a *better* bound than in the Schrödinger case as the one-dimension disadvantage gets “square-rooted” by exploiting randomness as explained above. This then

gives

$$|\mathcal{F}_x u^{(2)}(k)| \lesssim N^{-(2r+1)\alpha-1} N^{rd-\frac{1}{2}},$$

which leads to a *lower* probabilistic scaling threshold, namely $s_p^{\text{wave}} = -\frac{3}{4r}$.

However, unlike Schrödinger, there is also a “high-high to low” interaction, namely $|k| \sim 1$ in (1.30), that needs to be addressed. A similar calculation using randomness and counting bounds yields heuristically that

$$|\mathcal{F}_x u^{(2)}(k)| \lesssim N^{-(2r+1)\alpha} N^{rd-\frac{1}{2}},$$

which leads to the restriction $s \geq s'_p := \frac{-d-1}{2(2r+1)}$. Thus it is reasonable to conjecture that the wave equation is almost surely locally well-posed in $H^s \times H^{s-1}$ for

$$s > \max(s_p^{\text{wave}}, s'_p) = \max\left(-\frac{3}{4r}, -\frac{d+1}{4r+2}\right).$$

In particular, when $(d, r) = (3, 1)$, the conjectured threshold is $H^{-\frac{2}{3}}$, which is below $H^{-\frac{1}{2}-}$ where the Gibbs measure is supported, consistent with the recent positive result [16].

1.4.2. Other dispersion relations and/or nonlinearities. For more general dispersion relations on $\mathbb{T}^d$, say $\Lambda = \Lambda(k)$, the only thing above that changes is the counting bound for the set

$$\begin{aligned} S_m &= \{(k_1, \dots, k_{2r+1}) : k_1 - \dots + k_{2r+1} = k, \\ \Sigma &:= \Lambda(k) - \Lambda(k_1) + \dots - \Lambda(k_{2r+1}) \in [m, m+1]\}. \end{aligned}$$

In contrast to parabolic equations (see [48]) where the exact form of the elliptic operator is irrelevant once the order is fixed, here the properties of S_m depend crucially on the choice of Λ , and they have to be analyzed on a case by case basis. For simple dispersion relations like Schrödinger, wave or gravity water wave (where $\Lambda(k) = \sqrt{|k|}$) this is doable, but when Λ gets more complicated (say a high degree polynomial), determining the optimal local well-posedness threshold requires getting sharp bounds for $\#S_m$, which in itself may be a hard problem in analytic number theory.

For derivative nonlinearities, the scaling heuristics can still be carried out and the value of s_p can be calculated in the same way as before (since such heuristics essentially take into account only the high-high interactions). However the actual almost sure well-posedness threshold may be strictly higher than s_p due to high-low interactions and derivative loss (in the same way that the deterministic theory for quasilinear equations does not quite reach scaling — see, e.g., [56], [74]), which may be worth looking at first in some simple models. There is also the possibility of exponential nonlinearities but they are more of an “endpoint” nature and will not be discussed here.

1.4.3. *Stochastic equations.* We may also consider wave and heat equations with additive noise (Schrödinger is also possible but has worse behavior), say of form

$$(1.31) \quad (\partial_t^2 - \Delta)u = |u|^{2r}u + \zeta, \quad \text{or} \quad (\partial_t - \Delta)u = |u|^{2r}u + \zeta,$$

where ζ is the spacetime white noise that is essentially (after discretizing the time Fourier variable)

$$\zeta = \sum_{k,\xi} g_{k,\xi} e^{i(k \cdot x + \xi t)},$$

where $g_{k,\xi}$ are independent normalized Gaussians.

The heat case of (1.31) has been studied extensively; see [53] and the references in Section 1.2.2. In this case we can confirm that the scaling heuristics of Section 1.3.1 are consistent with that of [48]. Indeed, note that for (1.31), the linear evolution $e^{it\Delta}u_{\text{in}}$ in Section 1.3.1 is replaced by the linear noise term

$$\psi(t) = \int_0^t e^{(t-t')\Delta} \zeta(t') dt' \sim \sum_{k,\xi} \frac{g_{k,\xi}}{|k|^2 + |\xi|} e^{i(k \cdot x + \xi t)},$$

which belongs to $C_t^0 H_x^{s-}$ for $s = -\frac{d}{2} + 1$. The goal would then be to guarantee that the second iteration

$$u^{(1)}(t) = \int_0^t e^{(t-t')\Delta} (|\psi(s)|^{2r} \psi(t')) dt'$$

belongs to the same space. By similar arguments, this time also taking into account the time Fourier variable, we can show that this leads to the restriction $r(d-2) < 2$, which coincides with the subcriticality condition introduced in [48] in the case of (1.31).

For the wave case of (1.31), similar calculations lead to the subcriticality condition $r(d-2) < \frac{3}{2}$, which is consistent with the results in [39], [40]. In both cases, due to the particular choice of white noise, the high-to-low interactions studied in (1) above give the same condition on (d, r) .

1.4.4. *Other geometries.* Considering more general geometries in addition to $\mathbb{T}^d$ will lead to different scenarios. For compact manifold, the canonical randomization would be based on the spectral expansion of the Laplacian, in which case the probabilistic scaling depends on the *global geometry* of the underlying manifold. This is because the randomized data is *not* localized and has the same amplitude at each point of the domain. In comparison, the *deterministic scaling* threshold does not depend on the geometry because it corresponds to the data zoomed out at a point, which is localized.

For non-compact manifolds (say $\mathbb{R}^d$), the canonical randomization based on Laplacian eigenfunctions (i.e., $e^{i\xi \cdot x}$) would lead to initial data with infinite

L^2 mass, which is not compatible⁷ with the Schrödinger equation (1.1). There is another kind of “Wiener randomization” based on dividing the Fourier space into unit boxes and randomizing on each box (see, e.g., [83], [59], [7]), which produces localized initial data (which is not preserved by the linear flow). In the $\mathbb{R}^d$ case, this leads to the critical threshold $s_p = -3/(4r)$ that is lower than $\mathbb{T}^d$.

1.5. *Notation and choice of parameters.* In this section we collect some of the notation and conventions that will be used in the proof. Throughout the paper, the space and time Fourier transforms will be respectively fixed as

$$(1.32) \quad (\mathcal{F}_x u)(k) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{-ik \cdot x} f(x) dx, \quad (\mathcal{F}_t u)(\xi) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi t} f(x) dt.$$

We will be working in (k, t) or (k, ξ) variables, instead of the x variable; so we will abbreviate $(\mathcal{F}_x u)(k)$ simply as u_k and will abuse notation and write $u = u_k(t)$. The symbol $\widehat{u}$ will always represent time Fourier transform (or, for the second formula in (1.34) below, the corresponding two-dimensional time Fourier transform), so $(\mathcal{F}_{t,x} u)(k, \xi) = \widehat{u}_k(\xi)$.

Let the space mean $\mathcal{A}$ be defined by $\mathcal{A}u = (\mathcal{F}_x u)(0) = u_0$. (This may depend on time t if u does.) Define the twisted spacetime Fourier transform

$$(1.33) \quad \widetilde{u}_k(\lambda) = \widetilde{u}(k, \lambda) = \widehat{u}_k(\lambda - |k|^2).$$

We also need to study functions $h_{kk^*}(t)$ of variables $k, k^* \in \mathbb{Z}^2$ and $t \in \mathbb{R}$, and $\mathfrak{h}_{kk'}(t, t')$ of variables $k, k' \in \mathbb{Z}^2$ and $t, t' \in \mathbb{R}$; for these we will define

$$(1.34) \quad \widetilde{h}_{kk^*}(\lambda) = \widehat{h_{kk^*}}(\lambda - |k|^2) \quad \text{and} \quad \widetilde{\mathfrak{h}}_{kk'}(\lambda, \lambda') = 2\pi \widehat{\mathfrak{h}_{kk'}}(\lambda - |k|^2, |k'|^2 - \lambda'),$$

where λ and λ' are Fourier variables corresponding to t and t' respectively.

Recall that $\langle k \rangle = \sqrt{|k|^2 + 1}$, and $\mathbf{1}_E$ is the indicator function. The cardinality of a finite set E will be denoted by $|E|$ or by $\#E$. We will be using smooth cutoff functions $\chi = \chi(z)$ that equal 1 for $|z| \leq 1$ and equal 0 for $|z| \geq 2$. For any Schwartz function φ and any $0 < \tau \ll 1$, we will define $\varphi_\tau(t) = \varphi(\tau^{-1}t)$.

For a complex number z , define $z^+ := z$ and $z^- := \bar{z}$; we will also use the notation z^ι , where ι will always be $\pm$. In the proof we will encounter tuples $(k_1, \dots, k_n)$, or maybe $(k_1^*, \dots, k_n^*)$, with associated signs $\iota_1, \dots, \iota_n \in \{\pm\}$; they are usually linked by some equation $\iota_1 k_1 + \dots + \iota_n k_n = d$ or $\iota_1 |k_1|^2 + \dots + \iota_n |k_n|^2 = \alpha$, where d and α are given, or by some expression $g_{k_1^*}^{\iota_1} \dots g_{k_n^*}^{\iota_n}$.

⁷It is, however, compatible with the wave equation due to finite speed of propagation; in particular, the result of [16] is expected to be true also for the Gibbs measure on $\mathbb{R}^3$. We will not discuss this here, but see [16].

Definition 1.10. In the above context we say that (k_i, k_j) is a *pairing* in $\{k_1, \dots, k_n\}$ if $k_i = k_j$ and $\iota_i = -\iota_j$. We say a pairing is *over-paired* if $k_i = k_j = k_\ell$ for some $\ell \notin \{i, j\}$. Pairings and over-pairings in $\{k_1^*, \dots, k_n^*\}$ are defined similarly.

For example, suppose $k = k_1 - k_2 + k_3 + d$. If $k_1 = k_2$, then (k_1, k_2) is a pairing in $\{k_1, k_2, k_3\}$; if $k = k_1$, then (k, k_1) is a pairing in $\{k, k_1, k_2, k_3\}$. If $k = k_1 = k_2 \neq k_3$, then (k_1, k_2) is over-paired if considered as a pairing in $\{k, k_1, k_2, k_3\}$, but not if considered a pairing in $\{k_1, k_2, k_3\}$.

Recall [Remark 1.6\(2\)](#) that for the truncation Π_N defined in [\(1.4\)](#), N will be a power of two that is also $\gtrsim 1$. The same applies to other capital letters like M , L , R , etc. Define also $\Pi_N^\perp = \text{Id} - \Pi_N$ and $\Delta_N = \Pi_N - \Pi_{\frac{N}{2}}$, so that

$$(\Delta_N u)_k = \mathbf{1}_{N/2 < \langle k \rangle \leq N} \cdot u_k.$$

Let $\mathcal{V}_N$ and $\mathcal{V}_N^\perp$ be the ranges of Π_N and $\Pi_N^\perp$. For $N_1, \dots, N_n$, we will define $\max^{(j)}(N_1, \dots, N_n)$ to be the j -th maximal element among them; we denote it by $N^{(j)}$.

Definition 1.11. For any N as above, we denote by $\mathcal{B}_{\leq N}$ the σ -algebra generated by the random variables g_k for $\langle k \rangle \leq N$, and by $\mathcal{B}_{\leq N}^+$ the smallest σ -algebra containing both $\mathcal{B}_{\leq N}$ and the σ -algebra generated by the random variables $|g_k|^2$ for $k \in \mathbb{Z}^2$.

Recall that ε is fixed by [Remark 1.4](#). Let $1 \gg \delta_0 \gg \delta$ be two fixed small positive constants depending on r and ε (think of $\delta_0 = \delta^{1/50}$). Define the parameters

$$(1.35) \quad \begin{aligned} \gamma &= \delta^{\frac{3}{4}}, \quad \gamma_0 = \delta^{\frac{5}{4}}, \quad \kappa = \delta^{-4}, \\ b &= \frac{1}{2} + \delta^4, \quad b_1 = b + \delta^4, \quad b_2 = b - \delta^6, \quad a_0 = 2b - 10\delta^6. \end{aligned}$$

Then we have the following hierarchy:

$$(1.36) \quad \varepsilon \gg \delta_0 \gg \gamma \gg \delta \gg \gamma_0 \gg \delta\gamma_0 \gg b - \frac{1}{2} = b_1 - b = \kappa^{-1} \gg \delta^6.$$

Denote by θ any positive quantity that is small enough depending on δ (for example $\theta \ll \delta^{50}$). This θ may take different values at different places. Let C be any large absolute constant depending only on r , and let C_θ be any constant depending on θ . Unless otherwise stated, the constants in the $\lesssim$, $\ll$ and $O(\cdot)$ symbols will depend on C_θ . Finally, if some statement S about a random variable holds with probability $\mathbb{P}(S) \geq 1 - C_\theta e^{-A^\theta}$ for some quantity $A > 0$ and with given θ and C_θ independent of A , we will say this S is *A-certain*.

The rest of the paper is organized as follows. In [Section 2](#) we introduce the gauge transform and reduce to a favorable nonlinearity, and we define the norms that will be used in the proof below. In [Section 3](#) we identify the precise structure of the solution according to the ideas of [Section 1.3](#), and reduce the local well-posedness to some multilinear estimates, namely [Proposition 3.4](#). In [Section 4](#) we then set up the necessary tools (large deviation and counting estimates) needed in the proof of [Proposition 3.4](#), and [Section 5](#) contains the proof itself. Finally in [Section 6](#) we apply an adapted version of Bourgain's argument to extend local solutions to global ones and finish the proof of [Theorem 1.3](#).

1.6. *Acknowledgment.* The second author thanks Hendrik Weber for helpful comments regarding references on the ϕ^4 model.

2. Equations, measures and norms

In this section we make some preparations for the proof of [Theorem 1.3](#). These include definitions of Wick ordering and gauge transform, properties of Gibbs and Gaussian measures and their finite dimensional truncations, and choices of function and operator norms and linear estimates.

2.1. *Wick ordering and a gauge transform.* We start by defining the Wick ordering and the gauge transform. Consider a general polynomial $\mathcal{M}_n(u)$ or $\mathcal{H}_n(u)$ of degree n , defined by

$$(2.1) \quad [\mathcal{M}_n(u)]_k = \sum_{\iota_1 k_1 + \dots + \iota_n k_n = k} a_{kk_1 \dots k_n} u_{k_1}^{\iota_1} \dots u_{k_n}^{\iota_n},$$

$$(2.2) \quad [\mathcal{H}_n(u)]_{kk'} = \sum_{\iota_1 k_1 + \dots + \iota_n k_n + \iota k' = k} a_{kk'k_1 \dots k_n} u_{k_1}^{\iota_1} \dots u_{k_n}^{\iota_n},$$

where $a_{kk_1 \dots k_n}$ and $a_{kk'k_1 \dots k_n}$ are constants. Recall the definition of pairings in [Definition 1.10](#).

Definition 2.1. We say that the polynomial in (2.1) is *input-simple* if $a_{kk_1 \dots k_n} = 0$, unless each pairing in $\{k_1, \dots, k_n\}$ is over-paired. Similarly we say it is *simple* if $a_{kk_1 \dots k_n} = 0$ unless each pairing in $\{k, k_1, \dots, k_n\}$ is over-paired, and we say the polynomial in (2.2) is *simple* if $a_{kk'k_1 \dots k_n} = 0$ unless each pairing in $\{k, k', k_1, \dots, k_n\}$ is over-paired. These notions also apply to multilinear forms.

For $m := \mathcal{A}|u|^2$, define the following polynomials of degree $n \in \{2p, 2p+1\}$ (this u may also be replaced by v):

$$\begin{aligned}
(2.3) \quad & :|u|^{2p}: = \sum_{j=0}^p (-1)^{p-j} \binom{p}{j} \frac{m^{p-j} p!}{j!} |u|^{2j}, \\
& :|u|^{2p} u: = \sum_{j=0}^p (-1)^{p-j} \binom{p+1}{p-j} \frac{m^{p-j} p!}{j!} |u|^{2j} u.
\end{aligned}$$

We will see in the proof of [Proposition 2.2](#) that each of these is input-simple.

Define a gauge transform $v_N = \mathcal{G}_N u_N$ associated with (1.7) by

$$(2.4) \quad v_N(t) = u_N(t) \cdot \exp \left((r+1)i \int_0^t \mathcal{A}[W_N^{2r}(u_N)] dt' \right).$$

Then u_N solves (1.7) if and only if v_N solves the gauged equation

$$(2.5) \quad \begin{cases} (i\partial_t + \Delta)v_N = \Pi_N \mathcal{Q}_N(v_N), \\ v_N(0) = \Pi_N u_{\text{in}}, \end{cases}$$

where

$$(2.6) \quad \mathcal{Q}_N(v) = W_N^{2r+1}(v) - (r+1)\mathcal{A}[W_N^{2r}(v)]v.$$

Since the gauge transform does not change the $t = 0$ data, we will write $v_{\text{in}} = u_{\text{in}}$. The inverse of $\mathcal{G}_N$ is given by

$$(2.7) \quad u_N(t) = v_N(t) \cdot \exp \left(- (r+1)i \int_0^t \mathcal{A}[W_N^{2r}(v_N)] dt' \right),$$

since by (1.6), if $v = e^{i\alpha}u$ where $\alpha \in \mathbb{R}$, then $W^n(u) = W^n(v)$ for even n and $W^n(u) = e^{i\alpha}W^n(v)$ for odd n . Now assume v_N is a solution to (2.5). Let m_N be the truncated mass, which is conserved under (2.5),

$$(2.8) \quad m_N := \mathcal{A}|v_N|^2 = \sum_{\langle k \rangle \leq N} |(u_{\text{in}})_k|^2,$$

and let $m_N^* := m_N - \sigma_N$, where σ_N is as in (1.5). Note that m_N and m_N^* are random terms if $u_{\text{in}} = f(\omega)$ as in (1.2). The following proposition give us a useful formula for $\mathcal{Q}_N$.

PROPOSITION 2.2. *We have*

$$(2.9) \quad \mathcal{Q}_N(v_N) = \sum_{l=0}^r \binom{r+1}{r-l} \frac{(m_N^*)^{r-l} r!}{l!} \mathcal{N}_{2l+1}(v_N),$$

where

$$(2.10) \quad \mathcal{N}_{2l+1}(v) = :|v|^{2l} v: - (l+1)(\mathcal{A} :|v|^{2l}:)v.$$

Here $\mathcal{N}_{2l+1}$ is a simple polynomial of degree $2l+1$. By standard procedure, we can define a $(2l+1)$ -multilinear form, which we still denote by $\mathcal{N}_{2l+1}$, such that it reduces to $\mathcal{N}_{2l+1}(v)$ when all inputs equal to v .

Proof. First we prove (2.9). By the definition of $\mathcal{Q}_N(v)$ (see (2.6)), it will suffice to obtain that

$$(2.11) \quad W_N^{2r+1}(v_N) = \sum_{l=0}^r \binom{r+1}{r-l} \frac{(m_N^*)^{r-l} r!}{l!} :|v_N|^{2l} v_N:$$

and

$$(2.12) \quad (r+1)W_N^{2r}(v_N) = \sum_{l=0}^r \binom{r+1}{r-l} \frac{(m_N^*)^{r-l} r!}{l!} (l+1) :|v_N|^{2l}:$$

By the definition of $:|v|^{2l}v:$ (see (2.3)), and combinatorial identities, we have that

$$(2.13) \quad \begin{aligned} & \text{the right-hand side of (2.11)} \\ &= \sum_{l=0}^r \binom{r+1}{r-l} (m_N^*)^{r-l} r! \sum_{k=0}^l (-1)^{l-k} \binom{l+1}{l-k} \frac{m_N^{l-k}}{k!} |v_N|^{2k} v_N \\ &= \sum_{k=0}^r (-1)^{r-k} \binom{r+1}{r-k} \frac{r!}{k!} |v_N|^{2k} v_N \sum_{l=k}^r \binom{r-k}{l-k} m_N^{l-k} (-m_N^*)^{r-l}, \end{aligned}$$

which implies (2.11) due to binomial expansion.

Similarly we can calculate

$$(2.14) \quad \begin{aligned} & \text{the right-hand side of (2.12)} \\ &= \sum_{l=0}^r \binom{r+1}{r-l} (m_N^*)^{r-l} r! (l+1) \sum_{k=0}^l (-1)^{l-k} \binom{l}{k} \frac{m_N^{l-k}}{k!} |v_N|^{2k} \\ &= (r+1) \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} \frac{r!}{k!} |v_N|^{2k} \sum_{l=k}^r \binom{r-k}{l-k} m_N^{l-k} (-m_N^*)^{r-l}, \end{aligned}$$

which implies (2.12).

Next we prove that $:|v|^{2p}:$ and $:|v|^{2p}v:$ are input-simple. Working in Fourier space, for any monomial

$$\mathfrak{X} := (v_{k_1})^{a_1} (\overline{v_{k_1}})^{b_1} \cdots (v_{k_n})^{a_n} (\overline{v_{k_n}})^{b_n},$$

where the k_j 's are different and a_j and b_j are nonnegative integers, we will calculate the coefficient of $\mathfrak{X}$ in the polynomial $:|v|^{2p}:$ and $:|v|^{2p}v:$, and we will prove that this coefficient is zero provided $a_1 = b_1 = 1$. Now clearly the coefficient of $\mathfrak{X}$ in $|v|^{2p}$ and $|v|^{2p}v$, denoted by $[\mathfrak{X}] (|v|^{2p})$ and $[\mathfrak{X}] (|v|^{2p}v)$, are

$$[\mathfrak{X}] (|v|^{2p}) = \frac{(p!)^2}{a_1! \cdots a_n! b_1! \cdots b_n!}, \quad [\mathfrak{X}] (|v|^{2p}v) = \frac{p!(p+1)!}{a_1! \cdots a_n! b_1! \cdots b_n!},$$

under the assumptions $b_1 + \cdots + b_n = p$ and $a_1 + \cdots + a_n = p$ (or $a_1 + \cdots + a_n = p+1$).

Recalling that $m = \mathcal{A}|v|^2$, we can calculate that

(2.15)

$$\begin{aligned} [\mathfrak{X}](:|v|^{2p}:) &= \sum_{l=0}^p (-1)^{p-l} \frac{p!}{l!} \binom{p}{l} (p-l)!(l!)^2 \sum_{c_1+\dots+c_n=p-l} \prod_{s=1}^n \frac{1}{c_s!(a_s-c_s)!(b_s-c_s)!} \\ &= (-1)^p (p!)^2 \sum_{l=0}^p (-1)^l \sum_{c_1+\dots+c_n=p-l} \prod_{s=1}^n \frac{1}{c_s!(a_s-c_s)!(b_s-c_s)!}. \end{aligned}$$

Now suppose that $a_1 = b_1 = 1$; then $c_1 \in \{0, 1\}$. Clearly the terms for l and with $c_1 = 0$ exactly cancel the terms for $l+1$ and with $c_1 = 1$, so $[\mathfrak{X}](:|v|^{2p}:) = 0$. Similarly we can prove $[\mathfrak{X}](:|v|^{2p}v:) = 0$.

Finally, we prove that $\mathcal{N}_{2p+1}(v) = :|v|^{2p}v: - (p+1)(\mathcal{A} :|v|^{2p}:)v$ is simple. By definition it suffices to prove that

$$\mathcal{A}(\mathcal{N}_{2p+1}(v)\bar{v}) = \mathcal{A}(\bar{v} :|v|^{2p}v:) - (p+1)m \mathcal{A} :|v|^{2p}:$$

is input-simple. We will actually show that this equals $\mathcal{A} :|v|^{2p+2}:$; whence the result will follow. In fact, by (2.3) we have

$$\begin{aligned} \mathcal{A} :|v|^{2p+2}: &= \sum_{l=0}^{p+1} (-1)^{p-l+1} \binom{p+1}{l} \frac{(p+1)!}{l!} m^{p-l+1} \mathcal{A}|v|^{2l}, \\ \mathcal{A}(\bar{v} :|v|^{2p}v:) &= \sum_{l=1}^{p+1} (-1)^{p-l+1} \binom{p+1}{l} \frac{p!}{(l-1)!} m^{p-l+1} \mathcal{A}|v|^{2l}, \\ -(p+1)m \mathcal{A} :|v|^{2p}: &= \sum_{l=0}^p (-1)^{p-l+1} (p+1) \binom{p}{l} \frac{p!}{l!} m^{p-l+1} \mathcal{A}|v|^{2l}, \end{aligned}$$

so the first line equals the sum of the second and third lines by direct calculation. $\square$

Remark 2.3. Later on we will consider general multilinear forms $\mathcal{N}_n$, which are simple and can be written as

$$(2.16) \quad [\mathcal{N}_n(v^{(1)}, \dots, v^{(n)})]_k = \sum_{\iota_1 k_1 + \dots + \iota_n k_n = k} a_{kk_1 \dots k_n} (v_{k_1}^{(1)})^{\iota_1} \dots (v_{k_n}^{(n)})^{\iota_n}.$$

We may assume that the coefficient $a_{kk_1 \dots k_n}$ is symmetric in the k_j 's for which $\iota_j = +$, and also symmetric in the k_j 's for which $\iota_j = -$. Moreover, we assume that this coefficient only depends on the *set of pairings* among $\{k, k_1, \dots, k_n\}$.

The multilinear form $\mathcal{N}_{2l+1}$ corresponding to (2.10) satisfies the above properties, and we will assume without any loss of generality that $\iota_j = +$ (i.e., $\mathcal{N}_{2l+1}$ is linear in $v^{(j)}$) for j odd, and $\iota_j = -$ (i.e., $\mathcal{N}_{2l+1}$ is conjugate linear in $v^{(j)}$) for j even.

2.2. *Finite and infinite dimensional measures.* We now summarize some properties of the infinite dimensional and finite dimensional (or truncated) Gaussian and Gibbs measures, which will be used later in the proof.

Recall that $\mathcal{V}_N$ and $\mathcal{V}_N^\perp$ are respectively the ranges of the projections Π_N and $\Pi_N^\perp$. We will identify $\mathcal{V}$ with $\mathcal{V}_N \times \mathcal{V}_N^\perp$. Let $d\rho_N$ and $d\rho_N^\perp$ be the Gaussian measures defined on $\mathcal{V}_N$ and $\mathcal{V}_N^\perp$ respectively, such that $d\rho = d\rho_N \times d\rho_N^\perp$. Define the measures $d\mu_N^\circ$ on $\mathcal{V}_N$ and $d\mu_N$ on $\mathcal{V}$ by

$$(2.17) \quad d\mu_N^\circ = Z_N^{-1} e^{-V_N[u]} d\rho_N, \quad d\mu_N = Z_N^{-1} e^{-V_N[u]} d\rho; \quad Z_N = \int_{\mathcal{V}_N} e^{-V_N[u]} d\rho_N(u).$$

Then we have that $d\mu_N = d\mu_N^\circ \times d\rho_N^\perp$. Recall also the measure $d\mu$ defined in [Proposition 1.2](#); all these are probability measures.

PROPOSITION 2.4. *When $N \rightarrow \infty$, we have $Z_N \rightarrow Z$, with $0 < Z < \infty$. The sequence $Z_N^{-1} e^{-V_N[u]}$ converges to $Z^{-1} e^{-V[u]}$ almost surely, and also in $L^q(d\rho)$ for any $1 \leq q < \infty$. The measure $d\mu_N$ converges to $d\mu$ in the sense that the total variation of $\mu_N - \mu$ converges to 0. Finally, the measure $d\mu_N^\circ$ is invariant under the flows of (1.7) and (2.5).*

Proof. The convergence results are proved in [69]. The measure $d\mu_N^\circ$ is invariant under (1.7), because the latter is a finite dimensional Hamiltonian system, and

$$d\mu_N^\circ(u_N) = \frac{1}{E_N} e^{-\mathcal{H}_N[u_N] - \mathcal{M}[u_N]} d\mathcal{L}_N(u_N)$$

is its Gibbs measure (weighted by another conserved quantity), where E_N is some positive constant, $\mathcal{H}_N$ and $\mathcal{M}$ are as in (1.10), and $d\mathcal{L}_N$ is the Lebesgue measure on the finite dimensional space $\mathcal{V}_N$.

To prove that $d\mu_N^\circ$ is invariant under (2.5), it suffices to show that it is preserved⁸ by the gauge transform $\mathcal{G}_N$. In fact, by (1.6) and (1.10) we know $\mathcal{H}_N[u_N] = \mathcal{H}_N[v_N]$ and $\mathcal{M}[u_N] = \mathcal{M}[v_N]$, so it suffices to prove that $\mathcal{G}_N$ preserves the Lebesgue measure $d\mathcal{L}_N$. Working in the coordinates $(r_k, \theta_k)_{\langle k \rangle \leq N}$ and $(r_k^*, \theta_k^*)_{\langle k \rangle \leq N}$, which are defined by $(u_N)_k = r_k e^{i\theta_k}$ and $(v_N)_k = r_k^* e^{i\theta_k^*}$, we can write the measure $d\mathcal{L}_N$ as

$$(2.18) \quad d\mathcal{L}_N = \prod_{\langle k \rangle \leq N} r_k dr_k d\theta_k.$$

If $v_N = \mathcal{G}_N u_N$, then we have $r_k^* = r_k$ and $\theta_k^* = \theta_k + F((r_j, \theta_j)_{\langle j \rangle \leq N})$, where F may also depend on t , but does not depend on k . Moreover, by (1.6) and (2.4)

⁸For fixed time t , we can view $\mathcal{G}_N$ as a mapping from $\mathcal{V}_N$ to itself, by requiring u_N to solve (1.7).

we know that F actually depends only on r_j and on the differences $\theta_j - \theta_\ell$, which are invariant under the mapping $\theta_k \mapsto \theta_k^*$. It then follows that the transformation $(r_k, \theta_k) \mapsto (r_k^*, \theta_k^*)$ preserves the measure (2.18), by a simple calculation of its Jacobian. $\square$

2.3. Function spaces and linear estimates. From now on we will work with equation (2.5) with the nonlinearity defined by (2.9) and (2.10), which has the form (2.16). Recall the well-known $X^{s,b}$ spaces (where b may be replaced by b_1 or b_2)

$$(2.19) \quad \|u\|_{X^{s,b}} = \|\langle k \rangle^s \langle \lambda \rangle^b \tilde{u}_k(\lambda)\|_{\ell_k^2 L_\lambda^2}.$$

We will mostly consider $s = 0$ and will denote $X^{0,b} = X^b$. In addition we introduce matrix norms that measure the functions $h = h_{kk^*}(t)$ and $\mathfrak{h} = \mathfrak{h}_{kk'}(t, s)$, namely,

$$(2.20) \quad \|h\|_{Y^b} = \|\langle \lambda \rangle^b \tilde{h}_{kk^*}(\lambda)\|_{\ell_{k^*}^2 \rightarrow \ell_k^2 L_\lambda^2}, \quad \|\mathfrak{h}\|_{Y^{b,b}} = \|\langle \lambda \rangle^b \langle \lambda' \rangle^{-b} \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda')\|_{\ell_{k'}^2 L_{\lambda'}^2 \rightarrow \ell_k^2 L_\lambda^2},$$

$$(2.21) \quad \|h\|_{Z^b} = \|\langle \lambda \rangle^b \tilde{h}_{kk^*}(\lambda)\|_{\ell_{k,k^*}^2 L_\lambda^2}, \quad \|\mathfrak{h}\|_{Z^{\tilde{b},b}} = \|\langle \lambda \rangle^{\tilde{b}} \langle \lambda' \rangle^{-b} \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda')\|_{\ell_{k,k'}^2 L_{\lambda,\lambda'}^2},$$

where $\tilde{b} \in \{b, b_1\}$, $\|\cdot\|_{\ell_{k^*}^2 \rightarrow \ell_k^2 L_\lambda^2}$ and $\|\cdot\|_{\ell_{k'}^2 L_{\lambda'}^2 \rightarrow \ell_k^2 L_\lambda^2}$ represent the operator norms of linear operators with the given kernels; for example,

$$(2.22) \quad \|\mathfrak{h}\|_{Y^{b,b}} = \sup \left\{ \left\| \sum_{k'} \int d\mu \cdot \langle \lambda \rangle^b \langle \lambda' \rangle^{-b} \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda') y_{k'}(\lambda') \right\|_{\ell_k^2 L_\lambda^2} : \|y_{k'}(\lambda')\|_{\ell_{k'}^2 L_{\lambda'}^2} = 1 \right\}.$$

By definition we can verify that

$$(2.23) \quad \|\mathfrak{h}\|_{Y^{b,b}} = \sup_{\|y\|_{X^b} = 1} \left\| \sum_{k'} \int dt' \cdot \mathfrak{h}_{kk'}(t, t') y_{k'}(t') \right\|_{X^b}.$$

For any of the above spaces, we can localize them in the standard way to a time interval J ,

$$(2.24) \quad \|u\|_{Z(J)} = \inf \{ \|v\|_Z : v \equiv u \text{ on } J \}.$$

We will need the following simple estimates.

PROPOSITION 2.5. *The norms $\|h\|_{Y^b}$ and $\|\mathfrak{h}\|_{Y^{b,b}}$ do not increase when $\tilde{h}$ or $\tilde{\mathfrak{h}}$ is multiplied by a function of (k, λ) , or a function of k^* (or (k', λ') for $\tilde{\mathfrak{h}}$), which is at most 1 in the $l^\infty L^\infty$ or l^∞ norms.*

Next, if H is defined by

$$(2.25) \quad \tilde{H}_{kk^*}(\lambda) = \sum_{k'} \int d\lambda' \cdot \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda') \tilde{h}_{k'k^*}(\lambda'),$$

where $\tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda')$ is supported in $|k - k'| \lesssim L$, then for any $\alpha > 0$, we have

$$(2.26) \quad \left\| \left(1 + \frac{|k - k^*|}{L} \right)^\alpha H \right\|_{Z^b} \lesssim \|\mathfrak{h}\|_{Y^{b,b}} \cdot \left\| \left(1 + \frac{|k' - k^*|}{L} \right)^\alpha h \right\|_{Z^b}.$$

Proof. The first statement follows directly from [definition \(2.22\)](#). Now let us prove [\(2.26\)](#). We may fix k^* , and by translation invariance, we may assume $k^* = 0$. Relabeling $\tilde{H}_{k0}(\lambda) =: \tilde{H}_k(\lambda)$ and $\tilde{h}_{k'0}(\lambda) =: \tilde{h}_{k'}(\lambda)$, we may decompose

$$\tilde{H}_k(\lambda) = \sum_{M \geq L} (\tilde{H}^M)_k(\lambda); \quad (\tilde{H}^M)_k(\lambda) = \begin{cases} \mathbf{1}_{|k| \sim M} \tilde{H}_k(\lambda), & M > L, \\ \mathbf{1}_{|k| \lesssim L} \tilde{H}_k(\lambda), & M = L, \end{cases}$$

and similarly for $\tilde{h}$, so that we have

$$(2.27) \quad \left\| \langle \lambda \rangle^b \left(1 + \frac{|k|}{L} \right)^\alpha \tilde{H}_k(\lambda) \right\|_{\ell_k^2 L_\lambda^2}^2 \sim \sum_{M \geq L} L^{-2\alpha} M^{2\alpha} \|\langle \lambda \rangle^b (\tilde{H}^M)_k(\lambda)\|_{\ell_k^2 L_\lambda^2}^2$$

and similarly for h . Since $\tilde{\mathfrak{h}}$ is supported in $|k - k'| \lesssim L$, we have

$$|(\tilde{H}^M)_k(\lambda)| \leq \sum_{M' \sim M} \left| \sum_{k'} \int d\lambda' \cdot \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda') (\tilde{h}^{M'})_{k'}(\lambda') \right|,$$

therefore

$$\begin{aligned} & \|\langle \lambda \rangle^b (\tilde{H}^M)_k(\lambda)\|_{\ell_k^2 L_\lambda^2}^2 \\ & \lesssim \|\langle \lambda \rangle^b \langle \lambda' \rangle^{-b} \tilde{\mathfrak{h}}_{kk'}(\lambda, \lambda')\|_{\ell_{k'}^2 L_{\lambda'}^2 \rightarrow \ell_k^2 L_\lambda^2}^2 \sum_{M' \sim M} \|\langle \lambda' \rangle^b (\tilde{h}^{M'})_{k'}(\lambda')\|_{\ell_{k'}^2 L_{\lambda'}^2}^2, \end{aligned}$$

which, combined with [\(2.27\)](#), implies [\(2.26\)](#). $\square$

Let χ be a smooth cutoff as in [Section 1.5](#), and define the time truncated Duhamel operator

$$(2.28) \quad \mathcal{IF}(t) := \chi(t) \int_0^t e^{i(t-t')\Delta} \chi(t') F(t') dt'.$$

LEMMA 2.6. *We have $2\mathcal{IF}(t) = \mathcal{JF}(t) - \chi(t)e^{it\Delta}\mathcal{JF}(0)$, where $\mathcal{J}$ is defined by*

$$(2.29) \quad \mathcal{JF}(t) = \chi(t) \left(\int_{-\infty}^t - \int_t^\infty \right) e^{i(t-t')\Delta} \chi(t') F(t') dt'.$$

Moreover, we have the formula

$$(2.30) \quad \widetilde{\mathcal{JF}}(k, \lambda) = \int_{\mathbb{R}} \mathcal{J}(\lambda, \mu) \widetilde{F}(k, \mu) d\mu, \quad |\partial_{\lambda, \mu}^\alpha \mathcal{J}(\lambda, \mu)| \lesssim_{\alpha, A} \frac{1}{\langle \lambda - \mu \rangle^A} \frac{1}{\langle \mu \rangle}.$$

For the proof of [Lemma 2.6](#), see the calculations in [\[28, Lemma 3.1\]](#).

PROPOSITION 2.7. *Let φ be any Schwartz function, and recall that $\varphi_\tau(t) = \varphi(\tau^{-1}t)$ for any $0 < \tau \ll 1$. Then for any $u = u_k(t)$ and $\mathfrak{h} = \mathfrak{h}_{kk'}(t, t')$, we have*

$$(2.31) \quad \|\varphi_\tau \cdot u\|_{X^{s,b}} \lesssim \tau^{b_1-b} \|u\|_{X^{s,b_1}}, \quad \|\varphi_\tau(t) \cdot \mathfrak{h}\|_{Z^{b,b}} \lesssim \tau^{b_1-b} \|\mathfrak{h}\|_{Z^{b_1,b}},$$

provided that $u_k(0) = \mathfrak{h}_{kk'}(0, t') = 0$.

Proof. Using the definition of the $Z^{\tilde{b},b}$ norms and fixing the (k', λ') variables, we can reduce the second inequality in (2.31) to the first, and by fixing k and conjugating by the linear Schrödinger flow, we can reduce the first to

$$\|\langle \xi \rangle^b (\widehat{\varphi_\tau} * \widehat{v})(\xi)\|_{L^2} \lesssim \tau^{b_1-b} \|\langle \eta \rangle^{b_1} \widehat{v}(\eta)\|_{L^2}$$

for v satisfying $v(0) = 0$. Let $\widehat{v} = g_1 + g_2$, where

$$g_1(\xi) = \mathbf{1}_{|\xi| \geq \tau^{-1}}(\xi) \widehat{v}(\xi), \quad g_2(\xi) = \mathbf{1}_{|\xi| < \tau^{-1}}(\xi) \widehat{v}(\xi).$$

We will prove that

$$(2.32) \quad \|\langle \xi \rangle^b (\widehat{\varphi_\tau} * g_j)(\xi)\|_{L^2} \lesssim \tau^{b_1-b} \|\langle \eta \rangle^{b_1} \widehat{v}(\eta)\|_{L^2}$$

for $j \in \{1, 2\}$. To prove (2.32) for $j = 1$, we can reduce it to the $L^2 \rightarrow L^2$ bound for the operator

$$g(\eta) \mapsto \int_{\mathbb{R}} R(\xi, \eta) g(\eta) d\eta, \quad R(\xi, \eta) = \mathbf{1}_{|\eta| \geq \tau^{-1}} \cdot \tau \widehat{\varphi}(\tau(\xi - \eta)) \frac{\langle \xi \rangle^b}{\langle \eta \rangle^{b_1}}.$$

Since

$$\mathbf{1}_{|\eta| \geq \tau^{-1}} \cdot \frac{\langle \xi \rangle^b}{\langle \eta \rangle^{b_1}} \lesssim \tau^{b_1-b} \frac{\langle T\xi \rangle^b}{\langle \tau\eta \rangle^{b_1}} \lesssim \tau^{b_1-b} \langle \tau(\xi - \eta) \rangle^b,$$

it follows from Schur's estimate that this $L^2 \rightarrow L^2$ bound is at most

$$\tau^{b_1-b} \|\tau \widehat{\varphi}(\tau\zeta) \langle \tau\zeta \rangle^b\|_{L^1_\xi} \lesssim \tau^{b_1-b},$$

which proves (2.32) for $j = 1$.

To prove (2.32) for $j = 2$, note that since $v(0) = 0$, we have $\int_{\mathbb{R}} \widehat{v}(\eta) d\eta = 0$, so

$$\begin{aligned} |(\widehat{\varphi_\tau} * g_2)(\xi)| &= \left| -\tau \widehat{\varphi}(\tau\xi) \int_{|\eta| \geq \tau^{-1}} \widehat{v}(\eta) d\eta \right. \\ &\quad \left. - \int_{|\eta| < \tau^{-1}} \tau \widehat{v}(\eta) [\widehat{\varphi}(\tau\xi) - \widehat{\varphi}(\tau(\xi - \eta))] d\eta \right| \\ &\lesssim \tau \langle \tau\xi \rangle^{-4} \int_{\mathbb{R}} \min(1, |\tau\eta|) |\widehat{v}(\eta)| d\eta, \end{aligned}$$

and by Hölder's inequality we have

$$\begin{aligned} &\int_{\mathbb{R}} \min(1, |\tau\eta|) |\widehat{v}(\eta)| d\eta \\ &\lesssim \|\langle \eta \rangle^{b_1} \widehat{v}(\eta)\|_{L^2} \cdot \|\min(1, |\tau\eta|) \langle \eta \rangle^{-b_1}\|_{L^2} \lesssim \tau^{b_1-\frac{1}{2}} \|\langle \eta \rangle^{b_1} \widehat{v}(\eta)\|_{L^2}. \end{aligned}$$

Using also the elementary bound

$$\|\tau \langle \tau \xi \rangle^{-4} \langle \xi \rangle^b\|_{L^2} \lesssim \tau^{\frac{1}{2}-b},$$

we deduce (2.32) for $j = 2$, and hence (2.31). $\square$

3. Structure of the solution: random averaging operators

We now fix a short time $0 < \tau \ll 1$, and focus on the local theory for (2.5), with initial data distributed according to the Gaussian measure $d\rho_N$, on $J := [-\tau, \tau]$. By definition, this is equivalent to considering (2.5) with random initial data $u_{\text{in}} = v_{\text{in}} = f(\omega)$, which we will assume from now on, until the end of Section 5. Most functions that appear in the proof will be random (i.e., depend on ω), whether or not we explicitly write ω in their expressions.

3.1. *The decomposition.* We start by writing down the ansatz of the solution to (2.5). Recall that the truncated mass m_N defined in (2.8) and the corresponding m_N^* are random variables given by

$$(3.1) \quad m_N = \sum_{\langle k \rangle \leq N} \frac{|g_k|^2}{\langle k \rangle^2}, \quad m_N^* = \sum_{\langle k \rangle \leq N} \frac{|g_k|^2 - 1}{\langle k \rangle^2}.$$

Note that they are Borel functions of $|g_k|^2$ for $\langle k \rangle \leq N$. Let $\nu_N := m_N^* - m_{\frac{N}{2}}^*$. By standard large deviation estimates, we have

$$(3.2) \quad \mathbb{P}(|\nu_N| \geq AN^{-1}) \leq Ce^{-C^{-1}A}$$

for any $A > 0$, where C is an absolute constant. In particular, by removing a set of measure $\leq C_\theta e^{-\tau^{-\theta}}$ (which will be done before proving any estimates) we may assume the following bounds, which are used below without any further mentioning:

$$(3.3) \quad |g_k| \lesssim \tau^{-\theta} \langle k \rangle^\theta, \quad |m_N^*| \lesssim \tau^{-\theta}, \quad |\nu_N| \lesssim \tau^{-\theta} N^{-1+\theta}.$$

Our goal here is to obtain a quantitative estimate for the difference $y_N := v_N - v_{\frac{N}{2}}$. By (2.5), this y_N satisfies the equation

$$(3.4) \quad \begin{cases} (i\partial_t + \Delta)y_N = \Pi_N \mathcal{Q}_N(y_N + v_{\frac{N}{2}}) - \Pi_{\frac{N}{2}} \mathcal{Q}_{\frac{N}{2}}(v_{\frac{N}{2}}), \\ y_N(0) = \Delta_N f(\omega). \end{cases}$$

By (2.9) we can rewrite the above equation as

$$(3.5) \quad \begin{cases} (i\partial_t + \Delta)y_N = \sum_{l=0}^r c_{rl} (m_N^*)^{r-l} \left\{ \Pi_N [\mathcal{N}_{2l+1}(y_N + v_{\frac{N}{2}}) \right. \\ \qquad \qquad \qquad \left. - \mathcal{N}_{2l+1}(v_{\frac{N}{2}})] + \Delta_N \mathcal{N}_{2l+1}(v_{\frac{N}{2}}) \right\} \\ \qquad \qquad \qquad + \sum_{j=0}^r c_{rl} [(m_{\frac{N}{2}}^* + \nu_N)^{r-l} - (m_{\frac{N}{2}}^*)^{r-l}] \cdot \Pi_{\frac{N}{2}} \mathcal{N}_{2l+1}(v_{\frac{N}{2}}), \\ y_N(0) = \Delta_N f(\omega), \end{cases}$$

where c_{rl} are constants that will not be important in the proof.

Define the set

$$(3.6) \quad \mathcal{K} := \{(N, L) \in (2\mathbb{Z})^2 : 2^{-1} \leq L < N^{1-\delta}\}.$$

For each $(N, L) \in \mathcal{K}$, we define the function $\psi_{N,L}$ as the solution to the (linear) equation

$$(3.7) \quad \begin{cases} (i\partial_t + \Delta)\psi_{N,L} = \sum_{l=0}^r (l+1) c_{rl} (m_N^*)^{r-l} \Pi_N \mathcal{N}_{2l+1}(\psi_{N,L}, v_L, \dots, v_L), \\ \psi_{N,L}(0) = \Delta_N f(\omega). \end{cases}$$

It is important to place $\psi_{N,L}$ in the first position of $\mathcal{N}_{2l+1}$ in (3.7); see Remark 3.5. By linearity, we have

$$(3.8) \quad (\psi_{N,L})_k = \sum_{k^*} H_{kk^*}^{N,L} \frac{g_{k^*}(\omega)}{\langle k^* \rangle},$$

where for $\frac{N}{2} < \langle k^* \rangle \leq N$ and $\langle k \rangle \leq N$, $H_{kk^*}^{N,L} = \varphi_k$ is the k -th mode of the solution φ to the equation

$$(3.9) \quad \begin{cases} (i\partial_t + \Delta)\varphi = \sum_{l=0}^r (l+1) c_{rl} (m_N^*)^{r-l} \Pi_N \mathcal{N}_{2l+1}(\varphi, v_L, \dots, v_L), \\ \varphi(0) = e^{ik^* \cdot x}, \end{cases}$$

and for other $(k, k^*) \in (\mathbb{Z}^2)^2$, define $H_{kk^*}^{N,L} = 0$. By definition these $H_{kk^*}^{N,L}$, as well as the $h_{kk^*}^{N,L}$ defined below, are $\mathcal{B}_{\leq N}$ measurable and $\mathcal{B}_{\leq L}^+$ measurable in the sense of Definition 1.11.

For any N , let L_0 be the largest L satisfying $(N, L) \in \mathcal{K}$. We further define

$$(3.10) \quad \zeta_{N,L} := \psi_{N,L} - \psi_{N, \frac{L}{2}}, \quad h^{N,L} := H^{N,L} - H^{N, \frac{L}{2}}; \quad z_N := y_N - \psi_{N, L_0}.$$

Note that $\psi_{N, \frac{1}{2}} = e^{it\Delta}(\Delta_N f(\omega))$, and that $H_{kk^*}^{N, \frac{1}{2}}$ is $e^{-i|k|^2 t} \mathbf{1}_{k=k^*}$ restricted to the frequency band $\frac{N}{2} < \langle k \rangle \leq N$. Moreover, z_N is $\mathcal{B}_{\leq N}$ measurable,

$z_N(0) = 0$, and satisfies the equation

(3.11)

$$\begin{aligned} (i\partial_t + \Delta)z_N &= \sum_{l=0}^r c_{rl}(m_N^*)^{r-l} \\ &\quad \cdot \Pi_N \left[\mathcal{N}_{2l+1}(z_N + \psi_{N,L_0} + v_{\frac{N}{2}}) - \mathcal{N}_{2l+1}(v_{\frac{N}{2}}) + \Delta_N \mathcal{N}_{2l+1}(v_{\frac{N}{2}}) \right] \\ &\quad - \sum_{l=0}^r c_{rl}(m_N^*)^{r-l} \cdot \Pi_N \left[(l+1) \mathcal{N}_{2l+1}(\psi_{N,L_0}, v_{L_0}, \dots, v_{L_0}) \right] \\ &\quad + \sum_{l=0}^r c_{rl} \left[(m_{\frac{N}{2}}^* + \nu_N)^{r-l} - (m_{\frac{N}{2}}^*)^{r-l} \right] \cdot \Pi_{\frac{N}{2}} \mathcal{N}_{2l+1}(v_{\frac{N}{2}}). \end{aligned}$$

Remark 3.1. With the above construction, if we let $v = \lim_{N \rightarrow \infty} v_N$ be the gauged version of the solution u to (1.1), we then have

$$(3.12) \quad v = e^{it\Delta} f(\omega) + \sum_{(N,L) \in \mathcal{K}} \zeta_{N,L} + z, \quad \text{where} \quad z = \sum_N z_N.$$

This is the ansatz (1.29) in Section 1.3.4, where $\zeta_{N,L}$ can be viewed as a random averaging operator $\mathcal{P}_{NL}$, whose kernel is essentially given by $h^{N,L}$, applied to the Gaussian free field $e^{it\Delta} f(\omega)$. There are, however, two differences: (1) our $\mathcal{P}_{NL}$ is not exactly the one in (1.26), but an infinite iteration of the latter, because (1.26) has no smoothing effect; and (2) our $\mathcal{P}_{NL}$ is not exactly a Borel function of $(g_k)_{\langle k \rangle \leq L}$ as it also depends on m_N^* , but as it turns out this does not affect any probabilistic estimates; see Lemma 4.1.

3.2. The a priori bounds. We now state the local well-posedness result for (2.5). Its proof will occupy the rest of this section and Sections 4 and 5.

PROPOSITION 3.2. *Recall the relevant constants defined in (1.35), and that $\tau \ll 1$, $J = [-\tau, \tau]$. Then, τ^{-1} -certainly, i.e., with probability bigger than or equal to $1 - C_\theta e^{-\tau^{-\theta}}$, the following estimates hold for all $(N, L) \in \mathcal{K}$:*

(3.13)

$$\|h^{N,L}\|_{Y^b(J)} \leq L^{-\delta_0}, \quad \|h^{N,L}\|_{Z^b(J)} \leq N^{\frac{1}{2} + \delta_{\frac{5}{4}}} L^{-\frac{1}{2}}, \quad \|z_N\|_{X^b(J)} \leq N^{-1+\gamma}.$$

3.2.1. The extensions. In proving Proposition 3.2 we will restrict z_N and $h^{N,L}$ to J and construct extensions of these restrictions that are defined for all time. This has to be done carefully so as to maintain the correct independence properties. We define these extensions inductively, as follows.

First, let $z_1^\dagger(t) := z_1(t)\chi_\tau(t)$ and $\psi_{N,\frac{1}{2}}^\dagger(t) := \chi(t)e^{it\Delta}(\Delta_N f(\omega))$, and define $H^{N,\frac{1}{2},\dagger}$ accordingly. Suppose $M \geq 1$ is a dyadic number and that we have defined $z_N^\dagger$ for all $N \leq M$ and $h^{N,L,\dagger}$ for all $(N, L) \in \mathcal{K}$ and $L < M$. Then for

$L \leq M$, we may define

$$v_L^\dagger = \sum_{L' \leq L} y_{L'}^\dagger, \quad \text{where} \quad y_L^\dagger = z_L^\dagger + \chi(t) e^{it\Delta} (\Delta_L f(\omega)) + \sum_{(L,R) \in \mathcal{K}} \zeta_{L,R}^\dagger,$$

which is acceptable since for $(L, R) \in \mathcal{K}$ we must have $R < L^{1-\delta} \leq M$, so $h^{L,R,\dagger}$ and $H^{L,R,\dagger}$ are well defined, hence $\zeta_{L,R}^\dagger$ and $\psi_{L,R}^\dagger$ can be defined by (3.8) and (3.10).

Next, for $(N, L) \in \mathcal{K}$ and $L = M$, we can define $H^{N,M,\dagger}$ such that for $\frac{N}{2} < \langle k^* \rangle \leq N$ and $\langle k \rangle \leq N$, $H_{kk^*}^{N,M,\dagger} = \varphi_k^\dagger$ is the k -th mode of the solution $\varphi^\dagger$ to the equation

$$(3.14) \quad \begin{aligned} \varphi^\dagger(t) &= \chi(t) e^{it\Delta} (e^{ik^* \cdot x}) - i\chi_\tau(t) \sum_{l=0}^r (l+1) c_{rl} (m_N^*)^{r-l} \\ &\quad \times \mathcal{I}\Pi_N \mathcal{N}_{2l+1}(\varphi^\dagger, v_M^\dagger, \dots, v_M^\dagger), \end{aligned}$$

provided this solution exists and is unique; otherwise simply define $H^{N,M,\dagger} = H^{N,M} \cdot \chi_\tau(t)$. This defines $H^{N,M,\dagger}$ and hence also $h^{N,M,\dagger}$, $\psi_{N,M}^\dagger$ and $\zeta_{N,M}^\dagger$.

Finally we will define z_{2M} . As $\psi_{2M,L_0}^\dagger$ is already defined, where $L_0 \leq M$ is the largest L such that $(2M, L) \in \mathcal{K}$, we can define $z_{2M}^\dagger$ to be the unique fixed point of the mapping

$$(3.15) \quad \begin{aligned} z \mapsto & -i\chi_\tau(t) \sum_{l=0}^r c_{rl} (m_{2M}^*)^{r-l} \cdot \mathcal{I}\Pi_{2M} \left\{ \mathcal{N}_{2l+1}(z + \psi_{2M,L_0}^\dagger + v_M^\dagger) - \mathcal{N}_{2l+1}(v_M^\dagger) \right\} \\ & + i\chi_\tau(t) \sum_{l=0}^r (l+1) c_{rl} (m_N^*)^{r-l} \cdot \mathcal{I}\Pi_{2M} \mathcal{N}_{2l+1}(\psi_{2M,L_0}^\dagger, v_{L_0}^\dagger, \dots, v_{L_0}^\dagger) \\ & - i\chi_\tau(t) \sum_{l=0}^r c_{rl} [(m_M^* + \nu_{2M})^{r-l} - (m_M^*)^{r-l}] \cdot \mathcal{I}\Pi_M \mathcal{N}_{2l+1}(v_M^\dagger) \\ & - i\chi_\tau(t) \sum_{l=0}^r c_{rl} (m_{2M}^*)^{r-l} \cdot \mathcal{I}\Delta_{2M} \mathcal{N}_{2j+1}(v_M^\dagger) \end{aligned}$$

on the set $\mathcal{Z} = \{z : \|z\|_{X^b} \leq (2M)^{-1+\gamma}\}$, provided that this mapping is a contraction mapping from $\mathcal{Z}$ to itself. If it is not a contraction mapping, then simply define $z_{2M}^\dagger = z_{2M} \cdot \chi_\tau(t)$. This completes the inductive construction. We may then easily verify the following:

- The $z_N^\dagger$ and $h^{N,L,\dagger}$ we constructed indeed coincide with z_N and $h^{N,L}$ on J .
- The $z_N^\dagger$ is supported in $\langle k \rangle \leq N$, and $h_{kk^*}^{N,L,\dagger}$ is supported in $\langle k \rangle \leq N$ and $\frac{N}{2} < \langle k^* \rangle \leq N$.

- The random variable $h^{N,L,\dagger}$ is $\mathcal{B}_{\leq L}^+$ measurable, and $z_N^\dagger$ and $h^{N,L,\dagger}$ are $\mathcal{B}_{\leq N}$ measurable.
- All the above are smooth and compactly supported in time $t \in [-2, 2]$.

We will prove [Proposition 3.2](#) by induction in M , but in the process we will need some auxiliary estimates. More precisely, we will prove the following result, which contains [Proposition 3.2](#).

PROPOSITION 3.3. *Recall the relevant constants defined in (1.35), and that $\tau \ll 1$. Consider the following statement, which we call $\text{Loc}(M)$ for $M \geq 1$: for any $(N, L) \in \mathcal{K}$ with $L < M$, we have*

$$(3.16) \quad \|h^{N,L,\dagger}\|_{Y^b} \leq L^{-\delta_0},$$

$$(3.17) \quad \|h^{N,L,\dagger}\|_{Z^b} \leq N^{\frac{1}{2}+\gamma_0} L^{-\frac{1}{2}},$$

$$(3.18) \quad \left\| \left(1 + \frac{|k - k^*|}{L} \right)^\kappa h_{kk^*}^{N,L,\dagger} \right\|_{Z^b} \leq N.$$

Define the operators⁹ (where $0 \leq l \leq r$)

$$(3.19)$$

$$\mathcal{P}^+(w) := \chi_\tau(t) \cdot \mathcal{I}\Pi_N [\mathcal{N}_{2l+1}(w, v_L^\dagger, \dots, v_L^\dagger) - \mathcal{N}_{2l+1}(w, v_{\frac{L}{2}}^\dagger, \dots, v_{\frac{L}{2}}^\dagger)],$$

$$(3.20)$$

$$\mathcal{P}^-(w) := \chi_\tau(t) \cdot \mathcal{I}\Pi_N [\mathcal{N}_{2l+1}(v_L^\dagger, w, v_L^\dagger, \dots, v_L^\dagger) - \mathcal{N}_{2l+1}(v_{\frac{L}{2}}^\dagger, w, v_{\frac{L}{2}}^\dagger, \dots, v_{\frac{L}{2}}^\dagger)].$$

Then for any $(N, L) \in \mathcal{K}$ as defined in (3.6) with $L < M$, we have

$$(3.21) \quad \|\mathcal{P}^\pm\|_{X^b \rightarrow X^b} \leq \tau^\theta L^{-\delta_0^{\frac{1}{2}}}.$$

Let the kernel of $\mathcal{P}^+$ be $\mathfrak{h}_{kk'}^{N,L}(t, t')$. Then for any $(N, L) \in \mathcal{K}$ with $L < M$, we have

$$(3.22) \quad \|\mathbf{1}_{|k|, |k'| \geq \frac{N}{4}} \cdot \mathfrak{h}_{kk'}^{N,L}(t, t')\|_{Z^{b,b}} \leq \tau^\theta N^{\frac{1}{2}+\gamma_0-\delta^3} L^{-\frac{1}{2}}.$$

Finally for any $N \leq M$, we have

$$(3.23) \quad \|z_N^\dagger\|_{X^b} \leq N^{-1+\gamma}.$$

Now suppose that the statement $\text{Loc}(M)$ is true for $\omega \in \Xi$, where Ξ is a set. Then the statement $\text{Loc}(2M)$ is true for $\omega \in \Xi'$ where Ξ' is another set such that $\mathbb{P}(\Xi \setminus \Xi') \leq C_\theta e^{-(\tau^{-1}M)^\theta}$. In particular, apart from a set of ω with probability $\leq C_\theta e^{-\tau^{-\theta}}$, the statement $\text{Loc}(M)$ is true for all M .

⁹In fact we will prove stronger bounds where the low frequency inputs in (3.19) are replaced by $v_{L_1}^\dagger, \dots, v_{L_{2r}}^\dagger$ with $\max(L_j) = L$, and similarly for (3.20). But for simplicity we will just write (3.19) and (3.20).

3.3. *The proof of Proposition 3.3: reduction to multilinear estimates.* The heart of the proof of Proposition 3.3 is a collection of (probabilistic) multilinear estimates for $\mathcal{N}_{2l+1}$. We will state them in Proposition 3.4 below and show that they imply Proposition 3.3. We leave the proof of Proposition 3.4 to Section 5.

PROPOSITION 3.4. *Recall the relevant constants defined in (1.35), and that $\tau \ll 1$. Let the multilinear form $\mathcal{N}_n$ be as in (2.16), where $1 \leq n \leq 2r+1$. We will also consider $\mathcal{N}_{n+1}$, in which we assume $\iota_1 = +$. For each $1 \leq j \leq n$, the input function $v^{(j)}$ satisfies one of the following:*

(i) *Type (G), where we define $L_j = 1$, and*

$$(3.24) \quad (\widetilde{v^{(j)}})_{k_j}(\lambda_j) = \mathbf{1}_{N_j/2 < \langle k_j \rangle \leq N_j} \frac{g_{k_j}(\omega)}{\langle k_j \rangle} \widehat{\chi}(\lambda_j).$$

(ii) *Type (C), where*

$$(3.25) \quad (\widetilde{v^{(j)}})_{k_j}(\lambda_j) = \sum_{N_j/2 < \langle k_j^* \rangle \leq N_j} h_{k_j k_j^*}^{(j)}(\lambda_j, \omega) \frac{g_{k_j^*}(\omega)}{\langle k_j^* \rangle},$$

with $h_{k_j k_j^}^{(j)}(\lambda_j, \omega)$ supported in the set $\{\langle k_j \rangle \leq N_j, \frac{N_j}{2} < \langle k_j^* \rangle \leq N_j\}$, $\mathcal{B}_{\leq N_j}$ measurable and $\mathcal{B}_{\leq L_j}^+$ measurable for some $L_j \leq N_j^{1-\delta}$, and satisfying the bounds*

$$(3.26) \quad \begin{aligned} \|\langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{\ell_{k_j^*}^2 \rightarrow \ell_{k_j}^2 L_{\lambda_j}^2} &\lesssim L_j^{-\delta_0}, \quad \|\langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{\ell_{k_j, k_j^*}^2 L_{\lambda_j}^2} \lesssim N_j^{\frac{1}{2} + \gamma_0} L_j^{-\frac{1}{2}}, \\ \left\| \langle \lambda_j \rangle^b \left(1 + \frac{|k_j - k_j^*|}{L_j}\right)^\kappa h_{k_j k_j^*}^{(j)}(\lambda_j) \right\|_{\ell_{k_j, k_j^*}^2 L_{\lambda_j}^2} &\lesssim N_j. \end{aligned}$$

(iii) *Type (D), where $(\widetilde{v^{(j)}})_{k_j}(\lambda_j)$ is supported in $\{|k_j| \lesssim N_j\}$, and satisfies*

$$(3.27) \quad \|\langle \lambda_j \rangle^b (\widetilde{v^{(j)}})_{k_j}(\lambda_j)\|_{\ell_{k_j}^2 L_{\lambda_j}^2} \lesssim N_j^{-(1-\gamma)}.$$

In each case, we will assume that derivatives in λ_j of these functions satisfy the same bounds. This can always be guaranteed, since in practice everything will be compactly supported in time.

Assume for $n_1 \leq n$ that $v^{(j)}$ are of type (D) for $n_1 + 1 \leq j \leq n$, and of type (G) or (C) for $1 \leq j \leq n_1$. Let $\mathcal{G}$ and $\mathcal{C}$ be the sets of j such that $v^{(j)}$ are of type (G) and (C) respectively, similarly denote by $\mathcal{D} := \{n_1 + 1, \dots, n\}$. Let $N^{(j)} = \max^{(j)}(N_1, \dots, N_n)$ as before, and let $1 \leq a \leq n$ be such that $N^{(1)} \sim N_a$. Given $N_ \geq 1$, the following hold $\tau^{-1}N_*$ -certainly. We emphasize that the exceptional set of ω removed does not depend on the choice of the functions $v_j(j \geq n_1 + 1)$.*

(1) If $a \geq n_1 + 1$ (say $a = n$) and $N_* \gtrsim N^{(2)}$, then we have (recall $b_1 = b + \delta^4$)

$$(3.28) \quad \|\mathcal{IN}_n(v^{(1)}, \dots, v^{(n)})\|_{X^{b_1}} \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} (N^{(1)})^{-1+\gamma} (N^{(2)})^{-\delta_0^{\frac{1}{3}}}.$$

Here the exceptional set does not depend on $N^{(1)}$.

(2) If $a \leq n_1$ and $N_* \gtrsim N^{(1)}$, then we have

$$(3.29) \quad \|\mathcal{IN}_n(v^{(1)}, \dots, v^{(n)})\|_{X^{b_1}} \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} (N^{(1)} N^{(2)})^{-\frac{1}{2}(1-\gamma_0)}.$$

Moreover, if $\iota_a = -$, then we have the stronger bound

$$(3.30) \quad \|\mathcal{IN}_n(v^{(1)}, \dots, v^{(n)})\|_{X^{b_1}} \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} (N^{(1)})^{-(1-\gamma_0)}.$$

Moreover, if $\iota_a = +$ and $N^{(2)} \lesssim (N^{(1)})^{1-\delta}$, then we have stronger bound for the projected term

$$(3.31) \quad \|\mathcal{I}\Pi_{N^{(1)}}^\perp \mathcal{N}_n(v^{(1)}, \dots, v^{(n)})\|_{X^{b_1}} \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} (N^{(1)})^{-(1-\frac{4\gamma}{5})}.$$

(3) Now consider the operator

$$(3.32) \quad \mathcal{Q}^+(w) := \mathcal{I}\Pi_{N_0} \mathcal{N}_{n+1}(w, v^{(1)}, \dots, v^{(n)}),$$

and let its kernel be $\mathfrak{h}_{kk'}(t, t')$. If $N^{(1)} \lesssim N_0^{1-\delta}$ and $N_* \gtrsim N_0$, then we have

$$(3.33) \quad \|\mathbf{1}_{|k|, |k'| \geq \frac{N_0}{4}} \cdot \mathfrak{h}_{kk'}(t, t')\|_{Z^{b_1, b}} \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} N_0^{\frac{1}{2}} (N^{(1)})^{-\frac{1}{2}+\gamma_0}.$$

Remark 3.5. The improvement (3.31) is due to the exact projection $\Pi_{N^{(1)}}^\perp$. In fact this implies that in the expression (2.16) there exist some $1 \leq a \leq n$ and some Γ , namely $\Gamma = (N^{(1)})^2 - 1$, such that

$$(3.34) \quad |k|^2 \geq \Gamma \geq |k_a|^2 \quad \text{or} \quad |k|^2 \leq \Gamma \leq |k_a|^2, \quad \text{and } N^{(1)} \sim N_a.$$

We call (3.34) the Γ -condition. If we put some other projections in $\mathcal{N}_n$ that also guarantee (3.34), for example $\Pi_M \mathcal{N}_n(\dots, \Pi_M^\perp v^{(a)}, \dots)$ where $N^{(1)} \sim N_a$, then the same improvement (3.31) will remain true.

In the proof below we will see that the Γ condition provides the needed improvements in the case $N^{(1)} \sim N_a$ and $\iota_a = +$. This is the reason why we place $\psi_{N, L}$ in the first position of $\mathcal{N}_{2l+1}$ in (3.7). On the other hand, the term where $\psi_{N, L}$ is placed in the second position can be handled using the improvement (3.30).

Proof of Proposition 3.3 assuming Proposition 3.4. To prove the statement $\text{Loc}(2M)$ we start with (3.21) and (3.22), and we may assume $L = M$. The proof for $\mathcal{P}^-$ in (3.21) is similar, so let us consider $\mathcal{P}^+$. Since $v_M^\dagger = \sum_{L' \leq M} y_{L'}^\dagger$, by definition we can write $\mathcal{P}^+$ as a superposition of forms

$$w \mapsto \chi_\tau(t) \cdot \mathcal{I}\Pi_N \mathcal{N}_{2l+1}(w, y_{N_2}^\dagger, \dots, y_{N_{2l+1}}^\dagger),$$

where $\max(N_2, \dots, N_{2l+1}) = M$. As we have the decomposition

$$(3.35) \quad y_{N_j}^\dagger = \chi(t) e^{it\Delta} (\Delta_{N_j} f(\omega)) + \sum_{L_j: (N_j, L_j) \in \mathcal{K}} \zeta_{N_j, L_j}^\dagger + z_{N_j}^\dagger,$$

by $\text{Loc}(M)$ we know that each $y_{N_j}^\dagger$ can be decomposed into terms of type (G), type (C) (corresponding to some $L_j \lesssim N_j^{1-\delta}$), and type (D). The bound (3.21) is then a consequence of (2.31) and (3.28), after removing a set of ω with measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$ that is independent of N . Note that by (2.23), the $L = M$ case of (3.21) is equivalent to

$$(3.36) \quad \|\mathfrak{h}^{N,M}\|_{Y^{b,b}} \leq \tau^\theta M^{-\delta_0^{\frac{1}{2}}}.$$

Similarly (3.22) follows from (2.31) and (3.33), because we have

$$\tau^{b_1-b} \tau^{-\theta} N^{C\kappa^{-1}} N^{\frac{1}{2}} M^{-\frac{1}{2}+\gamma_0} \ll \tau^\theta N^{\frac{1}{2}+\gamma_0-2\delta^3} M^{-\frac{1}{2}}$$

using the fact that $M \lesssim N^{1-\delta}$. The set of ω removed here will depend on N , but it will have measure $\leq C_\theta e^{-(\tau^{-1}N)^\theta}$, so summing in $N \geq M$ we still get a set of measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$.

Next we prove (3.16)–(3.18), again assuming $L = M$. By (3.21) and $\text{Loc}(M)$ we already know that the right-hand side of (3.14) gives a contraction mapping in X^b , so (3.14) does have a unique solution. Subtracting equation (3.14) with M and with $\frac{M}{2}$ instead of M , we deduce that

$$(3.37) \quad \begin{aligned} h_{kk^*}^{N,M,\dagger}(t) = & -i \sum_{l=0}^r (l+1) c_{rl} (m_N^*)^{r-l} \left\{ \sum_{L \leq M} \sum_{k'} \int dt' \cdot \mathfrak{h}_{kk'}^{N,L}(t, t') h_{k'k^*}^{N,M,\dagger}(t') \right. \\ & \left. + \sum_{L < M} \sum_{k'} \int dt' \cdot \mathfrak{h}_{kk'}^{N,M}(t, t') h_{k'k^*}^{N,L,\dagger}(t') + \int dt' \cdot \mathfrak{h}_{kk^*}^{N,M}(t, t') H_{k^*k^*}^{N,\frac{1}{2},\dagger}(t') \right\}, \end{aligned}$$

where $\mathfrak{h}_{kk'}^{N,L}(t, t')$ is the kernel corresponding to $\mathcal{P}^+$ in the (k, k', t, t') variables. Recall that we are already in a set where (3.3) is true, which allows us to control m_N^* . Now by the definition of Y^b and $Y^{b,b}$ norms, the statement $\text{Loc}(M)$ and (3.36), we conclude that

$$\begin{aligned} \|h^{N,M,\dagger}\|_{Y^b} & \lesssim \sum_{L \leq M} \|\mathfrak{h}^{N,L}\|_{Y^{b,b}} \cdot \|h^{N,M,\dagger}\|_{Y^b} + \|\mathfrak{h}^{N,M}\|_{Y^{b,b}} \left(\sum_{L < M} \|h^{N,L,\dagger}\|_{Y^b} + 1 \right) \\ & \lesssim \|h^{N,M,\dagger}\|_{Y^b} \cdot \sum_{L \leq M} \tau^\theta L^{-\delta_0^{\frac{1}{2}}} \\ & \quad + \sum_{L < M} \tau^\theta M^{-\delta_0^{\frac{1}{2}}} L^{-\delta_0} \lesssim \tau^\theta \|h^{N,M,\dagger}\|_{Y^b} + \tau^\theta M^{-\delta_0^{\frac{1}{2}}}, \end{aligned}$$

which implies (3.16) as desired. In the same way we can prove (3.18) by using (2.26), noting that $\widetilde{\mathfrak{h}^{N,L}}$ is supported in $|k - k'| \lesssim L$.

As for (3.17), recall that for

$$\widetilde{H}_{kk^*}(\lambda) = \sum_{k'} \int d\lambda' \cdot \mathfrak{h}_{kk'}(\lambda, \lambda') h_{k'k^*}(\lambda')$$

we have, by definition of the relevant norms, that

$$\|H\|_{\ell_{k,k^*}^2 L_\lambda^2} \leq \min(\|\mathfrak{h}\|_{\ell_{k,k'}^2 L_{\lambda,\lambda'}^2}, \|h\|_{\ell_{k^*}^2 \rightarrow \ell_{k'}^2 L_{\lambda'}^2}, \|\mathfrak{h}\|_{\ell_{k,k'}^2 \rightarrow L_{\lambda,\lambda'}^2}, \|h\|_{\ell_{k',k^*}^2 L_{\lambda'}^2}).$$

Now in (3.37) we may assume $|k - k^*| \leq 2^{-10}N$ and $|k' - k^*| \leq 2^{-10}N$ (otherwise the bound follows trivially from (3.18) which we just proved) so, in particular, $|k|, |k'| \geq \frac{N}{4}$ as $|k^*| \geq \frac{N}{2}$. Using the statement $\text{Loc}(M)$ and (3.22) we get

$$\begin{aligned} \|h^{N,M,\dagger}\|_{Z^b} &\lesssim \sum_{L \leq M} \|\mathfrak{h}^{N,L}\|_{Y^{b,b}} \cdot \|h^{N,M,\dagger}\|_{Z^b} \\ &\quad + \|\mathbf{1}_{|k|,|k'| \geq \frac{N}{4}} \cdot \mathfrak{h}_{kk'}^{N,M}(t, t')\|_{Z^{b,b}} \left(\sum_{L < M} \|h^{N,L,\dagger}\|_{Y^b} + 1 \right) \\ &\lesssim \|h^{N,M,\dagger}\|_{Z^b} \cdot \sum_{L \leq M} \tau^\theta L^{-\delta_0^{\frac{1}{2}}} + \sum_{L < M} \tau^\theta N^{\frac{1}{2} + \gamma_0 - \delta^3} M^{-\frac{1}{2}} L^{-\delta_0} \\ &\lesssim \tau^\theta \|h^{N,M,\dagger}\|_{Y^b} + \tau^\theta N^{\frac{1}{2} + \gamma_0 - \delta^3} M^{-\frac{1}{2}}, \end{aligned}$$

which proves (3.17).

Finally we prove (3.23) with $L = 2M$, by showing that the mapping defined in (3.15) is indeed a contraction mapping from the given set $\mathcal{Z} = \{z : \|z\|_{X^b} \leq (2M)^{-1+\gamma}\}$ to itself. Actually we will only prove that this mapping sends $\mathcal{Z}$ to $\mathcal{Z}$, as the difference estimate is done in the same way.

We will separate the right-hand side of (3.15) into six groups, each of which has the form

$$\chi_\tau(t) \cdot (m_{2M}^*)^{r-l} \mathcal{I} \Pi_{2M} \mathcal{N}_{2l+1}(v^{(1)}, \dots, v^{(2l+1)}),$$

where

- (a) at least two of the $v^{(j)}$ are equal to $z + \psi_{2M,L_0}^\dagger$, and others are either $z + \psi_{2M,L_0}^\dagger$ or $v_M^\dagger$;
- (b) we have $v^{(2)} = \psi_{2M,L_0}^\dagger$, and all others equal $v_M^\dagger$;
- (c) one of $v^{(1)}$ or $v^{(2)}$ equals z , and all others equal $v_M^\dagger$;
- (d) we have $v^{(1)} = \psi_{2M,L_0}^\dagger$, another $v^{(j)}$ equals $v_M^\dagger - v_{L_0}^\dagger$, and all others equal either $v_M^\dagger$ or $v_{L_0}^\dagger$;
- (e) the factor $(m_{2M}^*)^{r-l}$ is replaced by $(m_M^* + \nu_{2M})^{r-l} - (m_M^*)^{r-l}$ and all $v^{(j)}$ equal $v_M^\dagger$;
- (f) same as (a), but with Δ_{2M} instead of Π_{2M} , and all $v^{(j)}$ equal $v_M^\dagger$.

By (2.31), it suffices to prove that each of these terms in (a) through (f), but without the $\chi_\tau(t)$ factor, is bounded in X^{b_1} by $\tau^{-\theta}(2M)^{-1+\gamma}$. Let one such term be denoted by $\mathcal{M}$, and notice that we can decompose

$$v_M^\dagger = \sum_{L \leq M} y_L^\dagger, \quad v_{L_0}^\dagger = \sum_{L \leq L_0} y_L^\dagger, \quad v_M - v_{L_0} = \sum_{L_0 < L \leq M} y_L^\dagger,$$

$$\psi_{2M, L_0}^\dagger = \chi(t) e^{it\Delta}(\Delta_{2M} f(\omega)) + \sum_{L \leq L_0} \zeta_{2M, L}.$$

Moreover, by what we have proved so far, we know that $y_L^\dagger$ for $L \leq M$ can be decomposed into terms of types (G), (C) and (D), and that $\chi(t) e^{it\Delta}(\Delta_{2M} f(\omega))$ is of type (G), $\zeta_{2M, L}$ is of type (C), and z is of type (D). By such decomposition we can reduce $\mathcal{M}$ to the terms studied in Proposition 3.4, with various choices of N_j and L_j . We now proceed case by case.

Case (a): Here we have at least two inputs $v^{(j)}$ with $N_j = 2M$, so by either (3.28) or (3.29) we can bound

$$\|\mathcal{M}\|_{X^{b_1}} \lesssim \tau^{-\theta} (2M)^{-1+\gamma_0+C\kappa^{-1}}$$

by removing a set of measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$, which suffices.

Case (b): Here we have $v^{(2)} = N^{(1)} = 2M$, while $\iota_2 = -$. By (3.30) we have the same bound as above.

Case (c): This term, *with the $\chi_\tau(t)$ factor*, can be written as

$$\sum_{L \leq M} \mathcal{P}_L^\pm(z),$$

where $\mathcal{P}_L^\pm$ are defined as in (3.19) and (3.20). (Such expressions can be defined even if $(2M, L) \notin \mathcal{K}$, and we use subscript L to indicate L dependence.) If $L \leq L_0$, then by (3.21) we can bound

$$\|\chi_\tau(t) \cdot \mathcal{M}\|_{X^b} \leq (2M)^{-1+\gamma} \sum_{L \leq L_0} \tau^\theta L^{-\delta_0^{\frac{1}{2}}} \lesssim \tau^\theta (2M)^{-1+\gamma},$$

which suffices. Note that here no further set of ω needs to be removed. If $L > L_0$, then this term can be bounded in the same way as in case (d) below, by removing a set of measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$.

Case (d): Here we have, due to the factor $v_M^\dagger - v_{L_0}^\dagger$, that $N^{(1)} = 2M$ and $N^{(2)} \gtrsim M^{1-\delta}$; so by either (3.28) or (3.29), we can bound

$$\|\mathcal{M}\|_{X^{b_1}} \lesssim \tau^{-\theta} M^{-1+\frac{\delta}{2}+C\gamma_0}$$

by removing a set of measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$, which suffices.

Case (e): The bound for this term follows from the bound for ν_{2M} and the trivial bound (say (3.28) or (3.29)) for the $\mathcal{N}_{2l+1}$ term.

Case (f): We may assume $N^{(1)} = N_a \sim M$. If either $v^{(a)}$ is of type (D) or $N^{(2)} \gtrsim M^{1-\delta}$ or $\iota_a = -$, we can reduce to one of the previous cases (namely (c) or (d) or (b)) and close as before; if $v^{(a)}$ is of type (G) or (C), $\iota_a = +$ and $N^{(2)} \ll M^{1-\delta}$, then the Δ_{2M} projection allows us to apply the improvement (3.31), which leads to

$$\|\mathcal{M}\|_{X^{b_1}} \lesssim \tau^{-\theta} M^{-1+\frac{4}{5}\gamma+C\kappa^{-1}}$$

by removing a set of measure $\leq C_\theta e^{-(\tau^{-1}M)^\theta}$, which suffices. This completes the proof. $\square$

4. Large deviation and counting estimates

Proposition 3.4 will be proved in [Section 5](#). In this section we make some preparations for the proof, namely we introduce two large deviation estimates and some counting estimates for integer lattice points.

4.1. *Large deviation estimates.* We first prove the following large deviation estimate for multilinear Gaussians, which as far as we know is new.

LEMMA 4.1. *Let $E \subset \mathbb{Z}^2$ be a finite subset, and let $\mathcal{B}$ be the σ -algebra generated by $\{g_k : k \in E\}$. Let $\mathcal{C}$ be a σ -algebra independent with $\mathcal{B}$, and let $\mathcal{C}^+$ be the smallest σ -algebra containing both $\mathcal{C}$ and the σ -algebra generated by $\{|g_k|^2 : k \in E\}$. Consider the expression*

$$(4.1) \quad F(\omega) = \sum_{(k_1, \dots, k_n) \in E^n} a_{k_1 \dots k_n}(\omega) \prod_{j=1}^n g_{k_j}(\omega)^{\iota_j},$$

where $n \leq 2r+1$, $\iota_j \in \{\pm\}$ and the coefficients $a_{k_1 \dots k_n}(\omega)$ are $\mathcal{C}^+$ measurable. Let $A \geq \#E$. Then A -certainly we have

$$(4.2) \quad |F(\omega)| \leq A^\theta M(\omega)^{\frac{1}{2}},$$

where

$$(4.3) \quad M(\omega) = \sum_{(X,Y)} \sum_{(k_m): m \notin X \cup Y} \left(\sum_{\text{pairing } (k_{i_s}, k_{j_s}): 1 \leq s \leq p} |a_{k_1 \dots k_n}(\omega)| \right)^2.$$

In the summation (4.3) we require that all $k_j \in E$, and that $X := \{i_1, \dots, i_p\}$ and $Y := \{j_1, \dots, j_p\}$ are two disjoint subsets of $\{1, 2, \dots, n\}$. Recall also the definition of pairing in [Definition 1.10](#).

Proof. Write in polar coordinates $g_k(\omega) = \rho_k(\omega)\eta_k(\omega)$ where $\rho_k = |g_k|$ and $\eta_k = \rho_k^{-1}g_k$. Then all the ρ_k and η_k are independent, and each η_k is uniformly distributed on the unit circle of $\mathbb{C}$. We may write

$$(4.4) \quad F(\omega) = \sum_{(k_1, \dots, k_n) \in E^n} b_{k_1 \dots k_n}(\omega) \prod_{j=1}^n \eta_{k_j}(\omega)^{\ell_j}, \quad b_{k_1 \dots k_n}(\omega) := a_{k_1 \dots k_n}(\omega) \prod_{j=1}^n \rho_{k_j}(\omega).$$

Since $a_{k_1 \dots k_n}(\omega)$ are $\mathcal{C}^+$ measurable, we know that the collection $\{b_{k_1 \dots k_n}\}$ is independent with the collection $\{\eta_k : k \in E\}$. The goal is to prove that

$$(4.5) \quad \mathbb{P}(|F(\omega)| \geq BM_1(\omega)^{\frac{1}{2}}) \leq Ce^{-B^{1/n}},$$

where C is an absolute constant, and $M_1(\omega)$ is the same as $M(\omega)$ but with the coefficients a replaced by the coefficients b . In fact, as $A \geq \#E$, we have A -certainly that $|b_{k_1 \dots k_n}(\omega)| \leq A^\theta |a_{k_1 \dots k_n}(\omega)|$, so (4.5) implies the desired bound.

We now prove (4.5). By independence, we may condition on the σ -algebra generated by $\{b_{k_1 \dots k_n}\}$ and prove (4.5) for the conditional probability, and then take another expectation; therefore we may assume that $b_{k_1 \dots k_n}$ are constants (so $M_1(\omega) = M_1$ is a constant). Now let $\{h_k : k \in E\}$ be another set of independent and identically distributed normalized complex Gaussian random variables, and define

$$(4.6) \quad G = \sum_{(k_1, \dots, k_n) \in E^n} |b_{k_1 \dots k_n}| \prod_{j=1}^n h_{k_j}^{\ell_j}.$$

We want to compare F and G and show $\mathbb{E}|F|^{2d} \leq \mathbb{E}|G|^{2d}$ for any positive integer d . In fact,

$$(4.7) \quad \mathbb{E}(|F|^{2d}) = \sum_{(k_j^i, \ell_j^i : 1 \leq i \leq d, 1 \leq j \leq n)} \prod_{i=1}^d b_{k_1^i \dots k_n^i} \overline{b_{\ell_1^i \dots \ell_n^i}} \mathbb{E} \left(\prod_{i=1}^d \prod_{j=1}^n \eta_{k_j^i}^{\ell_j^i} \overline{\eta_{\ell_j^i}^{\ell_j^i}} \right),$$

$$(4.8) \quad \mathbb{E}(|G|^{2d}) = \sum_{(k_j^i, \ell_j^i : 1 \leq i \leq d, 1 \leq j \leq n)} \prod_{i=1}^d |b_{k_1^i \dots k_n^i}| |b_{\ell_1^i \dots \ell_n^i}| \mathbb{E} \left(\prod_{i=1}^d \prod_{j=1}^n h_{k_j^i}^{\ell_j^i} \overline{h_{\ell_j^i}^{\ell_j^i}} \right).$$

The point is that we always have

$$\left| \mathbb{E} \left(\prod_{i=1}^d \prod_{j=1}^n \eta_{k_j^i}^{\ell_j^i} \overline{\eta_{\ell_j^i}^{\ell_j^i}} \right) \right| \leq \text{Re} \mathbb{E} \left(\prod_{i=1}^d \prod_{j=1}^n h_{k_j^i}^{\ell_j^i} \overline{h_{\ell_j^i}^{\ell_j^i}} \right).$$

In fact, by collecting all different factors we can write the expectations as

$$\mathbb{E} \left(\prod_{\alpha} (\eta_{k(\alpha)})^{x_\alpha} (\overline{\eta_{k(\alpha)}})^{y_\alpha} \right) \quad \text{and} \quad \mathbb{E} \left(\prod_{\alpha} (h_{k(\alpha)})^{x_\alpha} (\overline{h_{k(\alpha)}})^{y_\alpha} \right),$$

where the $k^{(\alpha)}$ are pairwise distinct. If $x_\alpha \neq y_\alpha$ for some α , both expectations will be 0; if $x_\alpha = y_\alpha$ for each α , then the first expectation will be 1 and the second expectation will be $\prod_\alpha x_\alpha! \geq 1$.

Now, since G is an exact multilinear Gaussian expression, by the standard hypercontractivity estimate (see [69]), we have

$$\mathbb{E}|F|^{2d} \leq \mathbb{E}|G|^{2d} \leq (2d-1)^{nd} (\mathbb{E}|G|^2)^d,$$

so for any $D > 0$ by using Chebyshev's inequality and optimizing in d , we have

$$\mathbb{P}(|F(\omega)| \geq D) \leq \min_d \left\{ (2d-1)^{nd} \left(\frac{\mathbb{E}|G|^2}{D^2} \right)^d \right\} \leq C \exp \left\{ \frac{-1}{2e} \left(\frac{D^2}{\mathbb{E}|G|^2} \right)^{\frac{1}{n}} \right\}$$

with some constant C depending only on n . It then suffices to prove $\mathbb{E}|G|^2 \lesssim M_1$ with constants depending only on n .

By dividing the sum (4.6) into finitely many terms and rearranging the subscripts, we may assume

$k_1 = \dots = k_{j_1}$, $k_{j_1+1} = \dots = k_{j_2}$, $\dots$, $k_{j_{r-1}+1} = \dots = k_{j_r}$, $1 \leq j_1 < \dots < j_r = n$, and the k_{j_s} are different for $1 \leq s \leq r$. Such a monomial that appears in (4.6) has the form

$$\prod_{s=1}^r h_{k_{j_s}}^{\beta_s} (\overline{h_{k_{j_s}}})^{\gamma_s}, \quad \beta_s + \gamma_s = j_s - j_{s-1} \quad (j_0 = 0),$$

where the factors for different s are independent. We may also assume $\beta_s = \gamma_s$ for $1 \leq s \leq q$ and $\beta_s \neq \gamma_s$ for $q+1 \leq s \leq r$, and that ι_j has the same sign as $(-1)^j$ for $1 \leq j \leq j_q$. Then we can further rewrite this monomial as a linear combination of

$$\prod_{s=1}^p \beta_s! \prod_{s=p+1}^q (|h_{k_{j_s}}|^{2\beta_s} - \beta_s!) \prod_{s=q+1}^r h_{k_{j_s}}^{\beta_s} (\overline{h_{k_{j_s}}})^{\gamma_s}$$

for $1 \leq p \leq q$. Thus, G is a finite linear combination of expressions of the form

$$\sum_{k_{j_1}, \dots, k_{j_r}} |b_{k_{j_1}, \dots, k_{j_1}, \dots, k_{j_r}, \dots, k_{j_r}}| \prod_{s=1}^p \beta_s! \prod_{s=p+1}^q (|h_{k_{j_s}}|^{2\beta_s} - \gamma_s!) \prod_{s=q+1}^r h_{k_{j_s}}^{\beta_s} (\overline{h_{k_{j_s}}})^{\gamma_s}.$$

Due to independence and the fact that $\mathbb{E}(|h|^{2\beta} - \beta!) = \mathbb{E}(h^\beta (\overline{h})^\gamma) = 0$ for a normalized Gaussian h and $\beta \neq \gamma$, we conclude that

$$\mathbb{E}|G|^2 \lesssim \sum_{k_{j_{p+1}}, \dots, k_{j_r}} \left(\sum_{k_{j_1}, \dots, k_{j_p}} |b_{k_{j_1}, \dots, k_{j_1}, \dots, k_{j_r}, \dots, k_{j_r}}| \right)^2,$$

which is bounded by M_1 choosing $X = \{1, 3, \dots, j_p-1\}$ and $Y = \{2, 4, \dots, j_p\}$, since by our assumptions (k_{2i-1}, k_{2i}) is a pairing for $2i \leq j_p$. This completes the proof. $\square$

For the purpose of [Section 5](#) we will also need the following lemma, which is a more general large deviation-type estimate restricted to the no-pairing case.

LEMMA 4.2. *Let δ be as in [Section 1.5](#), let $n \leq 2r + 1$, and consider the following expression:*

$$(4.9) \quad M(\omega) = \sum_{(k_1, \dots, k_n)} \sum_{(k_1^*, \dots, k_n^*)} \int d\lambda_1 \cdots d\lambda_n \cdot a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n) \cdot \prod_{j=1}^n g_{k_j^*}(\omega)^{\iota_j} h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^{\pm},$$

where $a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n)$ is a given (deterministic) function of $(k_1, \dots, k_n)$ and $(\lambda_1, \dots, \lambda_n)$. Moreover, in the summation we assume that there are no pairings among $\{k_1^*, \dots, k_n^*\}$, that $\langle k_j \rangle \leq N_j$ and $\frac{N_j}{2} < \langle k_j^* \rangle \leq N_j$, and that $h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)$, as a random variable, is $\mathcal{B}_{\leq N_j^{1-\delta}}^+$ measurable. Let $N_* \geq \max(N_1, \dots, N_n)$. Then N_* -certainly (the exceptional set removed will depend on the coefficients a) we have

$$(4.10) \quad |M(\omega)| \lesssim (N_*)^\theta \prod_{j=1}^n \|h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)\|_{\ell_{k_j^*}^2 \rightarrow \ell_{k_j}^2 L_{\lambda_j}^2} \cdot \|a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n)\|_{\mathcal{L}},$$

where $\mathcal{L}$ is an auxiliary norm defined (where $\partial_\lambda = (\partial_{\lambda_1}, \dots, \partial_{\lambda_n})$) by

$$(4.11) \quad \|a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n)\|_{\mathcal{L}}^2 := \sum_{k_1, \dots, k_n} \int d\lambda_1 \cdots d\lambda_n \cdot \left(\max_{1 \leq j \leq n} \langle \lambda_j \rangle \right)^{\delta_6} (|a|^2 + |\partial_\lambda a|^2).$$

Proof. Consider the big box $\{|\lambda_j| \leq (N_*)^{\delta^{-7}}\}$, and divide it into small boxes of size $(N_*)^{-\delta^{-1}}$. By exploiting the weight $(\max_{1 \leq j \leq n} \langle \lambda_j \rangle)^{\delta_6}$ in (4.11) and using Poincaré's inequality, we can find a function b that is supported in the big box and is constant on each small box, such that

$$(4.12) \quad \sup_{k_j, k_j^*} \|a - b\|_{L_{\lambda_1, \dots, \lambda_n}^2} \lesssim (N_*)^{-\delta^{-1}} \|a\|_{\mathcal{L}}.$$

Exploiting this $(N_*)^{-\delta^{-1}}$ gain, summing over $|k_j|, |k_j^*| \lesssim N_*$ and using the simple bound

$$\begin{aligned} & \sup_{k_j, k_j^*} \left| \int d\lambda_1 \cdots d\lambda_n \cdot a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n) \prod_{j=1}^n h_{k_j k_j^*}^{(j)}(\lambda_j)^\pm \right| \\ & \lesssim \sup_{k_j, k_j^*} \|a\|_{L_{\lambda_1, \dots, \lambda_n}^2} \prod_{j=1}^n \sup_{k_j, k_j^*} \|h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{L_{\lambda_j}^2} \end{aligned}$$

suffices to bound the contribution with a replaced by $a - b$; thus we may now replace a by b (or equivalently, assume a is supported in the big box and is

constant on each small box) and will prove (4.10) N_* -certainly, with the $\mathcal{L}$ norm replaced by the $l^2 L^2$ norm, by induction.

By symmetry we may assume $N_1 \geq \dots \geq N_n$. Choose the smallest q such that $N_q > 2^{10} N_{q+1}$. Then $N_1 \sim N_q$ with constant depending only on n . Unless $N_1 \leq C$, in which case (4.10) is trivial, we can conclude that

$$h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^\pm, 1 \leq j \leq n, \text{ are } \mathcal{B}_{\leq N_1^{1-\delta}}^+ \text{ measurable; } N_1^{1-\delta} \leq 2^{-10} N_q,$$

$$g_{k_j^*}(\omega)^\pm, q+1 \leq j \leq n, \text{ are } \mathcal{B}_{\leq N_{q+1}} \text{ measurable; } N_{q+1} \leq 2^{-10} N_q.$$

Note that in this case, there is no pairing among $\{k_1^*, \dots, k_n^*\}$ if and only if there is no pairing among $\{k_1^*, \dots, k_q^*\}$ and no pairing among $\{k_{q+1}^*, \dots, k_n^*\}$. We can then write $M(\omega)$ as

$$(4.13) \quad M(\omega) = \sum_{k_1^*, \dots, k_q^*} b_{k_1^* \dots k_q^*}(\omega) \cdot \prod_{j=1}^q g_{k_j^*}(\omega)^{\iota_j},$$

where

$$(4.14) \quad \begin{aligned} b_{k_1^* \dots k_q^*}(\omega) &= \sum_{k_1, \dots, k_q} \int d\lambda_1 \dots d\lambda_q \prod_{j=1}^q h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^\pm \\ &\times \sum_{(k_{q+1}, \dots, k_n)} \sum_{(k_{q+1}^*, \dots, k_n^*)} \int d\lambda_{q+1} \dots d\lambda_n \cdot a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n) \\ &\times \prod_{j=q+1}^n g_{k_j^*}(\omega)^{\iota_j} h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^\pm \end{aligned}$$

are $\mathcal{B}_{\leq 2^{-10} N_q}^+$ measurable. We then apply Lemma 4.1 and conclude that, after removing a set of ω with probability $\leq C_\theta e^{-(N_*)^\theta}$, we have

$$(4.15) \quad |M(\omega)|^2 \lesssim (N_*)^\theta \sum_{k_1^*, \dots, k_q^*} |b_{k_1^* \dots k_q^*}(\omega)|^2.$$

Now by (4.14) we have

$$(4.16) \quad \begin{aligned} \sum_{k_1^*, \dots, k_q^*} |b_{k_1^* \dots k_q^*}(\omega)|^2 &\lesssim \prod_{j=1}^q \|h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)\|_{\ell^2 \rightarrow \ell^2 L^2}^2 \sum_{k_1, \dots, k_q} \int d\lambda_1 \dots d\lambda_q \\ &\times \left| \sum_{(k_{q+1}, \dots, k_n)} \sum_{(k_{q+1}^*, \dots, k_n^*)} \int d\lambda_{q+1} \dots d\lambda_n \cdot a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n) \right. \\ &\times \left. \prod_{j=q+1}^n g_{k_j^*}(\omega)^{\iota_j} h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^\pm \right|^2; \end{aligned}$$

by induction hypothesis, we get that

$$\begin{aligned}
(4.17) \quad & \left| \sum_{(k_{q+1}, \dots, k_n)} \sum_{(k_{q+1}^*, \dots, k_n^*)} \int d\lambda_{q+1} \cdots d\lambda_n a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n) \right. \\
& \quad \times \left. \prod_{j=q+1}^n g_{k_j^*}(\omega)^{t_j} h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)^\pm \right|^2 \\
& \lesssim (N_*)^\theta \prod_{j=q+1}^n \|h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)\|_{\ell_{k_j^*}^2 \rightarrow \ell_{k_j}^2 L_{\lambda_j}^2}^2 \\
& \quad \times \sum_{k_{q+1}, \dots, k_n} \int d\lambda_{q+1} \cdots d\lambda_n \cdot |a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n)|^2,
\end{aligned}$$

up to a set of ω with probability $\leq C_\theta e^{-(N_*)^\theta}$, for any *fixed* (k_j, λ_j) for $1 \leq j \leq q$. By our assumption on the coefficients a , the function

$$(k_{q+1}, \dots, k_n, \lambda_{q+1}, \dots, \lambda_n) \mapsto a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n),$$

which depends on the parameters (k_j, λ_j) for $1 \leq j \leq q$, has only $(N_*)^{C\delta^{-7}}$ different possibilities, so by removing a set of ω with probability $\leq C_\theta e^{-(N_*)^\theta}$, we may assume (4.17) holds for *all* (k_j, λ_j) , $1 \leq j \leq q$. Thus we can sum (4.17) over k_j and integrate over λ_j , and combine with (4.15) and (4.16) to get that

$$\begin{aligned}
(4.18) \quad & |M(\omega)|^2 \lesssim (N_*)^\theta \prod_{j=1}^n \|h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)\|_{\ell_{k_j^*}^2 \rightarrow \ell_{k_j}^2 L_{\lambda_j}^2}^2 \\
& \quad \times \sum_{k_1, \dots, k_n} \int d\lambda_1 \cdots d\lambda_n \cdot |a_{k_1 \dots k_n}(\lambda_1, \dots, \lambda_n)|^2.
\end{aligned}$$

This completes the proof. $\square$

Remark 4.3. In the proof of Lemma 4.2 above, the first step involves approximating the function a by another function b that is supported in a big box of size $(N_*)^{\delta^{-7}}$ and is constant in each small box of size $(N_*)^{-\delta^{-1}}$. This reduces the infinitely many choices for the λ_j variables to essentially at most $(N_*)^{2\delta^{-7}}$ choices, which allows us to bound the probability of the tail event in question by $C_\theta e^{-(N_*)^\theta}$. This trick, which we refer to as the *meshing argument*, will be used frequently below (especially in Section 5) without further explanation.

4.2. Counting estimates for lattice points. We start with a simple lemma and then state the main integer lattice point counting bounds that will be used in the proof below.

LEMMA 4.4. (1) Let $\mathcal{R} = \mathbb{Z}$ or $\mathbb{Z}[i]$. Then, given $0 \neq m \in \mathcal{R}$, and $a_0, b_0 \in \mathbb{C}$, the number of choices for $(a, b) \in \mathcal{R}^2$ that satisfy

$$(4.19) \quad m = ab, \quad |a - a_0| \leq M, \quad |b - b_0| \leq N$$

is $O(M^\theta N^\theta)$ with constant depending only on $\theta > 0$.

(2) Given dyadic numbers $N_1 \gtrsim N_2 \gtrsim N_3$, consider the set

$$(4.20) \quad S = \{(x, y, z) \in (\mathbb{Z}^2)^3 : \iota_1 x + \iota_2 y + \iota_3 z = d, \quad \iota_1 |x|^2 + \iota_2 |y|^2 + \iota_3 |z|^2 = \alpha, \\ |x - a| \lesssim N_1, |y - b| \lesssim N_2, |z - c| \lesssim N_3\}.$$

Assume also there is no pairing in S . Then, uniformly in $(a, b, c, d) \in (\mathbb{Z}^2)^4$ and $\alpha \in \mathbb{Z}$, we have $\#S \lesssim N_2^{1+\theta} N_3$. Moreover, if $\iota_1 = \iota_2$, then we have the stronger bound $\#S \lesssim N_2^\theta N_3^2$.

Proof. (1) This strengthened divisor estimate is essentially proved in [28, Lemma 3.4]. We know that $\mathcal{R}$ has unique factorization and satisfies the standard divisor estimate, namely the number of divisors of $0 \neq m \in \mathcal{R}$ is $O(|m|^\theta)$. Now suppose $\max(|a_0|, M) \geq \max(|b_0|, N)$; then $|m| \lesssim \max(|a_0|, M)^2$. We may assume $M_1 \sim |a_0| \gg M^4$, and hence $|m| \lesssim M_1^2$.

We then claim that the number of divisors a of m that satisfies $|a - a_0| \leq M$ is at most two. In fact, suppose a, b, c are different divisors of m that belong to the ball $|x - a_0| \leq M$. Then by unique factorization, we have $\text{lcm}(a, b, c) |m|$, hence

$$\frac{abc}{\gcd(a, b) \gcd(b, c) \gcd(c, a)}$$

divides m . As $|a| \sim M_1$ etc., and $|\gcd(a, b)| \leq |a - b| \lesssim M$ etc., we conclude that

$$M_1^2 \gtrsim |m| \geq \left| \frac{abc}{\gcd(a, b) \gcd(b, c) \gcd(c, a)} \right| \gtrsim M_1^3 M^{-3},$$

contradicting the assumption $M_1 \gg M^4$.

(2) Let $x = (x_1, x_2)$, etc. If $\iota_1 = \iota_2$, then with fixed z (which has $O(N_3^2)$ choices), $x + y$ will be constant. Let $x - y = w$. Then

$$(w_1 + iw_2)(w_1 - iw_2) = |w|^2 = 2(|x|^2 + |y|^2) - |x + y|^2$$

is constant. As w belongs to a ball of radius $O(N_2)$ in $\mathbb{R}^2$, by (1) we know that the number of choices for w is $O(N_2^\theta)$, hence $\#S = O(N_2^\theta N_3^2)$. Below we will assume that $\iota_1 = +$ and $\iota_2 = -$.

(a) Suppose $\iota_3 = +$. Then we have that

$$(d_1 - z_1)(z_1 - y_1) + (d_2 - z_2)(z_2 - y_2) = (d - z) \cdot (z - y) = \frac{|d|^2 - \alpha}{2}$$

is constant. If $(d_1 - z_1)(z_1 - y_1) \neq 0$ (or similarly if $(d_2 - z_2)(z_2 - y_2) \neq 0$), then with fixed (y_2, z_2) (which has $O(N_2 N_3)$ choices), $(d_1 - z_1)(z_1 - y_1)$ will be constant. As $d_1 - z_1$ belongs to an interval of size $O(N_3)$ in $\mathbb{R}$, and $z_1 - y_1$ belongs to an interval of size $O(N_2)$ in $\mathbb{R}$, by (1) we know that the number of choices for (y_1, z_1) is $O(N_2^\theta)$, so $\#S \lesssim N_2^\theta N_2 N_3$.

If $(d_1 - z_1)(z_1 - y_1) = 0$ and $(d_2 - z_2)(z_2 - y_2) = 0$, as there is no pairing, we may assume that $d_1 = z_1$ and $z_2 = y_2$ (or $z_1 = y_1$ and $d_2 = z_2$, which is treated similarly), so $z_1 = d_1$ and $x_2 = d_2$ are fixed, z_2 has $O(N_3)$ choices and $x_1 = y_1$ has $O(N_2)$ choices, which implies $\#S \lesssim N_2 N_3$.

(b) Suppose $\iota_3 = -$. Then similarly we have that

$$(d_1 + z_1)(d_1 + y_1) + (d_2 + z_2)(d_2 + y_2) = (d + z) \cdot (d + y) = \frac{|d|^2 + \alpha}{2}$$

is constant. If $(d_1 + z_1)(d_1 + y_1) \neq 0$ (or similarly if $(d_2 + z_2)(d_2 + y_2) \neq 0$), then with fixed (y_2, z_2) (which has $O(N_2 N_3)$ choices), $(d_1 + z_1)(d_1 + y_1)$ will be constant. As $d_1 + z_1$ belongs to an interval of size $O(N_3)$ in $\mathbb{R}$, and $d_1 + y_1$ belongs to an interval of size $O(N_2)$ in $\mathbb{R}$, by (1) we know that the number of choices for (y_1, z_1) is $O(N_2^\theta)$, so $\#S \lesssim N_2^\theta N_2 N_3$.

If $(d_1 + z_1)(d_1 + y_1) = 0$ and $(d_2 + z_2)(d_2 + y_2) = 0$, as there is no pairing, we may assume that $d_1 + z_1 = 0$ and $d_2 + y_2 = 0$ (or $d_2 + z_2$ and $d_1 + y_1 = 0$, which is treated similarly), so $z_1 = -d_1$ and $y_2 = -d_2$ are fixed, z_2 has $O(N_3)$ choices and y_1 has $O(N_2)$ choices, which implies $\#S \lesssim N_2 N_3$. $\square$

PROPOSITION 4.5. *Recall the relevant constants defined in (1.35). The following bounds are uniform in all parameters. Given $d, d', k^0 \in \mathbb{Z}^2$ and $k_j^0 \in \mathbb{Z}^2$ with $1 \leq j \leq n$, let $\alpha, \Gamma \in \mathbb{R}, \iota_j \in \{+1, -1\}$ for $1 \leq j \leq n$, and $2p \leq n$. Also, let dyadic numbers M, N_j (for $0 \leq j \leq n$) and R_i (for $1 \leq i \leq p$) be such that for $1 \leq i \leq p$, we have*

$$N_{2i-1} \sim N_{2i}, \iota_{2i-1} = -\iota_{2i}, \text{ and } R_i \lesssim N_{2i-1}^{1-\delta}.$$

Let $N^{(j)} = \max^{(j)}(N_1, \dots, N_n)$, $N_{PR} = \max(N_1, \dots, N_{2p})$ and $N_ \gtrsim \max(N_0, N^{(1)})$. Also fix a subset A of $\{1, \dots, n\}$ that contains $\{1, \dots, 2p\}$, and recall the definition of the Γ -condition (3.34). Consider the sets*

(4.21)

$$S_1 = \left\{ (k, k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+1} : \sum_{j=1}^n \iota_j k_j = k + d, \quad \sum_{j=1}^n \iota_j |k_j|^2 = |k|^2 + \alpha, \right. \\ \left. |k_j - k_j^0| \lesssim N_j \ (1 \leq j \leq n), \quad |k_{2i-1} - k_{2i}| \lesssim R_i (N_*)^{C\kappa^{-1}} \ (1 \leq i \leq p) \right\},$$

(4.22)

$$\begin{aligned}
S_2 = & \left\{ (k, k', k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+2} : \iota k' + \sum_{j=1}^n \iota_j k_j = k + d, \right. \\
& \iota |k'|^2 + \sum_{j=1}^n \iota_j |k_j|^2 = |k|^2 + \alpha, \quad |k|, |k'| \lesssim N_0, \quad |k_j - k_j^0| \lesssim N_j \quad (1 \leq j \leq n), \\
& \left. |k_{2i-1} - k_{2i}| \lesssim R_i (N_*)^{C\kappa^{-1}} \quad (1 \leq i \leq p) \right\},
\end{aligned}$$

(4.23)

$$\begin{aligned}
S^+ = & \left\{ (k, k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+1} \text{ and} \right. \\
& \left. (k, k', k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+2} : \sum_{j \in A} \iota_j k_j = d' \right\},
\end{aligned}$$

(4.24)

$$\begin{aligned}
S_3 = & \left\{ (k, k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+1} : \sum_{j=1}^n \iota_j k_j = k + d, \right. \\
& \left| |k|^2 - \sum_{j=1}^n \iota_j |k_j|^2 - \alpha \right| \lesssim M, \quad |k| \lesssim N_0, \quad |k_j| \lesssim N_j \quad (1 \leq j \leq n), \\
& \left. |k_{2i-1} - k_{2i}| \lesssim R_i (N_*)^{C\kappa^{-1}} \quad (1 \leq i \leq p), \quad \text{and (3.34) holds} \right\}.
\end{aligned}$$

Assume that there is no pairing among the variables k, k' and k_j in the sets above. Let $S_j^+ = S_j \cap S^+$. Then for S_1 , we have

$$(4.25) \quad (\#S_1) \cdot \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \lesssim (N_{PR})^{2\gamma_0} (N_*)^{C\kappa^{-1}} (N^{(1)} N^{(2)})^{-1} \prod_{j=1}^n N_j^2.$$

If $N^{(1)} \sim N_a$ and $\iota_a = -$, then we have

$$(4.26) \quad (\#S_1) \cdot \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \lesssim (N_{PR})^{2\gamma_0} (N_*)^{C\kappa^{-1}} (N^{(1)})^{-2} \prod_{j=1}^n N_j^2.$$

For S_2 and S_3 , we have

$$(4.27) \quad (\#S_2) \cdot \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \lesssim (N_{PR})^{2\gamma_0} (N_*)^{C\kappa^{-1}} N_0 (N^{(1)})^{-1} \prod_{j=1}^n N_j^2,$$

(4.28)

$$(\#S_3) \cdot \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \lesssim (N_{PR})^{2\gamma_0} (N_*)^{C\kappa^{-1}} M \frac{\max((N^{(2)})^2, |\alpha|)}{(N^{(2)})^2} (N^{(1)})^{-2} \prod_{j=1}^n N_j^2.$$

Finally, suppose we replace any of these S_j by the set S_j^+ . Then (4.25)–(4.27) hold with the right-hand side multiplied by an extra factor

$$(4.29) \quad \left[\min(N^{(2)}, \max_{j \in A, j \geq 2p+1} N_j) \right]^{-1}.$$

If $N^{(1)} \sim N_a$ and $a \in A$, then the stronger bound (4.26) holds for S_1^+ with the right-hand side multiplied by an extra factor (4.29), regardless of whether $\iota_a = -$ or not. If $N^{(1)} \sim N_a$ and $2p+1 \leq a \in A$, then (4.27) holds for S_2^+ with the right-hand side multiplied by an extra factor $(N^{(1)})^{-1}$. As for S_3^+ , either it satisfies (4.28) with the same extra factor (4.29), or it satisfies

$$(4.30) \quad (\#S_3^+) \cdot \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \lesssim (N_*)^{C\kappa^{-1}} M(N_{PR})^{2\gamma_0} \min\left(\frac{\max((N^{(2)})^2, |\alpha|)}{(N^{(1)}N^{(2)})^2}\right. \\ \left. \times \left(\max^{(2)}\{N_j : 2p+1 \leq j \in A\}\right)^{-1}, (N^{(1)})^{-1}(N^{(2)})^{-2}\right) \prod_{j=1}^n N_j^2.$$

Proof. Let a, b, c be such that $N^{(1)} \sim N_a$, $N^{(2)} \sim N_b$, and $\max(\{N_j : 2p+1 \leq j \in A\}) \sim N_c$. In the proof below any factor that is $\lesssim (N_*)^{C\kappa^{-1}}$ will be negligible, so we will pretend they are 1. For simplicity, let us first also ignore all $N_{2i-1}^{2\gamma_0}$ factors; at the end of the proof we will explain how to put them back.

(1) We start with (4.25). If $a, b \geq 2p+1$, we may fix all $k_j (j \notin \{a, b\})$, and then apply Lemma 4.4(2) to count the triple (k, k_a, k_b) . This gives

$$(4.31) \quad (\#S_1) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, b\}} N_j^2 \right) N_a N_b,$$

which proves (4.25) as $R_j \lesssim N_{2j-1}^{1-\delta}$. If $a \geq 2p+1$ and $b \leq 2p$ (say $b=1$), we may fix all $k_j (j \notin \{1, a\})$, and then apply Lemma 4.4(2) to count the triple (k, k_1, k_a) , noticing that k_1 belongs to a disc of radius $O(R_1(N_*)^{C\kappa^{-1}})$ once k_2 is fixed. This gives

$$(4.32) \quad (\#S_1) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(N_1^2 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2 \right) N_a R_1 \frac{N_1}{R_1},$$

which proves (4.25). Finally, if $a \leq 2p$ (say $a=1$), then we may assume $b=2$. We may fix all $k_j (j \geq 3)$, and then apply Lemma 4.4(2) to count the triple (k, k_1, k_2) , noticing that k belongs to a disc of radius $O(R_1(N_*)^{C\kappa^{-1}})$ once all $k_j (j \geq 3)$ are fixed. This gives

$$(4.33) \quad (\#S_1) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=2}^p N_{2i-1}^3 R_i \prod_{j \geq 2p+1} N_j^2 \right) N_1 R_1 \frac{N_1}{R_1},$$

which proves (4.25).

As for (4.26) and (4.27) we only need to consider a . If $a \geq 2p+1$, we may fix all $k_j (j \neq a)$, and then apply Lemma 4.4(2) to count the pair (k, k_a) for (4.26) (using the fact $\iota_a = -$) and the triple (k, k', k_a) for (4.27). Then we get

$$(4.34) \quad (\#S_1) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2 \right),$$

$$(4.35) \quad (\#S_2) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2 \right) N_0 N_a,$$

which proves (4.26) and (4.27). If $a \leq 2p$ (say $a = 1$), we may assume $b = 2$ — in particular, $N^{(1)} \sim N^{(2)}$ and (4.26) follows from (4.25); for (4.27), we may fix all $k_j (j \geq 2)$ and then apply Lemma 4.4(2) to count the triple (k, k', k_1) , noticing that k_1 belongs to a disc of radius $O(R_1(N_*)^{C\kappa^{-1}})$ once k_2 is fixed. Then we get

$$(4.36) \quad (\#S_2) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(N_1^2 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{j \geq 2p+1} N_j^2 \right) N_0 R_1 \frac{N_1}{R_1},$$

which proves (4.27).

(2) Next we prove the improvements to (4.25)–(4.27) for S_j^+ . We start with (4.25). If $a, b \notin A$, we may fix all $k_j (c \neq j \in A)$ and apply (4.25) to the rest variables and get

$$(4.37) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim (N_a N_b)^{-1} \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{j \in A, c \neq j \geq 2p+1} N_j^2,$$

which gains a factor N_c^{-2} upon (4.25). If $a \notin A$ and $2p+1 \leq b \in A$, we may fix all $k_j (b \neq j \in A)$ and apply (4.25) to the rest variables and get

$$(4.38) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim N_a^{-1} \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{j \in A, b \neq j \geq 2p+1} N_j^2,$$

which gains a factor N_b^{-1} upon (4.25). If $a \notin A$ and $b \leq 2p$ (say $b = 1$), we may fix all $k_j (2 \leq j \neq a)$, noticing that k_c belongs to a ball of radius $\min(N_c, R_1(N_*)^{C\kappa^{-1}})$ once all $k_j (3 \leq j \notin \{a, c\})$ are fixed. We then apply Lemma 4.4(2) to count the pair (k, k_a) . Then we get

$$(4.39) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(N_1^2 \min(N_c^2, R_1^2) \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, c\}} N_j^2 \right) N_a \frac{N_1}{R_1},$$

which gains a factor N_c^{-1} upon (4.25).

Now if $2p+1 \leq a \in A$ and either $b \notin A$ or $2p+1 \leq b \in A$, we may fix all $k_j (j \notin \{a, b\})$ and apply Lemma 4.4(2) to count the pair (k, k_b) (if $b \notin A$) or (k_a, k_b) (if $2p+1 \leq b \in A$). Then we get

$$(4.40) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, b\}} N_j^2 \right) N_b,$$

which gains a factor N_b^{-1} upon the stronger bound (4.26). If $2p+1 \leq a \in A$ and $b \leq 2p$ (say $b = 1$), we may fix all $k_j (j \notin \{a, 1, 2\})$ and apply Lemma 4.4(2) to count the triple (k_a, k_1, k_2) , noticing that k_a belongs to a disc of radius $O(R_1(N_*)^{C\kappa^{-1}})$ once all $k_j (j \notin \{a, 1, 2\})$ are fixed. Then we get

$$(4.41) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2 \right) N_1 R_1 \frac{N_1}{R_1},$$

which gains a factor N_b^{-2} upon the stronger bound (4.26). Finally, if $a \leq 2p$ (say $a = 1$), then we may assume $b = 2$. We may fix all $k_j (j \geq 3)$, noticing that k_c belongs to a disc of radius $\min(N_c, R_1(N_*)^{C\kappa^{-1}})$ once all $k_j (3 \leq j \neq c)$ are fixed, and then apply Lemma 4.4(2) to count the pair (k_1, k_2) . This gives

$$(4.42) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(\min(N_c, R_1)^2 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq c} N_j^2 \right) N_1 \frac{N_1}{R_1},$$

which gains a factor N_c^{-1} upon the stronger bound (4.26).

As for (4.26) and (4.27) we only need to consider a . If $a \notin A$, we may fix all $k_j (c \neq j \in A)$ and apply (4.26) (if $\iota_a = -$) or (4.27) to the rest variables. Then we get

$$(4.43) \quad (\#S_1^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim N_a^{-2} \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{j \in A, c \neq j \geq 2p+1} N_j^2,$$

$$(4.44) \quad (\#S_2^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim N_0 N_a^{-1} \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{j \in A, c \neq j \geq 2p+1} N_j^2,$$

which gains a factor N_c^{-2} upon (4.26) or (4.27). If $a \in A$, then (4.26) follows from the above proof for (4.25); for (4.27), if $2p+1 \leq a \in A$, we may fix all $k_j (a \neq j \in A)$ and apply (4.27) to the rest variables and get

$$(4.45) \quad (\#S_2^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim N_0 \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{j \in A, a \neq j \geq 2p+1} N_j^2,$$

which gains a factor N_a^{-1} upon (4.27); if $a \leq 2p$ (say $a = 1$), we may fix all $k_j (2 \leq j \in A)$, noticing that k_c belongs to a ball of radius $\min(N_c, R_1(N_*)^{C\kappa^{-1}})$

once all $k_j (j \in A \setminus \{1, 2, c\})$ are fixed, and apply (4.27) to the rest variables and get

$$(4.46) \quad (\#S_2^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \min(N_c, R_1)^2 N_1^2 \frac{N_1}{R_1} N_0 \prod_{j \notin A} N_j^2 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{j \in A, c \neq j \geq 2p+1} N_j^2,$$

which gains a factor N_c^{-1} upon (4.27).

(3) Now we consider (4.28) and its improvement. We may assume $\iota_a = +$ and $N^{(1)} \gg N^{(2)}$ (so $a \geq 2p+1$), since otherwise (4.28) follows from (4.26) or (4.25) and similarly for the improvement. Now let $M_0 = \max(|\alpha|, (N^{(2)})^2)$. If $M \gg M_0$, then we have

$$||k|^2 - |k_a|^2| \leq |\alpha| + \sum_{j \neq a} |k_j|^2 + M \lesssim M;$$

combining with (3.34) and Lemma 4.4(1) we conclude that the number of choices for $|k_a|^2$, and thus k_a , is $O(M)$. We may fix k_a and then count $k_j (j \neq a)$ to get

$$(4.47) \quad (\#S_3) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim M \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2,$$

which proves (4.28). As for S_3^+ , if $a \in A$, then the improvement of (4.28) follows from the improvement of (4.26); if $a \notin A$, we may fix k_a and count $k_j (j \notin \{a, c\})$ to get

$$(4.48) \quad (\#S_3^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim M \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, c\}} N_j^2,$$

which gains a factor N_c^{-2} upon (4.28).

Assume now $M \lesssim M_0$. Then just like above we have $||k|^2 - |k_a|^2| \lesssim M_0$, so k has at most $O(M_0)$ choices, and similarly k_a has at most $O(M_0)$ choices. If $b \geq 2p+1$, we may assume $\iota_b = +$ (otherwise switch the roles of k and k_a), then fix k and $k_j (j \notin \{a, b\})$ and apply Lemma 4.4(2) to count the pair (k_a, k_b) to get

$$(4.49) \quad (\#S_3) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim M_0 \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, b\}} N_j^2 \right) M,$$

which proves (4.28); if $b \leq 2p$ (say $b = 1$), we may fix k and $k_j (j \notin \{1, 2, a\})$ and apply Lemma 4.4(2) to count the triple (k_1, k_2, k_a) , noticing that k_a belongs to a disc of radius $O(R_1(N_*)^{C\kappa^{-1}})$ once k and $k_j (j \notin \{1, 2, a\})$ are fixed, and get

$$(4.50) \quad (\#S_3) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim \left(M_0 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq a} N_j^2 \right) N_1 R_1 M \frac{N_1}{R_1},$$

which proves (4.28).

It remains to prove the improvement of (4.28) for S_3^+ . We may assume $a \notin A$, since otherwise it follows from the improvement of (4.26). Now if $b \notin A$, we may fix all $k_j (c \neq j \in A)$ and apply (4.28) to the rest variables and get

$$(\#S_3^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim MM_0(N_a N_b)^{-2} \prod_{j \notin A} N_j^2 \prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \neq c, j \in A} N_j^2, \quad (4.51)$$

which gains a factor N_c^{-2} upon (4.28). If $b \leq 2p$, say $b = 1$, we may fix k and $k_j (3 \leq j \neq a)$, noticing that k_c belongs to a disc of radius $\min(N_c, R_1(N_*)^{C\kappa^{-1}})$ once all $k_j (3 \leq j \notin \{a, c\})$ are fixed, and then apply Lemma 4.4(2) to count the pair (k_1, k_2) and get

$$(\#S_3^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim MM_0 \left(\min(N_c, R_1)^2 \prod_{i=2}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, c\}} N_j^2 \right) N_1 \frac{N_1}{R_1}, \quad (4.52)$$

which gains a factor N_c^{-1} upon (4.28). Finally, assume $2p+1 \leq b \in A$. Then we will prove (4.30). Let $\max^{(2)}\{N_j : 2p+1 \leq j \in A\} \sim N_d$. We may fix k and $k_j (j \notin \{a, b, d\})$, then apply Lemma 4.4(2) to count the pair (k_b, k_d) and get

$$(\#S_3^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim MM_0 \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, b, d\}} N_j^2 \right) N_d; \quad (4.53)$$

alternatively we may choose to fix $k_j (j \notin \{a, b\})$ and then apply Lemma 4.4(2) to count the pair (k, k_a) and get

$$(\#S_3^+) \prod_{i=1}^p \frac{N_{2i-1}}{R_i} \lesssim M \left(\prod_{i=1}^p N_{2i-1}^3 R_i \prod_{2p+1 \leq j \notin \{a, b\}} N_j^2 \right) N_a. \quad (4.54)$$

Combining (4.53) and (4.54) yields (4.30).

In the last part we will explain how to put back the powers $N_{2i-1}^{2\gamma_0}$. In fact, in each estimate above we have the product $\prod_{i \geq 2} N_{2i-1}^3 R_i$. As $R_i \lesssim N_{2i-1}^{1-\delta}$ and $N_{2i-1} \sim N_{2i}$, we have

$$N_{2i-1}^3 R_i \lesssim N_{2i-1}^2 N_{2i}^2 \cdot N_{2i-1}^{-2\gamma_0},$$

which allows us to incorporate the extra factor $N_{2i-1}^{2\gamma_0}$ for $i \geq 2$. Thus we lose at most a factor $N_1^{2\gamma_0}$, which is acceptable as $N_1 \lesssim N_{PR}$. $\square$

COROLLARY 4.6. *Recall $a_0 > 1$ defined in (1.35), and let all the parameters $(d, N_j, \iota_j, \text{etc.})$ be as in Proposition 4.5. From the sets $S_j (1 \leq j \leq 3)$ in Proposition 4.5 we may construct the quantities $\mathcal{E}_j$ as follows: each $\mathcal{E}_j$ is a sum over a set S_j^{lin} . This S_j^{lin} is formed from S_j by removing from its defining properties the one that involves the quadratic algebraic sum Σ (this Σ is $\iota_1|k_1|^2 + \dots + \iota_n|k_n|^2 - |k|^2$ for S_1 and S_3 , and $\iota_1|k_1|^2 + \dots + \iota_n|k_n|^2 + \iota|k'|^2 - |k|^2$*

for S_2), and the summand is simply $\langle \Sigma - \alpha \rangle^{-a_0}$. Similarly define $\mathcal{E}_j^+$ by replacing S_j with S_j^+ .

Then, [inequalities \(4.25\)–\(4.30\)](#), as well as their improvements, hold with $\#S_j$ replaced by $\mathcal{E}_j$ ($\#S_j^+$ replaced by $\mathcal{E}_j^+$), and with the factor M on the right-hand sides of [\(4.28\)](#) and [\(4.30\)](#) removed.

Proof. This is straightforward, by applying [Proposition 4.5](#) for each value of Σ and summing up using $a_0 > 1$ for [\(4.25\)–\(4.27\)](#), and by dyadically decomposing $\langle \Sigma - \alpha \rangle$ and applying [Proposition 4.5](#) for each dyadic piece for [\(4.28\)–\(4.30\)](#). $\square$

5. Proof of the multilinear estimates

In this section we will prove [Proposition 3.4](#), thus completing the local theory. We start with an estimate for general multilinear forms without pairing.

Given $d \in \mathbb{Z}^2$ and $\alpha \in \mathbb{R}$, consider the following expressions:

(5.1)

$$\begin{aligned} \mathcal{X} := & \sum_{\substack{(k, k_1, \dots, k_n) \\ \iota_1 k_1 + \dots + \iota_n k_n = k + d}} \int d\lambda d\lambda_1 \cdots d\lambda_n \\ & \times \eta \left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \alpha \right) \overline{v_k(\lambda)} \prod_{j=1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j}, \end{aligned}$$

(5.2)

$$\begin{aligned} \mathcal{Y} := & \sum_{\substack{(k, k', k_1, \dots, k_n) \\ \iota_1 k_1 + \dots + \iota_n k_n + \iota k' = k + d}} \int d\lambda d\lambda' d\lambda_1 \cdots d\lambda_n \cdot \eta \left(\lambda, \lambda - |k|^2 - \iota(\lambda' - |k'|^2) \right. \\ & \left. - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \alpha \right) y_{kk'}(\lambda, \lambda') \prod_{j=1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j}, \end{aligned}$$

where $d \in \mathbb{Z}^2$ and $\alpha \in \mathbb{R}$ are fixed and η is a function that satisfies

$$(5.3) \quad |\eta(\lambda, \mu)| + |\partial_{\lambda, \mu} \eta(\lambda, \mu)| \lesssim \langle \mu \rangle^{-10}.$$

In the summation we always assume that there is *no pairing*¹⁰ among the variables k , k' and k_j .

¹⁰This requirement appears in the form of coefficients that are indicator functions of sets of form $\{k_j \neq k_l\}$. Such coefficients may lead to slightly different multilinear Gaussian expressions in the estimates below, but there will be at most $(N_*)^C$ possibilities where N_* is a parameter to be defined below, and will not affect any estimates since our exceptional sets will always have measure at most $C_\theta e^{-(N_*)^\theta}$.

We assume that the input functions $v^{(j)}$ are as in [Proposition 3.4](#), where $v^{(j)}$ are of type (G) or (C) for $1 \leq j \leq n_1$, and of type (D) for $n_1 + 1 \leq j \leq n$. Since we are working exclusively in the λ_j spaces, we will abuse notation here and write $(v_{k_j}^{(j)})(\lambda_j)$ instead of $(\widetilde{v^{(j)}})_{k_j}(\lambda_j)$.

Let the parameters N_j , L_j , $N^{(j)}$ etc., and the sets $\mathcal{G}$ and $\mathcal{C}$ be as in [Proposition 3.4](#). We further assume that the functions $v_k(\lambda)$ and $y_{kk'}(\lambda, \lambda')$ satisfy

$$(5.4) \quad \|\langle \lambda \rangle^b v_k(\lambda)\|_{\ell_k^2 L_\lambda^2} \lesssim 1, \quad \|\langle \lambda \rangle^b \langle \lambda' \rangle^b y_{kk'}(\lambda, \lambda')\|_{\ell_{k,k'}^2 L_{\lambda,\lambda'}^2} \lesssim 1,$$

and that $v_k(\lambda)$ is supported in $\{|k| \lesssim N_0\}$ and $y_{kk'}(\lambda, \lambda')$ is supported in $\{|k|, |k'| \lesssim N_0\}$.

PROPOSITION 5.1. *Recall the relevant constants defined in (1.35), and that $\tau \ll 1$. Under all the above assumptions, there exist p and q , and $N_{2p+l} \gtrsim R_{2p+l} \gtrsim L_{2p+l}(1 \leq l \leq q)$ such that $2p + q \leq n_1$, that for $1 \leq i \leq p$ we must have $N_{2i-1} \sim N_{2i}$ and $\iota_{2i-1} = -\iota_{2i}$, and that $2i-1$ and $2i$ do not both belong to $\mathcal{G}$. Define $R_i = \max(L_{2i-1}, L_{2i})$, and let N_* be fixed. Then the following estimates hold $\tau^{-1}N_*$ -certainly. Here, as in [Proposition 3.4](#), the exceptional set of ω removed does not depend on the choice of the functions $v^{(j)}$ ($j \geq n_1 + 1$) or v or w .*

(1) Assume $N_* \gtrsim \max(N_0, N^{(1)})$. Then we have

$$(5.5) \quad |\mathcal{X}|^2 \lesssim \tau^{-\theta}(N_*)^{C\kappa^{-1}} \mathcal{E}_1 \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \prod_{j=1}^n N_j^{-2} \prod_{j \geq n_1+1} N_j^{2\gamma} \prod_{2p+1 \leq j \leq n_1} L_j^{-2\delta_0},$$

and similarly

$$(5.6) \quad |\mathcal{Y}|^2 \lesssim \tau^{-\theta}(N_*)^{C\kappa^{-1}} \mathcal{E}_2 \prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \prod_{j=1}^n N_j^{-2} \prod_{j \geq n_1+1} N_j^{2\gamma} \prod_{2p+1 \leq j \leq n_1} L_j^{-2\delta_0},$$

where $\mathcal{E}_1$ and $\mathcal{E}_2$ are the quantities defined in [Corollary 4.6](#), with $a_0 = 2b - 10\delta^6$, and for some choice of the parameters in that corollary that do not appear in the assumptions of the current proposition. Moreover, if in the sum defining $\mathcal{X}$ we also assume the Γ -condition (3.34), then (5.5) holds with $\mathcal{E}_1$ replaced by $\mathcal{E}_3$. (See that cor for the relevant definitions.)

(2) Assume $N_* \gtrsim \max(N_0, N^{(1)})$. Then we have

$$(5.7) \quad |\mathcal{X}|^4 \lesssim \tau^{-\theta}(N_*)^{C\kappa^{-1}} \mathcal{E}_1 \mathcal{E}_1^+ \left(\prod_{i=1}^p \frac{N_{2i-1}^{1+2\gamma_0}}{R_i} \right)^2 \prod_{j=1}^n N_j^{-4} \prod_{j \geq n_1+1} N_j^{4\gamma} \prod_{2p+1 \leq j \leq n_1} L_j^{40n^2},$$

and similarly

(5.8)

$$|\mathcal{Y}|^4 \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} \mathcal{E}_2 \mathcal{E}_2^+ \left(\prod_{i=1}^p \frac{N_i^{1+2\gamma_0}}{R_i} \right)^2 \prod_{j=1}^n N_j^{-4} \prod_{j \geq n_1+1} N_j^{4\gamma} \prod_{2p+1 \leq j \leq n_1} L_j^{40n^2},$$

where $\mathcal{E}_j$ and $\mathcal{E}_j^+$ are the quantities defined in [Corollary 4.6](#), again for some choice of the parameters in that corollary that do not appear in the assumptions of the current proposition. In the set S^+ in (4.23), the set A will contain $\{1, 2, \dots, 2p\} \cup \{n_1 + 1, \dots, n\}$. Moreover, if in the sum defining $\mathcal{X}$ we also assume the Γ -condition (3.34), then (5.7) holds with $\mathcal{E}_1 \mathcal{E}_1^+$ replaced by $\mathcal{E}_3 \mathcal{E}_3^+$.

(3) Assume, in addition, that $N^{(1)} \sim N_n$ and $N_* \gtrsim N^{(2)}$. Then (5.7) is true, with N_n replaced by $N^{(2)}$ in both quantities $\mathcal{E}_1$ and $\mathcal{E}_1^+$. Moreover, we have

$$(5.9) \quad |\mathcal{X}|^4 \lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}} (N^{(1)})^{-4(1-\gamma)} (N^{(2)})^{C\gamma} \\ \times \widetilde{\mathcal{E}}_1 \widetilde{\mathcal{E}}_1^+ \left(\prod_{i=1}^p \frac{N_i^{1+2\gamma_0}}{R_i} \right)^2 \prod_{j=1}^{n-1} N_j^{-4} \prod_{j=2p+1}^{2p+q} N_j^2 R_j^{-2},$$

where $\widetilde{\mathcal{E}}_1$ is the quantity defined in [Corollary 4.6](#) for some choice of the parameters in that corollary that do not appear in the assumptions of the current proposition, but with N_n replaced by $N^{(2)}$. Similarly $\widetilde{\mathcal{E}}_1^+$ is the quantity $\mathcal{E}_1^+$ defined in [Corollary 4.6](#) with $A = \{1, \dots, n\}$, but with N_n replaced by $N^{(2)}$ and N_{2p+l} replaced by $R_{2p+l} (N_*)^{C\kappa^{-1}}$ for $1 \leq l \leq q$. Moreover, the exceptional set of ω removed is independent of $N^{(1)}$.

5.1. Proof of [Proposition 5.1](#). We will prove [Proposition 5.1](#) in this section. We will only prove the bounds for $\mathcal{X}$ without Γ -condition; with obvious modifications the proof also works for $\mathcal{Y}$ and for the version with Γ -condition. For simplicity we will omit the ω dependence, and may ignore any factors that are $\lesssim \tau^{-\theta} (N_*)^{C\kappa^{-1}}$.

Our proof will roughly follow an algorithm, indicated by the following steps:

- (1) Distinguish between the inputs $j \in \mathcal{D}$, where $v^{(j)}$ are bounded in $\ell^2 L^2$, with $j \in \mathcal{G} \cup \mathcal{C}$.
- (2) Identify the pairings among $k_j (j \in \mathcal{G})$ and $k_j^* (j \in \mathcal{C})$, and reduce the sum of products of the $h^{(j)}$ functions over the paired variables to some functions $P^{(i)}$ (see (5.11)) that are also bounded in $\ell^2 L^2$.
- (3) Estimate the sum in unpaired variables using [Lemma 4.2](#). (In [Section 5.1.2](#) we will skip step (2) and estimate the whole sum including paired and unpaired variables using [Lemma 4.1](#).)
- (4) Apply Cauchy-Schwartz to handle all the factors in $\ell^2 L^2$, and then reduce to the $\mathcal{E}_j$ type quantities in [Corollary 4.6](#).

- (5) When necessary, apply a $\mathcal{T}^*\mathcal{T}$ argument and repeat the previous steps for the resulting kernel.

As the proof will be notation heavy, the reader may do a first reading making the following simplifications without missing the core parts of the proof: (1) omit integration in any λ_j and pretend $\lambda_j = 0$ (so $v^{(j)}$ is a function of k_j only and $h^{(j)}$ is a function of k_j and k_j^* only); (2) when identifying pairings, restrict to only simple pairings where (say) $k_i^* = k_j^*$ and does not equal any other k_l^* . These will make formulas like (5.12) simpler and the proofs more transparent.

Throughout the proof we will fix the sets $U = \{1, 2, \dots, n\}$ and $V = \{1, 2, \dots, n-1\}$. We will (in this section only) introduce a shorthand notation for vectors: for a finite set X , define $k_{[X]}$ to be the vector $(k_j : j \in X)$; similarly define $\lambda_{[X]}$, $k_{[X]}^*$, etc., and define $d\lambda_{[X]} = \prod_{j \in X} d\lambda_j$.

5.1.1. *A simple bound.* We first prove (5.5). By definition we expand

$$\begin{aligned}
 \mathcal{X} = & \sum_{(k, k_{[U]}): \iota_1 k_1 + \dots + \iota_n k_n = k+d} \sum_{k_{[U]}^*} \int d\lambda d\lambda_{[U]} \\
 (5.10) \quad & \times \eta\left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \alpha\right) \\
 & \times \overline{v_k(\lambda)} \prod_{j=1}^{n_1} \frac{g_{k_j^*}^{\iota_j}}{\langle k_j^* \rangle} h_{k_j k_j^*}^{(j)}(\lambda_j)^{\pm} \prod_{j=n_1+1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j};
 \end{aligned}$$

recall that $k_{[U]}$ means $(k_1, \dots, k_n)$, etc. The sum in $k_{[U]}^*$ is restricted to $\frac{N_j}{2} < \langle k_j^* \rangle \leq N_j$, and $h_{k_j k_j^*}^{(j)}(\lambda_j, \omega)$ is defined as in (3.25) for $j \in \mathcal{C}$ and is defined to be $\mathbf{1}_{k_j = k_j^*} \widehat{\chi}(\lambda_j)$ for $j \in \mathcal{G}$.

Consider now the sum in $k_{[U]}^*$. By identifying all pairings among them (recall the definition of pairings in Definition 1.10), we may assume there are p sets Y_i ($1 \leq i \leq p$, $2p \leq n_1$) and a set Z that partitions $\{1, \dots, n_1\}$, such that

- (i) each Y_i contains a pairing;
- (ii) the k_j^* takes a single value for j in each Y_i ;
- (iii) this value is different for different Y_i and is different from k_j^* for $j \in Z$;
- and
- (iv) there is no pairing in $\{k_j^* : j \in Z\}$.

Then we manipulate this sum and rewrite it as a combination of two types of sums, namely,

- (1) where we only require¹¹ that k_j^* takes a single value for j in each Y_i and that there is no pairing in $\{k_j^* : j \in Z\}$; and

¹¹That is, we relax the requirement (iii) above, keeping only requirements (ii) and (iv).

(2) where there are more pairings in addition to case (1), namely when the value for Y_i equals the value for some other $Y_{i'}$ or some k_j^* for $j \in Z$.

Since there are strictly more pairings in case (2) than in the sum we started with, we may repeat this process and eventually reduce to sums of type (1) only. The purpose of this manipulation is to ensure that the sum in k_j^* ($j \in Y_i$) gives *exactly*

$$(5.11) \quad P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]}) = \sum_{k^*} \prod_{j \in Y_i} h_{k_j k^*}^{(j)}(\lambda_j)^\pm \langle k^* \rangle^{-q_3} (g_{k^*})^{q_1} (\overline{g_{k^*}})^{q_2},$$

where $q_1 + q_2 = q_3 = |Y_i|$.

Note that N_j for $j \in Y_i$ are all comparable. Without loss of generality we may assume $\{2i-1, 2i\} \subset Y_i$ and $\iota_{2i-1} = -\iota_{2i}$. As (k_{2i-1}, k_{2i}) is *not* a pairing, $2i-1$ and $2i$ cannot both belong to $\mathcal{G}$. Now we may assume $|k_{2i-1} - k^*| \lesssim L_{2i-1}(N_*)^{C\kappa^{-1}}$ and similarly for k_{2i} , since otherwise we gain a power $(N_*)^{-200n^2}$ due to the last bound in (3.26), which cancels any summation in any (k_j, k_j^*) . The estimate will then follow immediately.

Let $R_i = \max(L_{2i-1}, L_{2i})$, say $R_i = L_{2i-1}$. Then we have $|k_{2i-1} - k_{2i}| \lesssim R_i(N_*)^{C\kappa^{-1}}$ for $1 \leq i \leq p$. For (5.11) using the first two bounds in (3.26), we have that

$$(5.12) \quad \begin{aligned} & \|P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]}) \prod_{j \in Y_i} \langle \lambda_j \rangle^b\|_{\ell_{k_{[Y_i]}}^2 L_{\lambda_{[Y_i]}}^2}^2 \\ & \lesssim \prod_{j \in Y_i} N_j^{-2} \cdot \|\langle \lambda_{2i} \rangle^b h_{k_{2i} k^*}^{(2i)}(\lambda_{2i})\|_{\ell_{k^*}^2 \rightarrow \ell_{k_{2i}}^2 L_{\lambda_{2i}}^2}^2 \sum_{k^*} \prod_{2i \neq j \in Y_i} \|\langle \lambda_j \rangle^b h_{k_j k^*}^{(j)}(\lambda_j)\|_{\ell_{k_j}^2 L_{\lambda_j}^2}^2 \\ & \lesssim \|\langle \lambda_{2i-1} \rangle^b h_{k_{2i-1} k^*}^{(2i-1)}(\lambda_{2i-1})\|_{\ell_{k_{2i-1} k^*}^2 L_{\lambda_{2i-1}}^2}^2 \prod_{j \in Y_i} N_j^{-2} \prod_{\substack{j \in Y_i \\ j \neq 2i-1, 2i}} L_j^{-2\delta_0} \\ & \lesssim N_{2i-1}^{1+2\gamma_0} R_i^{-1} \prod_{j \in Y_i} N_j^{-2} \prod_{\substack{j \in Y_i \\ j \neq 2i-1, 2i}} L_j^{-2\delta_0}. \end{aligned}$$

Now we have reduced the expression for $\mathcal{X}$ to

$$(5.13) \quad \begin{aligned} \mathcal{X} = & \sum_{\substack{(k, k_{[U]}): \\ \iota_1 k_1 + \dots + \iota_n k_n = k+d}} \sum_{k_{[Z]}^*} \int d\lambda d\lambda_{[U]} \cdot \eta\left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j(\lambda_j - |k_j|^2) - \alpha\right) \\ & \times \overline{v_k(\lambda)} \prod_{i=1}^p P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]}) \prod_{j \in Z} \frac{g_{k_j^*}^{\iota_j}}{\langle k_j^* \rangle} h_{k_j k_j^*}^{(j)}(\lambda_j)^\pm \prod_{j=n_1+1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j}. \end{aligned}$$

Compared to (5.10) it is important that there is no pairing in $k_{[Z]}^*$. For simplicity of notation, we will write

$$(5.14) \quad \mathcal{X} = \sum_{(1)} \int \mathfrak{F} \cdot \mathfrak{G},$$

where the symbol $\sum_{(1)} \int$ represents the sum in k and $k_{[U \setminus Z]}$ and integration in λ and $\lambda_{[U \setminus Z]}$, the factor $\mathfrak{F}$ is

$$(5.15) \quad \mathfrak{F} := \overline{v_k(\lambda)} \prod_{i=1}^p P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]}) \prod_{j=n_1+1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j},$$

and the multilinear Gaussian $\mathfrak{G}$ is given by

$$(5.16) \quad \mathfrak{G} := \sum_{k_{[Z]}} \sum_{k_{[Z]}^*} \int d\lambda_{[Z]} \cdot \prod_{j \in Z} \frac{g_{k_j^*}^{\iota_j}}{\langle k_j^* \rangle} \langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)^{\pm} \cdot \mathfrak{A},$$

with coefficient $\mathfrak{A}$ of form

$$(5.17) \quad \mathfrak{A} := \mathbf{1}_{\sum_{j \in Z} \iota_j k_j = d_0} \cdot \eta \left(\lambda, \alpha_0 - \sum_{j \in Z} \iota_j (\lambda_j - |k_j|^2) \right) \prod_{j \in Z} \langle \lambda_j \rangle^{-b},$$

where

$$(5.18) \quad d_0 := k + d - \sum_{j \notin Z} \iota_j k_j \in \mathbb{Z}^2, \quad \alpha_0 := \lambda - |k|^2 - \alpha - \sum_{j \notin Z} \iota_j (\lambda_j - |k_j|^2) \in \mathbb{R}.$$

The goal now is to estimate $\mathfrak{G}$. For fixed values of $(k, \lambda, k_{[U \setminus Z]}, \lambda_{[U \setminus Z]})$, we may apply Lemma 4.2; in order to make this uniform, we will apply the meshing argument in Remark 4.3. This allows us to reduce to at most $(N_*)^{C\delta^{-7}}$ choices, so in the end, after removing a set of probability $\leq C_\theta e^{-(\tau^{-1}N_*)^\theta}$ we can apply Lemma 4.2 for *all* choices of $(k, \lambda, k_{[U \setminus Z]}, \lambda_{[U \setminus Z]})$ and use the first bound in (3.26) to get that

$$(5.19) \quad \begin{aligned} |\mathfrak{G}|^2 &\lesssim \prod_{j \in Z} N_j^{-2} \prod_{j \in Z} L_j^{-2\delta_0} \cdot \|\mathfrak{A}\|_{\mathcal{L}}^2 \\ &\lesssim \prod_{j \in Z} N_j^{-2} \prod_{j \in Z} L_j^{-2\delta_0} \sum_{k_{[Z]} : \sum_{j \in Z} \iota_j k_j = d_0} \left\langle \alpha_0 + \sum_{j \in Z} \iota_j |k_j|^2 \right\rangle^{-a_0}. \end{aligned}$$

Finally applying Cauchy-Schwartz in the variables $(k, \lambda, k_{[U \setminus Z]}, \lambda_{[U \setminus Z]})$, we deduce that

$$|\mathcal{X}|^2 \lesssim \left(\sum_{(1)} \int \langle \lambda \rangle^{2b} \prod_{j \in U \setminus Z} \langle \lambda_j \rangle^{2b} \cdot |\mathfrak{F}|^2 \right) \left(\sum_{(1)} \int \langle \lambda \rangle^{-2b} \prod_{j \in U \setminus Z} \langle \lambda_j \rangle^{-2b} \cdot |\mathfrak{G}|^2 \right),$$

where the first parenthesis (together with some factors from the second parenthesis) gives the product of all factors in (5.5) except $\mathcal{E}_1$, by using (3.27),

(5.4) and (5.12); the second parenthesis, after applying (5.19), integrating in $(\lambda, \lambda_{[U \setminus Z]})$ and plugging in (5.18), reduces to

$$\sum_{\substack{(k, k_{[U]}): \\ \iota_1 k_1 + \dots + \iota_n k_n = k+d}} \langle \Sigma - \alpha \rangle^{-a_0} \lesssim \mathcal{E}_1,$$

where $\Sigma = \iota_1 |k_1|^2 + \dots + \iota_n |k_n|^2 - |k|^2$ as in Corollary 4.6. This proves (5.5).

5.1.2. *A general $\mathcal{T}^* \mathcal{T}$ argument.* Now we prove (5.7), starting from (5.13). Note that due to (3.27) and (5.4), the bound for $\mathcal{X}$ would follow from the $\ell_{k_n}^2 L_{\lambda_n}^2 \rightarrow \ell_k^2 L_\lambda^2$ bound of the linear operator $\mathcal{T}$ with kernel

$$\begin{aligned} (5.20) \quad & \mathcal{T}_{k k_n}(\lambda, \lambda_n) \\ &= \sum_{k_{[V]}: \iota_1 k_1 + \dots + \iota_n k_n = k+d} \sum_{k_{[Z]}^*} \int d\lambda_{[V]} \cdot \eta \left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \alpha \right) \\ & \quad \times \left(\prod_{i=1}^p P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]}) \prod_{j \in Z} \frac{g_{k_j^*}^{\iota_j}}{\langle k_j^* \rangle} h_{k_j k_j^*}^{(j)}(\lambda_j)^\pm \prod_{j=n_1+1}^{n-1} [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j} \right) \langle \lambda \rangle^{-b} \langle \lambda_n \rangle^{-b}. \end{aligned}$$

We then calculate the kernel of $\mathcal{O} = \mathcal{T}^* \mathcal{T}$, which (similar to (5.14)) can be written as

$$(5.21) \quad \mathcal{O}_{k_n k'_n}(\lambda_n, \lambda'_n) = \langle \lambda_n \rangle^{-b} \langle \lambda'_n \rangle^{-b} \sum_{(2)} \int \mathfrak{F} \cdot \mathfrak{G},$$

where the symbol $\sum_{(2)} \int$ represents the sum in $(k_{[V \setminus Z]}, k'_{[V \setminus Z]})$ and integration in $(\lambda_{[V \setminus Z]}, \lambda'_{[V \setminus Z]})$, the factor $\mathfrak{F}$ is independent of $(k_n, k'_n, \lambda_n, \lambda'_n)$, and is now defined as

$$(5.22) \quad \mathfrak{F} := \prod_{i=1}^p \overline{P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]})} P_{k'_{[Y_i]}}^{(i)}(\lambda'_{[Y_i]}) \prod_{j=n_1+1}^{n-1} \overline{[v_{k_j}^{(j)}(\lambda_j)]^{\iota_j}} [v_{k'_j}^{(j)}(\lambda'_j)]^{\iota_j},$$

and the multilinear Gaussian $\mathfrak{G}$ is now given by

$$(5.23) \quad \mathfrak{G} := \sum_{(k_{[Z]}^*, k'_{[Z]}^*)} \prod_{j \in Z} g_{k_j^*}^{-\iota_j} g_{k'_j}^{\iota_j} \sum_{\substack{(k, k_{[Z]}, k'_{[Z]}): \\ \sum_{j \in Z} \iota_j k_j = k+d_0 \\ \sum_{j \in Z} \iota_j k'_j = k+d'_0}} \mathfrak{C},$$

with coefficient $\mathfrak{C}$ of form

(5.24)

$$\begin{aligned} \mathfrak{C} := & \int d\lambda d\lambda_{[Z]} d\lambda'_{[Z]} \cdot \langle \lambda \rangle^{-2b} \eta \left(\lambda, \lambda - |k|^2 - \sum_{j \in Z} \iota_j (\lambda_j - |k_j|^2) - \alpha_0 \right) \\ & \times \eta \left(\lambda, \lambda - |k|^2 - \sum_{j \in Z} \iota_j (\lambda'_j - |k'_j|^2) - \alpha'_0 \right) \prod_{j \in Z} \overline{\frac{1}{\langle k_j^* \rangle} h_{k_j k_j^*}^{(j)}(\lambda_j)^\pm} \frac{1}{\langle k'_j \rangle} h_{k'_j k'^*_j}^{(j)}(\lambda'_j)^\pm, \end{aligned}$$

where we now have

$$\begin{aligned} d_0 &:= d - \sum_{j \notin Z} \iota_j k_j, \quad d'_0 := d - \sum_{j \notin Z} \iota_j k'_j, \\ \alpha_0 &:= \alpha + \sum_{j \notin Z} \iota_j (\lambda_j - |k_j|^2), \quad \alpha'_0 := \alpha + \sum_{j \notin Z} \iota_j (\lambda'_j - |k'_j|^2). \end{aligned} \quad (5.25)$$

As in [Section 5.1.1](#) we may assume $|k_{2i-1} - k_{2i}| \lesssim R_i(N_*)^{C\kappa^{-1}}$ for $1 \leq i \leq p$ and similarly for k'_{2i-1} and k'_{2i} . The goal now is to estimate $\mathfrak{G}$ in (5.23). Let $L_+ = \max\{L_j : j \in Z\}$. In view of the power $(L_+)^{40n^2}$ on the right-hand side of (5.7), we may assume $N_j \gg (L_+)^2$ for each $j \in Z$; otherwise we simply sum over (k_j, k_j^*) and (k'_j, k'^*_j) and get rid of these variables. By the meshing argument in [Remark 4.3](#), we may reduce to $\lesssim (N_*)^{C\delta^{-7}}$ choices for $(k_{[U \setminus Z]}, k'_{[U \setminus Z]}, \lambda_{[U \setminus Z]}, \lambda'_{[U \setminus Z]})$; for each single choice, as $\mathfrak{C}$ is $\mathcal{B}_{\leq L_+}^+$ measurable and there is no pairing in $k_{[Z]}^*$ or $k'_{[Z]}^*$, we may apply [Lemma 4.1](#) and get

$$|\mathfrak{G}|^2 \lesssim \sum_{(k_{[Z \setminus W]}^*, k'_{[Z \setminus W]})} \left(\sum_{k_{a_l}^* = k_{b_l}^* (1 \leq l \leq s)} \sum_{\substack{(k, k_{[Z]}, k'_{[Z]}): \\ \sum_{j \in Z} \iota_j k_j = k + d_0 \\ \sum_{j \in Z} \iota_j k'_j = k + d'_0}} |\mathfrak{C}| \right)^2, \quad (5.26)$$

where $W = \{a_1, \dots, a_s\}$ and $W' = \{b_1, \dots, b_s\}$ are subsets of Z , and we have $N_{a_l} \sim N_{b_l}$ and $\iota_{a_l} = \iota_{b_l}$ for $1 \leq l \leq s$. As before we may assume $|k_j - k_j^*| \lesssim L_+(N_*)^{C\kappa^{-1}}$ and similarly for $k'_j - k'^*_j$, and due to the $(L_+)^{40n^2}$ factor we may then fix the values of $k_j - k_j^* = e_j$ and $k'_j - k'^*_j = e'_j$. Therefore $k'_{b_l} - k_{a_l} = e'_{b_l} - e_{a_l} := f_l$ is also fixed.

Now the outer sum in (5.26) can be viewed as a sum over $k_{[Z \setminus W]}$ and $k'_{[Z \setminus W]}$, and the inner sums can be viewed as a sum over $(k, k_{a_l}, k'_{b_l} : 1 \leq l \leq s)$ that satisfies $k'_{b_l} - k_{a_l} = f_l$. When all these k -variables are fixed, we have

$$\sup_{k_j, k_j^*} \|\langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{L_{\lambda_j}^2} \lesssim 1, \quad \sup_{k'_j, k'^*_j} \|\langle \lambda'_j \rangle^b h_{k'_j k'^*_j}^{(j)}(\lambda'_j)\|_{L_{\lambda'_j}^2} \lesssim 1,$$

due to the first bound in (3.26). Using the algebra property of the norm $\|\langle \lambda \rangle^b h(\lambda)\|_{L^2}$ under convolution, we have

$$\begin{aligned}
 |\mathfrak{C}| &\lesssim \prod_{j \in Z} N_j^{-2} \int d\lambda \cdot \langle \lambda \rangle^{-2b} \\
 (5.27) \quad &\times \left\langle \lambda - |k|^2 + \sum_{j \in Z} \iota_j |k_j|^2 - \alpha_0 \right\rangle^{-b} \left\langle \lambda - |k|^2 + \sum_{j \in Z} \iota_j |k'_j|^2 - \alpha'_0 \right\rangle^{-b} \\
 &\lesssim \prod_{j \in Z} N_j^{-2} \cdot \left\langle |k|^2 - \sum_{j \in Z} \iota_j |k_j|^2 + \alpha_0 \right\rangle^{-b} \left\langle |k|^2 - \sum_{j \in Z} \iota_j |k'_j|^2 + \alpha'_0 \right\rangle^{-b}.
 \end{aligned}$$

With (5.27) we can now bound $\mathfrak{G}$ by

$$\begin{aligned}
 (5.28) \quad |\mathfrak{G}|^2 &\lesssim (L_+)^{40n^2} \prod_{j \in Z} N_j^{-4} \\
 &\times \sum_{(k_{[Z \setminus W]}, k'_{[Z \setminus W']})} \left(\sum_{(k, k_{a_l}, k'_{b_l} : 1 \leq l \leq s)} \left\langle |k|^2 - \sum_{j \in Z \setminus W} \iota_j |k_j|^2 - \sum_{l=1}^s \iota_{a_l} |k_{a_l}|^2 + \alpha_0 \right\rangle^{-b} \right. \\
 &\times \left. \left\langle |k|^2 - \sum_{j \in Z \setminus W'} \iota_j |k'_j|^2 - \sum_{l=1}^s \iota_{b_l} |k'_{b_l}|^2 + \alpha'_0 \right\rangle^{-b} \right)^2.
 \end{aligned}$$

Multiplying out the square we get

$$\begin{aligned}
 (5.29) \quad |\mathfrak{G}|^2 &\lesssim (L_+)^{40n^2} \prod_{j \in Z} N_j^{-4} \sum_{(k_{[Z \setminus W]}, k'_{[Z \setminus W']})} \sum_{\substack{(k, k_{a_l}, k'_{b_l} : 1 \leq l \leq s) \\ (\mathring{k}, \mathring{k}_{a_l}, \mathring{k}'_{b_l} : 1 \leq l \leq s)}} \langle \Upsilon \rangle^{-b} \cdot \langle \mathring{\Upsilon} \rangle^{-b} \cdot \langle \Upsilon' \rangle^{-b} \cdot \langle \mathring{\Upsilon}' \rangle^{-b},
 \end{aligned}$$

where

$$\begin{aligned}
 \Upsilon &= |k|^2 - \sum_{j \in Z \setminus W} \iota_j |k_j|^2 - \sum_{l=1}^s \iota_{a_l} |k_{a_l}|^2 + \alpha_0, \\
 \mathring{\Upsilon} &= |\mathring{k}|^2 - \sum_{j \in Z \setminus W} \iota_j |\mathring{k}_j|^2 - \sum_{l=1}^s \iota_{a_l} |\mathring{k}_{a_l}|^2 + \alpha_0, \\
 \Upsilon' &= |k|^2 - \sum_{j \in Z \setminus W'} \iota_j |k'_j|^2 - \sum_{l=1}^s \iota_{b_l} |k'_{b_l}|^2 + \alpha'_0, \\
 \mathring{\Upsilon}' &= |\mathring{k}|^2 - \sum_{j \in Z \setminus W'} \iota_j |\mathring{k}'_j|^2 - \sum_{l=1}^s \iota_{b_l} |\mathring{k}'_{b_l}|^2 + \alpha'_0,
 \end{aligned}$$

with $\iota_{a_l} = \iota_{b_l}$, α_0 and α'_0 as in (5.25). The variables in the summation (5.29) verify the following linear equations:

$$\begin{aligned}
(5.30) \quad & \sum_{j \in Z \setminus W} \iota_j k_j + \sum_{l=1}^s \iota_{a_l} k_{a_l} - k = \sum_{j \in Z \setminus W} \iota_j k_j + \sum_{l=1}^s \iota_{a_l} \overset{\circ}{k}_{a_l} - \overset{\circ}{k} = d_0, \\
& \sum_{j \in Z \setminus W'} \iota_j k'_j + \sum_{l=1}^s \iota_{b_l} k'_{b_l} - k = \sum_{j \in Z \setminus W'} \iota_j k'_j + \sum_{l=1}^s \iota_{b_l} \overset{\circ}{k}'_{b_l} - \overset{\circ}{k} = d'_0,
\end{aligned}$$

with d_0 and d'_0 as in (5.25), as well as $k'_{b_l} - k_{a_l} = \overset{\circ}{k}'_{b_l} - \overset{\circ}{k}_{a_l} = f_l$.

By Cauchy-Schwartz, we may replace the summand on the right-hand side of (5.29) by $\langle \Upsilon \rangle^{-2b} \cdot \langle \overset{\circ}{\Upsilon}' \rangle^{-2b}$ (or by $\langle \overset{\circ}{\Upsilon} \rangle^{-2b} \cdot \langle \Upsilon' \rangle^{-2b}$, which is treated similarly by symmetry). Now going back to (5.21) and applying Cauchy-Schwartz in the variables $(k_{[V \setminus Z]}, k'_{[V \setminus Z]}, \lambda_{[V \setminus Z]}, \lambda'_{[V \setminus Z]})$, we get

$$\begin{aligned}
|\mathcal{X}|^4 & \lesssim N_n^{-4(1-\gamma)} \sum_{k_n, k'_n} \int d\lambda_n d\lambda'_n |\mathcal{O}_{k_n k'_n}(\lambda_n, \lambda'_n)|^2 \\
& \lesssim N_n^{-4(1-\gamma)} \left(\sum_{(2)} \int \prod_{j \in V \setminus Z} \langle \lambda_j \rangle^{2b} \langle \lambda'_j \rangle^{2b} \cdot |\mathfrak{F}|^2 \right) \\
& \quad \times \left(\sum_{k_n, k'_n} \int d\lambda_n d\lambda'_n \cdot \langle \lambda_n \rangle^{-2b} \langle \lambda'_n \rangle^{-2b} \sum_{(2)} \int \prod_{j \in V \setminus Z} \langle \lambda_j \rangle^{-2b} \langle \lambda'_j \rangle^{-2b} \cdot |\mathfrak{G}|^2 \right).
\end{aligned}$$

The first parenthesis (together with some factors from the second parenthesis) give the product of all factors in (5.7) except $\mathcal{E}_1 \mathcal{E}_1^+$, by using (3.27) and (5.12). The second parenthesis, after applying (5.29) with the summand $\langle \Upsilon \rangle^{-b} \cdot \langle \overset{\circ}{\Upsilon} \rangle^{-b} \cdot \langle \Upsilon' \rangle^{-b} \cdot \langle \overset{\circ}{\Upsilon}' \rangle^{-b}$ replaced by $\langle \Upsilon \rangle^{-2b} \langle \overset{\circ}{\Upsilon}' \rangle^{-2b}$, integrating in $\lambda_{[U \setminus Z]}$ and $\lambda'_{[U \setminus Z]}$, and plugging in (5.25), reduces to

$$(5.31) \quad \sum_{(k_{[U \setminus W]}, k'_{[U \setminus W']})} \sum_{\substack{(k, k_{a_l}, k'_{b_l} : 1 \leq l \leq s) \\ (\overset{\circ}{k}, \overset{\circ}{k}_{a_l}, \overset{\circ}{k}'_{b_l} : 1 \leq l \leq s)}} \langle \Sigma - \alpha \rangle^{-2b} \langle \overset{\circ}{\Sigma}' - \alpha \rangle^{-2b},$$

where Σ and $\overset{\circ}{\Sigma}'$ are respectively

$$(5.32) \quad \Sigma = \sum_{j \notin W} \iota_j |k_j|^2 + \sum_{l=1}^s \iota_{a_l} |k_{a_l}|^2 - |k|^2, \quad \overset{\circ}{\Sigma}' = \sum_{j \notin W'} \iota_j |k'_j|^2 + \sum_{l=1}^s \iota_{b_l} |\overset{\circ}{k}'_{b_l}|^2 - |\overset{\circ}{k}|^2,$$

and the variables in the summation satisfy

$$\begin{aligned}
(5.33) \quad & \sum_{j \notin W} \iota_j k_j + \sum_{l=1}^s \iota_{a_l} k_{a_l} - k = \sum_{j \notin W} \iota_j k_j + \sum_{l=1}^s \iota_{a_l} \overset{\circ}{k}_{a_l} - \overset{\circ}{k} = d, \\
& \sum_{j \notin W'} \iota_j k'_j + \sum_{l=1}^s \iota_{b_l} k'_{b_l} - k = \sum_{j \notin W'} \iota_j k'_j + \sum_{l=1}^s \iota_{b_l} \overset{\circ}{k}'_{b_l} - \overset{\circ}{k} = d,
\end{aligned}$$

as well as $k'_{b_l} - k_{a_l} = k'_{b_l} - k_{a_l} = f_l$. Now, when $k_{[U \setminus W]}$ and $(k, k_{a_l} : 1 \leq l \leq s)$ are fixed, the sum of $\langle \Sigma' - \alpha \rangle^{-2b}$ over $k'_{[U \setminus W']}$ and $(\overset{\circ}{k}, \overset{\circ}{k}_{b_l} : 1 \leq l \leq s)$ can be bounded by $\mathcal{E}_1^+$ with $A = U \setminus W'$ in (4.23) due to (5.33); on the other hand, the sum of $\langle \Sigma - \alpha \rangle^{-2b}$ over $k_{[U \setminus W]}$ and $(k, k_{a_l} : 1 \leq l \leq s)$ can be bounded by $\mathcal{E}_1$. This bounds the sum (5.31) by $\mathcal{E}_1 \mathcal{E}_1^+$ and proves (5.7).

5.1.3. *A special $\mathcal{T}^* \mathcal{T}$ argument.* Assume now $N_n = N^{(1)}$ and $N_* \gtrsim N^{(2)}$. Again we only need to study the operator $\mathcal{T}$ given by the kernel (5.20); note that $\mathcal{T}_{kk_n}(\lambda, \lambda_n)$ is supported in the set $\{(k, k_n) : |k - \iota_n k_n + d| \lesssim N^{(2)}\}$. By the standard orthogonality argument it suffices to prove the same operator bound for $\widetilde{\mathcal{T}}$ which is $\mathcal{T}$ restricted to the set $\{k : |k - f| \leq N^{(2)}\}$, *uniformly* in $f \in \mathbb{Z}^2$. Below we will fix an f and denote $\widetilde{\mathcal{T}}$ still by $\mathcal{T}$, so that in any summations below we may assume $|k - f| \lesssim N^{(2)}$ and $|k_n - \iota_n(f + d)| \lesssim N^{(2)}$ (same for k'_n). At this point the parameter $N^{(1)}$ or N_n no longer explicitly appears in the estimate, so the set of ω we remove will be independent of it. Also we can prove (5.7), with N_n replaced by $N^{(2)}$ in both $\mathcal{E}_1$ and $\mathcal{E}_1^{\text{ex}}$, by essentially repeating the proof in Section 5.1.2 above (and making the bound uniform in f in exactly the same way as below); it remains to prove (5.9).

We start with (5.21) and now look for further pairings in $(k_{[Z]}^*, k_{[Z]}^*)$ in the expression $\mathfrak{G}$ given by (5.23). By repeating the same reduction step in Section 5.1.1, we can find two partitions $(X_1, \dots, X_q, W)$ and $(X'_1, \dots, X'_q, W')$ ¹² of the set Z , where $2p + q \leq n_1$, such that N_j are all comparable for j in each $X_l \cup X'_l$, and further reduce (5.21) to a sum

$$(5.34) \quad \mathcal{O}_{k_n k'_n}(\lambda_n, \lambda'_n) = \langle \lambda_n \rangle^{-b} \langle \lambda'_n \rangle^{-b} \sum_{(3)} \int \mathfrak{F}^+ \cdot \mathfrak{G}^+,$$

where the symbol $\sum_{(3)} \int$ represents the sum in $k_{[V \setminus W]}$ and $k'_{[V \setminus W']}$ and integration in $\lambda_{[V \setminus W]}$ and $\lambda'_{[V \setminus W']}$, the factor $\mathfrak{F}^+$ is independent of $(k_n, k'_n, \lambda_n, \lambda'_n)$,

$$(5.35) \quad \begin{aligned} \mathfrak{F}^+ &= \prod_{i=1}^p \overline{P_{k_{[Y_i]}}^{(i)}(\lambda_{[Y_i]})} P_{(k'_{[Y_i]})}^{(i)}(\lambda'_{[Y_i]}) \\ &\times \prod_{j=n_1+1}^{n-1} \overline{[v_{k_j}^{(j)}(\lambda_j)]^{\iota_j}} [v_{k'_j}^{(j)}(\lambda'_j)]^{\iota_j} \prod_{l=1}^q Q_{k_{[X_l]}, k'_{[X'_l]}}^{(l)}(\lambda_{[X_l]}, \lambda'_{[X'_l]}), \end{aligned}$$

$$(5.36) \quad \begin{aligned} &Q_{k_{[X_l]}, k'_{[X'_l]}}^{(l)}(\lambda_{[X_l]}, \lambda'_{[X'_l]}) \\ &:= \sum_{k^*} \prod_{j \in X_l} h_{k_j k^*}^{(j)}(\lambda_j)^{\pm} \prod_{j \in X'_l} h_{k'_j k^*}^{(j)}(\lambda'_j)^{\pm} \langle k^* \rangle^{-q_3} (g_{k^*})^{q_1} (\overline{g_{k^*}})^{q_2}, \end{aligned}$$

¹²This W and W' are different from the W and W' of Section 5.1.2.

where $q_1 + q_2 = q_3 = |X_l| + |X'_l|$, and the P factors are defined as in (5.11). We may also fix $a_l \in X_l$ and $b_l \in X'_l$ such that $\iota_{a_l} = \iota_{b_l}$; without loss of generality, assume $b_l = 2p + l$ for $1 \leq l \leq q$. We can bound (5.36) just like we bound (5.11) in (5.12), except that now it is possible to have $X_l \cup X'_l \subset \mathcal{G}$. Let $R_{2p+l} = \max\{L_j : j \in X_l \cup X'_l\} \gtrsim L_{2p+l}$. Then the same argument as in (5.12) gives

$$\begin{aligned}
 & \|Q_{k_{[X_l]}, k'_{[X'_l]}}^{(l)}(\lambda_{[X_l]}, \lambda'_{[X'_l]}) \prod_{j \in X_l} \langle \lambda_j \rangle^b \prod_{j \in X'_l} \langle \lambda'_j \rangle^b\|_{\ell^2_{k_{[X_l]}, k'_{[X'_l]}}}^2 L_{\lambda_{[X_l]}, \lambda'_{[X'_l]}}^2 \\
 (5.37) \quad & \lesssim \prod_{j \in X_l} N_j^{-2} \prod_{j \in X'_l} N_j^{-2} \cdot N_{2p+l}^{2+2\gamma_0} R_{2p+l}^{-2}.
 \end{aligned}$$

Finally, the multilinear Gaussian $\mathfrak{G}^+$ is given by

$$\begin{aligned}
 \mathfrak{G}^+ &= \sum_{(k_{[W]}, k'_{[W']})} \sum_{(k_{[W]}^*, k'_{[W']}^*)} \int d\lambda_{[W]} d\lambda'_{[W']} \\
 (5.38) \quad & \times \prod_{j \in W} \frac{g_{k_j^*}^{\iota_j}}{\langle k_j^* \rangle} \langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)^\pm \prod_{j \in W'} \frac{g_{k'_j}^{\iota_j}}{\langle k'_j \rangle} \langle \lambda'_j \rangle^b h_{k'_j k'_j}^{(j)}(\lambda'_j)^\pm \cdot \mathfrak{A}^+,
 \end{aligned}$$

where there is no pairing among $(k_{[W]}^*, k'_{[W']}^*)$, and coefficient $\mathfrak{A}^+$ of form

$$\begin{aligned}
 \mathfrak{A}^+ &= \sum_{\substack{k: \iota_1 k_1 + \dots + \iota_n k_n = k+d \\ \iota_1 k'_1 + \dots + \iota_n k'_n = k+d}} \int d\lambda \cdot \eta\left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j(\lambda_j - |k_j|^2) - \alpha\right) \\
 & \times \langle \lambda \rangle^{-2b} \eta\left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j(\lambda'_j - |k'_j|^2) - \alpha\right) \prod_{j \in W} \langle \lambda_j \rangle^{-b} \prod_{j \in W'} \langle \lambda'_j \rangle^{-b},
 \end{aligned}
 \tag{5.39}$$

where the sum is over a single variable k . As before we may also assume $|k_j - k^*| \lesssim R_{2p+l}(N_*)^{C\kappa^{-1}}$ for $j \in X_l$ in (5.36) and similarly for k'_j and $j \in X'_l$ so, in particular, $|k_{a_l} - k'_{2p+l}| \lesssim R_{2p+l}(N_*)^{C\kappa^{-1}}$.

The goal now is to estimate $\mathfrak{G}^+$. As before, we need to reduce to $(N_*)^{C\delta^{-7}}$ choices for $(k_{[U \setminus W]}, \lambda_{[U \setminus W]})$ and $(k'_{[U \setminus W]}, \lambda'_{[U \setminus W]})$, and f . By the meshing argument (Remark 4.3), we may assume $|\lambda| \lesssim (N_*)^{\delta^{-7}}$ and $|\lambda_j| \lesssim (N_*)^{\delta^{-7}}$ for $j \notin W$ (similarly for λ'_j) and get rid of these parameters; in the same way we may also fix $k_{[V \setminus W]}$ and $k'_{[V \setminus W]}$, as well as $f + d - \iota_n k_n$ and $f + d - \iota_n k'_n$.

Letting $k = f + g$, we can rewrite

$$\begin{aligned}
\mathfrak{A}^+ &= \sum_{\substack{|g| \leq N^{(2)} \\ \sum_{j \in W} \iota_j k_j = g + d_0 \\ \sum_{j \in W'} \iota_j k'_j = g + d'_0}} \mathbf{1} \\
&\times \int d\lambda \cdot \eta \left(\lambda, \lambda - 2f \cdot g - |g|^2 - \sum_{j \in V} \iota_j (\lambda_j - |k_j|^2) + \beta(f) + \gamma \right) \\
(5.40) \quad &\times \langle \lambda \rangle^{-2b} \eta \left(\lambda, \lambda - 2f \cdot g - |g|^2 - \sum_{j \in V} \iota_j (\lambda'_j - |k'_j|^2) + \beta'(f) + \gamma' \right) \\
&\times \prod_{j \in W} \langle \lambda_j \rangle^{-b} \prod_{j \in W'} \langle \lambda'_j \rangle^{-b},
\end{aligned}$$

where $\gamma, \gamma' \in [0, 1]$ are fixed, $d_0, d'_0 \in \mathbb{Z}^2$, and $\beta(f), \beta'(f)$ are fixed integer-valued functions of f . We may assume $|d_0|, |d'_0| \lesssim N^{(2)}$ since otherwise $\mathfrak{A}^+ \equiv 0$. Then we may fix them and see that f enters the whole expression only through the function $-2f \cdot g + \beta(f)$; moreover we may restrict g to the set where $|-2f \cdot g + \beta(f)| \leq (N_*)^{\delta-7}$ since otherwise either λ or one λ_j must be large and we close as before. The reduction to finitely many cases can then be done by invoking the following claim, which will be proved at the end of this section.

CLAIM 5.2. *Let the function $F_{f,\beta}(g) := -2f \cdot g + \beta$, with the particular domain $\text{Dom}(F_{f,\beta}) = \{|g| \leq N^{(2)} : |-2f \cdot g + \beta| \leq (N_*)^{\delta-7}\}$. Then when $f \in \mathbb{Z}^2$ and $\beta \in \mathbb{Z}$ varies, the function $F_{f,\beta}$ (together with its domain) has finitely many, and in fact $\lesssim (N_*)^{C\delta-7}$, possibilities.*

From now on we may fix the value of f . By removing a set of probability $\leq C_\theta e^{-(\tau^{-1}N_*)^\theta}$, we can apply [Lemma 4.2](#) and (recalling that $a_0 = 2b - 10\delta^6$) conclude that

$$\begin{aligned}
|\mathfrak{G}^+|^2 &\lesssim \prod_{j \in W} N_j^{-2} \prod_{j \in W'} N_j^{-2} \sum_{(k_{[W]}, k'_{[W']})} \int d\lambda_{[W]} d\lambda'_{[W']} \\
(5.41) \quad &\times \prod_{j \in W} \langle \lambda_j \rangle^{-a_0} \prod_{j \in W'} \langle \lambda'_j \rangle^{-a_0} \left[\sum_{\substack{k: \iota_1 k_1 + \dots + \iota_n k_n = k + d \\ \iota_1 k'_1 + \dots + \iota_n k'_n = k + d}} \int d\lambda \right. \\
&\times \langle \lambda \rangle^{-2b} \eta \left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \alpha \right) \\
&\left. \times \eta \left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda'_j - |k'_j|^2) - \alpha \right) \right]^2.
\end{aligned}$$

The integral over λ gives a factor

$$\left\langle \Sigma - \sum_{j=1}^n \iota_j \lambda_j - \alpha \right\rangle^{-b} \left\langle \Sigma' - \sum_{j=1}^n \iota_j \lambda'_j - \alpha \right\rangle^{-b},$$

where Σ and Σ' are defined as

$$\Sigma = \sum_{j=1}^n \iota_j |k_j|^2 - |k|^2, \quad \Sigma' = \sum_{j=1}^n \iota_j |k'_j|^2 - |k|^2.$$

Since there is only one value of k in the summation, we can reduce

(5.42)

$$\begin{aligned} |\mathfrak{G}^+|^2 &\lesssim \sum_k \sum_{(k_{[W]}, k'_{[W']})} \int d\lambda_{[W]} d\lambda'_{[W']} \prod_{j \in W} \langle \lambda_j \rangle^{-a_0} \prod_{j \in W'} \langle \lambda'_j \rangle^{-a_0} \\ &\quad \times \prod_{j \in W} N_j^{-2} \prod_{j \in W'} N_j^{-2} \left\langle \Sigma - \sum_{j=1}^n \iota_j \lambda_j - \alpha \right\rangle^{-2b} \left\langle \Sigma' - \sum_{j=1}^n \iota_j \lambda'_j - \alpha \right\rangle^{-2b} \\ &\lesssim \prod_{j \in W} N_j^{-2} \prod_{j \in W'} N_j^{-2} \\ &\quad \times \sum_k \sum_{(k_{[W]}, k'_{[W']})} \left\langle \Sigma - \sum_{j \notin W} \iota_j \lambda_j - \alpha \right\rangle^{-a_0} \left\langle \Sigma' - \sum_{j \notin W'} \iota_j \lambda'_j - \alpha \right\rangle^{-a_0}, \end{aligned}$$

where in the summation over k and $(k_{[W]}, k'_{[W']})$ we assume that $\iota_1 k_1 + \dots + \iota_n k_n = \iota_1 k'_1 + \dots + \iota_n k'_n = k + d$. Returning to (5.34), by applying Cauchy-Schwartz in the variables $(k_{[V \setminus W]}, \lambda_{[V \setminus W]})$ and $(k'_{[V \setminus W]}, \lambda'_{[V \setminus W]})$ we conclude as before that

$$\begin{aligned} |\mathcal{X}|^4 &\lesssim (N^{(1)})^{-4(1-\gamma)} \left(\sum_{(3)} \int \prod_{j \in V \setminus W} \langle \lambda_j \rangle^{2b} \prod_{j \in V \setminus W'} \langle \lambda'_j \rangle^{2b} \cdot |\mathfrak{F}|^2 \right) \\ &\times \left(\sum_{k_n, k'_n} \int d\lambda_n d\lambda'_n \cdot \langle \lambda_n \rangle^{-2b} \langle \lambda'_n \rangle^{-2b} \sum_{(3)} \int \prod_{j \in V \setminus W} \langle \lambda_j \rangle^{-2b} \prod_{j \in V \setminus W'} \langle \lambda'_j \rangle^{-2b} \cdot |\mathfrak{G}|^2 \right). \end{aligned}$$

The first parenthesis (together with some factors from the second parenthesis) gives the product of all factors in (5.9) except $\widetilde{\mathcal{E}}_1 \widetilde{\mathcal{E}}_1^+$, by using (3.27), (5.12) and (5.37). The second parenthesis, after applying (5.42) and integrating in $\lambda_{[U \setminus W]}$ and $\lambda'_{[U \setminus W']}$, reduces to

$$(5.43) \quad \sum_k \sum_{k_{[U]}: \iota_1 k_1 + \dots + \iota_n k_n = k+d} \sum_{k'_{[U]}: \iota_1 k'_1 + \dots + \iota_n k'_n = k+d} \langle \Sigma - \alpha \rangle^{-a_0} \langle \Sigma' - \alpha \rangle^{-a_0}.$$

Now when $(k, k_{[U]})$ are fixed, the sum of $\langle \Sigma' - \alpha \rangle^{-a_0}$ over $k'_{[U]}$ can be bounded by $\widetilde{\mathcal{E}}_1^+$ with $A = \{1, \dots, n\}$ in (4.23), due to the linear equation $\iota_1 k'_1 + \dots + \iota_n k'_n =$

$k+d$, the fact that k'_{2p+l} belongs to a disc of radius $O(R_{2p+l}(N_*)^{C\kappa^{-1}})$ once k_{al} is fixed, and the fact that $|\iota_n k'_n - f - d| \lesssim N^{(2)}$. Moreover, the sum of $\langle \Sigma - \alpha \rangle^{-a_0}$ over $(k, k_{[U]})$ can be bounded by $\widetilde{\mathcal{E}}_1$, due to the fact that $|\iota_n k_n - f - d| \lesssim N^{(2)}$. This then bounds (5.43) by $\widetilde{\mathcal{E}}_1 \mathcal{E}_1^+$ and proves (5.9).

Proof of Claim 5.2. Let $D = \text{Dom}(F_{f,\beta})$. If D contains three points g_1, g_2, g_3 that are not collinear, then we have $|f \cdot (g_1 - g_2)| \lesssim (N_*)^{\delta^{-7}}$, $|f \cdot (g_1 - g_3)| \lesssim (N_*)^{\delta^{-7}}$, and that $g_1 - g_2$ and $g_1 - g_3$ are linearly independent. This implies that $|f| \lesssim (N_*)^{2\delta^{-7}}$ and hence $|\beta| \lesssim (N_*)^{3\delta^{-7}}$ so the result is trivial. Now let us assume D is contained in a line ℓ ; we may assume ℓ contains at least two points in the set $\{g : |g| \leq N^{(2)}\}$, otherwise D is at most a singleton and the result is also trivial. Then the integer points in ℓ can be written as $p + q\sigma$, where $(p, q) \in (\mathbb{Z}^2)^2$ has at most $(N^{(2)})^{10}$ choices (so we may fix them), and hence $D = \{p + q\sigma : |p + q\sigma| \leq N^{(2)}, |a\sigma + b| \leq (N_*)^{\delta^{-7}}\}$, where a and b are integers. Again, as $|D| \geq 2$, we know that $|a| \lesssim (N_*)^{2\delta^{-7}}$ and $|b| \lesssim (N_*)^{3\delta^{-7}}$, so $F_{f,\beta}$ indeed has $\lesssim (N_*)^{C\delta^{-7}}$ possibilities, as claimed. $\square$

Remark 5.3. For later use, we will also consider the following variant of $\mathcal{X}$ (same for $\mathcal{Y}$):

$$(5.44) \quad \mathcal{X}^+ := \sum_{\substack{(k, k_1, \dots, k_n) \\ \iota_1 k_1 + \dots + \iota_n k_n = k+d}} \int d\lambda d\lambda_1 \dots d\lambda_n d\mu_1 \dots d\mu_s \\ \times \eta \left(\lambda, \lambda - |k|^2 - \sum_{j=1}^n \iota_j (\lambda_j - |k_j|^2) - \sum_{j=1}^s \mu_j - \alpha \right) \overline{v_k(\lambda)} \prod_{j=1}^n [v_{k_j}^{(j)}(\lambda_j)]^{\iota_j} \prod_{j=1}^s w_j(\mu_j),$$

where each w_j satisfies

$$\|\langle \mu_j \rangle^b w_j(\mu_j)\|_{L_{\mu_j}^2} \lesssim 1.$$

Then $\mathcal{X}^+$ will satisfy exactly the same estimates as $\mathcal{X}$ (same for $\mathcal{Y}$). In fact we can introduce a “virtual” variable l_j that takes a single value and view $w_j(\mu_j)$ as a function of l_j and μ_j that has type (D), and repeat all the above proofs with these new variables.

5.2. Proof of Proposition 3.4. Armed with Corollary 4.6 and Proposition 5.1, we can now prove Proposition 3.4. Recall that we will abuse notation and write $(v_{k_j}^{(j)})(\lambda_j)$ instead of $(\widetilde{v^{(j)}})_{k_j}(\lambda_j)$. We will proceed in three steps; note that as before, in the proof below we will ignore any factor $\lesssim \tau^{-\theta}(N_*)^{C\kappa^{-1}}$.

Step 1: Reduction to estimating $\mathcal{X}$ and $\mathcal{Y}$. First notice that, when the set of pairings among the variables involved in $\mathcal{N}_n$ is fixed, the coefficient in $\mathcal{N}_n$ will be a constant (see Remark 2.3). By Lemma 2.6, we may replace $\mathcal{I}$ by $\mathcal{J}$ in

all estimates. Now by definition of the relevant norms, the kernel bound (2.30) and duality, we can reduce the desired estimates to the estimates of quantities of form $\mathcal{X}$ (for parts (1) and (2)) and $\mathcal{Y}$ (for part (3)) defined in (5.1) and (5.2) with $d = \alpha = 0$. The only exceptions are that the functions v and y introduced by duality only satisfy weaker bounds

$$(5.45) \quad \|\langle \lambda \rangle^{1-b_1} v_k(\lambda)\|_{\ell_k^2 L_\lambda^2} \lesssim 1, \quad \|\langle \lambda \rangle^{1-b_1} \langle \lambda' \rangle^b y_{kk'}(\lambda, \lambda')\|_{\ell_{k,k'}^2 L_{\lambda,\lambda'}^2} \lesssim 1,$$

instead of (5.4), and that there may be pairings in $\mathcal{X}$ and $\mathcal{Y}$ (but they will always be over-paired). Now if $|\lambda| \leq (N_*)^{C_0}$, where C_0 is a large constant depending only on n , then since $b_1 - b \sim 2b - 1 \sim \kappa^{-1} \sim \delta^4$, we can replace the power $\langle \lambda \rangle^{1-b_1}$ by $\langle \lambda \rangle^b$ in (5.45) to match (5.4), at a price of losing a factor $(N_*)^{C\kappa^{-1}}$ which is acceptable. Now we will assume $|\lambda| \geq (N_*)^{C_0}$; below we will only consider part (1) of Proposition 3.4, since we have $N_* \gtrsim N^{(1)}$ in part (2) and $N_* \gtrsim N_0$ in part (3). The proof will be similar and much easier.

Here the point is to use the weight $\langle \lambda \rangle^{1-b_1}$ in (5.45) to gain a power $\geq (N_*)^{-\frac{C_0}{3}}$, after which we still can assume $\|v_k(\lambda)\|_{\ell_k^2 L_\lambda^2} \lesssim 1$. In view of this gain and the assumption $N_* \gtrsim N^{(2)}$, we may fix the values of k_j and/or k_j^* for each $1 \leq j \leq n-1$. Moreover, when k_j and k_j^* are fixed, the resulting function in λ_j (we will call them $w_j(\lambda_j)$) satisfies $\|\langle \lambda_j \rangle^b w_j(\lambda_j)\|_{L_{\lambda_j}^2} \lesssim 1$, which implies the corresponding $L_{\lambda_j}^1$ bound, so we may fix $\lambda_j (1 \leq j \leq n-1)$ also. Finally as

$$\int \langle \lambda_n \rangle^{2b} \|v_{k_n}^{(n)}(\lambda_n)\|_{\ell_{k_n}^2}^2 d\lambda_n = \|\langle \lambda_n \rangle^b v_{k_n}^{(n)}(\lambda_n)\|_{\ell_{k_n}^2 L_{\lambda_n}^2}^2 \lesssim (N^{(1)})^{-2(1-\gamma)},$$

we may also fix the value of λ_n , and reduce to

$$\mathcal{X} = \sum_k \int v_k(\lambda) d\lambda \cdot \eta(\lambda, \lambda - F(k)) G_{k-d'},$$

where $\|G\|_{\ell^2} \sim (N^{(1)})^{-1+\gamma}$ and $F(k)$ is a function of k which, as well as d' , depends on the choice of the other fixed variables. By first integrating in λ using Cauchy-Schwartz and (5.3), then summing in k using Cauchy-Schwartz again, we deduce that

$$|\mathcal{X}| \lesssim \|v_k(\lambda)\|_{\ell_k^2 L_\lambda^2} \cdot \|G\|_{\ell^2} \lesssim (N^{(1)})^{-1+\gamma},$$

which suffices in view of the gain $(N_*)^{-C_0/3}$.

Step 2: The no-pairing case. We have now reduced Proposition 3.4 to the estimates for the quantities $\mathcal{X}$ and $\mathcal{Y}$. If we assume there is no pairing, then we can apply Proposition 5.1 and then Corollary 4.6. Recall the new parameters such as p , q and R_j defined in Proposition 5.1; denote $L_+ = \max(L_{2p+1}, \dots, L_{n_1})$

and $N_+ = \max(N_{n_1+1}, \dots, N_n)$. Also when we talk about an estimate in [Proposition 4.5](#) we are actually talking about its counterpart in [Corollary 4.6](#).

In part (1), by combining (5.7), (4.25) and (4.25) with the improvement factor (4.29), with N_n replaced by $N^{(2)}$ in both places, we obtain that

$$|\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma} (N^{(2)})^{C\gamma} (L_+)^{40n^3} (N^{(2)})^{-\frac{1}{4}};$$

on the other hand, by combining (5.9), (4.25) and (4.25) with the improvement factor (4.29), with the changes adapted to $\widetilde{\mathcal{E}}_1$ and $\widetilde{\mathcal{E}}_1^{\text{ex}}$ indicated in [Proposition 5.1](#), we obtain that

$$|\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma} (N^{(2)})^{C\gamma} (L_+)^{-\frac{1}{4}},$$

noticing that $R_{2p+l} \gtrsim L_{2p+l}$ for $1 \leq l \leq q$ and $N_j \gtrsim L_j$ for $2p+q+1 \leq j \leq n_1$. Interpolating the above two bounds then gives (3.28).

In parts (2) and (3), we have $N^{(1)} \sim N_a$ and $a \in \mathcal{G} \cup \mathcal{C}$; in particular, the extra factor (4.29) is bounded by $(N_+)^{-1}$. Note that in case (3) we may have $a \in \mathcal{D}$ but in this case the extra factor will be replaced by $(N^{(1)})^{-1}$. By combining (5.5) and (4.25) we obtain (noticing that $N_{PR} \lesssim N^{(2)}$)

$$|\mathcal{X}| \lesssim (N^{(1)} N^{(2)})^{-\frac{1}{2}} (N^{(2)})^{\gamma_0} (N_+)^{C\gamma} (L_+)^{-\delta_0},$$

and by combining (5.7) and (4.25) together with the improvement factor (4.29) we obtain that

$$|\mathcal{X}| \lesssim (N^{(1)} N^{(2)})^{-\frac{1}{2}} (N^{(2)})^{\gamma_0} (N_+)^{C\gamma} (L_+)^{40n^3} (N_+)^{-\frac{1}{4}},$$

and interpolating the above two bounds gives (3.29). In the same way (3.30) follows from (5.5), (5.7), (4.26) and (4.26) with the improvement factor (4.29), and (3.33) follows from (5.6), (5.8), (4.27) and (4.27) with the suitable improvement factor.

Finally consider (3.31); here we will define $N' = \max^{(2)}(N_{n_1+1}, \dots, N_n)$. Note that $\alpha = 0$, so by combining (5.5) and (4.28) we get

$$(5.46) \quad |\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma_0} (N_+)^{\gamma} (N')^{C\gamma} (L_+)^{-\delta_0};$$

on the other hand, by combining (5.7) and either (4.28) with the improvement factor (4.29) or (4.30), we get that either

$$(5.47) \quad |\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma_0} (N_+)^{\gamma} (N')^{C\gamma} (L_+)^{40n^3} (N_+)^{-\frac{1}{4}},$$

or

$$(5.48) \quad |\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma_0} (N_+)^{\gamma} (N')^{C\gamma} (L_+)^{40n^3} \min((N')^{-\frac{1}{4}}, (N^{(1)})^{\frac{1}{4}} (N_+)^{-\frac{1}{2}}).$$

Clearly interpolating (5.46) and (5.47) gives (3.31); suppose instead we have (5.46) and (5.48). Now if $N_+ \geq (N^{(1)})^{\frac{2}{3}}$ and $L_+ \geq (N_+)^{\frac{1}{(40n)^4}}$, then (5.46)

implies (3.31); if $N_+ \geq (N^{(1)})^{\frac{2}{3}}$ and $L_+ \leq (N_+)^{\frac{1}{(40n)^4}}$, then (5.48) implies $|\mathcal{X}| \lesssim (N^{(1)})^{-1.01}$, which implies (3.31); if $N_+ \leq (N^{(1)})^{\frac{2}{3}}$, then interpolating (5.46) and (5.48) implies

$$|\mathcal{X}| \lesssim (N^{(1)})^{-1+\gamma_0} (N_+)^{\gamma} \lesssim (N^{(1)})^{-1+\gamma_0+\frac{2\gamma}{3}},$$

which implies (3.31).

Note also that for general α , due to the factor $\frac{\max((N^{(2)})^2, |\alpha|)}{(N^{(2)})^2}$ on the right-hand side of (4.28), the above argument gives the bound

$$(5.49) \quad |\mathcal{X}| \lesssim (N^{(1)})^{-1+\frac{4\gamma}{5}} \max\left(1, \frac{|\alpha|^{\frac{1}{2}}}{N^{(2)}}\right).$$

Step 3: The over-pairings. We will only consider $\mathcal{X}$; $\mathcal{Y}$ is similar and easier since there cannot be any pairing between $\{k, k'\}$ and any k_j due to $N^{(1)} \lesssim N_0^{1-\delta}$ and the restrictions $|k|, |k'| \geq \frac{N_0}{4}$ in (3.33). Now due to simplicity, any pairing in $\mathcal{X}$ must be an over-pairing; by collecting all these pairings we can find a partition $(A_1, \dots, A_p, B)$ of $\{1, \dots, n\}$ such that $|A_i| \geq 3$ and k_j takes a single value for j in each A_i , such that this value is different for different $1 \leq i \leq p$, and there is no over-pairing among $\{k_j : j \in B\}$. Then we can check that either there is no pairing among $\{k, k_j : j \in B\}$, or there is a unique over-pairing $k = k_{j_1} = k_{j_2}$ with $j_1, j_2 \in B$ and $(\iota_{j_1}, \iota_{j_2}) \neq (-, -)$. In the latter case denote $\{j_1, j_2\} = A_0$ and replace B by $B \setminus A_0$, so that there is no pairing among $\{k, k_j : j \in B\}$. Below we will focus on the first case, and leave the necessary changes caused by A_0 to the end.

Now $\mathcal{X}$ is reduced to

$$(5.50) \quad \mathcal{X} = \sum_{l_1, \dots, l_p} \int \prod_{i=1}^p \prod_{j \in A_i} (v_{l_i}^{(j)}(\lambda_j))^{\pm} d\lambda_j \cdot \mathcal{X}',$$

where l_i is the common value of k_j for $j \in A_i$ (so that $|l_i| \lesssim N^{(2)}$), and where $\mathcal{X}'$ is an expression of the same form as $\mathcal{X}$, but only involves the variables (k, k_j) and (λ, λ_j) for $j \in B$, with d being a fixed linear combination of l_i , and α being a fixed linear combination of $|l_i|^2$. This gives

$$|\mathcal{X}| \lesssim \sum_{l_1, \dots, l_p} \prod_{i=1}^p M_{l_i}^{(i)} \cdot \sup_{l_1, \dots, l_p} \left| \int \prod_{i=1}^p \frac{1}{M_{l_i}^{(i)}} \prod_{j \in A_i} (v_{l_i}^{(j)}(\lambda_j))^{\pm} d\lambda_j \cdot \mathcal{X}' \right|,$$

$$M_{l_i}^{(i)} := \prod_{j \in A_i} \|\langle \lambda_j \rangle^{b_2} v_{l_i}^{(j)}(\lambda_j)\|_{L_{\lambda_j}^2},$$

where we recall that $b_2 = b - \delta^6$. When each l_i is fixed, by [Remark 5.3](#), the expression

$$\int \prod_{i=1}^p \frac{1}{M_{l_i}^{(i)}} \prod_{j \in A_i} (v_{l_i}^{(j)}(\lambda_j))^{\pm} d\lambda_j \cdot \mathcal{X}'$$

can be estimated in the same way as $\mathcal{X}'$ (replacing b by b_2 will not change the proof), which is done in Step 2 above. We then only need to bound

$$\sum_{l_1, \dots, l_p} \prod_{i=1}^p M_{l_i}^{(i)} = \prod_{i=1}^p \sum_{l_i} M_{l_i}^{(i)},$$

which we establish in the following claim.

CLAIM 5.4. *Let $K_i = \max(N_j : j \in A_i)$ and $K'_i = \max^{(2)}(N_j : j \in A_i)$. Then $\tau^{-1}N_*$ -certainly we have that*

$$(5.51) \quad \sum_{l_i} M_{l_i}^{(i)} \lesssim \begin{cases} K_i^{-1+\gamma} (K'_i)^{-\frac{1}{3}}, & K_i \sim N_j, j \in \mathcal{D}, \\ K_i^{-1+\theta}, & K_i \sim N_j, j \in \mathcal{G} \cup \mathcal{C}. \end{cases}$$

Proof of Claim 5.4. Let $R_{l_i}^{(j)} = \|\langle \lambda_j \rangle^{b_2} v_{l_i}^{(j)}(\lambda_j)\|_{L_{\lambda_j}^2}$. Then we have $\|R^{(j)}\|_{\ell_{l_i}^\infty} \lesssim \|R^{(j)}\|_{\ell_{l_i}^2} \lesssim N_j^{-1+\gamma}$ if $j \in \mathcal{D}$, $\|R^{(j)}\|_{\ell_{l_i}^\infty} \lesssim N_j^{-1+\theta}$ and $\|R^{(j)}\|_{\ell_{l_i}^2} \lesssim N_j^\theta$ if $j \in \mathcal{G}$. If $j \in \mathcal{C}$, we will apply [Lemma 4.1](#) and again reduce to finitely many λ_j by restricting the size of λ_j and dividing into small intervals, and using the differentiability in λ_j of $h_{k_j k_j^*}^{(j)}(\lambda_j)$, which is assumed in the statement of [Proposition 3.4](#). In the same way as before, by removing a set of probability $\leq C_\theta e^{-(\tau^{-1}N_*)^\theta}$ and omitting any $\tau^{-\theta}(N_*)^{C_\theta-1}$ factors, we conclude that

$$\|R^{(j)}\|_{\ell_{l_i}^\infty} \lesssim \|R^{(j)}\|_{\ell_{l_i}^2} \lesssim N_j^{-1} \|\langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{\ell_{k_j, k_j^*}^2 L_{\lambda_j}^2} \lesssim N_j^{-\frac{1}{2}+\gamma_0} L_j^{-\frac{1}{2}},$$

as well as

$$\|R^{(j)}\|_{\ell_{l_i}^\infty} \lesssim N_j^{-1} L_j^2 \sup_{k_j, k_j^*} \|\langle \lambda_j \rangle^b h_{k_j k_j^*}^{(j)}(\lambda_j)\|_{L_{\lambda_j}^2} \lesssim N_j^{-1} L_j^2,$$

which also implies $\|R^{(j)}\|_{\ell_{l_i}^\infty} \lesssim N_j^{-0.55}$. Now let $K_i \sim N_j$ and $K'_i \sim N_s$. Then if $j \in \mathcal{D}$, (5.51) follows from applying Hölder and measuring $R^{(j)}$ and another factor other than $R^{(j)}$ or $R^{(s)}$ in $\ell_{l_i}^2$, and all other factors in $\ell_{l_i}^\infty$. If $j \in \mathcal{G} \cup \mathcal{C}$, then we may assume $|l_i| \sim K_i$ (otherwise (5.51) follows trivially from the third inequality in (3.26)), so the N_j for $j \in A_i$ must all be comparable. We may then assume $j \in \mathcal{G} \cup \mathcal{C}$ for each $j \in A_i$. Then (5.51) follows from applying Hölder and measuring two factors in $\ell_{l_i}^2$ and the rest in $\ell_{l_i}^\infty$, such that at least one $R^{(j)}$ with $j \in \mathcal{G}$ is measured in $\ell_{l_i}^\infty$ if there is any. This completes the proof. $\square$

The general case of [Proposition 3.4](#) then follows from the $\mathcal{X}'$ estimate, namely the no-pairing case in Step 2, combined with [Claim 5.4](#). More precisely, suppose $N^{(1)} = N_a$ with $a \in A_i$ for some i . Then if $a \in \mathcal{D}$, the bound [\(5.51\)](#) gives the power $(N^{(1)})^{-1+\gamma}$, while the power $(K'_i)^{-\frac{1}{3}}$ in [\(5.51\)](#), as well as the no-pairing case of the bounds [\(3.28\)](#) and [\(3.29\)](#), give the power $(N^{(2)})^{-\frac{1}{4}}$. If $a \in \mathcal{G} \cup \mathcal{C}$, then the power $(N^{(1)})^{-1+\theta}$ from [\(5.51\)](#) is already enough. If $N^{(1)} = N_a$ with $a \in B$, then we simply apply the no-pairing case and use [\(5.51\)](#) to gain decay in $N^{(2)}$ when $N^{(2)} = N_j$ and $j \notin B$. The only nontrivial case is [\(3.31\)](#), where there is no need to gain decay in $N^{(2)}$, but we have an extra factor $\lesssim \max(1, |\alpha|^{\frac{1}{2}})$ from [\(5.49\)](#), where α is a linear combination of $|l_i|^2$. By [Claim 5.4](#) we have

$$\langle \alpha \rangle^{\frac{1}{2}} \lesssim \max_{1 \leq i \leq p} \min_{j \in A_i} N_j, \quad \prod_{i=1}^p \sum_{l_i} M_{l_i}^{(i)} \lesssim (N_*)^\theta \langle \alpha \rangle^{-\frac{1}{2}},$$

which cancels this extra factor and proves [\(3.31\)](#).

Finally we consider the case with A_0 , say $k = k_{j_1} = k_{j_2}$ and $N_{j_1} \geq N_{j_2}$. Here we can check that $\mathcal{X}$ still has the form [\(5.50\)](#), except that in $\mathcal{X}'$ the input $v_k(\lambda)$ is replaced by (essentially)¹³

$$\tilde{v}_k(\lambda) = \int_{\pm \lambda_0 \pm \lambda_1 \pm \lambda_2 = \lambda} v_k(\lambda_0)^\pm \cdot v_k^{(j_1)}(\lambda_1)^\pm \cdot v_k^{(j_2)}(\lambda_2)^\pm d\lambda_1 d\lambda_2,$$

which due to the same proof as in [Claim 5.4](#), satisfies

$$\begin{aligned} & \| \langle \lambda \rangle^{b_2} \tilde{v}_k(\lambda) \|_{\ell_k^2 L_\lambda^2} \\ & \lesssim \| \langle \lambda_0 \rangle^{b_2} v_k(\lambda_0) \|_{\ell_k^2 L_{\lambda_0}^2} \| \langle \lambda_1 \rangle^{b_2} v_k^{(j_1)}(\lambda_1) \|_{\ell_k^\infty L_{\lambda_1}^2} \| \langle \lambda_2 \rangle^{b_2} v_k^{(j_2)}(\lambda_2) \|_{\ell_k^\infty L_{\lambda_2}^2} \\ (5.52) \quad & \lesssim \begin{cases} N_{j_1}^{-1+\gamma} N_{j_2}^{-\frac{1}{3}}, & j_1 \in \mathcal{D}, \\ N_{j_1}^{-1+\theta}, & j_1 \in \mathcal{G} \cup \mathcal{C}. \end{cases} \end{aligned}$$

The rest of proof now goes exactly as above using the additional bound [\(5.52\)](#), which has exactly the same gain as in [Claim 5.4](#), and the set A_0 is treated together with the other sets A_i . This completes the proof of [Proposition 3.4](#).

5.3. Stability and convergence. Recall that v_N is the solution to [\(2.5\)](#). [Proposition 3.2](#) already implies the convergence of v_N on the short time interval $[-\tau, \tau]$. For the purpose of proving global well-posedness, we need some additional results, namely a commutator estimate, a stability estimate and

¹³To deal with the η factor in the expression of $\mathcal{X}$ (see [\(5.1\)](#)), we only need to assume $\eta(\lambda, \mu) = \eta_1(\lambda)\eta_2(\mu)$ has factorized form. In general it is easy to write η as a linear combination of functions of such a form with summable coefficients, by invoking the explicit formula for η , which can be deduced from the calculations in the proof of [Lemma 2.6](#).

two convergence results laid out in [Propositions 5.5](#) and [5.6](#) below. For the notation involved in the proof, see [Sections 3.1](#) and [3.2](#).

PROPOSITION 5.5. *Recall the relevant constants defined in [\(1.35\)](#), and that $\tau \ll 1$, $J = [-\tau, \tau]$. The following two statements hold τ^{-1} -certainly:*

(1) (*Commutator estimate*). *For any $N \leq N'$, we have*

$$(5.53) \quad \|v_N - \Pi_N v_{N'}\|_{X^b(J)} \leq N^{-1+\gamma}.$$

(2) (*Stability*). *Let $w = \Pi_N w$ be a solution to [\(2.5\)](#) on J , but with data $w(t_0)$ assigned at some $t_0 \in J$ such that $\|w(t_0) - v_N(t_0)\|_{L^2} \leq AN^{-1+\gamma}(\log N)^\alpha$, where $\alpha \geq 0$ is an integer. Then we have*

$$(5.54) \quad \|w - v_N\|_{X^b(J)} \leq BN^{-1+\gamma}(\log N)^{\alpha+1},$$

where B depends only on A and α .

Proof. (1) It suffices to prove $\|\Pi_N v_{N'}^\dagger - v_N^\dagger\|_{X^b} \leq N^{-1+\gamma}$. We write

$$\Pi_N v_{N'}^\dagger - v_N^\dagger = \sum_{N < M \leq N'} \Pi_N y_M^\dagger = \sum_{N < M \leq N'} (\Pi_N \psi_{M, L_0(M)}^\dagger + \Pi_N z_M^\dagger),$$

where $L_0(M)$ is the largest L satisfying $(M, L) \in \mathcal{K}$. The bound for $\Pi_N z_M^\dagger$ follows from [Proposition 3.3](#), so it suffices to bound $\Pi_N \psi_{M, L_0}^\dagger$, where $L_0 = L_0(M)$. Let $\psi = \psi_{M, L_0}^\dagger$. Then we have

$$(5.55) \quad \begin{aligned} \psi(t) &= \chi(t) e^{it\Delta} (\Delta_M f(\omega)) - i\chi_\tau(t) \\ &\times \sum_{l=0}^r (l+1) c_{rl} (m_M^*)^{r-l} \cdot \mathcal{I} \Pi_M \mathcal{N}_{2l+1}(\psi, v_{L_0}^\dagger, \dots, v_{L_0}^\dagger). \end{aligned}$$

Since $N \leq \frac{M}{2}$, $\Pi_N \psi$ solves the equation

$$(5.56) \quad \begin{aligned} \Pi_N \psi(t) &= -i\chi_\tau(t) \sum_{l=0}^r (l+1) c_{rl} (m_M^*)^{r-l} \cdot \mathcal{I} \Pi_N \mathcal{N}_{2l+1}(\Pi_N \psi, v_{L_0}^\dagger, \dots, v_{L_0}^\dagger) \\ &- i\chi_\tau(t) \sum_{l=0}^r (l+1) c_{rl} (m_M^*)^{r-l} \cdot \mathcal{I} \Pi_N \mathcal{N}_{2l+1}(\Pi_N^\perp \psi, v_{L_0}^\dagger, \dots, v_{L_0}^\dagger). \end{aligned}$$

Now τ^{-1} -certainly we may assume [\(3.21\)](#) and the variant of [\(3.31\)](#) described in [Remark 3.5](#). Note that [\(3.21\)](#) allows us to control the first line of [\(5.56\)](#); the second line of [\(5.56\)](#) is controlled by using [\(2.31\)](#) and the variant of [\(3.31\)](#). In the end we get that

$$\|\Pi_N \psi\|_{X^b} \lesssim \tau^\theta \|\Pi_N \psi\|_{X^b} + \tau^\theta M^{-1+\frac{4\gamma}{5}},$$

which proves [\(5.53\)](#).

(2) If $N \leq O_{A,\alpha}(1)$, there is nothing to prove, so we may assume N is large depending on (A, α) . Let $\sigma = \mathcal{A}|w|^2 - \mathcal{A}|v_N|^2$ (this is conserved). Then we have

$$|\sigma| \lesssim (\|v_N(t_0)\|_{L^2} + \|w(t_0)\|_{L^2})\|w(t_0) - v_N(t_0)\|_{L^2} \lesssim \tau^{-\theta} AN^{-1+\gamma}(\log N)^{\alpha+1}.$$

Note the log loss due to the fact that $\|v_N\|_{L^2}^2 \lesssim \tau^{-\theta} \log N$. Recall that v_N and w satisfy the equations

$$(5.57) \quad \begin{cases} (i\partial_t + \Delta)v_N = \sum_{l=0}^r c_{rl}(m_N^*)^{r-l} \Pi_N \mathcal{N}_{2l+1}(v_N, \dots, v_N), \\ (i\partial_t + \Delta)w = \sum_{l=0}^r c_{rl}(m_N^* + \sigma)^{r-l} \Pi_N \mathcal{N}_{2l+1}(w, \dots, w) \end{cases}$$

on J , so $z = w - v_N$ satisfies the equation

$$(5.58) \quad \begin{aligned} (i\partial_t + \Delta)z &= \sum_{l=0}^r c_{rl}[(m_N^* + \sigma)^{r-l} - (m_N^*)^{r-l}] \Pi_N \mathcal{N}_{2l+1}(v_N, \dots, v_N) \\ &+ \sum_{l=0}^r c_{rl}(m_N^* + \sigma)^{r-l} \Pi_N [\mathcal{N}_{2l+1}(z + v_N, \dots, z + v_N) - \mathcal{N}_{2l+1}(v_N, \dots, v_N)] \end{aligned}$$

on J , and $z_0 = z(t_0)$ satisfies $\|z_0\|_{L^2} \leq AN^{-1+\gamma}(\log N)^\alpha$. In order to bound $\|z\|_{X^b(J)}$, it will suffice to prove that given z_0 and σ , the mapping

$$(5.59) \quad \begin{aligned} z^\dagger &\mapsto \chi(t - t_0)e^{i(t-t_0)\Delta} z_0 \\ &- i\chi_{2\tau}(t - t_0) \sum_{l=0}^r c_{rl}[(m_N^* + \sigma)^{r-l} - (m_N^*)^{r-l}] \mathcal{I}_{t_0} \Pi_N \mathcal{N}_{2l+1}(v_N^\dagger, \dots, v_N^\dagger) \\ &- i\chi_{2\tau}(t - t_0) \sum_{l=0}^r c_{rl}(m_N^* + \sigma)^{r-l} \mathcal{I}_{t_0} \Pi_N [\mathcal{N}_{2l+1}(z^\dagger + v_N^\dagger, \dots, z^\dagger + v_N^\dagger) \\ &- \mathcal{N}_{2l+1}(v_N^\dagger, \dots, v_N^\dagger)] \end{aligned}$$

is a contraction mapping from the set $\{z^\dagger : \|z^\dagger\|_{X^b} \leq AN^{-1+\gamma}(\log N)^{\alpha+1}\}$ to itself, where

$$\mathcal{I}_{t_0} F(t) = \mathcal{I}F(t) - \chi(t)e^{i(t-t_0)\Delta} \mathcal{I}F(t_0),$$

$$\mathcal{I}_{t_0} F(t) = \chi(t) \int_{t_0}^t \chi(t')e^{i(t-t')\Delta} F(t') dt'.$$

To this end we will decompose $v_N^\dagger = \sum_{N' \leq N} y_{N'}^\dagger$ and (τ^{-1} -certainly) apply the estimates (3.21) and (3.28), in the same way as in the proof of (3.23). More precisely, we may use (3.21) to control the terms in (5.59) that contain only one factor $z^\dagger$ (where we use the τ^θ gain to ensure smallness), and use (3.28) to control the terms in (5.59) that contain at least two factors $z^\dagger$ (where we use

the gain of powers of N to ensure smallness, noticing that N is large enough compared to A). Note that in applying these estimates we need to replace $\mathcal{I}$ by $\mathcal{I}_{t_0}$ and $\chi_\tau(t)$ by $\chi_{2\tau}(t - t_0)$. This can be done because in [Section 3.3](#) all estimates for $\chi_\tau(t) \cdot \mathcal{I}[\dots]$ are deduced from (2.31) and the corresponding estimates for $\mathcal{I}[\dots]$; here by definition we have $\|\mathcal{I}_{t_0} F\|_{X^{\tilde{b}}} \lesssim \|\mathcal{I} F\|_{X^{\tilde{b}}}$ for $\tilde{b} \in \{b, b_1\}$, which allows us to replace $\mathcal{I}$ by $\mathcal{I}_{t_0}$, and we have that $\mathcal{I}_{t_0} F(t_0) = 0$ so (2.31) is still applicable with $\mathcal{I}$ replaced by $\mathcal{I}_{t_0}$ and $\chi_\tau(t)$ replaced by $\chi_{2\tau}(t - t_0)$. The rest of the proof will be the same. $\square$

PROPOSITION 5.6 (Convergence). *Recall the relevant constants defined in (1.35), the ε fixed as in [Remark 1.4](#), and that $\tau \ll 1$ and $J = [-\tau, \tau]$. Then the following hold τ^{-1} -certainly:*

(1) *For any $N \leq N'$, we have*

$$(5.60) \quad \|v_N - v_{N'}\|_{X^{-\theta, b_2}(J)} \leq \tau^{-\theta} N^{-\frac{\theta}{2}},$$

$$(5.61) \quad \|(v_N - e^{it\Delta} v_N(0)) - (v_{N'} - e^{it\Delta} v_{N'}(0))\|_{X^{\frac{1}{2}-\gamma_0-\theta, b_2}(J)} \leq N^{-\frac{\theta}{2}}.$$

Note that the $X^{s, b}(J)$ bounds also imply the corresponding $C_t^0 H_x^s(J)$ bounds.

(2) *Let $\mathcal{N}_n(v)$ be a polynomial that is also viewed as a multilinear form $\mathcal{N}_n(v^{(1)}, \dots, v^{(n)})$, as in (2.16), but it is only assumed to be input-simple (instead of simple). Then for any $N \leq N'$, the distance in $C_t^0 H_x^\varepsilon(J)$ between any two of the following expressions,*

$$(5.62) \quad \mathcal{N}_n(v), \Pi_N \mathcal{N}_n(v), \Pi_{N'} \mathcal{N}_n(v) : v \in \{v_N, v_{N'}, \Pi_N v_{N'}\},$$

is bounded by $\tau^{-\theta} N^{-\gamma}$. The same conclusion holds if $\mathcal{N}_n$ is replaced by W_N^n or $W_{N'}^n$, or if v is perturbed by any w_N satisfying $\|w_N\|_{X^b(J)} \leq AN^{-1+\gamma}(\log N)^\alpha$. In the latter case the bound will be $O_{A, \tau, \alpha}(1)N^{-\gamma}$.

Proof. (1) We only need to prove (5.61). By taking a summation we may assume $N' = 2N$, and it suffices to prove that $\|y_{N'} - e^{it\Delta}(\Delta_{N'} f(\omega))\|_{X^{b_2}(J)} \leq (N')^{-\frac{1}{2}+\gamma_0+\frac{\theta}{2}}$. Now an extension of this function is given by

$$y^\dagger = \sum_L \zeta_{N', L}^\dagger + z_{N'}^\dagger$$

(see [Sections 3.1](#) and [3.2](#)), where $\zeta_{N', L}^\dagger$ is defined from $h^{N', L, \dagger}$ by (3.8) and (3.10), and $h^{N', L, \dagger}$ and $z_{N'}^\dagger$ satisfy (3.17) and (3.23). This controls the second term; to bound the first term, we use the $\mathcal{B}_{\leq L}^+$ measurability of $h^{N', L, \dagger}$ and [Lemma 4.1](#), and perform the same reduction step as in the proof of [Claim 5.4](#) using differentiability in λ of $\tilde{h}_{kk^*}(\lambda)$ with $h = h^{N', L, \dagger}$, to bound

$\tau^{-1}N'$ -certainly that

$$(5.63) \quad \sum_L \|\zeta_{N,L}^\dagger\|_{X^{b_2}} \lesssim \sum_L (N')^{-1} \|h^{N',L,\dagger}\|_{Z^b} \lesssim (N')^{-\frac{1}{2}+\gamma_0+\frac{\theta}{2}}.$$

The right-hand side of (5.63) may involve a $\tau^{-\theta}$ factor, but this loss can always be recovered by attaching χ_τ since $y^\dagger(0) = 0$.

(2) The bounds for W^n follows from the bounds for $\mathcal{N}_n$ and the formulas (2.11) and (2.12), noticing that $:|v|^{2r}v:$ and $:|v|^{2r}$ are input-simple. As for $\mathcal{N}_n$, by decomposing

$$v_N^\dagger = \sum_{N' \leq N} y_{N'}^\dagger, \quad y_{N'}^\dagger = \chi(t) e^{it\Delta} (\Delta_{N'} f(\omega)) + \sum_L \zeta_{N',L}^\dagger + z_{N'}^\dagger,$$

it suffices to bound $\mathcal{N}_n(v^{(1)}, \dots, v^{(n)})$, $\tau^{-1}N^{(1)}$ -certainly, by $\tau^{-\theta}(N^{(1)})^{-\gamma}$ in the space $C_t^0 H_x^{-\varepsilon}(J)$ for $v^{(j)}$ as in the assumptions of Proposition 3.4. The proof is a much easier variant of the arguments in Section 5.1.1, so we will only sketch the most important points.

First, since the ∂_t derivative of all the $v^{(j)}$'s are bounded by $(N^{(1)})^C$ (by restricting the size of λ_j variables as we did in Section 5.1.1), by dividing J into $(N^{(1)})^{\delta-1}$ intervals we may reduce to $(N^{(1)})^{C\delta-1}$ exceptional sets and thus fix a time $t \in J$. This gets rid of all the λ_j variables (so we are considering $v_{k_j}^{(j)}$ and $h_{k_j k_j^*}^{(j)}$), and by a simple $H_t^{\frac{1}{2}+} \hookrightarrow C_t^0$ argument, the estimates (3.26) and (3.27) remain true with the obvious changes. Now by repeating the arguments in Section 5.1.1 (in a simplified situation without λ_j integrations) and Section 5.2 (which deals with over-pairings) we get that

$$\|\Delta_{N_0} \mathcal{N}_n(v^{(1)}, \dots, v^{(n)})\|_{L^2}^2 \lesssim \tau^{-\theta} (N^{(1)})^{C\kappa-1} (N^{(1)})^{C\gamma} \cdot (\#S) \prod_{j=1}^n N_j^{-2} \prod_{i=1}^p \frac{N_{2i-1}}{R_i},$$

where $N_{2i-1} \sim N_{2i} \gtrsim R_i$ and

$$S = \left\{ (k, k_1, \dots, k_n) \in (\mathbb{Z}^2)^{n+1} : \sum_{j=1}^n \iota_j k_j = k, \quad |k| \lesssim N_0, \right. \\ \left. |k_j| \lesssim N_j \ (1 \leq j \leq n), \quad |k_{2i-1} - k_{2i}| \lesssim R_i (N^{(1)})^{C\kappa-1} \ (1 \leq i \leq p) \right\}.$$

A simple counting estimate yields

$$\begin{aligned} \|\Delta_{N_0} \mathcal{N}_n(v^{(1)}, \dots, v^{(n)})\|_{L^2}^2 &\lesssim \tau^{-\theta} (N^{(1)})^{C\gamma} \min(1, (N^{(1)})^{-1} N_0) \\ &\lesssim \tau^{-\theta} N_0^\varepsilon (N^{(1)})^{-\gamma} \end{aligned}$$

as by our choice $\gamma \ll \varepsilon$, which concludes the proof. With the w_N perturbations the proof works the same way, except that the constants may depend also on A and α . $\square$

Before ending this section, we would like to shift the point of view from the probability space $(\Omega, \mathcal{B}, \mathbb{P})$ to the spaces $\mathcal{V}$ and $\mathcal{V}_N$. Given $0 < \tau \ll 1$, all of the above proofs allow us to identify a Borel set E_τ of $\mathcal{V}$ with $\rho(E_\tau) \geq 1 - C_\theta e^{-\tau^{-\theta}}$, such that when $u_{\text{in}} = v_{\text{in}} \in E_\tau$, all the results in [Sections 3](#) and [5](#), including [Propositions 3.2](#), [3.3](#) and [5.5](#), are true.

In reality we will be using finite dimensional truncations of E_τ , namely $E_\tau^{\bar{N}} = \Pi_{\bar{N}} E_\tau$. Clearly when $\Pi_{\bar{N}} u_{\text{in}} \in E_\tau^{\bar{N}}$, all the quantitative estimates proved before will remain true if all the frequencies N, N', L , etc., are $\leq \bar{N}$. Moreover, we know that $\rho_{\bar{N}}(E_\tau^{\bar{N}}) \geq \rho(E_\tau) \geq 1 - C_\theta e^{-\tau^{-\theta}}$; since the Radon-Nikodym derivative $\frac{d\mu_{\bar{N}}^\circ}{d\rho_{\bar{N}}}$ is uniformly bounded in $L^2(d\rho_{\bar{N}})$, we have that

$$(5.64) \quad \mu_{\bar{N}}^\circ(E_\tau^{\bar{N}}) \geq 1 - C \sqrt{\rho_{\bar{N}}(\mathcal{V}_{\bar{N}} \setminus E_\tau^{\bar{N}})} \geq 1 - C_\theta e^{-\tau^{-\theta}}.$$

Finally, due to the gauge symmetry of [\(1.7\)](#) and [\(2.5\)](#), we may assume that E_τ (and $E_\tau^{\bar{N}}$) is rotation invariant, i.e., $e^{i\alpha} E_\tau = E_\tau$ for $\alpha \in \mathbb{R}$.

6. Global well-posedness and measure invariance

In this section we will prove [Theorem 1.3](#). Recall the sets $E_\tau^{\bar{N}}$ defined at the end of [Section 5.3](#). Denote the solution flow of [\(1.7\)](#) by Φ_t^N and the solution flow of [\(2.5\)](#) by Ψ_t^N , which are mappings from $\mathcal{V}_N$ to itself. Define successively the sets

$$(6.1) \quad F_{T,K}^{\bar{N}} = \bigcap_{|j| \leq K} (\Psi_{\frac{j}{K}}^{\bar{N}})^{-1} E_{\frac{T}{K}}^{\bar{N}},$$

$$(6.2) \quad G_{T,K,A,D}^{\bar{N},\alpha} = \left\{ v \in \mathcal{V}_{\bar{N}} : \exists t \in [-D, D] \text{ s.t. } \Psi_t^{\bar{N}} v = v' + v'', \right. \\ \left. v' \in F_{T,K}^{\bar{N}}, \ \|v''\|_{L^2} \leq A \bar{N}^{-1+\gamma} (\log \bar{N})^\alpha \right\},$$

$$(6.3) \quad \Sigma = \bigcup_{D \geq 1} \bigcap_{T \geq 2^{10} D} \bigcup_{K \gg T; A, \alpha \geq 1} \limsup_{\bar{N} \rightarrow \infty} \Pi_{\bar{N}}^{-1} G_{T,K,A,D}^{\bar{N},\alpha}.$$

Here $\Pi_{\bar{N}}^{-1} G = G \times \mathcal{V}_{\bar{N}}^\perp$ is the cylindrical set. We understand that T, K, A, D all belong to some given countable set (say powers of two), and α is an integer. All these sets are Borel, since in [\(6.2\)](#) we may replace the $\leq$ sign by the $<$ sign, and then restrict to rational t by continuity. We will start by proving global well-posedness and then measure invariance.

PROPOSITION 6.1. *The set Σ satisfies $\mu(\mathcal{V} \setminus \Sigma) = 0$, and $W^{2r+1}(u) \in H^{-\varepsilon}$ is well defined for $u \in \Sigma$. For any $u_{\text{in}} \in \Sigma$, the solutions $u_N(t) = \Phi_t^N \Pi_N u_{\text{in}}$ to (1.7) converge to some $u(t) = \Phi_t u_{\text{in}}$ in $C_t^0 H_x^{-\varepsilon}([-T, T])$ for any $T > 0$. This u is a distributional solution to (1.1), and $u(t) \in \Sigma$ for each t . The mappings $\Phi_t : \Sigma \rightarrow \Sigma$ satisfy $\Phi_0 = \text{Id}$ and $\Phi_{t+t'} = \Phi_t \Phi_{t'}$.*

Proof. We first prove $\mu(\mathcal{V} \setminus \Sigma) = 0$. By definition we have

$$(6.4) \quad \Sigma \supset \bigcap_{T \geq 2^{10}} \bigcup_{K \gg T} \limsup_{\bar{N} \rightarrow \infty} \Pi_{\bar{N}}^{-1} F_{T,K}^{\bar{N}},$$

so it suffices to prove for any fixed $T \geq 2^{10}$ that

$$\sup_{K \gg T} \mu \left(\limsup_{\bar{N} \rightarrow \infty} \Pi_{\bar{N}}^{-1} F_{T,K}^{\bar{N}} \right) = 1.$$

Now by Fatou's lemma and the fact that the total variation of $\mu - \mu_{\bar{N}}$ converges to 0, we have

$$\mu \left(\limsup_{\bar{N} \rightarrow \infty} \Pi_{\bar{N}}^{-1} F_{T,K}^{\bar{N}} \right) \geq \limsup_{\bar{N} \rightarrow \infty} \mu_{\bar{N}}(\Pi_{\bar{N}}^{-1} F_{T,K}^{\bar{N}}) = \limsup_{\bar{N} \rightarrow \infty} \mu_{\bar{N}}^{\circ}(F_{T,K}^{\bar{N}}).$$

By invariance of $d\mu_{\bar{N}}^{\circ}$ under the flow $\Psi_t^{\bar{N}}$ (Proposition 2.4) we know that

$$\mu_{\bar{N}}^{\circ}(F_{T,K}^{\bar{N}}) \geq 1 - (2K + 1) \mu_{\bar{N}}^{\circ}(\mathcal{V}_{\bar{N}} \setminus E_{\frac{T}{K}}^{\bar{N}}) \geq 1 - C_{\theta} K e^{-(KT^{-1})^{\theta}}$$

uniformly in $\bar{N}$, and the right-hand side converges to 1 as $K \rightarrow \infty$, so $\mu(\mathcal{V} \setminus \Sigma) = 0$.

Now suppose $u_{\text{in}} \in \Sigma$. By definition we may choose some D . Then for any $T \geq 2^{10}D$, we can find (K, A, α) such that $\Pi_{\bar{N}} u_{\text{in}} \in G_{T,K,A,D}^{\bar{N},\alpha}$ for infinitely many $\bar{N}$. We may fix this T (hence also (K, A, α)) and this $\bar{N}$, so that $\|\Psi_{t_0}^{\bar{N}} \Pi_{\bar{N}} u_{\text{in}} - v'\|_{L^2} \leq A \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha}$ for some $t_0 \in [-D, D]$ and $v' \in F_{T,K}^{\bar{N}}$. We proceed in three steps.

Step 1: Analyzing v' . We first prove that, for any $N \leq \bar{N}$ and $|j| \leq K$, there holds that

$$(6.5) \quad \|\Pi_N \Psi_{\frac{jT}{K}}^{\bar{N}} v' - \Psi_{\frac{jT}{K}}^N \Pi_N v'\|_{L^2} \leq B N^{-1+\gamma} (\log N)^{|j|}$$

for some B depending only on (T, K) . This is obviously true for $j = 0$; suppose this is true for j , since $\Psi_{\frac{jT}{K}}^{\bar{N}} v' \in E_{\frac{T}{K}}^{\bar{N}}$. By Proposition 5.5(1) we have

$$\|\Pi_N \Psi_{\frac{(j \pm 1)T}{K}}^{\bar{N}} v' - \Psi_{\frac{\pm T}{K}}^N \Pi_N \Psi_{\frac{jT}{K}}^{\bar{N}} v'\|_{L^2} \leq N^{-1+\gamma}$$

(note that as $K \gg T$ the local theory is applicable on intervals of length $\frac{T}{K}$), and

$$\|\Psi_{\frac{\pm T}{K}}^N \Pi_N \Psi_{\frac{jT}{K}}^{\bar{N}} v' - \Psi_{\frac{(j \pm 1)T}{K}}^N \Pi_N v'\|_{L^2} \leq B' N^{-1+\gamma} (\log N)^{|j|+1}$$

by Proposition 5.5(2) and (6.5), where B' depends only on B and (T, K) , so (6.5) holds also for $j \pm 1$, which concludes the inductive proof. By the same argument, we can show that (6.5) remains true with $\frac{jT}{K}$ replaced by any $t \in [-T, T]$ and $|j|$ replaced by K .

Similarly, since $\Psi_{\frac{jT}{K}}^{\bar{N}} v' \in E_{\frac{T}{K}}^{\bar{N}}$ for each $|j| \leq K$, by combining Propositions 5.5 and 5.6, we conclude that for any $N \leq N' \leq \bar{N}$,

$$(6.6) \quad \sup_{t \in [-T, T]} \|\Psi_t^N \Pi_N v' - \Psi_t^{N'} \Pi_{N'} v'\|_{H^{-\theta}} \leq O_{T, K}(1) N^{-\frac{\theta}{2}},$$

$$(6.7) \quad \sup_{t \in [-T, T]} \|W_N^n(\Psi_t^N \Pi_N v') - W_{N'}^n(\Psi_t^{N'} \Pi_{N'} v')\|_{H^{-\varepsilon}} \leq O_{T, K}(1) N^{-\gamma},$$

and the same is true if W_N^n and $W_{N'}^n$ in (6.7) is replaced by $\Pi_N W_N^n$ and $\Pi_{N'} W_{N'}^n$.

Step 2: Linking u_{in} to v' . Recall that

$$\|\Psi_{t_0}^{\bar{N}} \Pi_{\bar{N}} u_{\text{in}} - v'\|_{L^2} \leq A \bar{N}^{-1+\gamma} (\log \bar{N})^\alpha.$$

Since $|t_0| \leq D \ll T$, by iterating Proposition 5.5(2) we deduce that

$$\|\Pi_{\bar{N}} u_{\text{in}} - \Psi_{-t_0}^{\bar{N}} v'\|_{L^2} \leq A' \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+K},$$

where A' depends only on (T, K, A, α) . Writing $-t_0 = \frac{jT}{K} + t'$ with $|j| \leq 2^{-8}K$ and $|t'| \leq \frac{T}{K}$, we may apply Proposition 5.5(2) again and combine this with (6.5) and similar estimates to deduce for any $N \leq \bar{N}$ that

$$(6.8) \quad \sup_{t \in [-\frac{T}{2}, \frac{T}{2}]} \|\Psi_t^N \Pi_N u_{\text{in}} - \Psi_{t-t_0}^N \Pi_N v'\|_{L^2} \leq B N^{-1+\gamma} (\log N)^{\alpha+2K}$$

with B depending only on (T, K, A, α) . By (6.6), (6.7), (6.8) and Proposition 5.6, we conclude for all $N \leq N' \leq \bar{N}$ that

$$(6.9) \quad \sup_{t \in [-\frac{T}{2}, \frac{T}{2}]} \|\Psi_t^N \Pi_N u_{\text{in}} - \Psi_t^{N'} \Pi_{N'} u_{\text{in}}\|_{H^{-\theta}} \leq O_{T, K, A, \alpha}(1) N^{-\frac{\theta}{2}},$$

$$(6.10)$$

$$\sup_{t \in [-\frac{T}{2}, \frac{T}{2}]} \|W_N^n(\Psi_t^N \Pi_N u_{\text{in}}) - W_{N'}^n(\Psi_t^{N'} \Pi_{N'} u_{\text{in}})\|_{H^{-\varepsilon}} \leq O_{T, K, A, \alpha}(1) N^{-\gamma},$$

and the same is true for projections of W_N^n .

Step 3: Completing the proof. Now, for fixed (D, T, K, A, α) we know that there exists infinitely many $\bar{N}$ such that (6.9) and (6.10) are true for all $N \leq N' \leq \bar{N}$, so (6.9) and (6.10) are simply true for all $N \leq N'$. This implies the convergence of $\Psi_t^N \Pi_N u_{\text{in}}$ in $C_t^0 H_x^{-\varepsilon}([-\frac{T}{2}, \frac{T}{2}])$, and we will define $\Psi_t = \lim_{N \rightarrow \infty} \Psi_t^N \Pi_N$. Since by the definition of gauge transform we have

$$(6.11) \quad \Phi_t^N \Pi_N u_{\text{in}} = \Psi_t^N \Pi_N u_{\text{in}} \cdot e^{-iB_N(t)}, \quad B_N(t) = (r+1) \int_0^t \mathcal{A}[W_N^{2r}(\Psi_t^N \Pi_N u_{\text{in}})] dt',$$

(6.9) and (6.10) also imply the convergence of $\Phi_t^N \Pi_N u_{\text{in}}$ in $C_t^0 H_x^{-\varepsilon}([-\frac{T}{2}, \frac{T}{2}])$, as well as the convergence of $\Pi_N W_N^{2r+1}(\Phi_t^N \Pi_N u_{\text{in}})$ in the same space. As $u_N = \Phi_t^N \Pi_N u_{\text{in}}$ solves equation (1.7) with the right-hand side being $\Pi_N W_N^{2r+1}(u_N)$, we know that the limit $u = \lim_{N \rightarrow \infty} u_N$ solves (1.1) in the distributional sense. Let $u(t) = \Phi_t u$. The group properties of Φ_t follow from the group properties of Ψ_t^N and limiting arguments similar to the above.

Finally we prove that $\Phi_{t_1} u_{\text{in}} \in \Sigma$ for $u_{\text{in}} \in \Sigma$ and any t_1 . Let D be associated with the assumption $u_{\text{in}} \in \Sigma$, and fix $D_1 \gg D + |t_1|$. For any $T \geq 2^{10} D_1$, there exists (K, A, α) such that $\Pi_{\bar{N}} u_{\text{in}} \in G_{T,K,A,D}^{\bar{N},\alpha}$ for infinitely many $\bar{N}$. It suffices to show that for such $\bar{N}$, we must have $\Pi_{\bar{N}} \Phi_{t_1} u_{\text{in}} \in G_{T,K,B,D_1}^{\bar{N},\alpha+3K}$ with B depending only on (T, K, A, α) . Since $\Pi_{\bar{N}} \Phi_{t_1} u_{\text{in}}$ and $\Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}}$ only differ by a rotation and the sets we constructed are rotation invariant, we only need to prove the same thing for $\Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}}$.

Now, on the one hand we know for some $|t_0| \leq D$ that $\|\Pi_{\bar{N}} u_{\text{in}} - \Psi_{-t_0}^{\bar{N}} v'\|_{L^2} \leq A' \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+K}$ for some A' depending only on (T, K, A, α) (see Step 2) and similarly $\|\Psi_{t_1}^{\bar{N}} \Pi_{\bar{N}} u_{\text{in}} - \Psi_{t_1-t_0}^{\bar{N}} v'\|_{L^2} \leq A' \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+2K}$. On the other hand, by taking limits in (6.5) and (6.8) we also get $\|\Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}} - \Psi_{t_1}^{\bar{N}} \Pi_{\bar{N}} u_{\text{in}}\|_{L^2} \leq A' \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+2K}$, and hence

$$\|\Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}} - \Psi_{t_1-t_0}^{\bar{N}} v'\|_{L^2} \leq A' \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+2K}.$$

Applying Proposition 5.5(2) again we get that $\|\Psi_{t_0-t_1}^{\bar{N}} \Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}} - v'\|_{L^2} \leq B \bar{N}^{-1+\gamma} (\log \bar{N})^{\alpha+3K}$ with $|t_0 - t_1| \leq D_1$ and $B \leq B(T, K, A, \alpha)$, thus by definition $\Pi_{\bar{N}} \Psi_{t_1} u_{\text{in}} \in G_{T,K,B,D_1}^{\bar{N},\alpha+3K}$. This completes the proof. $\square$

PROPOSITION 6.2. *For any Borel subset $E \subset \Sigma$ and any $t_0 \in \mathbb{R}$, we have $\mu(E) = \mu(\Phi_{t_0} E)$.*

Proof. The map Φ_t is a limit of continuous mappings, so it is Borel measurable. By taking limits, we may assume the set E is compact in $H^{-\varepsilon}$ topology. We may also assume that $|t_0| \leq 1$, and that for some fixed (T, K, A, α, D) with $K \gg T \geq 2^{10} D$, we have $E \subset \limsup_{\bar{N} \rightarrow \infty} G_{T,K,A,D}^{\bar{N},\alpha}$. By the proof of Proposition 6.1 we can deduce that for $u \in E$ and $|t| \leq 2$,

$$(6.12) \quad \|\Psi_t^N \Pi_N u - \Pi_N \Psi_t u\|_{L^2} \lesssim N^{-1+\gamma} (\log N)^{\alpha+3K}$$

with constants depending on (T, K, A, α) (same below). Moreover, concerning the phase $B_N(t)$ involved in the gauge transform, namely

$$B_N(t) = (r+1) \int_0^t \mathcal{A}[W_N^{2r}(\Psi_t^N \Pi_N u)] dt',$$

we can show that as $N \rightarrow \infty$, $B_N(t)$ converges to its limit $B(t)$ at a rate $\|B_N(t) - B(t)\|_{C_t^0([-2,2])} \lesssim N^{-1+\gamma} (\log N)^{\alpha+4K}$. In fact we may first reduce to

short time intervals where local theory is applicable. Then notice that

$$\int_0^t \mathcal{A}[W_N^{2r}(\Psi_t^N \Pi_N u)] dt' = \mathcal{A}\mathcal{I}[W_N^{2r}(\Psi_t^N \Pi_N u)]$$

for $|t| \ll 1$, and apply [Proposition 3.4](#), more precisely (3.29), with the observation that the mean $\mathcal{A}$ restricts the two highest input frequencies in any multilinear expression $\mathcal{N}_n$ occurring in W_N^{2r} to be comparable, i.e., $N^{(1)} \sim N^{(2)}$. We omit the details.

With the explicit convergence rate of $B_N(t)$, we see that (6.12) holds with Ψ_t^N and Ψ_t replaced by Φ_t^N and Φ_t , and with $3K$ replaced by $4K$. For $|t| \leq 1$, this gives that

$$\begin{aligned} \Pi_N \Phi_t E &\subset \Phi_t^N \Pi_N E + \mathfrak{B}_{L^2}(A_1 N^{-1+\gamma} (\log N)^{\alpha+4K}) \\ &\subset \Phi_t^N (\Pi_N E + \mathfrak{B}_{L^2}(A_2 N^{-1+\gamma} (\log N)^{\alpha+5K})), \end{aligned}$$

where $A_{1,2}$ are constants depending only on (T, K, A, α) , and $\mathfrak{B}_{L^2}(R)$ is the ball of radius R in L^2 centered at the origin; note that the second subset relation follows from long-time stability, which is also a consequence of the proof of [Proposition 6.1](#). By invariance of $d\mu_N^\circ$ under Φ_t^N we have that

$$\begin{aligned} (6.13) \quad \mu_N(\Phi_{t_0} E) &\leq \mu_N^\circ(\Pi_N \Phi_{t_0} E) \leq \mu_N^\circ \Phi_{t_0}^N (\Pi_N E + \mathfrak{B}_{L^2}(A_2 N^{-1+\gamma} (\log N)^{\alpha+5K})) \\ &= \mu_N^\circ(\Pi_N E + \mathfrak{B}_{L^2}(A_2 N^{-1+\gamma} (\log N)^{\alpha+5K})). \end{aligned}$$

It then suffices to prove that

$$(6.14) \quad \limsup_{N \rightarrow \infty} \Pi_N^{-1} (\Pi_N E + \mathfrak{B}_{L^2}(A_2 N^{-1+\gamma} (\log N)^{\alpha+5K})) \subset E,$$

which would imply $\mu(\Phi_{t_0} E) \leq \mu(E)$, and conclude the proof by time reversibility.¹⁴ To prove (6.14), suppose u is such that $\|\Pi_N(u - u_N)\|_{L^2} \leq A_2 N^{-1+\gamma} (\log N)^{\alpha+5K}$ with $u_N \in E$ for infinitely many N . Then by compactness we may assume $u_N \rightarrow v \in E$ in $H^{-\varepsilon}$, so $u_N \rightarrow u$ coordinate-wise and $u_N \rightarrow v$ coordinate-wise, hence $u = v \in E$ and the proof is complete. $\square$

Remark 6.3. With a little more effort, we can show that the structural information for the short time solution, i.e., (3.12), is propagated for arbitrarily long time. This follows by iterating short time intervals using measure invariance, and applying (5.53) and (5.54) in [Proposition 5.5](#), in the same way as in the proof of global well-posedness ([Proposition 6.1](#)) above. We omit the details.

¹⁴By repeating the proof of [Proposition 6.1](#) we can show that $\Phi_t^N \Pi_N u \rightarrow \Phi_t u$ in $H^{-\varepsilon}$ uniformly for $u \in E$ and $|t| \leq 2$, so $\Phi_{t_0} E$ is also compact in $H^{-\varepsilon}$ and satisfies similar properties as E .

References

- [1] M. AIZENMAN, Geometric analysis of φ^4 fields and Ising models. I and II, *Comm. Math. Phys.* **86** no. 1 (1982), 1–48. MR 0678000. Zbl 0533.58034. <https://doi.org/10.1007/BF01205659>.
- [2] M. AIZENMAN and H. DUMINIL-COPIN, Marginal triviality of the scaling limits of critical 4D Ising and ϕ_4^4 models, *Ann. of Math.* (2) **194** no. 1 (2021), 163–235. MR 4276286. Zbl 1489.60150. <https://doi.org/10.4007/annals.2021.194.1.3>.
- [3] S. ALBEVERIO and A. B. CRUZEIRO, Global flows with invariant (Gibbs) measures for Euler and Navier-Stokes two-dimensional fluids, *Comm. Math. Phys.* **129** no. 3 (1990), 431–444. MR 1051499. Zbl 0702.76041. <https://doi.org/10.1007/BF02097100>.
- [4] I. BAILLEUL and F. BERNICOT, Heat semigroup and singular PDEs, *J. Funct. Anal.* **270** no. 9 (2016), 3344–3452, With an appendix by F. Bernicot and D. Frey. MR 3475460. Zbl 1416.35303. <https://doi.org/10.1016/j.jfa.2016.02.012>.
- [5] I. BAILLEUL and F. BERNICOT, High order paracontrolled calculus, *Forum Math. Sigma* **7** (2019), e44, 94. MR 4061967. Zbl 1473.60093. <https://doi.org/10.1017/fms.2019.44>.
- [6] I. BAILLEUL, F. BERNICOT, and D. FREY, Space-time paraproducts for paracontrolled calculus, 3D-PAM and multiplicative Burgers equations, *Ann. Sci. Éc. Norm. Supér.* (4) **51** no. 6 (2018), 1399–1456. MR 3940901. Zbl 1430.60053. <https://doi.org/10.24033/asens.2378>.
- [7] A. BÉNYI, T. OH, and O. POBOVICU, On the probabilistic Cauchy theory of the cubic nonlinear Schrödinger equation on $\mathbb{R}^d$, $d \geq 3$, *Trans. Amer. Math. Soc. Ser. B* **2** (2015), 1–50. MR 3350022. Zbl 1410.35195. <https://doi.org/10.1090/btran/6>.
- [8] A. BÉNYI, T. OH, and O. POBOVICU, Higher order expansions for the probabilistic local Cauchy theory of the cubic nonlinear Schrödinger equation on $\mathbb{R}^3$, *Trans. Amer. Math. Soc. Ser. B* **6** (2019), 114–160. MR 3919013. Zbl 1410.35195. <https://doi.org/10.1090/btran/29>.
- [9] V. I. BOGACHEV, *Gaussian Measures*, *Math. Surveys Monogr.* **62**, Amer. Math. Soc., Providence, RI, 1998. MR 1642391. Zbl 0913.60035. <https://doi.org/10.1090/surv/062>.
- [10] J. BOURGAIN, Periodic nonlinear Schrödinger equation and invariant measures, *Comm. Math. Phys.* **166** no. 1 (1994), 1–26. MR 1309539. Zbl 0822.35126. <https://doi.org/10.1007/BF02099299>.
- [11] J. BOURGAIN, Invariant measures for the 2D-defocusing nonlinear Schrödinger equation, *Comm. Math. Phys.* **176** no. 2 (1996), 421–445. MR 1374420. Zbl 0852.35131. <https://doi.org/10.1007/BF02099556>.
- [12] J. BOURGAIN, *Global Solutions of Nonlinear Schrödinger Equations*, *Amer. Math. Soc. Colloq. Publ.* **46**, Amer. Math. Soc., Providence, RI, 1999. MR 1691575. Zbl 0933.35178. <https://doi.org/10.1090/coll/046>.

- [13] J. BOURGAIN, Nonlinear Schrödinger equations, in *Hyperbolic Equations and Frequency Interactions* (Park City, UT, 1995), *IAS/Park City Math. Ser.* **5**, Amer. Math. Soc., Providence, RI, 1999, pp. 3–157. MR 1662829. Zbl 0952.35127. <https://doi.org/10.1090/coll/046>.
- [14] J. BOURGAIN and A. BULUT, Almost sure global well-posedness for the radial nonlinear Schrödinger equation on the unit ball II: the 3d case, *J. Eur. Math. Soc. (JEMS)* **16** no. 6 (2014), 1289–1325. MR 3226743. Zbl 1301.35145. <https://doi.org/10.4171/JEMS/461>.
- [15] B. BRINGMANN, Almost sure local well-posedness for a derivative nonlinear wave equation, *Int. Math. Res. Not. IMRN* no. 11 (2021), 8657–8697. MR 4266148. Zbl 1473.35361. <https://doi.org/10.1093/imrn/rnz385>.
- [16] B. BRINGMANN, Y. DENG, A. R. NAHMOD, and H. YUE, Invariant Gibbs measures for the three dimensional cubic nonlinear wave equation, *Invent. Math.* **236** no. 3 (2024), 1133–1411. MR 4743517. Zbl 07846295. <https://doi.org/10.1007/s00222-024-01254-4>.
- [17] D. C. BRYDGES and G. SLADE, Statistical mechanics of the 2-dimensional focusing nonlinear Schrödinger equation, *Comm. Math. Phys.* **182** no. 2 (1996), 485–504. MR 1447302. Zbl 0867.35090. <https://doi.org/10.1007/BF02517899>.
- [18] N. BURQ, L. THOMANN, and N. TZVETKOV, Remarks on the Gibbs measures for nonlinear dispersive equations, *Ann. Fac. Sci. Toulouse Math.* (6) **27** no. 3 (2018), 527–597. MR 3869074. Zbl 1405.35193. <https://doi.org/10.5802/afst.1578>.
- [19] N. BURQ and N. TZVETKOV, Random data Cauchy theory for supercritical wave equations. I. Local theory, *Invent. Math.* **173** no. 3 (2008), 449–475. MR 2425133. Zbl 1156.35062. <https://doi.org/10.1007/s00222-008-0124-z>.
- [20] R. CATELLIER and K. CHOUK, Paracontrolled distributions and the 3-dimensional stochastic quantization equation, *Ann. Probab.* **46** no. 5 (2018), 2621–2679. MR 3846835. Zbl 1433.60048. <https://doi.org/10.1214/17-AOP1235>.
- [21] A. CHANDRA and H. WEBER, Stochastic PDEs, regularity structures, and interacting particle systems, *Ann. Fac. Sci. Toulouse Math.* (6) **26** no. 4 (2017), 847–909. MR 3746645. Zbl 1421.81001. <https://doi.org/10.5802/afst.1555>.
- [22] J. COLLIANDER and T. OH, Almost sure well-posedness of the cubic nonlinear Schrödinger equation below $L^2(\mathbb{T})$, *Duke Math. J.* **161** no. 3 (2012), 367–414. MR 2881226. Zbl 1260.35199. <https://doi.org/10.1215/00127094-1507400>.
- [23] G. DA PRATO and A. DEBUSSCHE, Two-dimensional Navier-Stokes equations driven by a space-time white noise, *J. Funct. Anal.* **196** no. 1 (2002), 180–210. MR 1941997. Zbl 1013.60051. <https://doi.org/10.1006/jfan.2002.3919>.
- [24] G. DA PRATO and A. DEBUSSCHE, Strong solutions to the stochastic quantization equations, *Ann. Probab.* **31** no. 4 (2003), 1900–1916. MR 2016604. Zbl 1071.81070. <https://doi.org/10.1214/aop/1068646370>.
- [25] G. DA PRATO and L. TUBARO, Wick powers in stochastic PDEs: An introduction, in *Quantum and Stochastic Mathematical Physics, Springer Proc.*

- Math. Stat.* **377**, Springer, Cham, 2023, pp. 1–15. MR 4592556. Zbl 07819342. https://doi.org/10.1007/978-3-031-14031-0_1.
- [26] Y. DENG, Two-dimensional nonlinear Schrödinger equation with random radial data, *Anal. PDE* **5** no. 5 (2012), 913–960. MR 3022846. Zbl 1264.35212. <https://doi.org/10.2140/apde.2012.5.913>.
- [27] Y. DENG, Invariance of the Gibbs measure for the Benjamin-Ono equation, *J. Eur. Math. Soc. (JEMS)* **17** no. 5 (2015), 1107–1198. MR 3346690. Zbl 1379.37135. <https://doi.org/10.4171/JEMS/528>.
- [28] Y. DENG, A. R. NAHMOD, and H. YUE, Optimal local well-posedness for the periodic derivative nonlinear Schrödinger equation, *Comm. Math. Phys.* **384** no. 2 (2021), 1061–1107. MR 4259382. Zbl 1470.35326. <https://doi.org/10.1007/s00220-020-03898-8>.
- [29] Y. DENG, A. R. NAHMOD, and H. YUE, Random tensors, propagation of randomness, and nonlinear dispersive equations, *Invent. Math.* **228** no. 2 (2022), 539–686. MR 4411729. Zbl 1506.35208. <https://doi.org/10.1007/s00222-021-01084-8>.
- [30] Y. DENG, A. R. NAHMOD, and H. YUE, The probabilistic scaling paradigm, *Vietnam J. Math.* **52** no. 4 (2024), 1001Ð1015. MR 4772731. <https://doi.org/10.1007/s10013-023-00672-w>.
- [31] Y. DENG, N. TZVETKOV, and N. VISCIGLIA, Invariant measures and long time behaviour for the Benjamin-Ono equation III, *Comm. Math. Phys.* **339** no. 3 (2015), 815–857. MR 3385985. Zbl 1379.37136. <https://doi.org/10.1007/s00220-015-2431-8>.
- [32] B. DODSON, J. LÜHRMANN, and D. MENDELSON, Almost sure local well-posedness and scattering for the 4D cubic nonlinear Schrödinger equation, *Adv. Math.* **347** (2019), 619–676. MR 3920835. Zbl 1428.35503. <https://doi.org/10.1016/j.aim.2019.02.001>.
- [33] L. FRIEDLANDER, An invariant measure for the equation $u_{tt} - u_{xx} + u^3 = 0$, *Comm. Math. Phys.* **98** no. 1 (1985), 1–16. MR 0785258. Zbl 0576.35082. <https://doi.org/10.1007/BF01211041>.
- [34] P. K. FRIZ and M. HAIRER, *A Course on Rough Paths*, With an introduction to regularity structures, *Universitext*, Springer, Cham, 2014. MR 3289027. Zbl 1327.60013. <https://doi.org/10.1007/978-3-319-08332-2>.
- [35] J. FRÖHLICH, On the triviality of $\lambda\phi_d^4$ theories and the approach to the critical point in $d_{(-)} > 4$ dimensions, *Nuclear Phys. B* **200** no. 2 (1982), 281–296. MR 0643591. [https://doi.org/10.1016/0550-3213\(82\)90088-8](https://doi.org/10.1016/0550-3213(82)90088-8).
- [36] J. GLIMM and A. JAFFE, *Quantum Physics*, A functional integral point of view, second ed., Springer-Verlag, New York, 1987. MR 0887102. <https://doi.org/10.1007/978-1-4612-4728-9>.
- [37] M. GUBINELLI, P. IMKELLER, and N. PERKOWSKI, Paracontrolled distributions and singular PDEs, *Forum Math. Pi* **3** (2015), e6, 75. MR 3406823. Zbl 1333.60149. <https://doi.org/10.1017/fmp.2015.2>.

- [38] M. GUBINELLI, P. IMKELLER, and N. PERKOWSKI, A Fourier analytic approach to pathwise stochastic integration, *Electron. J. Probab.* **21** (2016), Paper No. 2, 37. [MR 3485344](#). [Zbl 1338.60139](#). <https://doi.org/10.1214/16-EJP3868>.
- [39] M. GUBINELLI, H. KOCH, and T. OH, Renormalization of the two-dimensional stochastic nonlinear wave equations, *Trans. Amer. Math. Soc.* **370** no. 10 (2018), 7335–7359. [MR 3841850](#). [Zbl 1400.35240](#). <https://doi.org/10.1090/tran/7452>.
- [40] M. GUBINELLI, H. KOCH, and T. OH, Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity, *J. Eur. Math. Soc. (JEMS)* **26** no. 3 (2024), 817–874. [MR 4721025](#). [Zbl 07834921](#). <https://doi.org/10.4171/jems/1294>.
- [41] M. GUBINELLI and N. PERKOWSKI, Lectures on singular stochastic PDEs, *Ensaaios Mat. [Mathematical Surveys]* **29** (2015), 1–89. [MR 3445609](#). [Zbl 1337.60005](#).
- [42] M. GUBINELLI and N. PERKOWSKI, KPZ reloaded, *Comm. Math. Phys.* **349** no. 1 (2017), 165–269. [MR 3592748](#). [Zbl 1388.60110](#). <https://doi.org/10.1007/s00220-016-2788-3>.
- [43] M. GUBINELLI and N. PERKOWSKI, Energy solutions of KPZ are unique, *J. Amer. Math. Soc.* **31** no. 2 (2018), 427–471. [MR 3758149](#). [Zbl 1402.60077](#). <https://doi.org/10.1090/jams/889>.
- [44] M. GUBINELLI and N. PERKOWSKI, An introduction to singular SPDEs, in *Stochastic Partial Differential Equations and Related Fields*, *Springer Proc. Math. Stat.* **229**, Springer, Cham, 2018, pp. 69–99. [MR 3828162](#). [Zbl 1405.60092](#). https://doi.org/10.1007/978-3-319-74929-7_4.
- [45] Z. GUO and T. OH, Non-existence of solutions for the periodic cubic NLS below L^2 , *Int. Math. Res. Not. IMRN* no. 6 (2018), 1656–1729. [MR 3801473](#). [Zbl 1410.35199](#). <https://doi.org/10.1093/imrn/rnw271>.
- [46] M. HAIRER, Solving the KPZ equation, *Ann. of Math. (2)* **178** no. 2 (2013), 559–664. [MR 3071506](#). [Zbl 1281.60060](#). <https://doi.org/10.4007/annals.2013.178.2.4>.
- [47] M. HAIRER, Singular stochastic PDEs, in *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. 1*, Kyung Moon Sa, Seoul, 2014, pp. 685–709. [MR 3728488](#). [Zbl 1373.60110](#).
- [48] M. HAIRER, A theory of regularity structures, *Invent. Math.* **198** no. 2 (2014), 269–504. [MR 3274562](#). [Zbl 1332.60093](#). <https://doi.org/10.1007/s00222-014-0505-4>.
- [49] M. HAIRER, Introduction to regularity structures, *Braz. J. Probab. Stat.* **29** no. 2 (2015), 175–210. [MR 3336866](#). [Zbl 1316.81061](#). <https://doi.org/10.1214/14-BJPS241>.
- [50] M. HAIRER, Regularity structures and the dynamical Φ_3^4 model, in *Current Developments in Mathematics 2014*, Int. Press, Somerville, MA, 2016, pp. 1–49. [MR 3468250](#). [Zbl 1506.60062](#). <https://doi.org/10.4310/CDM.2014.v2014.n1.a1>.

- [51] M. HAIRER, Renormalisation of parabolic stochastic PDEs, *Jpn. J. Math.* **13** no. 2 (2018), 187–233. MR 3855740. Zbl 1430.60055. <https://doi.org/10.1007/s11537-018-1742-x>.
- [52] M. HAIRER and C. LABBÉ, The reconstruction theorem in Besov spaces, *J. Funct. Anal.* **273** no. 8 (2017), 2578–2618. MR 3684891. Zbl 1380.46031. <https://doi.org/10.1016/j.jfa.2017.07.002>.
- [53] K. IWATA, An infinite-dimensional stochastic differential equation with state space $C(\mathbf{R})$, *Probab. Theory Related Fields* **74** no. 1 (1987), 141–159. MR 0863723. Zbl 0587.60044. <https://doi.org/10.1007/BF01845644>.
- [54] C. KENIG and D. MENDELSON, The focusing energy-critical nonlinear wave equation with random initial data, *Int. Math. Res. Not. IMRN* no. 19 (2021), 14508–14615. MR 4324723. Zbl 1481.35282. <https://doi.org/10.1093/imrn/rnz174>.
- [55] R. KILLIP, J. MURPHY, and M. VISAN, Almost sure scattering for the energy-critical NLS with radial data below $H^1(\mathbb{R}^4)$, *Comm. Partial Differential Equations* **44** no. 1 (2019), 51–71. MR 3933623. Zbl 1416.35241. <https://doi.org/10.1080/03605302.2018.1541904>.
- [56] S. KLAINERMAN, I. RODNIANSKI, and J. SZEFTTEL, The bounded L^2 curvature conjecture, *Invent. Math.* **202** no. 1 (2015), 91–216. MR 3402797. Zbl 1330.53089. <https://doi.org/10.1007/s00222-014-0567-3>.
- [57] A. KUPIAINEN, Renormalization group and stochastic PDEs, *Ann. Henri Poincaré* **17** no. 3 (2016), 497–535. MR 3459120. Zbl 1347.81063. <https://doi.org/10.1007/s00023-015-0408-y>.
- [58] J. L. LEBOWITZ, H. A. ROSE, and E. R. SPEER, Statistical mechanics of the nonlinear Schrödinger equation, *J. Statist. Phys.* **50** no. 3-4 (1988), 657–687. MR 0939505. Zbl 1084.82506. <https://doi.org/10.1007/BF01026495>.
- [59] J. LÜHRMANN and D. MENDELSON, Random data Cauchy theory for nonlinear wave equations of power-type on $\mathbb{R}^3$, *Comm. Partial Differential Equations* **39** no. 12 (2014), 2262–2283. MR 3259556. Zbl 1304.35788. <https://doi.org/10.1080/03605302.2014.933239>.
- [60] J.-C. MOURRAT and H. WEBER, The dynamic Φ_3^4 model comes down from infinity, *Comm. Math. Phys.* **356** no. 3 (2017), 673–753. MR 3719541. Zbl 1384.81068. <https://doi.org/10.1007/s00220-017-2997-4>.
- [61] J.-C. MOURRAT, H. WEBER, and W. XU, Construction of Φ_3^4 diagrams for pedestrians, in *From Particle Systems to Partial Differential Equations*, Springer Proc. Math. Stat. **209**, Springer, Cham, 2017, pp. 1–46. MR 3746744. Zbl 1390.81266. https://doi.org/10.1007/978-3-319-66839-0_1.
- [62] A. R. NAHMOD, T. OH, L. REY-BELLET, and G. STAFFILANI, Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS, *J. Eur. Math. Soc. (JEMS)* **14** no. 4 (2012), 1275–1330. MR 2928851. Zbl 1251.35151. <https://doi.org/10.4171/JEMS/333>.
- [63] A. R. NAHMOD, N. PAVLOVIĆ, G. STAFFILANI, and N. TOTZ, Global flows with invariant measures for the inviscid modified SQG equations, *Stoch. Partial Differ.*

- Equ. Anal. Comput.* **6** no. 2 (2018), 184–210. MR 3818404. Zbl 1406.35413. <https://doi.org/10.1007/s40072-017-0106-5>.
- [64] A. R. NAHMOD, L. REY-BELLET, S. SHEFFIELD, and G. STAFFILANI, Absolute continuity of Brownian bridges under certain gauge transformations, *Math. Res. Lett.* **18** no. 5 (2011), 875–887. MR 2875861. Zbl 1250.60018. <https://doi.org/10.4310/MRL.2011.v18.n5.a6>.
- [65] A. R. NAHMOD and G. STAFFILANI, Almost sure well-posedness for the periodic 3D quintic nonlinear Schrödinger equation below the energy space, *J. Eur. Math. Soc. (JEMS)* **17** no. 7 (2015), 1687–1759. MR 3361727. Zbl 1326.35353. <https://doi.org/10.4171/JEMS/543>.
- [66] E. NELSON, A quartic interaction in two dimensions, in *Mathematical Theory of Elementary Particles* (Proc. Conf., Dedham, Mass., 1965), MIT Press, Cambridge, MA, 1966, pp. 69–73. MR 0210416.
- [67] E. NELSON, Construction of quantum fields from Markoff fields, *J. Functional Analysis* **12** (1973), 97–112. MR 0343815. Zbl 0252.60053. [https://doi.org/10.1016/0022-1236\(73\)90091-8](https://doi.org/10.1016/0022-1236(73)90091-8).
- [68] T. OH, Invariant Gibbs measures and a.s. global well posedness for coupled KdV systems, *Differential Integral Equations* **22** no. 7-8 (2009), 637–668. MR 2532115. Zbl 1240.35477. <https://doi.org/10.57262/die/1356019542>.
- [69] T. OH and L. THOMANN, A pedestrian approach to the invariant Gibbs measures for the 2-*d* defocusing nonlinear Schrödinger equations, *Stoch. Partial Differ. Equ. Anal. Comput.* **6** no. 3 (2018), 397–445. MR 3844655. Zbl 1421.35340. <https://doi.org/10.1007/s40072-018-0112-2>.
- [70] T. OH and L. THOMANN, Invariant Gibbs measures for the 2-*d* defocusing nonlinear wave equations, *Ann. Fac. Sci. Toulouse Math. (6)* **29** no. 1 (2020), 1–26. MR 4133695. Zbl 1443.35101. <https://doi.org/10.5802/afst.1620>.
- [71] N. PERKOWSKI, Paracontrolled distributions and singular SPDEs, 2020, available at author’s website.
- [72] G. RICHARDS, Invariance of the Gibbs measure for the periodic quartic gKdV, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **33** no. 3 (2016), 699–766. MR 3489633. Zbl 1342.35310. <https://doi.org/10.1016/j.anihpc.2015.01.003>.
- [73] B. SIMON, *The $P(\phi)_2$ Euclidean (Quantum) Field Theory*, Princeton Ser. Physics, Princeton Univ. Press, Princeton, NJ, 1974. MR 0489552. Zbl 1175.81146. Available at <https://www.jstor.org/stable/j.ctt13x16st>.
- [74] H. F. SMITH and D. TATARU, Sharp local well-posedness results for the nonlinear wave equation, *Ann. of Math. (2)* **162** no. 1 (2005), 291–366. MR 2178963. Zbl 1098.35113. <https://doi.org/10.4007/annals.2005.162.291>.
- [75] A.-S. DE SUZZONI, Invariant measure for the cubic wave equation on the unit ball of $\mathbb{R}^3$, *Dyn. Partial Differ. Equ.* **8** no. 2 (2011), 127–147. MR 2857361. Zbl 1237.35118. <https://doi.org/10.4310/DPDE.2011.v8.n2.a4>.

- [76] M. SY, Invariant measure and long time behavior of regular solutions of the Benjamin-Ono equation, *Anal. PDE* **11** no. 8 (2018), 1841–1879. [MR 3812859](#). [Zbl 1388.35175](#). <https://doi.org/10.2140/apde.2018.11.1841>.
- [77] L. THOMANN, Invariant Gibbs measures for dispersive PDEs, 2016, Lecture Notes from *Hamiltonian dynamics, PDEs and waves on the Amalfi coast*, available at available at author’s website.
- [78] L. THOMANN and N. TZVETKOV, Gibbs measure for the periodic derivative nonlinear Schrödinger equation, *Nonlinearity* **23** no. 11 (2010), 2771–2791. [MR 2727169](#). [Zbl 1204.35154](#). <https://doi.org/10.1088/0951-7715/23/11/003>.
- [79] N. TZVETKOV, Invariant measures for the nonlinear Schrödinger equation on the disc, *Dyn. Partial Differ. Equ.* **3** no. 2 (2006), 111–160. [MR 2227040](#). [Zbl 1142.35090](#). <https://doi.org/10.4310/DPDE.2006.v3.n2.a2>.
- [80] N. TZVETKOV, Construction of a Gibbs measure associated to the periodic Benjamin-Ono equation, *Probab. Theory Related Fields* **146** no. 3-4 (2010), 481–514. [MR 2574736](#). [Zbl 1188.35183](#). <https://doi.org/10.1007/s00440-008-0197-z>.
- [81] W. WANG and H. YUE, Almost sure existence of global weak solutions to the Boussinesq equations, *Dyn. Partial Differ. Equ.* **17** no. 2 (2020), 165–183. [MR 4071980](#). [Zbl 1434.35067](#). <https://doi.org/10.4310/DPDE.2020.v17.n2.a4>.
- [82] H. YUE, Almost sure well-posedness for the cubic nonlinear Schrödinger equation in the super-critical regime on $\mathbb{T}^d$, $d \leq 3$, *Stoch. Partial Differ. Equ. Anal. Comput.* **9** no. 1 (2021), 243–294. [MR 4218792](#). [Zbl 1475.35324](#). <https://doi.org/10.1007/s40072-020-00183-6>.
- [83] T. ZHANG and D. FANG, Random data Cauchy theory for the generalized incompressible Navier-Stokes equations, *J. Math. Fluid Mech.* **14** no. 2 (2012), 311–324. [MR 2925111](#). [Zbl 1294.35076](#). <https://doi.org/10.1007/s00021-011-0069-7>.
- [84] P. E. ZHIDKOV, An invariant measure for a nonlinear wave equation, *Nonlinear Anal.* **22** no. 3 (1994), 319–325. [MR 1264088](#). [Zbl 0803.35098](#). [https://doi.org/10.1016/0362-546X\(94\)90023-X](https://doi.org/10.1016/0362-546X(94)90023-X).

(Received: October 29, 2019)

(Revised: July 5, 2023)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTHERN CALIFORNIA,
LOS ANGELES, CA, USA
E-mail: yudeng@usc.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MASSACHUSETTS,
AMHERST MA, USA
E-mail: nahmod@math.umass.edu

INSTITUTE OF MATHEMATICAL SCIENCES, SHANGHAI TECH UNIVERSITY,
SHANGHAI, CHINA
E-mail: yuehaitian@shanghaitech.edu.cn