ON THE DERIVATION OF THE HOMOGENEOUS KINETIC WAVE EQUATION

CHARLES COLLOT AND PIERRE GERMAIN

ABSTRACT. The nonlinear Schrödinger equation in the weakly nonlinear regime with random Gaussian fields as initial data is considered. The problem is set on the torus in any dimension greater than two. A conjecture in statistical physics is that there exists a kinetic time scale depending on the frequency localisation of the data and on the strength of the nonlinearity, on which the expectation of the squares of moduli of Fourier modes evolve according to an effective equation: the so-called kinetic wave equation. When the kinetic time for our setup is $1$, we prove this conjecture up to an arbitrarily small polynomial loss. When the kinetic time is larger than $1$, we obtain its validity on a more restricted time scale. The key idea of the proof is the use of Feynman interaction diagrams both in the construction of an approximate solution and in the study of its nonlinear stability. We perform a truncated series expansion in the initial data, and obtain bounds in average in various function spaces for its elements. The linearised dynamics then involves a linear Schrödinger equation with a corresponding random potential. We bound the expectation of the operator norm in Bourgain spaces using diagrams and random matrix tools. This gives a new approach for the analysis of nonlinear wave equations out of equilibrium, and gives hope that refinements of the method could help settle the conjecture.

CONTENTS

1. Introduction 2
1.1. Presentation of the problem 2
1.2. Relevant time scales, parameters range 3
1.3. Rescaling to a large torus 3
1.4. Background 4
1.5. Main result 4
1.6. Strategy of the proof and plan of the article 5
2. Notations 8
2.1. Time range 8
2.2. Inequalities 8
2.3. Fourier transform 8
2.4. Lebesgue spaces 8
2.5. Bourgain spaces 8
2.6. Randomization 9
3. Proof of the main theorem 9
3.1. The frequency truncation for Wick ordering 9
3.2. Approximate solution and error 9
3.3. Bounds on the expansion 10
3.4. Bounds on the linear, bilinear and trilinear terms 10
3.5. Control of the error, proof of the first part of Theorem 1.1 11
3.6. Comparison to the kinetic wave equation, proof of the second part of Theorem 1.1 11

2010 Mathematics Subject Classification. primary, 35Q55 70K70 35C20, secondary 35B34 81Q30.
Key words and phrases. Nonlinear Schrödinger equation, random initial data, kinetic wave equation, Feynman graphs.
1. Introduction

1.1. Presentation of the problem. We consider

\[ \begin{cases} \partial_t u + \Delta u = \lambda^2 |u|^2 u \\ u(t = 0) = u_0. \end{cases} \]  

(NLS)

set on the torus: \( x \in \mathbb{T}^d = \mathbb{R}^d/(2\pi \mathbb{Z}^d) \), where \( d \geq 2 \), and with initial data

\[ u_0(x) = \epsilon^{d/2} \sum_{k \in \mathbb{Z}^2} A(\epsilon k)G(k), \]  

(1.1)

where \( A \in C_0^\infty(\mathbb{R}^2, \mathbb{R}) \) and \( (G(k))_{k \in \mathbb{Z}^d} \) are independent standard centred complex Gaussians (see Section 2). The normalization is such that on average \( \|u_0\|_{L^2} \sim 1 \), and, by Khinchine’s inequality, \( \|u_0\|_{L^p} \sim 1 \) on average as well, for any \( p < \infty \).

Heuristic derivations, to which we will come back, show that, as \( \epsilon \lambda \to 0 \) (weakly nonlinear regime) and \( \epsilon \to 0 \) (high frequency limit),

\[ e^{-dE} \left| \hat{u}_k \left( \frac{t}{T_{\text{kin}}} \right) \right|^2 \to \rho(t, k), \quad \text{for } T_{\text{kin}} = \frac{1}{\epsilon^2 \lambda^4} \]  

(1.2)

where \( \rho \) solves the kinetic wave equation

\[ \begin{cases} \partial_t \rho(t, k) = C[\rho](k) \\ \rho(0, k) = |A(k)|^2. \end{cases} \]  

(KWE)
with the collision operator given by
\[ C[r](k) = \int_{\mathbb{R}^d} \delta(k + \ell - m - n) \delta(|k|^2 + |\ell|^2 - |m|^2 - |n|^2) \]
\[ \rho(k) \rho(\ell) \rho(m) \rho(n) \left[ \frac{\rho(k)}{\rho(\ell)} + \frac{1}{\rho(\ell)} - \frac{1}{\rho(m)} - \frac{1}{\rho(n)} \right] d\ell dm dn. \] (1.3)

The present paper is an attempt to show that the prediction (1.2) is verified on non-trivial time scales, even though the kinetic time scale \( T_{kin} \) is for the moment out of reach.

### 1.2. Relevant time scales, parameters range.

- \( T_{kin} = \frac{1}{\epsilon^2 \lambda^4} \): characteristic time scale for the kinetic wave equation.
- \( T_{lin} = \epsilon^2 \): characteristic time scale for the linear part of the equation.
- \( T_{nonlin} = \frac{1}{\lambda^2} \): characteristic time-scale for the nonlinear part of the equation (given that we expect \( \|u\|_{L^\infty} \sim 1 \) - up to logarithmic losses).

We will consider the regime where
\[ T_{lin} \ll T_{kin} \iff T_{nonlin} \ll T_{kin} \iff T_{lin} \ll T_{nonlin} \iff \lambda \epsilon \ll 1, \]

which means that the regime we are considering is weakly nonlinear. In particular, the linear time is much smaller than the deterministic nonlinear time that is in turn much smaller than the kinetic time:
\[ T_{lin} \ll T_{nonlin} \ll T_{kin}. \]

We consider a power type relation between the strength of the nonlinearity and the frequency
\[ \lambda = \epsilon^{-\gamma} \]
and choose
\[ 0 < \gamma < \frac{1}{2} \]
so that:
\[ T_{nonlin} < 1 < T_{kin}. \]

Finally, let us consider resonances. The resonance modulus is classically given by
\[ \Omega(k, \ell, m, n) = |k|^2 - |\ell|^2 + |m|^2 - |n|^2. \]

Since \( k, \ell, m, n \in \mathbb{Z}^d \), \( \Omega \) takes integer values; and since \( |k|, |\ell|, |m|, |n| \lesssim \frac{1}{\epsilon} \), it satisfies \( |\Omega| \lesssim \frac{1}{\epsilon^2} \). This means that only time scales \( T \) such that
\[ \epsilon^2 \ll T \ll 1 \] (1.4)
are susceptible to yield the kinetic wave equation in an asymptotic regime. Indeed, if \( T \lesssim \epsilon^2 \), resonances are hardly playing any role; while if \( T \gtrsim 1 \), the resonance moduli \( \Omega \) cannot be equidistributed modulo \( \frac{1}{T} \), which prevents from taking the discrete to continuous limit in frequency.

### 1.3. Rescaling to a large torus.

In the problem formulated above, the equation is set on a torus of size 1, and \( u_0 \) has size \( \sim 1 \) in any \( L^p \) (neglecting logarithmic factors if \( p = \infty \)), and varies on a typical scale \( \sim \epsilon \). For the reader’s convenience, we show how it can be rescaled to fit the setup adopted in [12, 7, 8].

We now let \( \epsilon = L^{-1} \), and rescale \( u \) by setting
\[ u'(t', x') = u' \left( \frac{t}{\epsilon^2}, \frac{x}{\epsilon} \right) = \epsilon^{d/2} u(t, x). \]

The equation solved by \( U \) is now
\[ i \partial_t u' + \Delta_x u' = (\lambda')^2 |u'|^2 u' \quad \text{with} \quad \lambda' = \lambda \epsilon^{1-d/2}. \] (NLS2)
In this new setting, the domain has size $L$, $u'$ is of size $\sim L^d \frac{d}{d'}$ in $L^p$, and varies on a typical scale $\sim 1$. These orders of magnitude coincide with the framework adopted in \[.\]

To convert results from (NLS) to (NLS2), observe that the time scale $t_0$ for (NLS) corresponds to $t_0' = \frac{t_0}{c}$ for (NLS2). In particular, the kinetic time scale $T_{\text{kin}} = \frac{1}{\lambda^d}$ for (NLS) corresponds to the kinetic time scale $T'_{\text{kin}} = \frac{1}{(\lambda')^{d'}} = \frac{t^{d'}}{(\lambda')^d}$ for (NLS2).

1.4. Background. The kinetic wave equation was first derived by Peierls \[22\] in the context of Quantum Mechanics, and, in a different form and independently by Hasselmann \[15\] \[16\], who worked on water waves. The theory was then revived by Zakharov and his collaborators \[27\], giving a very versatile framework, which applies to a number of Hamiltonian systems satisfying the basic assumptions of weak nonlinearity, high frequency (or infinite volume limit), and phase randomness. Introductions to this research field can be found in Nazarenko \[20\] and Newell-Rumpf \[21\].

As far as rigorous mathematics go, a fundamental work is due to Lukkarinen and Spohn \[18\], who tackled in \[8\], which is the source of a number of ideas which we extend here (construction of an approximate solution and control of the remainder in Bourgain spaces instead of Strichartz spaces).

The question of the derivation of (KWE) for random data out of statistical equilibrium was first reached in \[2\], \[3\] \[8\] \[10\] \[11\], and in another form and independently by Hasselmann \[15\] \[16\], who worked on water waves. The theory was then revived by Zakharov and his collaborators \[27\], giving a very versatile framework, which applies to a number of Hamiltonian systems satisfying the basic assumptions of weak nonlinearity, high frequency (or infinite volume limit), and phase randomness. Introductions to this research field can be found in Nazarenko \[20\] and Newell-Rumpf \[21\].

The kinetic wave equation is to phonons, or linear waves, what the Boltzmann equation is to classical particles. The derivation of the Boltzmann equation was put on a rigorous mathematical foundation with the foundational work of Lanford \[17\] and its more recent clarification by Gallagher-Saint-Raymond-Texier \[13\]. A few articles deal with the derivation of kinetic models for quantum particles \[1\] \[2\] \[3\] \[12\]: this question is closely related to the derivation of the kinetic wave equation, but is harder, since (NLS) can be thought of as an intermediary step between a quantum mechanical model with a large number of particles, and kinetic theory.

Another strand of research focuses on linear dispersive models with random potential, from which one can derive the linear Boltzmann equation on a short time scale \[23\], and the heat equation on a longer time scale \[10\] \[11\].

Finally, let us mention the possibility of deriving Hamiltonian models for NLS with deterministic data spreading in the infinite volume, or big box, limit \[12\] \[7\].

1.5. Main result. Our main result is the existence of a solution over the nontrivial time range $[0, 1]$, and the validity of the approximation by the kinetic wave equation on this time interval. As $\gamma$ approaches $1/2$ we get close, up to an arbitrarily small polynomial polynomial loss, from the kinetic time. The existence of (NLS) up to time 1 (on the complement of an exceptional set) is non trivial, and given by the first part of the Theorem; as for the local well-posedness of (KWE), it has been established for instance in \[14\].

**Theorem 1.1.** Pick any $0 < \gamma < \frac{1}{2}$ and $\eta > 0$. Then there exist universal constants $c, C > 0$ independent of $\gamma$ and $\eta$, $0 < \epsilon^*(\gamma, \eta) \le 1$ and $\delta(\gamma, \eta) > 0$ such that for all $0 < \epsilon \le \epsilon^*$, a set $\epsilon E = E_{\eta, \gamma, \epsilon}$ of measure $\mathbb{P}(E) > 1 - C\epsilon^{\eta_2}$ exists such that

- Existence of solution: On $E$, the equation (NLS) with initial datum (1.1) admits on $[0, 1]$ a solution $u \in C^\infty$, which is furthermore $O(\epsilon^{C\eta})$ in any $L^p_{[0, 1]} L^q_{T, 2d}$.

- Validity of kinetic wave equation: Furthermore, denoting $\mathcal{A} = |A|^2$, for any $\lambda^{-2} \le t \le 1$:

$$\sum_{k \in \mathbb{Z}^d} \left| \mathbb{E} \left[ 1_E \left| \hat{u}_k(1) \right|^2 - \frac{t}{T_{\text{kin}}} C(A) \right] \right| \lesssim \frac{t}{T_{\text{kin}}} \left| t + \epsilon^\delta \right|,$$

where $C$ is the collision operator defined in (1.3).
Remark 1.2. On the range of parameters:

- The condition \( \gamma > 0 \) gives \( T_{\text{nonlin}} = \lambda^{-2} \leq 1 \), so that the time interval \([0,1]\) exceeds the nonlinear time and the existence result is nontrivial. The condition \( \lambda^{-2} \leq t \leq 1 \) in (1.3) is due to the following. Before the nonlinear time \( \lambda^{-2} \), the nonlinear effects did not kick in, so that the kinetic wave equation is irrelevant on \([0,\lambda^{-2}]\). For \( t \) of order 1, the resonance moduli are not equidistributed and the underlying oscillatory integrals degenerate so that one needs \( t \ll 1 \) in (1.3) to obtain that the right hand side is of lower order compared to the size of the correction \( \frac{1}{T_{\text{kin}}}C(A) \) in the left hand side.

- The above shows that the prediction (1.2) is satisfied on time scales \( \ll 1 \); namely, the expectation of \( e^{-d}|\hat{u}_k|^2 \) and the solution of (KWE) are as close as expected. In particular, we are able to treat kinetic time scale of order \( e^{-\mu} \) for \( \mu > 0 \) as close as we want to 0 (that is, for \( 0 < \frac{1}{2} - \gamma \ll 1 \)). This shows how \( T_{\text{kin}} = 1 \) is attainable with arbitrarily small polynomial loss.

- For a similar reason of equidistribution for the dispersion relation of the Schrödinger equation, we cannot control the solution on a time interval \([0,T]\) with \( T \gg 1 \).

- In the case \( T_{\text{kin}} \leq 1 \), the bounds we obtain for the approximate solution are powers of \( \frac{1}{T_{\text{kin}}} \) due to degenerate low frequencies effects. This prevents both the convergence of our series expansion and its linear stability, so that we cannot cover this case.

Remark 1.3. Relaxing the set up: Many assumptions in the above theorem could be relaxed. The randomization through Gaussians is convenient since it allows the use of Wick’s formula, but other randomizations should also be possible. The function \( A \) was taken in \( C_0^\infty \) to simplify the proof as much as possible, but much milder hypotheses should suffice. Other dispersion relations could be considered as well, and if particular under better equidistribution properties our present analysis could be strengthen and intervals of time \([0,T]\) with \( T \gg 1 \) could be considered. The torus was chosen to be rational, but our proof applies verbatim to irrational tori - thanks to the work of Bourgain and Demeter [6], which gives Strichartz estimates for them. It should be possible to make the size of the exceptional set exponentially small in \( \epsilon \). Finally, as should be expected, whether the equation is focusing or defocusing does not change anything in our argument.

1.6. Strategy of the proof and plan of the article. We present here a caricature of the proof, describing only the main ideas. In particular, we simplify formulas by omitting the less important terms.

Approximation and error: The heart of the proof is to build a sufficiently good approximation of the solution. It is obtained by renormalising the phase (Wick ordering) and iterating Duhamel’s formula, yielding a truncated series expansion. Wick ordering can be explained formally as follows. Assuming formally that \( u \) remains Gaussian, one has that \( \mathbb{E}(\langle |u|^2 u, v \rangle) \approx 2\mathbb{E}(\langle |u|^2 L^2 u, v \rangle) \) for \( v \) a test Gaussian field via Wick formula (2.3). This and mass conservation \( \|u\|_{L^2} = \|u_0\|_{L^2} \) produces the first approximation \( iu_t + \Delta u = 2\lambda^2 \|u_0\|^2_{L^2} u \) for which the sole effect is a phase modulation of factor \( e^{-it\lambda^2\|u_0\|^2_{L^2}} \). This in particular does not change the statistical properties of the solution. We therefore first renormalise the solution by the phase \( e^{-it\lambda^2\|u_0\|^2_{L^2}} \). Forgetting about this phase in what follows, we define \( u^0 = e^{it\Delta}u_0 \) and, for \( n \geq 1 \),

\[
\begin{align*}
\left\{ \begin{array}{l}
i\partial_t u^n + \Delta u^n = \sum_{i+j+k=n-\lambda} P(u^i, \overline{u}^j, u^k) \\
u(t=0) = u_0,
\end{array} \right.
\end{align*}
\]

Here \( P \) is a trilinear operator which cancels the worst interactions responsible for Wick ordering (from the point of view of resonances and graph analysis, see below). The solution splits into
approximate solution and error:

$$u = u^{app} + u^{err}, \quad \text{with} \quad u^{app} = \sum_{n=0}^{N} u^n.$$  

The approximate solution satisfies the equation up to an error $\mathcal{E}$

$$i\partial_t u^{app} + \Delta u^{app} = |u^{app}|^2 u^{app} + \mathcal{E},$$

and the equation satisfied by $u^{err}$ is then

$$i\partial_t u^{err} + \Delta u^{err} = 2|u^{app}|^2 u^{err} + (u^{app})^2 u^{err} + \mathcal{L}(u^{err}) + \mathcal{B}(u^{err}) + \mathcal{T}(u^{err}) + 2|u^{err}|^2 u^{err},$$

where $\mathcal{L}$, $\mathcal{B}$, and $\mathcal{T}$ stand for the linear, bilinear, and trilinear terms respectively.

There remains to apply a fixed point argument in a (slightly modified) Bourgain space $X^{s,b}$ in order to show that $u^{err}$ is sufficiently small. In order to carry out this plan, we need bounds on $u^{app}$, $\mathcal{L}$, $\mathcal{B}$ and $\mathcal{T}$. Obtaining these bounds is the aim of the remainder of the paper.

Feynman diagrams A formula for $u^1$ is easy to write: in physical space

$$u^1 = i\lambda^2 \int_0^t e^{it-s} \Delta P(e^{it}u^0, e^{it}u^0, e^{it}u^0) \chi ds,$$

and in frequency

$$\hat{u}^1(k) = i\lambda^2 e^{\frac{4d}{2\pi}} e^{-|k|^2} \sum_{k=|k_0,1|+|k_0,2|+|k_0,3|} e^{i\Omega} G(-k_0,1) G(k_0,2) G(k_0,3) A(-\epsilon k_0,1) A(-\epsilon k_0,2) A(\epsilon k_0,3) \chi ds,$$

where $\chi$ is the symbol of $P$, and $\Omega = |k|^2 - |k_0,1|^2 + |k_0,1|^2 - |k_0,1|^2$. However, writing the formula giving $u^3$ or $u^4$ is already extremely lengthy. This motivates the introduction of Feynman diagrams, which provide an intuitive and analytically efficient way of representing iterates. For instance, $u^1$ is represented by

We refer to Section 4 for a full presentation of Feynman diagrams.

Bound on the iterates In order to understand the approximate solution, we are led to estimating

$$E(\|u^n\|_{L^2}), \quad \text{and more generally} \quad E(\|u^n\|_{L^p}), E(\|u^n\|^2_{X^{s,b}})$$

(expectations are always taken with respect to the randomization of the data; an application of the Bienaymé-Chebyshev inequality allows to transfer these bounds to most data, up to a small loss,
and excluding an exceptional set). Applying Wick’s formula (2.3) given the choice of the data (1.1), it appears that taking the expectation forces initial frequencies to be pairwise equal. At the level of Feynman diagrams, this can be represented intuitively as follows

Root vertices are identified

Using this representation, and some graph combinatorics, we can show that $\mathbb{E}(\|u^n\|_{L^2}^2)$ is bounded by a factor times $\left(1 + \frac{t}{T_{\text{kin}}^2}\right)^n$. Though the factor depends on $n$, this is indicative of the fact that the series $u^n$ converges on the right time-scale. This is achieved in Section 5.

Bound on the error $\mathcal{E}$ As for the error $\mathcal{E}$, it is bounded by a factor times $\left(1 + \frac{t}{T_{\text{kin}}^2}\right)^{\frac{N^2}{2}}$ which is arbitrarily small by taking $N$ arbitrarily large.

Bound on the linear term $\mathcal{L}$ Turning to the linear term $\mathcal{L}$, we need to show that $\int_0^t e^{i(t-s)\Delta} L \, ds$ has operator norm $\ll 1$ in $X^{s,b}$, so that it can be absorbed in a Neumann series type argument. By the theory of $X^{s,b}$ spaces, 

$$
\left\| \int_0^t e^{i(t-s)\Delta} L \, ds \right\|_{X^{s,b} \to X^{s,b}} \lesssim \|L\|_{X^{s,b} \to X^{s,b-1}},
$$

so it suffices to focus on the operator norm of $\mathcal{L}$ from $X^{s,b}$ to $X^{s,b-1}$. Using that $v^{app}$ is supported in Fourier on a ball of radius $C\epsilon^{-1}$, it suffices to bound $\mathcal{L}$ when it is furthermore localized on cubes of comparable size.

The idea is now to view $\mathcal{L}$ as a random operator, and rely on a classical trick in random matrix theory. First, consider $\mathcal{L}^* \mathcal{L}$, which is positive, and thus has its operator norm bounded by its trace. Since this operator is self-adjoint, we can estimate its norm by raising it to a high power, and taking the trace. Finally, taking in addition the expectation, we obtain that

$$
\mathbb{E}\|\mathcal{L}\| \lesssim \left[ \mathbb{E} \text{Tr}(\mathcal{L}^* \mathcal{L})^N \right]^{\frac{1}{2N}}.
$$

The key is now that $\mathbb{E}\|((\mathcal{L}^* \mathcal{L})^N)_{X^{s,b} \to X^{s,b}}$ can be represented through Feynman graphs, and estimated using the same tools as for the quantities we already discussed. This estimate is performed in Section 6.

Bound on the nonlinear terms $\mathcal{B}$ and $\mathcal{T}$ It is achieved by using the classical nonlinear theory of $X^{s,b}$ spaces in Section 7.
2. Notations

2.1. Time range. Throughout the rest of this paper, we assume that we study a solution over a
time interval $[0, T]$ with:
\[ \epsilon^c \leq T \lesssim 1. \tag{2.1} \]

2.2. Inequalities. For any two quantities $X$ and $Y$, we denote $X \lesssim Y$ if there exists a universal
constant $C$ such that $X \leq CY$. We denote $X \lesssim Z Y$ if the constant $C$ is allowed to depend on a
further quantity $Z$. To keep notations under control, we do not systematically record the dependence
of all the constants; in particular, we always omit the dependence on obvious quantities, such as
the dimension $d$, or the size of the support of $A$ through which the data is defined.

Most estimates are valid up to subpolynomial factors in $\epsilon$. This is recorded by a small constant
$\kappa > 0$, for instance we denote $X \lesssim \kappa \epsilon^{-\kappa} Y$. For ease of notation, we allow the value of kappa to change from one line to the next (provided it
can always be taken arbitrarily small), and we sometimes denote $\lesssim$ instead of $\lesssim \kappa$.

Due to the assumption (2.1), subpolynomial factors in $T$ give subpolynomial factors in $\epsilon$.

2.3. Fourier transform. Given a function $f(x)$ on the torus $T^d = \mathbb{R}^d / (2\pi \mathbb{Z}^d)$, the Fourier trans-
mform in space is denoted by
\[ \mathcal{F}(f)_k = \hat{f}_k = \frac{1}{(2\pi)^{d/2}} \int_{T^d} f(x) e^{-ik \cdot x} \, dx, \quad f(x) = \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathbb{Z}^d} \hat{f}_k e^{ik \cdot x}. \]

Given a function $g(t,x)$ on $\mathbb{R} \times T^d$, the Fourier transform in space-time is
\[ \mathcal{F}(g)(\tau, k) = \frac{1}{(2\pi)^{d+1/2}} \int_{\mathbb{R}} \int_{T^d} g(t,x) e^{-i(\tau t + k \cdot x)} \, dx \, dt, \]
\[ g(t,x) = \frac{1}{(2\pi)^{d+1/2}} \sum_{k \in \mathbb{Z}^d} \int_{\mathbb{R}} \mathcal{F}(g)(\tau, k) e^{2\pi i(\tau t + k \cdot x)} \, d\tau, \]

The Fourier multiplier $m(D)$ acts only in the space variable, through the definition
\[ m(D)f = \mathcal{F}^{-1}[m(k)\hat{f}_k]. \]

For $\epsilon > 0$, $N \in 2^N$, and $n \in \mathbb{Z}^d$, we now define $C^n_{\epsilon,N}$ to be the cuboid of side $N \epsilon$ and center $N \epsilon n$.

The characteristic function of this cube is denoted $1_{C^n_{\epsilon,N}}$, and enable us to define the projection
operators
\[ P_{\epsilon,N} = 1_{C^n_{\epsilon,N}}(D) - 1_{C^n_{\epsilon,N}}(D) \quad \text{and} \quad Q_{\epsilon,N} = 1_{C^n_{\epsilon,N}}(D). \]

These operators are bounded on $L^p$ spaces, $1 < p < \infty$ and provide decompositions of the identity:
\[ \sum_{N \in 2^N} P_{\epsilon,N} = \text{Id}, \quad \sum_{n \in \mathbb{Z}^d} Q^n_{\epsilon,N} = \text{Id}. \]

2.4. Lebesgue spaces. The scalar product on $L^2$ is defined by
\[ \langle f, g \rangle = \int \overline{f} g \, dx. \]

Space-time Lebesgue spaces on $[0, T] \times T^d$ are given by
\[ \|u\|_{L^p_T L^q} = \|u\|_{L^p([0,T],L^q(T^d))} = \left\| \|u(t,x)\|_{L^q(T^d)} \right\|_{L^p(0,T)}. \]
2.5. Bourgain spaces. We use scaled Sobolev spaces
\[ \|f\|_{H^s_x} = \|\langle \epsilon D\rangle^s f\|_{L^2} \]
and their associated scaled Bourgain spaces:
\[ \|u\|_{X^{s,b}} = \|e^{-it\Delta}u(t)\|_{L^2 H^s_x} = \|\langle \epsilon k\rangle^s (\tau + |k|^2)^b \hat{u}(\tau,k)\|_{L^2(\mathbb{R} \times \mathbb{Z}^d)}. \]  
Basic properties of these spaces are given in Appendix A.

2.6. Randomization. We denote \(P\) and \(E\) for the probability of an event and expectation of a random variable. The random variables \((G(k))_{k \in \mathbb{Z}^d}\) are independent centred standard complex Gaussians: they satisfy for all \(k, \ell \in \mathbb{Z}^d\),
\[ \mathbb{E}G(k) = 0, \quad \mathbb{E}[G(k)G(\ell)] = \delta(k - \ell) \]
and the Wick formula
\[ \mathbb{E}[G(k_1) \ldots G(k_r)\overline{G(k_{r+1})} \ldots G(k_{r+s})] = \# P, \]  
where \(P\) is the set of admissible pairings, that is to say partitions of \(\{1, \ldots, 2r\}\) into sets of the form \(\{i, j\}\) (unordered pairs). In particular, the above is zero if \(r \neq s\).

3. Proof of the main theorem

3.1. The frequency truncation for Wick ordering. Define the truncation operator
\[ \mathcal{F}[P](k) = \frac{1}{(2\pi)^d} \sum_{k_1+k_2+k_3=k} \hat{a}(k_1)\hat{b}(k_2)\hat{c}(k_3)(1 - \delta(k_1 + k_2) - \delta(k_2 + k_3)). \]
This gives the decomposition of the product
\[ abc = P(a, b, c) + \langle \pi, b \rangle c + a \langle \bar{b}, c \rangle. \]

3.2. Approximate solution and error. The approximate solution is defined through the following iterative resolution scheme
\[ u^0 = e^{it\Delta}u_0 \quad \text{and if } n \geq 1, \quad \begin{cases} \frac{d}{dt}u^n + \Delta u^n = \lambda^2 \sum_{i+j+k = n-1} P(u^i, \bar{u}^j, u^k), \\ u^n(0) = 0. \end{cases} \]  
Defining further
\[ V^{i,j} = \langle u^i, u^j \rangle, \quad V = \sum_{i,j \leq N} V^{i,j} = \left\| \sum_{n=0}^N u^n \right\|_{L^2}^2, \quad \omega(t) = -2t\|u_0\|_{L^2}^2, \]
our approximate solution will be
\[ u^{\text{app}} = \chi(t)e^{i\lambda^2 \omega}v^{\text{app}} \quad \text{with} \quad v^{\text{app}} = \sum_{n=0}^N u^n. \]
(notice how we add a smooth cutoff function in the definition of \(v^{\text{app}}\). Here, \(\chi\) is a function in \(C^\infty_0\) such that \(\chi = 1\) on \(B(0,1)\) and \(\chi = 0\) on \(B(0,2)^c\); as a result, the cutoff only affects \(t > 1\), and the equation is unchanged for \(t < 1\)). Extracting the factor \(e^{i\lambda^2 \omega}\) is sometimes called Wick ordering, and is classically used for random data problems, see for instance [5].

Since the flow of \([\text{NLS}]\) preserves the mass \(\|u(t)\|_{L^2} = \|u_0\|_{L^2}\), the approximate solution satisfies
\[ \frac{d}{dt}u^{\text{app}} + \Delta u^{\text{app}} - \lambda^2 |u^{\text{app}}|^2 u^{\text{app}} = E^N e^{i\lambda^2 \omega} + 2\lambda^2 \left(2 \mathcal{R}(u - u^{\text{app}}, u^{\text{app}}) + \|u - u^{\text{app}}\|_{L^2}^2\right) u^{\text{app}}, \]
where the error terms are given by
\[ E^N = \lambda^2 \left[ - \sum_{i,j,k \leq N} u^j \overline{u}^k + 2 \sum_{i,j,k \leq N} V^i j u^k \right]. \]

The solution \( u \) can now be decomposed into approximation plus error
\[ u = u^{app} + u^{err}, \quad \text{with} \quad u^{err} = \epsilon \lambda^2 \omega u^{err}, \]
where \( v^{err} \) solves
\[ i \partial_t v^{err} + \Delta v^{err} + 2\lambda^2 \| u \|_{L^2}^2 v^{err} = \lambda^2 \left( |v^{err} + u^{app}|^2 (v^{err} + v^{app}) - |v^{app}|^2 v^{app} \right) - E^N \]
\[ - 2\lambda^2 \left( 2Re(v^{err}, v^{app}) + \| v^{err} \|_{L^2}^2 \right) v^{app} \]
which can be written, for \( t < 1 \),
\[ i \partial_t v^{err} + \Delta v^{err} = \mathcal{L}(v^{err}) + \mathcal{B}(v^{err}) + \mathcal{T}(v^{err}) + \mathcal{E} \]
where the linear, bilinear, trilinear, and error terms are given by
\[ \mathcal{L}(u) = \chi(t) \lambda^2 \left[ 2|v^{app}|^2 u - 2V u - 2(v^{app}, u) v^{app} + (v^{app})^2 \pi - 2(u, v^{app}) v^{app} \right] \]
\[ \mathcal{B}(u) = \chi(t) \lambda^2 \left[ 2|u|^2 v^{app} - 2\| u \|_{L^2}^2 v^{app} - 2(u, v^{app}) u + u^2 v^{app} - 2(v^{app}, u) u \right] \]
\[ \mathcal{T}(u) = \chi(t) \lambda^2 \left( \| u \|^2 - 2 \| u \|_{L^2}^2 \right) \]
\[ \mathcal{E} = -\chi(t) E^N. \]

Notice how we once again added smooth cutoff functions in the definitions of the terms above; \( \chi \) is still a function in \( C^\infty_0 \) such that \( \chi = 1 \) on \( B(0,1) \) and \( \chi = 0 \) on \( B(0,2) \).

3.3. Bounds on the expansion. Assuming \( \epsilon \leq T \leq 1 \), Proposition 5.1 combined with the Bienaymé-Chebyshev inequality imply the following Corollary.

**Corollary 3.1.** Given \( 1 \leq t, T > \epsilon, N \in \mathbb{N}, \mu > 0, s > 0, p \geq 2 \), there exists \( b > \frac{1}{2} \) and a set \( E = E_{t,T,N,\mu,s,p} \) with probability \( \mathbb{P}(E) \geq 1 - C e^\mu \) such that on \( E \), and if \( n \leq N \),
\[ \| u^n(t) \|_{L^2} \lesssim_{n,\mu} e^{-\mu t^2} \left( \frac{1}{T_{kin}} \right)^n/2 \]
\[ \| u^n(t) \|_{L^p} \lesssim_{n,\mu,p} e^{-\mu (\epsilon + T^2)} \left( \frac{1}{T_{kin}} \right)^n/2 (T^{1/2} e^{-1})^{1/2} \]
\[ \left\| \chi(t) \int_0^t e^{i(t-s)\Delta} E^N ds \right\|_{X^s_b} \lesssim e^{-\mu} e^{-\frac{1}{2} - \frac{d}{2}} \left( \frac{1}{T_{kin}} \right)^{N/2} \]

**Remark 3.2.** Note a loss factor in the \( L^p \) estimates for \( p > 2 \) in the above. It could be removed by a further refined analysis, but there is no need for it in the present paper as the error shows an arbitrarily large polynomial gain.

3.4. Bounds on the linear, bilinear and trilinear terms. The following proposition gives a bound on \( \mathcal{L} \), if one excludes an exceptional set.

**Proposition 3.3.** If \( N \in \mathbb{N}, \mu > 0, s > 0 \), there exists \( b > \frac{1}{2} \) and a set \( E_{N,\mu,s} \) of probability \( \mathbb{P}(E_{N,\mu,s}) > 1 - C e^\mu \) on which the operator norm of \( \mathcal{L} \) can be bounded as follows:
\[ \left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \mathcal{L} ds \right\|_{X^{s,b} \to X^{s,b}} \lesssim e^{-\mu} \sqrt{\frac{1}{T_{kin}}} \]
Turning to the nonlinear terms, they will be controlled by the two following propositions.

**Proposition 3.4.** If \( N \in \mathbb{N}, \mu > 0, s > \frac{d}{2} - 1 \), there exists \( \epsilon > 1 \) and a set \( E_{N,\mu,s,b} \) with probability \( \mathbb{P}(E_{N,\mu,s,b}) \geq 1 - Ce^\mu \) such that on \( E_{N,\mu,s,b} \),
\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \mathcal{B}(u) \, ds \right\|_{X_s^b} \lesssim \lambda^2 \epsilon^2 \left( \frac{1}{2} - \frac{d}{2} - \eta \right) \|u\|^2_{X_s^b}.
\]

**Proposition 3.5.** Given \( s > \frac{d}{2} - 1 \) and \( \kappa > 0 \), there exists \( \beta > 0 \) such that
\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \mathcal{T}(u) \, ds \right\|_{X_s^b} \lesssim \lambda^2 \epsilon^{2-d-\eta} \|u\|^3_{X_s^b}.
\]

### 3.5. Control of the error, proof of the first part of Theorem 1.1.

Our aim is to apply the Banach fixed point theorem in \( B_{X_s^b}(0,\rho) \), where \( s > \frac{d}{2} - 1 \), and \( \rho > 0 \) will be fixed shortly, to the mapping
\[
\Phi : u \mapsto \int_0^t e^{i(t-s)\Delta} \left[ \mathcal{L}(u) + \mathcal{B}(u) + \mathcal{T}(u) + \mathcal{E} \right] \, ds.
\]
Note that \( \frac{1}{\epsilon} \Delta u \equiv \epsilon^4 \gamma^{-2} \) with \( 4\gamma - 2 > 0 \). Applying Corollary 3.1 the error term can be made smaller than \( N \) sufficiently big. This leads to the choice \( \rho = 2e^N(\frac{1}{4} - \gamma) \). Applying Proposition 3.3 with \( \mu < 2\gamma - 1 \), it appears that the linear term has an operator norm \( \ll 1 \). Similarly, applying propositions 3.4 and 3.5 one checks easily that the bilinear and cubic term act as contractions on \( B(0,\rho) \) thanks to the \( e^{N(\frac{1}{4} - \gamma)} \) size of \( v_{err} \). Therefore, the Banach fixed point theorem gives a solution \( v_{err} \), with norm \( \|v_{err}\|_{X_s^b} \lesssim e^{N(\frac{1}{4} - \gamma)} \).

In order to apply Corollary 3.1, Proposition 3.3 and Proposition 3.4, we had to exclude a set of size \( 2^N/4 \).

### 3.6. Comparison to the kinetic wave equation, proof of the second part of Theorem 1.1.

Fix \( \eta > 0 \), and let \( E \) be the exceptional set obtained in the previous subsection. Expanding \( u \) gives
\[
u = e^{i\lambda^2 \omega} \left[ u^0 + u^1 + u^2 + \sum_{n=3}^N w_n + v_{err} \right],
\]
so that, following the computation in \( \text{[8]} \),
\[
E \left( 1_E \left| \tilde{u}_k^2 - \epsilon^d |A(\epsilon k)|^2 \right| \right) = E \left( 1_E \left| \tilde{u}_k^1 \right|^2 \right) + 2 \Re E \left( 1_E \left| \tilde{u}_k^1 \tilde{u}_k^2 \right|^2 \right)
\]
main term
\[
+ \sum_{i,j \leq N} \sum_{i+j \geq 4} E \left( 1_E \left| \tilde{u}_k^i \tilde{u}_k^j \right|^2 \right) + 2 \sum_{i \leq N} \Re E \left( 1_E \left| \tilde{u}_k^i \tilde{u}_k^{v_{err}} \right|^2 \right) + |v_{err}|^2
\]
higher order
\[
(\text{notice that, due to a cancellation for } i+j = 3, \text{ the first sum of the higher order term only involves } i+j \geq 4). \text{ Up to excluding an exceptional set of size } \lesssim \epsilon^{\eta/4} \text{, using the fact that } X_s^b \text{ is continuously embedded in } C([0,T],L^2) \text{, and that } \lambda^2 \epsilon^2 \lesssim \epsilon \text{ the higher order terms can be bounded using the results of the previous subsections and } \text{(5.1)} \text{ by}
\]
\[
\|\text{higher order}\|_{L_k^2} \lesssim \sum_{i+j \geq 4} \|u^i(t)\|_{L^2} \|u^j(t)\|_{L^2} + \|v_{err}(t)\|_{L^2} \sum_{0 \leq n \leq N} \|u^n(t)\|_{L^2} + \|v_{err}(t)\|_{L^2}^2 \lesssim \epsilon^{-\eta} \frac{t 

Forgetting for a moment about \( \mathbb{1}_E \), the main term, following \cite{8}, can be written
\[
\text{main term} = \varepsilon^d \lambda^4 \sum_{k+\ell=m+n} \left| \frac{\sin(t\Omega(k, \ell, m, n))}{\Omega(k, \ell, m, n)} \right|^2 |A(ek)|^2 |A(\ell)|^2 |A(em)|^2 |A(en)|^2 \\
\left[ \frac{1}{|A(k)|^2} + \frac{1}{|A(\ell)|^2} - \frac{1}{|A(m)|^2} - \frac{1}{|A(n)|^2} \right].
\]
where
\[
\Omega(k, \ell, m, n) = |k|^2 + |\ell|^2 - |m|^2 - |n|^2.
\]
Viewing the sum above as a Riemann sum,
\[
\text{main term} = \lambda^4 \varepsilon^d \left( \int_{(\mathbb{R}^d)^3} \delta(k+\ell-m-n) \left| \frac{\sin(te^{-2\Omega(k, \ell, m, n)})}{e^{-2\Omega(k, \ell, m, n)}} \right|^2 |A(k)|^2 |A(\ell)|^2 |A(m)|^2 |A(n)|^2 \\
\left[ \frac{1}{|A(k)|^2} + \frac{1}{|A(\ell)|^2} - \frac{1}{|A(m)|^2} - \frac{1}{|A(n)|^2} \right] d\ell dm dn \right) (1 + O(t)).
\]
Since \( \int \frac{\sin(x)}{x^2} \text{d}x = \pi^2 \), there holds, for \( f \in C_0^\infty \) as \( \tau \to \infty \),
\[
\int \left| \frac{\sin(\tau\Omega)}{\Omega} \right|^2 f(\Omega) \text{d}\Omega = \pi^2 \tau^{-1} f(0) + O(\tau^{-2}),
\]
and therefore (recalling \( T_{kin} = \frac{1}{\tau \varepsilon} \))
\[
\text{main term} = \varepsilon^d \frac{t}{T_{kin}} \int_{(\mathbb{R}^d)^3} \delta(k+\ell-m-n) \delta(\Omega(k, \ell, m, n))|A(k)|^2 |A(\ell)|^2 |A(m)|^2 |A(n)|^2 \\
\left[ \frac{1}{|A(k)|^2} + \frac{1}{|A(\ell)|^2} - \frac{1}{|A(m)|^2} - \frac{1}{|A(n)|^2} \right] d\ell dm dn + O_{t_k}^1 \left( \frac{t}{T_{kin}} \left( \frac{t}{\varepsilon^2} + t \right) \right).
\]
Notice how the condition \( \varepsilon^2 \ll T \ll 1 \), already mentioned in \cite{14}, appears naturally here. Moreover, for \( t \geq \lambda^{-2} \), one has \( \frac{t}{\tau} \leq \lambda^2 \varepsilon^2 \).

This concludes the proof of the main theorem, except that we need to put back the characteristic function \( \mathbb{1}_E \). But one can check that the random variable \( |u^1_k|^2 \) enjoys better integrability properties: this is raised by using it to a high power, and taking the expectation. Therefore, the error resulting from \( \mathbb{1}_E \) is at most \( O(\varepsilon^m) \).

4. Encoding correlations by Feynman diagrams

Our strategy is to relate the computation of the quantities involved in Proposition \cite{5,1} to the computation of integrals with oscillatory phases in high dimension, whose structure can be encoded by Feynman diagram. We will give all details for the computation of \( \| u^m \|_{L^2} \). Other quantities will be also estimated using similar diagrams, and their construction and associated notations will naturally adapt. The notation and graph analysis follows essentially that of \cite{19}.

4.1. Diagrammatic representation. Recall that \( u^n \) is defined recursively by \cite{3,1}. To obtain a formula for \( u^n \), we use diagrams. We first define so-called interaction diagrams which encode three properties:

- As \( u^n \) is the solution of a forced Schrödinger equation, solved via Duhamel formula, the diagram possesses time slices corresponding to a specific choice of ordering of the time variables.
- As \( u^n \) involves a sum over triplets \( (i, j, k) \) with \( i + j + k = n - 1 \): the edges and vertices of the diagram corresponds to a particular choice of triplet at each recursive iteration.
The sum defining \( u^n \) involves complex conjugation: to each edge is associated a sign \( \sigma \) which records conjugation.

More precisely, following [19], given an integer \( n \), we first define the index set \( I_n = \{1, 2, ..., n\} \). A graph with \( n \) interaction has the total time \( t \) divided into \( n + 1 \) time slices of length \( s_i, i = 0, 1, ..., n \) whose index label the time ordering: from bottom to top in the graph. Associated with a time slice \( i \) there are \( 1 + 2(n - i) \) "waves", with three of them merging into a single one for the next time slice. Each wave in each time slice is represented by an edge \( e_{i,j} \) in the graph, with index set \( \mathcal{I}_n = \{(i, j), \ 0 \leq i \leq n, 1 \leq j \leq 1 + 2(n - i)\} \). The interaction history is encoded by a vector \( \ell = (\ell_1, \ldots, \ell_n) \in \mathcal{G}_n := I_{2n-1} \times I_{2n-3} \times \ldots \times I_1 \). Edges of the \( i \)-th time slice are related to edges of the \( i+1 \)-th one via vertices. An edge with index \((i, k)\) for \( k < \ell_i \) is matched with the edge with index \((i + 1, k)\) above it, the three edges with indexes \((i, \ell_{i+1}), (i, \ell_{i+1} + 1)\) and \((i, \ell_{i+1} + 2)\) merge to form the edge with index \((i + 1, \ell_i)\), and an edge with index \((i, k)\) for \( k > \ell_i + 2 \) is matched with the edge with index \((i + 1, k - 2)\). The corresponding interaction vertices are labelled according to the time ordering: \((v_i)_{1 \leq i \leq n}\). Complex conjugation is encoded by the "parity" \( \sigma_{i,j} \in \{\pm 1\} \) associated to each edge. Parity is defined recursively from top to bottom: \( \sigma_{n,1} = +1 \) or \( \sigma_{n,1} = -1 \) if the graph corresponds to \( u^n \) or to \( \overline{u}^n \). Then, parity is kept unchanged from an edge to the one below in absence of merging, and in case of a merging we require that \( \sigma_{i,\ell_i} = -1, \sigma_{i,\ell_{i+1}} = \sigma_{i+1,\ell_i}, \) and \( \sigma_{i,\ell_{i+2}} = +1 \).

At the initial time slice \( s_0 \), below each of the edges of index \((0, j)\) for \( 1 \leq j \leq 2n + 1 \) an initial vertex \( v_{0,j} \) is placed. At the final time slice slice \( s_n \), a final vertex \( v_R \) is placed. The graph obtained this way is a tree, and edges are oriented from bottom to top. The natural ordering between vertices is: \( v \leq v' \) if there is an oriented path from \( v \) to \( v' \). We define for each interaction vertex \( v_i \) the set of its initial vertices below as \( I_n(v_i) = \{j \in (0, 2n + 1), \ v_{0,j} \leq v_i\} \). Given a vertex \( v \) and an edge \( e \), we also write \( v \leq e \) if \( v \leq v' \) where \( v' \) is the top vertex of the edge \( e \), and \( I_n(e) = I_n(v') \). An example is given below:

We now associate to each edge \( e_{i,j} \) in the extended graph a frequency \( k_{i,j} \). At each vertex corresponds a \( \delta \) function ensuring that the sum of the momenta associated to the edges below is equal to that of the frequencies for edges above, and that frequencies associated with Wick ordering are removed. These are the Kirchhoff rules for the graph. At the final vertex \( v_R \), we impose the Dirac \( \delta(k_{n,1} - k_R) \) where \( k_R \) denotes the total output frequency. This gives the following formula for \( \hat{u}^n(k) \), where \( p \) is the number of vertices \( v_i \) whose edge above them carries a \(-1\) parity sign:

\[
\hat{u}^n(t, k_R) = e^{-it|k_R|^2} I^n(\lambda^2) \sum_{\ell \in \mathcal{G}_n} \sum_{k \in \mathbb{Z}^d} \mathcal{G}_n \cdot e^{-i\Omega k \sum_{j=0}^{n-1} s_j} \Delta_{\ell} |k_R| \delta(t - \sum_{i=0}^{n-1} s_i) ds
\]
where we used the shorthand notations

- \( k = (k_{i,j})_{(i,j) \in \mathbb{Z}_n} \in \mathbb{R}^{\# \mathbb{Z}_n} \)
- \( u_0 = (u_0, \ldots, u_n) \in \mathbb{R}^{n+1} \)
- \( u_0(k, +1) = \bar{u_0}(k) \) and \( u_0(k, -1) = \bar{u_0}(k) = \bar{u_0}(-k) \)
- \( \Omega_k = |k_{k-1,\ell_k+2}|^2 - |k_{k-1,\ell_k}|^2 + \sigma_k,\ell_k \) \(|k_{k-1,\ell_k+1}^2 - |k_{k,\ell_k}|^2\)

and, finally, \( \Delta_{\ell,P} \) encapsulates the Kirchhoff law and frequency truncation at each vertex, as well as the pairing of initial frequencies:

\[
\Delta_{\ell}(k, k_R) = \Delta_{\ell}(k) \delta(k_{0,1} - k_R)
\]

with

\[
\Delta_{\ell}(k) = \prod_{i=1}^{n} \prod_{j=1}^{\ell_i-1} \delta(k_{i,j} - k_{i-1,j}) \left( 1 - \delta(k_{i-1,\ell_i} + k_{i-1,\ell_i+2}) - \delta(k_{i-1,\ell_i+1} + k_{i-1,\ell_i+1+\sigma_i,\ell_i}) \right) \delta(k_{i,\ell_i}) - \sum_{j=0}^{2(n-i) - 2} \delta(k_{i,\ell_i+j}) \prod_{j=\ell_i+1}^{1+2(n-i)} \delta(k_{i,j} - k_{i,j+2}) \right). \quad (4.2)
\]

### 4.2. Expectation, cancellation of degenerate pairings

Given an integer \( n' \), and for \( 1 \leq i \leq 1 + 2n' \) frequencies \( k_{0,1+2n+i} \) and parities \( \sigma_{0,1+2n+i} \in \{\pm 1\} \) with \( \sum_{i=1}^{2n'+1} \sigma_{0,1+2n+i} = -\sigma_{0,1} \), we now want to evaluate expressions of the form

\[
\mathbb{E}(u_{n'} \prod_{i=1}^{1+2n'} \hat{u}_0(k_{0,1+2n+i}, \sigma_{0,1+2n+i})).
\]

This, via Wick’s formula \( (2.3) \), will induce pairings among the initial vertices. These pairings are encoded by a pairing \( P \) that is a partition of \( I_{2+2(n+n')} \) into pairs satisfying \( \sigma_{0,i}\sigma_{0,j} = -1 \) if \( \{i, j\} \in P \), and we denote by \( \mathcal{P} \alpha(n, n') \) the set of such pairings. We get via Wick’s formula;

\[
\mathbb{E} \left( \prod_{i \in I_{2n+1}} \hat{u}_0(k_{0,i}, \sigma_{0,i}) \prod_{i=1}^{1+2n'} \hat{u}_0(k_{0,1+2n+i}, \sigma_{0,1+2n+i}) \right) = \epsilon^{d(1+n+n')} \sum_{P \in \mathcal{P} \alpha(n, n')} \prod_{\{i, j\} \in P} |A(\epsilon k_{0,i})|^2.
\]

We say that a pairing \( P \) pairs an initial vertex \( v_{0,i} \) with another \( v_{0,j} \) if \( \{i, j\} \in P \). We say a pairing \( P \) has a degeneracy of index \( \{i, \{j, k\}\} \) if \( \sigma_{i-1,\ell_i+j}\sigma_{i-1,\ell_i+k} = -1 \) and if the vertices of \( \mathcal{I}_{n}(e_{i-1,\ell_i+j}) \cup \mathcal{I}_{n}(e_{i-1,\ell_i+k}) \) are all paired together by \( P \). We say a pairing \( P \) is degenerate if it has a degeneracy. We denote by \( \mathcal{P}(\ell, \ell', n, n') \) the set of all non-degenerate pairings of \( \{1, \ldots, 2(n + n' + 1)\} \). We aim here at explaining the following fact: degenerate pairings are responsible for the phase modulation \( e^{i\lambda \omega} \) of our approximate solution. Wick renormalisation, that cancels out this phase, is responsible for the presence of the frequency truncation \( 1 - \delta(k_{i-1,\ell_i} + k_{i-1,\ell_i+2}) - \delta(k_{i-1,\ell_i+1} + k_{i-1,\ell_i+1+\sigma_i,\ell_i}) \) in the Kirchhoff law \( (4.2) \) above. We claim that degenerate pairings are precisely cancelled by this truncation in a sense made precise in the next Lemma. While the Lemma itself is not used in the present analysis, it is extremely relevant to understand the problem at hand and how graph combinatorics are related to degeneracies and nondegeneracies in oscillatory phases of Feynman diagrams. The following Lemma states that if a pairing is degenerate, the Kirchhoff law \( \Delta_{\ell}(k, k_R) \) reduces further the dimension of the sum to be performed. This fact appears in Lemma \( \ref{B.2} \).

**Lemma 4.1.** Given integers \( n \) and \( n' \), \( \ell \in \mathcal{G}_n \), and for \( 1 \leq i \leq 2(1 + n + n') \) frequencies \( k_{0,i} \) and parities \( \sigma_{0,i} \in \{\pm 1\} \) with \( \sum_{i=1}^{2(n+n')} \sigma_{0,i} = 0 \), if the pairing \( P \) has a degeneracy of index \( \{i, \{j, k\}\} \) then:

\[
|\Delta_{\ell}(k, k_R)| \lesssim \delta(k_{i-1,\ell_i} + \sigma_{i-1,\ell_i+1}k_{i-1,\ell_i+1}) \delta(k_{i-1,\ell_i} + k_{i-1,\ell_i+2}).
\]
Proof. Assume that the pairing $P$ has a degeneracy of index $(i, \{j, k\})$. Assume without loss of generality that $j = 0$ and $k = 1$ (hence $\sigma_{i-1,\ell_i+2} = +1$). Then consider the edges and vertices below the two edges $e_{i-1,\ell_i+j}$ and $e_{i-1,\ell_i+k}$. The sum of all frequencies of the corresponding initial vertices is 0:
\[
\sum_{i \in I_{2n+1}} k_{0,i} = 0.
\]
This is because the pairing of $i$ with $j$ forces $k_{0,i} = -k_{0,j}$, and because all initial vertices in the sum above are paired together from the degeneracy assumption. Next, this sum is preserved among all time slices because the Kirchhoff laws enforce for each $0 \leq m \leq i - 1$:
\[
\sum_{m' \in I_{2(n-m)+1}} k_{m+1,m'} = \sum_{m' \in I_{2(n-m)+1}} k_{m,m'}.
\]
From the initial value, when reaching $m = i - 1$ we obtain:
\[
k_{i-1,\ell_i+j} + k_{i-1,\ell_i+k} = k_{i-1,\ell_i} + k_{i-1,\ell_i+1} = 0.
\]
As a consequence, one has the following identity for the truncation appearing in the definition (4.2) of $\Delta \ell(k)$:
\[
1 - \delta(k_{i-1,\ell_i} + k_{i-1,\ell_i+2}) - \delta(k_{i-1,\ell_i+1} + k_{i-1,\ell_i+1+\sigma_{i,\ell_i}}) = -\delta(k_{i-1,\ell_i} + k_{i-1,\ell_i+2})
\]
Combining the two identities above, we get the desired result.

\[\blacksquare\]

Remark 4.2. Below is an example of a degenerate pairing, which is ruled out by the above Lemma:

As a direct consequence of the above, one obtains the following corollary.

Corollary 4.3. Let $n \in \mathbb{N}$, $t > 0$. Then one has the following formula:
\[
\mathbb{E}\|u^n\|_{L^2}^2 = \sum_{G,P} F(G, P)
\]
where the sum is performed over all possible combinations of:

\[
\text{This is a degenerate pairing:}
\]

\[
\text{the initial vertices below } e_1, e_2
\]

\[
\text{form an isolated subpartition}
\]
- \( G = G_{\ell,e} \) is the tree \( G \) which we now describe. It is composed of one left and one right sub-tree, which have depth \( n \) and interaction histories \( \ell \) and \( e \) respectively. The root vertices of the left and right sub-trees, \( v_R \) and \( v'_R \), are merged into a single root vertex \( v_R \), and to them is attached an edge with frequency \( k_R \).

- \( P \) a pairing of \( \{1, \ldots, 2(2n+1)\} \) that is non-degenerate for \( \ell \) and \( e \), that is, given any vertex \( v \) and two edges \( e \) and \( e' \) below \( v \), the initial vertices associated with \( e \) and \( e' \) are not all paired together by \( P \).

Frequencies are denoted \( k_{i,j} \) for the left sub-tree, and \( k'_{i,j} \) for the right sub-tree, except when considering the pairing, in which case it is convenient to concatenate \( (k_{0,i}) \) and \( (k'_{0,i}) \) into a vector, which is still denoted \( (k_{0,i}) \), but has length \( 2(2n+1) \):

\[
(k_{0,1}, \ldots, k_{0,4n+2}) = (k_{0,1}, \ldots, k_{0,2n+1}, k'_{0,1}, \ldots, k'_{0,2n+1}).
\]

The formula is:

\[
\mathcal{F}(G, P) = \lambda^{4n} \xi^{d(2n+1)} \sum_{k_R \in \mathbb{Z}_+^{d}} \int_{\mathbb{R}_+^{n+1} \times \mathbb{R}_+^{n+1}} \prod_{(i,j) \in P} |A(\ell k_{0,i})|^2 \prod_{k=1}^n e^{-i\Omega_k \sum_{j=0}^{k-1} s_j} \prod_{k=1}^n e^{-i\Omega'_k \sum_{j=0}^{k-1} s'_j} \delta(t - \sum_{i=0}^n s_i) \delta(t - \sum_{i=0}^n s'_i) ds ds' \tag{4.4}
\]

where

- \( k = (k_{i,j})_{(i,j) \in \mathbb{Z}_n^+} \) or \( \ell = -1, j \in \mathbb{Z}_n^+ \) \( \in \mathbb{Z}_+^{2n+2N+1} \) and similarly for \( k' \),
- \( s = (s_0, \ldots, s_n) \in \mathbb{R}_+^n \) and similarly for \( s' \),
- \( \Omega_k, \Omega'_k \) were defined previously below (4.1),

and finally

\[
\Delta_{\ell,e', P}(k, k', k_R) = \Delta_{\ell}(k) \Delta_{e'}(k) \delta(k_R - k_{n,1} + k'_{n,1}) \Delta_P(k)
\]

with

\[
\Delta_P(k) = \prod_{(i,j) \in P} \delta(k_{0,i} - k_{-1,i}) \delta(k_{0,j} - k_{-1,j}) \delta(k_{-1,i} + k_{-1,j}).
\]

The above formula is encoded by the Diagram below:
Lemma 4.3 has the following natural total ordering. We write 

\[ 2(2 \text{ being induced by the pairing}) \]

Therefore, we will choose vector space given by the Kirchhoff rules for the 

\[ \text{they are continued unchanged (no merging) from one slice to another:} \]

the values of all other wave numbers. The first trivial simplification is to identify two edges when

\[ (\text{minimal collection of edges from which, given their associated wave number}) \]

Kirchhoff's law at each vertex, as appearing in the diagram of Lemma 4.3. We aim at finding a

\[ \text{Construction of the spanning tree.} \]

Lemma 4.4. Let \( m \in \mathbb{N}, e_1, \ldots, e_m \in \mathbb{R}, \) and \( \eta > 0. \) Then

\[
\int_{\mathbb{R}^+} \prod_{k=1}^{m} e^{-is_k e_k R} \left( \sum_{k=1}^{m} s_k - t \right) ds_1 \ldots ds_m = \frac{\eta}{2\pi} \int_{\mathbb{R}} \prod_{k=1}^{m} \frac{i}{\alpha - e_k + i\eta} d\alpha.
\]

Proof. We use the identity \( \delta(x) = \frac{1}{2\pi} \int e^{i\alpha x} d\alpha, \) together with the fact that \( e^{\eta(t-\sum s_k)} = 1 \) on the support of the integral, to write

\[
\int_{\mathbb{R}^+} \prod_{k=1}^{m} e^{-is_k e_k R} \left( \sum_{k=1}^{m} s_k - t \right) ds_1 \ldots ds_m = \int_{\mathbb{R}} \prod_{k=1}^{m} e^{-is_k e_k R} \frac{1}{2\pi} \int e^{i\alpha(\sum s_k - t)} d\alpha ds_1 \ldots ds_m
\]

By Fubini’s theorem, this is

\[
\ldots = \frac{\eta}{2\pi} \int e^{-i\alpha} \prod_{k=1}^{m} \int_{0}^{\infty} e^{is(\alpha - e_k + i\eta)} ds_k d\alpha = \frac{\eta}{2\pi} \int e^{-i\alpha} \prod_{k=1}^{m} \frac{i}{\alpha - e_k + i\eta} d\alpha.
\]

Choosing \( \eta = 1/t \) in the previous lemma, the identity (4.4) is transformed into:

\[
\mathcal{F}(G, P) = (-1)^{p'+p} e^{2\lambda n} \lambda^{(2n+1)} \sum_{k, k' \in \mathbb{Z}^{2(2n+1)}} \int_{\mathbb{R}^+} \prod_{\{i, j\} \in P} |A(\epsilon k)|^2 \ e^{-i(\alpha + \alpha')t} \prod_{k=1}^{2n} \alpha - \sum_{j=k}^{n} \Omega_j + \frac{1}{t} \prod_{k=1}^{2n} \alpha' - \sum_{j=k}^{n} \Omega_j' + \frac{1}{t} \Delta_{\ell, \ell', P}(k, k', 0) \ d\alpha d\alpha'.
\]

4.3. The resolvent identity.

4.4. Construction of the spanning tree. In the formula (4.5), the variables \( k_{i,j} \) are related by Kirchhoff’s law at each vertex, as appearing in the diagram of Lemma 4.3. We aim at finding a minimal collection of edges from which, given their associated wave number \( k_{i,j}, \) one can retrieve the values of all other wave numbers. The first trivial simplification is to identify two edges when they are continued unchanged (no merging) from one slice to another: \( (e_{i,k}, k_{i,k}) \) is identified with \( (e_{i+1,k}, k_{i+1,k}) \) if \( k < \ell_{i+1}, \) and \( (e_{i,k}, k_{i,k}) \) is identified with \( (e_{i+1,k}, k_{i+1,k-2}) \) if \( k > \ell_{i+1} + 2. \) The vector space given by the Kirchhoff rules for the \( k_{i,j} \) has dimension \( 2n + 1: \) indeed, there are \( 2(2n + 1) \) initial frequencies, and \( 2n + 1 \) pairings (the condition that the frequencies add up to zero being induced by the pairing). Therefore, we will choose \( 2n + 1 \) free edges, from whose frequencies all other frequencies can be reconstructed; these free edges will be determined by a spanning tree which will be constructed shortly. Edges that are not free are called integrated. The graph of Lemma 4.3 has the following natural total ordering. We write \( e \leq e' \) if:

\[ \bullet \ e = e_{i,j} \text{ is a root pairing edge and } e' \text{ is any other edge.} \]

\[ \bullet \ e = e_{-1,i} \text{ is an upper pairing edge, and } e' \neq e_{i,j} \text{ is not a root pairing edge.} \]
• $e$ and $e'$ are not pairing edges, $e$ belongs to the interaction graph on the right and $e'$ to that on the left.

• $e$ and $e'$ are not pairing edges and belong to the same interaction graph (left or right), and end at two time slices ($s_i$ and $s_j$ or $s'_i$ and $s'_j$) with $i \leq j$.

Notice that, for this order, edges which lie below the same vertex are considered equal. This is the natural time ordering inside each graph (left or right), and corresponds to the fact that we will integrate over the graph according to the natural time ordering of the time slices, first over the graph on the right, then over the graph on the left. To integrate the frequency constraints, we first define a minimal spanning tree the following way, following Lukkarinen and Spohn [19].

**Theorem 4.5.** Consider a frequency graph $G$. There exists a complete integration of the frequency constraints (4.2), determined by a certain spanning tree of the graph, in the following sense. There exists a subset of free interaction edges:

$$E^f = \{e^f_1, ..., e^f_{2n+1}\},$$

with associated frequencies $(k^f_i)_{1 \leq i \leq 2n+1}$, satisfying the following properties.

- **Spanning family:** On the vectorial subspace of $\mathbb{R}^{2\#I_n+4n+2}_{\mathbb{I}_{n+1}}$ determined by the Kirchhoff rules encoded in $\Delta_{\ell,\ell'}$, the family $(k^f_i)_{1 \leq i \leq 2n+1}$ is a free family. Moreover, any other wave number $k_{i,j}$ associated to an edge $e$ can be written as a unique linear combination of elements the elements $(k^f_i)_{1 \leq i \leq 2n+1}$.

- **Time ordering for the spanning:** In the decomposition above, if $1 \leq i \leq j \leq 2n + 1$ then $e^f_i \leq e^f_j$ for the natural time ordering of the graph. If $e$ is not a free edge, then its associated wave number can be uniquely written as:

$$k_e = \sum_{1 \leq i \leq 2n+1} c_{i,e} k^f_i \quad \text{with} \quad c_{i,e} \in \{-1, 0, 1\},$$

and $c_{i,e} = 0$ whenever $e^f_i < e$ for the natural time ordering of the diagram.

**Proof.** The spanning tree is constructed iteratively via the following algorithm. This again follows [19]. First all upper pairing edges $(e_{-1,j})_{1 \leq i \leq 2n+1}$ are added to the spanning graph. Then the interaction vertices are considered one by one, according to the natural ordering of the original graph.
At step $3k + 1$, the edge on the right is added to the spanning tree.

At step $3k + 2$, we consider the middle edge: if adding it does not create a loop with the spanning tree obtained after step $3k + 1$, we add it to the spanning tree, and if it does we leave it free.

At step $3k + 3$, we consider the right edge: if adding it does not create a loop with the spanning tree obtained after step $3k + 2$, we add it, and if it does we leave it free.

We move next to the $k + 1$-th interaction vertex for Step $3(k + 1) + 1$ and repeat this. Once we have finished with the graph on the right we add the edge above and move to the graph on the left. Once done with it we add the root edge on top of the diagram.

After the procedure explained above is completed, we declare all the edges which have been added to the graph integrated, and the ones which have not free.

The graph obtained at the end of the algorithm is a tree: each vertex is connected to the root vertex by a unique path. We call this tree the spanning tree. It carries a natural orientation, defined as follows: an integrated edge $e = \{v', v\}$ goes from $v'$ to $v$ if $v$ belongs to the path from $v'$ to the root vertex. This also defines a partial order: we say that $v' \leq v$ if $v$ belongs to the path from $v'$ to the root vertex. We denote by $P(v) = \{v', v' \prec v\}$ the set of vertices $v'$ such that $v$ belongs to the path from $v'$ to the origin. Given an integrated edge $e$ going from $v'$ to $v$, the number of edges on the path from $v$ to the root (counting $v$ and the root) is the distance to the root. An integrated edge $e$ going from $v'$ to $v$ is a leaf of the spanning tree if $P(v) = \emptyset$. 
Momenta of the edges belonging to the spanning tree are expressed in function of free frequencies. Given an edge $e$, and $v$ one of its vertices we define the parity of the edge with respect to the vertex as:

$$\sigma_v(e) = \begin{cases} +1 & \text{if } e \text{ is one of the edges above } v \text{ for the natural ordering,} \\ -1 & \text{if } e \text{ is one of the edges below } v \text{ for the natural ordering.} \end{cases}$$

Given a vertex $v$, $\mathcal{F}(v)$ denotes the set of free edges $f$ that have one extremity at $v$. Given $e = \{v_1, v_2\}$ an integrated edge going from $v_2$ to $v_1$, the formula for its associated frequency is then

$$k_e = \sum_{v \in \mathcal{P}(v_1), f \in \mathcal{F}(v)} (-\sigma_{v_1}(e)\sigma_v(f))k_f.$$  \hfill (4.6)

This formula can be proven by induction starting from the leaves of the spanning tree, and then step by step advancing toward the root vertex. If $e$ is a leaf, let us write $e = \{v_1, v_2\}$ with $e$ going from $v_1$ to $v_2$. All other edges having $v$ as an extremity are free, and $\mathcal{P}(v) = \{v\}$. The Kirchhoff law at $v$ is then:

$$\sigma_v(e)k_e + \sum_{f \in \mathcal{F}(v)} \sigma_v(f)k_f = 0$$

which implies (4.6) for the leaf $e$. Next, let $D \in \mathbb{N}$ be the maximal distance in the graph between an integrated edge and the root vertex. We prove formula (4.6) by induction on $1 \leq D' \leq D$. It is true for $D' = D$ as all integrated edges are leaves. Now assume it is true for $2 \leq D' \leq D$ and consider an integrated edge $e = \{v, v'\}$ going from $v$ to $v'$ at distance $D' - 1$ from the root. Let us denote by $\mathcal{I}(v)$ the integrated edges ending at $v$ (for the orientation of the spanning tree). Then Kirchhoff law at $v$ gives:

$$\sigma_v(e)k_e + \sum_{f \in \mathcal{F}(v)} \sigma_v(f)k_f + \sum_{e' \in \mathcal{I}(v)} \sigma_v(e')k_{e'} = 0.$$
as \( e' \) is at distance \( D' \) from the root, then gives the identity \( [4.6] \) at \( v \). The result follows for any \( D' \) by induction.

To finish the proof of Theorem 4.5, we need to show that if \( e \) is an integrated edge, then it is only a linear combination of free frequencies appearing after \( e \) for the natural time ordering of the diagram. Assume \( f = \{ u', u \} \) is a free edge, with \( u' \) before \( u \) for the natural time ordering. This means that during the construction of the spanning tree, at the step where the vertex \( u \) is considered, \( f \) is not added as this would create a loop in the spanning tree in construction. At that step, all edges in the spanning tree are before \( f \) for the natural time ordering. Hence there exists a path \( \tilde{p} \) from \( u \) to \( u' \), and all its edges are before \( f \) for the natural time ordering. Also, there exist unique paths \( p \) and \( p' \) going from \( u \) to the root and from \( u' \) to the root respectively. These paths intersect at a vertex \( v \). By their uniqueness, \( v \) has to belong to \( \tilde{p} \). Consider now the formula above: \( k_f \) can only appear in the integrated frequencies on the paths from \( u \) and \( u' \) to the root. Moreover, after the vertex \( v \), the two contributions from \( u \) and \( u' \) in this formula cancel. Hence \( k_f \) can only appear in the integrated frequencies on the path from \( u \) to \( v \), and in the integrated frequencies on the path from \( u' \) to \( v \). These belong to \( \tilde{p} \) hence are indeed before \( k_f \) for the natural time ordering. 

4.5. Recovering all frequencies from the free frequencies. We attach the interaction edges of the graph to their upper vertex and classify interaction vertices according to the number of free edges that are attached to them:

- **Degree 0 vertex**: if no free edge is attached to this vertex.
- **Degree 1 vertex**: if one free edge is attached to this vertex.
- **Degree 2 vertex**: if two free edges are attached to this vertex.

Any vertex is necessarily of degree 0, 1 or 2. Indeed, in the determination of the free edges in Theorem 4.5 by construction, the edges on the right below each vertex always belong to the minimal spanning tree and hence are not free. The following Lemma describes how the frequencies associated to free edges below one interaction vertex appear in the decomposition of the frequencies associated to the other integrated edges below this vertex.

**Lemma 4.6.** The following holds true.

- **Degree 1 vertex**: Assume \( v \) is a degree one vertex. Then the edges below it are always of the form \( \{ f, e, e' \} \) (unordered list), where \( f \) is the free edge, and where the formulas giving \( k_e \) and \( k_{e'} \) in terms of the free edges are \( k_e = -k_f + G \) and \( k_{e'} = G' \), where \( G \) and \( G' \) are independent of \( k_f \).

- **Degree 2 vertex**: Assume \( v \) is a degree two vertex. Then the edges below it are always of the form \( \{ f, f', e \} \) (unordered list), where \( f \) and \( f' \) are the free edges, and where the formula giving \( k_e \) and \( k_{e'} \) in terms of the free edges is \( k_e = -k_f - k_{f'} + G \) where \( G \) is independent of \( k_f \) and \( k_{f'} \).

**Proof.** Assume \( v \) is of degree one, and denote by \( f = \{ v', v \} \) the free edge below. As explained in the proof of Theorem 4.5, the path \( \tilde{p} \) going from \( v' \) to \( v \) in the spanning tree is made of edges appearing before \( v \) for the natural time ordering, and also, the path \( p' \) going from \( v' \) to the root vertex and the path \( p \) going from \( v \) to the root vertex intersect at a vertex \( v_0 \) belonging to \( \tilde{p} \). Let us call \( e = \{ u, v \} \) the one of the integrated below \( v \) that belongs to \( \tilde{p} \) and \( e' = \{ u', v \} \) the other one. There are several cases to distinguish.

**In case 1.** \( e \) goes from \( v \) to \( u \) for the orientation of the tree, hence \( e' \) goes from \( u' \) to \( v \) and belongs to neither of \( p \) or \( p' \). We get \( v', v' \notin P(u') \) so from the formula \( [4.6] \) \( k_{e'} = G' \) is independent of \( k_f \).

**In case 2.** \( e \) goes from \( u \) to \( v \) for the orientation of the tree, with \( e' \) going from \( u' \) to \( v \). We get \( v' \notin P(v) \) so from \( [4.6] \) \( k_e = -k_f + G \) with \( G \) independent of \( k_f \).
In case 3, $e$ goes from $u$ to $v$ for the orientation of the tree, with $e'$ going from $v$ to $u'$. We get $v' \in \mathcal{P}(u)$ so from the formula (4.6) $k_e = -k_f + G$, $G$ independent of $k_f$. We get $v' \in \mathcal{P}(v)$ so (4.6) implies $k_e' = G'$ independent of $k_f$ (the two contributions from $v$ and $v'$ giving a $k_f$ term come with opposite signs and cancel).

For a degree two vertex, the very same reasoning applies: there are two free edges $f = \{v', v\}$ and $f = \{v'', v\}$ below $v$ on the left and middle, and one integrated edge $e = \{u, v\}$ below $v$ on the right. There are unique paths $p, p'$ and $p''$ going from $v, v'$ and $v''$ to the root, with $p'$ intersecting $p$ at a vertex $v_0'$ and $p''$ intersecting $p$ at a vertex $v_0''$, with both $v_0'$ and $v_0''$ before $v$ for the natural time ordering of the graph. Moreover, either $e$ belongs to $p$ and to neither of $p'$ and $p''$, or $e$ belongs to both $p'$ and $p''$ and not to $p$. In each case, applying the same reasoning as for a degree one vertex gives $k_e = -k_f - k_{f'} + G$ with $G$ independent of $k_f$ and $k_{f'}$. □

Lemma 4.7. Denoting by $n_i$ the number of interaction edges of degree $i$ for $i = 0, 1, 2, 3$, one has the following relations:

$$n_0 + n_1 + n_2 = 2n$$

(4.7)

$$n_1 + 2n_2 = 2n$$

(4.8)

Proof. The first relation [4.7] comes from the fact that there are $2n$ interaction vertices, $n$ in each subgraph below the root vertex. For the second, recall that there are $2n + 1$ free edges (excluding the pairing edges). Moreover, there is always one edge below the root vertex that is free, and one that is not. Hence, there are $2n + 1 - 1 = 2n$ free variables below the interaction edges, which on the other hand equals $n_1 + 2n_2$ by definition of the degree. This proves (4.8). □

Lemma 4.8. The last interaction vertex, that on top of the left graph, is always of degree 2.

Proof. In other words, we claim that the last three free edges obtained in Theorem 4.5 are $e_{2n-1}^l = e_{n-1,1}, e_{2n}^l = e_{n-1,2}$ and $e_{2n+1}^l = e_{n,1}$.

To show it, let us for simplicity call $v$ the last interaction vertex, $e_1 = e_{n-1,1}, e_2 = e_{n-1,2}$ and $e_3 = e_{n-1,3}$ the three edges below it (left, center, right), and $\tilde{v}$ the last interaction vertex of the right graph. Let us call $S_1, S_2, S_3$ the set of initial vertices below $e_1, e_2, e_3$, and $\tilde{S}$ the set of initial vertices of the right graph (below $\tilde{v}$). Let us also call $G^m$ the minimal spanning tree at the start of the penultimate step, when $v$ and its edges below are considered.

By the algorithm used to construct the spanning tree, there exists paths in $G^m$ between any two initial vertices in $S_1$, and also in $S_2$ and in $S_3$. As $\tilde{S}$ contains an odd number of elements, there exists $i \in \{1, 2, 3\}$ such that one vertex of $\tilde{S}$ is paired with another one of $S_i$. This implies that there exists a path in $G^m$ between any two vertices in $\tilde{S} \cup S_i$. Let us call $j, k \neq i$ the remaining vertices in $\{1, 2, 3\}$. By non-degeneracy of $P, S_j$ and $S_k$ cannot be fully paired one with another, so that $S_j$ (without loss of generality) has one vertex paired with either $S_i$ or $\tilde{S}$. Hence, there exists a path between any two vertices in $\tilde{S} \cup S_i \cup S_j$. As $S_k$ has to be paired at least with one element in $\tilde{S}, S_i$ or $S_j$, we finally find that at the before last step, there already exists a path between any two initial vertices. So after $e_3$ is added to the spanning tree, adding either $e_1$ or $e_2$ would create a loop: these vertices are then free. □
5. Bounds on the approximate solution

5.1. Main result.

Proposition 5.1. For any $p \in \mathbb{N}$, under the hypothesis \([2.1]\),

$$
E\|u^n(t)\|_{L^2}^{2p} \lesssim_{n,p,\epsilon} \epsilon^{-\kappa} \left\{ \begin{array}{ll}
(\lambda^2 t)^{2p} & \text{for } t \leq \epsilon,
(\lambda^2 t)^{2p} t^{1/2} (t^2 \epsilon^{-1})^{p-1} & \text{for } \epsilon \leq t \lesssim 1.
\end{array} \right.
$$

(5.1)

$$
E\|u^n\|_{L^2}^{2p} \lesssim_{n,p,\epsilon} \epsilon^{-\kappa} \left\{ \begin{array}{ll}
T(\lambda^2 T)^{2p} & \text{for } T \leq \epsilon,
(\epsilon + T^2)(\lambda^2 t)^{2p} (T^2 \epsilon^{-1})^{p-1} & \text{for } \epsilon \leq T \lesssim 1.
\end{array} \right.
$$

(5.2)

As a consequence,

$$
E \left\| \chi(t) \int_0^t e^{i(t-s)\Delta} E^N ds \right\|_{X^{s,b}} \lesssim \epsilon^{-\frac{1}{4}-\frac{d}{4}-\kappa} \left( \frac{T}{T_{kin}} \right)^{\frac{N}{2}}
$$

In addition, for $b > \frac{1}{2}$ and $s \in \mathbb{R}$,

$$
E \left\| \chi \left( \frac{t}{T} \right) u^n \right\|_{X^{s,b}}^2 \lesssim_{n,s,b,\epsilon} \epsilon^{-\kappa} (\lambda^2 \epsilon)^{2n} \text{ for } \epsilon \lesssim T \lesssim 1
$$

(5.3)

5.2. Proof of the $L^2$ bound. The trivial bound on the time integral. For $t \leq \epsilon$, we use the identity \([4.4]\) and use the rough estimates $|e^{-it\Omega_k \sum_{j=0}^{k-1} s_j}| \leq 1$:

$$
|\mathcal{F}(G, P)| \lesssim \lambda^{4n} \epsilon^{d(2n+1)} \sum_{k^l} \int_{\mathbb{R}_+^{n+1}} 1(|k|, |k'| \lesssim \epsilon^{-1}) \Delta \epsilon^{i\sigma P}(k,k',0) \delta(t - \sum_{i=0}^{n} s_i) \delta(t - \sum_{i=0}^{n} s_i') \, ds \, ds'
$$

After resolution of the momenta constraints, Theorem \([4.5]\) we compute the above integral by integrating over the free variables $(k^l)^i_{1 \leq i \leq 2n+1}$ given by this Theorem, giving a factor $\epsilon^{-d(2n+1)}$. Then we integrate over the temporal variables, giving a $t\epsilon$ factor, so that:

$$
|\mathcal{F}(G, S, P)| \lesssim \lambda^{4n} \epsilon^{d(2n+1)} \epsilon^{-d(2n+1)} t^{2n} = \lambda^{4n} t^{2n}.
$$

Splitting the resolvent integral. We use the identity \([4.5]\) for the oscillatory factors, instead of \([4.4]\) as in Step 0. We resolve the momenta constraints using Theorem \([4.5]\) so that the integral over all frequencies reduces to the integral over free frequencies $(k^l)^i_{1 \leq i \leq 2n+1}$. This gives

$$
|\mathcal{F}(G, P)| \lesssim \lambda^{4n} \epsilon^{d(2n+1)} \sum_{k^l} 1(|k|, |k'| \lesssim \epsilon^{-1}) \Delta \epsilon^{i\sigma P}(k,k',0)
$$

$$
= \lambda^{4n} \epsilon^{d(2n+1)} \int_{|\alpha,\alpha'| \leq \frac{K}{2}} \frac{1}{\alpha - \sum_{j=k}^{n} \Omega_j + \frac{i}{2}} \prod_{k=1}^{n} \frac{1}{\alpha' - \sum_{j=k}^{n} \Omega_j' + \frac{i}{2}} \, d\alpha \, d\alpha' + \lambda^{4n} \epsilon^{d(2n+1)} \int_{|\alpha,\alpha'| > \frac{K}{2}} \, d\alpha \, d\alpha',
$$

(5.4)

where above the variables $k_{i,j}$ and the resonance moduli $\Omega_j$ are obtained from the variables $\eta_{i,j}$ and $k^l_i$ via Theorem \([4.5]\) and where we split the integral into the two regions $|\alpha,\alpha'| \leq K \epsilon^{-2}$ and $|\alpha,\alpha'| > K \epsilon^{-2}$ for some large constant $K \gg 1$ in the last line.

The bound for $(\alpha, \alpha')$ small: $\mathcal{F}_1(G, P)$. In order to bound this term, we integrate over the free variables in the following order: we consider each interaction vertex iteratively according to the natural time ordering of the graph, and each time we integrate over the free variables below it (see below for
the details of this operation). This results in an integration over the variables \((k_i^j)_{1 \leq i \leq 2n+1}\). Then we integrate over the \(\alpha\) and \(\alpha'\) variables which contribute a subpolynomial factor.

When integrating at each edge we obtain the following bounds. Below we treat the case for which the edge belong to the left graph for simplicity.

- If \(v_k\) is of degree 0, then we use the rough bound

\[
\left| \frac{1}{\alpha - \sum_{j=k}^{n} \Omega_j + \frac{t}{t}} \right| \leq t.
\]

- If \(v_k\) is of degree 1, we write:

\[
\alpha - \sum_{j=k}^{n} \Omega_j = -\Omega_k + \alpha - \sum_{j=k+1}^{n} \Omega_j
\]

\[
= |k_{k-1,\ell_k}|^2 - |k_{k-1,\ell_k+1}|^2 - |k_{k,\ell_k}|^2 + \alpha - \sum_{j=k+1}^{n} \Omega_j.
\]

There exists one free variable among \(k_{k-1,\ell_k}\) and \(k_{k-1,\ell_k+1}\), that we denote by \(k_i^j\) (i.e. this is the \(i_k\)-th free variable given by Theorem 4.5). Above, from the time ordering property of Theorem 4.5 one notices that the quantities \(|k_{k,\ell_k}|^2\) and \(\alpha - \sum_{j=k+1}^{n} \Omega_j\) are independent of \(k_i^j\). By Lemma 4.6 and Lemma B.2 since \(\epsilon \lesssim t\),

\[
\left| \sum_{k_i^j} 1(\| \mathbf{k} \| \lesssim \epsilon^{-1}) \Delta_{\ell,\ell',\mathbf{P}}(\mathbf{k}, \mathbf{k}', 0) \frac{1}{\alpha - \sum_{j=k}^{2n} \Omega_j + \frac{t}{t}} \right| \lesssim \epsilon^{1-d-\kappa}.
\]

- If \(v_k\) is of degree 2 we perform a similar analysis. We integrate over its associated free variables (that we denote here by \(k_i^j\) and \(k_i^j\)) and estimate by Lemma B.1

\[
\left| \sum_{k_i^j, k_i^j} \Delta_{\ell,\ell',\mathbf{P}}(\mathbf{k}, \mathbf{k}', 0) \frac{1(\| \mathbf{k} \| \lesssim \epsilon^{-1})}{\alpha - \sum_{j=k}^{2n} \Omega_j + \frac{t}{t}} \right| \lesssim \epsilon^{2-2d-\kappa}.
\]

Recall the notation \(n_t\) from Lemma 4.7. Once the above integration procedure is completed, we integrate over the last free variable ending at the root vertex using solely that it is restricted to a ball of radius \(\epsilon^{-1}\), yielding a factor \(\epsilon^{-d}\). We thus obtain the following bound for the first part of the integral in (5.4):

\[
|\mathcal{F}_1(G, P)| \lesssim_{n, \kappa} \lambda^{4n} \epsilon^{-\kappa} \epsilon^{d(2n+1)} t^{n_0} (\epsilon^{-d})^{n_1} (\epsilon^{2-2d})^{n_2} \epsilon^{-d} \int_{|\alpha, \alpha'| \leq \frac{K}{\epsilon^2} (|\alpha| + \frac{1}{t})} \frac{d\alpha}{|\alpha + \frac{1}{t}|} d\alpha'
\]

\[
\lesssim \epsilon^{-\kappa} \lambda^{4n} \epsilon^{2n_0} t^{n_0} \epsilon^{(1-d)n_1} \epsilon^{(2-2d)n_2}.
\]

Using successively that \(n_1 + 2n_2 = 2n\) and \(n_0 + \frac{n_1}{2} = n\), as follows from (4.7) and (4.8), this is

\[
\ldots = \epsilon^{-\kappa} \lambda^{4n} \epsilon^{2n_0} t^{n_0} t^{-\frac{n_1}{2}} \lesssim \epsilon^{-\kappa} \lambda^{4n} \epsilon^{2n_1} t.
\]  (5.5)

In order to obtain the last inequality, we used that the worst case is \(n_1 = 2n - 2\) for \(t \leq 1\) (using Lemma 4.8 to rule out the case \(n_1 \geq 2n - 2\)), and \(n_1 = 0\) for \(t \geq 1\).

The bound for \((\alpha, \alpha')\) large: \(\mathcal{F}_2(G, P)\). Let us for simplicity only consider the case where \(|\alpha'| > K\epsilon^{-2}\), \(|\alpha| < K\epsilon^{-2}\) (by symmetry, the only other case to consider is \(|\alpha|, |\alpha'| > K\epsilon^{-2}\), which is
simpler). Noticing that $|Ω_k| \lesssim ε^{-2}$ for the interactions we consider, there holds for $|α'| > Kε^{-2}$ if $K$ has been taken large enough:

$$\frac{1}{|α' - \sum_{j=k}^{2k} Ω'_j + \frac{i}{t}|} \lesssim \frac{1}{|α'|}.$$ 

Let us denote by $n'_0$, $n'_1$ and $n'_2$ the numbers of degree 0, 1 and 2 vertices in the right graph. In particular, there are $n'_1 + 2n'_2$ free variables in the right graph, and $n'_0 + n'_1 + n'_2 = n$. Using the trivial support estimate $|k'_j| \lesssim ε^{-1}$, summing over the free variables of the right graph and then integrating with respect to $α'$ yields:

$$\sum_{k'j'} 1(|k'_j, k'_j'| \lesssim ε^{-1}) \Delta t, G, P(k, k'_j', 0) \int_{|α'| ≥ Kε^{-2}} \frac{\frac{1}{α' - \sum_{j=k}^{k} Ω'_j + \frac{i}{t}}}{|α'|^{n+1}} dα' \lesssim ε^{-2(d-1)n'_2(1-d)n'_1 t'n'_0} (ε^2 t^{-1})^{2n'_0} K\lesssim ε^{-2(d-1)n'_2(1-d)n'_1 t'n'_0},$$

where $T ≥ ε^2$ were used. Above, note that $ε^{-2(d-1)n'_2(1-d)n'_1 t'n'_0}$ is the contribution of the right graph in the case $|α'| \leq Kε^{-2}$ done previously. Hence we get a better estimate comparing with the previous case. We integrate next over the free variables of the left graph as in the $|α|, |α'| \leq Kε^{-2}$ case, and obtain in fine a smaller upper bound for this term than (5.5). This shows that $\mathcal{F}_2(G, P)$ admits (5.5) as an upper bound as well.

5.3. Proof of the $L^p$ bound. We use the relation:

$$\|w^n\|_{L^{2p}}^2 = \mathcal{F} ((w^n)^p (w^n)^p) (0)$$

to use the same framework as the proof for $p = 1$. What changes is that the identity corresponding to (4.3) is now:

$$\mathbb{E}\|w^n\|_{L^{2p}}^2 = \sum_{G, P} \mathcal{F}(G, P)$$

where one has the following analogue of (4.3) where the notation $'$ to distinguish between the first and second graph is now replaced by the superscript $m \in \{1, ..., 2p\}$ to distinguish between the $2p$ different graphs:

$$\mathcal{F}(G, P) = (-1)^{\sum_{m=1}^{2p} m} e^{2p} \lambda^{4p} \epsilon^{dp(2n+1)} \sum_k \int_{\mathbb{R}^{2p}} \prod_{j \in P} |A(εk_{0j})|^2 e^{-i\sum_{m=1}^{2p} 2p \prod_{m=1}^{n} \prod_{k=1}^{2p} \frac{1}{α^m - \sum_{j=k}^{2k} Ω'_j + \frac{i}{t}}}
\prod_{m=1}^{2p} e^{-i\sum_{m=1}^{2p} 2p \prod_{m=1}^{n} \Delta_{k, P}(k, k_R) dα},$$

with an an obvious definition for $\Delta_{k, P}(k, k_R)$. This can be represented by the following graph:
\( n_3 \) denotes the number of free edges towards the root.

\[ \xi_R = 0 \] corresponds to \( E\|u_n\|_{L^2}^{2p} \)

\( n_3 \) denotes the number of free edges towards the root. \( 2p \) interaction graphs indexed by \( m = 1, \ldots, 2p \)

\[ \widehat{u^i}, \widehat{\bar{u}^i} \]

Initial vertices paired with \( P \)

We introduce the following new notation: \( n_3 \in \{1, \ldots, 2p-1\} \) denotes the number of free edges joining the last interaction vertices to the root. Lemma 4.7 naturally adapts and we get the following relations:

\( n_0 + n_1 + n_2 = 2pn, \quad \text{and} \quad n_1 + 2n_2 + n_3 = p(2n + 1), \quad 1 \leq n_3 \leq 2p - 1. \) (5.6)

The first one is the decomposition of all \( 2pn \) vertices into those of degree 0, 1 and 2 as for \( p = 1 \), the second one is the total number of free variables, and the last one expresses the fact that at least one of the last edges reaching the root is free, and that at least one is not. We perform the very same strategy of the proof of the bound for \( \mathcal{F}(G, P) \) as for the case \( p = 1 \). There is one exception: after integrating all free variables below interaction edges, we integrate over the remaining last \( n_3 \) free variables and estimate using the trivial support estimate \( |k_{n,1}^{m}| \leq \varepsilon^{-1} \). This produces:

\( |\mathcal{F}(G, P)| \lesssim \varepsilon^{-\kappa} \lambda^{4\rho p} e^{4p(2n+1)\epsilon n_0} \epsilon^{1-\delta} n_1 \epsilon (2-2d)n_2 \epsilon^{-dn_3}. \)

Since \( 2n_2 = p(2n + 1) - n_1 - n_3 \) and \( n_0 + \frac{n_1}{2} = pn + \frac{n_3-p}{2} \), this is

\[ \ldots = \varepsilon^{-\kappa} \lambda^{4\rho p} e^{2pn} e^{-(n_3-p)} (t^4 \epsilon^{-1})^{n_3-p} \lesssim \varepsilon^{-\kappa} \lambda^{4\rho p} e^{2pn} e^{(t^4 \epsilon^{-1})^{p-1}} \]

where we used in that if \( t \leq 1 \) then \( t^4 \epsilon^{-1} \leq t^{1-p} \) as \( n_1 \leq 2pn - 2 \) (using Lemma 4.8 to rule out the case \( n_1 > 2pn - 2 \)), and for the last inequality that the worst possible contribution of \( (t^4 \epsilon^{-1})^{p-1} \) occurs for \( n_3 = 2p - 1 \) as \( \varepsilon \geq 1 \).

5.4. Proof of the \( X^{s,b} \) bound. The proof follows the same strategy as that of the \( L^2 \) norm. We will solely use the space-time Fourier transform of \( w^n \), which will only produce minor changes. From (4.1) and the resolvent identity Lemma 4.4 with \( \eta = 1/T \), one obtains the following expression for the spacetime Fourier transform of \( w^n \) (where the \( e^T \) factor has been absorbed in the cut-off \( \chi(t/T) \) in the right hand side to simplify notations):

\[ \mathcal{F} \left( \chi \left( \frac{t}{T} \right) u^n \right)(\tau, k_R) = i^n (-1)^p (-i \lambda^2)^n \sum_{\ell \in \mathcal{G}_n} \sum_i \hat{u}_0(k_{0,i}, \sigma_{0,i}) \]

\[ \int_{\mathbb{R}} T\tilde{\chi}(T(\tau - \tau_1)) \frac{1}{-\tau_1 - \frac{|k_n|^2}{2} + i \frac{1}{T} \prod_{k=1}^n (-\tau_1 - |k_R|^2 - \sum_{j=k}^n \Omega_j + i \frac{1}{T}) \Delta_{\ell,p}(\hat{k}, k_R) d\tau_1. \]
The identity corresponding to (4.3) is now:

\[ E \left\| \left( \frac{t}{T} \right) w^{n} \right\|_{X^{s,b}}^{2} = \sum_{G,P} \mathcal{F}(G, P) \]

where one has the following analogue of (4.5):

\[ \mathcal{F}(G, P) = T^{2} \lambda^{4n} e^{d(2n+1)} \sum_{k_{f},k'_{f},k_{R}} \prod_{(i,j) \in P} |A(\epsilon k_{0,i})|^{2} \Delta_{t',e',P}(k_{f},k'_{f},k_{R}) \]

\[ \int_{\mathbb{R}^{3} \times \mathbb{R}^{3}} \frac{1}{-\tau_{1} - |k_{R}|^{2} + \frac{i}{T}} \frac{1}{-\tau_{2} - |k_{R}|^{2} + \frac{i}{T}} \frac{1}{-\tau_{2} - |k_{R}|^{2} - \sum_{j=k}^{n} \Omega_{j}} \]

\[ \langle \tau + |k_{R}|^{2} \rangle^{2b} \check{\chi}(T(\tau - \tau_{1})) \check{\chi}(T(\tau - \tau_{2})) d\tau_{1} d\tau_{2} d\tau \]

\[ = T^{2} \lambda^{4n} e^{d(2n+1)} \int_{|\tau| \leq K^{\epsilon-2}} [...] + T^{2} \lambda^{4n} e^{d(2n+1)} \int_{|\tau| > K^{\epsilon-2}} [...]. \]

For the first part \( \mathcal{F}_{1} \), after summing over free variables (except the last one) like in the proof for \( p = 2 \) done in Subsection 5.2, we are left with the same estimate, times the following quantity:

\[ \sum_{|k_{R}| \leq 1} \int_{\mathbb{R}^{3}} \frac{1}{-\tau - |k_{R}|^{2} + \frac{i}{T}} \frac{1}{-\tau - |k_{R}|^{2} - \sum_{j=k}^{n} \Omega_{j}} \langle \tau + |k_{R}|^{2} \rangle^{2b} \check{T}^{2} \check{\chi}(T(\tau - \tau_{1})) \check{\chi}(T(\tau - \tau_{2})) 1(\tau < K^{\epsilon-2}) d\tau_{1} d\tau_{2} d\tau. \]

Since the convolution of \( \frac{1}{|\tau + \epsilon + \frac{i}{T}|} \) with \( \check{T} \) is bounded by a multiple of \( \frac{1}{|\tau + \epsilon + \frac{i}{T}|} \), the above can be bounded by

\[ \sum_{|k_{R}| \leq 1} \int_{\mathbb{R}^{3}} \frac{1}{-\tau - |k_{R}|^{2} + \frac{i}{T}} \frac{1}{-\tau - |k_{R}|^{2} - \sum_{j=k}^{n} \Omega_{j}} \langle \tau + |k_{R}|^{2} \rangle^{2b} 1(\tau < K^{\epsilon-2}) d\tau. \]

Summing in \( k_{R} \) before integrating in \( \tau \), this is

\[ \cdots \lesssim \epsilon^{-d} \int_{\mathbb{R}} \frac{1}{|\tau + \epsilon + \frac{i}{T}|^{2}} \langle \tau \rangle^{2b} 1(\tau < K^{\epsilon-2}) d\tau \lesssim \epsilon^{-d+2(2b-1)}. \]

Hence this part yields same estimate as \( \mathcal{F}_{1} \) in the proof for \( p = 2 \), up to a loss of \( \epsilon^{2(2b-1)} \):

\[ |\mathcal{F}_{1}(G, P)| \lesssim \epsilon^{-K^{\epsilon-2}} \lambda^{4n} e^{2n+1} t \quad (5.7) \]

For the second part \( \mathcal{F}_{2} \), we similarly get rid of the convolution (integration over \( \tau_{1} \) and \( \tau_{2} \) at a total cost of \( T^{-2} \)), and use the fact that since all frequencies \( k_{f}^{i} \) are constrained \( |k_{f}^{i}| \lesssim \epsilon^{-1} \) then for \( \tau \geq K^{\epsilon-2} \) for \( K \) large enough,

\[ \left| \frac{1}{-\tau - |k_{R}|^{2} + \frac{i}{T}} \prod_{k_{R}}^{n} \frac{1}{-\tau - |k_{R}|^{2} - \sum_{j=k}^{n} \Omega_{j}} \right| \lesssim \tau^{-n-1}. \]

Injecting these estimates in \( \mathcal{F}_{2} \), we see that

\[ |\mathcal{F}_{2}| \lesssim \lambda^{4n} e^{d(2n+1)} \sum_{|k_{f}|,|k_{f}^{i}| \leq 1} \int_{|\tau| > K^{\epsilon-2}} \frac{1}{\tau^{2n+2-2b}} d\tau \lesssim T^{2} \lambda^{4n} e^{d(2n+1)} \epsilon^{-d(2n+1)} \epsilon^{2((2n+1)-2b)} \]

\[ \lesssim \lambda^{4n} e^{2n} T^{n} (2T^{-1})^{2n} \epsilon^{2(1-2b)} \ll \lambda^{4n} e^{2n} T^{n} \epsilon^{2(1-2b)}, \]

which is smaller than the estimates obtained on \( \mathcal{F}_{1} \).
5.5. **Bound on the error.** First, notice that the Fourier support of the approximate solution makes the choice of $s$ irrelevant in our scaled Sobolev and Bourgain spaces. Second, we discard the second summand in the definition of $E_s^N$, since it is easier to estimate. Next, by (A.2)

$$
\left\| \int_0^t e^{i(t-s)\Delta} E_s^N \, ds \right\|_{X_{s,b}^b} \lesssim \sum_{i,j,k \leq N} u_i^j w_{i+j+k}^k.
$$

We now proceed as usual by interpolating $X_{s,b}^{b-1}$ between $X_{s,b}^{b'}$, with $b' < \frac{1}{2}$, and a trivial but lossy bound in $X_{s,0}$. Omitting details for the latter, we focus on the former and choose $b' < -\frac{1}{2}$. Then, by (A.4),

$$
\left\| \chi(t) \sum_{i,j,k \leq N} u_i^j w_{i+j+k}^k \right\|_{X_{s,b}^{b-1}} \lesssim \left\| \chi(t) \sum_{i,j,k \leq N} u_i^j w_{i+j+k}^k \right\|_{L_t^{4/3}L_y^{4/3}} \lesssim \sum_{i,j,k \leq N} \left\| u_i^j \right\|_{L_t^{4}L_y^{4}} \left\| w_{i+j+k}^k \right\|_{L_t^{4}L_y^{4}}.
$$

By Hölder’s inequality and the $L^p$ estimates of Proposition 5.1

$$
\mathbb{E} \left\| \chi(t) \sum_{i,j,k \leq N} u_i^j w_{i+j+k}^k \right\|_{X_{s,b}^{b-1}} \lesssim \left( \frac{1}{T_{kin}} \right)^{N \frac{2}{d}}
$$

6. **Control of the linearization around $u_{\text{app}}$: proof of Proposition 3.3**

We consider here the operator

$$
\mathcal{L}(u) = 2\mathcal{L} + \mathcal{L}'
$$

where

$$
\mathcal{L} : f \mapsto \lambda^2 \chi(t) \left( f |u_{\text{app}}|^2 - fV - \langle u_{\text{app}}, f \rangle u_{\text{app}} \right)
$$

and

$$
\mathcal{L}' : f \mapsto \lambda^2 \chi(t) \left( (u_{\text{app}})^2 u - 2 \langle u, u_{\text{app}} \rangle u_{\text{app}} \right)
$$

and aim at proving Proposition 3.3. We only prove the corresponding bound for the operator $\mathcal{L}$. Indeed, the proof for $\mathcal{L}'$ is verbatim the same.

6.1. **Reduction to elementary operators.** The operator $\mathcal{L}$ will be decomposed into

$$
\mathcal{L} = \sum_{i,j=1}^{n} \mathcal{L}_{i,j} f, \quad \mathcal{L}_{i,j} = \lambda^2 \chi(t) \left( f u_i^j u^j - f \langle u^j, u^j \rangle - \langle u^j, f \rangle u^j \right).
$$

and each $\mathcal{L}_{i,j}$ is localised in frequency by letting

$$
\mathcal{L}_{i,j,n} = \mathcal{L}_{i,j} Q_{\epsilon,N}^n
$$
(the value of $N$ will be fixed shortly). The following Lemma gives an upper bound for these operators.

**Lemma 6.1.** For any $0 < \kappa \ll 1$, for $\epsilon$ small enough, there exists a set of measure greater than $1 - \epsilon^\kappa$ such that on this set, for all $n \in \mathbb{Z}^d$ and $(i, j) \in \{0, N\}^2$ we have the following estimates for the operator norms:

$$\| \mathfrak{L}_{i,j,n} \|_{X^{0,0,0,-\frac{1}{2}} \to X^{0,0,0,-\frac{1}{2}}} \lesssim \left( \frac{1}{T_{kin}} \right)^{\frac{i+j+1}{2}} \epsilon^{-\kappa}$$

With the help of the above Lemma, we are able to prove Proposition 3.3.

**Proof of Proposition 3.3 from Lemma 6.1.** Almost locality. We decompose the input and output function in frequency cubes according to:

$$\mathfrak{L}_{i,j} u = \sum_{n,n' \in \mathbb{Z}^d} Q_{\epsilon,N}^{n'} \mathfrak{L}_{i,j} Q_{\epsilon,N}^{n} u.$$  

Since $\mathfrak{L}_{i,j}$ corresponds to convolution in frequency with a kernel localized in a frequency ball of size $C\epsilon^{-1}$, the operator is almost local: for $N$ sufficiently large,

$$Q_{\epsilon,N}^{n'} \mathfrak{L}_{i,j} Q_{\epsilon,N}^{n} u = 0 \quad \text{if} \quad |n - n'| \geq 2.$$  

**Bound from $X_{s,\frac{1}{2}}$ to $X_{s,-\frac{1}{2}}$.** In this case, by almost locality,

$$\| \mathfrak{L}_{i,j} Q_{\epsilon,N}^{n} \|_{X_{s,\frac{1}{2}} \to X_{s,-\frac{1}{2}}} \sim \| \mathfrak{L}_{i,j} Q_{\epsilon,N}^{n} \|_{X_{s,\frac{1}{2}} \to X_{s,-\frac{1}{2}}}$$

and by almost orthogonality,

$$\| \mathfrak{L}_{i,j} u \|_{X_{s,-\frac{1}{2}}} \lesssim \left[ \sum_{n \in \mathbb{Z}^d} \| \mathfrak{L}_{i,j} Q_{\epsilon,N}^{n} u \|^2_{X_{s,-\frac{1}{2}}} \right]^{1/2} \lesssim \left( \sup_{n \in \mathbb{Z}^d} \| \mathfrak{L}_{i,j,n} \|_{X_{s,\frac{1}{2}} \to X_{s,-\frac{1}{2}}} \right) \left[ \sum_{n \in \mathbb{Z}^d} \| Q_{\epsilon,N}^{n} u \|^2_{X_{s,\frac{1}{2}}} \right]^{1/2}$$

on the set $E_\kappa$.

**Bound from $X_{s,0}^* \to X_{s,0}^*$ and interpolation.** Since $X_{s,0}^*$ is simply $L_2^2 H_s^*$, and since $u_i$ and $u_j$ are localized in frequency in a ball of radius $C\epsilon^{-1}$, the operator norm of $\mathfrak{L}_{i,j}$ is less than $\| u_i \|_{L^\infty} \| u_j \|_{L^\infty}$, which in turn can be bounded very roughly by $\epsilon^{-(2i+2j+2)d}$. Interpolating between this very rough bound and the $X_{s,\frac{1}{2}} \to X_{s,-\frac{1}{2}}$ bound, we obtain a bound from $X_{s,\frac{1}{2}-\delta}$ to $X_{s,-\frac{1}{2}+\delta}$ with a loss $\epsilon^{-\kappa}$, where $\kappa$ can be made arbitrarily small by choosing $\delta$ sufficiently small. There remains to choose $b > \frac{1}{2}$ such that $b - 1 < -\frac{1}{2} + \delta$. □

6.2. **The trace of the iterated operator and its diagrammatic representation.** The key idea will be to use the inequality $\| T \| \lesssim \| (T^* T)^N \|^{1/2N}$ to relate the control of $T$ to that of a nonnegative self-adjoint operator, and then to use the control of the operator norm by the trace $\| (T^* T)^N \| \leq Tr(T^* T)^N$, together with $N \to \infty$ to eliminate the size of the cube $\epsilon^{-d}$ contribution in the trace.
By transferring the weight from the function space to the operator, and getting rid of irrelevant constants, the operator norm of $\mathcal{L}_{t,j,n}$ from $X^0 \frac{1}{2}$ to $X^0 \frac{1}{2}$ equals that of the operator $\mathcal{R}$

$$\mathcal{R} : L^2(\mathbb{R} \times \mathbb{Z}^d) \to L^2(\mathbb{R} \times \mathbb{Z}^d)$$

with kernel

$$K(\tau_3, \tau_0, k_3, k_0) = \lambda^2(\tau_0 + |k_0|^2)^{-\frac{1}{2}}(\tau_3 + |k_3|^2)^{-\frac{1}{2}} \sum_{k_0 + k_1 + k_2 = k_3} 1_{C^{n,N}}(k_0)$$

$$\int_{\mathbb{R}^2} \tilde{u}^*(k_1, \tau_1) \tilde{\omega}^*(k_2, \tau_2)(1 - \delta(k_1 + k_2) - \delta(k_2 + k_0)) \tilde{\chi}(\tau_3 - \tau_0 - \tau_1 - \tau_2) d\tau_1 d\tau_2.$$

To compute the adjoint kernel, we change variables by setting $(k_0', k_1', k_2', k_3') = (k_3, -k_2, -k_1, k_0)$ and $(\tau_0', \tau_1', \tau_2', \tau_3') = (\tau_3, -\tau_2, -\tau_1, \tau_0)$. Getting rid of primes gives the following formula for the adjoint kernel:

$$K^*(\tau_3, \tau_0, k_3, k_0) = \lambda^2(\tau_0 + |k_0|^2)^{-\frac{1}{2}}(\tau_3 + |k_3|^2)^{-\frac{1}{2}} \sum_{k_0 + k_1 + k_2 = k_3} 1_{C^{n,N}}(k_3)$$

$$\tilde{u}^*(k_1, \tau_1) \tilde{\omega}^*(k_2, \tau_2)(1 - \delta(k_1 + k_2) - \delta(k_2 + k_0)) \tilde{\chi}(\tau_3 - \tau_0 - \tau_1 - \tau_2) d\tau_1 d\tau_2$$

(here we are using that $\chi$ is even, which can be assumed without loss of generality). Transposing and iterating this operator, we see that the operator $(\mathfrak{M})^N = ((\mathcal{R})^* \mathcal{R})^N$ has kernel

$$M^N(\tau_{6N}, \tau_0, k_{6N}, k_0)$$

$$= \lambda^{4N} \sum_{k_1, \ldots, k_{6N-1}} \int_{\mathbb{R}^{6N-1}} \langle \tau_{6N} + |k_{6N}|^2 \rangle^{-1/2} \langle \tau_0 + |k_0|^2 \rangle^{-1/2} \Delta(k) \prod_{m=0}^{N-1} 1_{C^{n,N}}(k_{6m})$$

$$\prod_{m=0}^{N-1} \tilde{u}^*(k_{6m+1}, \tau_{6m+1}) \tilde{\omega}^*(k_{6m+2}, \tau_{6m+2}) \tilde{u}^*(k_{6m+4}, \tau_{6m+4}) \tilde{\omega}^*(k_{6m+5}, \tau_{6m+5})$$

$$\prod_{m=1}^{2N-1} \langle \tau_{3m} + |k_{3m}|^2 \rangle^{-1/2} \prod_{m=0}^{2N-1} \tilde{\chi}(\tau_{3m+3} - \tau_{3m} - \tau_{3m+1} - \tau_{3m+2}) d\tau_1 \ldots d\tau_{6N-1}$$

where $k = (k_0, \ldots, k_{6N})$ and

$$\Delta(k) = \prod_{m=0}^{2N-1} \delta(k_{3m+3} - k_{3m} - k_{3m+1} - k_{3m+2})(1 - \delta(k_{3m+1} + k_{3m+2}) - \delta(k_{3m+1} + k_{3m}))$$

Setting

$$\omega_{3m} = \tau_{3m} + |k_{3m}|^2, \quad \omega_{3m+1} = \tau_{3m+1} + |k_{3m+1}|^2, \quad \omega_{3m+2} = \tau_{3m+2} - |k_{3m+2}|^2$$

$$\Omega_m = \tau_{3m+3} - \tau_{3m} - \tau_{3m+1} - \tau_{3m+2}$$

$$= - |k_{3m+3}|^2 + |k_{3m}|^2 + |k_{3m+1}|^2 - |k_{3m+2}|^2 + \omega_{3m+3} - \omega_{3m} - \omega_{3m+1} - \omega_{3m+2},$$
this becomes

\[ M^N(\tau_0, k_0, k_{0N}) = \lambda^4N \sum_{k_{1},\ldots,k_{6N-1}} \langle \omega_0 \rangle^{\frac{1}{2}} \langle \omega_{6N} \rangle^{\frac{1}{2}} \Delta(k) \prod_{m=0}^{N} 1_{\epsilon,N}(k_{6m}) \]

\[ \int_{\mathbb{R}^{6N-1}} \prod_{m=0}^{N-1} \langle \omega_{6m} \rangle^{-1} \prod_{m=0}^{2N-1} \chi(\Omega_m) d\omega \]

Taking the trace gives

\[ \text{Tr}(\mathcal{M})^N = \lambda^4N \sum_{k} \int_{\mathbb{R}^{n+1}} \Delta(k) \delta(\omega_0 - \omega_{6N}) \delta(k_0 - k_{6N}) \prod_{m=0}^{N} 1_{\epsilon,N}(k_{6m}) \int_{\mathbb{R}^{6N-1}} \prod_{m=0}^{N-1} \langle \omega_{6m} \rangle^{-1} \prod_{m=0}^{2N-1} \chi(\Omega_m) d\omega \]

This can be represented by the following interaction diagram, in which the input and output frequencies are \( k_0 \) and \( k_{0N} \) respectively, and are equal since the trace was taken.

Using (4.5) (without loss of generality we place a \( \chi(t) \) in front of each \( u^i \) and \( w^j \)), once the expectation is taken, getting rid of all irrelevant constants, and using the localisation \( A(\epsilon k) \lesssim 1(|k| \leq \epsilon^{-1}) \), we end up with the following upper bound:

\[ \mathbb{E} \text{Tr} \mathcal{M}^N = \mathbb{E} \text{Tr} \mathcal{M}^N = \lambda^4N(i+j+1) 2^{N} d(i+j+1) \sum_{\vec{l}=(l_1,l_2,\ldots,l_{4N}) \in \mathcal{G}_l \times \cdots \times \mathcal{G}_j} \sum_{P} \mathcal{F}(\vec{l}, P), \quad (6.1) \]
where $\bar{\ell} = (\ell_1, ..., \ell_{4N})$ represents all possible interaction histories for the $u^i$ and $u^j$ terms, and $P$ is summed over all possible non-degenerate pairings of the initial vertices, and where (for $T = 1$):

$$F(\bar{\ell}, P) = \sum_{k} \int d\tau_0 d\tau_1 ... d\tau_{6N} P(k) \delta(\tau_0 - \tau_{6N}) \delta(k_0 - k_{6N}) \prod_{\{i,j\} \in P} 1(|ek_{(i,j)}|) \leq C e^{-1}$$

$$\prod_{m=0}^{2N-1} \chi(\Omega_m) \chi(\bar{\omega}_m) - 1 \frac{\chi(\bar{\omega}_m+1)}{\omega_m + 1 + \frac{1}{T}} |\omega_m + 2 - \bar{\omega}_m| ^{\frac{1}{T}} d\omega_m$$

where for $n = 1, 2$:

$$\bar{\omega}_{3m+n} = \tau_{3m+n} - \bar{\tau}_{3m+n},$$

and where $\Delta_{\ell_1, ..., \ell_{6N-1}, P(k)}$ records all Kirchhoff laws of the graph. Once a pairing has been fixed, we can represent formula (6.2) as an interaction diagram in which initial vertices are paired. This is done the same way as in Subsection 4.1. Then we construct a spanning tree for the graph the same way as in Theorem 4.5. We proceed to the construction interaction vertices by interaction vertices, first covering each graph of $u^i$ and $u^j$ one after another, from right to left. Once we are done we proceed to the examination of the interaction vertices between $k_{2m}$, $k_{2m+1}$ and $k_{2m+2}$, going from right to left. The edge representing the output $k_{6N}$ is always let free.

We denote by $\tilde{n}$ the number of free edges ending to or starting from the top vertices. We denote by $k_1^f, ..., k_{\tilde{n}}^f$ the corresponding free frequencies, where the order is from right to left in the graph. The spanning tree then yields a collection of free edges $(k_1^f)_{1 \leq i \leq \tilde{n}}$ and $(\bar{k}_i^f)_{1 \leq i \leq \tilde{n}}$ from which all other edges are recovered in the graph using the Kirchhoff laws. An example for free edges for the $\Omega_m$’s vertices is as follows:

Not that since in the spanning tree algorithm top vertices are considered after the vertices associated to the graphs of $u^i$ and $u^j$, we have that integrated frequencies below top vertices are only sums of free frequencies associated to top vertices. Namely, for all $m \in \{0, ..., 6N\}$, if $k_m$ is not a free frequency, then $k_m$ is a sum of free frequencies among $k_1^f, ..., k_{\tilde{n}}^f$ only, that moreover appear
after \( k_m \) for the natural ordering of the graph.

We define degree zero, one and two vertices as in Subsection 4.5 and Lemma 4.6 holds true. We denote by \( n_0, n_1 \) and \( n_2 \) the number of degree zero, one and two interaction vertices inside a graph for either \( u^i \) or \( u^j \). The other vertices are those associated with the phases \( \Omega_1, \ldots, \Omega_{2N - 1} \) on top of the graph and we denote by \( \tilde{n}_0, \tilde{n}_1 \) and \( \tilde{n}_2 \) the number of degree zero, one and two vertices among these vertices. In particular we get:

\[
n_0 + \tilde{n}_0 + n_1 + \tilde{n}_1 + n_2 + \tilde{n}_2 = 2N(i + j + 1), \quad \text{and} \quad n_1 + \tilde{n}_1 + 2n_2 + 2\tilde{n}_2 = 2N(i + j + 1).
\]

6.3. Expectation of the trace. We are now ready to estimate (6.2). First, we notice that as in the proof for the \( X^{s,b} \) estimate for \( u^i \), if one takes \( |\omega_{3m+1}| \geq \epsilon^{-K} \) or \( |\omega_{3m+2}| \geq \epsilon^{-K} \) then the contribution of the third line in (6.2), once integrated, is negligible of irrelevant size \( \epsilon^{K/2} \). Hence we focus on the case for which \( |\omega_{3m+1}|, |\omega_{3m+2}| \leq \epsilon^{-K} \) for \( m = 0, \ldots, 2N - 1 \). We first integrate over all vertices inside the graphs for \( u^i \) and \( u^j \), according the natural time ordering of the graph, from right to left. Using Lemmas B.1 and B.2 this produces for any \( K \gg 1 \):

\[
\mathcal{F}(\tilde{f}, P) \lesssim (\ln \epsilon)^C (\epsilon^{-2(d-1)} n_2 e^{-(d-1)n_1}) \sum_{k^f} \left\{ \Delta_{i_1, \ldots, i_{6N-1}, P(k)} \delta(\tau_0 - \tau_{6N}) \delta(k_0 - k_{6N}) \right\}
\]

\[
\prod_{m=0}^{2N-1} \hat{\chi}(\Omega_m) (\langle \omega_{3m} \rangle^{-1} \hat{\chi}(\omega_{3m+1}) - \omega_{3m+1} + \frac{1}{T}) (\omega_{3m+2} - \omega_{3m+2} + \frac{1}{T})
\]

\[
d\omega_{3m} d\omega_{3m+1} d\omega_{3m+2} d\hat{\omega}_{3m+1} d\hat{\omega}_{3m+2} \sum_{m=0}^{2N-1} 1(|\omega_{3m+1}|, |\omega_{3m+2}| \leq \epsilon^{-K}) 1(|k_f| \leq C \epsilon^{-1}) + O(\epsilon^K)
\]

where \( \hat{K} \gg K \) is another very large constant. The first part \( \mathcal{F}_1 \) gives an irrelevant contribution. Indeed, we write:

\[
\mathcal{F}_1(G, P) \leq \sum_{m=0}^{2N-1} \int_{S_m} [...]\]

where we define \( S_m = \{ |\omega_{3m}| \geq \epsilon^{-K} \} \). Notice that on \( S_m \) that we have \( \Omega_m = \omega_{3m+3} - \omega_{3m} + O(C \epsilon^{-K}) \) so that if all momenta are kept fixed:

\[
\int_{|\omega_{3m+3}| \leq |\omega_{3m}|} \langle \omega_{3m+3} \rangle^{-1} \hat{\chi}(\Omega_m) d\omega_{3m+3} \lesssim \langle \omega_{3m} \rangle^{-1}.
\]

Hence, on \( S_m \), integrating first with respect to \( \omega_{3m+3} \) using the above estimate, then integrating over all \( \omega_{3m'} \) for \( m' \neq m \) producing a \( (\ln \epsilon) \) factor, then integrating over all \( k_f \) producing an \( \epsilon^{-d(n_1 + 2\tilde{n}_2)} \) factor, and over \( k_{6N} \) producing an \( \epsilon^{-d} \) factor, we arrive at:

\[
\int_{S_m} [...] \lesssim (\ln \epsilon)^C \epsilon^{-O(N)} \int_{|\omega_{3m}| \geq \epsilon^{-K}} \langle \omega_{3m} \rangle^{-2} d\omega_{3m} \lesssim \epsilon^{\frac{K}{2}} = O(\epsilon^K).
\]
This shows that $\mathcal{F}_1$ gives an irrelevant contribution:

$$
\mathcal{F}_1(G, P) = O(\epsilon^K).
$$

We now claim that for any $P$ and $\vec{l}$,

$$
\mathcal{F}_2(\vec{l}, P) \lesssim \langle \ln \epsilon \rangle^C \epsilon^{-2Nd(i+j+1)+2N(i+j+1)-d}
$$

To estimate $\mathcal{F}_2$, we integrate iteratively at interaction vertices from left to right, each time over the free variables below it. When the vertex is of degree zero we get a factor 1, when it is of degree one we use Lemma 3.2 and when it is of degree 2 we use Lemma B.1.

At the end of the integration process, we integrate over the last free variable $k_{6N}$ which produces an $\epsilon^{-d}$ factor, and integrate $\langle \omega_i \rangle^{-1} d\omega_i$ and $\hat{\chi}(\omega_i) d\omega_i$ over the $\omega_i$ and $\hat{\omega}_i$'s variables which, from the constraint $|\omega_i| \leq \epsilon^{-K}$ for all $i$, produces a factor $\langle \ln \epsilon \rangle^{2N}$. This produces the following estimate:

$$
\mathcal{F}(\vec{l}, P) \lesssim \langle \ln \epsilon \rangle^C \epsilon^{-2(d-1)} n_2 \epsilon^{-(d-1)n_1} \epsilon^{-2(d-1)n_2} \epsilon^{-(d-1)n_1} \epsilon^{-d} \langle \ln \epsilon \rangle^C \epsilon^{-2Nd(i+j+1)+2N(i+j+1)-d}
$$

(6.4)

where we used (6.3) and the fact that $\tilde{n}_2 + \tilde{n}_1 + \tilde{n}_0 = 2N$ and $n_0 + \tilde{n}_0 + (n_1 + \tilde{n}_1)/2 = N(i + j + 1)$. This estimate concludes the proof of the Lemma.

6.4. End of the proof of Lemma 6.1. We can now end the proof of the Lemma. From the identity (6.1) and (6.4), we obtain that the trace of $\mathfrak{M}^N$ is independent of $n$ and satisfies:

$$
\mathbb{E} \text{Tr} \mathfrak{M}^N \lesssim (\frac{1}{T_{\text{kin}}})^N(i+j+1) \langle \ln \epsilon \rangle^C \epsilon^{-2Nd(i+j+1)+2N(i+j+1)-d}
$$

Hence, via Bienaymé-Tchebychev, for any $\kappa$, there exists a set of measure greater than $1 - \epsilon^\kappa$ such that:

$$
\text{Tr} \mathfrak{M}^N \lesssim (\frac{1}{T_{\text{kin}}})^N(i+j+1) \epsilon^{-d-\kappa} \langle \ln \epsilon \rangle^C \text{ for all } n \in \mathbb{Z}^d.
$$

Then, for all $n \in \mathbb{Z}^d$, on this set:

$$
\|\Theta_{i,j,n}\|_{X^{n,\frac{1}{2}} \to X^{0,-\frac{1}{2}}} \leq (\text{Tr} \mathfrak{M}^N)^{\frac{1}{2N}} \lesssim (\frac{1}{T_{\text{kin}}})^{\frac{i+j+1}{2}} \epsilon^{-\frac{d+\kappa}{2N}} \langle \ln \epsilon \rangle^C \lesssim (\frac{1}{T_{\text{kin}}})^{\frac{i+j+1}{2}} \epsilon^{-\kappa}
$$

for $N$ large enough.

7. The nonlinear terms: proof of propositions 3.4 and 3.5

We will consider in this section that

$$
\mathcal{B}(u) = \lambda^2 \left[ |u|^2 v_{pp} + u^2 v_{ppp} \right]
$$

$$
\mathcal{T}(u) = \lambda^2 \left( |u|^2 u - 2 \|u\|_{L^2}^2 u \right).
$$

Indeed, the additional terms in the definition of $\mathcal{B}$ and $\mathcal{T}$ in Section (3) can be treated similarly (actually, even more simply). We will fix the time $T = 1$ and adapt accordingly the notation $L^p_t L^q$ to $L^p_1 L^q$. 
7.1. A simplified case: restricting to low frequencies. We add here a projector on freq \( \lesssim \frac{1}{\epsilon} \), which is the range of interest physically. It greatly simplifies nonlinear estimates and in particular

\[
\| Q^0_{\epsilon,1} f \|_{X^s,b} \sim \| f \|_{X^{0,0}}
\]

(the projection operator \( Q^0_{\epsilon,1} \) is defined in Section 2). We will denote

\[
u_{\epsilon} = Q^0_{\epsilon,1} u.
\]

**Proposition 7.1.** For any \( \kappa > 0 \), there exists \( b > \frac{1}{2} \) such that

\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \, ds \right\|_{X^{0,b}} \lesssim \epsilon^{2-d-\kappa} \| u_{\epsilon} \|_{X^{0,0}}^3
\]

**Proof.** First, by (A.2), for \( b > \frac{1}{2} \),

\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \, ds \right\|_{X^{0,b}} \lesssim_b \| \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \|_{X^{0,b-1}}.
\]

Estimate \((X^{0,b})^3 \to X^{0,b'}\), with \( b > \frac{1}{2}, b' < -\frac{1}{2}\). Applying successively (A.4), Hölder’s inequality, and (A.3), we get for \( b > \frac{1}{2} \)

\[
\left\| \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \right\|_{X^{0,b'}} \lesssim_{b',\kappa} \epsilon^{\frac{1}{4} - \frac{d}{2} - \kappa} \| u_{\epsilon} \|^3_{L^4 L^4} \lesssim_{b',\kappa} \epsilon^{\frac{1}{2} - \frac{d}{4} - \kappa} \| u_{\epsilon} \|^3_{L^4 L^4} \lesssim_{b,\kappa} \epsilon^{2-d-4\kappa} \| u_{\epsilon} \|^3_{X^{0,0}}.
\]

Estimate \((X^{0,b})^3 \to X^{0,0}\) for \( b > \frac{1}{2} \). Using successively the definition of \( X^{s,b} \), Hölder’s inequality, the Sobolev embedding theorem and (A.3),

\[
\left\| \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \right\|_{X^{0,0}} \lesssim \| |u_{\epsilon}|^2 u_{\epsilon} \|_{L^2 L^2} \leq \| u_{\epsilon} \|^3_{L^\infty L^6} \lesssim \epsilon^{-d} \| u_{\epsilon} \|^3_{L^\infty L^2} \lesssim_b \epsilon^{-d} \| u_{\epsilon} \|^3_{X^{0,0}}.
\]

Interpolation. Interpolating between the two estimates above, taking first \( b' \) close enough to \(-1/2\) and then \( b \) close enough to \(1/2\), gives that for any \( \kappa > 0 \), there exists \( b > \frac{1}{2} \) such that

\[
\left\| \chi(t) |u_{\epsilon}|^2 u_{\epsilon} \right\|_{X^{0,b-1}} \lesssim_{b,\kappa} \epsilon^{2-d-\kappa} \| u_{\epsilon} \|^3_{X^{0,0}}.
\]

**Proposition 7.2.** For any \( \kappa > 0 \), if \( \mu \) in Corollary 3.1 is chosen sufficiently small, there exists \( b > \frac{1}{2} \) such that

\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \chi(t) |u_{\epsilon}|^2 e^{\mu app} \, ds \right\|_{X^{0,b}} \lesssim \epsilon^{-\kappa} T^{3/4} \epsilon^{\frac{1}{2} - \frac{d}{2}} \| u_{\epsilon} \|^2_{X^{0,0}}.
\]

The same result holds if \( |u_{\epsilon}|^2 e^{\mu app} \) is replaced by \( (u_{\epsilon})^2 e^{\mu app} \).
Proof. First, by (A.2), for $b > \frac{1}{2}$,
\[
\left\| \chi(t) \int_0^t e^{i(t-s)\Delta} \chi(t)|u_\epsilon|^2 v_{\text{app}} \, ds \right\|_{X^{0,b}} \lesssim_b \left\| \chi(t)|u_\epsilon|^2 v_{\text{app}} \right\|_{X^{0,b-1}}.
\]

Estimate $(X^{0,b})^3 \to X^{0,b'}$, with $b > \frac{1}{2}$, $b' < -\frac{1}{2}$, applying successively Bernstein’s inequality, Hölder’s inequality, and Corollary 3.1, we get for $p > q$,
\[
\|v_{\text{app}}\|_{L^1_t L^\infty_x} \lesssim \epsilon^{-d/p}\|v_{\text{app}}\|_{L^1_t L^p_x} \lesssim \epsilon^{-\frac{d}{2}}\|v_{\text{app}}\|_{L^1_t L^p_x} \lesssim \epsilon^{-\mu\epsilon^{-\frac{2}{2}} (\epsilon-1) \frac{1}{2} - \frac{1}{p}} = \epsilon^{-\frac{1}{2} + \frac{d+\mu}{p}}.
\]
As a consequence, given $\kappa > 0$, if $\mu > 0$ is chosen sufficiently small and $p$ sufficiently large,
\[
\|v_{\text{app}}\|_{L^p_t L^\infty_x} \lesssim \epsilon^{-\frac{1}{2} - \kappa}.
\]
Therefore, taking $q = 1/2$, by (A.4), Hölder’s inequality, and (A.3),
\[
\|\chi(t)|u_\epsilon|^2 v_{\text{app}}\|_{X^{0,b'}} \lesssim \|\chi(t)|u_\epsilon|^2 v_{\text{app}}\|_{L^1_t L^2_x} \lesssim \|v_{\text{app}}\|_{L^2_t L^\infty_x} \|u_\epsilon\|_{L^4_t L^4_x}^2 \lesssim \epsilon^{-\frac{1}{2} - \mu} \epsilon^{\frac{1}{2} - \kappa} \|u_\epsilon\|_{X^{0,b}}^2.
\]

Estimate $(X^{0,b})^3 \to X^{0,0}$ for $b > \frac{1}{2}$,
\[
\|\chi(t)|u_\epsilon|^2 v_{\text{app}}\|_{X^{0,0}} \lesssim \|\chi(t)|u_\epsilon|^2 v_{\text{app}}\|_{L^2_t L^2_x} \lesssim \|v_{\text{app}}\|_{L^\infty_t L^6_x} \|u_\epsilon\|_{L^\infty_t L^6_x}^2 \lesssim \epsilon^{-2d/3} \|u_\epsilon\|_{L^\infty_t L^6_x}^2 \lesssim \epsilon^{-2d/3} \|u_\epsilon\|_{X^{0,0}}^2.
\]

Interpolation. Interpolating between the two estimates above gives the desired result if $\mu$ is chosen sufficiently small and $p$ sufficiently large. \hfill $\square$

7.2. Proof of Proposition 3.5: the trilinear bound. We aim at proving that for $s > \frac{d}{2} - 1$, for any $\kappa > 0$ there exists $b > \frac{1}{2}$ such that
\[
\|\chi(t)\int_0^t e^{i(t-s)\Delta} \chi(t)|u|^2 u \, ds\|_{X^{s,b}} \lesssim \epsilon^{2-d-\kappa} \|u\|_{X^{s,b}}^3.
\]

The starting point is (A.2), which gives, for $b > \frac{1}{2}$,
\[
\|\chi(t)\int_0^t e^{i(t-s)\Delta} \chi(t)|u|^2 u \, ds\|_{X^{s,b}} \lesssim_b \|\chi(t)|u|^2 u\|_{X^{s,b-1}}.
\]

By duality, it will therefore suffice to show that
\[
\sup_{\|v\|_{X^{-s,1-b}} \leq 1} \iint \chi(t)|u|^2 uv \, dx \, dt \lesssim \epsilon^{2-d-\kappa} \|u\|_{X^{s,b}}^3.
\]
As a preparation for this estimate, we will use the following lemma. Recall that the projection operators $P_{\epsilon,N}$ and $Q_{\epsilon,N}^0$ are defined in Section 2.

Lemma 7.3. If $N_1 \leq N_2$, for any $\kappa > 0$ there exists $b_0 < \frac{1}{2}$ such that
\[
\|P_{\epsilon,N_1} u P_{\epsilon,N_2} v\|_{L^2_t L^2_x} \lesssim N_1^{\frac{d}{2} - 1 + \kappa} \epsilon^{-\frac{d}{2} + 1 - \kappa} \|P_{\epsilon,N_1} u\|_{X^{0,b_0}} \|P_{\epsilon,N_2} v\|_{X^{0,b_0}}.
\]
The same holds if $u$ or $v$ are replaced by their complex conjugates.
Proof. Step 1: the estimate \((X_t^{0.4+})^2 \to L_t^6 L_x^2\). By almost orthogonality followed by Hölder’s inequality,

\[ \| P_{\epsilon,N_1} u P_{\epsilon,N_2} v \|_{L_t^6 L_x^2}^2 \lesssim \sum_{n \in \mathbb{Z}^d} \| P_{\epsilon,N_1} u Q_{\epsilon,N_1} P_{\epsilon,N_2} v \|_{L_t^4 L_x^4}^2 \lesssim \sum_{n} \| P_{\epsilon,N_1} u \|_{L_t^4 L_x^4}^2 \| Q_{\epsilon,N_1} P_{\epsilon,N_2} v \|_{L_t^4 L_x^4}^2. \]

Applying now the Strichartz estimates, followed once again by almost orthogonality, we get for \( b > \frac{1}{2} \)

\[ \cdots \lesssim_b N_1^{-d-1} \epsilon^{-d+1} \sum_n \| P_{\epsilon,N_1} u \|_{X_t^{0,b}}^2 \| Q_{\epsilon,N_1} P_{\epsilon,N_2} v \|_{X_t^{0,b}}^2 \lesssim N_1^{-d-1} \epsilon^{-d+1} \| P_{\epsilon,N_1} u \|_{X_t^{0,b}}^2 \| P_{\epsilon,N_2} v \|_{X_t^{0,b}}^2. \]

Step 2: the estimate \((X_t^{0.4+})^2 \to L_t^6 L_x^2\) Interpolating between the inequalities \(\|v\|_{L_t^\infty L_x^2} \lesssim \|v\|_{X_t^{0,b}}\) if \( b > \frac{1}{2} \), and \(\|v\|_{L_t^2 L_x^2} = \|v\|_{X_t^{0,0}}\) gives for any \( \delta > 0 \)

\[ \|v\|_{L_t^4 L_x^2} \lesssim_b \|v\|_{X_t^{0,\frac{1}{4}+\delta}}. \]

Applying Hölder’s inequality, followed by Sobolev embedding and the above inequality,

\[ \| P_{\epsilon,N_1} u P_{\epsilon,N_2} v \|_{L_t^4 L_x^2} \lesssim \| P_{\epsilon,N_1} u \|_{L_t^4 L_x^{\infty}} \| P_{\epsilon,N_2} v \|_{L_t^4 L_x^2} \lesssim \epsilon^{-d/2} N_1^{-d/2} \| P_{\epsilon,N_1} u \|_{L_t^4 L_x^2} \| P_{\epsilon,N_2} v \|_{L_t^4 L_x^2} \lesssim \epsilon^{-d/2} N_1^{-d/2} \| P_{\epsilon,N_1} u \|_{X_t^{0,\frac{1}{4}+\delta}} \| P_{\epsilon,N_2} v \|_{X_t^{0,\frac{1}{4}+\delta}}. \]

Step 3: interpolation Interpolating between the results of Step 1 and Step 2 gives the desired result. \(\square\)

We can now proceed with the proof of (7.2). Assuming in what follows that \(\|v\|_{X_t^{-s,1-b}} \leq 1\), and omitting complex conjugation in the notations for simplicity, we split all the functions into dyadic frequency projections to obtain

\[ \iint \chi(t) |u|^2 \overline{v} \, dx \, dt = \sum_{N_1, N_2, N_3, N_4} \iint \chi(t) P_{\epsilon,N_1} u P_{\epsilon,N_2} u P_{\epsilon,N_3} v P_{\epsilon,N_4} v \, dx \, dt. \]

Without loss of generality, we can assume that \( N_1 \leq N_2 \leq N_3 \). We distinguish two cases.

Case 1: \( N_3 \sim N_4 \) By Cauchy-Schwarz’ inequality followed by Lemma (7.3)

\[ \left| \iint \chi(t) P_{\epsilon,N_1} u P_{\epsilon,N_2} u P_{\epsilon,N_3} v P_{\epsilon,N_4} v \, dx \, dt \right| \leq \| P_{\epsilon,N_1} u P_{\epsilon,N_3} u \|_{L_t^2 L_x^2} \| P_{\epsilon,N_2} u P_{\epsilon,N_4} v \|_{L_t^2 L_x^2} \leq \epsilon^{-d+2-\kappa} N_1^{-d-1+\kappa} N_2^{-d-1+\kappa} \| P_{\epsilon,N_1} u \|_{X_t^{0,b_0}} \| P_{\epsilon,N_2} u \|_{X_t^{0,b_0}} \| P_{\epsilon,N_3} u \|_{X_t^{0,b_0}} \| P_{\epsilon,N_4} v \|_{X_t^{0,b_0}} \leq \epsilon^{-d+2-\kappa} N_1^{-d-1+\kappa-s} N_2^{-d-1+\kappa-s} N_3^{-s} N_4^{-s} \| P_{\epsilon,N_1} u \|_{X_t^{s,b_0}} \| P_{\epsilon,N_2} u \|_{X_t^{s,b_0}} \| P_{\epsilon,N_3} u \|_{X_t^{s,b_0}} \| P_{\epsilon,N_4} v \|_{X_t^{-s,b_0}} \sim 1. \]

It is now easy to conclude that

\[ \sum_{N_1 \leq N_2 \leq N_3} \left| \iint P_{\epsilon,N_1} u P_{\epsilon,N_2} u P_{\epsilon,N_3} v P_{\epsilon,N_4} v \, dx \, dt \right| \lesssim \epsilon^{-d+2-\kappa} \|u\|_{X_t^{s,b_0}}^3 \|v\|_{X_t^{-s,b_0}}. \]

Indeed, the variables \( N_1 \) and \( N_2 \) simply contribute a geometric series as \( s > d/2 - 1 \), while the sum over \( N_3 \sim N_4 \) can be bounded by the Cauchy-Schwarz inequality.
Case 2: $N_4 \ll N_3$ By the same arguments as in Case 1,
\[
\left| \int \int \chi(t) P_{\epsilon, N_1} u \ P_{\epsilon, N_2} u \ P_{\epsilon, N_3} u \ u \ dx \ dt \right| \leq \| P_{\epsilon, N_1} u \ P_{\epsilon, N_3} u \ u \|_{L_{t}^{2} L^{2}} \| P_{\epsilon, N_4} u \ u \|_{L_{t}^{2} L^{2}} \\
\leq e^{-d+2-\kappa} N_1^{-\frac{d}{2}-1-\kappa} N_4^{-\frac{d}{2}-1-\kappa} \| P_{N_1} u \|_{X_{s, \epsilon_0}^{0, b}} \| P_{N_2} u \|_{X_{s, \epsilon_0}^{0, b}} \| P_{N_3} u \|_{X_{s, \epsilon_0}^{0, b}} \| P_{N_4} u \ u \|_{X_{s, \epsilon_0}^{0, b}} \\
\leq e^{-d+2-\kappa} N_1^{-\frac{d}{2}-1-\kappa-s} N_4^{-\frac{d}{2}-1-\kappa+s} N_2^{-s} N_3^{-s} \| P_{N_1} u \|_{X_{s, \epsilon_0}^{s, b}} \| P_{N_2} u \|_{X_{s, \epsilon_0}^{s, b}} \| P_{N_3} u \|_{X_{s, \epsilon_0}^{s, b}} \| P_{N_4} u \ u \|_{X_{s, \epsilon_0}^{s, b}}.
\]
The condition $N_4 \ll N_3$ and quasi-orthogonality implies $N_2 \sim N_3$, and
\[
\sum_{N_1 \leq N_2 \leq N_3 \atop N_4 \ll N_3} \left| \int \int \chi(t) P_{\epsilon, N_1} u \ P_{\epsilon, N_2} u \ P_{\epsilon, N_3} u \ u \ dx \ dt \right| \lesssim \epsilon^{-d+2-\kappa} \| u \|_{X_{s, \epsilon_0}^{s, b}}^{3} \| v \|_{X_{s, \epsilon_0}^{s, b}}.
\]
The estimate (7.2) now follows by combining Case 1 and Case 2, and by choosing $b$ such that $\frac{1}{2} < b < 1 - b_0$.

7.3. Proof of Proposition 3.4: the bilinear bound. Recall the identity $X^{0, b} = X_{s}^{0, b}$. First note that, by interpolating between $\| v \|_{L_{t}^{\infty} L^{2}} \lesssim \| v \|_{X_{s}^{0, b}}$, for $b > \frac{1}{2}$, and $\| v \|_{L_{t}^{2} L^{2}} = \| v \|_{X_{s}^{0, 0}}$, one obtains that for any $r < \infty$, there exists $\bar{b} < \frac{1}{2}$ such that
\[
\| v \|_{L_{t}^{r} L^{2}} \lesssim \| v \|_{X_{s}^{0, \bar{b}}}.
\]
Proceeding as in the previous subsection, we are to bound
\[
\sup_{\| v \|_{X_{s, \epsilon_0}^{0, \bar{b}}} \lesssim 1} \int \int \chi(t) v^{app} \| u \|^{2} v \ dx \ dt.
\]
Next, we localize $v, \bar{u}$ and $v$ in frequency, at scales $N_1$, $N_2$, and $N_3$ respectively. By symmetry, we can assume that $N_1 \leq N_2$; and since $v^{app}$ is localized in Fourier on $B(0, C^{-1})$, we can assume that $N_3 \ll N_2$. Therefore, it suffices to bound
\[
\sum_{N_1 \leq N_2 \leq N_3 \atop N_4 \ll N_3} \left| \int \int \chi(t) v^{app} P_{N_1} u P_{N_2} u P_{N_3} v \ dx \ dt \right|.
\]
Applying successively Hölder’s inequality with $\frac{1}{r} + \frac{1}{q} = \frac{1}{2}$, Lemma 7.3 and inequalities (7.1) and (7.3), and the above inequality this is
\[
\cdots \lesssim \| v^{app} \|_{L_{t}^{r} L^{\infty}} \sum_{N_1 \leq N_2 \atop N_3 \ll N_2} \| P_{N_1} u P_{N_2} u \|_{L_{t}^{2} L^{2}} \| P_{N_3} v \|_{L_{t}^{r} L^{2}} \\
\lesssim e^{-d+\frac{1}{2}-\kappa} \sum_{N_1 \leq N_2 \atop N_3 \ll N_2} N_1^{d-1-\kappa} \| P_{N_1} u \|_{X_{s, \epsilon_0}^{\epsilon_0, b}} \| P_{N_2} u \|_{X_{s, \epsilon_0}^{\epsilon_0, b}} \| P_{N_3} v \|_{X_{s, \epsilon_0}^{0, \bar{b}}} \\
\lesssim e^{-d+\frac{1}{2}-\kappa} \sum_{N_1 \leq N_2 \atop N_3 \ll N_2} N_1^{d-1-\kappa-s} N_2^{-s} N_3^{\kappa} \| P_{N_1} u \|_{X_{s, \epsilon_0}^{s, \bar{b}}} \| P_{N_2} u \|_{X_{s, \epsilon_0}^{s, \bar{b}}} \| P_{N_3} v \|_{X_{s, \epsilon_0}^{s, \bar{b}}}.
\]
Summing the geometric series in $N_1$, and applying Cauchy-Schwarz in $N_2$ and $N_3$, this is
\[
\cdots \lesssim e^{-d+\frac{1}{2}-\kappa} \| u \|_{X_{s, \epsilon_0}^{s, \bar{b}}}^{2} \| v \|_{X_{s, \epsilon_0}^{s, \bar{b}}}.
\]
There remains to fix $q$ as close to $2$ as desired, choose $\tilde{b} < \frac{1}{2}$ which allows for (7.3), and finally choose $b > \frac{1}{2}$ such that $1 - b > \tilde{b}$.

**Appendix A. Basics of $X^{s,b}$ spaces**

These spaces were introduced in [4]. We quickly review their properties, referring the reader to [25], Section 2.6, for details.

**Definition** Let

$$\|f\|_{H^s} = \|\langle \epsilon D \rangle^s f\|_{L^2}$$

and

$$\|u\|_{X^{s,b}} = \|e^{-it\Delta} u(t)\|_{L^2 H^s} = \|\langle \epsilon k \rangle^s (\tau + |k|^2)^{b/2} \tilde{u}(\tau, k)\|_{L^2(\mathbb{R} \times \mathbb{Z}^2)}$$

**Time continuity** For $b > \frac{1}{2}$,

$$\|u\|_{C^H^s} \lesssim \|u\|_{X^{s,b}}.$$  \hspace{1cm} (A.1)

**Inverting the linear Schrödinger equation** Assume that $u$ solves

$$\begin{cases} 
  i\partial_t u - \Delta u = F \\
  u(t=0) = 0
\end{cases}$$

Then, denoting $\chi$ for a smooth cutoff function, supported on $B(0, 2)$ and equal to $1$ on $B(0, 1)$,

$$\|\chi(t) u\|_{X^{s,b-1}} \lesssim \|F\|_{X^{s,b}}.$$  \hspace{1cm} (A.2)

**From group to $X^{s,b}$ estimates** Assume that, uniformly in $\tau_0 \in \mathbb{R}$,

$$\|e^{it\tau_0} e^{it\Delta} f\|_{Y} \leq C_0(\epsilon) \|\langle \epsilon D \rangle^s f\|_{L^2}$$

Then, if $b > \frac{1}{2}$,

$$\|u\|_{Y} \lesssim_b C_0(\epsilon) \|u\|_{X^{s,b}}$$

**Strichartz estimates** We want to apply the previous statement to Strichartz estimates: it was proved in [4] that, for $s > 0$ and $\kappa > 0$,

$$\|e^{it\Delta} f\|_{L^4_t L^4_x} \lesssim_{s,\kappa} \epsilon^{-\kappa} \|\langle \epsilon D \rangle^s f\|_{L^2} \quad \text{if } d = 2$$

$$\|e^{it\Delta} f\|_{L^4_t L^4_x} \lesssim_{s,\kappa} \epsilon^{\frac{1}{2} - \frac{1}{4}} \|\langle \epsilon D \rangle^{\frac{1}{2} - \frac{1}{4}} f\|_{L^2} \quad \text{if } d \geq 3.$$

As a consequence, if $s > 0$, $\kappa > 0$, $b > \frac{1}{2}$,

$$\|u\|_{L^4_t L^4_x} \lesssim_{s,\kappa,b} \epsilon^{-\kappa} \|u\|_{X^{s,b}} \quad \text{if } d = 2$$

$$\|u\|_{L^4_t L^4_x} \lesssim_{s,\kappa,b} \epsilon^{\frac{1}{2} - \frac{1}{4}} \|u\|_{X^{s,b-1}} \quad \text{if } d \geq 3.$$  \hspace{1cm} (A.3)

It will also be useful to localize Strichartz estimates through frequency projectors: if $d \geq 3$,

$$\|P_{\epsilon,N} e^{it\Delta} f\|_{L^4_t L^4_x} \lesssim \left( \frac{N}{\epsilon} \right)^{\frac{d}{4} - \frac{1}{2}} \|P_{\epsilon,N} f\|_{L^2}$$

$$\|P_{\epsilon,N} u\|_{L^4_t L^4_x} \lesssim \left( \frac{N}{\epsilon} \right)^{\frac{d}{4} - \frac{1}{2}} \|u\|_{X^{0,b}},$$

with an additional $\kappa$ loss if $d = 2$, and identical statements if $P_{\epsilon,N}$ is replaced by $Q_{\epsilon,N}^n$. 

Duality The dual of $X_t^{s,b}$ is $X_t^{-s,-b}$. Therefore, the previous inequalities implies that, if $s' < 0$, \( \kappa > 0 \), \( b' < -\frac{1}{2} \),

\[
\|\chi(t)u\|_{X_{t}^{s',b'}} \lesssim_{\kappa,s',b'} e^{-\kappa}\|u\|_{L^{4/3}_{t}L^{4/3}_{x}} \quad \text{if } d = 2
\]
\[
\|\chi(t)u\|_{X_{t}^{-s,-b'}} \lesssim_{b'} e^{\frac{1}{2}-\frac{d}{2}}\|u\|_{L^{4/3}_{t}L^{4/3}_{x}} \quad \text{if } d \geq 3.
\]

Similarly, the dual of the inequality (A.1) is, for any \( b' < \frac{1}{2} \),

\[
\|u\|_{X_{s}^{b'}} \lesssim \|u\|_{L^{1}_{t}H^{s}_{x}}
\]

Interpolation If \( 0 \leq \theta \leq 1 \), \( s = \theta s_0 + (1-\theta)s_1 \) and \( b = \theta b_0 + (1-\theta)b_1 \),

\[
\|u\|_{X_{s}^{b}} \leq \|u\|_{X_{s_0,b_0}}^{\theta}\|u\|_{X_{s_1,b_1}}^{1-\theta}.
\]

Appendix B. Summing at a vertex

**Lemma B.1** (Degree two vertex). For any \( \epsilon \leq 1 \), \( k_0 \in \mathbb{Z}^d \) with \( |k_0| \leq \epsilon^{-1} \), \( \alpha \in \mathbb{R} \), \( \beta \geq 1 \), \( \sigma, \sigma' \in \{-1, 1\} \), and \( \kappa > 0 \),

\[
\sum_{k,k' \in \mathbb{Z}^d} \frac{1}{||k||^2 + |\sigma| |k'|^2 + |\sigma| |k_0 - k - k'|^2 - \alpha + i\beta|} \lesssim_{\kappa} \epsilon^{-2d-\kappa}.
\]

In particular,

\[
\# \{(k,k') \text{ such that } |k|, |k'| < \epsilon^{-1} \text{ and } ||k||^2 + |\sigma| |k'|^2 + |\sigma| |k_0 - k - k'|^2 - \alpha + i\beta| \leq \beta \} \lesssim_{\kappa} \beta \epsilon^{-2d-\kappa}.
\]

**Proof.** Denoting \( Q(k,k') = ||k||^2 + |\sigma| |k'|^2 + |\sigma| |k_0 - k - k'|^2 \), the above left-hand side can be bounded by

\[
\cdots \lesssim \sum_{j \in \mathbb{N}} (2^j \beta)^{-1} \# \{(k,k') \text{ such that } |k|, |k'| < \epsilon^{-2} \text{ and } Q(k,k') - \alpha + i\beta \leq 2^j \beta\}.
\]

Therefore, it suffices to show that

\[
\# \{(k,k') \text{ such that } |k|, |k'| < \epsilon^{-1} \text{ and } Q(k,k') - \alpha \leq 2^j \beta\} \lesssim_{\kappa} (2^j \beta)^{2d-2\kappa},
\]

which follows from

\[
\# \{(k,k') \text{ such that } |k|, |k'| < \epsilon^{-1} \text{ and } Q(k,k') = m\} \lesssim_{\epsilon} \epsilon^{2-2d-\kappa}
\]

(where \( m \in \mathbb{Z} \)). We distinguish two cases.

**Case 1: \( \sigma = \sigma' = 1 \).** In this case, \( Q(k,k') = 2 \left| k - \frac{k_0}{2} \right|^2 + 2 \left| k' - \frac{k_0}{2} \right|^2 \). Since \( |k_0| < \epsilon^{-1} \), we can translate the variables, and reduce matters to the case where \( Q(k,k') = |k|^2 + |k'|^2 \), in other words, \( Q \) is the sum of 2d squares. We claim that, for any \( n \), the number of solutions of \( Q(k,k') = m \) is \( \lesssim \epsilon^{2d-2\kappa} \), which gives the desired bound. To see why this claim is correct, fix \( n \) and suppose that the first \( 2d-2 \) variables, namely \( k_1, \ldots, k_{d-1}, k'_1, \ldots, k'_{d-2} \) are chosen freely, which gives \( \epsilon^{2d-2} \) possibilities. Then, there remains to choose \( k'_d \) and \( k'_d \), which have to solve an equation of the form \( (k'_d)^2 + (k'_d)^2 = m \); but this has \( \lesssim \epsilon^{-\kappa} \) solutions, as follows from the divisor bound in \( \mathbb{Z}[i] \).

**Case 2: \( \sigma = 1, \sigma' = -1 \).** In this case, \( Q(k,k') = 2(k - k_0) \cdot (k + k') + |k_0|^2 \). It suffices to show that \( Q(k,k') = m \) has \( \lesssim_{\kappa} \epsilon^{2-2d-\kappa} \) solutions, which can be proved as in Case 1: one picks \( k_1, \ldots, k_{d-1}, k'_1, \ldots, k'_{d-1}, \ldots, k'_d \),
which gives $\epsilon^{2-2d}$ possibilities. There remains to pick $k_1$ and $k_1'$ such that their product is a fixed number; by the divisor bound, this contributes a further $\epsilon^{-\kappa}$.

Cases 3 and 4: $\sigma = -1, \sigma' = 1$ and $\sigma = 1, \sigma' = -1$. In the former case, $Q(k, k') = 2(k - k_0) \cdot (k_0 - k') + |k_0|^2$, and in the latter, $Q(k, k') = -|k_0|^2 + 2(k + k') \cdot (k_0 - k')$. Either way, the proof of Case 2 applies.

\begin{proof}
Lemma B.2 (Degree one vertex). For any $\epsilon \leq 1$, $k_0 \in \mathbb{Z}^d$ with $|k_0| \leq \epsilon^{-1}$, $\alpha \in \mathbb{R}$, $\beta \geq 1$, and $\kappa > 0$,

$$\sum_{k \in \mathbb{Z}^d, |k| < \epsilon^{-1}} \frac{1}{|k|^2 + |k_0 - k|^2 - \alpha + i\beta} \lesssim_{\kappa} \epsilon^{1-d-\kappa} \left( \frac{1}{\beta} + 1 \right).$$

If moreover $k_1 \in \mathbb{Z}^d$ is such that $|k_1| \lesssim \epsilon^{-1}$:

$$\sum_{k \in \mathbb{Z}^d, |k| < \epsilon^{-1}} \frac{|1 - \delta(k_0) - \delta(k_0 + k_1 - k)|}{|k|^2 - |k_0 - k|^2 - \alpha + i\beta} \lesssim_{\kappa} \epsilon^{1-d-\kappa} \left( \frac{1}{\beta} + 1 \right).$$

In particular, for any $\sigma \in \{\pm 1\}$, assuming $k_0 \neq 0$ for $\sigma = -1$:

$$\# \{k \text{ such that } |k| < \epsilon^{-1} \text{ and } |k|^2 + \sigma|k_0 - k|^2 - \alpha + i\beta| \leq \beta \} \lesssim_{\kappa} \epsilon^{1-d-\kappa} (1 + \beta).$$

Proof. The first estimate can be dealt with as in Case 1 of the proof of Lemma B.1 and gives a bound $\epsilon^{2-d-\kappa} \lesssim \epsilon^{1-d+\kappa}$ on the right-hand side. We then turn to the second estimate to be proved, in which case $|k|^2 - |k_0 - k|^2 = k_0 \cdot (2k - k_0)$, which we will denote $Q(k)$.

Case 1: $k_0 \neq 0$ We first assume $k_0 \neq 0$. In order to obtain the desired bound, it suffices to show that, for $j \in \mathbb{N}$,

$$\# \{k \text{ such that } |k| < \epsilon^{-1} \text{ and } |Q(k) - \alpha| \leq 2^j \beta \} \lesssim_{\kappa} \epsilon^{2^j \beta + 1}.$$ But an elementary argument shows that the number of solutions of $|k_0 \cdot k - \alpha| < 2^j \beta$ is $\lesssim \left( \frac{2^j \beta}{|k_0|} + 1 \right) \epsilon^{1-d} \lesssim (2^j \beta + 1) \epsilon^{1-d}$.

Case 2: $k_0 = 0$ We now assume $k_0 = 0$. In that case, notice that the numerator forces:

$$|1 - \delta(k_0) - \delta(k_0 + k_1 - k)| = \delta(k_0 + k_1) - k_0.$$ Therefore, the sum is trivial equal to 1 as it is made of the only element $k_0 + k_1$. The desired bound holds as $1 \lesssim \epsilon^{1-d-\kappa} \left( \frac{1}{\beta} + 1 \right)$ due to the dimensional assumption $d \geq 2$.
\end{proof}

References

[1] D. Benedetto, F. Castella, R. Esposito, M. Pulvirenti, Some considerations on the derivation of the nonlinear quantum Boltzmann equation. J. Stat. Phys. 116 (2004), 381–410.

[2] D. Benedetto, F. Castella, R. Esposito, M. Pulvirenti, Some considerations on the derivation of the nonlinear quantum Boltzmann equation. II. The low density regime. J. Stat. Phys. 124 (2006), no. 2-4, 951–996.

[3] D. Benedetto, F. Castella, R. Esposito, M. Pulvirenti, From the N-body Schrödinger equation to the quantum Boltzmann equation: a term-by-term convergence result in the weak coupling regime. Comm. Math. Phys. 277 (2008), no. 1, 1–44.

[4] J. Bourgain, Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. Geom. Funct. Anal. 3 (1993), no. 2, 107–156.
[5] J. Bourgain, Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. *Comm. Math. Phys.* 176 (1996), no. 2, 421–445.
[6] J. Bourgain, C. Demeter, The proof of the $l^2$ decoupling conjecture. *Ann. of Math.* (2) 182 (2015), no. 1, 351–389.
[7] T. Buckmaster, P. Germain, Z. Hani, J. Shatah, Effective dynamics of the nonlinear Schrödinger equation on large domains. *Comm. Pure Appl. Math.* 71 (2018), no. 7, 1407–1460.
[8] T. Buckmaster, P. Germain, Z. Hani, J. Shatah, Onset of the kinetic wave turbulence description of the long-time behavior of the nonlinear Schrödinger equation. arXiv:1907.03667.
[9] Y. Deng, Two-dimensional nonlinear Schrödinger equation with random radial data. *Anal. PDE* 5 (2012), no. 5, 913–960.
[10] L. Erdös, M. Salmhofer, H.-T. Yau, Quantum diffusion of the random Schrödinger evolution in the scaling limit. II. The recollision diagrams. *Comm. Math. Phys.* 271 (2007), no. 1, 1–53.
[11] L. Erdös, M. Salmhofer, H.-T. Yau, Horng-Tzer Quantum diffusion of the random Schrödinger evolution in the scaling limit. *Acta Math.* 200 (2008), no. 2, 211–277.
[12] E. Faou, P. Germain, Z. Hani, The weakly nonlinear large-box limit of the 2D cubic nonlinear Schrödinger equation. *J. Amer. Math. Soc.* 29 (2016), no. 4, 915–982.
[13] I. Gallagher, L. Saint-Raymond, B. Texier, *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.
[14] P. Germain, A. Ionescu, and M.-B. Tran, Optimal local well-posedness theory for the kinetic wave equation. arXiv e-prints, Nov. 2017.
[15] K. Hasselmann. On the non-linear energy transfer in a gravity-wave spectrum. I. General theory. *J. Fluid Mech.*, 12:481–500, 1962.
[16] K. Hasselmann. On the non-linear energy transfer in a gravity wave spectrum. II. Conservation theorems; wave particle analogy; irreversibility. *J. Fluid Mech.* 15:273–281, 1963.
[17] O. Lanford III, Time evolution of large classical systems. *Lecture Notes in Physics* 38 (1975), 1–111, E. J. Moser ed., Springer-Verlag (1975).
[18] J. Lukkarinen, H. Spohn, Not to normal orde - notes on the kinetic limit for weakly interacting quantum fluids. *J. Stat. Phys.* 134 (2009), no. 5-6, 1133–1172.
[19] J. Lukkarinen, H. Spohn, Weakly nonlinear Schrödinger equation with random initial data. *Invent. Math.* 183 (2011), no. 1, 79–188.
[20] S. Nazarenko, *Wave turbulence*. Lecture Notes in Physics, 825. Springer, Heidelberg, 2011.
[21] A. C. Newell, B. Rumpf, Wave turbulence: a story far from over. *Advances in wave turbulence*, volume 83 of World Sci. Ser. Nonlinear Sci. Ser. A Monogr. Treatises, pages 1–151. World Sci. Publ., Hackensack, NJ, 2013.
[22] R. Peierls. Zur kinetischen theorie der Wärmeleitung in kristallen. *Annalen der Physik*, 395(8):1055–1101, 1929.
[23] H. Spohn, Derivation of the transport equation for electrons moving through random impurities. *J. Stat. Phys.* 124 (2006), no. 11, 2771–2791.
[24] H. Spohn, The phonon Boltzmann equation, properties and link to weakly anharmonic lattice dynamics. *J. Stat. Phys.* 124 (2006), no. 2-4, 1041–1104.
[25] T. Tao, *Nonlinear dispersive equations. Local and global analysis*. CBMS Regional Conference Series in Mathematics, 106. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006.
[26] L. Thomann, N. Tavetkov, Gibbs measure for the periodic derivative nonlinear Schrödinger equation. *Nonlinearity* 23 (2010), no. 11, 2771–2791.
[27] V. E. Zakharov, V. S. L’vov, G. Falkovich, *Kolmogorov spectra of turbulence I: Wave turbulence*. Springer Science & Business Media, 2012.