INVARIANT GIBBS MEASURE AND GLOBAL STRONG SOLUTIONS FOR THE HARTREE NLS EQUATION IN DIMENSION THREE

YU DENG\textsuperscript{1}, ANDREA R. NAHMOD\textsuperscript{2}, AND HAITIAN YUE\textsuperscript{3}

ABSTRACT. In this paper we consider the defocusing Hartree nonlinear Schrödinger equations on $\mathbb{T}^3$ with real valued and even potential $V$ and Fourier multiplier decaying like $|k|^{-\beta}$. By relying on the method of random averaging operators \cite{NS2}, we show that there exists $\frac{1}{2} \ll \beta_0 < 1$ such that for $\beta > \beta_0$ we have invariance of the associated Gibbs measure and global existence of strong solutions in its statistical ensemble. In this way we extend Bourgain’s seminal result \cite{B1} which requires $\beta > 2$ in this case.

1. INTRODUCTION

In this paper we study the invariant Gibbs measure problem for the nonlinear Schrödinger (NLS) equation on $\mathbb{T}^3$ with Hartree nonlinearity. Such equation takes the form

\begin{equation}
\begin{cases}
(i\partial_t + \Delta)u = (|u|^2 * V)u, \\
u(0) = u_{in},
\end{cases}
\end{equation}

where $V$ is a convolution potential. We will assume that it satisfies the following properties:

- That $V$ is real-valued and even, and so is $\hat{V}$;
- That \ref{eq:1.1} is defocusing, i.e. $V \geq 0$;
- That $V$ acts like $\beta$ antiderivatives, i.e. $\hat{V}(0) = 1$ and $|\hat{V}(k)| \lesssim (k)^{-\beta}$ for some $\beta \geq 0$.

A typical example for such $V$ is the Bessel potential $\langle \nabla \rangle^{-\beta}$; note that when $V$ is the $\delta$ function (and $\beta = 0$) we recover the usual cubic NLS equation. Our main result, see Theorem \ref{thm:1.3} below, establishes invariance of Gibbs measure for \ref{eq:1.1} when $\beta < 1$ and is close enough to 1, greatly improving the previous result of Bourgain \cite{B1} which assumes $\beta > 2$.

1.1. Background. The equation \ref{eq:1.1} can be viewed as a regularized or tempered version of the cubic NLS equation, and both naturally arise in the limit of quantum many-body problems for interacting bosons (see e.g. \cite{Se, St} and references therein). An important question, both physically and mathematically, is to study the construction and dynamics of the Gibbs measure for \ref{eq:1.1}, which is a Hamiltonian system.

1.1.1. Gibbs measure construction. The Gibbs measure, which we henceforth denote by $d\nu$, is formally expressed as

\begin{equation}
d\nu = e^{-H[u]} \prod_{x \in \mathbb{T}^3} du(x),
\end{equation}

This article is dedicated to the memory of Jean Bourgain.

\textsuperscript{1}Y. D. is funded in part by NSF DMS-1900251.

\textsuperscript{2}A.N. is funded in part by NSF DMS-1800852 and the Simons Foundation Collaborations Grant on Wave Turbulence (Nahmod’s Award ID 651469).
where $H[u]$ is the renormalization of the Hamiltonian
\[
\int_{\mathbb{T}^3} |\nabla u|^2 + \frac{1}{2} |u|^2 (V * |u|^2) \, dx.
\]
Rigorously making sense of (1.2) is closely linked to the construction of the $\Phi^4_3$ measure in quantum field theory, which has attracted a lot of interest since the 70-80’s [28, 32, 1, 21, 23, 27] and in recent years [3, 4, 22, 33]. In the case of (1.1), the answer actually depends on the value of $\beta$.

When $0 < \beta < 1/2$, the measure $d\nu$ can be defined as a weighted version of the Gaussian measure $d\rho$, namely
\[
d\nu = e^{-\int_{\mathbb{T}^3} \frac{1}{2} |u|^2 (V * |u|^2) \, dx} \, d\rho, \quad d\rho \sim e^{-\frac{1}{2} \int_{\mathbb{T}^3} |\nabla u|^2 \prod_{x \in \mathbb{T}^3} \, dx},
\]
where $|u|^2 (V * |u|^2)$ is a suitable renormalization of the nonlinearity (see (1.12) for a precise definition), and the Gaussian free field $d\rho$ is defined as the law of distribution for the random variable
\[
f(\omega) = \sum_{k \in \mathbb{Z}^3} \frac{g_k(\omega)}{\langle k \rangle} \, e^{ik \cdot x},
\]
with $\{g_k(\omega)\}$ being i.i.d. normalized centered complex Gaussians. On the other hand, if $0 < \beta \leq 1/2$, then $d\nu$ is a weighted version of a shifted Gaussian measure, which is singular with respect to $d\rho_1$. These results were proved recently by Bringmann [12] and Oh-Okamoto-Tolomeo [29] by adapting the variational methods of Barashkov-Gubinelli [3].

We remark that, in either case above, it can be shown that the Gibbs measure $d\nu$ is supported in $H^{-1/2}(-\mathbb{T}^3)$, the same space as $d\rho_1$. In particular the typical element in the support of $d\nu$ has infinite mass, which naturally leads to the renormalizations in the construction of $d\nu$ alluded above, see Section 1.2 below. From the physical point of view it is also worth mentioning that, in the same way (1.1) is derived from quantum many-body systems, the Gibbs measure $d\nu$, with the correct renormalizations, can also be obtained by taking the limit of thermal states of such systems, at least when $V$ is sufficiently regular (see [22, 33]).

1.1.2. Gibbs measure dynamics and invariance. Of same importance as the construction of the Gibbs measure is the study of its dynamics and rigorous justification of its invariance under the flow of (1.1). The question of proving invariance of Gibbs measures for infinite dimensional Hamiltonian systems, with interest from both mathematical and physical aspects, has been extensively studied over the last few decades. In fact, it is the works [27, 5, 6]—which attempted to answer this question in some special cases—that mark the very beginning of the subject of random data PDEs.

The literature is now extensive, so we will only review those related to NLS equations. After the construction of Gibbs measures in [27], the first invariance result was due to Bourgain [5], which applies in one dimension for focusing sub-quintic equations, and for defocusing equations with any power nonlinearity. Bourgain [6] then extended the defocusing result to two dimensions, but only for the cubic equation; the two-dimensional case with arbitrary (odd) power nonlinearity was recently solved by the authors [18]. For the case of Hartree nonlinearity (1.1) in three dimensions, Bourgain [7] obtained invariance for $\beta > 2$. We also mention the works of Tzvetkov [35, 36] and

---

1. We will not study the focusing case $V \leq 0$, where the measure can be constructed only when $\beta > 2$; see [29].

2. Actually the law of (1.3) requires another factor, which is $e^{-\|u\|_{L^2}^2}$, in (1.2) and (1.3), which does not make a big difference because the $L^2$ norm is also conserved under (1.1).

---
of Bourgain-Bulut [8, 9] which concern the NLS equation inside a disc or ball, the construction of non-unique weak solutions by Oh-Thomann [31] following the scheme in [2, 16, 14], and the relevant works on wave equations [15, 30, 12, 13, 29]. In particular the recent work of Bringmann [13] establishes Gibbs measure invariance for the wave equation with the Hartree nonlinearity (1.1) for arbitrary $\beta > 0$.

The main mathematical challenge in proving invariance of Gibbs measure is the low regularity of the support of the measure, especially in two or more dimensions. For example, for the two dimensional NLS equation with power nonlinearity, the support of the Gibbs measure $d\nu$ lies in the space of distributions $H^{0-}(\mathbb{T}^2)$, while the scaling critical space is $H^{1/2}(\mathbb{T}^2)$ for the quintic equation, and approaches $H^1(\mathbb{T}^2)$ for equations with high power nonlinearities. This gap is a major reason why the two-dimensional quintic and higher cases have remained open for so many years. In the case of (1.1) a similar gap is present, namely between the support of $d\nu$ at $H^{-1/2-}(\mathbb{T}^3)$ and the scaling critical space $H^{(1-\beta)/2}(\mathbb{T}^3)$ which is higher than $H^0(\mathbb{T}^3)$ with $\beta < 1$.

On the other hand, it is known since the pioneering work of Bourgain [6] that with random initial data, one can go below the classical scaling critical threshold and obtain almost-sure well-posedness results. In the recent works [18, 19] of the authors, an intuitive probabilistic scaling argument was performed. This leads to the notion of the probabilistic scaling critical index $s_{pr} := -1/(p - 1)$ which is much lower than the classical scaling critical index $s_{cr} := (d/2) - 2/(p - 1)$ in the case of $p$-th power nonlinearity in $d$ dimensions. In [19] we proved that almost-sure local well-posedness indeed holds in $H^s$ in the full probabilistic subcritical range when $s > s_{pr}$, in any dimensions and for any (odd) power nonlinearity.

For the case of (1.1), a similar argument as in [18, 19] yields that the probabilistic scaling critical index for (1.1) is $s_{pr} = (-1 - \beta)/2$ which is lower than $-1/2$, so it is reasonable to think that almost-sure well-posedness would be true. However the situation here is somewhat different from [18, 19] due to the asymmetry of the nonlinearity (1.1) compared to the power one, which leads to interesting modifications of the methods in these previous works, as we will discuss in Section 1.3 below.

1.1.3. Probabilistic methods. The first idea in proving almost-sure well-posedness was due to Bourgain [6] and to Da Prato-Debussche [16], the latter in the setting of parabolic SPDEs, which can be described as a linear-nonlinear decomposition. Namely, the solution is decomposed into a linear, random evolution (or noise) term, and a nonlinear term that has strictly higher regularity, thanks to the smoothing effect of randomization. If the linear term has regularity close to scaling criticality, then the nonlinear term can usually be bounded sub-critically, hence a fixed point argument applies. However this idea has its limitations in that the nonlinear term may not be smooth enough, and in practice it is usually limited to slightly supercritical cases (relative to deterministic scaling) and does not give optimal results.

In [18], inspired partly by the regularity structures theory of Hairer and the para-controlled calculus by Gubinelli-Imkeller-Perkowski in the parabolic SPDE setting, we developed the theory of random averaging operators. The main idea is to take the high-low interaction, which is usually the worst contribution in the nonlinear term described above, and express them as a para-product type linear operator—called the random averaging operator—applied to the random initial data. Moreover, this linear operator is independent from the initial data it applies to, and has a randomness structure which includes the information of the solution at lower scales, see Section 1.3.
This structure is then shown to be preserved from low to high frequencies by an induction on scales argument, and eventually leads to improved almost-sure well-posedness results. We refer the reader to [34] for an example of a recent application of the method of random averaging operators of [18] to weakly dispersive NLS.

In [19], the random averaging operators is extended to the more general theory of random tensors. In this theory, the linear operators are extended to multilinear operators which are represented by tensors, and whole algebraic and analytic theories are then developed for these random tensors. For NLS equations with odd power nonlinearity, this theory leads to the proof of optimal almost-sure well-posedness results, see [19]. We remark that, while the theory of random tensors is more powerful than random averaging operators, the latter has a simpler structure, is less notation-heavy, and is already sufficient in many situations (especially if one is not very close to probabilistic criticality).

Finally, we would like to mention other probabilistic methods, developed in the recent works of Gubinelli-Koch-Oh [26], Bringmann [11, 13], and Oh-Okamoto-Tolomeo [29]. These methods also go beyond the linear-nonlinear decomposition, and are partly inspired by the parabolic theories. They have important similarities and differences compared to our methods in [18, 19], but they mostly apply for wave equations instead of Schrödinger equations, so we will not further elaborate here, but refer the reader to the above papers for further explanation.

1.2. Setup and the main result. We start by fixing the i.i.d. normalized (complex) Gaussian random variables \( \{g_k(\omega)\}_{k \in \mathbb{Z}^3} \), so that \( E g_k = 0 \) and \( E |g_k|^2 = 1 \). Let

\[
 f(\omega) = \sum_{k \in \mathbb{Z}^3} g_k(\omega) \frac{1}{\langle k \rangle} e^{ik \cdot x}, 
\]

it is easy to see that \( f(\omega) \in H^{-1/2-} (\mathbb{T}^3) \) almost surely. Let \( V : \mathbb{T}^3 \to \mathbb{R} \) is a potential such that \( V \) is even, nonnegative, and \( V_0 = 1, |V_k| \lesssim \langle k \rangle^{-\beta} \) as described above. Here and below we will use \( u_k \) to denote the Fourier coefficients of \( u \) and use \( \hat{u} \) to represent time Fourier transform only. In this paper we fix \( \beta < 1 \) and sufficiently close\(^1\) to 1. Let \( N \in 2^\mathbb{Z}_{\geq 0} \cup \{0\} \) be a dyadic scale, define projections \( \Pi_N \) such that \( (\Pi_N u)_k = 1_{\langle k \rangle \leq N} \cdot u_k \), and \( \Delta_N = \Pi_N - \Pi_{N/2} \), and define

\[
 f_N(\omega) = \Pi_N f(\omega), 
 F_N(\omega) = \Delta_N f(\omega) = f_N(\omega) - f_{N/2}(\omega). 
\]

we introduce the following truncated and renormalized version of (1.1), with truncated random initial data, namely:

\[
 \begin{cases}
 i \partial_t u_N + \Delta u_N = \Pi_N[(|u_N|^2 * V) \cdot u_N] - \sigma_N u_N - C_N u_N \\
 u_N(0) = \Pi_N u_{in}.
\end{cases} 
\]  

Here in (1.7) we fix

\[
 \sigma_N = E \int_{\mathbb{T}^3} |f_N(\omega)|^2 \, dx = \sum_{\langle k \rangle \leq N} \frac{1}{\langle k \rangle^{2}}, 
\]

and \( C_N \) is a Fourier multiplier,

\[
 (C_N u)_k = (C_N)_k \cdot u_k, \quad (C_N)_k := \sum_{\langle \ell \rangle \leq N} \frac{V_{k-\ell}}{\langle \ell \rangle^{2}}. 
\]

\(^1\)This is a specific value but we do not track it below.
Note that $u_N$ is supported in $\langle k \rangle \leq N$ for all time. The first counterterm in (1.7), namely $-\sigma_N u_N$, corresponds to the standard Wick ordering, where one fixes $k_1 = k_2$ in the expression

$$[(|u|^2 * V) \cdot u]_k = \sum_{k_1 - k_2 + k_3 = k} V_{k_1 - k_2} \cdot u_{k_2} u_{k_3},$$

(1.10)

plugs in $u = f_N(\omega)$, and takes expectations. The second term $-\mathcal{C}_N u_N$ corresponds to fixing $k_2 = k_3$, which is present due to the asymmetry of the nonlinearity $(|u|^2 * V) \cdot u$. Note that $(\mathcal{C}_N)_k$ is uniformly bounded, and thus is unnecessary, if $\beta > 1$ (in particular this is the case of Bourgain [7]); if $\beta < 1$ this becomes a divergent term which needs to be subtracted.

The equation (1.7) is a finite dimensional Hamiltonian equation with Hamiltonian

$$H_N[u] := \int_{T^3} \left( |\nabla u|^2 + \frac{1}{2} |u|^2 (V * |u|^2) - \sigma_N |u|^2 - \frac{1}{2} \mathcal{C}_N u \cdot \overline{u} + \frac{1}{2} \sigma_N - \frac{1}{2} \gamma_N \right),$$

(1.11)

where $\gamma_N = \sum_{\langle k \rangle, \langle \ell \rangle \leq N} \frac{V_{k \cdot \ell}}{|\ell|^2}.$

Remark 1.1. In fact, the Hamiltonian $H_N[u]$ can be also expressed as $\int_{T^3} \left( |\nabla u|^2 + |u|^2 (V * |u|^2) \right)$, where the suitable renormalized nonlinearity $:|u|^2 (V * |u|^2):$ is defined

$$:|u|^2 (V * |u|^2) := |u|^2 (V * |u|^2) - \sigma_N (V * |u|^2) - \sigma_N |u|^2 - \mathcal{C}_N u \cdot \overline{u} + \sigma_N - \gamma_N.$$

(1.12)

Notice that $\int_{T^3} \sigma_N (V * |u|^2) = \int_{T^3} \sigma_N |u|^2$ since $\hat{V}(0) = 1$.

We can define the corresponding truncated and renormalized Gibbs measure, namely

$$d\eta_N(u) = \frac{1}{Z_N} e^{-H_N[u] - ||u||^2_{L^2}} \prod_{\langle k \rangle \leq N} du_k du_k$$

(1.13)

where $Z_N > 0$ is a normalization constant making $d\nu_N$ a probability measure. Clearly $d\nu_N$ is invariant under the finite dimensional flow (1.7). Note that we can also write

$$d\eta_N(u) = \frac{1}{Z_N^*} e^{-H_N^{pot}[u]} d\rho_N(u),$$

(1.14)

where $Z_N^*$ is another positive constant, $d\rho_N$ is the law of distribution for the linear Gaussian random variable $f_N(\omega) := \Pi_N f(\omega)$, and $H_N^{pot}[u]$ represents the potential energy

$$H_N^{pot}[u] = \int_{T^3} \left( \frac{1}{2} |u|^2 (V * |u|^2) - \sigma_N |u|^2 - \frac{1}{2} \mathcal{C}_N u \cdot \overline{u} + \frac{1}{2} \sigma_N - \frac{1}{2} \gamma_N \right).$$

(1.15)

Now, define $\Pi_N^\perp = 1 - \Pi_N$, let $\mathcal{V}_N$ and $\mathcal{V}_N^\perp$ be the ranges of the projections $\Pi_N$ and $\Pi_N^\perp$, and define $d\rho$ and $d\rho_N^\perp$ be the laws of distribution for $f(\omega)$ and $\Pi_N^\perp f(\omega)$ respectively. Then we have $d\rho = d\rho_N \times d\rho_N^\perp$, moreover we define

$$d\nu_N = d\eta_N \times d\rho_N^\perp = G_N(u) \cdot d\rho,$$

where $G_N(u) := \frac{1}{Z_N} e^{-H_N^{pot}[\Pi_N u]}$.

We have the following result. Recall that in this paper we are fixing $\beta < 1$ close enough to 1, in particular $\beta > 1/2$.

**Proposition 1.2.** Suppose $\beta > 1/2$, then $G_N(u)$ converges to a limit $G(u)$ in $L^q(d\rho)$ for all $1 \leq q < \infty$, and the sequence of measures $d\nu_N$ converges to a probability measure $d\nu$ in the sense of total variations. The measure $d\nu$ is call the Gibbs measure associated with the system (1.7).
Proof. This is proved in the recent works of Bringmann \[12\] and Oh-Okamoto-Tolomeo \[29\]. Strictly speaking they are dealing with the case of real-valued $u$ (as they are concerned about the wave equation), but the proof can be readily adapted to the complex-valued case here. \qed

Now we can state our main theorem.

**Theorem 1.3.** Let $\beta < 1$ be close enough to 1. There exists a Borel set $\Sigma \subset H^{-1/2-}(\mathbb{T}^3)$ such that $\nu(\Sigma) = 1$, and the following holds. For any $u_{\text{in}} \in \Sigma$, let $u_N(t)$ be defined by (1.7), then

$$\lim_{N \to \infty} u_N(t) = u(t)$$

exists in $C_t^0 H^{-1/2-}_x(\mathbb{R} \times \mathbb{T}^3)$, and $u(t) \in \Sigma$ for all $t \in \mathbb{R}$. This $u(t)$ solves (1.1) with a suitably renormalized nonlinearity, and defines a mapping $\Phi_t : \Sigma \to \Sigma$ for each $t \in \mathbb{R}$. These mappings satisfy the group properties $\Phi_{t+s} = \Phi_t \Phi_s$, and keeps the Gibbs measure $d\nu$ invariant, namely $\nu(E) = \nu(\Phi_t(E))$ for any $t \in \mathbb{R}$ and Borel set $E \subset \Sigma$.

**Remark 1.4.** As in \[18, 19\], the sequence $\{u_N\}$ can be replaced by other canonical approximation sequences, for example with the sharp truncations $\Pi_N$ on initial data replaced by smooth truncations, or with the projection $\Pi_N$ on the nonlinearity in (1.7) omitted. The limit obtained does not depend on the choice of such sequences, and the proof will essentially be the same.

1.2.1. **Regarding the range of $\beta$.** The range of $\beta$ obtained in Theorem 1.3 is clearly not optimal. In fact, the equation (1.1) with Gibbs measure data is probabilistically subcritical as long as $\beta > 0$, and one should expect the same result at least when $\beta > 1/2$ (so the Gibbs measure is absolutely continuous with the Gaussian free field).

The purpose of this paper, however, is to provide an example where the method of random averaging operators \[18\] is applied so that one can significantly improve the existing probabilistic results ($\beta$ close but smaller than 1 versus $\beta > 2$ in \[17\]), while keeping the presentation relatively short. In order to treat $\beta > 1/2$ one would need to adapt the sophisticated theory of random tensors \[19\] which will considerably increase the length of this work, so we decide to leave this part to a next paper.

As for the case $0 < \beta < 1/2$, one would need to deal with the mutual singularity between the Gibbs measure and the Gaussian free field (of course, if one studies the local well-posedness problem with Gaussian initial data as in (1.5), which is of course different from Gibbs, then a modification of the random tensor theory \[19\] would also likely work for all $\beta > 0$). The recent work of Bringmann \[13\] provides a nice example where this issue is solved in the context of wave equations, and it would be interesting to see whether this can be extended to Schrödinger equations. Finally, the case $\beta = 0$, which is the famous Gibbs measure invariance problem for the three-dimensional cubic NLS equation, still remains an outstanding open problem as of now. It is probabilistically critical, which presumably would require completely new techniques to solve.

1.3. **Main ideas.** Due to the absolute continuity of the Gibbs measure in Proposition 1.2 in order to prove Theorem 1.3, we only need to consider initial data distributed according to $d\rho$ for (the renormalized version of) (1.1), and the initial data distributed according to $d\rho_N$ for (1.7). In other words, we may assume $u(0) = f(\omega)$ for (1.1), and $u_N(0) = f_N(\omega)$ for (1.7).

\[1\]We remark that Bringmann has an unpublished proof for the same result assuming $\beta > 3/2$. 

1.3.1. Random averaging operators. Let us focus on (1.7); for simplicity we will ignore the renormalization terms. The approach of Bourgain and of Da Prato-Debussche corresponds to decomposing

\[ u_N(t) = e^{it\Delta} f_N(\omega) + v(t), \]

where \( f_N \) is as in (1.6), and \( v(t) \) is the nonlinear evolution. In particular this \( v(t) \) contains a trilinear Gaussian term

\[ v^*(t) = \int_0^t e^{i(t-t')\Delta} \Pi_N[(|e^{it'\Delta} f_N(\omega)|^2 * V)e^{it'\Delta} f_N(\omega)] \, dt'. \]

This term turns out to only have \( H^{0-} \) regularity, which is not regular enough for a fixed point argument (note that the classical scaling critical threshold is \( H^{(1-\beta)/2} \)). Therefore this approach does not work.

Nevertheless, one may observe that the only contribution to \( v^* \) that has worst \( (H^{0-}) \) regularity is when the first two input factors are at low frequency and the third factor is at high frequency, such as

\[ \int_0^t e^{i(t-t')\Delta} \Pi_N[(|e^{it'\Delta} f_{N'}(\omega)|^2 * V)e^{it'\Delta} F_N(\omega)] \, dt', \]

for \( N' \ll N \) and \( F_N \) as in (1.6). Moreover this low frequency component \( f_{N'} \) may also be replaced by the corresponding nonlinear term at frequency \( N' \), so it makes sense to separate the low-low-high interaction term \( \psi^N \) defined by

\[ \begin{cases} 
(i\partial_t + \Delta) \psi^N = \Pi_N[(|u_{N/2}|^2 * V)\psi^N], \\
\psi^N(0) = F_N(\omega)
\end{cases} \quad (1.16) \]

as the singular part of \( y_N := u_N - u_{N/2} \), so that \( y_N - \psi^N \) has higher regularity.

The idea of considering high-low interactions is consistent with the para-controlled calculus in [24, 25, 26]. However in those works the singular term \( \psi^N \) and the regular term \( y_N - \psi^N \) are characterized only by their regularity (for example one is constructed via fixed point argument in \( H^{0-} \) and the other in \( H^{1/2-} \)), which, as pointed out in [18], is not enough in the context of Schrödinger equations. Instead, it is crucial that one studies the operator, referred to as the random averaging operator in [18], which maps \( z \) to the solution to the equation

\[ \begin{cases} 
(i\partial_t + \Delta) \psi = \Pi_N[(|u_{N/2}|^2 * V)\psi], \\
\psi(0) = z.
\end{cases} \quad (1.17) \]

Note that the kernel of this operator, which we denote by \( H^N = (H^N)_{kk'}(t) \), is a Borel function of \( \{g_k(\omega)\}_{(k) \leq N/2} \) and is independent from \( F_N(\omega) \). Moreover, this \( H^N \) encodes the whole randomness structure of \( u_{N/2} \), which is captured in two particular matrix norm bounds for \( H^N \). Essentially, they involve the \( \ell^2_k \rightarrow \ell^2_{k'} \) operator norm and the \( \ell^2_{kk'} \) Hilbert-Schmidt norm for fixed time \( t \) (or fixed Fourier variable \( \lambda \)), see Section 2.2.2 for details.

This is the main idea of the random averaging operators in [18]. Basically, it allows one to fully exploit the randomness structure of the solution at all scales, which is necessary for the proof in the setting of Schrödinger equations in the lack of any smoothing effect.
The special term $\rho^N$: a ‘critical’ component. In addition to the ansatz introduced in Section 1.3.1, it turns out that an extra term is necessary due to the structure (especially the asymmetry) of the nonlinearity (1.1). Recall that $(|u|^2 * V)u$ can be expressed as in (1.10); for simplicity we will ignore any resonances (which are cancelled by the renormalizations), i.e. assume $k_2 \notin \{k_1, k_3\}$ in (1.10). Here, if $|k_1 - k_2| \gtrsim N\varepsilon$ for some small constant $\varepsilon$, then the potential $V_{k_1-k_2}$, which is bounded by $(k_1-k_2)^{-\beta}$, will transform into a derivative gain, which allows one to close easily using the random averaging operator ansatz in Section 1.3.1.

However, suppose $|k_1 - k_2|$ is very small, say $|k_1 - k_2| \sim 1$ in (1.10), then the potential does not lead to any gain of derivatives, and we will see that this particular term in fact exhibits some (probabilistically) “critical” feature. To see this, let us define $u$ as in (1.1), it follows from simple calculations that $u(0)$ has regularity $H^{-1/2-}$, while each $u^{(m)}$, where $m \geq 1$ has exactly regularity $H^{1/2-}$. Therefore, although $u^{(1)}$ is indeed more regular than $u(0)$, the higher order iterations are not getting smoother despite all input functions (which are $F_N(\omega)$) having the same (and high) frequency. This is in contrast with the “genuinely (probabilistically) subcritical” situations (for the standard NLS) in [19], where for fixed positive constants $\varepsilon$ and $c$, the $m$-th iteration $u^{(m)}$, assuming all input frequencies are the same, will have increasing and positive regularity in $H^{2m-\varepsilon}$ as $m$ grows and becomes large. Similarly, one may consider the linear operator

$$z \mapsto \int_0^t e^{i(t-t')\Delta} N(z(t'), e^{it'\Delta} F_N(\omega), e^{it'\Delta} F_N(\omega)) \, \mathrm{d}t',$$

with $N$ as in (1.18) and in typical subcritical cases the norm of this operator from a suitable $X^{s,b}$ space to itself would be $N^{-\alpha}$ for some $\alpha > 0$, see [18, 19]. However here (for Hartree) one can check that the corresponding norm is in fact $\sim 1$, and may even exhibit a logarithmic divergence if one adds up different scales.

Therefore, it is clear that the contribution $N$ as in (1.18) needs a special treatment in addition to the ansatz in Section 1.3.1. Fortunately, this term does not depend on the value of $\beta$ and was already treated in Bourgain’s work [7]. In this work, we introduce an extra term $\rho^N$, which corresponds to the term treated in Bourgain [7], by defining $\xi^N$ such that

$$\begin{cases} (i\partial_t + \Delta)\xi^N = \Pi_N([(u_{N/2}|^2 * V)\xi^N + \Pi_{N^\varepsilon}((|u_N|^2 - |u_{N/2}|^2) * V)])\xi^N, \\
\xi^N(0) = F_N(\omega) \end{cases}$$

and defining $\rho^N = \xi^N - \psi^N$, where $\Pi_{N^\varepsilon}$ is a smooth truncation at frequency $N^\varepsilon$ for some small $\varepsilon$. This term is then measured at regularity $H^s$ for some $s < 1/2$, while the remainder term $z_N := y_N - \xi^N$, where $y_N = u_N - u_{N/2}$, is measured at regularity $H^{s'}$ for some $s < s' < 1/2$. See Section 3.1.3 for the solution ansatz and Proposition 3.1 for the precise formulations.
1.3.3. Additional remark. Note that the precise definitions of the equations satisfied by $\psi^N$ and $\xi^N$, see (3.2) and (3.8), involve projection $\Delta_N$ on the right hand sides; this is to make sure that $(\psi^N)_k$ and $(\xi^N)_k$ are exactly supported in $N/2 < \langle k \rangle \leq N$, so that one can exploit the cancellation due to the unitarity of the matrices $H^N$ (corresponding to $\psi^N$), as well as the matrices $M^N$ which corresponds to the term $\xi^N$. This unitarity comes from the mass conservation property of the linear equations defining these matrices, and already plays a key role in Bourgain’s work [7]. See Section 3.2 for details.

2. Preparations

2.1. Reduction of the equation. We start with the system (1.7) with initial data $u_N(0) = f_N(\omega)$. Clearly $(u_N)_k$ is supported in $\langle k \rangle \leq N$. If we denote the right hand side of (1.7) by $\Pi_NN(u_N)$, then in Fourier space we have

$$\mathcal{N}(u)_k = \mathcal{N}^\circ(u)_k + u_k \cdot \left( \sum_\ell |u_{\ell}|^2 - \alpha_N \right) + u_k \cdot \sum_{\ell \neq k, \langle \ell \rangle \leq N} V_{k-\ell} \left( |u_{\ell}|^2 - \frac{1}{\langle \ell \rangle^2} \right) - \frac{u_k}{\langle k \rangle^2}, \quad (2.1)$$

$$\mathcal{N}^\circ(u)_k = \sum_{k_1 - k_2 + k_3 = k \atop k_2 \notin \{k_1, k_3\}} V_{k_1 - k_2} \cdot u_{k_1} v_{k_2} v_{k_3}. \quad (2.2)$$

We will extend $\mathcal{N}^\circ(u)$, which is a cubic polynomial of $u$, to an $\mathbb{R}$-trilinear operator $\mathcal{N}^\circ(u, v, w)$ in the standard way. Note that

$$\sum_\ell |(u_N)_\ell|^2 = \int_{\mathbb{T}^3} |u_N|^2 \, dx$$

is conserved under the flow (1.7), we may get rid of the second term on the right hand side of (2.1) by a gauge transform

$$u_N \to e^{iB_N t} u_N, \quad B_N := \sum_{\langle \ell \rangle \leq N} \frac{|g_\ell|^2 - 1}{\langle \ell \rangle^2}. \quad$$

If we further define the profile $v_N$ by

$$(v_N)_k(t) = e^{-it|k|^2} e^{iB_N t} (u_N)_k(t),$$

then $v$ will satisfy the integral equation

$$(v_N)_k(t) = (f_N)_k - i \int_0^t \Pi_N \mathcal{M}^\circ(v_N, v_N, v_N)_k(s) \, ds$$

$$- i \int_0^t (v_N)_k(s) \sum_{\ell \neq k, \langle \ell \rangle \leq N} V_{k-\ell} \left( |(v_N)_\ell(s)|^2 - \frac{1}{\langle \ell \rangle^2} \right) \, ds + i \int_0^t \frac{1}{\langle k \rangle^2} \, ds \quad (2.3)$$

where

$$\mathcal{M}^\circ(u, v, w)_k(s) = \sum_{k_1 - k_2 + k_3 = k \atop k_2 \notin \{k_1, k_3\}} e^{is\Omega} V_{k_1 - k_2} u_{k_1}(s) \overline{v_{k_2}(s)} w_{k_3}(s), \quad \Omega := |k_1|^2 - |k_2|^2 + |k_3|^2 - |k|^2. \quad (2.4)$$

Below we will focus on the system (2.3)--(2.4).

2.2. Notations and norms. We setup some basic notations and norms needed later in the proof.
2.2.1. Notations. As denoted above, we will use $v_k$ to denote Fourier coefficients, and $\mathcal{F}v_k = \hat{v}_k = \hat{v}_k(\lambda)$ denotes the Fourier transform in time. For a finite index set $A$, we will write $k_A = (k_j : j \in A)$ where each $k_j \in \mathbb{Z}^3$ and denote by $h_{k_A}$ a tensor $h : (\mathbb{Z}^3)^A \to \mathbb{C}$. We may also define tensors involving $\lambda$ variables where $\lambda \in \mathbb{R}$.

We fix the parameters, to be used in the proof, as follows. Let $\varepsilon > 0$ be sufficiently small absolute constant. Let $\varepsilon_1$ and $\varepsilon_2$ be fixed such that $\varepsilon_2 < \varepsilon_1 < \varepsilon$. Let $\beta < 1$ be such that $1 - \beta < \varepsilon_2$, and choose $\delta$ such that $\delta < 1 - \beta$, and $\kappa$ such that $\kappa > \delta^{-1}$. We use $\theta$ to denote any generic small positive constant such that $\theta < \delta$ (which may be different at different places). Let $b = 1/2 + \kappa^{-1}$, so $1 - b = 1/2 - \kappa^{-1}$. Finally, let $\tau$ be sufficiently small compared to all the above parameters, denote $J = [-\tau, \tau]$. Fix a smooth cutoff function $\chi(t)$ which equals 1 for $|t| \leq 1$ and equals 0 for $|t| \geq 2$, and define $\chi_{\tau}(t) := \chi(\tau^{-1}t)$. We use $C$ to denote any large absolute constant, and $C_\theta$ for any large constant depending on $\theta$. If some event happens with probability $\geq 1 - C_\theta e^{-A^\theta}$, where $A$ is a large parameter, we say this event happens $A$-certainly.

2.2.2. Norms. If $(B, C)$ is a partition of $A$, namely $B \cap C = \emptyset$ and $B \cup C = A$, we define the norm $\|h\|_{k_B \to k_C}$ such that

$$\|h\|_{k_B \to k_C}^2 = \sup \left\{ \sum \left| \sum h_{k_B} z_{k_B} \right|^2 : \sum \left| z_{k_B} \right|^2 = 1 \right\}. \quad (2.5)$$

The same notation also applies for tensors involving the $\lambda$ variables. For functions $u = u_k(t)$ and $h = h_{kk'}(t)$, and $0 < c < 1$, we also define the norms

$$\|u\|_{X^c}^2 := \int \langle \lambda \rangle^{2c} \|\hat{u}_k(\lambda)\|^2_k \, d\lambda,$$

$$\|h\|_{Y^c}^2 := \int \langle \lambda \rangle^{2c} \|\hat{h}_{kk'}(\lambda)\|^2_{k \to k'} \, d\lambda,$$

$$\|h\|_{Z^c}^2 := \int \langle \lambda \rangle^{2c} \|\hat{h}_{kk'}(\lambda)\|^2_{kk'} \, d\lambda. \quad (2.6)$$

For any interval $I$, define the corresponding localized norms

$$\|u\|_{X^c(I)} := \inf \left\{ \|v\|_{X^c} : v = u \text{ on } I \right\} \quad (2.7)$$

and similarly define $Y^c(I)$ and $Z^c(I)$. By abusing notations, we will call the above $u$ an extension of $u$, though it’s actually an extension of the restriction of $u$ to $I$.

2.3. Preliminary estimates. Here we record some basic estimates. Most of them are standard, or are in our previous works [18] [19].

2.3.1. Linear estimates. Define the original and truncated Duhamel operators

$$Iv(t) = \int_0^t v(t') \, dt', \quad \mathcal{I}_\lambda v(t) = \chi(t) \int_0^t \chi(t') v(t') \, dt'. \quad (2.8)$$

Lemma 2.1. We have the formula

$$\widehat{\mathcal{I}_\lambda v}(\lambda) = \int_R I(\lambda, \lambda') \hat{v}(\lambda') \, d\lambda', \quad (2.9)$$

where the kernel $I$ satisfies that

$$|I| + |\partial_{\lambda, \lambda'} I| \lesssim \left( \frac{1}{\langle \lambda \rangle^3} + \frac{1}{(\lambda - \lambda')^3} \right) \frac{1}{\langle \lambda \rangle} \lesssim \frac{1}{\langle \lambda \rangle} \langle \lambda - \lambda' \rangle. \quad (2.10)$$
Proof. See [17], Lemma 3.1 whence by a similar proof, one can also prove (2.10) for $|\partial_{\lambda_M} I|$. □

Proposition 2.2 (Short time bounds). Let $\varphi$ be any Schwartz function, recall that $\varphi_t(t) = \varphi(\tau^{-1} t)$ for $\tau \ll 1$. Then for any $u = u_k(t)$ we have

$$\| \varphi_t \cdot u \|_{X^c} \lesssim \| u \|_{X^{c_1}}$$

(2.11)

provided either $0 < c \leq c_1 < 1/2$, or $u_k(0) = 0$ and $1/2 < c \leq c_1 < 1$. The same result also holds if $u = u(t)$ is measured in norms other than $\ell^2$, so (2.11) is true with $X$ replaced by $Y$ or $Z$.

Proof. See [19], Lemma 4.2. □

Lemma 2.3 (Suitable extensions). Suppose $f(x,t)$ is a function defined in $t \in [-\tau, \tau] = J$ with $|\tau| \ll 1$. Define that

$$g(t) = \begin{cases} 
  f(t) & \text{if } |t| \leq \tau \\
  f(\tau) & \text{if } t > \tau \\
  f(-\tau) & \text{if } t < -\tau.
\end{cases}$$

(2.12)

For any Schwartz function $\varphi$, we have

$$\| \varphi(t) \cdot g(t) \|_{X^b} \lesssim \| f \|_{X^b_1(J)} + \| f \|_{L^\infty L^2(J)}$$

(2.13)

provided either $0 < b < b_1 < 1/2$ or $1/2 < b < b_1 < 1$. When $1/2 < b < b_1 < 1$, we have

$$\| \varphi(t) \cdot g(t) \|_{X^b} \lesssim \| f \|_{X^b_1(J)}.$$ 

(2.14)

Proof. We only need to bound locally-in-time the function $f^*(t)$, which equals $f(0)$ for $t \geq 0$ and $f(t)$ for $t < 0$; in fact $g$ is obtained by performing twice the transformation from $f$ to $f^*$, first at center $\tau$ and then at center $-\tau$.

We can decompose $f$ into two parts, $f_1$ which is smooth and equals $f(0)$ near 0, and $f_2$ such that $f_2(0) = 0$. Clearly we only need to consider $f_2$, so that $f^*$ equals $f_2$ multiplied by a smooth truncation of $1_{[0, +\infty)}$, with $f_2(0) = 0$.

We may replace $1_{[0, +\infty)}$ by the sign function, and then apply Proposition 2.2. Note that for an even smooth cutoff function $\chi$,

$$\chi(x) \cdot sgn(x) = \sum_{N \geq 1} \Delta_N \chi \cdot sgn(x)$$

where $\Delta_N$ are the standard Littlewood-Paley projections. Moreover $\Delta_N (\chi \cdot sgn)(x)$ can be viewed as a rescaled Schwartz function of the same form as in Proposition 2.2 with $\tau = N^{-1}$ (due to the expression of the Fourier transform of $\text{sgn}$ and simple calculations), so the desired result follows from Proposition 2.2. □

2.3.2. Counting estimates. Here we list some counting estimates and the resulting tensor norm bounds.

Lemma 2.4. (1) Let $R = \mathbb{Z}$ or $\mathbb{Z}[i]$. Then, given $0 \neq m \in R$, and $a_0, b_0 \in \mathbb{C}$, the number of choices for $(a, b) \in R^2$ that satisfy

$$m = ab, \ |a - a_0| \leq M, \ |b - b_0| \leq N$$

(2.15)

is $O(M^9 N^9)$ with constant depending only on $\theta > 0$. 

For dyadic numbers $N_1, N_2, N_3, R > 0$ and some fixed number $\Omega_0$.

\[
S^R = \left\{ (k, k_1, k_2, k_3) \in (\mathbb{Z}^3)^4, \ k_2 \notin \{k_1, k_3\} \right\},
\]

(2.16)

and then $S^R_k$ is the set of $(k, k_1, k_2, k_3) \in S^R$ when $k$ is fixed and etc. We have the following counting estimates

\[
|S^R| \leq \min (N_1^3 N_3^2 (N_2 \wedge N)^{1+\theta}, N_3 N_2^2 (N_1 \wedge N_3)^{1+\theta}, N_2^3 (RN_3)^{2+\theta}, N_3^3 (RN_1)^{2+\theta});
\]

(2.17)

\[
|S^R_k| \leq \min (N_2^3 (N_1 \wedge N_3)^{1+\theta}, (N_1 N_3)^{2+\theta}, (RN_1)^{2+\theta});
\]

(2.18)

\[
|S_{k_3}^R| \leq \min (N_3^2 (N_2 \wedge N)^{1+\theta}, (N_2 N)^{2+\theta}, (RN_2)^{2+\theta});
\]

(2.19)

\[
|S_{k_2}^R| \leq \min (N_3^2 (N_1 \wedge N_3)^{1+\theta}, (N_1 N_3)^{2+\theta}, (RN_3)^{2+\theta});
\]

(2.20)

\[
|S_{k_1 k_3}^R| \leq \min (N_2 N_3, R)^{2+\theta}; \quad |S_{k_1 k_3}^R| \leq \min (N_1, N_1, R)^{2+\theta}.
\]

(2.21)

\[
|S_{k_1 k_2}^R| \leq \min (N_1, N_3, R)^{1+\theta}; \quad |S_{k_1 k_2}^R| \leq \min (N_2, N, R)^{1+\theta};
\]

(2.22)

\[
|S_{k_1 k_2}^R| \leq \min (N_1, N_2, R)^{2+\theta}; \quad |S_{k_1 k_2}^R| \leq \min (N, N_3, R)^{2+\theta}.
\]

(2.23)

Proof. (1) It is the same as the part (1) of Lemma 4.3 in [18]. (2) We consider $|S^R|$. First the number of choices of $k_1$ and $k_3$ is $N_1^3 N_3^2$. After fixing the choice of $k_1$ and $k_3$ to count $(k, k_2)$, it is equivalent to count $k_2$ satisfying the restriction $|k_2|^2 + |k_2 + c_1|^2 = c_2$ or to count $k$ satisfying the restriction $|k|^2 + |k + c_3|^2 = c_4$ for some fixed numbers $c_1, \ldots, c_4$ and hence we have $|S^R| \leq N_1^3 N_3^2 (N_2 \wedge N)^{1+\theta}$. Similarly if we first fix $k$ and $k_2$, we have $|S^R| \leq N_2^3 N_3^2 (N_1 \wedge N_3)^{1+\theta}$. Also if we fix $k_2$ first, then to count $(k, k_1, k_3)$ is equivalent to count $(k_1, k_3)$ with the restriction $(k_2 - k_1) \cdot (k_2 - k_3) = c$ for some fixed number $c$. By fixing the first two components of $(k_1, k_3)$ and using part (1), we have $|S^R_{k_1 k_3}| \leq N_3^3 (RN_3)^{2+\theta}$. Similarly we also have $|S^R| \leq N_3^3 (RN_1)^{2+\theta}$. The proofs of (2.18)–(2.23) are similar.

\[\square\]

2.3.3. Probabilistic and tensor estimates.

**Proposition 2.5** (Proposition 4.11 in [19]). Consider two tensors $h_{kA_1}^{(1)}$ and $h_{kA_2}^{(2)}$, where $A_1 \cap A_2 = C$. Let $A_1 \Delta A_2 = A$, define the semi-product

\[
H_{kA} = \sum_{kC} h_{kA_1}^{(1)} h_{kA_2}^{(2)}.
\]

(2.24)

Then, for any partition $(X, Y)$ of $A$, let $X \cap A_1 = X_1$, $Y \cap A_1 = Y_1$ etc., we have

\[
\|H\|_{kX \rightarrow kY} \leq \|h_{kX_1 \cup \bar{C}}^{(1)}\|_{kX_1 \rightarrow kY_1} \cdot \|h_{kY_2 \rightarrow kC \cup \bar{Y}_2}^{(2)}\|_{kY_2 \rightarrow kC \cup \bar{Y}_2}.
\]

(2.25)

**Proposition 2.6** (Proposition 4.12 in [19]). Let $A_j (1 \leq j \leq m)$ be index sets, such that any index appears in at most two $A_j$’s, and let $h^{(j)} = h_{kA_j}^{(j)}$ be tensors. Let $A = A_1 \Delta \cdots \Delta A_m$ be the set of indices that belong to only one $A_j$, and $C = (A_1 \cup \cdots \cup A_m) \setminus A$ be the set of indices that belong to two different $A_j$’s. Define the semi-product

\[
H_{kA} = \sum_{kC} \prod_{j=1}^{m} h_{kA_j}^{(j)}.
\]

(2.26)
Let \((X, Y)\) be a partition of \(A\). For \(1 \leq j \leq m\) let \(X_j = X \cap A_j\) and \(Y_j = Y \cap A_j\), and define

\[ B_j := \bigcup_{\ell \geq j} (A_j \cap A_\ell), \quad C_j = \bigcup_{\ell < j} (A_j \cap A_\ell), \tag{2.27} \]

then we have

\[ \|H\|_{k_X \to k_Y} \leq \prod_{j=1}^m \|h^{(j)}\|_{k_{X_j} \cup B_j \to k_{Y_j} \cup C_j}. \tag{2.28} \]

For the proofs of Propositions 2.5 and 2.6 see 19. In that work the full power of (2.25) and (2.28) is needed, but here we only need some specific cases, mainly those of the following form (where \(q \leq r\))

\[ \left\| \sum_{k_1, \ldots, k_q} H_{k_1 \cdots k_q} h^{(1)}_{k_1} \cdots h^{(q)}_{k_q} \right\|_{k_{A'} \to k_{B'}} \leq \|H\|_{k_{A'} \to k_B} \prod_{j=1}^q \|h^{(j)}\|_{k_j \to k'_j}. \tag{2.29} \]

where \((k_{A'}, k_{B'})\) is a partition of the variables \((k_1', \ldots, k_q', k_{q+1}', \ldots, k_r')\) and \((k_A, k_B)\) is a partition of the variables \((k_1, \ldots, k_r)\) where each \(k_j'(1 \leq j \leq q)\) is replaced by \(k_j\) in \((k_A', k_{B'}).\)

**Proposition 2.7** (Proposition 4.14 in 19). Let \(A\) be a finite set and \(h_{bck_A} = h_{bck_A}(\omega)\) be a tensor, where each \(k_j \in \mathbb{Z}^d\) and \((b, c) \in (\mathbb{Z}^d)^q\) for some integer \(q \geq 2\). Given signs \(\zeta_j \in \{\pm\}\), we also assume that \(\langle b \rangle, \langle c \rangle \lesssim M\) and \(\langle k_j \rangle \lesssim M\) for all \(j \in A\), where \(M\) is a dyadic number, and that in the support of \(h_{bck_A}\) there is no pairing in \(k_A\). Define the tensor

\[ H_{bc} = \sum_{k_A} h_{bck_A} \prod_{j \in A} \eta_{k_j}^j, \tag{2.30} \]

where we restrict \(k_j \in E\) in (2.30), \(E\) being a finite set such that \(\{h_{bck_A}\}\) is independent with \(\{\eta_k : k \in E\}\). Then \(\tau^{-1} M\)-certainly, we have

\[ \|H_{bc}\|_{b \to c} \lesssim \tau^{-\theta} M^\theta \cdot \max_{(B, C)} \|h\|_{b_{B'} \to c_{B'}}, \tag{2.31} \]

where \((B, C)\) runs over all partitions of \(A\). The same results holds if we do not assume \(\langle b \rangle, \langle c \rangle \lesssim M\), but instead that (i) \(b, c \in \mathbb{Z}^3\) and \(|b - c| \lesssim M\) and \(|b|^2 - |c|^2| \lesssim M^3\), and (ii) \(h_{bck_A}\) can be written as a function of \(b - c\), \(|b|^2 - |c|^2\) and \(k_A\).

For the proof of Proposition 2.7 see 19, Propositions 4.14 and 4.15.

**Proposition 2.8** (Weighted bounds). Suppose the matrices \(h = h_{kk''}\), \(h^{(1)} = h^{(1)}_{kk'}\) and \(h^{(2)} = h^{(2)}_{kk''}\) satisfy that

\[ h_{kk''} = \sum_{k'} h^{(1)}_{kk'} h^{(2)}_{k'k''}, \]

and \(h^{(1)}_{kk}\) is supported in \(|k - k'| \lesssim L\), then we have

\[ \left\| \left(1 + \frac{|k - k''|}{L} \right)^\kappa h_{kk''} \right\|_{\ell^2_{kk''}} \lesssim \|h^{(1)}\|_{k \to k'} \left\| \left(1 + \frac{|k' - k''|}{L} \right)^\kappa h^{(2)}_{k'k''} \right\|_{\ell^2_{kk''}}. \]

For the proof of Proposition 2.8 see 18, Proposition 2.5 or 19, Lemma 4.3 (there are different versions of this bound, but the proofs are the same).
3. The ansatz

3.1. The structure of \( y_N \). Start with the system \((2.3)-(2.4)\). Let \( y_N = v_N - v_{N/2} \), then \( y_N \) satisfies the integral equation

\[
(y_N)_k(t) = (F_N)_k - i \sum_{\max(N_1,N_2,N_3) = N} \int_0^t \Pi_N \mathcal{M}^\delta(y_{N_1}, y_{N_2}, y_{N_3})_k(s) \, ds
\]

\[
+ i \int_0^t \frac{(y_N)_k(s)}{(k)^2} \, ds - i \sum_{\max(N_1,N_2,N_3) \leq N/2} \int_0^t \Delta_N \mathcal{M}^\delta(y_{N_1}, y_{N_2}, y_{N_3})_k(s) \, ds
\]

\[
- i \int_0^t \left[ (v_N)_k(s) \sum_{\ell \neq k \ell \leq N} V_{k-\ell} \left( |(v_N)_\ell(s)|^2 - \frac{1}{\ell^2} \right) \right. \\
- \left. (v_{N/2})_k(s) \sum_{\ell \neq k \ell \leq N/2} V_{k-\ell} \left( |(v_{N/2})_\ell(s)|^2 - \frac{1}{\ell^2} \right) \right] \, ds.
\]  

(3.1)

3.1.1. The term \( \psi^{N,L} \). For any \( L \leq N/2 \), consider the linear equation for \( \Psi = \Psi_k(t) \):

\[
\partial_t \Psi_k(t) = -i \Delta_N \mathcal{M}^\delta(v_L, v_L, \Psi)_k(t),
\]

(3.2)

where we define, with \( \delta \ll 1 \),

\[
\mathcal{M}^\delta(u, v, w)_k(t) := \sum_{k_1 - k_2 + k_3 = k \atop k_2 \notin \{k_1, k_3\}} e^{it\Omega} \cdot \eta \left( \frac{k_1 - k_2}{N^{1-\delta}} \right) V_{k_1 - k_2} \cdot u_{k_1}(t) \overline{v_{k_2}(t)} w_{k_3}(t);
\]

(3.3)

define also \( \mathcal{M}^> := \mathcal{M}^\delta - \mathcal{M}^\leq \). If \((3.2)\) has initial data \( \Psi_k(0) = \Delta_N \phi_k \), then the solution may be expressed as

\[
\Psi_k(t) = \sum_{k'} H^{N,L}_{kk'}(t) \phi_{k'}.
\]

(3.4)

where \( H^{N,L} = H^{N,L}_{kk'} \) is the kernel of a linear operator (or a matrix). Define also

\[
(\psi^{N,L})_k(t) = \sum_{k'} H^{N,L}_{kk'}(t)(F_N)_{k'},
\]

(3.5)

and similarly

\[
h^{N,L} := H^{N,L} - H^{N,L/2}, \quad \zeta^{N,L} := \psi^{N,L} - \psi^{N,L/2};
\]

(3.6)

note that when \( L = 1 \) we will replace \( L/2 \) by 0, so for example \((\psi^{N,0})_k(t) = (F_N)_k \). For simplicity denote

\[
H^N := H^{N,N/2} \quad \text{and} \quad \psi^N := \psi^{N,N/2}.
\]

(3.7)

Note that each \( h^{N,L} \) and \( H^{N,L} \) is a Borel function of \((g_k(\omega))_{(k) \leq N/2}\), and is thus independent from the Gaussians in \( F_N \).

3.1.2. The terms \( \xi^N \) and \( \rho^N \). Next, similar to \((3.2)\), we consider the linear equation

\[
\partial_t \Xi_k(t) = -i \Delta_N \left[ \mathcal{M}^\leq(v_{N/2}, v_{N/2}, \xi^N) + \mathcal{M}^\leq(v_N, v_N, \xi^N) - \mathcal{M}^\leq(v_{N/2}, v_{N/2}, \Xi) \right]_k(t),
\]

(3.8)

where \( \mathcal{M}^\leq \) is defined by

\[
\mathcal{M}^\leq(u, v, w)_k(s) := \sum_{k_1 - k_2 + k_3 = k \atop k_2 \notin \{k_1, k_3\}} e^{is\Omega} \cdot \eta \left( \frac{k_1 - k_2}{N^\varepsilon} \right) V_{k_1 - k_2} \cdot u_{k_1}(s) \overline{v_{k_2}(s)} w_{k_3}(s).
\]

(3.9)
If the initial data is $\Xi_k(0) = \Delta_N \phi_k$, then we may write the solution as

$$\Xi_k(t) = \sum_{k'} M_{kk'}^N(t) \phi_{k'},$$

(3.10)

which defines the matrix $M^N = M_{kk'}^N$. We then define $\xi^N$ and $\rho^N$ by

$$(\xi^N)_k(t) := \sum_{k'} M_{kk'}^N(t) (F_N)_{k'}, \quad \rho^N := \xi^N - \psi^N.$$  

(3.11)

3.1.3. The ansatz. Now we introduce the ansatz

$$(y_N)_k(t) = (\xi^N)_k(t) + (z_N)_k(t).$$

(3.12)

where $z_N$ is a remainder term. We can calculate that $z_N$ solves the equation (recall $y_N = v_N - v_{N/2}$),

$$(z_N)_k(t) = - i \sum_{\text{max}(N_1,N_2,N_3) = N} \int_0^t \Pi_N M^>(y_{N_1}, y_{N_2}, y_{N_3})_k(s) \, ds$$

$$+ i \int_0^t \frac{(y_N)_k(s)}{\langle k \rangle^2} \, ds - i \sum_{\text{max}(N_1,N_2,N_3) \leq N/2} \int_0^t \Delta_N M^0(y_{N_1}, y_{N_2}, y_{N_3})_k(s) \, ds$$

$$- i \sum_{\text{max}(N_1,N_2) = N; N_3 \leq N} \int_0^t \Pi_N (M^< - M^<)(y_{N_1}, y_{N_2}, y_{N_3})_k(s) \, ds$$

$$- i \int_0^t \Pi_{N/2} M^< (v_{N/2}, v_{N/2}, y_N)_k(s) \, ds - i \int_0^t \Delta_N M^< (v_{N/2}, v_{N/2}, z_N)_k(s) \, ds$$

$$- i \sum_{\text{max}(N_1,N_2) = N} \int_0^t \Pi_{N/2} M^< (y_{N_1}, y_{N_2}, y_N)_k(s) \, ds - i \sum_{\text{max}(N_1,N_2) = N} \Delta_N M^< (y_{N_1}, y_{N_2}, z_N)_k(s) \, ds$$

$$- i \int_0^t \left[ (v_N)_k(s) \sum_{\ell \neq k, \langle \ell \rangle \leq N} V_{k-\ell} \left( |(v_N)_\ell(s)|^2 - \frac{1}{\langle \ell \rangle^2} \right) 

- (v_{N/2})_k(s) \sum_{\ell \neq k, \langle \ell \rangle \leq N/2} V_{k-\ell} \left( |(v_{N/2})_\ell(s)|^2 - \frac{1}{\langle \ell \rangle^2} \right) \right] \, ds.$$  

(3.13)

3.2. Unitarity of matrices $H^{N,L}$ and $M^N$. The following properties of $H$ and $M$ will play a fundamental role. This idea goes back to Bourgain [4]. Recall that for $L \leq N/2$ the matrix $H^{N,L}$ is defined by (3.2) and (3.3). Note that if $\Psi$ solves (3.3), then $\Psi_k(t)$ is supported in $N/2 < \langle k \rangle \leq N$, and we have

$$\partial_t \sum_k |\Psi_k(t)|^2 = 2 \cdot \text{Im} \sum_k \Psi_k(t) \cdot \sum_{k_1-k_2+k_3 = k \atop k_2 \notin \{k_1,k_3\}} e^{it(k_1^2-k_2^2+k_3^2-k^2)}$$

$$\times \eta \left( \frac{k_1-k_2}{N^{1-\delta}} \right) V_{k_1-k_2} \cdot (v_L)_{k_1}(t) \overline{(v_L)_{k_2}(t)} \Psi_{k_3}(t).$$

(3.14)

The sum on the right hand side may be replaced by two terms, namely $S_1$ where we only require $k_1 \neq k_2$ in the summation, $S_2$ where we require $k_1 \neq k_2$ and $k_2 = k_3$ in the summation. For $S_1$ by swapping $(k,k_1,k_2,k_3) \mapsto (k_3, k_2, k_1, k)$ we also see that $S_1 \in \mathbb{R}$ and hence $\text{Im}(S_1) = 0$; moreover

$$S_2 = \sum_{k \neq k_2} \eta \left( \frac{k-k_2}{N^{1-\delta}} \right) V_{k-k_2} \overline{\Psi_k(t)} (v_L)_k(t) \cdot \Psi_{k_2}(t) (v_L)_{k_2}(t).$$
which is also real valued by swapping \((k, k_2) \mapsto (k_2, k)\). This means that \(\sum_k |\Psi_k(t)|^2\) is conserved in time. Therefore for each fixed \(t\), the matrix \(H^{N,L}_k = H_{kk'}^{N,L}\) is unitary, hence we get the identity \[\sum_{k'} H_{kk'}^{N,L} \cdot \overline{H_{kk'}^{N,L}} = \delta_{k,k'}\] (3.15)

with \(\delta_{k_1,k_2}\) being the Kronecker delta. This in particular holds for \(L = N/2\). In the same way, the matrix \(M^N\) defined by (3.8) and (3.10) also satisfies (3.15).

3.3. The a priori estimates. We now state the main a priori estimate, and prove that this implies Theorem [1.3]

**Proposition 3.1.** Given \(0 < \tau \ll 1\), and let \(J = [-\tau, \tau]\). Recall the parameters defined in Section. For any \(M\), consider the following statements, which we call Local\((M)\):

1. For the operators \(h^{N,L}\), where \(L < M\) and \(N > L\) is arbitrary, we have

\[
\|h^{N,L}\|_{Y^{1-b}(J)} + \sup_{t \in J} \|h^{N,L}(t)\|_{\ell^2 \to \ell^2} \leq L^{-1/2+3\varepsilon_1}, \quad \|h^{N,L}\|_{Z^b(J)} \leq N^{1+\delta}L^{-1/2+2\varepsilon_1},
\]

as well as

\[
\left\| \left(1 + \frac{|k-k'|}{\min(L,N^{1-\theta})} \right)^\kappa h^{N,L}_{kk'} \right\|_{Z^b(J)} \leq N^{3/2}.
\]

(2) For the terms \(\rho^N\) and \(z_N\), where \(N \geq M\), we have

\[
\|\rho^N\|_{X^b(J)} \leq N^{-1/2+\varepsilon_1+\varepsilon_2}, \quad \|z_N\|_{X^b(J)} \leq N^{-1/2+\varepsilon_1}.
\]

(3) For any \(L_1, L_2 < M\), the operator defined by

\[
(\mathcal{L} z)_k(t) = -i \int_0^t \Delta_N \mathcal{M}(y_{L_1}, y_{L_2}, z)k(t') dt'
\]

has an extension, which we still denote by \(\mathcal{L}\) for simplicity. The kernel \(\mathcal{L}_{kk'}(t,t')\) has Fourier transform \(\hat{\mathcal{L}}_{kk'}(\lambda, \lambda')\), which satisfies

\[
\int_{\mathbb{R}^2} \langle \lambda \rangle^{2(1-b)} \langle \lambda' \rangle^{-2b} \|\hat{\mathcal{L}}\|_{kk'}^2 d\lambda d\lambda' \leq L^{-1+6\varepsilon_1-2\varepsilon_2}
\]

and

\[
\int_{\mathbb{R}^2} \langle \lambda \rangle^{2b} \langle \lambda' \rangle^{-2(1-b)} \|\hat{\mathcal{L}}\|_{kk'}^2 d\lambda d\lambda' \leq L^{2+2\delta} L^{-1+4\varepsilon_1-2\varepsilon_2},
\]

where \(L = \max(L_1, L_2)\).

Now, with the above definition, we have that

\[
\mathbb{P}(\text{Local}(M/2) \land \neg(\text{Local}(M))) \leq C_{\theta} e^{-\tau^{-1}M}\theta
\]

**Proof of Theorem 1.3** By Proposition 3.1 in particular we know that \(\tau^{-1}\)-certainly, the event Local\((M)\) happens for any \(M\). By (3.4), (3.11) and (3.12) we have

\[
y_N = F_N + \sum_{L \leq N/2} \zeta^{N,L} + \rho^N + z_N.
\]
Exploiting independence between $h^{N,L}$ and $F_N$ and using Proposition 2.7 combined with (3.16), we can show that $\|\zeta^{N,L}\|_{C^2_t H^0_x(J)} \lesssim N^\delta L^{-1/3}$. Summing over $L$ and noticing that $\zeta^{N,L}$ is supported in $N/2 < (k) \leq N$, we see that

$$\left\| \sum_{L \leq N/2} \zeta^{N,L} \right\|_{C^0_t H^0_x(J)} \lesssim N^{-\gamma/2}$$

for any $\gamma > 0$. Using also (3.18) we can see that the sequence $\{v_N - f_N\}$ converges in $C^0_t H^0_x(J)$, hence $\{v_N\}$ converges in $C^0_t H^{-1/2}_x(J)$, and so does the original sequence $\{u_N\}$.

Therefore, the solution $u_N$ to (1.7) converges to a unique limit as $N \to \infty$, up to an exceptional set with probability $\geq 1 - C_9 e^{-\tau \delta}$. This proves the almost-sure local well-posedness of (1.1) with Gibbs measure initial data. Since the truncated Gibbs measure $d\eta_N$ defined by (1.13) is invariant under (1.7), and the truncated Gibbs measures converge strongly to the Gibbs measure $d\nu$ as in Proposition 1.2, we can apply the standard local-to-global argument of Bourgain, where the a priori estimates in Proposition 3.1 allows us to prove the suitable stability bounds needed in the process, in exactly the same way as in [18]. The almost-sure global existence and invariance of Gibbs measure then follows.

3.4. A few remarks and simplifications. From now on we will focus on the proof of Proposition 3.1 and assume that the bounds involved in Local$(M/2)$ are already true. The goal is to recover (3.16)-(3.18), and (3.20)-(3.21) for $M$. Before proceeding, we want to remark on a few simplifications that we would like to make in the proof below. These are either standard, or are the same as in [18, 19], and we will not detail out these arguments in the proof below.

(1) In proving these bounds we will use the standard continuity argument, which involves a smallness factor. Here this factor is provided by the short time $\tau \ll 1$. In particular, we can gain a positive power $\tau^\theta$ by using Proposition 2.2 at the price of changing the $c$ exponent in the $X^c$ (or $Y^c$ or $Z^c$) norm by a little. It can be checked in the proof below that all the estimates allow for some room in $c$, so this is always possible.

(2) In each proof below, we can actually gain an extra power $M^{\delta/10}$ compared to the desired estimate, so any loss which is $M^{C\delta^{-1}}$ will be acceptable. In fact, in the proof below we will frequently encounter losses of at most $M^{C\delta^{-1}}$ due to manipulations of the $c$ exponent in various norms as in (1), and due to application of probabilistic bounds such as Proposition 2.7 were we lose a small $\theta$ power.

(3) In the course of the proof, we will occasionally need to obtain bounds of quantities of form $\sup_\lambda G(\lambda)$, where $\lambda$ ranges in an interval, and for each fixed $\lambda$, the quantity $|G(\lambda)|$ can be bounded, apart from a small exceptional set; moreover, here $G$ will be differentiable and $G'(<\lambda)$ will satisfy a weaker but unconditional bound. Then we can apply the meshing argument in [18, 19], where we divide the interval into a large number of subintervals, approximate $G$ on each small interval by a sample (or an average), control the error term using $G'$, and add up the exceptional sets corresponding to the sample in each interval. In typical cases, where $M$-certainly $|G(\lambda)| \leq M^\theta$ for each fixed $\lambda$, $|I| \leq M^C$ and $|G'(\lambda)| \leq M^C$ unconditionally, we can deduce that $M$-certainly, $\sup_\lambda |G(\lambda)| \leq M^\theta$, because the number of subintervals is $O(M^C)$ so the total probability for the union of exceptional sets is still sufficiently small.

1In the case $c > 1/2$ we also need $u(0) = 0$ in Proposition 2.7 but as we will only estimate the Duhamel terms of form $u = I(N)$ or $u = \mathcal{T}_x N$, see (2.8), we do indeed have $u(0) = 0$. 
4. The random averaging operator

4.1. The operator \( \mathcal{L} \). We start by proving (4.20)–(4.21) for \( L = M/2 \). We need to construct an extension of \( \mathcal{L} \) defined in (3.19). This is done first using Lemma 2.3 to find extensions of each component of \( y_{L_j} \) and \( y_{L_j} \) (note that \( \max(L_1, L_2) = M/2 \)), such that these extension terms satisfy (3.16)–(3.18) with the localized \( X^b(J) \) etc. norms replaced by the global \( X^b \) etc. norms, at the expense of some slightly worse exponents. The change of value in exponents will play no role in the proof below so we will omit it. Then, by attaching to \( \mathcal{L} \) a factor \( \chi(\tau^{-1}t) \) and using Lemma 2.2 (see Section 3.4) we can gain a smallness factor \( \tau \) at the price of further worsening the exponents. These operations are standard so we will not repeat them below.

Note that the extension defined in Lemma 2.3 preserves the independence between the matrices \( h_{L_j, R_j} \) and \( F_{L_j} \) for \( R_j \leq L_j/2 \).

Recall that \( \mathcal{L}_{kk'}(\lambda, \lambda') \) is the Fourier transform of the kernel \( \mathcal{L}_{kk'}(t, t') \) of \( \mathcal{L} \), we have

\[
(\mathcal{L} z)_k(\lambda) = \sum_{k'} \int \mathcal{L}_{kk'}(\lambda, \lambda') \tilde{z}_{k'}(\lambda') \, d\lambda',
\]

(4.1)

Now we consider the different cases.

(1) Suppose in (3.19) we replace \( y_{L_j} \) by \( \rho_{L_j} + z_{L_j} \) for \( j \in \{1, 2\} \), then in particular we may assume that \( \|y_{L_j}\|_{X^b} \lesssim L_j^{-1/2 + \varepsilon_1 + \varepsilon_2} \) due to (3.18). By (3.19) and (4.1), we have

\[
\mathcal{L}_{kk'}(\lambda, \lambda') = \sum_{k_1 - k_2 = k - k'} \int_{\mathbb{R}^2} I(\lambda, \Omega + \lambda_1 - \lambda_2 + \lambda') \cdot V_{k_1 - k_2}(y_{L_1})_{k_1}(\lambda_1) \cdot \overline{(y_{L_2})_{k_2}(\lambda_2)} \, d\lambda_1 d\lambda_2
\]

(4.2)

where \( \Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k'|^2 \) and \( I = I(\lambda, \mu) \) as in (2.10); we will omit the factor \( \eta((k_1 - k_2)/N^{1/\delta}) \) in the definition of \( M^< \) in (3.3) as it does not play a role. We may also assume that \( |k_1 - k_2| \sim R \lesssim L \). In the above expression, let \( \mu := \lambda - (\Omega + \lambda_1 - \lambda_2 + \lambda') \), in particular we have \( |I| \lesssim |\lambda|^{-1} |\mu|^{-1} \) by (2.10). By a routine argument, in proving (3.20) we may assume \( |\lambda_j| \leq L_{100} \) and \( |\mu| \lesssim L_{100} \); in fact, if say \( |\lambda_1| \) is the maximum of these values and \( |\lambda_1| \geq L_{100} \) (the other cases being similar), then we may fix the values of \( k_j \), and hence \( k - k' \), at a loss of at most \( L_{12} \), and reduce to estimating

\[
|\mathcal{L}_{kk'}(\lambda, \lambda')| \lesssim \int_{\mathbb{R}^3} \frac{1}{|\lambda| |\lambda - \lambda_1 + \lambda_2 - \lambda_3 - \Omega|} \w_1(\lambda_1) \w_2(\lambda_2) \, d\lambda_1 d\lambda_2,
\]

with \( |\lambda_1| \sim K \geq L_{100} \) and \( \|\lambda_j \w_j\|_{L^2} \lesssim 1 \) for each \( j \). By estimating \( w_1 \) in the unweighted \( L^2 \) norm we can gain a power \( K^{-1/2} \), and using the \( L_{12}^1 \) integrability of \( \w_2 \) which follows from the weighted \( L^2 \) norm we can fix the value of \( \lambda_2 \). In the end this leads to estimating

\[
\sup_{k,k'} |\mathcal{L}_{kk'}(\lambda, \lambda')| \lesssim |\lambda|^{-1} K^{-1/2}
\]

and hence

\[
\|\langle \lambda \rangle^{1-b} \langle \lambda' \rangle^{-b} \sup_{k,k'} |\mathcal{L}_{kk'}(\lambda, \lambda')|\|_{L^2_{\lambda, \lambda'}} \lesssim K^{-1/3} \lesssim L_{30},
\]

which is more than enough, because \( \|\mathcal{L}\|_{k \to k'} = \sup_{k,k'} |\mathcal{L}_{kk'}| \) if \( \mathcal{L} \) is supported where \( k - k' \) is constant.

Now we may assume \( |\lambda_j| \leq L_{100} \) for \( j \in \{1, 2\} \) and \( |\mu| \leq L_{100} \), we may also assume \( |\lambda| + |\lambda'| \leq L_{100} \) as otherwise we gain from the weights \( \langle \lambda \rangle^{2(1-b)} \) and \( \langle \lambda' \rangle^{-2b} \) in (3.20). Similarly, in proving (3.21) we may assume \( |\lambda_j| \leq N_{100} \) for \( j \in \{1, 2\} \), \( |\mu| \leq N_{100} \), and \( |\lambda| + |\lambda'| \leq N_{100} \) (otherwise we may also fix
$(k, k')$ and argue as above). Therefore, in proving (4.21) we may replace the unfavorable exponents $\langle \lambda \rangle^{2b}(\lambda')^{-2(1-b)}$ by the favorable ones $\langle \lambda \rangle^{2(1-b)}(\lambda')^{-2b}$ at a price of $N^{C\kappa^{-1}}$; this will be acceptable since in the proof we will be able to gain a power $N^{-\delta/2}$. We remark that in the proof below (though not here), we may use the $Y^{1-b}$ norm as in (3.16) for the matrices in the decomposition of $y_L$; using the bounds of $\lambda_j$ as above, we may replace the exponent $1 - b$ by $b$ (which then implies $L_{\lambda_j}$ integrability) again at a loss of either $L^{C\kappa-1}$ or $N^{C\kappa-1}$ depending on whether we are proving (3.20) or (3.21), which is acceptable. See also Section 3.4.

This then allows us to fix the values of $\lambda_j$ in (4.2) using the $L_{\lambda_j}^1$ integrability coming from the weighted norms; moreover, by using the bound $|I| \lesssim \langle \lambda \rangle^{-1} \langle \mu \rangle^{-1}$, upper bounds for $\lambda$ and $\mu$ as above, and the weights in (3.20)–(3.21), we may also fix the values of $\lambda$, $\lambda'$ and $|\mu|$, and reduce to estimating the quantity

$$Q_{kk'} = \sum_{k_1-k_2=k-k'} h_{kk_1k_2k'}^b(w_1_1)(w_1_2),$$

(4.3)

where the tensor (which we call the base tensor) $h^b = h_{kk_1k_2k'}^b = V_{k_1-k_2} \cdot 1_{k_1-k_2+k'=k} \cdot 1_{|k|^2-|k_1|^2+|k_2|^2-|k'|^2=\Omega_0}$ with some value $\Omega_0$ determined by $\lambda_j$, $\lambda$, $\lambda'$ and $|\mu|$. Here we also assume $|k_j| \lesssim L_j$ and $|k_1-k_2| \sim R \lesssim L$, and $|w_j|_{\ell^2} \lesssim L_j^{-1/2+\epsilon_1+\epsilon_2}$.

Now (4.3) is easily estimated by using Proposition 2.6 that

$$||Q||_{kk'} \lesssim ||h^b||_{kk_1k_2k'} \cdot ||w_1||_{k_1} \cdot ||w_2||_{k_2} \lesssim R^{-\beta} \cdot L_1^{-1/2+\epsilon_1+\epsilon_2} L_2^{-1/2+\epsilon_1+\epsilon_2} \lesssim L^{-1/2+2\epsilon_1-\epsilon_2},$$

which is enough for (3.20) (namely we multiply this by the factor $\langle \lambda \rangle^{-1}$ coming from $I$, and the weight $\langle \lambda \rangle^{1-b} \langle \lambda' \rangle^{-b}$ in (3.20), then take the $L^2$ norm in $\lambda$ and $\lambda'$ to get (3.20); the same happens below). For the $||Q||_{kk'}$ norm we have

$$||Q||_{kk'} \lesssim ||h^b||_{kk_1k_2k'} \cdot ||w_1||_{k_1} \cdot ||w_2||_{k_2} \lesssim R^{-\beta} \cdot N R \cdot L_1^{-1/2+\epsilon_1+\epsilon_2} L_2^{-1/2+\epsilon_1+\epsilon_2} \lesssim NL^{-1/2+2\epsilon_1-\epsilon_2},$$

which is enough for (3.21). Note that all the bounds for $h^b$ we use here follow from Lemma 2.4.

(2) Suppose $y_{L}$ is replaced by $y_{L}^2 + z_{L}$, and $y_{L_2}$ is replaced by $y_{L_2}^2$. We may further decompose $y_{L_1}$ into $\zeta_{Lz_2}$ for $R_2 \leq L_2/2$, (including the case $R_2 = 0$ by which we mean $\zeta_{Lz_2} = F_{L_2}$) and perform the same arguments as above fixing the $\lambda$ variables, and reduce to estimating the quantity

$$Q_{kk'} = \sum_{k_1-k_2=k-k'} h_{kk_1k_2k'}^b(w_1_1) \sum_{k_2'} h_{k_2k_2'}^{(2)}(F_{L_2})k_2',$$

(4.4)

where $||w_1||_{\ell^2} \lesssim L_1^{-1/2+\epsilon_1+\epsilon_2}$, $h^b$ is independent from $F_{L_2}$, and is either the identity matrix or satisfies $||h^b||_{k_2'k_2} \lesssim R_2^{-1/2+3\epsilon_1}$ and $||h^b||_{k_2k_2'} \lesssim L_2^{1+\delta} R_2^{-1/2+2\epsilon_1}$. We then estimate (4.4) by

$$||Q||_{kk'} \lesssim L_1^{-1} (||h^b||_{kk_1k_2k'} + ||h^b||_{kk_1k_2k'}) ||w_1||_{k_1} ||h^b||_{k_2k_2'} \lesssim R^{-\beta} \cdot R \min(L_1, L_2) \cdot L_2^{-1} L_1^{-1/2+\epsilon_1+\epsilon_2} \lesssim L^{-1/2+2\epsilon_1-\epsilon_2},$$

(4.5)

using Propositions 2.5 and 2.7 which is enough for (3.20). Note that here $h^b$ depends on $k$ and $k'$ only via $k - k'$ and $|k|^2 - |k'|^2$, and that $||k||^2 - |k'|^2 \leq L_k^3$ given the assumptions, so Proposition

1This reduction step actually involves a meshing argument as the estimate for $Q$ is probabilistic, see Section 3.4.
\(2.7\) is applicable. Similarly for the \(\ell^2_{kk'}\) norm we have

\[
\|Q\|_{kk'} \lesssim L_2^{-1} (\|h^b\|_{kk' \rightarrow k_1k_2} + \|h^b\|_{kk_1k_2' \rightarrow k_2})\|w_1\|_{k_1} \|h^{(2)}\|_{k_2 \rightarrow k_2'} \lesssim R^{-\beta} \cdot N (\min(L_1, L_2) + \min(L_1, R)) \cdot L_2^{-1} L_1^{-1/2+\varepsilon_1+\varepsilon_2} \lesssim NL^{-1/2+2\varepsilon_1-\varepsilon_2}, \tag{4.6}
\]

which is enough for (3.21).

(3) Suppose \(y_{L_j}\) is replaced by \(\psi^{L_j}\) for \(j \in \{1, 2\}\). In this case we will start from (3.19) and expand

\[
(\psi^{L_j})_{k_j} = \sum_{k_j'} (H^{L_j})_{k_jk_j'} (F_{L_j})_{k_j'}
\]

for \(j \in \{1, 2\}\). There are then two cases, namely when \(k_1' = k_2'\) or otherwise.

If \(k_1' \neq k_2'\), then we can repeat the above argument (including further decomposing \(\psi^{L_j}\) into \(\zeta_{L_j, R_j}\) using (3.6) and (3.7)) and fix the time Fourier variables, and reduce to estimating a quantity

\[
Q_{kk'} = \sum_{k_1' = k_2' = k_2} h^b_{k_1k_1'} h^b_{k_1k_2'} h^b_{k_1k_2} \cdot (F_{L_1})_{k_1'} \cdot h^{(2)}_{k_2k_2'} \cdot (F_{L_2})_{k_2'}, \tag{4.7}
\]

where \(h^{(j)}\) is independent from \(F_{L_j}\), and is either the identity matrix or satisfies \(\|h^{(j)}\|_{k_j \rightarrow k_j'} \lesssim R_j^{-1/2+3\varepsilon_1}\) and \(\|h^{(j)}\|_{k_1k_1'} \lesssim L_j^{1-\delta} R_j^{-1/2+2\varepsilon_1}\). Since \(k_1' \neq k_2'\), we can apply Proposition (2.7) either in \((k_1, k_2')\) jointly (if \(L_1 = L_2\)) or first in \(k_1'\) then in \(k_2'\) (if, say, \(L_1 \geq 2L_2\)) and get that

\[
\|Q\|_{k \rightarrow k'} \lesssim (L_1L_2)^{-1} \max (\|h^b\|_{k \rightarrow k_1k_1'}, \|h^b\|_{k_1 \rightarrow k_2k_2'}, \|h^b\|_{k_2 \rightarrow k_1k_1'}, \|h^b\|_{kk_1k_2 \rightarrow k_2}) \times \|h^{(1)}\|_{k_1 \rightarrow k_1'} \|h^{(2)}\|_{k_2 \rightarrow k_2'} \lesssim R^{-\beta} (L_1L_2)^{-1} \cdot R \min(L_1, L_2)^{-2/3}, \tag{4.8}
\]

which is enough for (3.20). As for \(\ell^2_{kk'}\) norm we have

\[
\|Q\|_{kk'} \lesssim (L_1L_2)^{-1} \|h^b\|_{kk_1k_2k_3} \cdot \|h^{(1)}\|_{k_1 \rightarrow k_1'} \|h^{(2)}\|_{k_2 \rightarrow k_2'} \lesssim (L_1L_2)^{-1} R^{-\beta} \cdot \min(L_1, L_2)^{-2/3} NR
\]

which is enough for (3.21).

Finally assume \(k_1' = k_2'\), then \(L_1 = L_2 = L\). In (3.15) we the summation in \(k_1' = k_2'\) gives

\[
\sum_{k_1'} \frac{1}{(k_1')^2} (H^{L_1})_{k_1k_1'} (t') (H^{L_2})_{k_2k_2'} (t').
\]

Using the cancellation (3.15) since \(k_1 \neq k_2\), we can replace the factor \(1/(k_1')^2\) in the above expression by \(1/(k_1')^2 - 1/(k_1)^2\); then by further decomposing \(H^{L_j}\) into \(h^{L_j, R_j}\) by (3.7) and repeating the above arguments, we can reduce to estimating the quantity

\[
Q_{kk'} = \sum_{k_1 - k_2 = k - k'} h^b_{kk_1k_2} \cdot (\bar{h})_{k_1k_2} = \sum_{k_1} \left( \frac{1}{(k_1')^2} - \frac{1}{(k_1)^2} \right) h^{(1)}_{k_1k_1'} h^{(2)}_{k_2k_2'}, \tag{4.9}
\]

where \(h^{(j)}\) is either the identity matrix or satisfies \(\|h^{(j)}\|_{k_j \rightarrow k_j'} \lesssim R_j^{-1/2+3\varepsilon_1}\) and \(\|h^{(j)}\|_{k_jk_j'} \lesssim L_j^{1+\delta} R_j^{-1/2+2\varepsilon_1}\). Note that we may assume \(|k_j - k_j'| \lesssim R_j L^\delta\) using the bound (3.17), so in particular we have

\[
\left| \frac{1}{(k_1')^2} - \frac{1}{(k_1)^2} \right| \lesssim \frac{R + \min(R_1, R_2)}{L^3}.
\]
up to a loss of $L^C\delta$ (which is acceptable as in this case we can gain at least $L^{\varepsilon_2}$). Using these, we estimate, assuming without loss of generality that $R_1 \geq R_2$:

$$\|Q\|_{k \to k'} \lesssim \|h^b\|_{kk \to kk'} \|\tilde{h}\|_{kk} \lesssim \frac{R + R_2}{L^3} \cdot L^{1+\delta} R_1^{-1/2+2\varepsilon_1} R_2^{-1/2+3\varepsilon_1} \cdot R^{-\beta} L \min(R_1, R) \lesssim L^{-1/2+3\varepsilon_1 - \varepsilon_2}, \quad (4.10)$$

This completes the proof for $(3.20)$ and $(3.21)$.

4.2. The matrices $H^{N,L}$ and $h^{N,L}$. We now prove $(3.16)$–$(3.17)$. Let $\mathcal{L}^{N,L}$ be the linear operator defined by

$$z \mapsto -i \int_0^t \Delta_N \mathcal{M}^<(v_L, v_L, z) (t') \, dt',$$

we also extend its kernel in the same way as we do for $\mathcal{L}$ in Section 4.1. Let $\widehat{\mathcal{L}}^{N,L} = \mathcal{L}^{N,L} - \mathcal{L}^{N,L}/2$, then by induction hypothesis and the proof in Section 4.1 we know that $\widehat{\mathcal{L}}^{N,L}$ also satisfies the estimates $(3.20)$–$(3.21)$. Clearly $(3.20)$ implies that $\|\widehat{\mathcal{L}}^{N,L}\|_{X^b \to X^{1-b}} \lesssim L^{-1/2+3\varepsilon_1 - \varepsilon_2}$; moreover it is easy to see that

$$\|\mathcal{L}^{N,L} z\|_{X^1} \lesssim \|\mathcal{M}^<(v_L, v_L, z)\|_{L^2_{z,x}} \lesssim L^{12} \|z\|_{X^0},$$

due to $(3.20)$ and $\varepsilon_2$. Hence $\|\mathcal{L}^{N,L}\|_{X^0 \to X^1} \lesssim L^{12}$ and hence holds for $\widehat{\mathcal{L}}^{N,L}$. By interpolation we obtain that

$$\|\widehat{\mathcal{L}}^{N,L}\|_{X\to X^\alpha} \lesssim L^{-1/2+3\varepsilon_1}$$

for $\alpha \in \{b, 1-b\}$ (note that we can always gain a positive power of $\tau$ using Lemma 2.2, see Section 3.4). Moreover, consider the kernel $(\mathcal{F}\widehat{\mathcal{L}}^{N,L})_{kk'}(\lambda, \lambda')$, then we also have the bound

$$\int_{\mathbb{R}} \langle \lambda \rangle^{2(1-b)} \langle \lambda' \rangle^{-b} (\mathcal{F}\widehat{\mathcal{L}}^{N,L})_{kk'}(\lambda, \lambda') \|_{k' \to k} \, d\lambda \lesssim L^{-1+6\varepsilon_1 - 2\varepsilon_2}$$

which follows from $(3.20)$. If we replace the factor $\langle \lambda' \rangle^{-b}$ by 1, then a simple argument shows that

$$\|\mathcal{F}\mathcal{L}^{N,L}\|_{kk'}(\lambda, \lambda') \|_{k' \to k} \lesssim L^{12} \langle \lambda \rangle^{-1}$$

(and the same for $\widehat{\mathcal{L}}^{N,L}$) by using that

$$\|\mathcal{F}\mathcal{L}^{N,L}_z\|_{k} \lesssim \langle \lambda \rangle^{-1} \int_{\mathbb{R}} \langle \lambda - \mu \rangle^{-1} \|\mathcal{F}\mathcal{M}^<(v_L, v_L, z)\|_{k} \, d\mu \lesssim \langle \lambda \rangle^{-1} \|\mathcal{M}^<(v_L, v_L, z)\|_{L^2}$$

and then fixing the Fourier modes of $v_L$. Interpolating again, we get that

$$\int_{\mathbb{R}} \langle \lambda \rangle^{2(1-b)} \|\langle \lambda' \rangle^{-1-b} (\mathcal{F}\mathcal{L}^{N,L})_{kk'}(\lambda, \lambda') \|_{k' \to k} \, d\lambda \lesssim L^{-1+6\varepsilon_1}. \quad (4.12)$$

A similar interpolation gives

$$\int_{\mathbb{R}} \langle \lambda' \rangle^{-2b} \|\langle \lambda \rangle^b (\mathcal{F}\mathcal{L}^{N,L})_{kk'}(\lambda, \lambda') \|_{k' \to k} \, d\lambda' \lesssim L^{-1+6\varepsilon_1}. \quad (4.13)$$

Clearly $\mathcal{L}^{N,L}$ satisfies $(4.12)$–$(4.13)$ with right hand sides replaced by 1.

Now let

$$\mathcal{H}^{N,L} = (1 - \mathcal{L}^{N,L})^{-1} = \sum_{n=0}^{\infty} (\mathcal{L}^{N,L})^n,$$
it is easy to see that $\mathcal{H}^{N,L} - 1$ satisfies the same bounds (4.12)–(4.13) with right hand sides replaced by 1; for example (4.12) follows from iterating the bound

$$
\| \langle \lambda \rangle^{1-b} \| \langle \lambda'' \rangle^{-(1-b)} \langle \mathcal{A} \hat{B} \rangle_{kk''} (\lambda, \lambda'') \|_{k'' \rightarrow k} \|_{L^2}^2
\lesssim \| \langle \lambda \rangle^{1-b} \| \langle \lambda' \rangle^{-(1-b)} \langle \mathcal{A} \hat{k'} \rangle (\lambda, \lambda'') \|_{k' \rightarrow k} \|_{L^2} \cdot \| \mathcal{B} \|_{X^{1-b} \rightarrow X^{1-b}}
$$

provided

$$
\langle \mathcal{A} \hat{B} \rangle_{kk''} (\lambda, \lambda'') = \sum_{k'} \int_{\mathbb{R}} \mathcal{A}_{kk'} (\lambda, \lambda') \mathcal{B}_{k'k''} (\lambda', \lambda'') \, d\lambda',
$$

and (4.13) is proved similarly. Defining further

$$
\mathcal{H}^{N,L} = \mathcal{H}^{N,L} - \mathcal{H}^{N,L/2} = \sum_{n=1}^{\infty} (-1)^{n-1} (\mathcal{H}^{N,L} \mathcal{F}^{N,L} \mathcal{H}^{N,L})^n \mathcal{H}^{N,L}.
$$

By iterating the $X^\alpha \rightarrow X^\alpha$ bounds and using also (3.21) for $\mathcal{F}^{N,L}$ we can show that

$$
\int_{\mathbb{R}^2} \langle \lambda \rangle^{2b} \langle \lambda' \rangle^{-2(1-b)} \| \langle \mathcal{F} \mathcal{H}^{N,L} \rangle_{kk'} (\lambda, \lambda') \|_{kk'}^2 \, d\lambda d\lambda' \lesssim N^{2+2\delta} L^{-1+4\epsilon_1}.
$$

The weighted bound

$$
\int_{\mathbb{R}^2} \langle \lambda \rangle^{2b} \langle \lambda' \rangle^{-2(1-b)} \left(1 + \frac{|k - k'|}{\min(L, N^{1-\delta})}\right)^{\kappa} \langle \mathcal{F} \mathcal{H}^{N,L} \rangle_{kk'} (\lambda, \lambda') \|_{kk'}^2 \, d\lambda d\lambda' \lesssim N^3
$$

is shown in the same way but using Proposition 2.8.

In addition, we can also show that

$$
\int_{\mathbb{R}^2} \langle \lambda \rangle^{2(1-b)} \langle \lambda' \rangle^{-2b} \| \langle \mathcal{F} \mathcal{H}^{N,L} \rangle_{kk'} (\lambda, \lambda') \|_{kk'}^2 \, d\lambda d\lambda' \lesssim L^{-1+6\epsilon_1}.
$$

This can be proved using (4.12)–(4.13), by iterating the bounds

$$
\| \langle \lambda \rangle^{1-b} \langle \lambda'' \rangle^{-b} \| \langle \mathcal{A} \hat{B} \rangle_{kk'} (\lambda, \lambda'') \|_{k'' \rightarrow k'} \|_{L^2_{k', \lambda'}}
\lesssim \| \langle \lambda \rangle^{1-b} \langle \lambda' \rangle^{-b} \langle \mathcal{A} \hat{k'} \rangle (\lambda, \lambda'') \|_{k' \rightarrow k} \|_{L^2_{k, \lambda}} \cdot \| \langle \lambda'' \rangle^{-b} \| \langle \mathcal{B} \rangle_{k'' \rightarrow k'} (\lambda', \lambda'') \|_{k'' \rightarrow k'} \|_{L^2_{k', \lambda'}}
$$

and similarly

$$
\| \langle \lambda \rangle^{1-b} \langle \lambda'' \rangle^{-b} \| \langle \mathcal{A} \hat{k'} \rangle (\lambda, \lambda'') \|_{k' \rightarrow k''} \|_{L^2_{k, \lambda}}
\lesssim \| \langle \lambda \rangle^{1-b} \langle \lambda' \rangle^{-1-b} \langle \mathcal{A} \hat{k'} \rangle (\lambda, \lambda') \|_{k' \rightarrow k} \|_{L^2_{k, \lambda}} \cdot \| \langle \lambda'' \rangle^{-b} \| \langle \mathcal{B} \rangle_{k'' \rightarrow k'} (\lambda', \lambda'') \|_{k'' \rightarrow k'} \|_{L^2_{k', \lambda'}}
$$

assuming (4.14).

Now we can finally prove (3.16)–(3.17). In fact, by definition of $\mathcal{H}^{N,L}$ and $\mathcal{F}^{N,L}$, there exists an extension of $h^{N,L}$ such that

$$
(h^{N,L})_{kk'} (\lambda) = \int_{\mathbb{R}} \langle \mathcal{F} \mathcal{H}^{N,L} \rangle_{kk'} (\lambda, \lambda') \chi (\lambda') \, d\lambda',
$$

so the $Y^{1-b}$ and $Z^b$ bounds in (3.16), as well as (3.17), can be deduced directly from (4.15)–(4.17). The bound $\sup_t \| h^{N,L}(t) \|_{k \rightarrow k'}$ is also easily controlled by $\| \mathcal{F}^{N,L} \|_{X^b \rightarrow X^b}$ using the embedding $L^\infty_t L^2 \hookrightarrow X^b$. This completes the proof for (3.16)–(3.17).
In this section we prove the first bound in (3.18) regarding $\rho^N$, assuming $N = M$. Recall that from (3.2), (3.4) and (3.8) we deduce that $\rho^N$ satisfies the equation

$$(\rho^N)_k(t) = -i\int_0^t \Delta_N \mathcal{M}^<(v_{N/2}, v_{N/2}, \rho^N)_k(t') dt'$$

$$- i\Delta_N [\mathcal{M}^<(v_N, v_N, \psi^N + \rho^N) - \mathcal{M}^<(v_{N/2}, v_{N/2}, \psi^N + \rho^N)]_k(t') dt'$$

with initial data $(\rho^N)_k(0) = 0$. Let $\mathcal{L}^{N,L}$ be defined as in (4.11), and denote $\mathcal{L}^N := \mathcal{L}^{N,N/2}$ from Section 4.2 we know that $(1 - \mathcal{L}^N)^{-1} := \mathcal{H}^N$ is well-defined, and has kernel $(\mathcal{H}^N)_{kk'}(t, t')$ in physical space and $(\mathcal{F}\mathcal{H}^N)_{kk'}(\lambda, \lambda')$ in Fourier space. Then (5.1) can be reduced to

$$(\rho^N)_k(t) = \sum_{k'} \int_0^t (\mathcal{H}^N)_{kk'}(t, t') W_{k'}(t') dt'$$

where

$$W_k(t) = -i\Delta_N \int_0^t \sum_{w_1, w_2, w_3} \mathcal{M}^<(w_1, w_2, w_3)_k(t') dt'.$$

In order to prove the bound for $\rho^N$ in (3.17), we will apply a continuity argument, namely assuming (5.17) and then improving it with a smallness factor. This can be done as long as we bound

$$\|W\|_{X^b(J)} \leq \tau^\theta N^{-1/2+\varepsilon_1+\varepsilon_2},$$

since from Section 4.2 we know $\mathcal{H}^N$ is bounded from $X^b(J)$ to $X^b(J)$. In fact we will prove (5.4) with an extra gain $N^{-\varepsilon_2/2}$ which will allow us to ignore any possible $N^{C\delta}$ loss in the process. The smallness factor $\tau^\theta$ will be provided by Lemma 2.2 as in Section 3.4 so we will not worry about it below. We divide the right hand side of (5.3) into three terms:

- Term I: when $w_3 = \rho^N$;
- Term II: when $w_3 = \psi^N$ and $z_N \in \{w_1, w_2\}$ for some $N' \geq N/2$;
- Term III: when $w_3 = \psi^N$ and $w_1, w_2 \in \{\psi^N, \rho^N, \psi^N/2, \rho^N/2\}$.

Note that these are the only possibilities, since if (say) $N_1 = N$, $w_1 \in \{\psi^N, \rho^N\}$ and $N_2 \leq N/2$, then we must have $N_2 = N/2$ due to the support condition for $\psi^N$ and $\rho^N$, as well as the restriction $|k_1 - k_2| \lesssim N^\varepsilon$ in $\mathcal{M}^\leq$. Moreover, the estimate of term I follows from the operator norm bound

$$\|I\Delta_N \mathcal{M}^\leq(y_{N_1}, y_{N_2}, z)\|_{X^b(J)} \lesssim \tau^\theta \max(N_1, N_2)^{-1/3} \|z\|_{X^b(J)}$$

which is proved by repeating the arguments in Section 4.1 (the proof that works for $\mathcal{M}^\leq$ certainly also works for $\mathcal{M}^\leq$). In the next two sections we will deal with terms II and III respectively.

5.1. **Term II.** Assume without loss of generality that $w_1 = z_{N'}$. There are then two cases to consider, when $w_2 \in \{\rho^{N_2}, z_{N_2}\}$ or when $w_2 = \psi^{N_2}$.
5.1.1. The case $w_2 \in \{\rho_{N_2}, z_{N_2}\}$. If $w_2 \in \{\rho_{N_2}, z_{N_2}\}$, then we in particular have $\|w_2\|_{X^b(J)} \lesssim N_2^{-1/2+\varepsilon_1+\varepsilon_2}$. By Lemma 2.3, we may fix an extension of $w_1$ and $w_2$ that satisfy the same bounds as they do but with $X^b(J)$ replaced by $X^b$; moreover they satisfy the same measurability conditions as $w_1$ and $w_2$. For simplicity we will still denote them by $w_1$ and $w_2$. The same thing is done for $w_3 = \psi^N$, as well as the corresponding matrices.

Now, by (5.3) and Lemma 2.1, we can find an extension of $\Pi$, which we still denote by $\Pi$ for simplicity, such that

$$\hat{\Pi}_k(\lambda) = \sum_{k_1-k_2+k_3=k} \int_{\mathbb{R}^3} I(\lambda, \Omega + \lambda_1 - \lambda_2 + \lambda_3) \cdot \overline{(\overline{w_1})_{k_1}(\lambda_1)} \cdot \overline{(\overline{w_2})_{k_2}(\lambda_2)} \times \eta(\frac{k_1-k_2}{N^\varepsilon}) V_{k_1-k_2} \sum_{k_3} \overline{(H_N)_{k_3k_3'}}(\lambda_3)(F_N)_{k_3'} d\lambda_1 d\lambda_2 d\lambda_3$$

(5.6)

where $\Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2$ and $I = I(\lambda, \mu)$ is as in (2.10). In the above expression, let $\mu := \lambda - (\Omega + \lambda_1 - \lambda_2 + \lambda_3)$, in particular we have $|I| \lesssim \langle \lambda \rangle^{-1} \langle \mu \rangle^{-1}$ by (2.10). By a routine argument we may assume $|\lambda| \leq N^{100}$ and similarly for $\mu$ and each $\lambda_j$; in fact, if say $|\lambda_1|$ is the maximum of these values and $|\lambda_1| \geq N^{100}$, then we may fix the values of $k$ and all $k_j$ at a loss of at most $N^{12}$, and reduce to estimating (with the value of $\Omega$ fixed)

$$|\hat{\Pi}(\lambda)| \lesssim \int_{\mathbb{R}^3} \frac{1}{|\lambda|} |(\overline{\lambda_1})_{\lambda_1 + \lambda_2 - \lambda_3 - \Omega_j}| \overline{w_1}(\lambda_1) \overline{w_2}(\lambda_2) \overline{w_3}(\lambda_3) d\lambda_1 d\lambda_2 d\lambda_3,$$

with $|\lambda_1| \sim K \geq N^{100}$ and $\|\langle \lambda_j \rangle^{ib} \overline{w_j}\|_{L^2} \lesssim 1$ for each $j$. By estimating $w_1$ in the unweighted $L^2$ norm we can gain a power $K^{-1/2}$, and using the $L^1$ integrability of $\overline{w_j}$ which follows from the weighted $L^2$ norms we can fix the values of $\lambda_j$ for $j \in \{2,3\}$. In the end this leads to

$$|\hat{\Pi}(\lambda)| \lesssim 1_{|\lambda| \leq K} \langle \lambda \rangle^{-1} K^{-1/2}$$

and hence $\|\langle \lambda \rangle^{ib} \hat{\Pi}\|_{L^2} \lesssim K^{-1/3} \lesssim N^{-30}$, which is more than enough for (3.18).

Now, with $|\lambda| \leq N^{100}$ etc., we may apply the bounds (3.16)–(3.18), but for the extensions and global norms, and replace the $Y^{1-b}$ norm (if any) by the $Y^b$ norm at a loss of $N^{C\kappa^{-1}}$ which will be neglected as stated above. Similarly, as $|\lambda| \leq N^{100}$, we also only need to estimate $\Pi$ in the $X^{1-b}$ instead of $X^b$ norm again at a loss of $N^{C\kappa^{-1}}$. Then, using $L^1$ integrability in $\lambda_j$ (together with a meshing argument, see Section 3.3) provided by the weighted bounds (3.16)–(3.18), and the (almost) summability in $\mu$ due to the $\langle \mu \rangle^{-1}$ factor in (2.10), we may fix the values of $\lambda$, $\lambda_j$ ($1 \leq j \leq 3$) and $|\mu|$ (and hence the value of $\Omega \in \mathbb{Z}$) and reduce to estimating the $L^2_k$ norm of the following quantity

$$Q_k := \sum_{k_1-k_2+k_3=k} \eta(\frac{k_1-k_2}{N^\varepsilon}) V_{k_1-k_2} \cdot (\overline{w_1})_{k_1} (\overline{w_2})_{k_2} \cdot \sum_{k_3} H_{k_3k_3'}(F_N)_{k_3'}$$

(5.7)

Here in (5.7) we assume that $|k_1| \leq N_2$, $|k_2| \leq N_2$, $|k_1-k_2| \lesssim N^\varepsilon$ and $N/2 < \langle k_3 \rangle, \langle k_3' \rangle \leq N$, and $\Omega_0 \in \mathbb{Z}$ is fixed, and the inputs satisfy that

$$\|\overline{w_1}\|_{L^2} \lesssim N^{-1/2+\varepsilon_1}, \quad \|\overline{w_2}\|_{L^2} \lesssim N_2^{-1/2+\varepsilon_1+\varepsilon_2}, \quad \|H\|_{k_3 \rightarrow k_3'} \lesssim 1.$$

To estimate $Q$, we may assume $|k_1-k_2| \sim R \lesssim N^\varepsilon$, and define the base tensor

$$h^b_{kk_1k_2k_3} = \eta(\frac{k_1-k_2}{N^\varepsilon}) V_{k_1-k_2} \cdot 1_{k_1-k_2+k_3=k} \cdot 1_{|k|^2-|k_1|^2+|k_2|^2-|k_3|^2=\Omega_0},$$
with also the restrictions on $k_j$ as above. Then we have
\[
Q_k = \sum_{k_1,k_2,k_3,k_3'} h_{k_1k_2k_3}^b \cdot (\bar{w}_1)_{k_1}(\bar{w}_2)_{k_2} \cdot H_{k_3k_3'}(F_N)_{k_3'}
\]
and hence
\[
\|Q\|_{\ell^2} \lesssim N^{-1/2+\varepsilon_1} N_2^{-1/2+\varepsilon_1+\varepsilon_2} \left\| \sum_{k_3,k_3'} h_{k_1k_2k_3}^b H_{k_3k_3'}(F_N)_{k_3'} \right\|_{kk_2\to k_1}.
\]
By Lemma 2.7 and the independence between $H_{k_3k_3'}$ and $(F_N)_{k_3'}$, we get that
\[
\left\| \sum_{k_3,k_3'} h_{k_1k_2k_3}^b H_{k_3k_3'}(F_N)_{k_3'} \right\|_{kk_2\to k_1} \lesssim N^\delta \cdot \max \left[ N^{-1} \left\| \sum_{k_3} h_{k_1k_2k_3}^b H_{k_3k_3'} \right\|_{kk_2\to k_1}, \left\| \sum_{k_3} h_{k_1k_2k_3}^b H_{k_3k_3'} \right\|_{kk_2\to k_1} \right]
\]
\[
\lesssim N^\delta \cdot \|H\|_{k_3\to k_3'} \cdot \max \left( \left\| \sum_{k_3} h_{k_1k_2k_3}^b \right\|_{kk_2\to k_1}, \left\| \sum_{k_3} h_{k_1k_2k_3}^b \right\|_{kk_2\to k_1} \right) \lesssim N^\delta R^{-\beta} \cdot N \cdot \min(N_2,R).
\]
Since also $\|H\|_{k_3\to k_3'} \lesssim 1$, we conclude that
\[
\|Q\|_{\ell^2} \lesssim N^{-1/2+\varepsilon_1} N_2^{-1/2+\varepsilon_1+\varepsilon_2} \cdot N^2 R^{-\beta} \min(N_2,R) \lesssim N^{-1/2+\varepsilon_1+\varepsilon_2}/2,
\]
which is enough for (5.18). This concludes the proof for term II when $w_2 \in \{\rho^{N_2},z_{N_2}\}$. Note that the above argument also works for the case when $w_1 = \rho^N$ and $w_2 = \rho^{N_2}$, because here we must have $N_2 \geq N/2$ due to the support condition of $\rho^N$ and the assumption $|k_1 - k_2| \lesssim N^\varepsilon$, and the above arguments give (in fact better) estimates.

5.1.2. The case $w_2 = \psi_{N_2}$. In this case, by repeating the first part of the arguments in Section 5.1.1, we can reduce to estimating the $\ell^2_k$ norm of the quantity
\[
Q_k := \sum_{k_1,k_2,k_3,k_3'} \eta \left( \frac{k_1 - k_2}{N^\varepsilon} \right) V_{k_1-k_2} \cdot (\bar{w}_1)_{k_1} \cdot \sum_{k_2'} H_{k_2k_2'}(F_N)_{k_2'} \cdot \sum_{k_3'} H_{k_3k_3'}(F_N)_{k_3'}.
\]
Here in (5.9) we assume that $|k_1| \leq N$, $|k_2| \leq N_2$, $|k_1 - k_2| \sim R \lesssim N^\varepsilon$, $N_2/2 < \langle k_2 \rangle$, $\langle k_3 \rangle \leq N_2$, and $N/2 < \langle k_3 \rangle$, $\langle k_3 \rangle \leq N$, and $\Omega_0 \in \mathbb{Z}$ is fixed, and the inputs satisfy that
\[
\|\bar{w}_1\|^2 \lesssim N^{-1/2+\varepsilon_1}, \quad \|H^{(j)}\|_{k_j\to k_j'} \lesssim 1 \quad (j = 2,3).
\]
Moreover, this $H^{(j)}$ is such that either $H^{(j)} = \text{Id}$ or $\|H^{(j)}\|_{k_j \to k_j'} \lesssim N^{1+\delta}$ with $N_3 = N$. The sum in (5.9) can be decomposed into a term where $k_2' \neq k_3'$ and a term where $k_2' = k_3'$.

Case 1: $k_2' \neq k_3'$. Let $h_{k_1k_2k_3}^b$ be defined as above, it suffices to estimate the $\ell^2_{k_1} \to \ell^2_{k_2}$ norm of the tensor
\[
(k,k_1) \mapsto \sum_{k_2,k_3} h_{k_1k_2k_3}^b \sum_{k_2',k_3'} H_{k_2k_2'}^{(2)}(F_N)_{k_2'} \cdot H_{k_3k_3'}^{(3)}(F_N)_{k_3'},
\]
by using the $\ell^2$ norm of $w_1$. If $N_3 = N$, then the tensors $h^{b}, H^{(2)}$ and $H^{(3)}$ are independent from $(F_N)_{k_2'}$ and $(F_N)_{k_3'}$, and $k_2' \neq k_3'$, so we can apply Lemma 2.7 if $N_2 \leq N/2$, then $h^{b}, H^{(2)}$ and $H^{(3)}$ and $(F_N)_{k_2'}$ are all independent from $(F_N)_{k_3'}$, and moreover $h^{b}$ and $H^{(2)}$ are independent from $(F_N)_{k_2'}$, so we can apply Lemma 2.7 iteratively, first for the sum in $(k_3,k_3')$, and then for the sum...
in \((k_2, k_2')\). In either case, by applying Lemma 2.4 and combining it with Lemma 2.6 and estimating \(H^{(\ell)}\) in the \(k_j \to k_j'\) norm, we obtain \(N\)-certainly that the desired \(l_{k_1}^2 \to l_{k_2}^2\) norm of the tensor is bounded by

\[
N^\delta N_2^{-1} N^{-1} \cdot \max \left( \|h^b\|_{kk_2k_3 \to k_1}, \|h^b\|_{kk_2 \to k_1k_3}, \|h^b\|_{kk_3 \to k_1k_2}, \|h^b\|_{k \to k_1k_2k_3} \right).
\]

Using the fact that \(|k_2| \leq N_2\) and \(|k_3| \leq N\) in the support of \(h^b\), and Lemma 2.4 as above, we can show that

\[
\max \left( \|h^b\|_{kk_2k_3 \to k_1}, \|h^b\|_{kk_2 \to k_1k_3}, \|h^b\|_{kk_3 \to k_1k_2}, \|h^b\|_{k \to k_1k_2k_3} \right) \lesssim R^{-\beta} N^\delta \cdot N N_2,
\]

hence

\[
\|Q\|_{\ell^2} \lesssim N^{-1/2+\epsilon_1} \cdot R^{-\beta} N^{2\delta} \lesssim N^{-1/2+\epsilon_1+\epsilon_2/2},
\]

which is enough for (3.18).

**Case 2:** \(k_2' = k_3'\). In this case we must have \(N_2 = N\), and we can reduce (5.9) to the expression

\[
Q_k = \sum_{k_1 - k_2 + k_3 = k} \eta \left( \frac{k_1 - k_2}{N^\epsilon} \right) V_{k_1 - k_2} \cdot (\tilde{w}_1)_{k_1} \cdot (\tilde{H})_{k_2k_3}
\]

(5.11)

where

\[
(\tilde{H})_{k_2k_3} = \sum_{k_2'} \frac{1}{|k_2'|^2} \overline{H}_{k_2k_2'} \cdot H_{k_3k_2'}.
\]

As \(k_2 \neq k_3\) in (5.9) due to the definition of \(\mathcal{M}^\infty\), we know that either \(H^{(2)}\) or \(H^{(3)}\) must not be identity, hence we have \(\|\tilde{H}\|_{\ell^2_{k_2k_3}} \lesssim N^{-1+\delta}\). By (5.11) we then simply estimate

\[
\|Q\|_{\ell^2_k} \lesssim \|\tilde{w}_1\|_{\ell^2} \cdot \|\tilde{H}\|_{\ell^2_{k_2k_3}} \cdot \|h^b\|_{kk_2k_3 \to k_1} \lesssim N^{-1/2+\epsilon_1} \cdot N^{-1+\delta} \cdot R^{-\beta} \cdot NR \lesssim N^{-1/2+\epsilon_1+\epsilon_2/2}
\]

(5.12)

using Lemma 2.4 noticing that \(|k_1 - k_2| \lesssim R\) and \(|k_3| \leq N\). This completes the proof for term II.

5.2. **Term III.** Here we assume \(w_3 = \psi^N\) and \(w_1, w_2 \in \{\psi^N, \rho^N, \psi^{N/2}, \rho^{N/2}\}\). We consider two possibilities, when \(w_1, w_2 \in \{\psi^N, \psi^{N/2}\}\), which we call term IV, and when \(w_j \in \{\rho^N, \rho^{N/2}\}\) for some \(j \in \{1, 2\}\), which we call term V.

5.2.1. **Term IV.** Suppose \(w_1, w_2 \in \{\psi^N, \psi^{N/2}\}\). We may also decompose them into \(\psi^{N_j, L_j}\) for \(L_j \leq N_j/2\), and reduce to

\[
IV_k(t) = -i \Delta_N \int_0^t \sum_{k_1 - k_2 + k_3 = k} \epsilon^{it} \Omega \eta \left( \frac{k_1 - k_2}{N^\epsilon} \right) V_{k_1 - k_2} \times \sum_{k_1', k_2', k_3'} (h^{N_1, L_1}_{k_1'k_1'}(t')(h^{N_2, L_2}_{k_2'k_2'}(t')(h^{N_3, L_3}_{k_3'k_3'}(t')(F_{N_1})_{k_1'}(F_{N_2})_{k_2'}(F_{N_3})_{k_3'} dt',
\]

(5.13)

where \(N_1, N_2 \in \{N, N/2\}\) and \(N_3 = N\). In (5.13) we consider two cases, depending on whether there is a pairing \(k_1' = k_2'\) or \(k_2' = k_3'\), or not.

---

1Here we are simplifying by replacing \(|g_k|^2\) by 1 (we will do the same below). This is because \(\mathbb{E}(|g_k|^2 - 1) = 0\), so any large deviation estimate satisfied by linear combinations of \(g_k\), which is the only thing we rely on, will hold also for linear combinations of \(|g_k|^2 - 1\), so the contribution of \(|g_k|^2 - 1\) can always be treated in the same way as the \(k_2' \neq k_3'\) case.
Case 1: no pairing. Assume that \( k_2' \not\in \{k_1', k_3'\} \), then we take the Fourier transform in the time variable \( t \), and repeat the first part of the arguments in Section 5.1.1 to reduce to estimating the \( \ell^2_k \) norm of the quantity

\[
Q_k := \sum_{k_1 - k_2 + k_3 = k} \eta \left( \frac{k_1 - k_2}{N^\varepsilon} \right) V_{k_1 - k_2} \times \sum_{k_1'} h^{(1)}_{k_1' k_1} (F_{N_1})_{k_1'} \sum_{k_2'} h^{(2)}_{k_2' k_2} (F_{N_2})_{k_2'} \sum_{k_3'} h^{(3)}_{k_3' k_3} (F_{N_3})_{k_3'}
\]

In (5.14) we assume that \( |k_j| \sim N \) and \( |k_1 - k_2| \sim R \lesssim N^\varepsilon \), and that the matrices \( h^{(j)} \) is either identity or satisfies that

\[
\|h^{(j)}\|_{k_j \to k'_j} \lesssim L^{-1/2+3\varepsilon_1}_j, \quad \|h^{(j)}\|_{k_j \to k'_j} \lesssim N^{1+\delta} L^{-1/2+2\varepsilon_1}_j,
\]

and moreover we may assume \( h^{(j)} \) is supported in \( |k_j - k'_j| \lesssim L_j N^{\delta} \) by inserting a cutoff exploiting (3.17). The \( \ell^2_k \) norm for \( Q_k \) can be estimated using Proposition 2.7 in the same way as in Section 5.1.2 either jointly in \( (k_1', k_2', k_3') \) if each \( N_j = N \) or first in those \( k_j \) with \( N_j = N \) and then in those \( k_j \) with \( N_j = N/2 \), so that \( N \)-certainly we have (with the base tensor \( h^b \) defined as in Sections 5.1.1 and 5.1.2 above)

\[
\|Q\|_{\ell^2} \lesssim N^\delta \cdot N^{-3} \|h^b\|_{kk_1 k_2 k_3} \prod_{j=1}^3 \|h^{(j)}\|_{k_j \to k'_j} \lesssim N^{-3+\delta} \cdot N^{3/2} \cdot R \cdot R^{-\beta} \lesssim N^{-1/2+\varepsilon_1/2}
\]

using Lemma 2.4, which is enough for (3.18).

Case 2: pairing. We now consider the cases when \( k_1' = k_2' \) or \( k_2' = k_3' \). First, if \( k_2' = k_3' \), then we can apply the reduction arguments as above and reduce to

\[
Q_k := \sum_{k_1 - k_2 + k_3 = k} \eta \left( \frac{k_1 - k_2}{N^\varepsilon} \right) V_{k_1 - k_2} \sum_{k_1'} h^{(1)}_{k_1' k_1} (F_{N_1})_{k_1'} \tilde{h}_{k_2 k_3}
\]

where

\[
(h)_{k_2 k_3} = \sum_{k_2'} \frac{1}{(k_2')^2} h^{(2)}_{k_2' k_2 \cdot h^{(3)}_{k_3 k_3'}} \quad \|\tilde{h}\|_{k_2 k_3} \lesssim N^{-2} \min(\|h^{(2)}\|_{k_2 \to k_2'}, \|h^{(3)}\|_{k_3 k_3'}, \|h^{(2)}\|_{k_2 k_2'}, \|h^{(3)}\|_{k_3 \to k_3'}). 
\]

Note that \( h^{(2)} \) and \( h^{(3)} \) cannot both be identity as \( k_2 \neq k_3 \). Now if max(\( L_2, L_3 \)) \( \leq N/2 \) then due to independence, applying similar arguments as before we can estimate \( N \)-certainly that

\[
\|Q\|_{\ell^2} \lesssim N^\delta N^{-1} \cdot \|h^{(1)}\|_{k_1 \to k_1'} \|h^b\|_{kk_1 k_2 k_3} \|\tilde{h}\|_{k_2 k_3} \lesssim N^{-2+2\varepsilon+4\delta},
\]

using the constraint \( |k_1 - k_2| \lesssim N^\varepsilon \), which is enough for (3.18); if max(\( L_2, L_3 \)) \( = N \) then we can gain a negative power of this value and view \( F_{N_1} \) simply as an \( H^{-1/2-} \) function (without considering randomness) and bound

\[
\|Q\|_{\ell^2} \lesssim N^\delta N^{1/2+\delta} \cdot \|h^{(1)}\|_{k_1 \to k_1'} \|h^b\|_{kk_1 k_2 k_3} \|\tilde{h}\|_{k_2 k_3} \lesssim N^{-1+2\varepsilon+4\delta},
\]

which is also enough for (3.18).
Finally consider the case $k_1' = k_2'$, so in particular $N_1 = N_2$. We will sum over $L_1$ and $L_2$ in order to exploit the cancellation \[(3.15)\] (as $k_1 \neq k_2$); this leads to the expression

$$
\sum_{k_1'} \frac{1}{(k_1')^2} (H^{N_1})_{k_1 k_1'} (t') (H^{N_2})_{k_2 k_1'} (t)
$$

where again we have replaced $|g_{k_1'}|^2$ by 1 as before. Since $k_1 \neq k_2$, by \[(3.15)\], we may replace $(k_1')^{-2}$ in the above expression by $(k_1')^{-2} - (k_1)^{-2}$. Then, decomposing in $L_1$ and $L_2$ again and taking Fourier transform in $t$ and repeating the reduction steps as before, we arrive at the quantity

$$
Q_k := \sum_{k_1 - k_2 + k_3 = k} \eta \left( \frac{k_1 - k_2}{N^\varepsilon} \right) V_{k_1 - k_2} \cdot \tilde{h}_{k_1 k_2} \cdot \sum_{k_3'} h_{k_3 k_3'}^{(3)} (F_N)_{k_3'}
$$

where

$$
\tilde{h}_{k_1 k_2} = \sum_{k_1'} \left( \frac{1}{(k_1')^2} - \frac{1}{(k_1)^2} \right) h_{k_1 k_1'}^{(1)} h_{k_1 k_2} \overline{h_{k_1 k_1'}}
$$

with $h^{(j)}$ as above. Note that may assume $|k_1 - k_2| \lesssim N^\delta \min(L_1, N^\varepsilon + L_2) \lesssim N^{\varepsilon+\delta} \min(L_1, L_2)$ in view of $|k_1 - k_2| \lesssim N^\varepsilon$, it is easy to show, assuming min($L_1, L_2$) = $L$, that

$$
||h||_{k_1 k_2} \lesssim N^{\varepsilon+\delta} \cdot L N^{-3} ||h^{(1)}||_{k_1 k_1'} ||h^{(2)}||_{k_2 \rightarrow k_2'} \lesssim N^{-2} N^{\varepsilon+\delta+4\varepsilon_1}.
$$

Since max($L_1, L_2$) $\leq N/2$, using independence and arguing as before, we can estimate that $N$-certainly,

$$
||Q||_{\ell^2} \lesssim N^{\delta} \cdot N^{-1} ||h^{(3)}||_{k_3 \rightarrow k_3'} \cdot ||h^{b}||_{k_3 \rightarrow k_1 k_2} \cdot ||\tilde{h}||_{k_1 k_2} \lesssim N^{-1+\varepsilon+2\delta+4\varepsilon_1}
$$

which is enough for \[(3.18)\]. This completes the estimate for term IV.

5.2.2. Properties of the matrix $M^N - H^N$. Before studying term V, we first establish some properties of the matrix $Q^N := M^N - H^N = (Q^N)_{k}k'(t)$ such that

$$
(\rho^N)_k(t) = \sum_{k'} (Q^N)_{kk'}(t)(F_N)_{k'}.
$$

\[\textbf{Lemma 5.1.}\] Let $\varepsilon' := \sqrt{\varepsilon}$ so that ($\varepsilon_1 \ll \varepsilon \ll \varepsilon' \ll 1$). Then we have

$$
||Q^N||_{Y^1-b(J)} + \sup_{t \in J} ||(Q^N)_{kk'}(t)||_{k \rightarrow k'} \lesssim N^{-1/2+3\varepsilon_1}, \quad ||Q^N||_{Z^b(J)} \lesssim N^{1/2+2\varepsilon_1}.
$$

Moreover we can decompose $Q^N = Q^{N,\ll} + Q^{N,\text{rem}}$ such that $||Q^{N,\text{rem}}||_{Z^b(J)} \lesssim N^{1/2+2\varepsilon_1-\varepsilon'/4}$, and that

$$
||\rho^{N,\text{rem}}||_{X^b(J)} \lesssim N^{-1/2+2\varepsilon_1-\varepsilon'/4}, \quad \text{where } (\rho^{N,\text{rem}})_k(t) = \sum_{k'} (Q^{N,\text{rem}})_{kk'}(t)(F_N)_{k'}.
$$

Moreover $Q^{N,\ll}$ can be decomposed into at most $N^{C\varepsilon'}$ terms. For each term $Q$ there exist vectors $\ell^*, m^*$ such that $|\ell^*|, |m^*| \lesssim N^{\varepsilon'}$, and that $(Q)_{kk'}(\lambda)$ is a linear combination (in the form of some integrals), with summable coefficients, of expressions of form

$$
1_{k' = k'\ell^*} \cdot \mathcal{Y}_{\ell^*, m^*}(k, \lambda) \cdot \mathcal{R}_{\ell^*, m^*}(k)
$$

\[\text{1Strictly speaking this means } (Q)_{kk'}(\lambda) = \int a(\mu) 1_{k' = k'\ell^*} \cdot \mathcal{Y}_{\ell^*, m^*}(k, \lambda, \mu) \cdot \mathcal{R}_{\ell^*, m^*}(k, \mu) \text{ where the integration is taken over some Euclidean space, } a(\mu) \in L^1, \text{ and the bounds for } \mathcal{R} \text{ and } \mathcal{Y} \text{ are uniform in } \mu.\]
where $\mathcal{Y}$ is independent with $(F_N)_k$, and $|\mathcal{Y}| \lesssim 1$ and $\mathcal{R}(k)$ depends only on $m^* \cdot k$, moreover we have
\[ \|\mathcal{R}\|_{L_k^2}^2 \lesssim N^{1/2 + 2\varepsilon_1 + C_{\varepsilon'}} , \quad \|\langle \lambda \rangle^b \mathcal{Y}\|_{L_k^2}^2 \lesssim N^{C_{\varepsilon'}} . \tag{5.21} \]

Remark 5.2. Lemma 5.1 plays an important role in Section 5.2.3 when estimating Term V. In particular, we will exploit the one-dimensional extra independence of $\mathcal{R}(k)$ with $(F_N)_k$, since $\mathcal{R}(k)$ depends only on $m^* \cdot k$ instead of on $k$.

Proof. By definition of $\xi^N$ and $\psi^N$ in (5.11) and (5.1)–(5.3), as well as the associated matrices, we have the identity
\[ (Q^N)_{kk'}(t) = \sum_{k_1} \int_0^t (\mathcal{H}^N \mathcal{M})_{kk_1}(t,t_1)(H^N - Q^N)_{k_1,k'}(t_1) \, dt_1 \]
and hence we have
\[ (Q^N)_{kk'}(t) = \sum_{n=1}^\infty (-1)^{n-1} \sum_{k_1} \int_0^t \{ (\mathcal{H}^N \mathcal{M})^n \}_{kk_1}(t,t_1)(H^N)_{k_1,k'}(t_1) \, dt_1, \tag{5.22} \]
where $\mathcal{H}^N = \mathcal{H}^{N,N/2}$ is defined in Section 1.2 and $\mathcal{M}$ denotes the operator
\[ z \mapsto \sum_{\max(N_1,N_2)=N} \Delta_{N,N}(y_{N_1},y_{N_2},z). \tag{5.23} \]

The bounds in (5.18) then follow from iterating like in Section 3.2 using the bounds (4.12)–(4.17) (together with the $X^\alpha \to X^\alpha$ bounds) for the operators $\mathcal{H}^N$ and $\mathcal{M}$, where the bounds for $\mathcal{M}$ is proved in the same way as in Sections 4.1 and 4.2. Moreover, in (5.22) if we assume $n \geq 2$ or replace $\mathcal{H}^N$ by $\mathcal{H}^N - \mathcal{H}^{N,N_{\varepsilon'}}$ (or $H^N$ by $H^N - H^{N,N_{\varepsilon'}}$) then the corresponding bounds can be improved by $N^{-\varepsilon'/4}$, and the resulting terms can be put in $Q^{\text{rem}}$. As for the remaining contribution, we can write
\[ (Q^N)_{kk'}(t) = \sum_{k_1,k_2} \int (\mathcal{H}^N_{kk_1}(t,t_1) \mathcal{M}_{k_1,k_2}(t_1,t_2) \mathcal{F}^{N,N_{\varepsilon'}}_{k_2,k'}(t_2) \, dt_1 dt_2, \]

hence
\[ (\hat{Q}^{N,N_{\varepsilon'}})_{kk'}(\lambda) = \sum_{k_1,k_2} \int (\mathcal{F}^{N,N_{\varepsilon'}}_{kk_1}(\lambda,\lambda_1) \mathcal{F} \mathcal{M}_{k_1,k_2}(\lambda_1,\lambda_2) \mathcal{F}^{N,N_{\varepsilon'}}_{k_2,k'}(\lambda_2) \, d\lambda_1 d\lambda_2. \tag{5.24} \]

We may assume $|k-k_1| \lesssim N^{\varepsilon'}$ and the same for $k_1-k_2$, so at a loss of $N^{C_{\varepsilon'}}$ we may fix the values of $k-k_1$, $k_1-k_2$ and $k_2-k'$. Note that the matrices $\mathcal{F}^{N,N_{\varepsilon'}}$ and $\mathcal{F} \mathcal{M}$ satisfy the bounds (4.15)–(4.17); moreover in (4.17) we may replace the unfavourable exponents $\langle \lambda \rangle^{2(1-b)} \langle \lambda \rangle^{-2b}$ by the favorable ones $\langle \lambda \rangle^{2b} \langle \lambda \rangle^{-2(1-b)}$, at the price of replacing the right hand side by a small positive power $N^{C_{\varepsilon'}-1}$, by repeating the interpolation argument in Section 4.2. Using these bounds, we then see that the integral (5.24) provides the required linear combination. Here summability of coefficients follows from the estimate
\[ \int_{\mathbb{R}^2} A(\lambda_1)B(\lambda_1,\lambda_2)C(\lambda_2) \, d\lambda_1 d\lambda_2 \lesssim \|\langle \lambda \rangle^{-(1-b)} A(\lambda_1)\|_{L^2} \cdot \|\langle \lambda_1 \rangle^b \langle \lambda \rangle^{-(1-b)} B(\lambda_1,\lambda_2)\|_{L^2} \cdot \|\langle \lambda_2 \rangle^b C(\lambda_2)\|_{L^2} \tag{5.25} \]

Footnote 1. To prove (5.19), we may repeat the proofs above for terms I–IV, and then treat V in the same way as II. This leads to a loss of $N^{O(\varepsilon_1)}$, which will be negligible compared to the gain $N^{\varepsilon'/4}$.
and the improved versions of (4.15)–(4.17). Recall that $k - k_1$, $k_1 - k_2$ and $k_2 - k'$ are all fixed. We set that $\ell^* := (k_1 - k) + (k_2 - k_1) + (k' - k_2) = k' - k$ and $m^* := k_1 - k_2$. Finally, for fixed $(\lambda_1, \lambda_2)$, we set $\mathcal{Y}(\ell^*, m^*, k, \lambda) := (\mathcal{F} \mathcal{M} \mathcal{N}^c \mathcal{N}^e)_{k_1 k_2} (\lambda_1, \lambda_2) (\mathcal{F} \mathcal{N}^N \mathcal{N}^e)_{k_2 k'} (\lambda_2)$ and $\mathcal{R}(\ell^*, m^*, k) := (\mathcal{F} \mathcal{M})_{k_1 k_2} (\lambda_1, \lambda_2)$. The factors coming from $\mathcal{H}^N \mathcal{N}^e$ and $\mathcal{N}^N \mathcal{N}^e$ are independent from $(\mathcal{F} \mathcal{N})_k$, while the factor coming from $\mathcal{M}$ depends on $k_1$ only via the quantity $|k_1|^2 - |k_2|^2$ in view of the definition (5.22), hence the desired decomposition is valid because $|k_1|^2 - |k_2|^2$ equals $m^* \cdot k$ plus a constant once the above-mentioned difference vectors are all fixed. Also the bounds of $\mathcal{R}$ and $\mathcal{Y}$ in (5.21) can be easily proved by the above setting of $\mathcal{R}$ and $\mathcal{Y}$ together with the bounds (4.12)–(4.17) (together with the $X^\alpha 	o X^\alpha$ bounds) for the operators $\mathcal{H}^N$ and $\mathcal{M}$.

5.2.3. Term $V$. Now, let us consider term $V$ as defined in the introduction of Section 5.2. In the following proof of estimating term $V$, we will fully use the cancellation in (3.15) together with (3.5) and (3.11) and that

$$\sup_{k_3} \left| \sum_{k'_3} h^{(3)}_{k_3 k'_3} (F_{N_3})_{k'_3} \right| \lesssim N^{O(\varepsilon_1)} \cdot N^{-1} L^{-1/2} \cdot \|h\|_{k_1 k_2 + k_3 k'_3} \lesssim N^{2+C\varepsilon},$$

and we have these bounds

$$\|h\|_{k_1 k_2 + k_3 k'_3} \lesssim N^{O(\varepsilon_1)} \cdot N^{-1} L^{-1/2}, \quad \|h\|_{k_1 k_2} \lesssim N^{O(\varepsilon_1)} \cdot L_2 N^{-3} N^{1/2} L_2^{-1/2}$$

In fact, to reach the heart of the matter easily, we don’t show all details about the $\lambda_1, \lambda_2$ here but it could be seen by using (5.26) and its above argument about $\lambda$’s.
(with $L_2 = N$ if $P$ is coming from $Q^N$), where the first bound above follows from Proposition 2.7 for each fixed $k_3$, and the second bound follows from estimating $\|h\|_{k_1k_2} \leq L_2N^{-3}\|Q\|_{k_1k_1'}\|P\|_{k_2\rightarrow k_2'}$. This leads to

$$\|Q\|_{\ell^2} \lesssim N^{O(\varepsilon)} \cdot N^2N^{-1} \cdot L_2N^{-3} \cdot N^{1/2}L_2^{-1/2} \lesssim N^{-1+C\varepsilon}$$

which is enough.

Now we consider (5.27). If $P$ is coming from $Q^N$, then in (5.27) we may remove the condition $k_1' \neq k_2'$, reducing essentially to the expression in (5.3) with both $w_1$ and $w_2$ replaced by $\rho^N$, which is estimated in the same way as in Section 5.1.1. On the other hand, the term when $k_1' = k_2'$ can be estimates in the same way as (5.28) above. The same argument applies if $P$ is coming from $h^{N,L_2}$ and $\max(L_2,L_3) \geq N^{\varepsilon'}$, where we can gain a power $N^{-\varepsilon'/4}$ from either $L_2$ or $L_3$, or if $Q$ is coming from $Q^N,rem$, where we can gain extra powers $N^{-\varepsilon'/4}$ using Lemma 5.1.

Finally, consider (5.27), assuming $\max(L_2,L_3) \leq N^{\varepsilon'}$, and that $Q$ comes from $Q^{N,rem}$ in Lemma 5.1. By losing at most $N^{C\varepsilon'}$ we may fix the values of $k_1 - k_2$, $k - k_3$, $k_2 - k_2'$, $k_3 - k_3'$, and consider one single component of $Q^{N,rem}$ as described in Lemma 5.1. Then there are only two independent variables—namely $k$ and $k_1'$—and we essentially reduce (5.27) to

$$Q_k = \tilde{g}_k \cdot \sum_{k_1': \ell(k+k_1)=\Omega_0} A \cdot 1_{k_1' = \ell'} \cdot \frac{1}{(k_1')(k_2')} \langle k_1' \rangle \langle k_2' \rangle R(k_1)P_{k_2k_2'} \cdot g_{k_1'} \tilde{g}_{k_2'}. \quad (5.29)$$

Here $|A| \lesssim 1$ is a non-probabilistic factor, $|\ell|, |\ell'| \lesssim N^{\varepsilon'}$ are fixed vectors, $Y = Y(k_1)$ and $Y = Y(k_1)$ are as in Lemma 5.1 and $P = P_{k_2k_2'}$ is defined as above. Moreover we know that $Y$ and $P$ are independent from $g_{k_1'}$ and $g_{k_2'}$, that $R(k_1)$ depends only on $m^* \cdot k_1$ for some fixed vector $|m^*| \lesssim N^{\varepsilon'}$, and that $|P| \lesssim N^{O(\varepsilon)}$, $|Y| \lesssim N^{O(\varepsilon)}$ and $\|P\|_{\ell^2} \lesssim N^{1/2+O(\varepsilon)}$ (after fixing $\lambda$ as before). Finally $\tilde{g}_k$ in (5.29) is $\sum_{k_3} h^{(3)}_{k_3k_3'}(F_{N_3})k_3'$ bounded by $|\tilde{g}_k| \lesssim N^{-1}$.

Since $R(k_1)$ only depends on $m^* \cdot k_1$, if we fix the value of $m^* \cdot k_1$ in the above summation, then $R(k_1)$ can be extracted as a common factor and for the rest of the sum we can apply independence (using Proposition 2.7) and get

$$|Q_k| \lesssim |\tilde{g}_k| \cdot \sum_a |R(a)| \cdot \left( \sum_{k_1 \in S_{a,k}} \left| \frac{1}{(k_1')(k_2')} \langle k_1' \rangle \langle k_2' \rangle R(k_1)P_{k_2k_2'} \right|^2 \right)^{1/2} \lesssim N^{-3+O(\varepsilon)} \cdot \sum_a |R(a)| \cdot |S_{a,k}|^{1/2},$$

where $R(a) = R(k_1)$ for any $k_1 \cdot m^* = a$ and $S_{a,k} := \{k_1 \in \mathbb{Z}^3 : \ell \cdot (k_1 + k) = \Omega_0, k_1 \cdot m^* = a \}$. Note that in the above estimate we are dividing the set of possible $k_1$’s into subsets $S_{a,k}$ where $\ell \cdot k_1$ equals some constant, and $m^* \cdot k_1$ equals another constant, and that $S_{a,k}$ is either empty or has cardinality $\geq N^{1-C\varepsilon'}$. When $S_{a,k} = \emptyset$, $|Q_k| = 0$. When $S_{a,k} \neq \emptyset$, we have $|S_{a,k}| \geq N^{1-C\varepsilon'}$ and hence

$$|Q_k| \lesssim N^{C\varepsilon' - 7/2} \cdot \sum_a |R(a)| \cdot |S_{a,k}| = N^{C\varepsilon' - 7/2} \cdot \sum_{k_1: \ell(k+k_1) = \Omega_0} |R(k_1)|,$$

Then, using Schur’s bound, we get that

$$\|Q\|_{\ell^2} \lesssim N^{C\varepsilon' - 7/2}N^2 \|R\|_{\ell^2} \lesssim N^{-1+C\varepsilon'}$$

which is enough for (5.18). This completes the proof for $\rho^N$. 

5.3. An extra improvement. For the purpose of Section 4 we need an improvement for the \( \rho^N \) bound in (3.18), namely the following.

**Proposition 5.3.** Let \( N = M, Y \in \mathbb{R} \) be any constant, and consider \( \rho^* \) defined by
\[
(\rho^*)_k(t) = (\rho^N)_k(t) \cdot 1_{|k|^2 \leq Y+N^{c'}}
\]
then \( N \)-certainly we can improve (3.18) to \( \|\rho^*\|_{X^0(J)} \leq N^{-1/2+\varepsilon_1/2} \). Note that this bound is better than the bound for \( z_N \) in (3.18) (which is better than the bound for \( \rho^N \) in (3.18)).

**Proof.** We only need to examining the terms I~V in the above proof. For terms I and IV and V (hence also III), in the above proof we already obtain bounds better than \( N^{-1/2+\varepsilon_1/2} \), so these terms are acceptable, and we need to study term II. Note that the definition of \( \rho^* \) restricts \( k \) to a set \( E \) of cardinality \( \leq N^{1+c'} \) by the standard divisor counting bound.

Let \( h^b = h^b_{kk_1k_2k_3} \) be the base tensor, which is supported in \( |k_j| \lesssim N_j \lesssim N \) and \( |k_1-k_2| \sim R \), such that in the support of \( h^b \) we have \( k-k_1+k_2-k_3 = 0 \) and \( |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2 = \Omega_0 \).

There are three cases in term II that need consideration:

(1) The case in Section 5.1.1. Here the bound (5.8) suffices unless \( \max(N_2, R) \leq N^{C\varepsilon'} \); if this happens, note that in the above proof, (5.8) follows from the estimate
\[
\max \left( \|h^b_{kk_1k_2k_3}\|_{\ell^\infty_{kk_1k_2k_3}}, \|h^b_{kk_1k_2k_3}\|_{\ell^\infty_{kk_2k_3-k_1k}} \right) \lesssim N^{1+\delta}
\]
assuming \( \max(N_2, R) \leq N^{C\varepsilon'} \). However if we further require \( k \in E \), then the right hand side of the above bound can be improved to \( |E|^{1/2} = N^{1/2+C\varepsilon'} \), which leads to the desired improvement of (3.18).

(2) The case 1 in Section 5.1.2. Here the bound (5.10) suffices unless \( R \leq N^{C\varepsilon'} \); if this happens, note that (5.10) follows from the estimate
\[
\max \left( \|h^b\|_{\ell^\infty_{kk_2k_3-k_1k}}, \|h^b\|_{\ell^\infty_{kk_2k_1k}}, \|h^b\|_{\ell^\infty_{kk_3k_2-k_1k}}, \|h^b\|_{\ell^\infty_{kk_1k_2k}} \right) \lesssim N^{1+\delta}N_2
\]
assuming \( R \leq N^{C\varepsilon'} \). However if we further require \( k \in E \), then the right hand side can be improved to \( |E|^{1/2}N_2 = N^{1/2+C\varepsilon'}N_2 \), which allows for the improvement.

(3) The case 2 in Section 5.1.2. Here (5.12) follows from the estimate \( \|h^b\|_{\ell^\infty_{kk_2k_3-k_1}} \lesssim R^{-\beta}NR \). However if we further require \( k \in E \), then the right hand side can be improved to \( R^{-\beta}|E|^{1/2}R = R^{-\beta}N^{1/2+C\varepsilon'}R \), which allows for the improvement. This finishes the proof.

\[ \Box \]

6. The remainder terms

Now we will prove the \( z_N \) part of the bound (3.18), assuming \( N = M \). We will prove it by a continuity argument, so we may assume (3.18) and only need to improve it using the equation (3.13); note that the smallness factor is automatic as long as we use (3.13), as explained before. As such, we can assume that each input factor \( w, \hat{s} \) on the right hand side of (3.13) has one of the following four types, where in all cases we have \( N_j \leq N \):

(i) Type (G), where we define \( L_j = 1 \), and
\[
(\bar{w}_j)_{k_j}(\lambda_j) = 1_{N_j/2 < (k_j) \leq N_j} \frac{g_{k_j}(\omega)}{(k_j)} \hat{\chi}(\lambda_j).
\]

(6.1)
(ii) Type (C), where
\[(\hat{w}_j)_{k_j}(\lambda_j) = \sum_{N_j/2 < \langle k_j' \rangle \leq N_j} h^{(j)}_{k_j k_j'}(\lambda_j, \omega) \frac{g_{k_j'}(\omega)}{\langle k_j' \rangle}, \quad (6.2)\]
with \(h^{(j)}_{k_j k_j'}(\lambda_j, \omega)\) supported in the set \(\{ \frac{N_j}{2} < \langle k_j \rangle \leq N_j, \frac{N_j}{2} < \langle k_j' \rangle \leq N_j \}\), \(B_{\leq L_j}\) measurable for some \(L_j \leq N_j/2\), and satisfying the bounds (where in the first bound we first fix \(\lambda_j\), take the operator norm, and take the \(L^2\) norm in \(\lambda_j\))
\[\begin{align*}
\|\langle \lambda_j \rangle^{-1-b} h^{(j)}_{k_j k_j'}(\lambda_j)\|_{L^2_{\lambda_j}([\ell^2_{k_j} \rightarrow \ell^2_{k_j}])} &\lesssim L_j^{-1/2+3\epsilon_1}, \\
\|\langle \lambda_j \rangle^{b} h^{(j)}_{k_j k_j'}(\lambda_j)\|_{L^2_{k_j} L^2_{\lambda_j}} &\lesssim N_j^{1+\delta} L_j^{-1/2+2\epsilon_1}. \quad (6.3)
\end{align*}\]
Moreover using (3.17) we may assume \(h^{(j)}\) is supported in \(| k_j - k_j' | \leq N^\delta L_j\). Note that if \(w_j\) is of type (G), \((\hat{w}_j)_{k_j}(\lambda_j)\) can be also expressed in the same form as (6.2) but with \(h^{(j)}_{k_j k_j'} = 1_{k_j = k_j'} \cdot \hat{x}(\lambda_j)\), except the second equation in (6.3) is not true in this case.

(iii) Type (L), where \((\hat{w}_j)_{k_j}(\lambda_j)\) is supported in \(| k_j | \sim N_j \}, and satisfies
\[\|\langle \lambda_j \rangle^b (\hat{w}_j)_{k_j}(\lambda_j)\|_{L^2_{\lambda_j} [\ell^2_{k_j}]^3} \lesssim N_j^{-1/2+\epsilon_1+\epsilon_2}. \quad (6.4)\]
Also such \(w_j\) is a solution to the equation (5.1).

(iv) Type (D), where \((\hat{w}_j)_{k_j}(\lambda_j)\) is supported in \(| k_j | \lesssim N_j \}, and satisfies
\[\|\langle \lambda_j \rangle^b (\hat{w}_j)_{k_j}(\lambda_j)\|_{L^2_{k_j} L^2_{\lambda_j}} \lesssim N_j^{-1/2+\epsilon_1}. \quad (6.5)\]

Now, let the multilinear forms \(\mathcal{M}^\circ, \mathcal{M}^\prec, \mathcal{M}^\succ\) and \(\mathcal{M}^{\ll}\) be as in (2.4), (3.3) and (3.9). The terms on the right hand side of (3.13) are, from the first line of (3.13) which is trivially bounded, the followings:

(1) The term
\[I = \mathcal{I}_\chi \Pi_N \mathcal{M}^\prec(w_1, w_2, w_3)\]
where \(w_j\) can be any type and \(\text{max}(N_1, N_2, N_3) = N\).

(2) The term
\[II = \mathcal{I}_\chi \Pi_N (\mathcal{M}^\prec - \mathcal{M}^{\ll})(w_1, w_2, w_3)\]
where \(w_j\) can be any type and \(\text{max}(N_1, N_2) = N\).

(3) The term
\[\text{III} = \mathcal{I}_\chi \Delta_N \mathcal{M}^\circ(w_1, w_2, w_3)\]
where \(w_j\) can be any type and \(\text{max}(N_1, N_2, N_3) \leq N/2\).

(4) The term
\[IV = \mathcal{I}_\chi \Pi_{N/2} \mathcal{M}^\prec(w_1, w_2, w_3)\]
where \(w_j\) can be any type and \(\text{max}(N_1, N_2) \leq N/2\) and \(N_3 = N\).

(5) The term
\[V = \mathcal{I}_\chi \Pi_{N/2} \mathcal{M}^{\ll}(w_1, w_2, w_3)\]
where \(w_j\) can be any type and \(\text{max}(N_1, N_2) = N_3 = N\).

(6) The term
\[VI = \mathcal{I}_\chi \Delta_N \mathcal{M}^\prec(w_1, w_2, w_3)\]
where \(w_1\) and \(w_2\) can be any type, \(w_3\) has type (D) and \(\text{max}(N_1, N_2) \leq N/2\) and \(N_3 = N\).
(7) The term
\[ \text{VII} = \mathcal{I}_\chi \Delta_N \mathcal{M}^\infty(w_1, w_2, w_3) \]
where \( w_1 \) and \( w_2 \) can be any type, \( w_3 \) has type (D) and \( \max(N_1, N_2) = N_3 = N \).

(8) The term VIII, which is the last two lines of the right hand side of (3.13).

Our goal is to recover the bound for \( z_N \) in (3.18) for each of the terms I–VIII above. In doing so we will consider two cases. First is the no-pairing case, where if \( w_1 \) and \( w_2 \) are of type (C) or (G) and hence expanded as in (6.2), then we assume \( k'_1 \neq k'_2 \); similarly if \( w_2 \) and \( w_3 \) are of type (C) or (G) then we assume \( k'_2 \neq k'_3 \). The second case is the pairing case which is when \( k'_1 = k'_2 \) or \( k'_2 = k'_3 \) (the over-pairing case where \( k'_1 = k'_2 = k'_3 \) is easy and we shall omit it). We will deal with the no-pairing case for terms I–VII in Sections 6.1–6.3, the pairing case for these terms in Section 6.4, and term VIII in Section 6.5.

6.1. No-pairing case. We start with the no-pairing case.

6.1.1. Preparation of the proof. We start with some general reductions in the no-pairing case. Recall as in Section 3.4 that we can always gain a smallness factor from the short time \( \tau \ll 1 \), and can always ignore losses of \( (N_3)^{C_k^{-1}} \) provided we can gain a power \( N^{-\epsilon/10} \) (which will be clear in the proof). We will consider \( \mathcal{I}_\chi \mathcal{M}^*(w_1, w_2, w_3)_k(\lambda) \) where \( \mathcal{M}^*(\cdot) \) can be one of \( \Pi \mathcal{M}^0, \Pi \mathcal{M}^\infty, \Pi \mathcal{M}^\geq, \Pi \mathcal{M}^\leq \) and \( \Pi(\mathcal{M}^\infty - \mathcal{M}^\leq) \) with \( \Pi \) being a general notation for projections for \( \Pi_N, \Pi_{N/2} \) and \( \Delta_N \),

\[
\mathcal{I} \mathcal{M}^*(w_1, w_2, w_3)_k(\lambda) = \sum_{(k_1, k_2, k_3), k_2 \neq k_1 + k_3, k_2 \neq k_1 - k_3} \int d\lambda_1 d\lambda_2 d\lambda_3 I(\lambda, \Omega + \lambda_1 - \lambda_2 + \lambda_3) \\
\times V_{k_1 - k_2} \cdot (\hat{w}_1)_k(\lambda_1) (\hat{w}_2)_k(\lambda_2) (\hat{w}_3)_k(\lambda_3), \quad (6.6)
\]

where \( \Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2 \) and \( \sum \) is directly defined based on the definitions of \( \mathcal{M}^0, \mathcal{M}^\infty, \mathcal{M}^\geq \) and \( \mathcal{M}^\leq \) and the selection of \( \Pi \). For example, if \( \mathcal{M}^*(\cdot) = \Pi_N \mathcal{M}^\geq \), then there will be two more restrictions \( |k| \leq N \) and \( (k_1 - k_2) > N^{1-\delta} \) in the sum \( \sum \). The other \( \sum \) will defined in the similar ways.

Before going into the different estimates for I–VII, we first make a few remarks.

• If a position \( w_j \) has type (L) or (D), then in most cases we only need to consider type (L) terms since (6.5) is stronger than (6.4); there are exceptions that will be treated separately later.

• the \( w_j \) of type (G) can be considered as a special case of type (C) when \( \ell^{(j)}_{k_j k'_j} (\lambda_j) = 1_{N_j/2 < (k_j) \leq N_j} \cdot 1_{k_j = k'_j} \cdot \hat{\chi}(\lambda_j) \); if we avoid using the \( \ell^2_{k_j k'_j} \) norm in (6.3), then we only need to consider type (C) terms.

• Term I can be estimates in the same way as term II. In fact the definition of \( \mathcal{M}^\geq \) implies \( \max(N_1, N_2) \geq N^{1-\delta} \), so we are essentially in (special case of) term II up to a possible loss \( NC^3 \) which will be negligible compared to the gain. Moreover, term V can be estimated similarly as term IV, see Section 6.3.

• Terms VI and VII are readily estimated using the \( X^\alpha \to X^\alpha \) bounds for the linear operator (3.19) proved in Sections 4.1 and 4.2.
Based on these remarks, from now on we will consider terms II–IV (and VIII at the end), where the possible cases for the types of \( (w_1, w_2, w_3) \) are (a) \((C, C, C)\), (b) \((C, C, L)\), (c) \((C, L, C)\), (d) \((L, C, C)\), (e) \((L, L, C)\), (f) \((C, L, L)\), (g) \((L, C, L)\), and (h) \((L, L, L)\).

In Section 6.2 we will estimate term II, which can be understood as high-high interactions in view of \( \max(N_1, N_2) = N \), and noticing that assuming \( k \) is the high frequency, then either \( k_3 \) is also high frequency or \( |k_1 - k_2| \) must be large. In Section 6.3 we will estimate terms III and IV by using a counting technique in a special situation called \( \Gamma \)-condition (see (6.18)). In Section 6.4 we consider the pairing case.

### 6.2. High-high interactions.

We will estimate term II in this subsection. First we can repeat the arguments for \( \lambda \), \( \lambda_j \) and the Duhamel operator \( I \) in (6.6) as in Section 4 and 5. Namely, we first restrict to \( |\lambda_j| \leq N^{100} \) and \( |\lambda|, |\mu| \leq N^{100} \) where \( \mu = \lambda - (\Omega + \lambda_1 - \lambda_2 + \lambda_3) \), and replace the unfavorable exponents \((1 - b) \) or \( b \) depending on the context) by the favorable ones \((b \) or \( 1 - b) \), then exploit the resulting integrability in \( \lambda_j \) to fix the values of \( \lambda \), \( \lambda_j \) and \( |\mu| \). Then we reduce to the following expression where \( \Omega_0 \) is a fixed integer:

\[
X_k := \sum_{k_1 - k_2 + k_3 = k} h^b_{k_1 k_2 k_3} (\hat{w}_1)_{k_1} (\hat{w}_2)_{k_2} (\hat{w}_3)_{k_3},
\]

where \( h^b \) is the base tensor which contains the factors

\[
V_{k_1 - k_2} \cdot 1_{k_1 - k_2 + k' = k} \cdot 1_{|k_1|^2 - |k_1|^2 + |k_2|^2 - |k'|^2 = \Omega_0}.
\]

We assume \( h^b \) is supported in the set where \( |k_j| \leq N_j \) and \( (k_1 - k_2) \sim R \) where \( R \) is a dyadic number. Moreover we assume that \( R \) and the support of \( h^b \) satisfies the conditions associated with the definition of some \( \mathcal{M}^{(*)} \). In view of the factor \( |V_{k_1 - k_2}| \sim R^{-\beta} \) in \( h^b \), we also define \( h^{R,(*)} := R^\beta \cdot h^b \), which is essentially the characteristic function of the set

\[
S^R = \left\{ (k_1, k_2, k_3) \in (\mathbb{Z}^3)^4, \quad k_2 \notin \{k_1, k_3\} \right\},
\]

possibly with extra conditions determined by the definition of \( \mathcal{M}^{(*)} \). We also define \( S^R_k \) to be the set of \((k_1, k_2, k_3) \in \mathcal{S}^R \) with fixed \( k \), and similarly define \( S^R_{k_1 k_2} \) etc. Noticing that when \( w_j \) has type \((G), (C) \) or \((L) \), we can further assume that \( |k_j| > N/2 \) in the definition of \( S^R \).

The goal now is to bound the norm \( \|X_k\|_\ell^2 \) or abbreviated \( \|X_k\|_k \), assuming \( w_j \) satisfy the bounds (6.1)–(6.5) but without the \( \lambda_j \) component, for example (6.5) becomes \( \|w_j\|_{k_j} \lesssim N_j^{-1/2 + \epsilon} \).

#### 6.2.1. Case (a): \((C, C, C)\).

In this case we have

\[
X_k = R^{-\beta} \sum_{(k_1, k_2, k_3)} h^{R,(*)}_{k_1 k_2 k_3} \cdot \sum_{(k_1', k_2', k_3')} g_{k_1 k_2 k_3} g_{k_1' k_2' k_3'} \cdot \frac{g_{k_1} g_{k_2} g_{k_3}}{(k_1') (k_2') (k_3')},
\]

where \( h^{(j)}_{k_j k'_j} = h^{(j)}_{k_j k'_j} (\omega) \) satisfies (6.3) with some \( N_j \) and \( L_j \leq N_j/2 \) for \( 1 \leq j \leq 3 \) and \( h^{R,(*)}_{k_1 k_2 k_3} \) is defined as above.

To estimate \( \|X\|_k \) we would like to apply Proposition 2.7 and then Proposition 2.6. Like in Sections 4.1 and 5.1 the way we apply Proposition 2.7 depends on the relative sizes of \( N_j \) \((1 \leq j \leq 3)\).
For example, if $N_1 = N_2 = N_3$ we shall apply Proposition 2.7 jointly in the $(k'_1, k'_2, k'_3)$ summation in (6.9); if $N_1 = N_3 > N_2$ we will first apply Proposition 2.7 jointly in the $(k'_1, k'_2)$ summation, then apply it in the $k'_3$ summation, if $N_3 > N_1 > N_2$ we will apply first in the $k'_3$ summation, then in the $k'_1$ summation, and then in the $k'_2$ summation, etc. The results in the end will be the same in all cases, so for example we will consider the case $N_3 > N_1 > N_2$. Now we have

$$\|X_k\|_k = R^{-\beta} \left\| \sum_{k'_3} \left( \sum_{k_3} \tilde{H}_{kk_3} h^{(3)}_{k_3 k'_{3}} \right) g_{k'_3} \left( k'_3 \right) \right\|_k,$$

(6.10)

where

$$\tilde{H}_{kk_3} := \sum_{k'_1} \left( \sum_{k_1} \tilde{H}_{kk_1} h^{(1)}_{k_1 k'_1} \right) g_{k'_1} \left( k'_1 \right), \quad \tilde{H}_{kk_3} := \sum_{k'_2} \left( \sum_{k_2} h^{R, (\ast)}_{kk_2} h^{(2)}_{k_2 k'_2} \right) g_{k'_2} \left( k'_2 \right).$$

(6.11)

By the independence between $g_{k'_3}$ and $\tilde{H}_{kk_3} h^{(3)}_{k_3 k'_3}$ since $N_3 > N_1 > N_2$, we apply Proposition 2.7 and Proposition 2.6 and get $\tau^{-1} N_4$-certainly that

$$\|X_k\|_k \leq R^{-\beta} N_3^{-1} \cdot \left\| \sum_{k_3} \tilde{H}_{kk_3} h^{(3)}_{k_3 k'_3} \right\|_{k'_3} \leq R^{-\beta} N_3^{-1} \cdot \|h^{(3)}_{k_3 k'_3}\|_{k'_3 \to k_3} \|\tilde{H}_{kk_3}\|_{kk_3}$$

(6.12)

Similarly, by the independence between $g_{k'_1}$ and $\tilde{H}_{kk_1} h^{(1)}_{k_1 k'_1}$ since $N_1 > N_2$, and also by the independence between $g_{k'_2}$ and $h^{R, (\ast)}_{kk_2} h^{(2)}_{k_2 k'_2}$, once again we can apply Proposition 2.7 and Proposition 2.6 to $\|\tilde{H}_{kk_3}\|_{kk_3}$ and then to $\|\tilde{H}\|_{kk_1 k_3}$. As a consequence, we have $\tau^{-1} N_4$-certainly

$$\|X_k\|_k \lesssim R^{-\beta} (N_1 N_2 N_3)^{-1} \cdot \left( \prod_{j=1}^{3} \|h^{(j)}_{k_j k'_j}\|_{k'_j \to k_j} \right) \cdot \|h^{R, (\ast)}_{kk_1 k_3}\|_{kk_1 k_3}.$$

(6.13)

In the other cases we get the same bound. Without loss of generality we may assume $N_1 = N$, then using Lemma 2.3 we can estimate

$$\|h^{R, (\ast)}_{kk_1 k_3}\|_{kk_1 k_3} \lesssim N^\delta : N_3^{3/2} : RN_2$$

which implies that $\|X_k\|_k \lesssim N^{1+C\delta} N_{3}^{1/2} \lesssim N^{-1/2+C\delta}$, which is enough for (3.18).

6.2.2. Case (b): $(C, C, L)$. In this case we have

$$X_k = R^{-\beta} \sum_{(k_1, k_2, k_3)} h^{R, (\ast)}_{kk_1 k_2 k_3} \cdot \sum_{(k'_1, k'_2, k'_3)} h^{(1)}_{k'_1 k_1} h^{(2)}_{k'_2 k_2} \frac{g_{k'_1} g_{k'_2}}{\langle k'_1 \rangle \langle k'_2 \rangle} (w_3)_{k_3},$$

(6.14)

where $h^{(j)}_{k_j k'_j}(\omega)$ satisfies (6.3) with some $N_j$ and $L_j \leq N_j/2$ for $1 \leq j \leq 2$ and the base tensor $h^{R, (\ast)}_{kk_1 k_3}$ is defined as before. Clearly $\|X_k\|_k$ can be bounded by $N_3^{-1/2+\varepsilon_1+\varepsilon_2}$ times the norm

$$R^{-\beta} \left\| \sum_{(k_1, k_2, k_3)} h^{R, (\ast)}_{kk_1 k_2 k_3} \cdot \sum_{(k'_1, k'_2, k'_3)} h^{(1)}_{k'_1 k_1} h^{(2)}_{k'_2 k_2} \frac{g_{k'_1} g_{k'_2}}{\langle k'_1 \rangle \langle k'_2 \rangle} \right\|_{k \to k_3}.$$
By applying Propositions 2.7 and 2.6 again, in the same manner as (6.2.1), we get that the above norm is bounded by

\[ R^{-\beta} \cdot (N_1 N_2)^{-1} \max(\| h^{R,(\ast)}_{k \rightarrow k_1 k_2 k_3} \|, \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \|, \| h^{R,(\ast)}_{k k_2 \rightarrow k_1 k_3} \|, \| h^{R,(\ast)}_{k k_1 k_2 \rightarrow k_3} \|). \]

By Lemma 2.4 we can conclude that

\[ \max(\| h^{R,(\ast)}_{k \rightarrow k_1 k_2 k_3} \|, \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \|, \| h^{R,(\ast)}_{k k_2 \rightarrow k_1 k_3} \|, \| h^{R,(\ast)}_{k k_1 k_2 \rightarrow k_3} \|) \lesssim R \cdot \min(N_1, N_2), \]

hence we easily get \( \| X_k \| \lesssim N^{-1+C\varepsilon_1} \), which is enough for (3.18).

6.2.3. Cases (c): (C,L,C) and (d): (L,C,C). The estimates of Case (c) and Case (d) are similar with Case (b), so we will state the estimates in Case (c) and Case (d) without proofs. In Case (c) we get

\[ \| X_k \| \lesssim N^{-1/2+\varepsilon_1+\varepsilon_2} R^{-\beta} (N_1 N_2)^{-1} \max(\| h^{R,(\ast)}_{k \rightarrow k_1 k_2 k_3} \|, \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \|, \| h^{R,(\ast)}_{k k_2 \rightarrow k_1 k_3} \|, \| h^{R,(\ast)}_{k k_1 k_2 \rightarrow k_3} \|) \]

and in case (d) we get a similar bound, but with the subindices 1 and 2 switched.

Now by Lemma 2.4 we can obtain that

\[ \max(\| h^{R,(\ast)}_{k \rightarrow k_1 k_2 k_3} \|, \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \|, \| h^{R,(\ast)}_{k k_2 \rightarrow k_1 k_3} \|, \| h^{R,(\ast)}_{k k_1 k_2 \rightarrow k_3} \|) \lesssim N^{C_2} \cdot N_3 \cdot \min(N_1, N_2), \]

\[ \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \| \lesssim N^{C_2} \cdot \min(R \cdot \min(N_1, N_2), N_1 N_3), \]

\[ \| h^{R,(\ast)}_{k k_1 k_3 \rightarrow k_2} \| \lesssim \min(R, N_1) N_3. \]

In the first case we directly get

\[ \| X_k \| \lesssim N_1^{-1} N_2^{-1/2+\varepsilon_1+\varepsilon_2} R^{-\beta} \cdot \min(N_1, N_2) \]

which is enough for (3.18) as \( \max(N_1, N_2) = N \) and \( R \geq N^{\varepsilon} \) in view of the definition of \( \mathcal{M}_c - \mathcal{M}_\varepsilon \).

In the second case we get

\[ \| X_k \| \lesssim \min( R^{-\beta} N_1^{-1/2+\varepsilon_1+\varepsilon_2}, N_2^{-1/2+\varepsilon_1+\varepsilon_2} R^{-\beta} ) \]

which is also enough for (3.18) as \( \max(N_1, N_2) = N \) and \( R \geq N^{\varepsilon} \). In the third case we get

\[ \| X_k \| \lesssim N_2^{-1/2+\varepsilon_1+\varepsilon_2} \max(R, N_1)^{-1} \]

which is also enough for (3.18). By switching the indices 1 and 2 we also get the same estimates in case (d).

6.2.4. Case (e): (L,L,C). In this case we have

\[ X_k = \sum_{k_3} \sum_{(k_1, k_2, k_3)} h_{k k_1 k_2 k_3}^{R,(\ast)} (w_1 k_1 \overline{(w_2 k_2) k_3} h_{k k_1 k_2 k_3}^{(3)} g_{k_3}^{2} \overline{A}^{(3)}_{k} (k_3)). \]  

(6.15)

where \( h_{k k'_3}^{(3)} = h_{k k'_3}^{(3)} (\omega) \) satisfies (6.3) with some \( N_3 \) and \( L_3 \leq N_3 / 2 \) and the base tensor \( h^{R,(\ast)}_{k k_1 k_2 k_3} \) is defined as before. By symmetry we may assume \( N_1 \leq N_2 \), then by the same argument as above, using Propositions 2.6 and 2.7 we can bound

\[ \| X_k \| \lesssim (N_1 N_2)^{-1/2+\varepsilon_1+\varepsilon_2} N_3^{-1} R^{-\beta} \cdot \max(\| h^{R,(\ast)}_{k k_1 k_3 \rightarrow k_2} \|, \| h^{R,(\ast)}_{k k_1 \rightarrow k_2 k_3} \|). \]

By Lemma 2.4 both tensor norms are bounded by \( \min(N_1, R) N_3 \); as \( N_1 \leq N_2 \) (and hence \( N_2 = N \)) and \( R \geq N^{\varepsilon} \), it is easy to check that this bound is enough for (3.18).
6.2.5. Cases (f): \((C,L,L)\) and (g): \((L,C,L)\). The estimates of Case (f) and (g) are similar with Case (e), so we will state the estimates of Case (f) and (g) directly. Again the two cases here only differs by switching indices 1 and 2, so we only consider case (f). Like in case (e) we get two bounds:
\[
\|X_k\|_k \lesssim (N_2 N_3)^{-1/2+\varepsilon_1+\varepsilon_2} N_1^{-1} R^{-\beta} \max(\|h^{R,(*)}_{kk_1 k_2 \to k_3}\|, \|h^{R,(*)}_{kk_2 \to k_1 k_3}\|)
\]
and
\[
\|X_k\|_k \lesssim (N_2 N_3)^{-1/2+\varepsilon_1+\varepsilon_2} N_1^{-1} R^{-\beta} \max(\|h^{R,(*)}_{kk_1 k_2 \to k_3}\|, \|h^{R,(*)}_{k k_2 \to k_1 k_3}\|)
\]
Now if \(N_3 \geq N^\varepsilon\) we will apply the first bound and use that
\[
\max(\|h^{R,(*)}_{kk_1 k_2 \to k_3}\|, \|h^{R,(*)}_{k k_2 \to k_1 k_3}\|) \lesssim R \min(N_1, N_2),
\]
so the factor \(N_1^{-1} N_2^{-1/2+\varepsilon_1+\varepsilon_2} \min(N_1, N_2)\), together with \(N_3^{-1/2+\varepsilon_1+\varepsilon_2}\) where \(N_3 \geq N^\varepsilon\), provides the bound that is enough for (3.18). Moreover, the same bound also works if \(N_2 \leq N^{1-\varepsilon^2}\) (since in this case \(N_1 = N\)).

If \(N_3 \leq N^\varepsilon^2\) and \(N_2 \geq N^{1-\varepsilon^2}\) we will apply the second bound and use that
\[
\max(\|h^{R,(*)}_{kk_1 k_2 \to k_3}\|, \|h^{R,(*)}_{k k_2 \to k_1 k_3}\|) \lesssim N^{C\varepsilon^2} N_1
\]
assuming \(N_3 \leq N^\varepsilon^2\). This is also enough for (3.18) assuming \(N_2 \geq N^{1-\varepsilon^2}\) and \(R \geq N^\varepsilon\).

6.2.6. Case (h): \((L,L,L)\). In this case we have
\[
X_k := \sum_{(k_1, k_2, k_3)} h^{R,(*)}_{kk_1 k_2 k_3} \cdot \overline{(w_1)(w_2)_{k_2}(w_3)_{k_3}},
\]
where the base tensor \(h^{R,(*)}_{kk_1 k_2 k_3}\) is defined as before. Then simply using Proposition 2.6 we get
\[
\|X_k\|_k \lesssim R^{-\beta} \cdot (N_1 N_2 N_3)^{-1/2+\varepsilon_1+\varepsilon_2} \cdot \max(\|h^{R,(*)}_{kk_1 k_2 k_3}\|_{kk_2 \to k_1 k_3}).
\]
By Lemma 2.4 we have \(\|h^{R,(*)}_{kk_1 k_2 k_3}\|_{kk_2 \to k_1 k_3} \lesssim (R \min(N_1, N_2))^{1/2}\), which implies
\[
\|X_k\|_k \lesssim R^{-\beta+1/2} \max(N_1, N_2)^{-1/2+C\varepsilon_1}
\]
which is enough for (3.18) because \(\max(N_1, N_2) = N\) and \(R \geq N^\varepsilon\).

6.3. The \(\Gamma\) condition terms. In this section we estimate terms III and IV. These two terms are actually similar, and the key property they satisfy is the so-called \(\Gamma\) condition. Namely, due to the projections and assumptions on the inputs in terms III and IV, we have that
\[
|k|^2 \geq \Gamma \geq |k_3|^2, \quad \text{for all } (k, k_1, k_2, k_3) \in S
\]
or
\[
|k|^2 \leq \Gamma \leq |k_3|^2, \quad \text{for all } (k, k_1, k_2, k_3) \in S
\]
for some real number \(\Gamma\), where \(S\) is the support of the base tensor \(h^b\) (note that in term IV we may assume \(w_3\) is not of type (D) as otherwise the bound follows from what we have already done, so here we may choose \(\Gamma = (N/2)^2 - 1\)).

To proceed, we return to \(T^{(\mu)}_{\lambda} \mathcal{N}^{(\lambda)}(w_1, w_2, w_3, \lambda)\) in (6.5) where \(\Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2\) and suppose \(\mu = \lambda - (\Omega + \lambda_1 - \lambda_2 + \lambda_3)\) and then we have \(|I| \lesssim (\lambda)^{-\mu} (\mu)^{-1}\) by (2.10). Following the same reduction steps as before, we can assume \(|\lambda|, |\lambda_j| (j = 1, 2, 3), |\mu| \leq N^{100}\) and may replace the unfavorable exponents by the favorable ones. Now, instead of fixing each \(\lambda_j\) and \(\lambda\) and \(|\mu|\), we do the following.
Without loss of generality, we may assume $|\lambda_3|$ is the maximum of all the parameters $|\lambda_j|$ and $|\lambda|$ and $|\mu|$; the other cases are treated similarly. We may fix a dyadic number $K$ and assume $|\lambda_3| \sim K$. Then, we may fix $\lambda_j$ ($j \neq 3$) and $\lambda$ and $|\mu|$, again using integrability in these variables, and exploit the weight $\langle \lambda_3 \rangle^b$ in the weighted norms in which $w_3$ is bounded, and reduce to an expression

\[ X_k := R^{-\beta} K^{-b} \sum_{k_1-k_2+k_3=k} \int d\lambda_3 \cdot h_{k_1,k_2,k_3}^{R,k}(\lambda_3) \cdot (w_1)_{k_1} (w_2)_{k_2} \cdot (\tilde{w}_3)_{k_3}(\lambda_3), \]  

(6.19)

where $(\tilde{w}_3)_{k_3}(\lambda_3) = K^b(\tilde{w}_3)_{k_3}(\lambda_3)$ and $h_{k_1,k_2,k_3}^{R,k}(\lambda_3)$ is essentially the characteristic function of the set (with possibly more restrictions according to the definition of $M^{(\epsilon)}$)

\[ S_{R,K} = \left\{ (k, k_1, k_2, k_3, \lambda_3) \in (\mathbb{Z}^3)^4 \times \mathbb{R}, \quad k_2 \notin \{k_1, k_3\} \right\}, \]  

(6.20)

where $\Omega_0$ is a fixed number such that $|\Omega_0| \lesssim K$. We also define the sets $S_{k,R,M}^{R,M}$ to be the set of $(k_1, k_2, k_3, \lambda_3)$ such that $(k, k_1, k_2, k_3, \lambda_3) \in S_{R,K}$ for fixed $k$ etc.. Note that when $w_j$ is of type (C), (G) or (L), we can further assume $N_j < |k_j| \leq N_j$.

The idea in estimating (6.19) is to view $(k_3, \lambda_3)$ as a whole (say denote it by $\tilde{k}_3$), which will allow us to gain using the $\Gamma$ condition in estimating the norms of the base tensor $h_{R,k}(\lambda_3)$. Though our tensors here involve the variable $\lambda_3 \in \mathbb{R}$, it is clear that Propositions 2.5 and 2.6 still hold for such tensors, and Proposition 2.7 can also be proved by using a meshing argument (see Section 3.4, where the derivative bounds in $\lambda_3$ is easily proved as all the relevant functions are compactly supported in physical space). Moreover, by the induction hypothesis and the manipulation above (for example with $Y^{1-b}$ norm replaced by $Y^b$ norm) we can also deduce corresponding bounds for $w_3 = (w_3)_{k_3}$ and the corresponding matrices such that $\tilde{h}_{k_3,k_3'}^{(3)}$, for example $\|h_{k_3,k_3'}^{(3)}\|_{k_3,k_3'} \lesssim \mathcal{L}_{1/2+\epsilon_3}^{-1}$. Because of this, in the proof below we will simply write $\sum_{k_3} \int d\lambda_3$, so the proof has the same format as the previous ones.

We now consider the input functions. In term III, clearly max$(N_1, N_2, N_3) \gtrsim N$; if $N_3 \ll N$, then we must have max$(N_1, N_2) \gtrsim N$ and $|k_1 - k_2| \gtrsim N$, hence this term can be treated in the same way as term II. Therefore we may assume $N_3 \sim N$, and clearly the same happens for term IV. If max$(N_1, N_2) \gtrsim N$, then again using term II estimate we only need to consider the case where $|k_1 - k_2| \lesssim N^\epsilon$. This term can be treated using similar arguments as below and is much easier due to the smallness of $|k_1 - k_2|$, so we will only consider the case max$(N_1, N_2) \ll N$. In the same way we will not consider term V here. Finally, if $w_3 = z_{N_3}$ with $N_3 \sim N$, then (3.18) directly follows from the linear estimate proved in Section 4.11 and the $\Gamma$ condition is not needed.

There are two cases: when $w_3$ has type (L) or or $w_3$ has type (C) (or (G)). In the latter case there are four further cases for the types of $w_1$ and $w_2$, which we will discuss below.

6.3.1. The type (L) case. Suppose $w_3$ has type (L). Clearly if max$(N_1, N_2) \geq N^{100\epsilon_2}$ then (3.18) also follows from the linear estimates in Section 4.11 (because the difference between the $\rho^N$ bound and the $z_N$ bound in (3.18) is at most $N^{\epsilon_2}$), so we may assume max$(N_1, N_2) \leq N^{100\epsilon_2}$. Then in
we may further fix the values of \((k_1, k_2)\) at the price of \(N^{C\varepsilon_2}\), hence we may write
\[
\mathcal{X}_k = R^{-\beta} K^{-b} \sum_{\tilde{k}_3} h(k, \tilde{k}_3) \cdot (\tilde{w}_3)_{\tilde{k}_3}
\]
and by definition it is easy to see that \(|h|_{k \rightarrow \tilde{k}_3} \lesssim 1\). Then, \((3.18)\) follows, using the bound for \(w_3\), if \(K \geq N^{\varepsilon_2}\). Finally, if \(K \leq N^{\varepsilon_2}\), then we have \(|\Omega| \lesssim N^{\varepsilon_2}\) where \(\Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2\). Using the \(\Gamma\) condition \((6.18)\) we conclude that \(|k_3|^2\) belongs to an interval of length \(N^{O(\varepsilon_2^2)}\), so we can apply Proposition \((5.3)\) to gain a power \(N^{-\varepsilon_1/2}\), which covers the loss \(N^{O(\varepsilon_2 + \varepsilon_2^2)}\) and is enough for \((3.18)\).

6.3.2. The type \((C,C,C)\) case. Now suppose \(w_1, w_2\) and \(w_3\) has type \((C,C,C)\). By symmetry we may assume \(N_1 \leq N_2\). Then by the same argument as in Section \(6.2.1\) we obtain that
\[
||\mathcal{X}_k||_k \lesssim R^{-\beta} K^{-b} (N_1 N_2 N)^{-1} \|h^{R,K,(\ast)}\|_{kk_1 k_2 k_3} \cdot \|h^{(1)}\|_{k_1 \rightarrow k_1'} \|h^{(2)}\|_{k_2 \rightarrow k_2} \|h^{(3)}\|_{k_3' \rightarrow \tilde{k}_3}.
\]
The last three factors are easily bounded by 1, so it suffices to bound the tensor \(h^{R,K,(\ast)}\).

By definition, this is equivalent to counting the number of lattice points \((k, k_1, k_2, k_3)\) such that \(k_1 - k_2 + k_3 = k\) (and also satisfying the inequalities listed above) and \(|\Omega| \lesssim K\). Note that
\[
|k_1|^2 - |k_2|^2| \lesssim R \cdot \max(N_1, R) := K_1,
\]
so when \(K \leq K_1\), by the \(\Gamma\) condition, \(|k|^2\) has at most \(K_1\) choices, hence \(k\) has at most \(K_1 N\) choices. Once \(K\) is fixed, the number of choices fo \((k_1, k_2, k_3)\) is at most \(K N_1^2 R^2\), which leads to the bound
\[
||h^{R,K,(\ast)}\|^2_{kk_1 k_2 k_3} \lesssim N^{C\delta} \cdot K K_1 N N_1^2 R^2.
\]
If instead \(K \geq K_1\), then \(k\) has at most \(K N\) choices, and once \(k\) is fixed the number of choices for \((k_1, k_2, k_3)\) is at most \(N_1^3 R^3\), so we get
\[
||h^{R,K,(\ast)}\|^2_{kk_1 k_2 k_3} \lesssim N^{C\delta} \cdot K N N_1^3 R^3.
\]
In either way we get
\[
||\mathcal{X}_k||_k \lesssim N^{C\varepsilon_2} N^{-1/2} \cdot \max(R, R^{1/2} N_1^{1/2}) N_2^{-1}
\]
which is enough for \((3.18)\) as \(\max(R, N_1) \lesssim N_2\).

6.3.3. The type \((L,L,C)\) case. Now suppose \(w_1, w_2\) and \(w_3\) has type \((L,L,C)\). First assume \(N_1 \leq N_2\). The same arguments in Section \(6.2.4\) yields
\[
||\mathcal{X}_k||_k \lesssim (N_1 N_2)^{-1/2 + \varepsilon_1 + \varepsilon_2} N^{-1} R^{-\beta} \cdot \max(||h^{R,K,(\ast)}||_{kk_1 k_3 \rightarrow k_2}, ||h^{R,K,(\ast)}||_{kk_1 \rightarrow k_2 k_3}).
\]
The second norm above is easily bounded by \(K^{1/2} R N_1\) using Lemma \(2.4\) which is clearly enough for \((3.18)\); for the first norm there are two ways to estimate. 

The first way is to use Lemma \(2.4\) directly, without using \(\Gamma\) condition, to get
\[
||h^{R,K,(\ast)}||_{kk_1 k_3 \rightarrow k_2} \lesssim K^{1/2} \min(R, N_1) N.
\]
The second way is to use the \(\Gamma\) condition and first fix the value of \(|k|^2\) and hence \(k\), then count \((k_1, k_3)\). This yields
\[
||h^{R,K,(\ast)}||_{kk_1 k_3 \rightarrow k_2} \lesssim K^{1/2} N^{1/2} (R + R^{1/2} N_1^{1/2}) \min(R, N_1)^{1/2}
\]
assuming $K \leq K_1$, and a better bound assuming $K \geq K_1$. Now, plugging in the second bound yields
\[
\|X_k\| \lesssim (N_1 N_2)^{-1/2+\epsilon_1+\epsilon_2} N^{-1} R^{-\beta} K^{-b} \cdot K^{1/2} N^{1/2} (R + R^{1/2} N_1^{1/2}) \min(R, N_1)^{1/2},
\]
which can be shown to be $\lesssim N^{-1/2}$ using the fact $\max(R, N_1) \leq N_2$ and by considering whether $R \geq N_1$ or $R \leq N_1$. Moreover the same estimate can be checked to work if $N_1 \leq N_1^{1,1}$. If $N_1 \geq N_2^{1,1}$ we can switch the subscripts 1 and 2, in which case we have the weaker bound
\[
\|X_k\| \lesssim (N_1 N_2)^{-1/2+\epsilon_1+\epsilon_2} N^{-1} R^{-\beta} K^{-b} \cdot K^{1/2} N^{1/2} (R + R^{1/2} N_2^{1/2}) \min(R, N_2),
\]
without the 1/2 power in the last factor, however this is still $\lesssim N^{-1/2}$ provided $N_1 \geq N_2^{1,1}$.

6.3.4. The type $(L,C,C)$ and $(C,L,C)$ cases. Now suppose $w_1$, $w_2$ and $w_3$ has type (L,C,C); the case (C,L,C) is treated similarly. Here the same arguments in Section 6.2.3 implies
\[
\|X_k\| \lesssim N_1^{-1} N_2^{-1/2+\epsilon_1+\epsilon_2} N^{-1} R^{-\beta} K^{-b} \times \max(\|h_{R,K,(*)}\|_{kk_{k_3} \rightarrow k_2}, \|h_{R,K,(*)}\|_{k_{k_3} \rightarrow k_{k_3}}, \|h_{R,K,(*)}\|_{k_{k_3} \rightarrow k_{k_2}}).
\]
(6.21)
The two norms $k \rightarrow k_{k_3} k_3$ and $k_{k_1} \rightarrow k_{k_3}$ can be estimated by $K^{1/2} R \min(N_1, N_2)$, using Lemma 2.4 only and without $\epsilon$, which is clearly enough for (3.18). For the $kk_{k_3} \rightarrow k_2$ norm we can use the estimates in Section 6.3.3 and get
\[
\|h_{R,K,(*)}\|_{kk_{k_3} \rightarrow k_2} \lesssim K^{1/2} N^{1/2} (R + R^{1/2} N^{1/2}) \min(R, N_1) \lesssim K^{1/2} N^{1/2} R N_1
\]
up to $N^{C\beta}$ losses, which yields
\[
\|X_k\| \lesssim R^{1-\beta} N^{-1/2} N_2^{-1/2+\epsilon_1+\epsilon_2}
\]
and is also enough for (3.18). Finally we consider the $k_{k_3} \rightarrow k_{k_1}$ norm. By Schur's bound and using the $\Gamma$ condition we can get
\[
\|h_{R,K,(*)}\|_{k_{k_3} \rightarrow k_{k_1}} \lesssim \min(N_1, N_2) \cdot (R + R^{1/2} \min(N_1, N_2)^{1/2}) N^{1/2}
\]
(note the absence of $K$ on the right hand side) if $K \leq K_1 := R^2 + R \min(N_1, N_2)$, and
\[
\|h_{R,K,(*)}\|_{k_{k_3} \rightarrow k_{k_1}} \lesssim \min(N_1, N_2) \cdot K^{1/2} N^{1/2}
\]
if $K \geq K_1$. The second bound is obviously enough for (3.18); by examining the relation between $N_1$ and $N_2$, we see that the first bound is also enough if $\max(K, R) \geq N^{\epsilon}$.

Finally, suppose $K, R \leq N^{\epsilon}$, then by losing $N^{C\epsilon}$ we may fix the values of $k_1 - k_2$ and $\Omega = |k|^2 - |k_1|^2 + |k_2|^2 - |k_3|^2$. Here we will improve the bound on the $k_{k_3} \rightarrow k_{k_1}$ norm. Namely, when $(k_1, k_2)$ is fixed, let $\ell = k_1 - k_2$ with $0 < |\ell| \leq N^{\epsilon}$, then the value of $k \cdot \ell$ is also fixed. Moreover, by the $\Gamma$ condition we know that $|k|^2$ belongs to an interval of length $O(\min(N_1, N_2))$. Once $|k|^2$ and $k \cdot \ell$ are fixed, $k$ will be determined by a lattice point on a two-dimensional ellipse of radius $O(N)$, and the number of such points is at most $N^{2/3}$ by a classical geometric argument (see for example [20], Lemma 4.1). This leads to the improved bound
\[
\|h_{R,K,(*)}\|_{k_{k_3} \rightarrow k_{k_1}} \lesssim N^{C\epsilon} \min(N_1, N_2)^{3/2} N^{1/3},
\]
which is then enough for (3.18).
6.4. The pairing case. Now we consider the pairing case where we may expand some \( w_j \) as in \( 6.2 \) and assume either \( k_1' = k_2' \) or \( k_2' = k_3' \). In this case we will use the cancellation \( 6.15 \) as before. First consider term II; there are four different cases.

6.4.1. Case (C,C,C): \( k_1' = k_2' \). Suppose each \( w_j \) has type (C) (or (G)) and assume \( k_1' = k_2' \), so in particular \( N_1 = N_2 \). Since we are considering term II, we must have \( N_1 = N_2 = N \). Then exploiting \( 6.15 \) like before, we can reduce to estimating the quantity

\[
X'_k = R^{-\beta} \sum_{(k_1,k_2,k_3)} h^{R,\ast}_{k_1k_2k_3} \cdot \sum_{N_j / 2 < |k_j'| \leq N_j} \left( \frac{1}{\langle k_1' \rangle^2} - \frac{1}{\langle k_1 \rangle^2} \right) h^{(1)}_{k_1k_1'} h^{(2)}_{k_2k_2'} h^{(3)}_{k_3k_3'} g_{k_3' k_2' k_1'}. \tag{6.22}
\]

Note that we may assume \( |k_1 - k_1'| \lesssim L_1 N^\delta \) and \( |k_2 - k_1'| \lesssim L_2 N^\delta \), and \( |k_1 - k_2| \sim R \), so at a loss of \( N^C \) we may assume \( R \lesssim \max(L_1, L_2) \), and

\[
\left| \frac{1}{\langle k_1' \rangle^2} - \frac{1}{\langle k_1 \rangle^2} \right| \lesssim N^{-3} (R + \min(L_1, L_2)).
\]

Therefore, the matrix

\[
\tilde{h}_{k_1k_2} = \sum_{k_1'} \left( \frac{1}{\langle k_1' \rangle^2} - \frac{1}{\langle k_1 \rangle^2} \right) h^{(1)}_{k_1k_1'} h^{(2)}_{k_2k_2'} \tag{6.23}
\]

is bounded (up to loss \( N^{C_\delta} \)) by

\[
\| \tilde{h} \|_{k_1k_2} \lesssim N^{-2} (R + \min(L_1, L_2)) \cdot (L_1 L_2)^{-1/2 + 3\varepsilon_1}.
\]

Note that here \( h^{(1)} \) and \( h^{(2)} \) cannot both be identity, so we may always estimate the non-identity one in the Hilbert-Schmidt \( \ell^2 \) norm. Using Propositions \( 2.6 \) and \( 2.7 \) we can estimate

\[
\| X_k \|_k \lesssim R^{-\beta} N^{-1} \| \tilde{h} \|_{k_1k_2} \cdot \| h^{(3)} \|_{k_3 \to k_3'} \cdot \max(\| h^{R,\ast} \|_{k \to k_1k_2k_3}, \| h^{R,\ast} \|_{kk_3 \to k_1k_2}).
\]

Both norms are bounded by \( NN_3 \), so

\[
\| X_k \|_k \lesssim N^{-1} R^{-\beta} (R + \min(L_1, L_2)) \cdot (L_1 L_2)^{-1/2 + 3\varepsilon_1}
\]

which is enough for \( 6.18 \).

6.4.2. Case (C,C,C): \( k_2' = k_3' \). Now suppose each \( w_j \) has type (C) (or (G)), and assume \( k_2' = k_3' \), then \( N_2 = N_3 \) and \( \max(N_1, N_2) = N \). In this case we do not need to use the cancellation \( 6.15 \). The same argument as above yields

\[
\| X_k \|_k \lesssim N_1^{-1} R^{-\beta} \| \tilde{h} \|_{k_2k_3} \| h^{(1)} \|_{k_1 \to k_1'} \cdot \max(\| h^{R,\ast} \|_{k \to k_1k_2k_3}, \| h^{R,\ast} \|_{kk_3 \to k_1k_2}).
\]

where \( \tilde{h}_{k_2k_3} \) is the matrix

\[
\tilde{h}_{k_2k_3} = \sum_{k_2'} \frac{1}{\langle k_2' \rangle^2} h^{(2)}_{k_2k_2'} h^{(3)}_{k_3k_3'}
\]

and satisfies \( \| \tilde{h} \|_{k_2k_3} \lesssim N_2^{-1} \). As both norms of \( h^{R,\ast} \) are bounded by \( R \min(N_1, N_2) \), we get that

\[
\| X_k \|_k \lesssim N^{C_\delta} \cdot (N_1 N_2)^{-1} \min(N_1, N_2)
\]

which is enough for \( 6.18 \) as \( \max(N_1, N_2) = N \).
6.4.3. Case \((C,C,L)\): \(k_1' = k_2'\). Here assume that \(w_1\) and \(w_2\) has type \((C)\) or \((G)\), \(w_3\) has type \((L)\) (or \((D)\)), and \(k_1' = k_2'\). Then we have

\[
\mathcal{X}_k = R^{-\beta} \sum_{(k_1,k_2,k_3)} h_{kk_1k_2k_3}^{R(\ast)} \cdot \sum_{N_1/2 < |k_1'| \leq N_1} \left( \frac{1}{\langle k_1' \rangle^2} - \frac{1}{\langle k_1 \rangle^2} \right) h_{k_1k_1'}^{(1)} \overline{h}_{k_2k_2'}^{(2)} (w_3)_{k_3}.
\]

(6.24)

Hence we can easily estimate

\[
\|\mathcal{X}_k\|_k \lesssim R^{-\beta} \|\overline{h}\|_{k_1k_2} \|w_3\|_{k_3} \cdot \|h^{R(\ast)}\|_{k' \rightarrow k_1k_2k_3}
\]

where \(\overline{h}\) is defined as in (6.23). This yields

\[
\|\mathcal{X}_k\|_k \lesssim N^{-1}(R + \min(L_1,L_2))(L_1L_2)^{-1/2+3\epsilon_1} R^{-\beta} N_3^{-1/2+\epsilon_1+\epsilon_2} \min(R,N_3)
\]

as \(N_1 = N_2 = N\), using also Lemma 2.4. Since \(R \lesssim \max(L_1,L_2)\), by considering the relative sizes between \(R\) and \(\min(L_1,L_2)\) we can check that this term is always bounded by \(N^{-1+C\epsilon_1}\), which is enough for (3.18).

6.4.4. Case \((L,C,C)\): \(k_2' = k_3'\). Here assume that \(w_2\) and \(w_3\) has type \((C)\) or \((G)\), \(w_1\) has type \((L)\) (or \((D)\)), and \(k_2' = k_3'\); then \(N_2 = N_3\) and \(\max(N_1,N_2) = N\). In this case we will need to use the cancellation (3.15). Like before we can reduce to estimating the quantity

\[
\mathcal{X}_k = R^{-\beta} \sum_{(k_1,k_2,k_3)} h_{kk_1k_2k_3}^{R(\ast)} \cdot (w_1)_{k_1} \sum_{N_2/2 < |k_2'| \leq N_2} \left( \frac{1}{\langle k_2' \rangle^2} - \frac{1}{\langle k_2 \rangle^2} \right) h_{k_2k_2'}^{(2)} \overline{h}_{k_3k_3'}^{(3)}.
\]

(6.25)

Denote

\[
\overline{h}_{k_2k_3} = \sum_{N_2/2 < |k_2'| \leq N_2} \left( \frac{1}{\langle k_2' \rangle^2} - \frac{1}{\langle k_2 \rangle^2} \right) h_{k_2k_2'}^{(2)} \overline{h}_{k_3k_3'}^{(3)},
\]

then similarly we have

\[
\|\overline{h}\|_{k_2k_3} \lesssim N^{-2} \max(L_2,L_3) \cdot N_2(L_2L_3)^{-1/2+3\epsilon_1} \lesssim N^{-3/2+C\epsilon_1},
\]

hence

\[
\|\mathcal{X}_k\|_k \lesssim N_1^{-1/2+\epsilon_1+\epsilon_2} N_2^{-3/2+C\epsilon_1} R^{-\beta} \|h^{R(\ast)}\|_{k' \rightarrow k_1k_2k_3}.
\]

Using that

\[
\|h^{R(\ast)}\|_{k' \rightarrow k_1k_2k_3} \lesssim \min(N_2,R) \cdot \min(N_1,N_2)
\]

and that \(R \geq N\), by considering the relative size between \(N_1, N_2\) and \(R\), it is easy to check that this bound is enough for (3.18) when \(\max(N_1,N_2) = N\).

6.4.5. The Gamma condition term. Finally we consider terms III and IV with pairing. Note that as in Section 6.3 we may assume \(N_3 \sim N\) and \(\max(N_1,N_2) \ll N\), hence the only possibility of pairing is \(k_1' = k_2'\) (so \(N_1 = N_2\)). Moreover if \(w_3\) has type \((L)\) or \((D)\) the proof can be done as in Section 6.3 above, so we only need to consider the case of type \((C,C,C)\) and \(k_1' = k_2'\) (in particular \(N_1 = N_2\)). Like in Section 6.3 we can reduce to the quantity

\[
\mathcal{X}_k = R^{-\beta} K^{-b} \sum_{(k_1,k_2,k_3)} h_{kk_1k_2k_3}^{R,K(\ast)} \sum_{k_3} \overline{h}_{k_1k_2k_3} \overline{h}_{k_3k_3'}^{(3)} (F_{N_3})_{k_3'}
\]

where \(\overline{h}\) is defined as in (6.23). Using Propositions 2.6 and 2.7 we get

\[
\|\mathcal{X}_k\|_k \lesssim R^{-\beta} K^{-b} N^{-1} \|\overline{h}\|_{k_1k_2} \|\overline{h}^{(3)}\|_{k_3} \cdot \max(\|h^{R,K(\ast)}\|_{k' \rightarrow k_1k_2k_3}, \|h^{R,K(\ast)}\|_{k_3 \rightarrow k_1k_2}).
\]
The \( k \to k_1 k_2 \tilde{k}_3 \) norm can be bounded by \( K^{1/2} R N_1 \) which is clearly enough for \( 3.18 \); for the \( k \tilde{k}_3 \to k_1 k_2 \) norm we use the bound obtained in Section 6.3.4 to get
\[
\| h^{R,K, \ell(\cdot)} \|_{k \tilde{k}_3 \to k_1 k_2} \lesssim R^{1/2} N_1^{3/2} N^{1/2} K^{1/2}
\]
hence
\[
\| X_k \|_k \lesssim R^{-\beta} K^{-b} N^{-1} \cdot N_1^{-2} (R + \min(L_1, L_2))(L_1 L_2)^{-1/2+3\varepsilon_1} \cdot R^{1/2} N_1^{3/2} N^{1/2} K^{1/2}
\]
with a possible loss of \( N^C\delta \), which is enough for \( 3.18 \) by considering the relative size between \( R \) and \( \min(L_1, L_2) \), using also that \( R \lesssim \max(L_1, L_2) \).

6.5. **Term VIII.** Finally, we will estimate term VIII, which is the last two lines of (3.13). This is an easier term and most part of this term can be estimated using similar arguments as the above proof, so we will not detail them out. In fact, this term can be decomposed into expressions in the form
\[
(VIII)_k(t) = -i \int_0^t \sum_{\ell} V_{k-\ell}(w_1)_k(t')(w_2)_\ell(t')(w_3)_\ell(t').
\]
(6.26)
Here, if \( w_2 \) and \( w_3 \) both have type (C) or (G), then we can expand them as in (6.2) and exploit the independence if \( k'_2 \neq k'_3 \), and exploit the cancellation (3.15) if \( k'_2 = k'_3 \) (note that here \( k_2 = k_3 = \ell \), so the right hand side of (3.15) is in fact 1 instead of 0, but this cancels with the term \(-1/(\ell)^2 \) which is subtracted in (3.13); this is also the reason why the renormalization term), and the rest of proof can go just like before.

The hardest term in VIII in fact is the term where \( w_1 \) has type (G), one of \( w_2 \) and \( w_3 \) has type (G), and the other has type (L) or (D) in (6.20). For such terms, standard estimates will fail short by a power \( N^{1-\beta} \) as \( \beta < 1 \); however since \( 1 - \beta \ll \varepsilon_2 \) by our choice, this can be controlled if we gain a power \( N^{\varepsilon_2/10} \) from elsewhere. If either \( w_2 \) or \( w_3 \) has type (L), then we can plug in the equation satisfies by \( \rho^N \) and estimate like Section 5 to gain this extra power.

So the only bad term is when \( w_1 \) has type (G), and when one of \( w_2 \) and \( w_3 \) has type (G), and the other has type (D). Let this term be \( z_N \), then \( z_N - z_N^* \) satisfies (3.18); due to the symmetry between \( w_2 \) and \( w_3 \) in (6.20), we see that \( (z_N^*)_k \in (-i g_k) \cdot \mathbb{R} \). Then, if we replace the type (D) term (which is \( z_{N'} \) for some \( N' \)) with \( z_N - z_N^* \), the resulting contribution will satisfy (3.18), while if we replace this term by \( z_N^* \), the net contribution to term VIII, after exploiting symmetry between \( w_2 \) and \( w_3 \), will be
\[
\text{Re}(\langle \bar{g}_k \cdot (z_N^*)_\ell(t') \rangle) = 0.
\]
Therefore, in any case, we can control this term by (3.18). This finishes the proof of Proposition 3.1 and hence Theorem 1.3.

**References**

[1] M. Aizenman. Geometric analysis of \( \Phi^4 \) fields and Ising models. Part I and II. Comm. Math. Phys. 86 (1982), issue 1, 1–48.
[2] S. Albeverio, A. Cruzeiro. Global flows with invariant (Gibbs) measures for Euler and Navier-Stokes two dimensional fluids. Comm. Math. Phys. 129 (1990) 431–444.
[3] N. Barashkov and M. Gubinelli. A variational method for \( \Phi_4^4 \). Duke Math. J. 169 (2020), no. 17, 3339–3415.
[4] N. Barashkov and M. Gubinelli. The \( \Phi_4^4 \) measure via Girsanov’s theorem. arXiv preprint 2004.01513 (2020).

---

\(^1\) There is also a term which is essentially \( \mathcal{M}^{N,N'} \) applied to \( z_N \), which can also be treated using the Hilbert-Schmidt bound for the matrices \( \mathcal{M}^{N,N'} \). We omit the details.
[5] J. Bourgain. Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys.* 166 (1994), 1–26.

[6] J. Bourgain. Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. *Comm. Math. Phys.* 176 (1996), 421–445.

[7] J. Bourgain. Invariant measures for the Gross-Piatevskii equation. *J. Math. Pures Appl.* (9) 76 (1997), no. 8, 649–702.

[8] J. Bourgain and A. Bulut. Almost sure global well posedness for the radial nonlinear Schrödinger equation on the unit ball I: the 2D case. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 31 (2014), no. 6, 1267–1288.

[9] 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 (2014), no. 6, 1289–1325.

[10] J. Bourgain and C. Demeter. The proof of the $l^2$ decoupling conjecture. *Ann. of Math.* (2) 182 (2015), no. 1, 351–389.

[11] B. Bringmann. Almost Sure Local Well-Posedness for a Derivative Nonlinear Wave Equation. *Int. Math. Res. Not. (IMRN)*, rnz385, https://doi.org/10.1093/imrn/rnz385.

[12] B. Bringmann. Invariant Gibbs measures for the three-dimensional wave equation with a Hartree nonlinearity I: Measures. *arXiv preprint* 2009.04609.

[13] B. Bringmann. Invariant Gibbs measures for the three-dimensional wave equation with a Hartree nonlinearity II: Dynamics. *arXiv preprint* 2009.04616.

[14] N. Burq, L. Thomann and N. Tzvetkov. Remarks on the Gibbs measures for nonlinear dispersive equations. *Annales de la Faculté des sciences de Toulouse: Mathématiques*, Série 6, Tome 27 (2018) no. 3, 527–597.

[15] N. Burq and N. Tzvetkov. Random data Cauchy theory for supercritical wave equations. II. A global existence result. *Invent. Math.* 173 (2008) no. 3, 477–496.

[16] G. Da Prato and A. Debussche. Two-dimensional Navier-Stokes equations driven by a space-time white noise. *J. Funct. Anal.* 196 (2002), no. 1, 180–210.

[17] Y. Deng, Andrea R. Nahmod and H. Yue. Optimal local well-posedness for the periodic derivative nonlinear Schrödinger equation. To appear in *Comm. Math. Phys.* (2020).

[18] Y. Deng, A. Nahmod and H. Yue. Invariant Gibbs measures and global strong solutions for nonlinear Schrödinger equations in dimension two. *arXiv preprint* 1910.08492.

[19] Y. Deng, A. Nahmod and H. Yue. Random tensors, propagation of randomness, and nonlinear dispersive equations. *arXiv preprint* 2006.09285.

[20] C. Fan, Y. Ou, G. Staffilani and H. Wang. 2D-Defocusing nonlinear Schrödinger equation with random data on irrational tori. *arXiv preprint* 1910.03199.

[21] J. Fröhlich. On the triviality of $\lambda \Phi^4_d$ theories and the approach to the critical point in $d_{\text{cr}} > 4$ dimensions. *Nuclear Physics B* 200 (1982), issue 2, 281–296.

[22] J. Fröhlich, A. Knowles, B. Schlein and V. Sohinger. Gibbs measures of nonlinear Schrödinger equations as limits of many-body quantum states in dimensions $d \leq 3$. *Comm. Math. Phys.* 356 (2017), no. 3, 883–980.

[23] J. Glimm and A. Jaffe. *Quantum physics, A functional integral point of view*, Second edition, Springer-Verlag, New York, 1987. xxii+535 pp.

[24] M. Gubinelli, P. Imkeller and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math Pi* 3 (2015), e6, 75 pp

[25] M. Hairer. A theory of regularity structures. *Invent. Math.* 198 (2014), no. 2, 269–504.

[26] M. Gubinelli, H. Koch and T. Oh. Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. *arXiv preprint* 1811.07808.

[27] J. Lebowitz, R. Rose and E. Speer. Statistical mechanics of the nonlinear Schrödinger equation. *J. Statist. Phys.* 50 (1988), 657–687.

[28] E. Nelson. Construction of quantum fields from Markoff fields, *J. Functional Analysis* 12 (1973), 97–112.

[29] T. Oh, M. Okamoto, L. Tolomeo. Focusing $\Phi^4_d$-model with a Hartree-type nonlinearity. *arXiv preprint* 2009.03251.

[30] T. Oh and L. Thomann. Invariant Gibbs measures for the 2-d defocusing nonlinear wave equations. *Ann. Fac. Sci. Toulouse Math.* (6) 29 (2020), no. 1, 1–26.

[31] 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 (2018), no. 3, 397–445.
[32] B. Simon. The $P(\phi)_2$ Euclidean (quantum) field theory, Princeton Series in Physics. Princeton University Press, Princeton, N.J., 1974. xx+392 pp.

[33] V. Sohinger. A microscopic derivation of Gibbs measures for nonlinear Schrödinger equations with unbounded interaction potentials, arXiv preprint 1904.08137 (2019).

[34] C. Sun and N. Tzvetkov. Refined probabilistic global well-posedness for the weakly dispersive NLS. arXiv preprint 2010.13065. (2020).

[35] N. Tzvetkov. Invariant measures for the nonlinear Schrödinger equation on the disk. Dyn. Partial Diff. Equ. 3 (2006), no. 2, 111–160

[36] N. Tzvetkov. Invariant measures for the defocusing nonlinear Schrödinger equation. Ann. Inst. Fourier (Grenoble) 58 (2008), no. 7, 2543–2604.

1 Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA
Email address: yudeng@usc.edu

2 Department of Mathematics, University of Massachusetts, Amherst MA 01003
Email address: nahmod@math.umass.edu

3 Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA
Email address: haitiyan@usc.edu