GLOBAL UNIFORM IN N ESTIMATES FOR SOLUTIONS OF A SYSTEM OF HARTREE–FOCK–BOGOLIUBOV TYPE IN THE GROSS–PITAVESKII REGIME.

XIAOQI HUANG

ABSTRACT. We extend the recent work of Chong et al., 2022 [10] to the critical case. More precisely, we prove global in time, uniform in N estimates for the solutions ϕ, Λ and Γ of a coupled system of Hartree–Fock–Bogoliubov type with interaction potential $\frac{1}{N}V_N(x-y) = N^2 v(N(x-y))$. We assume that the potential v is small which satisfies some technical conditions, and the initial conditions have finite energy. The main ingredient is a sharp estimate for the linear Schrödinger equation with potential in 6+1 dimension, which may be of interest in its own right.

1. **Introduction.**

Consider the N–body linear Schrödinger equation which governs the time-evolution of N boson systems

$$\Big(\frac{1}{i}\frac{\partial}{\partial_t} - \sum_{j=1}^{N} \Delta_{x_j} + \frac{1}{N} V_N(x_i - x_j)\Big)\psi_N(t,\cdot) = 0, \tag{1.1}$$

where $x_i \in \mathbb{R}^3$, N is large and $V_N(x) = N^3 v(Nx)$ for some Schwarz class potential v. The conditions on the potential will be discussed below. A physically appealing case concerns initial data forming a tensor product of the same one–particle state, in spirit of the Bose-Einstein condensation. We refer to [30] for extensive background on Bose-Einstein condensation.

The goal is to find a rigorous, simple approximation to ψ_N which is consistent with

$$\psi_{approx}(t, x_1, \dots, x_N) \sim \phi(t, x_1)\phi(t, x_2)\dots\phi(t, x_N) \tag{1.2}$$

in an appropriate sense, where ϕ is often called the mean–field limit.

In the stationary case, a survey of results concerning the ground state properties of the dilute bosonic gases can be found in [30]. In the time dependent case, in the work of Erdös, Schlein and Yau [14, 15, 16], by using the BBGKY hierarchies and the density matrix $\gamma_{N,t}$ formalism, the convergence of the exact dynamics to

2010 *Mathematics Subject Classification.* 35Q55, 81.
Key words and phrases. Hartree–Fock–Bogoliubov, Gross–Pitaevskii.

the mean–field limit is asserted in the trace norm as $N \to \infty$, provided that the the mean–field limit $\phi(t,x)$ satisfies the Gross–Pitaevskii equation

$$\frac{1}{i}\frac{\partial}{\partial_t}\phi - \Delta\phi + 8\pi a_0|\phi|^2\phi = 0. \tag{1.3}$$

Here a_0 is the scattering length of the potential v, physically, the scattering length measures the effective range of the potential V, see e.g., [16] for a precise definition of scattering length. Also in the recent work of Pickl [32], the Gross–Pitaevskii equation is derived using a different method. We refer to reader to [2, 3, 6, 13, 34] for more backgrounds on the problem of approximating the many-body Schrödinger dynamics and related results in the Gross–Pitaveskii regime.

The Fock space approach to study the problem originated in physics, with the papers by Bogoliubov [7], Lee, Huang and Yang [29] in the static case, and Wu [35] in the time dependent case. In the mathematical literature, it originates in the of work Hepp [24], Ginibre and Velo [18] and more recently by Rodnianski and Schlein [34] and Grillakis, Machedon and Margetis [22].

The Hartree-Fock-Bogoliubov type equations are derived in Fock space, which describes additional second order corrections (given by a Bogoliubov transformation) to the right hand side of the approximation (1.2). We briefly review the background of Focks space for the reader's convenience, see for instance [19] for more details and comments. The elements in $\mathcal{F}$ are vectors of the form

$$\psi = \big(\psi_0\ ,\ \psi_1(x_1)\ ,\ \psi_2(x_1,x_2)\ ,\ \dots\ \big)$$

where $\psi_0 \in \mathbb{C}$ and $\{\psi_k\}_{k\in\mathbb{N}}$ are symmetric complex-valued L^2 functions. The vacuum state is the vector defined by

$$\Omega := (1,0,0\dots)$$

which models a state with no particles, . The symmetric Fock space $\mathcal{F}$ has a norm induced by the inner product

$$\langle\varphi,\psi\rangle = \overline{\varphi_0}\,\psi_0 + \sum_{n=1}^{\infty}\int \overline{\varphi_n}\,\psi_n \tag{1.4}$$

The creation and annihilation distribution-valued operators at $x \in \mathbb{R}^3$, denoted by a_x^* and a_x, are defined by actions on vectors of the form $(0,\cdots,\psi_{n-1},0,\cdots)$ and $(0,\cdots,\psi_{n+1},0,\cdots)$ respectively as follow

$$\begin{aligned}
a_x^*(\psi_{n-1}) &:= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\delta(x-x_j)\psi_{n-1}(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n),\\
a_x(\psi_{n+1}) &:= \sqrt{n+1}\psi_{n+1}([x],x_1,\dots,x_n),
\end{aligned}$$

with $[x]$ indicating that the variable x is frozen.

For every $\phi \in L^2(\mathbb{R}^3)$, the creation and annilihation operators associated to ϕ are defined by

$$a(\overline{\phi}) := \int dx \left\{\overline{\phi}(x)\, a_x\right\} \quad \text{and} \quad a^*(\phi) := \int dx \left\{\phi(x)\, a_x^*\right\}$$

where by convention we associate a with $\overline{\phi}$ and a^* with ϕ. Let us also define the skew-Hermitian operator

$$\mathcal{A}(\phi) := \int dx \left\{\overline{\phi}(x) a_x - \phi(x) a_x^*\right\} \tag{1.5}$$

which is the Weyl operator used by Rodnianski and Schlein in [34]. The corresponding coherent state is

$$\psi(\phi) := e^{-\sqrt{N}\mathcal{A}(\phi)}\Omega \ . \tag{1.6}$$

Here ϕ is called the condensate wave function. It is not hard to check that

$$\psi(\phi) := e^{-\sqrt{N}\mathcal{A}(\phi)}\Omega = \left(\cdots c_n \prod_{j=1}^{n} \phi(x_j) \cdots\right) \text{ with } c_n = \left(e^{-N\|\phi\|_{L^2}^2} N^n/n!\right)^{\frac{1}{2}}$$

Also consider the following skew-Hermitian quadratic operator

$$\mathcal{B}(k(t)) := \frac{1}{2}\int dxdy \left\{\overline{k}(t,x,y) a_x a_y - k(t,x,y) a_x^* a_y^*\right\} \ . \tag{1.7}$$

Here k is a symmetric wave function, i.e. $k(t,x,y) = k(t,y,x)$, called the pair exicitation function. This particular construction and the corresponding unitary operator

$$\mathcal{M}(t) := e^{-\sqrt{N}\mathcal{A}(\phi(t))} e^{-\mathcal{B}(k(t))} = e^{-\sqrt{N}\mathcal{A}(t)} e^{-\mathcal{B}(t)}$$

were introduced in Grillakis, Machedon and Margetis [22]. The construction is in the spirit of Bogoliubov theory in physics, and the Segal–Shale–Weil representation in mathematics.

Consider the Fock Hamiltonian

$$\mathcal{H} = \int dxdy \left\{a_x^*\, \Delta_x \delta(x-y)\, a_y\right\} - \frac{1}{2N}\int dxdy \left\{v_N(x-y) a_x^* a_y^* a_y a_x\right\}, \tag{1.8}$$

where

$$v_N(x) = N^{3\beta} v\big(N^\beta x\big).$$

It can readily be checked that $\mathcal{H}$ is a diagonal operator on Fock space and it acts as a differential operator in n variables

$$H_{n,\,\mathrm{PDE}} = \sum_{j=1}^{n} \Delta_{x_j} - \frac{1}{N}\sum_{i<j}^{n} N^{3\beta} v\big(N^\beta (x_j - x_k)\big)$$

on the nth sector of $\mathcal{F}$.

The goal is to study the evolution of coherent initial conditions of the form

$$\psi_{\text{exact}}(t) = e^{it\mathcal{H}} e^{-\sqrt{N}\mathcal{A}(\phi_0)} e^{-\mathcal{B}(k_0)}\Omega \tag{1.9}$$

In the papers [20, 23, 22], Grillakis, Machedon and Margetis proposed an approximation of the form

$$\psi_{\text{appr}}(t) := e^{-\sqrt{N}\mathcal{A}(\phi(t))} e^{-\mathcal{B}(k(t))}\Omega \tag{1.10}$$

The strategy to study the approximation is to first evolve forward in time by the exact dynamics and then backward by the approximation dynamics. That is, to consider the reduced dynamics

$$\psi_{\text{red}}(t) = e^{\mathcal{B}(t)} e^{\sqrt{N}\mathcal{A}(t)} e^{it\mathcal{H}} e^{-\sqrt{N}\mathcal{A}(0)} e^{-\mathcal{B}(0)}\Omega \tag{1.11}$$

and compute the "reduced Hamiltonian"

$$\mathcal{H}_{\text{red}} = \frac{1}{i}\big(\partial_t \mathcal{M}^*\big)\mathcal{M} + \mathcal{M}^*\mathcal{H}\mathcal{M} \tag{1.12}$$

so that

$$\frac{1}{i}\partial_t \psi_{\text{red}} = \mathcal{H}_{\text{red}}\psi_{\text{red}} \ . \tag{1.13}$$

In [22], Grillakis, Machedon and Margetis imposed some Schrödinger type equations on ϕ (the Hartree equation), k, and proved the following results for $0 < \beta < 1/3$ by using an energy estimate based approach in Fock space and decay properties of ϕ

$$\left\| \psi_{\text{exact}}(t) - e^{iN\chi(t)}\psi_{\text{appr}}(t) \right\|_{\mathcal{F}} \le P(t) N^{\frac{3\beta-1}{2}}. \tag{1.14}$$

Here χ is some real phase function, and $P(t)$ is of polynomial growth in time.

This result was extended to $\beta < 1/2$ in Kuz [28], where the author also argued that the equations used in [22] can not provide an approximation for $\beta > 1/2$. The Hartree-Fock-Bogoliubov equations was later introduced in [19], in the hope of obtaining an approximation for higher values of β. There are several equivalent ways of writing these equations. Broadly speaking, the equations ensure that after Wick reordering, the reduced Hamiltonian has neither a or a^* linear terms, nor aa or a^*a^* quadratic terms.

The Hartree-Fock-Bogoliubov equations were also introduced independently in a different context in [1], and they were studied in [4, 10, 21]. In particular, local in time, uniform in N estimates for solutions to the HFB system were obtained in [21], and they were used in [12] to give a Fock space approximation of the form

$$\left\| \psi_{\text{exact}}(t) - e^{iN\chi(t)}\psi_{\text{appr}}(t) \right\|_{\mathcal{F}} \le C e^{P(t)} N^{\frac{\beta-1}{2}}, \tag{1.15}$$

for a polynomial $P(t)$ and $0 < \beta < 1$. The global in time estimate for HFS system in [10], as well as the main results in this paper, is in the hope of improving the $e^{P(t)}$ in (1.15) to some polynomial $P(t)$, and possibly extending to the case $\beta = 1$, we wish to address this problem in a future work [25].

We also mention that in Benedikter, de Oliveira and Schlein [2], a similar approach is considered, where the authors impose the Gross–Pitaevskii equation equation for ϕ and define k by an explicit formula, and give a rate of convergence result in terms of the marginal density.

In the remaining of this paper, we shall focus on the analysis of the systems of PDEs. The functions described by these equations are: the condensate $\phi(t,x)$ and the density matrices

$$\Gamma(t,x_1,x_2) = \frac{1}{N}\left(\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)\right)(t,x_1,x_2) + \bar{\phi}(t,x_1)\phi(t,x_2) \tag{1.16}$$

$$\Lambda(t,x_1,x_2) = \frac{1}{2N}\mathrm{sh}(2k)(t,x_1,x_2) + \phi(t,x_1)\phi(t,x_2) \tag{1.17}$$

where

$$\begin{aligned} \mathrm{sh}(k) &= k + \frac{1}{3!}k \circ \bar{k} \circ k + \dots \\ \mathrm{ch}(k) &= \delta(x-y) + \frac{1}{2!}\bar{k} \circ k + \dots \end{aligned} \tag{1.18}$$

Here $(u \circ v)(x,y) = \int u(x,z)v(z,y)dz$, the pair excitation function k is an auxiliary function, which does not explicitly appear in the system.

There are several equivalent ways of expressing the equations, in this section we shall use a compact, matrix formulation as in [10]. We separate the condensate part from the pair interaction part: define $\Gamma_c = \bar{\phi} \otimes \phi$, $\Lambda_c = \phi \otimes \phi$, $\Gamma_p = \frac{1}{N}\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)$ and $\Lambda_p = \frac{1}{2N}\mathrm{sh}(2k)$. Also denote $\rho(t,x) = \Gamma(t,x,x)$

To write the Hartree-Fock-Bogoliubov equations in matrix notation, denote

$$V_N(x-y) = N^3 v(N(x-y))$$

for some Schwarz class potential which will be discussed below. Define

$$\Omega = \begin{pmatrix} -\Gamma & -\bar{\Lambda} \\ \Lambda & \Gamma \end{pmatrix} = \Psi + \Phi$$

where

$$\Psi = \begin{pmatrix} -\Gamma_p & -\bar{\Lambda}_p \\ \Lambda_p & \Gamma_p \end{pmatrix}$$

$$\Phi = \begin{pmatrix} -\Gamma_c & -\bar{\Lambda}_c \\ \Lambda_c & \Gamma_c \end{pmatrix}$$

Finally, let

$$S_3 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$$

where I is the identity operator.

The evolution equations for Ψ and Φ are

$$\begin{aligned}
&\frac{1}{i}\partial_t \Phi - [\Delta_x \delta(x-y) S_3, \Phi] \\
&\quad = -[\big(V_N * \rho(t,x)\big)\delta(x-y)S_3, \Phi] - [V_N \Psi^*, \Phi]
\end{aligned} \tag{1.19}$$

$$\begin{aligned}
&\frac{1}{i}\partial_t \Psi - [\Delta_x \delta(x-y) S_3, \Psi] \\
&\quad = -[\big(V_N * \rho(t,x)\big)\delta(x-y)S_3, \Psi] - \frac{1}{2N}[S_3, V_N \Psi] \\
&\qquad - [V_N \Omega^*, \Psi] - \frac{1}{2N}[S_3, V_N \Phi]
\end{aligned} \tag{1.20}$$

In addition, the condensate ϕ satisfies

$$\begin{aligned}
&\left\{\frac{1}{i}\partial_t - \Delta_{x_1}\right\}\phi(x_1) \\
&\quad = -\int dy \{V_N(x_1 - y)\Gamma(y,y)\}\phi(x_1) \\
&\qquad - \int dy \{V_N(x_1 - y)\Gamma_p(y, x_1)\}\phi(y) \\
&\qquad + \int dy \{V_N(x_1 - y)\Lambda_p(x_1, y)\}\bar{\phi}(y)
\end{aligned} \tag{1.21}$$

Here $A^*(x,y) = \bar{A}(y,x), [A,B] = A \circ B - B \circ A$ and V_N acts as pointwise multiplication by $V_N(x-y)$. See (5.1)-(5.4) for a scalar form of the above equations.

The solutions ϕ, Λ_p, Λ_c, Γ_p and Γ_c all depend on N. This has been suppressed to simplify the notation. However, we will always keep track of dependence on N in our estimates.

Next we review the conserved quantities of these equations, see [19] for more details. The first conserved quantities is the total number of particles (normalized by division by N):

$$\mathrm{tr}\{\Gamma(t)\} = \|\phi(t,\cdot)\|^2_{L^2(dx)} + \frac{1}{N}\|\mathrm{sh}(k)(t,\cdot,\cdot)\|^2_{L^2(dxdy)} = 1 \tag{1.22}$$

From here we see that

$$\|\Lambda(t,\cdot,\cdot)\|_{L^2(dxdy)} \leq C \tag{1.23}$$

The second conserved quantity is the energy per particle

$$
\begin{aligned}
E(t) =&\operatorname{tr}\{\nabla_{x_1}\cdot\nabla_{x_2}\Gamma(t)\}\\
&+\frac{1}{2}\int dx_1dx_2\Big\{V_N(x_1-x_2)|\Lambda(t,x_1,x_2)|^2\Big\}\\
&+\frac{1}{2}\int dx_1dx_2\Big\{V_N(x_1-x_2)|\Gamma(t,x_1,x_2)|^2\Big\}\\
&+\frac{1}{2}\int dx_1dx_2\Big\{V_N(x_1-x_2)\Gamma(t,x_1,x_1)\Gamma(t,x_2,x_2)\Big\}\\
&-\int dx_1dx_2\Big\{V_N(x_1-x_2)|\phi(t,x_1)|^2|\phi(t,x_2)|^2\Big\}
\end{aligned} \tag{1.24}
$$

We shall assume that

$$
v \text{ is Schwarz } \sup_{1\le p\le\infty}\|v\|_{L^p}<\varepsilon,\ \textit{and}\ \operatorname{supp}\hat{v}\subset B_1(0). \tag{1.25}
$$

In addition, we also assume

$$
\begin{aligned}
&v \text{ is spherically symmetric and}\\
&v\ge 0, \frac{\partial v}{\partial r}(r)\le 0.
\end{aligned} \tag{1.26}
$$

Here $\hat{v}$ denotes the Fourier transform and $B_1(0)$ denotes the unit ball in $\mathbb{R}^3$, and ε is a fixed small constant to be specified later (see (5.32)) which is independent of N. The condition (1.26) allows us to use a priori estimates for Γ (the interaction Morawetz estimate, Lemma 6.2 in [10]), which is part of our main strategies in treating the HFB system, i.e., We regard the equation for Λ as a linear equation with non-local "coefficients" given by Γ and a forcing term involving ϕ. For Γ and ϕ, we will only use a priori estimates, given by conserved quantities and the interaction Morawetz estimate. It is still open to us if we can analysis the Γ directly without using the a priori estimates, the main difficulties come from the fact that the linear part $-\Delta_x+\Delta_y$ of the Γ equation that is anti-symmetric in x, y, see (1.39).

The smallness assumption on the potential is due to the perturbation based arguments we used in analysing the linear and nonlinear Schrödinger equations. The smallness and spherically symmetric assumptions on the interacting potential were also used in earlier work of Erdös, Schlein and Yau [16] and Boccato, Brennecke, Cenatiempo and Schlein [5] in the Gross–Pitaevskii regime, and Grillakis, Machedon and Margetis [22] in the mean field regime.

Despite the smallness assumption, the case $\beta=1$ is strictly harder than $\beta<1$, which allows large potentials if the parameter N is sufficiently large. For instance, in the case $\beta<1$, one has additional regularity in using Sobolev estimate, while in the case $\beta=1$, the criticalness of the scaling forces us to use and also develop new sharp estimates in the arguments, see the end of this section for more discussions on the difficulties in the critical setting.

The assumption $\operatorname{supp}\hat{v} \subset B_1(0)$ is only used in section 4, it is essentially not required for our proof, but it simplify the arguments greatly, see the beginning of section 4 for a brief discussion on how to remove this assumption.

For the initial conditions, we assume that

$$\begin{aligned} \operatorname{tr}\{\Gamma(0)\} &\leq C_0 \\ E(0) &\leq C_0 \end{aligned} \tag{1.27}$$

The size of the initial condition C_0 is important, as we shall see in the end of section 5, our assumption on the size of the potential v depends on the size of initial data. Thus, we keep track of the constant C_0 from now on.

Note that the kinetic energy is

$$\begin{aligned} \operatorname{tr}\{\nabla_{x_1} \cdot \nabla_{x_2} \Gamma(t)\} &= \int dx \Big\{ |\nabla_x \phi(t,x)|^2 \Big\} \\ &+ \frac{1}{2N} \int dx_1 dx_2 \Big\{ |\nabla_{x_1} \operatorname{sh}(k)(t,x_1,x_2)|^2 + |\nabla_{x_2} \operatorname{sh}(k)(t,x_1,x_2)|^2 \Big\} \end{aligned} \tag{1.28}$$

If we assume $E \leq C_0$, then we have an H^1 estimate for Λ, uniformly in time (and N):

$$\begin{aligned} &\int dx_1 dx_2 \Big\{ |\nabla_{x_1} \Lambda(t,x_1,x_2)|^2 + |\nabla_{x_2} \Lambda(t,x_1,x_2)|^2 \Big\} \leq CC_0 \\ &\frac{1}{N} \int dx_1 dx_2 \Big\{ |\nabla_{x_1,x_2} \operatorname{sh}(2k)(t,x_1,x_2)|^2 \Big\} \leq CC_0 \end{aligned} \tag{1.29}$$

Also, Γ satisfies the H^2 type estimate

$$\| |\nabla_{x_1}| |\nabla_{x_2}| \Gamma(t) \|_{L^2(dxdy)} \leq CC_0 \tag{1.30}$$

See [20], [19], as well as [1] for these conserved quantities.

By Plancherel theorem, we see that (1.29) implies that for all time t

$$\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_p(t) \|_{L^2(dxdy)} \leq CC_0. \tag{1.31}$$

as well as

$$\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_c(t) \|_{L^2(dxdy)} \leq CC_0. \tag{1.32}$$

Here, and also later on, $\langle \nabla_x \rangle^{\frac{1}{2}}$ means $(1-\Delta_x)^{\frac{1}{4}}$, which is a Fourier multiplier with symbol $(1+|\xi|^2)^{\frac{1}{4}}$, and similarly for $\langle \nabla_y \rangle^{\frac{1}{2}}$.

Similarly, by (1.28) and Plancherel, we also have

$$\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Gamma_p(t) \|_{L^2(dxdy)} \leq CC_0. \tag{1.33}$$

as well as

$$\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Gamma_c(t) \|_{L^2(dxdy)} \leq CC_0. \tag{1.34}$$

In order to state the main result for this paper in the simplest possible form, we define the following partial Strichartz norms:

$$\begin{aligned} &\|\Lambda\|_{\mathcal{S}_{x,y}} \\ &= \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dx)L^2(dy)} \\ &\quad + \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dy)L^2(dx)} \end{aligned} \tag{1.35}$$

Recall (p,q) are admissible in $3+1$ dimensions if $\frac{2}{p}+\frac{3}{q}=\frac{3}{2}$, $2 \le p \le \infty$.

The main result of this paper is

Theorem 1.1. *Let* $\Lambda = \Lambda_p + \Lambda_c$, $\Gamma = \Gamma_p + \Gamma_c$ *be solutions of* (1.19), (1.20), *while the potential satisfies* (1.25) *and* (1.26), *and the initial conditions satisfy* (1.27). *Then we have*

$$\|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda\|_{\mathcal{S}_{x,y}} + \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Gamma\|_{\mathcal{S}_{x,y}} \le C \tag{1.36}$$

for some constant C *independent of* N. *The above estimate still hold if we replace* Λ *by* Λ_c *or* Λ_p, *or replace* Γ *by* Γ_c *or* Γ_p.

We also have a theorem for $\mathrm{sh}(2k)$ (without dividing by N).

Theorem 1.2. *Let* Λ, Γ, ϕ *be solutions of* (1.19), (1.20), *while the potential satisfies* (1.25) *and* (1.26), *and the initial conditions satisfy* (1.27). *Assume also that*

$$\|\mathrm{sh}(2k)(0,\cdot,\cdot)\|_{L^2} + \|\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)(0,\cdot,\cdot)\|_{L^2} \le C$$

Then we have

$$\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}} + \|\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)\|_{\mathcal{S}_{x,y}} \le C \tag{1.37}$$

Remark 1.3. Here (1.37) improves the results in [10, Theorem 1.3] in two ways. First, the potential $N^2 v(N(x-y))$ we considered here represents a stronger interaction between particles, compared with the $N^{3\beta-1}v(N^\beta(x-y))$, $\beta<1$ type potentials considered in [10]. Second, the $\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}}$ norm stays bounded uniformly in N, compared with the $\log N$ growth in [10]. Although our argument is written for the case where the potential is $N^2 v(N(x-y))$, it also works for the $N^{3\beta-1}v(N^\beta(x-y))$ case. For example, it can be shown that the uniform in N estimates in (1.37) still hold for the case where the potential is equal to $N^{3\beta-1}v(N^\beta(x-y))$, $\beta<1$.

The above estimates also imply some estimates for $\mathrm{sh}(k)$. In particular,

$$\|\mathrm{sh}(k)\|_{L^p(dx)L^2(dy)} \le C\|\mathrm{sh}(2k)\|_{L^p(dx)L^2(dy)}$$

This is because $\mathrm{sh}(k) = \frac{1}{2}\mathrm{sh}(2k) \circ \mathrm{ch}(k)^{-1}$ and $\mathrm{ch}(k)^{-1}$ has bounded operator norm.

Finally, we also have estimates for ϕ. Define the standard Strichartz spaces

$$\|\phi\|_{\mathcal{S}} = \sup_{(p,q)\ admissible} \|\phi\|_{L^p(dt)L^q(dx)}.$$

Corollary 1.4. *Under the assumptions of Theorem 1.1, and the additional assumption* $\|\langle\nabla\rangle^{\frac{1}{2}}\phi(0,\cdot)\|_{L^2} \le C$ *, we have*

$$\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{\mathcal{S}} \le C. \tag{1.38}$$

We shall now brief describe the difficulties in the critical case $\beta = 1$ along with the strategies employed in addressing them in our arguments. Denote

$$\mathbf{S} = \frac{1}{i}\frac{\partial}{\partial t} - \Delta_x - \Delta_y$$

$$\mathbf{S}_{\pm} = \frac{1}{i}\frac{\partial}{\partial t} - \Delta_x + \Delta_y$$

Schematically, if we treat V_N as the δ potential and ignore the constants, the equations become

(1.39)

$$\begin{aligned}
&\mathbf{S}\Lambda_c = \Gamma(t,x,x)\Lambda_c(t,x,y) + \Lambda_p(t,x,x)\Gamma_c(t,x,y)\\
&\mathbf{S}_{\pm}\Gamma_c = \Gamma(t,x,x)\Gamma_c(t,x,y) + \bar{\Lambda}_p(t,x,x)\Lambda_c(t,x,y)\\
&\mathbf{S}\Lambda_p + \frac{V_N}{N}\Lambda_p = \Gamma(t,x,x)\Lambda_p(t,x,y) + \Lambda_p(t,x,x)\Gamma_p(t,x,y) - \frac{V_N}{N}\Lambda_c\\
&\quad + \Lambda_c(t,x,x)\Gamma_p(t,x,y)\\
&\mathbf{S}_{\pm}\Gamma_p = \Gamma(t,x,x)\Gamma_p(t,x,y) + \Lambda_p(t,x,x)\Lambda_p(t,x,y) + \bar{\Lambda}_c(t,x,x)\Lambda_p(t,x,y)\\
&\quad\left(\frac{1}{i}\partial_t - \Delta_{x_1}\right)\phi(x_1) = -\Gamma(x_1,x_1)\phi(x_1) - \Gamma_p(x_1,x_1)\phi(x_1) + \Lambda(x_1,x_1)\bar{\phi}(x_1)
\end{aligned}$$

Since $\Gamma_c = \bar{\phi}\otimes\phi$, $\Lambda_c = \phi\otimes\phi$, one can view that ϕ satisfies the cubic nonlinear Schrödinger equations with additional corrections.

Recall that $V_N(x) = N^3 v(Nx)$, thus $\frac{V_N}{N}$ in the Λ_p equation satisfies the critical scaling in the sense that $\frac{V_N}{N} \in L^{3/2}$ uniformly in $N \ge 1$. For Schrödinger operators $-\Delta_{\mathbb{R}^n} + V$, it is known that $V \in L^{n/2}_{\text{loc}}$ for dimension $n \ge 3$ is almost the minimal condition to ensure the Schrödinger operators $-\Delta + V$ is bounded from below and self-adjoint. There is a vast amount of literature in the study of Schrödinger operators $-\Delta + V$ with critically singular potentials from different aspects, e.g., Strichartz estimates, unique continuation and dispersive estimates, see [8, 26, 27, 33, 36].

It is known in the study of cubic nonlinear Schrödinger equations that $1/2$ derivative is the minimal regularity required on the initial data in order to have local or global well-poseness results (see [9]). Similarly, our arguments in treating the nonlinear terms require the Strichartz estimates for $\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Lambda_{p \text{ or } c}$ and $\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Gamma_{p \text{ or } c}$ with $\alpha \ge \frac{1}{2}$. In the case $\beta = 1$, the $\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}$ derivative is the threshold for the linear equation in the sense that if we apply $\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}$ to $\frac{V_N}{N}\Lambda_p$ or $\frac{V_N}{N}\Lambda_c$ in the Λ_p equation, we get a singularity which is essentially $N^{2\alpha-1}V_N\Lambda$. And since $\|N^{2\alpha-1}V_N\|_{L^p} \to \infty$ as $N \to \infty$ for any $1 \le p \le \infty$ if $\alpha > 1/2$, it would be hard to obtain any uniform in N estimates in this setting.

For the same reason, in order to get Strichartz type estimate for $\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_p$, we need to handle an inhomogeneous forcing term like $N^3v(N(x-y))\Lambda_p(x,y)$, if $\beta = 1$ and $N^{4\beta-1}v(N^\beta(x-y))\Lambda_p(x,y)$ if $\beta < 1$. As $N \to \infty$, $N^3v(N(x-y)) \to \delta$, the delta function up to some constant. Thus, to use a perturbative argument, we have to develop Strichartz estimate involving the L^1 norm in the $x-y$ direction, and a collapsing estimate which involves the L^∞ norm in the $x-y$ direction, see Theorem 3.2 and Lemma 3.8–3.9 respectively. The collapsing estimate (see (2.8) for the definition of the collapsing norm) is a natural generalization of the Morawetz inequalities in the study of nonlinear Schrödinger equations, see [31].

Theorem 3.2 generalizes Proposition 4.7 [10] by removing the frequency assumption there. The main difficulty in the proof is the lack of square function estimate, or equivalently sharp Sobolev estimate in the L^1 norm setting. The main idea is to use Littlewood–Paley estimate in the $x-y$ direction on the left hand side, and add up different frequency pieces in the $x+y$ or t directions on the right side. To do this, we need to keep track of the size of frequency variables, including their ratios and make explicit use of the magnitude of $\langle\nabla_{x+y}\rangle^{\frac{1}{2}}$ and $|\partial_t|^{\frac{1}{4}}$ derivatives. The proof of Lemma 3.8–3.9 is based similar ideas along with the use of Christ–Kiselev lemma, and they are essentially a dual version of the Strichartz estimate in Theorem 3.2.

Another main difficulty lies in the proof of Theorem 2.1, i.e., the proof of Strichartz type estimates for linear Schrödinger equations with interacting potentials. We need to choose the norms for the perturbation arguments properly under various frequency support assumptions, see section 4 case 1–case 3. We divide the cases based on the frequency support in the $x+y$ direction, since multiplication by an interacting potential may enlarge the frequency support in the $x-y$ direction after each iteration. The norms within each case may depend on each other. Due to the failure of sharp Sobolev estimates at L^∞ and the shift in frequency support after multiplications by the potential, it is hard to bound these norms independently, without the frequency support assumption or the use of other norms, for instance, the full collapsing norm and the full Strichartz norm including the endpoint pair (p,q) in the $x-y$ direction. Similar difficulties also arise in the analysis of full nonlinear equations, where we need to define the norms appropriately under various frequency support assumptions in order to close the bootstrap argument. In the case of $\beta < 1$, however, this can be remedied by the allowance of additional derivatives as discussed above.

The proof of Theorem 2.1 for different cases also relies on a collapsing estimate for the linear equations with interacting potentials under low frequency assumptions, which is Theorem 4.1. The low frequency assumptions allow us to use Bernstein type inequality at L^∞ instead of classical Sobolev estimate, which requires additional regularity. However, as discussed above, after each iteration step, the frequency support may expand due to multiplication by the potential, the main idea in the proof of Theorem 4.1 is to exploit the gain of small constant coming from the smallness assumption on the potential at each iteration step and use crude estimates after finitely many steps of iterations. Also, we prove a full collapsing norm estimates in the last section assuming the forcing term $H = 0$ in (2.11), which is based on Theorem 2.1 and the analysis of the wave operator W for the Schrödinger operators $-\Delta + V$ as in the work of Yajima [36].

The structure of the rest of the paper is the following. In section 2, we list the notations used in this paper and state our main estimate, Theorem 2.1 for the linear Schrödinger equation in $6+1$ dimension. In sections 3 and 4, we prove Theorem 2.1. In section 5, we prove Theorem 1.1 using the linear estimate Theorem 2.1. In section 6 and section 7, we use prove Theorem 1.2 and Corollary 1.4. In the last section, we use Theorem 2.1 to prove a "collapsing estimate" for the linear equation involving the interaction potential, which may be of interest on its own right.

Acknowledgements. The author is indebted to Manoussos Grillakis and Matei Machedon for both suggesting the problem and for many discussions and comments on various stages of this work, their roles on this work is no less than the author's one. The author would also like to thank Jacky Chong and Zehua Zhao for comments on an earlier version of this paper. The author is supported in part by an AMS–Simons Travel Grant.

2. **List of notations and statement of the main linear estimates.**

Let us define the partial Strichartz norms

$$\begin{aligned}\|\Lambda\|_{\mathcal{S}_{x,y}} &= \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dx)L^2(dy)} \\ &+ \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dy)L^2(dx)}.\end{aligned} \tag{2.1}$$

where the pair (p,q) is admissible in $3+1$ dimension if $\frac{2}{p}+\frac{3}{q}=\frac{3}{2}$, $2 \le p \le \infty$.

Define the full Strichartz norm

$$\begin{aligned}\|\Lambda\|_{\mathcal{S}} &= \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dx)L^2(dy)} \\ &+ \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(dy)L^2(dx)} \\ &+ \sup_{(p,q)\ admissible} \|\Lambda\|_{L^p(dt)L^q(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{2.2}$$

And define the restricted dual Strichartz norm, excluding the end-points $p' = 2$, $p' = 1$: let p_1 large and $p_0 > 2$ but close to 2, for admissible pairs (p,q), define

$$\|G\|_{\mathcal{S}'_r} = \inf_{p_1 \ge p \ge p_0} \{\|G\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}, \|G\|_{L^{p'}(dt)L^{q'}(dy)L^2(dx)}\}. \tag{2.3}$$

Let us also recall the standard Littlewood-Paley decomposition. Let $\phi(x)$ such that $\hat{\phi} \in C_0^\infty$ and $\hat{\phi}(\xi) = 1$ in $|\xi| < 1$, $\hat{\phi}(\xi) = 0$ in $|\xi| > 2$. Define the ϕ_k for $k \geq 0$ by $\hat{\phi}_k(\xi) = \hat{\phi}(\frac{\xi}{2^k})$ and denote

$$P_{|\xi|<2^k} f = f * \phi_k \tag{2.4}$$

so that the inverse Fourier transform of $\hat{\phi}(\frac{\xi}{2^k})\hat{f}$ is $P_{|\xi|<2^k} f$.

Next let $\psi_0 = \phi$ and define ψ_k for $k \geq 1$ by $\hat{\psi}_k(\xi) = \hat{\phi}(\frac{\xi}{2^k}) - \hat{\phi}(\frac{\xi}{2^{k-1}})$. We also denote

$$P_{|\xi|\sim 2^k} f = f * \psi_k \tag{2.5}$$

For later use, we shall also abuse our notation a bit and define, for an arbitrary positive constant M,

$$P_{|\xi|<M} f = f * \phi(Mx), \tag{2.6}$$

for any fixed constant M. And define,

$$P_{|\xi|\geq M} f = f - P_{|\xi|<M} f. \tag{2.7}$$

So for any two fixed constants $0 < M < N$,

$$P_{N\leq|\xi|<M} f = P_{|\xi|<M} f - P_{|\xi|<N} f.$$

Now we can define the following two "collapsing norms". Let

$$\|\Lambda\|_{collapsing} = \left\|\Lambda\right\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}. \tag{2.8}$$

And define

$$\begin{aligned} \|\Lambda\|_{low\ collapsing} =& \left\|P_{|\xi-\eta|<20N}\Lambda\right\|_{collapsing} + \left\|P_{|\xi|<20N}\Lambda\right\|_{collapsing} \\ &+ \left\|P_{|\eta|<20N}\Lambda\right\|_{collapsing}. \end{aligned} \tag{2.9}$$

If $A \lesssim B$, there is a constant C such that $A \leq CB$, and we use $A \sim B$ to denote the case when $A \lesssim B$ and $B \lesssim A$.

Define $\langle\nabla_x\rangle^{\frac{1}{2}} f = (1 - \Delta_x)^{\frac{1}{4}} f$ such that the Fourier transform of $\langle\nabla_x\rangle^{\frac{1}{2}} f$ is $(1+|\xi|^2)^{\frac{1}{4}}\hat{f}$, and similarly the Fourier transform of $\langle\nabla_y\rangle^{\frac{1}{2}} f$ is $(1+|\eta|^2)^{\frac{1}{4}}\hat{f}$ for any $f \in L^2(\mathbb{R}^6)$.

Let $x, y \in \mathbb{R}^3$, define

$$\begin{aligned} \mathbf{S} &= \frac{1}{i}\frac{\partial}{dt} - \Delta_x - \Delta_y \\ \mathbf{S}_\pm &= \frac{1}{i}\frac{\partial}{\partial t} - \Delta_x + \Delta_y \end{aligned} \tag{2.10}$$

Consider the equation

$$\begin{aligned}\mathbf{S}\Lambda(t,x,y) &= N^2 v\big(N(x-y)\big)\Lambda(t,x,y) + G(t,x,y)\\ &\qquad + N^2 v\big(N(x-y)\big)H(t,x,y)\\ \Lambda(0,\cdot) &= \Lambda_0\end{aligned} \tag{2.11}$$

The simplest form of theorem is

Theorem 2.1. *Let Λ satisfy* (2.11)*, and assume v satisfy* (1.25)*, we have*

$$\begin{aligned}&\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{\mathcal{S}_{x,y}} + \|\langle\nabla_{x+y}\rangle^{\frac12}\Lambda\|_{low\ collapsing} + \||\partial_t|^{\frac14}\Lambda\|_{low\ collapsing}\\ &\lesssim \|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}G\|_{\mathcal{S}'_r} + \varepsilon\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}H\|_{L^2(dt)L^6(x-y)L^2(d(x+y))}\\ &\quad + \varepsilon\|\langle\nabla_{x+y}\rangle^{\frac12}H\|_{collapsing} + \varepsilon\||\partial_t|^{\frac14}H\|_{collapsing}\\ &\quad + \varepsilon\|\langle\nabla_x\rangle^{\frac12}H\|_{collapsing} + \varepsilon\|\langle\nabla_y\rangle^{\frac12}H\|_{collapsing} + \|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda_0\|_{L^2}.\end{aligned} \tag{2.12}$$

Remark 2.2. The main difficulty in proving the theorem is the presence of the term $N^2v\big(N(x-y)\big)\Lambda(t,x,y)$, where $N^2v\big(N(x-y) \in L^{3/2}$ satisfies the critical scaling. The term $N^2v\big(N(x-y)\big)H(t,x,y)$ in (2.11) is a technical term which arises from the term $N^2v\big(N(x-y)\big)\Lambda(t,x,y)$, since in our application, we split $\Lambda=\Lambda_p+\Lambda_c$ and take $\Lambda=\Lambda_p$ and $H=\Lambda_c$ in (2.11). And the presence of $H(t,x,y)$ does not lead to any essential difficulty in the proof of the above theorem.

We also remark that in the case $H=0$, by using the above theorem plus an abstract argument, one can replace the $\|\cdot\|_{low\ collapsing}$ norm on the left side of (2.12) by the $\|\cdot\|_{collapsing}$ norm and the same result still holds. See the discussion in section 8 for more details.

All the implicit constants in $\lesssim$ are independent of N and ε, and the choice of the small constant ε in (1.25) will depend on the implicit constants and C_0 in (1.27).

3. **Preliminary estimates for solutions to the linear Schrödinger equation.**

We will use the following Strichartz estimate. In 6+1 dimensions,

Theorem 3.1 (Theorem 2.4, 2.5 of [11])**.** *Let $\mathbf{S}u = f+g$, $u(0,\cdot)=u_0$. Then*

$$\|u\|_{\mathcal{S}} \lesssim \|f\|_{L^2(dt)L^{\frac65}(x-y)L^2(d(x+y))} + \|g\|_{\mathcal{S}'_r} + \|u_0\|_{L^2}.$$

Now we shall present the main theorem of this section.

Theorem 3.2. *Let $\mathbf{S}u=f$ $u(0,\cdot)=0$ Then*

$$\|u\|_{\mathcal{S}_{x,y}} \lesssim \big\|\langle\nabla_{x+y}\rangle^{\frac12}f\big\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} + \big\||\partial_t|^{\frac14}f\big\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}.$$

Remark 3.3. A frequency localized version of the above Theorem appears in [10, Proposition 4.7], , which is also a motivation of the above theorem. The $\mathcal{S}_{x,y}$ norm is crucial here, we do not expect the above estimate to be true if we replace it by the full Strichartz norm $\mathcal{S}$ as in (2.2).

Proof. The main ideas is to divide the frequency support of u into several regions, and use Strichartz estimate for the regions where $\tau \sim |\xi|^2 + |\eta|^2$, and use Sobolev for the remaining regions. As we shall see later, the proof of Theorem 3.4-3.5 below uses essentially the same idea.

To begin with, we shall use the decomposition $u = \sum_{k=0}^{\infty} P_{|\xi-\eta|\sim 2^k} u$, where for the case $k = 0$, we are abusing notations a bit by letting $P_{|\xi-\eta|\sim 1} u$ to denote the operator $P_{|\xi-\eta|<1} u$. We have the square function estimate(see e.g., Lemma 3.5 in [10]).

$$\begin{aligned} \|u\|_{\mathcal{S}_{x,y}} &\sim \|\Big(\sum_{k=0}^{\infty} |P_{|\xi-\eta|\sim 2^k} u|^2\Big)^{\frac{1}{2}}\|_{\mathcal{S}_{x,y}} \\ &\lesssim \Big(\sum_{k=0}^{\infty} \|P_{|\xi-\eta|\sim 2^k} u\|^2_{\mathcal{S}_{x,y}}\Big)^{\frac{1}{2}}. \end{aligned} \tag{3.1}$$

We shall focus on the dyadic pieces where $k \geq 1$, since by the Strichartz estimate and the Sobolev estimate, one can easily show that

$$\|P_{|\xi-\eta|<1} u\|_{\mathcal{S}_{x,y}} \lesssim \big\|\langle \nabla_{x+y}\rangle^{\frac{1}{2}} f\big\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}.$$

Now let $u_k = P_{|\xi-\eta|\sim 2^k} u$, $f_k = P_{|\xi-\eta|\sim 2^k} f$, and decompose $u_k = u_k^1 + u_k^2 + u_k^3$, where

$$\begin{aligned} &\mathbf{S}\, u_k^1 = P_{10|\tau|^{\frac{1}{2}} \geq 2^k} f_k, \quad \text{with initial conditions } 0 \\ &\mathcal{F}\, u_k^2 = \frac{\mathcal{F}\Big(P_{10|\tau|^{\frac{1}{2}} \leq 2^k} f_k\Big)}{\tau + |\xi|^2 + |\eta|^2}, \text{ this no longer has initial conditions } 0 \\ &\mathbf{S}\, u_k^3 = 0, \quad \text{a correction so that } u_k^2 + u_k^3 \text{ has initial condition } 0. \end{aligned} \tag{3.2}$$

For u_k^1, by the Strichartz estimate

$$\begin{aligned} \|u_k^1\|_{\mathcal{S}} &\lesssim \|f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} \\ &\lesssim \|f_k\|_{L^{\frac{6}{5}}(d(x-y))L^2(dt)L^2(d(x+y))} \\ &\lesssim 2^{\frac{k}{2}} \|f_k\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}, \end{aligned}$$

where in the last line we used the fact that f_k is frequency supported in $|\xi-\eta| \sim 2^k$ and Bernstein's inequality, which is a (elementary) generalization of the classical Bernstein's inequality to L^2 valued functions.

Now we make another dyadic decomposition, write

$$f_k = \sum_{\ell \geq 0} P_{|\tau| \sim 2^{2k+\ell}} f_k = \sum_{\ell} f_{k,\ell}. \tag{3.3}$$

Note that for each fixed ℓ, k, we have

$$\begin{aligned}\|f_{k,\ell}\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} &\sim 2^{-\frac{\ell}{4}-\frac{k}{2}} \||\partial_t|^{\frac{1}{4}} f_{k,\ell}\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \\ &\lesssim 2^{-\frac{\ell}{4}-\frac{k}{2}} \||\partial_t|^{\frac{1}{4}} P_{|\tau|\sim 2^{2k+\ell}} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\end{aligned}$$

where in the first line we used Bernstein's inequality, and in the second line we used the fact that

$$\|P_{|\xi-\eta|\sim 2^k} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \lesssim \|f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}, \tag{3.4}$$

which can be proved, for example, using the generalized Young's inequality on the space of L^2 valued functions.

Thus, by Minkowski's inequality

$$\begin{aligned}\Big(\sum_{k=0}^{\infty} \|u_k^1\|_{\mathcal{S}_{x,y}}^2\Big)^{\frac{1}{2}} &\lesssim \Big(\sum_{k=0}^{\infty} 2^k \|\sum_{\ell} f_{k,\ell}\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}} \\ &\lesssim \sum_{\ell} \Big(\sum_{k=0}^{\infty} 2^k \|f_{k,\ell}\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}} \\ &\lesssim \sum_{\ell} 2^{-\frac{\ell}{4}} \Big(\sum_{k} \||\partial_t|^{\frac{1}{4}} P_{|\tau|\sim 2^{2k+\ell}} f\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}} \\ &\lesssim \sum_{\ell} 2^{-\frac{\ell}{4}} \Big(\|\big(\sum_{k} ||\partial_t|^{\frac{1}{4}} P_{|\tau|\sim 2^{2k+\ell}} f|^2\big)^{\frac{1}{2}}\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}} \\ &\lesssim \||\partial_t|^{\frac{1}{4}} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))},\end{aligned}$$

where in the last line we used the square function estimate in t variable.

For u_k^2, the denominator is comparable with $|\xi-\eta|^2 + |\xi+\eta|^2 \geq 2^{2k} \geq 100\tau$. Thus, by Sobolev's estimates at an angle, which is Lemma 3.2 in [10], we have

$$\begin{aligned}&\|u_k^2\|_{L^2(dt)L^6(dx)L^2(dy)} + \|u_k^2\|_{L^2(dt)L^6(dy)L^2(dx)} \\ &\lesssim \|\langle\nabla_{x+y}\rangle u_k^2\|_{L^2(dt)L^2(dx)L^2(dy)} \\ &\lesssim \|\Big(\langle\nabla_{x-y}\rangle + \langle\nabla_{x+y}\rangle\Big)^{-2} \langle\nabla_{x+y}\rangle f_k\|_{L^2(dt)L^2(dx)L^2(dy)}.\end{aligned} \tag{3.5}$$

This is also the place where we require the norm on the left side to be $\mathcal{S}_{x,y}$, since we do not have Sobolev-type estimates like

$$\|u_k^2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|\langle\nabla_{x+y}\rangle u_k^2\|_{L^2(dt)L^2(dx)L^2(dy)}.$$

Let us first assume $|\xi+\eta| \le |\xi-\eta|$, and make the decomposition

$$f_k = \sum_{j=0}^{k} P_{|\xi+\eta|\sim 2^{k-j}} f_k = \sum_j f_{k,j}. \tag{3.6}$$

Then for each fixed k, j, we have

$$\begin{aligned}
&\|\Big(\langle\nabla_{x-y}\rangle + \langle\nabla_{x+y}\rangle\Big)^{-2}\langle\nabla_{x+y}\rangle f_{k,j}\|_{L^2(dt)L^2(dx)L^2(dy)}\\
&\sim 2^{-\frac{j}{2}}\|\langle\nabla_{x-y}\rangle^{-\frac{3}{2}}\langle\nabla_{x+y}\rangle^{\frac{1}{2}} f_{k,j}\|_{L^2(dt)L^2(dx)L^2(dy)}\\
&\lesssim 2^{-\frac{j}{2}}\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}} f_{k,j}\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\lesssim 2^{-\frac{j}{2}}\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}} P_{|\xi+\eta|\sim 2^{k-j}} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))},
\end{aligned} \tag{3.7}$$

where in the third line we used Bernstein's inequality, and in the last line we used the fact that

$$\|f_k\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \lesssim \|f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}.$$

Thus, by Minkowski's inequality

$$\begin{aligned}
&\Big(\sum_{k=0}^{\infty}\|u_k^2\|^2_{L^2(dt)L^6(dx)L^2(dy)} + \|u_k^2\|^2_{L^2(dt)L^6(dy)L^2(dx)}\Big)^{\frac{1}{2}}\\
&\quad\lesssim \Big(\sum_{k=0}^{\infty}\|\sum_j \big(\langle\nabla_{x-y}\rangle + \langle\nabla_{x+y}\rangle\big)^{-2}\langle\nabla_{x+y}\rangle f_{k,j}\|^2_{L^2(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}}\\
&\quad\lesssim \sum_j \Big(\sum_k \|\big(\langle\nabla_{x-y}\rangle + \langle\nabla_{x+y}\rangle\big)^{-2}\langle\nabla_{x+y}\rangle f_{k,\ell}\|^2_{L^2(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}}\\
&\quad\lesssim \sum_j 2^{-\frac{j}{2}}\Big(\sum_k \|\langle\nabla_{x+y}\rangle^{\frac{1}{2}} P_{|\xi+\eta|\sim 2^{k-j}} f\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}}\\
&\quad\lesssim \sum_j 2^{-\frac{j}{2}}\Big(\|\big(\sum_k |\langle\nabla_{x+y}\rangle^{\frac{1}{2}} P_{|\xi+\eta|\sim 2^{k-j}} f|^2\big)^{\frac{1}{2}}\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac{1}{2}}\\
&\quad\lesssim \|\langle\nabla_{x+y}\rangle^{\frac{1}{2}} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}
\end{aligned}$$

where in the last line we used the square function estimate in $x+y$ variable.

The case $|\xi+\eta| \ge |\xi-\eta|$ is similar.

For the other endpoint $p=\infty$, define

$$u_k^2 = \sum_{0\le \ell\le k/2} P_{|\tau|\sim 2^{2k-\ell}} u_k^2 = \sum_\ell u_{k,\ell}^2 \tag{3.8}$$

and similarly

$$f_k = \sum_{0\le \ell\le k/2} P_{|\tau|\sim 2^{2k-\ell}} f_k = \sum_{\ell} f_{k,\ell}. \tag{3.9}$$

Note that for each fixed ℓ, k, we have

$$\begin{aligned}
\|u^2_{k,\ell}\|_{L^\infty(dt)L^2(dx)L^2(dy)} &\lesssim \| \int \Big|\frac{\mathcal{F}(f_{k,\ell})}{\tau+|\xi|^2+|\eta|^2}\Big| d\tau\|_{L^2(d\xi)L^2(d\eta)} \\
&\sim 2^{k/2-\ell/4}\|\langle\nabla_{x-y}\rangle^{-2}|\partial_t|^{\frac14} f_{k,\ell}\|_{L^2(dt)L^2(dx)L^2(dy)} \\
&\sim 2^{-\ell/4}\|\langle\nabla_{x-y}\rangle^{-\frac32}|\partial_t|^{\frac14} f_{k,\ell}\|_{L^2(dt)L^2(dx)L^2(dy)} \\
&\lesssim 2^{-\ell/4}\||\partial_t|^{\frac14} f_{k,\ell}\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\lesssim 2^{-\ell/4}\||\partial_t|^{\frac14} P_{|\tau|\sim 2^{2k-\ell}} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))},
\end{aligned} \tag{3.10}$$

where in the third line we used Bernstein's inequality, and in the last line we used the fact that

$$\|f_k\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \lesssim \|f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}.$$

By Minkowski's inequality

$$\begin{aligned}
&\Big(\sum_{k=0}^\infty \|u^2_k\|^2_{L^\infty(dt)L^2(dx)L^2(dy)}\Big)^{\frac12} \\
&= \Big(\sum_{k=0}^\infty \|\sum_{\ell} u^2_{k,\ell}\|^2_{L^\infty(dt)L^2(dx)L^2(dy)}\Big)^{\frac12} \\
&\lesssim \sum_{\ell=0}^\infty \Big(\sum_{k:k\ge 2\ell} \|u^2_{k,\ell}\|^2_{L^\infty(dt)L^2(dx)L^2(dy)}\Big)^{\frac12} \\
&\lesssim \sum_{\ell=0}^\infty 2^{-\ell/4}\Big(\sum_{k:k\ge 2\ell} \||\partial_t|^{\frac14} P_{|\tau|\sim 2^{2k-\ell}} f\|^2_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\Big)^{\frac12} \\
&\lesssim \||\partial_t|^{\frac14} f\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}
\end{aligned}$$

where in the last line we used the square function estimate in t variable.

To deal with u^3_k, note that since u^3_k is solution to free Schrödinger,

$$\begin{aligned}
\|u^3_k\|_{\mathcal{S}_{x,y}} &\lesssim \|u^3_k(0,\cdot,\cdot)\|_{L^2(dx)L^2(dy)} \\
&= \|u^2_k(0,\cdot,\cdot)\|_{L^2(dx)L^2(dy)} \\
&\le \|u^2_k\|_{L^\infty(dt)L^2(dx)L^2(dy)}.
\end{aligned} \tag{3.11}$$

Thus it can be treated as in the previous case.

□

Theorem 3.4. *Let* $\mathbf{S}u = f$ $u(0,\cdot) = 0$. *Then*

$$\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}u\|_{\mathcal{S}_{x,y}} \lesssim \big\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}f\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \big\||\partial_t|^{\frac{1}{4}}f\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}. \tag{3.12}$$

Proof. We shall use the decomposition $u = \sum_{k=0}^{\infty} P_{|\xi-\eta|\sim 2^k}u$, and the square function estimate

$$\begin{aligned}\|u\|_{\mathcal{S}_{x,y}} &\sim \|\Big(\sum_{k=0}^{\infty}|P_{|\xi-\eta|\sim 2^k}u|^2\Big)^{\frac{1}{2}}\|_{\mathcal{S}_{x,y}}\\ &\lesssim \Big(\sum_{k=0}^{\infty}\|P_{|\xi-\eta|\sim 2^k}u\|^2_{\mathcal{S}_{x,y}}\Big)^{\frac{1}{2}}.\end{aligned}$$

Since the right side of (3.12) only involves $L^{\frac{6}{5}}$-norm in the $x-y$ direction, we can add up the dyadic pieces in the right side using the square function estimate, thus it suffices to prove the Theorem for a single dyadic piece where $|\xi-\eta|\sim 2^k$(which would simplify our argument compare with the previous theorem). Here again for the case $k=0$, we are abusing notations a bit by letting $P_{|\xi-\eta|\sim 1}u$ to denote the operator $P_{|\xi-\eta|<1}u$, and the case $k=0$ is easy to handle by just using Strichartz.

Now let $u_k = P_{|\xi-\eta|\sim 2^k}u$ and $f_k = P_{|\xi-\eta|\sim 2^k}f$. We shall use the same decomposition as in (3.2), write $u_k = u_k^1 + u_k^2 + u_k^3$.

For u_k^1, by Strichartz and Bernstein's inequality

$$\begin{aligned}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}u_k^1\|_{\mathcal{S}} &\lesssim \|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\ &\lesssim \||\partial_t|^{\frac{1}{4}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{3.13}$$

For u_k^2, the denominator is comparable with $|\xi-\eta|^2+|\xi+\eta|^2 \geq 2^{2k} \geq 100\tau$. Thus, by Sobolev estimates at an angle, we have

$$\begin{aligned}&\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}u_k^2\|_{L^2(dt)L^6(dx)L^2(dy)} + \|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}u_k^2\|_{L^2(dt)L^6(dy)L^2(dx)}\\ &\lesssim \|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\langle\nabla_{x+y}\rangle u_k^2\|_{L^2(dt)L^2(dx)L^2(dy)}\\ &\lesssim \|\Big(\langle\nabla_{x-y}\rangle+\langle\nabla_{x+y}\rangle\Big)^{-1}\langle\nabla_{x+y}\rangle^{\frac{1}{2}}f_k\|_{L^2(dt)L^2(dx)L^2(dy)}\\ &\lesssim \|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{3.14}$$

For the other endpoint $p=\infty$, since we are in the region $10|\tau|^{\frac{1}{2}} \leq 2^k$, we have

$$
\begin{aligned}
&\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \\
&\lesssim 2^{k/2}\| \int \Big|\frac{\mathcal{F}\Big(P_{10|\tau|^{\frac{1}{2}}\leq\cdot 2^k} f_k\Big)}{\tau+|\xi|^2+|\eta|^2}\Big| d\tau\|_{L^2(d\xi)L^2(d\eta)} \\
&\lesssim 2^k \||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{-2} f_k\|_{L^2(dt)L^2(d(x-y))L^2(d(x+y))} \\
&\lesssim \||\partial_t|^{\frac{1}{4}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.
\end{aligned} \tag{3.15}
$$

To deal with u_k^3, if we repeat the argument in (3.11), we have

$$
\|u_k^3\|_{\mathcal{S}_{x,y}} \lesssim \|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)}.
$$

Thus it can be treated as in the previous case.

□

Theorem 3.5. *Let*

$$
\mathbf{S}u = f,\ u(0,\cdot) = u_0
$$

We have

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}} u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\lesssim \min\Big\{\big\||\partial_t|^{\frac{1}{4}} f\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}, \\
&\qquad \big\|\big(\langle\nabla_{x-y}\rangle^{\frac{1}{2}} + \langle\nabla_{x+y}\rangle^{\frac{1}{2}}\big) f\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}, \\
&\qquad \big\|\big(\langle\nabla_{x-y}\rangle^{\frac{1}{2}} + \langle\nabla_{x+y}\rangle^{\frac{1}{2}}\big) f\big\|_{\mathcal{S}'_r}\Big\} \\
&+ \big\|\big(\langle\nabla_{x-y}\rangle^{\frac{1}{2}} + \langle\nabla_{x+y}\rangle^{\frac{1}{2}}\big) u_0\big\|_{L^2}.
\end{aligned} \tag{3.16}
$$

Proof. For simplicity, we shall only treat the case where $|\xi+\eta| < |\xi-\eta|$, the case $|\xi+\eta| \geq |\xi-\eta|$ is similar and to some extent simpler. We shall use the decomposition $u = \sum_{k=0}^{\infty} P_{|\xi-\eta|\sim 2^k} u$, by using the square function estimate, It suffices to prove the Theorem for a single dyadic piece where $|\xi-\eta| \sim 2^k$. The argument below also works for $k<0$.

Now let $u_k = P_{|\xi-\eta|\sim 2^k} u$ and $f_k = P_{|\xi-\eta|\sim 2^k} f$. We decompose $u_k = u_k^1 + u_k^2 + u_k^3 + u_k^4$, where

$$\mathbf{S}\, u_k^1 = P_{|\tau|^{\frac{1}{2}} \sim 2^k} f_k, \quad \text{with initial conditions } 0$$
$$\mathcal{F}\, u_k^2 = \frac{\mathcal{F}\Big(P_{10|\tau|^{\frac{1}{2}} \leq 2^k} f_k + P_{|\tau|^{\frac{1}{2}} \geq 10\cdot 2^k} f_k\Big)}{\tau + |\xi|^2 + |\eta|^2}, \quad \text{this no longer has initial conditions } 0$$
$$\mathbf{S}\, u_k^3 = 0, \quad \text{a correction so that } u_k^2 + u_k^3 \text{ has initial condition } 0$$
$$\mathbf{S}\, u_k^4 = 0, \quad u_k^4(0,\cdot) = P_{|\xi-\eta|\sim 2^k} u_0.$$

It is easy to handle u_k^4, since in the case where $|\xi+\eta| < |\xi-\eta|$,

$$\begin{aligned} &\||\partial_t|^{\frac{1}{4}} u_k^4\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\ &\quad \sim 2^{k/2} \|e^{it\Delta} P_{|\xi-\eta|\sim 2^k} u_0\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}. \\ &\quad \lesssim \big\|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} P_{|\xi-\eta|\sim 2^k} u_0\big\|_{L^2}. \end{aligned} \tag{3.17}$$

For u_k^1, since for f_k, we have $|\xi-\eta|^2 + |\xi+\eta| \sim 2^{2k}$ and $\tau^{\frac{1}{2}} \sim 2^k$, it is straightforward to check that the dual variable τ to t for u_k^1 is also supported where $|\tau|^{1/2} \sim 2^k$, by Strichartz estimates

$$\begin{aligned} &\||\partial_t|^{\frac{1}{4}} u_k^1\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\ &\quad \lesssim 2^{k/2} \|u_k^1\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\ &\quad \lesssim 2^{k/2} \|P_{|\tau|^{\frac{1}{2}}\sim 2^k} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} \\ &\quad \lesssim \||\partial_t|^{\frac{1}{4}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}. \end{aligned} \tag{3.18}$$

The same argument also gives

$$\||\partial_t|^{\frac{1}{4}} u_k^1\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))},$$

as well as

$$\||\partial_t|^{\frac{1}{4}} u_k^1\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{\mathcal{S}_r'}.$$

For u_k^2, if $10|\tau|^{\frac{1}{2}} \leq 2^k$, the denominator is comparable with $|\xi-\eta|^2 + |\xi+\eta|^2 \sim 2^{2k}$. Thus, by Sobolev's inequality in the $x-y$ direction, we have

$$\||\partial_t|^{\frac{1}{4}} u_k^2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \||\partial_t|^{\frac{1}{4}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.$$

The same argument also gives

$$\||\partial_t|^{\frac{1}{4}} u_k^2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.$$

It remains to show that, if $10|\tau|^{\frac{1}{2}} \leq 2^k$,

$$\||\partial_t|^{\frac{1}{4}} u_k^2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{\mathcal{S}_r'}. \tag{3.19}$$

By interpolation, it suffices to show that

$$\| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \lesssim \| \langle \nabla_{x-y} \rangle^{\frac{1}{2}} f_k \|_{L^2(dt) L^{\frac{6}{5}}(dx) L^2(dy)}, \tag{3.20}$$

$$\| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \lesssim \| \langle \nabla_{x-y} \rangle^{\frac{1}{2}} f_k \|_{L^2(dt) L^{\frac{6}{5}}(dy) L^2(dx)}, \tag{3.21}$$

and

$$\| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \lesssim \| \langle \nabla_{x-y} \rangle^{\frac{1}{2}} f_k \|_{L^1(dt) L^2(dx) L^2(dy)}, \tag{3.22}$$

which would be stronger than (3.19) since it includes two endpoint cases.

The estimates (3.20) and (3.21) follow directly by Sobolev's estimates at an angle. To prove (3.22), first by Sobolev in the $x-y$ direction, we have

$$\begin{aligned} \| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} &\lesssim 2^{-k/2} \| f_k \|_{L^2(dt) L^2(d(x-y)) L^2(d(x+y))}. \\ &= 2^{-k/2} \| f_k \|_{L^2(dt) L^2(dx) L^2(dy)} \end{aligned} \tag{3.23}$$

and now by Bernstein's inequality in the t direction and $x-y$ direction.

$$\begin{aligned} 2^{-k/2} \| f_k \|_{L^2(dt) L^2(dx) L^2(dy)} &\lesssim 2^{k/2} \| f_k \|_{L^1(dt) L^2(dx) L^2(dy)} \\ &\lesssim \| \langle \nabla_{x-y} \rangle^{\frac{1}{2}} f_k \|_{L^1(dt) L^2(dx) L^2(dy)}. \end{aligned} \tag{3.24}$$

If $|\tau|^{\frac{1}{2}} \geq 10 \cdot 2^k$, the denominator is comparable with τ. Thus, by Sobolev in the $x-y$ direction, we have

$$\begin{aligned} &\| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\ &\qquad \lesssim \| |\partial_t|^{-\frac{3}{4}} \langle \nabla_{x-y} \rangle^2 f_k \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \\ &\qquad \lesssim \| |\partial_t|^{\frac{1}{4}} f_k \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))}. \end{aligned} \tag{3.25}$$

The same argument also gives,

$$\| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \lesssim \| \langle \nabla_{x-y} \rangle^{\frac{1}{2}} f_k \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))},$$

when $|\tau|^{\frac{1}{2}} \geq 10 \cdot 2^k$. It remains to show that (3.19) holds if $|\tau|^{\frac{1}{2}} \geq 10 \cdot 2^k$, which would be a consequence of (3.20)-(3.22). And as before, the estimates (3.20) and (3.21) in this case follow by Sobolev's estimates at an angle.

To prove (3.22) when $|\tau|^{\frac{1}{2}} \geq 10 \cdot 2^k$, first by Sobolev in the $x-y$ direction, we have

$$\begin{aligned} \| |\partial_t|^{\frac{1}{4}} u_k^2 \|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} &\lesssim \| \langle \partial_t \rangle^{-\frac{3}{4}} \langle \nabla_{x-y} \rangle f_k \|_{L^2(dt) L^2(d(x-y)) L^2(d(x+y))} \\ &= \| \langle \partial_t \rangle^{-\frac{3}{4}} \langle \nabla_{x-y} \rangle f_k \|_{L^2(dt) L^2(dx) L^2(dy)}. \end{aligned}$$

Now by Minkowski's inequality for dt integral,

$$\begin{aligned}
&\|\langle\partial_t\rangle^{-\frac{3}{4}}\langle\nabla_{x-y}\rangle f_k\|_{L^2(dt)L^2(dx)L^2(dy)} \\
&= \|\tau^{-\frac{3}{4}}\Big|\int e^{-it\tau}\langle\nabla_{x-y}\rangle f_k(t,\cdot)dt\Big|\|_{L^2(d\tau)L^2(dx)L^2(dy)} \\
&\lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle f_k\|_{L^1(dt)L^2(dx)L^2(dy)} \\
&\lesssim \|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^1(dt)L^2(dx)L^2(dy)}.
\end{aligned} \tag{3.26}$$

To deal with u_k^3, note that since u_k^3 is solution to free Schrödinger equation, and since we are assuming $|\xi-\eta|^2+|\xi+\eta|^2\sim 2^{2k}$

$$\begin{aligned}
\||\partial_t|^{\frac{1}{4}}u_k^3\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} &\lesssim 2^{k/2}\|u_k^3(0,\cdot,\cdot)\|_{L^2(dx)L^2(dy)} \\
&= 2^{k/2}\|u_k^2(0,\cdot,\cdot)\|_{L^2(dx)L^2(dy)} \\
&\le 2^{k/2}\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)}.
\end{aligned} \tag{3.27}$$

Thus, it suffices to control $\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)}$.

First, if $|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k$

$$\begin{aligned}
\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} &\lesssim \|\int\Big|\frac{\mathcal{F}\Big(P_{|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k}f_k\Big)}{\tau+|\xi|^2+|\eta|^2}\Big|d\tau\|_{L^2(d\xi)L^2(d\eta)} \\
&\lesssim 2^{-\frac{3k}{2}}\||\partial_t|^{\frac{1}{4}}f_k\|_{L^2(dt)L^2(dx)L^2(dy)} \\
&\lesssim 2^{-k/2}\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{-1}f_k\|_{L^2(dt)L^2(d(x-y))L^2(d(x+y))} \\
&\lesssim 2^{-k/2}\||\partial_t|^{\frac{1}{4}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.
\end{aligned} \tag{3.28}$$

Similarly,

$$\begin{aligned}
\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} &\lesssim \|\int\Big|\frac{\mathcal{F}\Big(P_{|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k}f_k\Big)}{\tau+|\xi|^2+|\eta|^2}\Big|d\tau\|_{L^2(d\xi)L^2(d\eta)} \\
&\lesssim 2^{-k}\|f_k\|_{L^2(dt)L^2(dx)L^2(dy)} \\
&\lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.
\end{aligned} \tag{3.29}$$

It remains to show that

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f_k\|_{\mathcal{S}'_r}. \tag{3.30}$$

In this case, we won't prove (3.30) by interpolation since we do not know if one can show that

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f_k\|_{L^1(dt)L^2(dx)L^2(dy)}, \tag{3.31}$$

which is also the reason why we have the restricted norm $\mathcal{S}'_r$ in the statement of the Theorem.

Instead, we shall prove (3.30) by showing that

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}, \tag{3.32}$$

as well as

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^{p'}(dt)L^{q'}(dy)L^2(dx)}, \tag{3.33}$$

for all admissible pairs (p,q), with $\frac{2}{p}=\frac{3}{2}-\frac{3}{q},\ 2\le p<\infty$.

To prove (3.32) when $|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k$, for admissible pair (p,q) with $p<\infty$,

$$\begin{aligned}
\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim& \|\int\Big|\frac{\mathcal{F}\Big(P_{|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k}f_k\Big)}{\tau+|\xi|^2+|\eta|^2}\Big|d\tau\|_{L^2(d\xi)L^2(d\eta)}\\
\lesssim& 2^{-\frac{2k}{p}}\|\mathcal{F}\Big(P_{|\tau|^{\frac{1}{2}}\ge 10\cdot 2^k}f_k\Big)\|_{L^2(d\xi)L^2(d\eta)L^p(d\tau)}\\
\lesssim& 2^{-\frac{2k}{p}}\|f_k\|_{L^{p'}(dt)L^2(dx)L^2(dy)}\\
\lesssim& 2^{-\frac{2k}{p}}2^{(\frac{3}{2}-\frac{3}{q})k}\|f_k\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}\\
\lesssim& \|f_k\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}\\
\lesssim& 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)},
\end{aligned} \tag{3.34}$$

where we used Hölder's inequality in the second line, the Hausdorff-Young inequality in the third line, and Bernstein's inequality at an angle in the fourth and last line. The proof of (3.33) is similar.

If $10|\tau|^{\frac{1}{2}}\le 2^k$,

$$\begin{aligned}
\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim& \|\int\Big|\frac{\mathcal{F}\Big(P_{10|\tau|^{\frac{1}{2}}\le\cdot 2^k}f_k\Big)}{\tau+|\xi|^2+|\eta|^2}\Big|d\tau\|_{L^2(d\xi)L^2(d\eta)}\\
\lesssim& 2^{-\frac{3k}{2}}\||\partial_t|^{\frac{1}{4}}f_k\|_{L^2(dt)L^2(dx)L^2(dy)}\\
\lesssim& 2^{-k/2}\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{-1}f_k\|_{L^2(dt)L^2(d(x-y))L^2(d(x+y))}\\
\lesssim& 2^{-k/2}\||\partial_t|^{\frac{1}{4}}f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.
\end{aligned} \tag{3.35}$$

The same argument also gives,

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.$$

It remains to show that

$$\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim 2^{-k/2}\|\langle\nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{\mathcal{S}_r'}, \tag{3.36}$$

for the case $10|\tau|^{\frac{1}{2}}\le 2^k$, which would be a consequence of (3.20)-(3.22).

To prove (3.20),

$$
\begin{aligned}
\|u_k^2\|_{L^\infty(dt)L^2(dx)L^2(dy)} \lesssim & \| \int \Big| \frac{\mathcal{F}\Big(P_{|\tau|^{\frac{1}{2}} \geq 10\cdot 2^k} f_k\Big)}{\tau + |\xi|^2 + |\eta|^2} \Big| d\tau \|_{L^2(d\xi)L^2(d\eta)} \\
\lesssim & 2^{-k} \|f_k\|_{L^2(dt)L^2(dx)L^2(dy)} \\
\lesssim & 2^{-k/2} \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^2(dt)L^{\frac{6}{5}}(dx)L^2(dy)}.
\end{aligned}
$$

The proof of (3.21) is similar.

To prove (3.22), we use Bernstein's inequality in the t direction

$$
\begin{aligned}
2^{-k}\|f_k\|_{L^2(dt)L^2(dx)L^2(dy)} \lesssim & \|f_k\|_{L^1(dt)L^2(dx)L^2(dy)} \\
\lesssim & 2^{-k/2} \|\langle \nabla_{x-y}\rangle^{\frac{1}{2}} f_k\|_{L^1(dt)L^2(dx)L^2(dy)}.
\end{aligned} \tag{3.37}
$$

□

Now we shall present several lemmas that involve the collapsing norm.

Lemma 3.6. *If* $\mathbf{S}u = g, u(0,\cdot) = u_0$. *Then*

$$
\|u\|_{collapsing} \lesssim \min\{\|\langle \nabla_x\rangle^{\frac{1}{2}} g\|_{\mathcal{S}_r'} + \|\langle \nabla_x\rangle^{\frac{1}{2}} u_0\|_{L^2}, \|\langle \nabla_y\rangle^{\frac{1}{2}} g\|_{\mathcal{S}_r'} + \|\langle \nabla_y\rangle^{\frac{1}{2}} u_0\|_{L^2}\}.
$$

We record that the above implies

Lemma 3.7. *If* $\mathbf{S}u = g, u(0,\cdot) = u_0$. *Then*

$$
\begin{aligned}
&\|\langle \nabla_x\rangle^{\frac{1}{2}} u\|_{collapsing} + \|\langle \nabla_y\rangle^{\frac{1}{2}} u\|_{collapsing} \\
&\qquad \lesssim \|\langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} g\|_{\mathcal{S}_r'} + \|\langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} u_0\|_{L^2}.
\end{aligned}
$$

We will also need

Lemma 3.8. *If* $\mathbf{S}u = g, u(0,\cdot) = u_0$. *Then*

$$
\|\langle \nabla_{x+y}\rangle^{\frac{1}{2}} u\|_{collapsing} \lesssim \|\langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} g\|_{\mathcal{S}_r'} + \|\langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} u_0\|_{L^2}. \tag{3.38}
$$

The proof of Lemma 3.6 and Lemma 3.8 are similar, for simplicity, we shall only present the proof of Lemma 3.8 here. The proof essentially follows from ideas in Lemma 5.1, 5.3 in [20].

Proof. We shall first prove the homogeneous estimate, let $\mathbf{S}u = 0$, with $u(0,\cdot) = u_0$. Our goal is to show

$$
\sup_{x-y} \||\nabla_{x+y}|^{\frac{1}{2}} u\|_{L^2(dt)L^2(d(x+y))} \lesssim \||\nabla_x|^{\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} u_0\|_{L^2}. \tag{3.39}
$$

This is stronger than desired, since $\||\nabla_x|^{\frac{1}{2}}|\nabla_y|^{\frac{1}{2}}u_0\|_{L^2} \lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}u_0\|_{L^2}$, and also by Lemma 3.6, we have

$$\sup_{x-y}\|u\|_{L^2(dt)L^2(d(x+y))} \lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}u_0\|_{L^2} \lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}u_0\|_{L^2}.$$

To prove (3.39) , let $\widetilde{\Lambda}$ denote the space-time Fourier. For fixed $x-y$, doing Cauchy-Schwarz with measures,

$$\begin{aligned}
&||\nabla_{x+y}|^{\frac{1}{2}}\Lambda(t, \widetilde{x-y, x}+y)(\tau, \xi+\eta)|^2 \\
&\lesssim \int \delta(\tau-|\xi|^2-|\eta|^2)\frac{|\xi+\eta|}{|\xi||\eta|}d(\xi-\eta) \\
&\lesssim \int \delta(\tau-|\xi|^2-|\eta|^2)|\widehat{\nabla_x^{\frac{1}{2}}\nabla_y^{\frac{1}{2}}\Lambda_0}(\xi,\eta)|^2 d(\xi-\eta).
\end{aligned} \tag{3.40}$$

In order to prove the estimate, we must show

$$\sup_{\tau,\xi}\int \delta(\tau-|\xi|^2-|\eta|^2)\frac{|\xi+\eta|}{|\xi||\eta|}d(\xi-\eta) \lesssim 1.$$

Without loss of generality, consider the region $|\xi| \le |\eta|$. If $|\xi| \sim |\eta|$, $\frac{|\xi+\eta|}{|\xi||\eta|} \lesssim \frac{1}{|\xi-\eta|}$ and the integral can be evaluated in polar coordinates. If $|\xi| << |\eta|$ then $|\xi+\eta| \sim |\xi-\eta|$ Writing $\frac{|\xi+\eta|}{|\xi||\eta|} \lesssim \frac{1}{|\xi|} \lesssim \frac{1}{|\xi-\eta|\sqrt{1-\cos(\theta)}}$ where θ is the angle between $\xi-\eta$ and $\xi+\eta$, we estimate

$$\sup_{\tau}\int_0^{\pi}\int \delta(\tau-\rho^2)\frac{1}{\rho\sqrt{1-\cos(\theta)}}\rho^2 d\rho\sin(\theta)d\theta \lesssim 1. \tag{3.41}$$

□

The inhomogeneous estimate (3.38) just follows from the homogeneous estimate (3.39) and the Christ–Kiselev lemma. More precisely, let $T_1 = e^{it(\Delta_x+\Delta_y)}$, so $T_1 : L^2(\mathbb{R}^6) \to L^p(dt)L^q(dx)L^2(dy)$ and

$$T_1^* : L^{p'}(dt)W^{\alpha,q'}(dx)H^{\alpha}(dy) \to H^{\alpha}(dx)H^{\alpha}(dy).$$

Fix $x-y$ and let $T_2 : H^{\alpha}(dx)H^{\alpha}(dy) \to L^2(dt)H^{\alpha}(d(x+y))$ be the operator $f \to \big(e^{it(\Delta_x+\Delta_y)}f\big)(t, x-y, x+y)$. Then the inhomogeneous estimate follows by applying the Christ–Kiselev lemma to $T_2T_1^*$.

Lemma 3.9. *If* $\mathbf{S}u = g, u(0,\cdot) = u_0$. *Then*

$$\||\partial_t|^{\frac{1}{4}}u\|_{collapsing} \lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}g\|_{\mathcal{S}'_r} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}u_0\|_{L^2}. \tag{3.42}$$

Proof. For the homogeneous estimate, it follows from the same argument as above. However, we can not apply Christ–Kiselev lemma here to get inhomogeneous estimate since $|\partial_t|^{\frac{1}{4}}$ does not commute with $1_{[0,t]}$ when we write out the solution using Duhamel's formula. Let $\mathbf{S}u = g$, with $u(0,\cdot) = 0$, it suffices to prove

$$\||\partial_t|^{\frac{1}{4}} u\|_{collapsing} \lesssim \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} g\|_{\mathcal{S}'_r}. \tag{3.43}$$

To prove this, we shall decompose the Fourier support τ and $|\xi+\eta|$ of u into finitely many regions.

Case 1: $\tau^{\frac{1}{2}} \leq 10(1+|\xi+\eta|)$.

In this case, we have

$$\||\partial_t|^{\frac{1}{4}} u\|_{collapsing} \lesssim \|\langle \nabla_{x+y} \rangle^{\frac{1}{2}} u\|_{collapsing},$$

thus the desired estimates follows from Lemma 3.8.

Case 2: If $|\tau|^{\frac{1}{2}} > 2(|\xi|+|\eta|)$.

Write $u = u^1 + u^2$, where

$$\begin{aligned} &\mathcal{F}\, u^1 = \frac{\mathcal{F} f}{\tau + |\xi|^2 + |\eta|^2}, \text{ this no longer has initial conditions 0} \\ &\mathbf{S}\, u^2 = 0, \quad \text{a correction so that } u^1 + u^2 \text{ has initial condition 0.} \end{aligned} \tag{3.44}$$

In this case, it suffices to control u_1 since u_2 is only supported where $|\tau| = |\xi|^2 + |\eta|^2$. The goodness about u_1 is that it has the same Fourier support with f. The strategy is based on

$$\begin{aligned} \||\partial_t|^{\frac{1}{4}} u_1\|_{L^\infty(d(x-y))L^2(d(x+y)dt)} &= \|\tau^{\frac{1}{4}} \frac{\mathcal{F} f}{\tau + |\xi|^2 + |\eta|^2}\|_{L^\infty(d(x-y))L^2(d\tau)d(\xi+\eta)} \\ &\lesssim \|\int |\tau|^{\frac{1}{4}} \Big|\frac{\mathcal{F}\Big(f\Big)}{\tau + |\xi|^2 + |\eta|^2}\Big| d(\xi-\eta)\|_{L^2(d\tau d(\xi+\eta))}. \end{aligned}$$

It suffices to show

$$\begin{aligned} &\|\int |\tau|^{\frac{1}{4}} \Big|\frac{\mathcal{F} f}{\tau + |\xi|^2 + |\eta|^2}\Big| d(\xi-\eta)\|_{L^2(d\tau d(\xi+\eta))} \\ &\qquad \lesssim \||\nabla_x|^{\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} f\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}, \end{aligned} \tag{3.45}$$

as well as

$$\begin{aligned} &\|\int |\tau|^{\frac{1}{4}} \Big|\frac{\mathcal{F} f}{\tau + |\xi|^2 + |\eta|^2}\Big| d(\xi-\eta)\|_{L^2(d\tau d(\xi+\eta))} \\ &\qquad \lesssim \||\nabla_x|^{\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} f\|_{L^{p'}(dt)L^{q'}(dy)L^2(dx)}, \end{aligned} \tag{3.46}$$

for all admissible pairs (p,q), where $\frac{2}{p} = \frac{3}{2} - \frac{3}{q}$, $2 \leq p < \infty$.

For $2 \le p < \infty$, by Cauchy-Schwarz, we have

$$\begin{aligned} \text{LHS(3.45)} &\lesssim \| \frac{|\tau|^{\frac{1}{4}}}{|\tau|} \int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} |\mathcal{F}f| d(\xi-\eta) \|_{L^2(d\tau d(\xi+\eta))} \\ &\lesssim A \big\| |\partial_t|^{\frac{1}{p}-\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} |\nabla_x|^{\frac{1}{2}-\frac{2}{p}} f \big\|_{L^2(dt)d(x-y)d(x+y)} \end{aligned} \tag{3.47}$$

where

$$A = \sup_{\tau, \xi+\eta} \frac{|\tau|^{\frac{3}{4}-\frac{1}{p}}}{|\tau|} \left(\int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} \frac{|\xi|^{\frac{4}{p}-1}}{|\eta|} d(\xi-\eta) \right)^{\frac{1}{2}}. \tag{3.48}$$

Changing variables, this is something like

$$A = \sup_{\tau, |u|<|\tau|^{\frac{1}{2}}} \frac{|\tau|^{\frac{3}{4}-\frac{1}{p}}}{|\tau|} \left(\int_{|v|<|\tau|^{\frac{1}{2}}} \frac{|u+v|^{\frac{4}{p}-1}}{|u-v|} dv \right)^{\frac{1}{2}}.$$

After a change of variables this is reduced to $\tau = 1$, and A is bounded.

Since by Sobolev,

$$\begin{aligned} &\big\| |\partial_t|^{\frac{1}{p}-\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} |\nabla_x|^{\frac{1}{2}-\frac{2}{p}} f \big\|_{L^2(dt)d(x-y)d(x+y)} \\ &= \big\| |\partial_t|^{\frac{1}{2}-\frac{1}{p}} |\nabla_x|^{\frac{3}{q}-\frac{3}{2}} |\nabla_y|^{\frac{1}{2}} |\nabla_x|^{\frac{1}{2}} f \big\|_{L^2(dtdxdy)} \\ &\lesssim \| |\nabla_x|^{\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} f \|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}. \end{aligned} \tag{3.49}$$

Thus the proof of (3.45) is complete, and the proof of (3.46) is similar.

Case 3: $|\xi| + |\eta| > 2|\tau|^{\frac{1}{2}}$

In this case, due to Case 1, we can assume additionally that $|\tau|^{\frac{1}{2}} > 10(1+|\xi+\eta|)$. Thus $|\xi-\eta| > |\xi+\eta|$, so also $|\xi-\eta| > |\tau|^{\frac{1}{2}}$. As before, it suffices to show (3.45) and (3.46).

$$\begin{aligned} \text{LHS(3.45)} &\lesssim \| \int_{2|\xi-\eta|>|\xi+\eta|+|\tau|^{\frac{1}{2}}} | \frac{|\tau|^{\frac{1}{4}} |\mathcal{F}f|}{|\xi-\eta|^2} d(\xi-\eta) \|_{L^2(d\tau d(\xi+\eta))} \\ &\lesssim A \big\| |\partial_t|^{\frac{1}{p}-\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} |\nabla_x|^{\frac{1}{2}-\frac{2}{p}} f \big\|_{L^2(dt)d(x-y)d(x+y)} \end{aligned} \tag{3.50}$$

In this case,

$$A^2 = \sup_{\xi+\eta, \tau} \int_{2|\xi-\eta|>|\xi+\eta|+|\tau|^{\frac{1}{2}}} \frac{|\tau|^{\frac{3}{2}-\frac{2}{p}}}{|\xi-\eta|^4} \frac{|\xi|^{\frac{4}{p}-1}}{|\eta|} d(\xi-\eta) \tag{3.51}$$

Again we scale to $|\tau|^{\frac{1}{2}} + |\xi+\eta| = 1$ and have to estimate

$$\int_{|v|>1} \frac{1}{|v|^4} \frac{|u+v|^{\frac{4}{p}-1}}{|u-v|} dv$$

This is bounded uniformly in $|u| < 1$. As before, the rest of the proof follow from Sobolev's inequality, and the proof of (3.46) is similar.

Case 4 :$\frac{1}{2}(|\xi| + |\eta|) < |\tau|^{\frac{1}{2}} < 2(|\xi| + |\eta|)$, In this case, due to Case 1, we can assume additionally that $|\tau|^{\frac{1}{2}} > 10(1 + |\xi + \eta|)$ so in this case $|\tau|^{\frac{1}{2}} \sim |\xi| \sim |\eta| \sim |\xi - \eta|$.

We shall use the decomposition $u = \sum_{k=0}^{\infty} P_{|\tau|\sim 2^k} u$, and the square function estimate

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}} u\|_{L^\infty(d(x-y))L^2(d(x+y)dt)} \\
&\sim \|\Big(\sum_{k=0}^{\infty} |P_{|\tau|\sim 2^k} |\partial_t|^{\frac{1}{4}} u|^2\Big)^{\frac{1}{2}}\|_{L^\infty(d(x-y))L^2(d(x+y)dt)} \\
&\lesssim \Big(\sum_{k=0}^{\infty} 2^{2k} \|P_{|\tau|\sim 2^k} u\|^2_{L^\infty(d(x-y))L^2(d(x+y)dt)}\Big)^{\frac{1}{2}}.
\end{aligned} \tag{3.52}
$$

For each fixed dyadic piece $P_{|\tau|\sim 2^k} u$, in the current case, we have

$$
P_{|\tau|\sim 2^k} u = P_{|\tau|\sim 2^k} P_{|\xi|\sim 2^{k/2}} P_{|\eta|\sim 2^{k/2}} u,
$$

which implies

$$
\begin{aligned}
\||P_{|\tau|\sim 2^k} u\|_{collapsing} &\lesssim \|\langle \nabla_x \rangle^{\frac{1}{2}} P_{|\xi|\sim 2^{k/2}} P_{|\eta|\sim 2^{k/2}} u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\lesssim 2^{-k/2} \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} P_{|\eta|\sim 2^{k/2}} u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}
\end{aligned}
$$

where we used Bernstein's inequality in rotated coordinates twice, see e.g., Lemma 3.1 in [10] for more details.

Thus,

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}} u\|_{L^\infty(d(x-y))L^2(d(x+y)dt)} \\
&\lesssim \Big(\sum_{k=0}^{\infty} \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} P_{|\eta|\sim 2^{k/2}} u\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\Big)^{\frac{1}{2}} \\
&\lesssim \left(\sum_{k=0}^{\infty} \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} P_{|\eta|\sim 2^{k/2}} f\|^2_{\mathcal{S}'}\right)^{\frac{1}{2}} \\
&\lesssim \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} f\|_{\mathcal{S}'}
\end{aligned} \tag{3.53}
$$

where we used Strichartz (Theorem 3.1) in the second line, and square function estimates in y the the last line.

□

4. **Proof of Theorem 2.1.**

In this section, we shall see how we can apply the theorems in the previous section to prove Theorem 2.1. Throughout this section, we shall use extensively the fact that

$$\operatorname{supp} \hat{v} \subset B_1(0)$$

This assumption implies that, the multiplication operator $Tu = v_N u$ can at most enlarge the Fourier support of u by a set of size N, which will greatly simplify our proof, especially the proof of Theorem 4.1 below. If we assume v satisfies (1.25) without this condition, one can follow similar steps in this section to get the same conclusion. In that case, the multiplication operator $Tu = v_N u$ can enlarge the Fourier support of u by a set of arbitrary large size, but with a rapidly decay constant if the new Fourier support deviates from the Fourier support of u by a large distance. For the sake of simplicity, we do not present the full details here for this general case.

To begin with, we shall first prove the following theorem involving collapsing norms at low frequency

Theorem 4.1. *Let* Λ *satisfy* (2.11)*, we have*

$$\begin{aligned}
&\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{\textit{low collapsing}} + \||\partial_t|^{\frac{1}{4}}\Lambda\|_{\textit{low collapsing}} \\
&\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r} + \varepsilon\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{\textit{collapsing}} + \varepsilon\||\partial_t|^{\frac{1}{4}}H\|_{\textit{collapsing}} \\
&\quad + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}
\end{aligned} \tag{4.1}$$

where the norms $\|\cdot\|_{\textit{collapsing}}$ *and* $\|\cdot\|_{\textit{low collapsing}}$ *are defined as in* (2.8) *and* (2.9)*, and* ε *is defined as in* (1.25)*.*

Remark 4.2. Due to the criticalness of the potential $N^2 v(N(x-y))$ in (2.11), it is still open to us if one can prove the above theorem without the frequency assumption on Λ, i.e., to replace on the left side the $\|\cdot\|_{\textit{low collapsing}}$ by the full collapsing norm. However, in the case where Λ satisfy (2.11) with $H = 0$, we do know how to control the full collapsing norm

$$\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{\textit{collapsing}} + \||\partial_t|^{\frac{1}{4}}\Lambda\|_{\textit{collapsing}}$$

by using a different argument, the details are given in Section 8. Also in [10], a stronger version of the theorem is proved for the case when one replace $N^2 v(N(x-y))$ by $N^{3\beta-1}v(N^\beta(x-y))$ for some $\beta < 1$.

Proof. We shall focus on $\||\partial_t|^{\frac{1}{4}}\Lambda\|_{\textit{low collapsing}}$, the proof for the term $\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{\textit{low collapsing}}$ is similar. And we will first show that

$$\begin{aligned}
&\left\|P_{|\xi-\eta|<20N}|\partial_t|^{\frac{1}{4}}\Lambda\right\|_{\textit{collapsing}} \\
&\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r} + \varepsilon\||\partial_t|^{\frac{1}{4}}H\|_{\textit{collapsing}} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}.
\end{aligned} \tag{4.2}$$

Let $\rho \in C_0^\infty(\mathbb{R}^3)$ be a smooth partition of unity, which satisfies

$$\sum_{j\in\mathbb{Z}^3} \rho(\xi - j) \equiv 1, \forall\, \xi \in \mathbb{R}^3.$$

We also assume that $0 \le \rho \le 1$, $\rho \equiv 1$ if $|\xi| \le \frac{1}{2}$, and $\operatorname{supp}\rho \in B_1(0)$, the unit ball in $\mathbb{R}^3$ centered at origin. Let $\psi_j(x-y)$ be the inverse Fourier transform of $\rho(\frac{\xi-\eta}{40N} - j)$. For each fixed $j \in \mathbb{Z}^3$, ψ_j is Fourier supported in a ball of radius $40N$ centered at $40N \cdot j$. Denote

$$P_j f = \psi_j * f$$

so that the Fourier transform of $P_j f = \rho(\frac{\xi-\eta}{40N} - j)\hat{f}$. In particular, let $P_0 f$ denote projection onto ball of radius $40N$ centered at origin, so that the Fourier transform of $P_0 f = \rho(\frac{\xi-\eta}{40N})\hat{f}$.

Define

$$\||\partial_t|^{\frac{1}{4}}\Lambda\|_{\mathcal{N}} = \sum_{k=0}^{N} 2^{-k}\Big(\sum_{k\le|j|<k+1} \big\|P_j|\partial_t|^{\frac{1}{4}}\Lambda\big\|_{collapsing}\Big). \tag{4.3}$$

It is clear that

$$\big\|P_{|\xi-\eta|<20N}|\partial_t|^{\frac{1}{4}}\Lambda\big\|_{collapsing} \lesssim \big\|P_0|\partial_t|^{\frac{1}{4}}\Lambda\big\|_{collapsing} \lesssim \||\partial_t|^{\frac{1}{4}}\Lambda\|_{\mathcal{N}}.$$

Thus it suffices to show that

$$\begin{aligned} &\||\partial_t|^{\frac{1}{4}}\Lambda\|_{\mathcal{N}} \\ &\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r} + \varepsilon\||\partial_t|^{\frac{1}{4}}H\|_{collapsing} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}. \end{aligned} \tag{4.4}$$

We shall first deal with the last term in the norm, where $k = N$. Note that for each fixed $j \in \mathbb{Z}^3$, the Fourier transform in $x-y$ direction of $P_j|\partial_t|^{\frac{1}{4}}\Lambda$ is supported in a ball of radius $40N$, by Bernstein's inequality,

$$\begin{aligned} &\|P_j|\partial_t|^{\frac{1}{4}}\Lambda\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))} \\ &\qquad\lesssim N^{\frac{1}{2}}\|P_j|\partial_t|^{\frac{1}{4}}\Lambda\|_{L^6(d(x-y))L^2(dt)L^2(d(x+y))} \\ &\qquad\lesssim N^{\frac{1}{2}}\||\partial_t|^{\frac{1}{4}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}. \end{aligned} \tag{4.5}$$

Recall that

$$\begin{aligned} \mathbf{S}\Lambda(t,x,y) &= N^2 v\big(N(x-y)\big)\Lambda(t,x,y) + G(t,x,y) \\ &\quad + N^2 v\big(N(x-y)\big)H(t,x,y) \\ \Lambda(0,\cdot) &= \Lambda_0. \end{aligned}$$

By Theorem 3.5, we have

$$
\begin{aligned}
&\| |\partial_t|^{\frac{1}{4}} \Lambda \|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\quad \leq C \big\| N^2 v\big(N(x-y)\big) |\partial_t|^{\frac{1}{4}} \Lambda \big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}, \\
&\quad + C \big\| N^2 v\big(N(x-y)\big) |\partial_t|^{\frac{1}{4}} H \big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}, \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) G \big\|_{\mathcal{S}'_r} \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) \Lambda_0 \big\|_{L^2}, \\
&\quad \leq C\varepsilon \big\| |\partial_t|^{\frac{1}{4}} \Lambda \big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}, \\
&\quad + C\varepsilon N^{-1/2} \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) G \big\|_{\mathcal{S}'_r} \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) \Lambda_0 \big\|_{L^2},
\end{aligned} \tag{4.6}
$$

where we used Hölder in the second inequality. By choosing ε small enough such that $C\varepsilon < 1/2$, we have

$$
\begin{aligned}
&\| |\partial_t|^{\frac{1}{4}} \Lambda \|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\quad \leq C\varepsilon N^{-1/2} \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) G \big\|_{\mathcal{S}'_r} \\
&\quad + C \big\| \big(\langle \nabla_{x-y}\rangle^{\frac{1}{2}} + \langle \nabla_{x+y}\rangle^{\frac{1}{2}}\big) u_0 \big\|_{L^2}, \\
&\quad \leq C\varepsilon N^{-1/2} \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\quad + C \big\| \langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} G \big\|_{\mathcal{S}'_r} + C \big\| \langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} u_0 \big\|_{L^2}.
\end{aligned} \tag{4.7}
$$

If we combine (4.5) and (4.7), we get

$$
\begin{aligned}
&\| P_j |\partial_t|^{\frac{1}{4}} \Lambda \|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\quad \leq C\varepsilon \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\quad + C N^{\frac{1}{2}} \big\| \langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} G \big\|_{\mathcal{S}'_r} + C N^{\frac{1}{2}} \big\| \langle \nabla_x\rangle^{\frac{1}{2}} \langle \nabla_y\rangle^{\frac{1}{2}} u_0 \big\|_{L^2}.
\end{aligned} \tag{4.8}
$$

Since the number of $j \in \mathbb{Z}^3$ such that $N \le |j| < N+1$ is bounded by CN^2, we have

$$\begin{aligned}
&\sum_{N \le |j| < N+1} 2^{-N} \big\| P_j |\partial_t|^{\frac{1}{4}} \Lambda \big\|_{collapsing} \\
&\qquad \le C 2^{-N} N^2 \varepsilon \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\qquad\quad + C 2^{-N} N^{\frac{5}{2}} \big\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \big\|_{\mathcal{S}'_r} \\
&\qquad\qquad + C 2^{-N} N^{\frac{5}{2}} \big\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} u_0 \big\|_{L^2}. \\
&\qquad \le C \varepsilon \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}, \\
&\qquad\quad + C \big\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \big\|_{\mathcal{S}'_r} + C \big\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} u_0 \big\|_{L^2}.
\end{aligned} \tag{4.9}$$

Now we shall control the terms $0 \le k \le N-1$ using a bootstrap argument. Write $\Lambda = \Lambda_1 + \Lambda_2 + \Lambda_3$, where

$$\begin{aligned}
&\mathbf{S}\,\Lambda_1 = G, \quad \Lambda_1(0, \cdot) = \Lambda_0 \\
&\mathbf{S}\,\Lambda_2 = N^2 v\big(N(x-y)\big) H(t,x,y), \quad \text{with initial conditions } 0 \\
&\mathbf{S}\,\Lambda_3 = N^2 v\big(N(x-y)\big) \Lambda(t,x,y), \quad \text{with initial conditions } 0.
\end{aligned}$$

By Lemma 3.9, we have

$$\begin{aligned}
&\sum_{k=0}^{N-1} 2^{-k} \Big(\sum_{k \le |j| < k+1} \big\| P_j |\partial_t|^{\frac{1}{4}} \Lambda_1 \big\|_{collapsing} \Big) \\
&\quad \lesssim \sum_{k=0}^{N-1} 2^{-k} \Big(\sum_{k \le |j| < k+1} \big\| |\partial_t|^{\frac{1}{4}} \Lambda_1 \big\|_{collapsing} \Big) \\
&\quad \lesssim \sum_{k=0}^{N-1} 2^{-k} \sum_{k \le |j| < k+1} \Big(\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \|_{\mathcal{S}'_r} + \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_0 \|_{L^2} \Big) \\
&\quad \lesssim \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \|_{\mathcal{S}'_r} + \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_0 \|_{L^2}.
\end{aligned} \tag{4.10}$$

And if we repeat the argument in (4.5)-(4.8), for each fixed j, we have

$$\begin{aligned}
&\| P_j |\partial_t|^{\frac{1}{4}} \Lambda_2 \|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\qquad\qquad \le C \varepsilon \big\| |\partial_t|^{\frac{1}{4}} H \big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}.
\end{aligned} \tag{4.11}$$

Thus,

$$\begin{aligned}
&\sum_{k=0}^{N-1} 2^{-k}\Big(\sum_{k\le|j|<k+1}\big\|P_j|\partial_t|^{\frac14}\Lambda_2\big\|_{collapsing}\Big)\\
&\lesssim \sum_{k=0}^{N-1}\varepsilon 2^{-k}\Big(\sum_{k\le|j|<k+1}\big\||\partial_t|^{\frac14}H\big\|_{collapsing}\Big)\\
&\lesssim\varepsilon\big\||\partial_t|^{\frac14}H\big\|_{collapsing}.
\end{aligned}\tag{4.12}$$

It remains to control the terms involving Λ_3, note that since $\hat{v}$ is supported in the unit ball centered at origin, we have

$$\begin{aligned}
\mathbf{S}\,P_j\Lambda_3 &= P_jN^2v\big(N(x-y)\big)\Lambda(t,x,y)\\
&= P_jN^2v\big(N(x-y)\big)\sum_{j':|j'-j|\le 1}P_{j'}\Lambda(t,x,y).
\end{aligned}$$

Thus we have the following analog of (4.11)

$$\begin{aligned}
&\|P_j|\partial_t|^{\frac14}\Lambda_3\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad\le C\sum_{j':|j'-j|\le1}\varepsilon\big\|P_{j'}|\partial_t|^{\frac14}\Lambda\big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))},
\end{aligned}\tag{4.13}$$

Thus,

$$\begin{aligned}
&\sum_{k=0}^{N-1} 2^{-k}\Big(\sum_{k\le|j|<k+1}\big\|P_j|\partial_t|^{\frac14}\Lambda_3\big\|_{collapsing}\Big)\\
&\lesssim \sum_{k=0}^{N-1}\varepsilon 2^{-k}\Big(\sum_{k\le|j|<k+1}\sum_{j':|j'-j|\le1}\big\|P_{j'}|\partial_t|^{\frac14}\Lambda\big\|_{collapsing}\Big).
\end{aligned}\tag{4.14}$$

By choosing ε small enough, the right hand side of (4.14) is bounded by $\frac12\||\partial_t|^{\frac14}\Lambda\|_{\mathcal{N}}$, thus the proof of (4.4) is complete.

To prove (4.2) when $P_{|\xi-\eta|<20N}$ is replace by $P_{|\xi|<20N}$ or $P_{|\eta|<20N}$, we use a complete similar argument, the only necessary change is to replace the use of Bernstein's inequality in (4.5) by Bernstein's inequality in rotated coordinates.

□

Proof of Theorem 2.1.

To prove Theorem 2.1, it remains to control $\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{\mathcal{S}_{x,y}}$, where $\|\cdot\|_{\mathcal{S}_{x,y}}$ is defined as in (2.1).

Case 1. Let us first assume $|\xi+\eta|\ge\frac{N}{10}$, the norms we are going to control are different for the case $|\xi+\eta|<\frac{N}{10}$, but in both cases the norms contains $\|\Lambda\|_{\mathcal{S}_{x,y}}$.

In this case, the norm with respect to which the potential is a perturbation should be

$$\|\Lambda\|_{\mathcal{N}} = \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{\mathcal{S}} \tag{4.15}$$

$$+ \|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda\|_{collapsing} \tag{4.16}$$

$$+ \|P_{|\xi|<10N}P_{|\eta|\geq 10N}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{collapsing} \tag{4.17}$$

where $\|\cdot\|_{\mathcal{S}}$ is defined in (2.2) and $\|\cdot\|_{collapsing}$ is defined as in (2.8). Here we are abusing notations a bit by using Λ to denote $P_{|\xi+\eta|\geq\frac{N}{10}}\Lambda$. The projection on $|\xi+\eta|\geq\frac{N}{10}$ is necessary, as we shall see later in the proof, we do not know how to control the collapsing norms in (4.16) and (4.17) without this assumption.

Let's first control the low frequency, let $P_{<N} = P_{|\xi|<10N}P_{|\eta|<10N}$ and $P_{>N} = I - P_{<N}$. Note that

$$\begin{aligned}
&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}\Lambda\\
&\quad=P_{<N}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}N^2v(N(x-y))P_{|\xi-\eta|<20N}\Lambda\\
&\qquad+\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}G(t,x,y)\\
&\qquad+P_{<N}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}N^2v(N(x-y))H(t,x,y)\\
&\quad\sim P_{<N}N^3v(N(x-y))P_{|\xi-\eta|<20N}\Lambda\\
&\qquad+\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}G(t,x,y)\\
&\qquad+P_{<N}N^3v(N(x-y))H(t,x,y)
\end{aligned} \tag{4.18}$$

Here we are abusing the notation a bit by writing

$$P_{<N}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}N^2v(N(x-y))P_{|\xi-\eta|<20N} \sim P_{<N}N^3v(N(x-y))P_{|\xi-\eta|<20N},$$

since, as a result of Bernstein's inequality, for all $1\leq p,q\leq\infty$, we have

$$\begin{aligned}
&\|P_{<N}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}N^2v(N(x-y))P_{|\xi-\eta|<20N}\Lambda\|_{L^p(x-y)L^q(x+y)}\\
&\qquad\lesssim\|P_{<N}N^3v(N(x-y))P_{|\xi-\eta|<20N}\Lambda\|_{L^p(x-y)L^q(x+y)},
\end{aligned} \tag{4.19}$$

which is harmless for our purposes. We shall use the same notation $\sim$ repeatly in the later arguments. Also, strictly speaking,

$$P_{<N}N^2v(N(x-y))\Lambda = P_{<N}N^2v(N(x-y))P_{|\xi-\eta|<21N}\Lambda,$$

instead of $P_{|\xi-\eta|<20N}\Lambda$, due to the fact that convolution with $\hat{v}$ will shift the frequency support. But it will not make a essential difference in our argument.

By Strichartz estimate (Theorem 3.1), we have

$$
\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}\Lambda\|_{\mathcal{S}}\\
&\lesssim \|P_{|\xi-\eta|<20N}N^3v(N(x-y))\Lambda\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\
&\quad+\|P_{|\xi-\eta|<20N}N^3v(N(x-y))H\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\
&\quad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}\\
&\lesssim \|v\|_{L^{\frac{6}{5}}}\left(\|P_{|\xi-\eta|<20N}\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{collapsing}+\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{collapsing}\right)\\
&\quad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2},
\end{aligned}
\tag{4.20}
$$

where in the second inequality we used Bernstein's inequality and the fact that $|\xi+\eta|\geq\frac{N}{10}$.

By Theorem 4.1, we have

$$
\begin{aligned}
\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}\Lambda\|_{\mathcal{S}}\lesssim{}&\varepsilon\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{collapsing}\\
&+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{<N}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}.
\end{aligned}
\tag{4.21}
$$

At high frequency, write

$$
\begin{aligned}
P_{>N}\Lambda={}&P_{|\xi|\geq 10N}P_{|\eta|\geq 10N}\Lambda+P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\\
&+P_{|\xi|<10N}P_{|\eta|\geq 10N}\Lambda\\
={}&I+II+III.
\end{aligned}
\tag{4.22}
$$

To handle the first term I, note that

$$
\begin{aligned}
&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|\geq 10N}\Lambda\\
&\quad\sim N^2v(N(x-y))\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 9N}P_{|\eta|\geq 9N}\Lambda\\
&\qquad+\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|\geq 10N}G(t,x,y)\\
&\qquad+N^2v(N(x-y))\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 9N}P_{|\eta|\geq 9N}H.
\end{aligned}
\tag{4.23}
$$

The lower bounds on $|\xi|$ and $|\eta|$ changed slightly due to convolution with $\hat{v}_N$, since $\hat{v}_N$ is compact supported in a set of size N.

Using Strichartz(Theorem 3.1) and Hölder's inequality, it is not hard to see that

$$
\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|\geq 10N}\Lambda\|_{\mathcal{S}}\\
&\lesssim\varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{\mathcal{S}}\\
&\quad+\varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\qquad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2},
\end{aligned}
\tag{4.24}
$$

where the ε comes from $\|v\|_{L^{3/2}}$ when applying Hölder's inequality.

To handle the second term II,

$$
\begin{aligned}
&\mathbf{S}\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\\
&\quad\sim N^{\frac52}v(N(x-y))\langle\nabla_x\rangle^{\frac12}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\\
&\qquad+N^2v(N(x-y))\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}P_{|\xi|\geq 10N}P_{10N\leq|\eta|<11N}\Lambda\\
&\qquad+N^3v(N(x-y))P_{9N\leq|\xi|<10N}P_{|\eta|<10N}\Lambda\\
&\qquad+N^{\frac52}v(N(x-y))\langle\nabla_x\rangle^{\frac12}P_{|\xi|\geq 9N}P_{|\eta|<11N}H\\
&\qquad+\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}P_{|\xi|\geq 10N}P_{|\eta|<10N}G(t,x,y).
\end{aligned}
\tag{4.25}
$$

Again the bounds on $|\xi|$ and $|\eta|$ changed slightly due to convolution with $\hat{v}_N$. By using Strichartz and Hölder's inequality,

$$
\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\|_{\mathcal{S}}\\
&\quad\lesssim\varepsilon\|\langle\nabla_x\rangle^{\frac12}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad+\varepsilon\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\qquad+\|N^3v(N(x-y))P_{9N\leq|\xi|<10N}P_{|\eta|<10N}\Lambda\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&\qquad+\varepsilon\|\langle\nabla_x\rangle^{\frac12}H\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad+\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}G\|_{\mathcal{S}'_r}+\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda_0\|_{L^2}.
\end{aligned}
\tag{4.26}
$$

Again, since we are assuming $|\xi+\eta|\geq\frac{N}{10}$, by Bernstein's inequality and Hölder's inequality, the third term on the right side can be controlled by

$$
\begin{aligned}
&\|N^3v(N(x-y))P_{9N\leq|\xi|\leq 10N}P_{|\eta|\leq 10N}\Lambda\|_{L^2(dt)L^{6/5}(d(x-y))L^2(d(x+y))}\\
&\quad\lesssim\|P_{|\xi-\eta|<20N}\langle\nabla_{x+y}\rangle^{\frac12}\Lambda\|_{collapsing}.
\end{aligned}
$$

The third term III can be handled in a similar way as the second term.

To finish to discussion for the case $|\xi+\eta|\geq\frac{N}{10}$, we are reduced to estimate

$$
\begin{aligned}
&\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac12}\Lambda\|_{collapsing}\\
&\quad+\|P_{|\xi|<10N}P_{|\eta|\geq 10N}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{collapsing}.
\end{aligned}
\tag{4.27}
$$

We shall focus on the first term, since the other term involving $\langle\nabla_y\rangle^{\frac12}$ can be dealt with similarly.

Write $\Lambda=\Lambda_1+\Lambda_2+\Lambda_3$, where

$$
\begin{aligned}
&\mathbf{S}\,\Lambda_1=G,\quad \Lambda_3(0,\cdot)=\Lambda_0\\
&\mathbf{S}\,\Lambda_2=N^2v\big(N(x-y)\big)H(t,x,y),\quad\text{with initial conditions }0\\
&\mathbf{S}\,\Lambda_3=N^2v\big(N(x-y)\big)\Lambda(t,x,y),\quad\text{with initial conditions }0.
\end{aligned}
\tag{4.28}
$$

By Lemma 3.7, we have

$$\begin{aligned}&\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_1\|_{collapsing},\\ &\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}.\end{aligned} \tag{4.29}$$

To handle Λ_2, since $|\eta|<10N$, by Bernstein's inequality at an angle, we have,

$$\begin{aligned}&\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_2\|_{collapsing}\\ &\qquad\lesssim N^{1/2}\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{4.30}$$

Note that

$$\begin{aligned}&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}N^{1/2}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda_2\\ &\qquad\sim N^{5/2}v(N(x-y))\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 9N}P_{|\eta|<11N}H.\end{aligned} \tag{4.31}$$

By Strichartz and Hölder's inequality

$$\begin{aligned}&N^{1/2}\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_2\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad\lesssim \|N^{\frac{5}{2}}v(N(x-y))P_{|\xi|\geq 9N}P_{|\eta|<11N}H\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\ &\quad\lesssim \varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}H\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}.\end{aligned} \tag{4.32}$$

To handle Λ_3, since $|\eta|<10N$, by Bernstein's inequality at an angle, we still have,

$$\begin{aligned}&\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_3\|_{collapsing}\\ &\qquad\lesssim N^{1/2}\|P_{|\xi|\geq 10N}P_{|\eta|<10N}\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_3\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{4.33}$$

Note that

$$\begin{aligned}&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}N^{1/2}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda_3\\ &\qquad\sim N^{\frac{5}{2}}v(N(x-y))\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\\ &\qquad+N^{\frac{5}{2}}v(N(x-y))\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{10N\leq|\eta|<11N}\Lambda\\ &\qquad+N^{3}v(N(x-y))P_{9N\leq|\xi|<10N}P_{|\eta|<10N}\Lambda.\end{aligned} \tag{4.34}$$

By Strichartz and Hölder's inequality

$$
\begin{aligned}
&N^{\frac{1}{2}}\|\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda_3\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\quad\lesssim \varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad+\varepsilon N^{\frac{1}{2}}\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\qquad+\|N^3v(N(x-y))P_{9N\leq|\xi|<10N}P_{|\eta|<10N}\Lambda\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))},\\
&\quad\lesssim \varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\geq 10N}P_{|\eta|<10N}\Lambda\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad+\varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\qquad+\|P_{|\xi-\eta|<20N}\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{collapsing},
\end{aligned}
\tag{4.35}
$$

where for the second term on the right side we used Bernstein's inequality and the fact that $|\eta|\geq 10N$, and for the third term in the right side, we used Bernstein's inequality and the fact that $|\xi+\eta|\geq\frac{N}{10}$.

Case 2. $|\xi+\eta|<\frac{N}{10}$, $|\xi-\eta|<10N$.

In this case, we are only able to control $\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{\mathcal{S}_{x,y}}$ instead of $\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda\|_{\mathcal{S}}$, recall that $\|\cdot\|_{\mathcal{S}_{x,y}}$ is defined as in (2.1). Note that

$$
\begin{aligned}
&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}\Lambda\\
&\quad\sim P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}N^3v(N(x-y))P_{|\xi-\eta|<20N}\Lambda\\
&\qquad+\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}G(t,x,y)\\
&\qquad+P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}N^3v(N(x-y))H(t,x,y)
\end{aligned}
\tag{4.36}
$$

By using Theorem 3.2 for the first and third term, and the Strichartz estimate (Theorem 3.1) for the second term on the right side of (4.36), we have

$$
\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}\Lambda\|_{\mathcal{S}_{x,y}}\\
&\lesssim\|P_{|\xi-\eta|<20N}N^3v(N(x-y))\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\quad+\|P_{|\xi-\eta|<20N}N^3v(N(x-y))|\partial_t|^{\frac{1}{4}}\Lambda\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\quad+\|P_{|\xi-\eta|<20N}N^3v(N(x-y))\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\quad+\|P_{|\xi-\eta|<20N}N^3v(N(x-y))|\partial_t|^{\frac{1}{4}}H\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\qquad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}\\
&\lesssim\|v\|_{L^1}\Big(\|P_{|\xi-\eta|<20N}\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{collapsing}+\|P_{|\xi-\eta|<20N}|\partial_t|^{\frac{1}{4}}\Lambda\|_{collapsing}\\
&\qquad\qquad+\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{collapsing}+\||\partial_t|^{\frac{1}{4}}H\|_{collapsing}\Big)\\
&\qquad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}_r'}+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2},
\end{aligned}
\tag{4.37}
$$

By Theorem 4.1, we have

$$\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|<10N}\Lambda\|_{\mathcal{S}} \lesssim \varepsilon\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}H\|_{collapsing} + \varepsilon\||\partial_t|^{\frac{1}{4}}H\|_{collapsing} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}_r'} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}. \tag{4.38}$$

Case 3. $|\xi+\eta|<\frac{N}{10}$, $|\xi-\eta|\geq 10N$.

In this case, $|\xi|\sim|\eta|\sim|\xi-\eta|$, define $P_{>N}=P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 10N}$ The norm with respect to which the potential is a perturbation should be

$$\|\Lambda\|_{\mathcal{N}} = \||\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}\Lambda\|_{\mathcal{S}_{x,y}} \tag{4.39}$$

$$+\|||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \tag{4.40}$$

$$+\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}, \tag{4.41}$$

where $\widetilde{P}_{>N}=P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 9N}$.

In this case,

$$\begin{aligned}&\mathbf{S}\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}\Lambda\\ &\quad\sim N^2v(N(x-y))\langle\nabla_{x-y}\rangle\widetilde{P}_{>N}\Lambda\\ &\quad+\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}G(t,x,y)+N^2v(N(x-y)\langle\nabla_{x-y}\rangle\widetilde{P}_{>N}H.\end{aligned} \tag{4.42}$$

Using Strichartz for the second and third terms on the right side, Theorem 3.4 for the first term on the RHS, and Hölder's inequality, we get

$$\begin{aligned}&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}\Lambda\|_{\mathcal{S}_{x,y}}\\ &\lesssim\varepsilon\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad+\varepsilon\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}G\|_{\mathcal{S}_r'}\\ &\quad+\varepsilon\|\langle\nabla_{x-y}\rangle\widetilde{P}_{>N}H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\lesssim\varepsilon\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad+\varepsilon\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad+\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{>N}G\|_{\mathcal{S}_r'}\\ &\quad+\varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{4.43}$$

Thus, it suffice to control

$$\begin{aligned}&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad+\||\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &=I+II.\end{aligned} \tag{4.44}$$

For simplicity, we shall only give the details for the first term I, the second term is easier and can be handled in a similar way.

Note that

$$
\begin{aligned}
&\mathbf{S}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda \\
&\qquad\sim N^2v(N(x-y))\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda \\
&\qquad\quad + N^{\frac{5}{2}}v(N(x-y))P_{8N\leq|\xi-\eta|<9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda \\
&\qquad\quad + \langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}G(t,x,y) \\
&\qquad\quad + N^2v(N(x-y))\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 8N}P_{|\xi+\eta|<\frac{N}{10}}H,
\end{aligned}
\tag{4.45}
$$

where the bounds on $|\xi-\eta|$ changed slightly due to convolution with $\hat{v}_N$.

By Theorem 3.5 and Hölder's inequality, we have

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\lesssim \varepsilon\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&+ \||N^{\frac{5}{2}}v(N(x-y))P_{8N\leq|\xi-\eta|<9N}P_{|\xi+\eta|<\frac{N}{10}}|\partial_t|^{\frac{1}{4}}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&+ \|\langle\nabla_{x-y}\rangle P_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}G\|_{\mathcal{S}'_r} \\
&+ \varepsilon\|\langle\nabla_{x-y}\rangle\widetilde{P}_{|\xi-\eta|\geq 8N}P_{|\xi+\eta|<\frac{N}{10}}H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&+ \|\langle\nabla_{x-y}\rangle\widetilde{P}_{|\xi-\eta|\geq 9N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda_0\|_{L^2} \\
&\lesssim \varepsilon\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}\widetilde{P}_{>N}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&+ \varepsilon\||\partial_t|^{\frac{1}{4}}P_{|\xi-\eta|\leq 20N}\Lambda\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))} \\
&+ \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r} + \varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\qquad + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2},
\end{aligned}
\tag{4.46}
$$

where the first two terms on the right side of the first inequality corresponds to the first term on the right side of (3.16), and the last three terms on the right side of the first inequality corresponds to the remaining terms on the right side of (3.16).

5. **Estimates for the nonlinear equation.**

Recall the notation

$$\mathbf{S}_\pm = \frac{1}{i}\frac{\partial}{\partial t} - \Delta_x + \Delta_y$$

From now on, $V_N(x-y) = N^3v(N(x-y))$.

Define $\Gamma = \Gamma_c + \Gamma_p$, $\Lambda = \Lambda_c + \Lambda_p$, where $\Gamma_c = \bar{\phi}\otimes\phi$, $\Lambda_c = \phi\otimes\phi$, $\Gamma_p = \frac{1}{N}\overline{\mathrm{sh}(k)}\circ\mathrm{sh}(k)$, and $\Lambda_p = \frac{1}{2N}\mathrm{sh}(2k)$. Let $\rho(t,x) = \Gamma(t,x,x)$.

The four relevant equations are

$$\mathbf{S}\Lambda_p + \{V_N * \rho, \Lambda_p\} + \frac{V_N}{N}\Lambda_p + \big((V_N\bar\Gamma_p)\circ\Lambda_p + (V_N\Lambda_p)\circ\Gamma_p\big)_{symm} + \big((V_N\bar\Gamma_c)\circ\Lambda_p + (V_N\Lambda_c)\circ\Gamma_p\big)_{symm} = -\frac{V_N}{N}\Lambda_c \tag{5.1}$$

$$\mathbf{S}_\pm\Gamma_p + [V_N * \rho, \Gamma_p] + \big((V_N\Gamma_p)\circ\Gamma_p + (V_N\bar\Lambda_p)\circ\Lambda_p\big)_{skew} + \big((V_N\Gamma_c)\circ\Gamma_p + (V_N\bar\Lambda_c)\circ\Lambda_p\big)_{skew} = 0 \tag{5.2}$$

$$\mathbf{S}\Lambda_c + \{V_N * \rho, \Lambda_c\} + \big((V_N\bar\Gamma_p)\circ\Lambda_c + (V_N\Lambda_p)\circ\Gamma_c\big)_{symm} = 0 \tag{5.3}$$

$$\mathbf{S}_\pm\Gamma_c + [V_N * \rho, \Gamma_c] + \big((V_N\Gamma_p)\circ\Gamma_c + (V_N\bar\Lambda_p)\circ\Lambda_c\big)_{skew} = 0. \tag{5.4}$$

Here $\big(A(x,y)\big)_{symm} = A(x,y) + A(y,x)$, $\big(A(x,y)\big)_{skew} = A(x,y) - \bar A(y,x)$,

$$\{V_N * \rho, \Lambda\}(x,y) = \int dz\,(V_N(x-z) + V_N(y-z))\,\rho(z)\Lambda(x,y),$$

and

$$[V_N * \rho, \Gamma](x,y) = \int dz\,(V_N(x-z) - V_N(y-z))\,\rho(z)\Gamma(x,y).$$

The norm used for Λ_p is called $\mathcal{N}_1(\Lambda)$ and is

$$\begin{aligned}\|\Lambda\|_{\mathcal{N}_1(\Lambda)} &= \|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{\mathcal{S}_{x,y}} + \|\langle\nabla_{x+y}\rangle^{\frac12}\Lambda\|_{low\ collapsing}\\ &+ \||\partial_t|^{\frac14}\Lambda\|_{low\ collapsing}\\ &+ \big\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}P_{|\xi|\ge 10N}P_{|\eta|\ge 10N}P_{|\xi+\eta|\ge\frac{N}{10}}\Lambda\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &+ \big\|\langle\nabla_{x+y}\rangle^{\frac12}\langle\nabla_{x-y}\rangle^{\frac12}P_{|\xi-\eta|\ge 10N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &+ \big\||\partial_t|^{\frac14}\langle\nabla_{x-y}\rangle^{\frac12}P_{|\xi-\eta|\ge 10N}P_{|\xi+\eta|<\frac{N}{10}}\Lambda\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))},\end{aligned} \tag{5.5}$$

where the norm $\|\cdot\|_{low\ collapsing}$ is defined as in (2.9). The last three norms in (5.5) does not appear in the statement of Theorem 2.1, but as one can see from the proof of Theorem 2.1 in the previous section, they satisfy the same bounds as the first three norms on the right side of (5.5).

The norm used for Λ_c is called $\mathcal{N}_2(\Lambda)$ and is

$$\begin{aligned}\|\Lambda\|_{\mathcal{N}_2(\Lambda)} &= \|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\|_{\mathcal{S}_{x,y}} + \|\langle\nabla_{x+y}\rangle^{\frac12}\Lambda\|_{collapsing}\\ &+ \||\partial_t|^{\frac14}\Lambda\|_{collapsing} + \|\langle\nabla_x\rangle^{\frac12}\Lambda\|_{collapsing}\\ &+ \|\langle\nabla_y\rangle^{\frac12}\Lambda\|_{collapsing} + \big\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned} \tag{5.6}$$

We will use the following a priori estimates for $\Gamma(t,x,x)$ (proved in Lemma 6.2 in [10]).

Lemma 5.1. *Let the potential v satisfies* (1.26)*, and the initial conditions satisfy* (1.27)*, we have for all $0 \le \alpha \le 1$,*

$$\|\langle \nabla_{x+y}\rangle^{\alpha}\Gamma\|_{L^8(dt)L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \lesssim 1. \tag{5.7}$$

The above estimates also hold for Γ_p, Γ_c and Λ_c separately.

We need to use a continuity argument, we have to localize our estimates to intervals $[0,T]$, where the right end of the interval must be a variable T. Define $\Lambda_{c,T}, \Lambda_{p,T}, \Gamma_{p,T}, \Gamma_{c,T}$ to be solutions to the standard equations with the RHS multiplied by $\chi_{[0,T]}$:

$$\begin{aligned}
&\mathbf{S}\Lambda_{p,T} + \frac{V_N}{N}\Lambda_{p,T} \\
&\quad = \chi_{[0,T]}\Big(- \{V_N * \rho, \Lambda_p\} - \big((V_N\bar{\Gamma}_p)\circ\Lambda_p + (V_N\Lambda_p)\circ\Gamma_p\big)_{symm} \\
&\quad - \big((V_N\bar{\Gamma}_c)\circ\Lambda_p + (V_N\Lambda_c)\circ\Gamma_p\big)_{symm} - \frac{V_N}{N}\Lambda_c\Big)
\end{aligned} \tag{5.8}$$

$$\begin{aligned}
&\mathbf{S}_{\pm}\Gamma_{p,T} = \chi_{[0,T]}\Big(- [V_N * \rho, \Gamma_p] - \big((V_N\Gamma_p)\circ\Gamma_p + (V_N\bar{\Lambda}_p)\circ\Lambda_p\big)_{skew} \\
&\quad - \big((V_N\Gamma_c)\circ\Gamma_p + (V_N\bar{\Lambda}_c)\circ\Lambda_p\big)_{skew}\Big)
\end{aligned} \tag{5.9}$$

$$\mathbf{S}\Lambda_{c,T} = \chi_{[0,T]}\Big(- \{V_N * \rho, \Lambda_c\} - \big((V_N\bar{\Gamma}_p)\circ\Lambda_c + (V_N\Lambda_p)\circ\Gamma_c\big)_{symm}\Big) \tag{5.10}$$

$$\mathbf{S}_{\pm}\Gamma_{c,T} = \chi_{[0,T]}\Big(- [V_N * \rho, \Gamma_c] - \big((V_N\Gamma_p)\circ\Gamma_c + (V_N\bar{\Lambda}_p)\circ\Lambda_c\big)_{skew}\Big) \tag{5.11}$$

with $\Lambda_{c,T}(0,\cdot) = \Lambda_c(0,\cdot)$, and similarly for the other three functions. Also, we $\Lambda_{c,T} = \Lambda_c$ in $[0,T]$ (but not outside this interval), and similarly for the other three functions.

Theorem 5.2. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$\begin{aligned}
\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} &\le C\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_p(0,\cdot)\|_{L^2} + C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} \\
&\quad + C\varepsilon\|\Gamma_{p,T}\|_{\mathcal{S}_{x,y}} + C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}\|\Gamma_{p,T}\|_{\mathcal{S}_{x,y}} + C\varepsilon\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)}.
\end{aligned} \tag{5.12}$$

The proof is based on Theorem 2.1, there exists a constant C such that

$$
\begin{aligned}
&\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}\\
&\le C\Big(\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\chi_{[0,T]}\big(\{V_N*\rho,\Lambda_p\}+\big((V_N\bar\Gamma_p)\circ\Lambda_p+(V_N\Lambda_p)\circ\Gamma_p\big)_{symm}\\
&\quad+\big((V_N\bar\Gamma_c)\circ\Lambda_p+(V_N\Lambda_c)\circ\Gamma_p\big)_{symm}\big)\|_{\mathcal{S}'_r}\\
&\quad+\varepsilon\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\chi_{[0,T]}\Lambda_c\|_{L^2(dt)L^6(x-y)L^2(d(x+y))}\\
&\quad+\varepsilon\|\langle\nabla_{x+y}\rangle^{\frac12}\chi_{[0,T]}\Lambda_c\|_{collapsing}+\varepsilon\||\partial_t|^{\frac14}\chi_{[0,T]}\Lambda_c\|_{collapsing}\\
&\quad+\varepsilon\|\langle\nabla_x\rangle^{\frac12}\chi_{[0,T]}\Lambda_c\|_{collapsing}+\varepsilon\|\langle\nabla_y\rangle^{\frac12}\chi_{[0,T]}\Lambda_c\|_{collapsing}\\
&\quad+C\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda_p(0,\cdot)\|_{L^2}\Big).
\end{aligned}
\tag{5.13}
$$

For all terms other than $\||\partial_t|^{\frac14}\chi_{[0,T]}\Lambda_c\|_{collapsing}$, the subscript T can be trivially added to Λ, Γ on the RHS. And we also have

$$
\begin{aligned}
&\||\partial_t|^{\frac14}\chi_{[0,T]}\Lambda_c\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&=\||\partial_t|^{\frac14}\chi_{[0,T]}\Lambda_{c,T}\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\lesssim\||\partial_t|^{\frac14}\Lambda_{c,T}\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&\lesssim\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)},
\end{aligned}
\tag{5.14}
$$

where in the third line we used the fact that

$$
\||\partial_t|^{\frac14}\chi_{[0,T]}F\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\lesssim\||\partial_t|^{\frac14}F\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))},
$$

for any interval $[0,T]$. As remarked in [10], this can be shown by using the equivalent definition

$$
\||\partial_t|^{\frac14}u\|_{L^2}=\int\int\frac{|u(t)-u(s)|^2}{|t-s|^{1+\frac12}}dtds,
\tag{5.15}
$$

and the generalized Hardy's inequality from [37].

In the lemma that follow, we estimate the norm of the nonlinear terms in suitable dual Strichartz norms, using the bound (5.1) whenever possible.

Lemma 5.3. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$
\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Big(\{V_N*\rho,\Lambda_{p,T}\}+(V_N\bar\Gamma)\circ\Lambda_{p,T}\Big)\|_{L^{\frac85}([0,T])L^{\frac43}(dx)L^2(dy)}.\\
&\le C\varepsilon\|\langle\nabla_x\rangle^{\frac12}\langle\nabla_y\rangle^{\frac12}\Lambda_{p,T}\|_{\mathcal{S}_{x,y}}.
\end{aligned}
\tag{5.16}
$$

Here Γ can be Γ_p or Γ_c. The result depends on the a priori bounds for Γ, but is true with $\Lambda_{p,T}$ replaced with any other function.

Proof. In this case, we essentially view V_N as a δ distribution, by Minkowski integral inequality, and it suffices to show that

$$\begin{aligned}(5.17)\quad &\sup_z \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big(\Gamma(t,x,x+z)\Lambda_{p,T}(t,x+z,y)\Big)\|_{L^{\frac{8}{5}}([0,T])L^{\frac{4}{3}}(dx)L^2(dy)}\\ &\le C\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_{p,T}\|_{\mathcal{S}_{x,y}},\end{aligned}$$

where the extra ε factor in (5.3) can be remedied by the smallness of $\|V_N\|_{L^1}$.

Using the fractional Leiniz rule from Theorem 5.1 in [10], we have the following estimate, uniformly in z:

$$\begin{aligned}&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big(\Gamma(t,x,x+z)\Lambda_{p,T}(t,x+z,y)\Big)\|_{L^{\frac{8}{5}}([0,T])L^{\frac{4}{3}}(dx)L^2(dy)}\\ &\le C\|\langle\nabla_x\rangle^{\frac{1}{2}}\Gamma(t,x,x+z)\|_{L^8(dt)L^{\frac{4}{3}+}(dx)}\|\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_{p,T}\|_{L^2(dt)L^{\infty-}(dx)L^2(dy)}\\ (5.18)\quad &+C\|\Gamma(t,x,x+z)\|_{L^8(dt)L^{\frac{12}{7}}(dx)}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_{p,T}\|_{L^2(dt)L^6(dx)L^2(dy)}\\ &\le C\|\langle\nabla_x\rangle^{\alpha}\Gamma(t,x,x+z)\|_{L^8(dt)L^{\frac{4}{3}}(dx)}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_{p,T}\|_{L^2(dt)L^6(dx)L^2(dy)}\\ &\le C\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x+z,y)\|_{L^2(dt)L^6(dx)L^2(dy)}.\end{aligned}$$

Here α is can be any number in $(\frac{1}{2},1]$, $\frac{4}{3}+$ is a number that is bigger than but can be arbitrary close to $\frac{4}{3}$, similarly $\infty-$ is any finite number but can be arbitrary large. We use $\infty-$ since we do not have the sharp Sobolev estimate from L^6 to L^∞. In the last inequality we used Lemma 5.1. □

Since Λ_c satisfies the same a priori estimates as Γ, by the exact same argument we get

Lemma 5.4. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$\begin{aligned}(5.19)\quad &\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big((V_N\Lambda_c)\circ\Gamma_{p,T}\Big)\|_{L^{\frac{4}{3}}([0,T])L^{\frac{3}{2}}(dx)L^2(dy)}\\ &\le C\varepsilon\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_{p,T}\|_{\mathcal{S}_{x,y}}.\end{aligned}$$

The result depends on the a priori bounds for Λ_c, but is true with $\Gamma_{p,T}$ replaced with any other function.

We continue estimating nonlinear terms.

Lemma 5.5. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$\begin{aligned}(5.20)\quad &\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big((V_N\Lambda_{p,T})\circ\Gamma_{p,T}\Big)\|_{L^{\frac{4}{3}}([0,T])L^{\frac{3}{2}}(dx)L^2(dy)}\\ &\le C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_{p,T}\|_{\mathcal{S}_{x,y}}.\end{aligned}$$

The result is still true if we replace $\Gamma_{p,T}$ with any other function.

Proof. In this case, we shall not treat V_N as a δ distribution. Recall that

$$\begin{aligned}
&\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big((V_N\Lambda_{p,T})\circ\Gamma_{p,T}\Big)\\
&=\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\int N^3v(N(x-z))\Lambda_{p,T}(x,z)\Gamma_{p,T}(z,y)dz\\
&=\int N^3v(Nz)\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big(\Lambda_{p,T}(x,x+z)\Gamma_{p,T}(x+z,y)\Big)dz.
\end{aligned}\tag{5.21}$$

For fixed z, the following holds, uniformly in z

$$\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Big(\Lambda_{p,T}(t,x,x+z)\Gamma_{p,T}(t,x+z,y)\Big)\|_{L^{\frac{4}{3}}([0,T])L^{\frac{3}{2}}(dx)L^2(dy)}\\
&\leq C\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_{p,T}(t,x,y)\|_{L^4(dt)L^6(dx)L^2(dy)}\\
&+C\|\Lambda_{p,T}(t,x,x+z)\|_{L^2(dt)L^3(dx)}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_{p,T}(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}\\
&\leq C\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_{p,T}(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}.
\end{aligned}$$

Thus, by Minkowski integral inequality it suffices to show that

$$\int|N^3v(Nz)|\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x,x+z)\|_{L^2(dt)L^2(dx)}dz\leq C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)},\tag{5.22}$$

which is equivalent to

$$\begin{aligned}
&\int|N^3v(N(x-y))|\big\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x,y)\big\|_{L^2(dt)L^2(d(x+y))}d(x-y)\\
&\leq C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}.
\end{aligned}\tag{5.23}$$

To see this, note that if for $\Lambda_{p,T}$, we have $|\xi-\eta|<20N$, $|\xi|<20N$ or $|\eta|<20N$, then the left side of (5.23) is easily controlled by

$$\begin{aligned}
&\int|N^3v(N(x-y))|\big\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda_{p,T}(t,x,y)\big\|_{L^2(dt)L^2(d(x+y))}d(x-y)\\
&\leq C\varepsilon\|\Lambda_{p,T}\|_{low\ collapsing}.
\end{aligned}\tag{5.24}$$

Thus if we denote $P_{>N}=P_{|\xi-\eta|\geq20N}P_{|\xi|\geq20N}P_{|\eta|\geq20N}$, we are further reduced to showing that

$$\begin{aligned}
&\int|N^3v(N(x-y))|\big\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}P_{>N}\Lambda_{p,T}(t,x,y)\big\|_{L^2(dt)L^2(d(x+y))}d(x-y)\\
&\leq C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}.
\end{aligned}\tag{5.25}$$

As in the proof of main theorem for the linear equation, we shall divide our discussion into two cases.

Case 1: $|\xi+\eta|<\frac{N}{10}$.

In this case, by Hölder's inequality,

$$\begin{aligned}
&\int |N^3 v(N(x-y))| \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} P_{>N} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^2(d(x+y)) d(x-y)} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^6(d(x-y)) L^2(dt) L^2(d(x+y))} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&\leq C\varepsilon \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} \langle \nabla_{x-y} \rangle^{\frac{1}{2}} P_{|\xi-\eta|\geq 10N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&\leq C\varepsilon \|\Lambda_p^T\|_{\mathcal{N}_1(\Lambda)},
\end{aligned}$$

where in the third inequality we used Bernstein's inequality along with the fact that $|\xi - \eta| \geq 20N$.

Case 2: $|\xi + \eta| \geq \frac{N}{10}$.

In this case, by Hölder's inequality,

$$\begin{aligned}
&\int |N^3 v(N(x-y))| \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} P_{>N} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^2(d(x+y)) d(x-y)} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|\geq\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^6(d(x-y)) L^2(dt) L^2(d(x+y))} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\| \langle \nabla_{x} \rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|\geq\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&\quad + C\varepsilon N^{\frac{1}{2}} \big\| \langle \nabla_{y} \rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|\geq\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&\leq C\varepsilon \big\| \langle \nabla_{x} \rangle^{\frac{1}{2}} \langle \nabla_{y} \rangle^{\frac{1}{2}} P_{|\xi|\geq 10N} P_{|\eta|\geq 10N} P_{|\xi+\eta|\geq\frac{N}{10}} \Lambda_{p,T}(t,x,y) \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&\leq C\varepsilon \|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}
\end{aligned}$$

where in the third inequality we used Bernstein's inequality in rotated coordinates along with the fact that $|\xi| \geq 20N$ and $|\eta| \geq 20N$.

□

We continue with estimates for $\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)}$. This is an easy version of the previous theorem. Using Lemma 3.7-Lemma 3.9 and Strichartz estimates (Theorem 3.1), and then applying Lemma 5.3-5.5 in this section to handle the nonlinear terms, we get

Theorem 5.6. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$\begin{aligned}
\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)} \leq C \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_c(0,\cdot)\|_{L^2} + C\varepsilon \|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)} \\
+ C\varepsilon \|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} \|\Gamma_{c,T}\|_{\mathcal{S}_{x,y}}.
\end{aligned} \tag{5.26}$$

Using Strichartz estimates for $\mathbf{S}_\pm$ and Lemma 5.3-5.5, we get

Theorem 5.7. *Let $[0,T]$ be as above, there exist a universal constant C such that*

$$\begin{aligned}
\|\Gamma_{c,T}\|_{\mathcal{S}_{x,y}} \leq C \|\langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Gamma_c(0,\cdot)\|_{L^2} + C\varepsilon \|\Gamma_{c,T}\|_{\mathcal{S}_{x,y}} \\
+ C\varepsilon \|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} \|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)}
\end{aligned} \tag{5.27}$$

$$\begin{aligned} \|\Gamma_{p,T}\|_{\mathcal{S}_{x,y}} \leq C\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Gamma_p(0,\cdot)\|_{L^2} + C\varepsilon\|\Gamma_{p,T}\|_{\mathcal{S}_{x,y}} \\ + C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} + C\varepsilon\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)}. \end{aligned} \tag{5.28}$$

For later use, let us denote

$$\begin{aligned} X(T) =&\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)} + \|\Gamma_{c,T}\|_{\mathcal{S}_{x,y}} \\ Y(T) =&\|\Lambda_{p,T}\|_{\mathcal{N}_1(\Lambda)} + \|\Gamma_{p,T}\|_{\mathcal{S}_{x,y}}. \end{aligned} \tag{5.29}$$

We want to show that $X(T)$, $Y(T)$ depends continuously on T. To see this, for any fixed $T \geq 0$, we have

$$\Big|\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)} - \|\Lambda_{c,T+\delta}\|_{\mathcal{N}_2(\Lambda)}\Big| \lesssim \|\Lambda_{c,T+\delta} - \Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)}.$$

And note that $\Lambda_{c,T+\delta} - \Lambda_{c,T}$ satisfy

$$\mathbf{S}(\Lambda_{c,T+\delta} - \Lambda_{c,T}) = \chi_{[T,T+\delta]}\Big(- \{V_N * \rho, \Lambda_c\} - \big((V_N\bar{\Gamma}_p)\circ\Lambda_c + (V_N\Lambda_p)\circ\Gamma_c\big)_{symm}\Big),$$

with 0 initial condition. It is not hard to see that, by crude energy estimates

$$\|\Lambda_{c,T+\delta} - \Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)} \leq \delta\, C(T,N) \tag{5.30}$$

for some constant $C(T,N)$ that depend on T, N, which implies the continuity of $\|\Lambda_{c,T}\|_{\mathcal{N}_2(\Lambda)}$. The continuity of other norms in $X(T)$ and $Y(T)$ can be proved similarly.

We can now state and prove the main theorem of this section

Theorem 5.8. *Assume Λ, Γ and ϕ are smooth solutions to the HFB system, with finite energy per particle, uniformly in N (see* (1.27)*), which implies* (1.31)*-*(1.34)*, we have*

$$\begin{aligned} \|\Lambda_c\|_{\mathcal{N}_2(\Lambda)} + \|\Gamma_c\|_{\mathcal{S}_{x,y}} \leq C \\ \|\Lambda_p\|_{\mathcal{N}_1(\Lambda)} + \|\Gamma_p\|_{\mathcal{S}_{x,y}} \leq C. \end{aligned} \tag{5.31}$$

Proof. By Theorem 5.2, 5.6, 5.7, and the size of initial conditions (1.31)-(1.34), we have

$$\begin{aligned} X(T) \leq& CC_0 + C\varepsilon X(T)Y(T) \\ Y(T) \leq& CC_0 + C\varepsilon Y(T)^2 + C\varepsilon X(T). \end{aligned} \tag{5.32}$$

At this stage, we will need to assume ε is small, where the smallness may depend on C_0. Without loss of generality, let's assume $CC_0 = 1$ and $C\varepsilon \leq \frac{1}{10}$, then we have the following simplified version of (5.32)

$$\begin{aligned} X(T) \leq& 1 + \frac{1}{10}X(T)Y(T) \\ Y(T) \leq& 1 + \frac{1}{10}Y(T)^2 + \frac{1}{10}X(T). \end{aligned} \tag{5.33}$$

If $Y(T) \leq 4$, by the first line, we have $X(T) \leq 2$, and if we plug this into the second line, we get $Y(T) \leq 3$. By continuity, since $Y(0) \leq 1$, we always have $Y(T) \leq 3$, and thus $X(T) \leq 2$, for all $T \geq 0$, which concludes our proof. $\square$

6. **Estimates for** $\mathrm{sh}(2k)$**,** $p_2 = \overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)$ **and** $\mathrm{sh}(k)$**.**

The equations for $\mathrm{sh}(2k) = N\Lambda_p$ and $p_2 = N\Gamma_p$ are

$$\mathbf{S}\,\mathrm{sh}(2k) + \{V_N * \rho, \mathrm{sh}(2k)\} + \big((V_N\bar{\Gamma}) \circ \mathrm{sh}(2k) + (V_N\Lambda) \circ p_2\big)_{symm} = -\frac{V_N}{2}\Lambda$$
$$\mathbf{S}_{\pm}\, p_2 + [V_N * \rho, p_2] + \big((V_N\Gamma) \circ p_2 + (V_N\bar{\Lambda}) \circ \mathrm{sh}(2k)\big)_{skew} = 0.$$

To handle the inhomogeneous term $-\frac{V_N}{2}\Lambda$, we shall need the following lemma.

Lemma 6.1. *Let* $\mathbf{S}u = -\frac{V_N}{2}\Lambda_p$ *with* $u(0,\cdot) = 0$*, we have*

$$\|u\|_{\mathcal{S}_{x,y}} \lesssim \|\Lambda_p\|_{\mathcal{N}_1}$$

where $\|\cdot\|_{\mathcal{N}_1}$ *is defined as in* (5.5).

Note that the above result also hold if one replace Λ_p by Λ_c and replace $\|\Lambda_p\|_{\mathcal{N}_1}$ by $\|\Lambda_c\|_{\mathcal{N}_2}$, which is a direct consequence of Theorem 3.2 and Hölder's inequality.

Proof. First note that if for Λ, we have $|\xi - \eta| < 20N$, $|\xi| < 20N$ or $|\eta| < 20N$, then by Theorem 3.2 and Hölder's inequality, we have

$$\|u\|_{\mathcal{S}_{x,y}} \leq C\varepsilon \|\Lambda_p\|_{low\ collapsing}.$$

Thus if we denote $P_{>N} = P_{|\xi-\eta|\geq 20N} P_{|\xi|\geq 20N} P_{|\eta|\geq 20N}$, if suffices to prove Lemma 6.1 with Λ replaced by $P_{>N}\Lambda$. As in the proof of main theorem for the linear equation, we shall divide our discussion into two cases.

Case 1: $|\xi + \eta| < \frac{N}{10}$.

In this case, by Theorem 3.2 and Hölder's inequality, we have

$$\begin{aligned}
&\|u\|_{\mathcal{S}_{x,y}} \\
&\leq C\big\|V_N \langle \nabla_{x+y}\rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\quad + C\big\|V_N |\partial_t|^{\frac{1}{4}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^1(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\|\langle \nabla_{x+y}\rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^6(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\quad + C\varepsilon N^{\frac{1}{2}} \big\||\partial_t|^{\frac{1}{4}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^6(d(x-y))L^2(dt)L^2(d(x+y))} \\
&\leq C\varepsilon N^{\frac{1}{2}} \big\|\langle \nabla_{x+y}\rangle^{\frac{1}{2}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\quad + C\varepsilon N^{\frac{1}{2}} \big\||\partial_t|^{\frac{1}{4}} P_{>N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\leq C\varepsilon \big\||\partial_t|^{\frac{1}{4}} \langle \nabla_{x-y}\rangle^{\frac{1}{2}} P_{|\xi-\eta|\geq 10N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\quad + C\varepsilon \big\|\langle \nabla_{x+y}\rangle^{\frac{1}{2}} \langle \nabla_{x-y}\rangle^{\frac{1}{2}} P_{|\xi-\eta|\geq 10N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\leq C\varepsilon \|\Lambda_p\|_{\mathcal{N}_1(\Lambda)}
\end{aligned} \tag{6.1}$$

where in the fourth inequality we used Bernstein's inequality along with the fact that $|\xi - \eta| \geq 20N$.

Case 2: $|\xi+\eta| \geq \frac{N}{10}$.

In this case, by Strichartz (Theorem 3.1) and Hölder's inequality,

$$
\begin{aligned}
&\|u\|_{\mathcal{S}_{x,y}} \\
&\leq C\big\|V_N P_{>N} P_{|\xi+\eta|>\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} \\
&\leq C\varepsilon N\big\|P_{>N} P_{|\xi+\eta|\geq\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\leq C\varepsilon\big\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}} P_{|\xi|\geq 10N} P_{|\eta|\geq 10N} P_{|\xi+\eta|<\frac{N}{10}} \Lambda_p\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \\
&\leq C\varepsilon\|\Lambda_p\|_{\mathcal{N}_1(\Lambda)}
\end{aligned}
\tag{6.2}
$$

where in the third inequality we used Bernstein's inequality in rotated coordinates along with the fact that $|\xi| \geq 20N$ and $|\eta| \geq 20N$. □

Now we shall estimate the other nonlinear terms in dual Strichartz norms

$$
\begin{aligned}
&\|\big(V_N * \rho(t,x)\big)\mathrm{sh}(2k)(t,x,y)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)} \\
&\qquad + \|\big((V_N\bar{\Gamma}) \circ \mathrm{sh}(2k)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)} \\
&\leq C\varepsilon \sup_z \|\Gamma(t,x+z,x)\|_{L^8(dt)L^{\frac{12}{7}}(dx)} \|\mathrm{sh}(2k)\|_{L^2(dt)L^6(dx)L^2(dy)} \\
&\leq C\varepsilon\|\mathrm{sh}(2k)\|_{L^2(dt)L^6(dx)L^2(dy)},
\end{aligned}
\tag{6.3}
$$

where we used Sobolev and Lemma 5.1 in the last inequality. The above estimate still hold if we replace $\bar{\Gamma}$ by $\bar{\Lambda}_c$. Similarly

$$
\begin{aligned}
&\|\big(V_N * \rho(t,x)\big)p_2(t,x,y)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)} \\
&\qquad + \|\big((V_N\Gamma) \circ p_2\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)} \\
&\leq C\varepsilon\|p_2\|_{L^2(dt)L^6(dx)L^2(dy)}.
\end{aligned}
\tag{6.4}
$$

And the above estimate still hold if we replace Γ by Λ_c. Also, since for fixed z, the following holds, uniformly in z

$$
\begin{aligned}
&\|\Lambda(t,x,x+z)p_2(t,x+z,y)\|_{L^{\frac{4}{3}}(dt)L^{\frac{3}{2}}(dx)L^2(dy)} \\
&\leq \|\Lambda(t,x,x+z)\|_{L^2(dt)L^3(dx)}\|p_2(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)} \\
&\leq C\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|p_2(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}.
\end{aligned}
\tag{6.5}
$$

If we repeat the argument in the proof of Lemma 5.5 and using (5.23), we have

$$
\begin{aligned}
&\|\big(V_N\Lambda_p\big) \circ p_2\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)} \\
&\leq C\int |N^3 v(Nz)|\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_p(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|p_2\|_{L^4(dt)L^3(dx)L^2(dy)}dz \\
&\leq C\varepsilon\|\Lambda\|_{\mathcal{N}_1(\Lambda)}\|p_2\|_{L^4(dt)L^3(dx)L^2(dy)}.
\end{aligned}
$$

Similarly, we also have

$$\begin{aligned}
&\|\big(V_N\bar{\Lambda}_p\big)\circ \mathrm{sh}(2k)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)}\\
&\leq C\int |N^3 v(Nz)|\|\langle\nabla_x\rangle^{\frac{1}{2}}\Lambda_p(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|\mathrm{sh}(2k)\|_{L^4(dt)L^3(dx)L^2(dy)}dz\\
&\leq C\varepsilon\|\Lambda\|_{\mathcal{N}_1(\Lambda)}\|\mathrm{sh}(2k)\|_{L^4(dt)L^3(dx)L^2(dy)}.
\end{aligned}$$

If we choose ε small enough such that $\varepsilon C\leq \frac{1}{10}$, combing the above estimates, and using the fact that $\|\Lambda_p\|_{\mathcal{N}_1(\Lambda)}+\|\Lambda_c\|_{\mathcal{N}_2(\Lambda)}\lesssim 1$ from the main theorem of the last section, we get, by Strichartz

$$\begin{aligned}
&\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}}+\|p_2\|_{\mathcal{S}_{x,y}}\\
&\leq C\Big(\|\mathrm{sh}(2k)(0,\cdot)\|_{L^2}+\|p_2(0,\cdot)\|_{L^2}\Big)+C.
\end{aligned}\tag{6.6}$$

7. Estimates for the condensate ϕ.

The non-linear equation for ϕ can be regarded as a linear equation on a background given by Γ and Λ, for which we already have estimates:

$$\begin{aligned}
&\Big\{\frac{1}{i}\partial_t-\Delta_{x_1}\Big\}\phi(x_1)\\
&=-\int dy\{V_N(x_1-y)\Gamma(y,y)\}\phi(x_1)\\
&\quad-\int dy\{V_N(x_1-y)\Gamma_p(y,x_1)\}\phi(y)\\
&\quad+\int dy\{V_N(x_1-y)\Lambda_p(x_1,y)\}\bar{\phi}(y).
\end{aligned}\tag{7.1}$$

Define the standard Strichartz spaces

$$\|\phi\|_{\mathcal{S}}=\sup_{(p,q)\ admissible}\|\phi\|_{L^p(dt)L^q(dx)}.$$

Proof. We shall estimate the right hand side of the equation for ϕ in dual Strichartz norms, if we repeat the proof of Lemma 5.3-5.5, it is not hard to show

(7.2)

$$\begin{aligned}
&\|\langle\nabla\rangle^{\frac{1}{2}}\int dy\{V_N(x_1-y)\Gamma(y,y)\}\phi(x_1)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)}\leq C\varepsilon\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{L^2(dt)L^6(dx)}\\
&\|\langle\nabla\rangle^{\frac{1}{2}}\int dy\{V_N(x_1-y)\Gamma_p(y,x_1)\}\phi(y)\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)}\leq C\varepsilon\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{L^2(dt)L^6(dx)}\\
&\|\langle\nabla\rangle^{\frac{1}{2}}\int dy\{V_N(x_1-y)\Lambda_p(x_1,y)\}\bar{\phi}(y)\|_{L^{\frac{4}{3}}(dt)L^{\frac{3}{2}}(dx)}\leq C\varepsilon\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{L^4(dt)L^3(dx)}.
\end{aligned}$$

Thus,

$$\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{\mathcal{S}}\leq C\|\langle\nabla\rangle^{\frac{1}{2}}\phi(0,\cdot)\|_{L^2}+3C\varepsilon\|\langle\nabla\rangle^{\frac{1}{2}}\phi\|_{\mathcal{S}}.$$

This gives us desired result by taking $\varepsilon C\leq\frac{1}{10}$. □

8. **Remarks on the collapsing norm of Λ.**

In this section, we shall see how we can use Theorem 2.1 to prove the following

Theorem 8.1. *Let Λ satisfy*

$$\mathbf{S}\Lambda(t,x,y) + N^2 v\big(N(x-y)\big)\Lambda(t,x,y) = G(t,x,y),\ \Lambda(0,\cdot) = \Lambda_0 \tag{8.1}$$

where we assume v satisfy (1.25)*, we have*

$$\begin{aligned} &\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{collapsing} + \||\partial_t|^{\frac{1}{4}}\Lambda\|_{collapsing} \\ &\qquad \lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}_r'} + \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\Lambda_0\|_{L^2}. \end{aligned} \tag{8.2}$$

Recall that we already have desired estimates if we replace $\|\cdot\|_{collapsing}$ by $\|\cdot\|_{low\ collapsing}$, by using the results in Theorem 2.1.

Proof. We shall first treat the homogeneous equation, let

$$\begin{aligned} &\mathbf{S}\Lambda(t,x,y) + N^2 v\big(N(x-y)\big)\Lambda(t,x,y) = 0 \\ &\Lambda(0,\cdot) = \Lambda_0. \end{aligned} \tag{8.3}$$

Let $H = -\Delta_x - \Delta_y + N^2 v(N(x-y))$, $H_0 = -\Delta_x - \Delta_y$. If we let W denote the wave operator in Yajima's paper [36], acting in $x-y$ direction, we have $e^{itH} = We^{itH_0}W^*$, where * denotes the dual operator. We also have, for the potential v satisfying (1.25), W is a bounded operator from $L^p \to L^p$ for any $1 \le p \le \infty$, with bound independent of N (see [15] Proposition 5.1). Moreover, by calculating the integral kernel of W explicitly, the $L^p \to L^p$ boundness of W extends to the space of L^2 valued function,). Thus,

$$\begin{aligned} &\sup_{x-y}\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}e^{-itH}\Lambda_0\|_{L^2(dtd(x+y))} + \sup_{x-y}\|||\frac{\partial}{\partial t}|^{1/4}e^{-itH}\Lambda_0\|_{L^2(dtd(x+y))} \\ &= \sup_{x-y}\|W\langle\nabla_{x+y}\rangle^{\frac{1}{2}}e^{itH_0}W^*\Lambda_0\|_{L^2(dtd(x+y))} + \sup_{x-y}\|W|\frac{\partial}{\partial t}|^{1/4}e^{itH_0}W^*\Lambda_0\|_{L^2(dtd(x+y))} \\ &\lesssim \sup_{x-y}\|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}e^{itH_0}W^*\Lambda_0\|_{L^2(dtd(x+y))} + \sup_{x-y}\|||\frac{\partial}{\partial t}|^{1/4}e^{itH_0}W^*\Lambda_0\|_{L^2(dtd(x+y))} \\ &\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}W^*\Lambda_0\|_{L^2(dxdy)}, \end{aligned}$$

where we used Lemma 3.8-3.9 in the last inequality.

Recall $W^* = \lim_{t\to\infty} e^{itH_0}e^{-itH}$, where the limit exists in the strong operator topology, see e.g., [15, Proposition 5.1] and also [17] for a proof of the existence of strong limit when v satisfy (1.25). Thus, for each fixed $g \in C_0^\infty(\mathbb{R}^6)$ with $\|g\|_2 = 1$

$$
\begin{aligned}
&\langle \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} W^* \Lambda_0, g \rangle \\
&= \langle W^* \Lambda_0, \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} g \rangle \\
&= \lim_{t \to \infty} \langle e^{itH_0} e^{-itH} \Lambda_0, \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} g \rangle \\
&= \lim_{t \to \infty} \langle e^{-itH} \Lambda_0, \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} e^{-itH_0} g \rangle \\
&= \lim_{t \to \infty} \langle \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} e^{-itH} \Lambda_0, e^{-itH_0} g \rangle \\
&\leq \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} e^{-itH} \Lambda_0 \|_{L^\infty(dt) L^2(dxdy)} \\
&\leq \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} e^{-itH} \Lambda_0 \|_{\mathcal{S}_{x,y}} \\
&\leq \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_0 \|_{L^2}
\end{aligned} \tag{8.4}
$$

where the second equality is a consequence of the existence of strong limit, and in the last inequality we used the special case of Theorem 2.1 with $G = H = 0$. By taking supremum among all choices of g, we have

$$
\| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} W^* \Lambda_0 \|_{L^2(dxdy)} \lesssim \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} \Lambda_0 \|_{L^2}.
$$

Now we shall consider the inhomogeneous equation, let

$$
\mathbf{S}\Lambda(t,x,y) + N^2 v\big(N(x-y)\big)\Lambda(t,x,y) = G(t,x,y), \ \Lambda(0,\cdot) = 0. \tag{8.5}
$$

If we let H be defined as above, we have by Duhamel's formula

$$
\Lambda = \int_0^t e^{-i(t-s)H} G(s,\cdot) ds.
$$

We shall first show that

$$
\begin{aligned}
&\sup_{x-y} \| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} \int_0^t e^{-i(t-s)H} G(s,\cdot) ds \|_{L^2(dtd(x+y))} \\
&\lesssim \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \|_{\mathcal{S}'_r}.
\end{aligned} \tag{8.6}
$$

To prove (8.6), it suffices to show that, for any fixed $x - y$ and fixed $T > 0$,

$$
\begin{aligned}
&\| \chi_{[0,T]}(t) \langle \nabla_{x+y} \rangle^{\frac{1}{2}} \int_0^t e^{-i(t-s)H} G(s,\cdot) ds \|_{L^2(dtd(x+y))} \\
&\lesssim \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \|_{\mathcal{S}'_r}
\end{aligned}
$$

which, by the Christ–Kiselev lemma, is a consequence of

$$
\begin{aligned}
&\| \langle \nabla_{x+y} \rangle^{\frac{1}{2}} \int_0^T e^{-i(t-s)H} G(s,\cdot) ds \|_{L^2(dtd(x+y))} \\
&\lesssim \| \langle \nabla_x \rangle^{\frac{1}{2}} \langle \nabla_y \rangle^{\frac{1}{2}} G \|_{\mathcal{S}'_r}.
\end{aligned} \tag{8.7}
$$

If we apply the homogeneous estimates proved above, the left side of (8.7) is bounded by

$$\begin{aligned}
&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\int_0^T e^{isH}G(s,\cdot)ds\|_{L^2(dxdy)}\\
&=\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}e^{iTH}\int_0^T e^{-i(T-s)H}G(s,\cdot)ds\|_{L^2(dxdy)}\\
&\lesssim\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}\int_0^T e^{-i(T-s)H}G(s,\cdot)ds\|_{L^2(dxdy)}\\
&\lesssim\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r},
\end{aligned}\tag{8.8}$$

where in the third line we used the special case of Theorem 2.1 with $G = H = 0$ (estimates for the homogeneous equation), and in the last line we used the special case of Theorem 2.1 with $\Lambda_0 = H = 0$(inhomogeneous equation with zero initial data).

To prove inhomogeneous estimate for $|\partial_t|^{\frac{1}{4}}$ derivative, as before we can not use the Christ–Kiselev lemma, we shall follow the ideas in the prove of Lemma 3.9. Write $\Lambda = \Lambda_1 + \Lambda_2$, where

$$\begin{aligned}
&\mathbf{S}\,\Lambda_1 = G(t,x,y), \quad \text{with initial conditions } 0\\
&\mathbf{S}\,\Lambda_2 = -N^2v\big(N(x-y)\big)\Lambda(t,x,y), \quad \text{with initial conditions } 0.
\end{aligned}$$

For both Λ_1 and Λ_2, it suffices to consider the region where

$$\tau^{\frac{1}{2}} \geq 10(1+|\xi+\eta|), |\xi| \geq 20N,\ |\eta| \geq 20N \text{ and } |\xi-\eta| \geq 20N,\tag{8.9}$$

since otherwise we have $\||\partial_t|^{\frac{1}{4}}\Lambda\|_{collapsing} \lesssim \|\langle\nabla_{x+y}\rangle^{\frac{1}{2}}\Lambda\|_{collapsing}$, or we already know how to control collapsing norm when $|\xi| < 20N$, $|\eta| < 20N$, or $|\xi-\eta| < 20N$.

The estimate for Λ_1 just follows from Lemma 3.9 directly, and for Λ_2, we claim that it suffices to show

Proposition 8.2. *Let* $\mathbf{S}u = f$ *with* $u(0,\cdot) = 0$*, then if the Fourier support* τ, ξ, η *of* u *satisfy* (8.9)*, we have*

$$\begin{aligned}
\||\partial_t|^{\frac{1}{4}}u\|_{collapsing} \lesssim \min\Big\{&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))},\\
&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + N^{-\frac{1}{2}}\||\partial_t|^{\frac{1}{4}}f\|_{L^2(dtdxdy)}\Big\}.
\end{aligned}\tag{8.10}$$

Remark 8.3. Note that the τ support of u may be different from f, so τ support of f may not satisfy (8.9) , but the ξ, η support of f does satisfy (8.9).

We shall first see how we can apply the Proposition to get desired results.

Case 1: $|\xi+\eta| \geq \frac{N}{10}$.

In this case,

$$\mathbf{S}\,P_{|\xi|\geq 20N}P_{|\eta|\geq 20N}\Lambda_2 \sim -N^2v\big(N(x-y)\big)P_{|\xi|\geq 19N}P_{|\eta|\geq 19N}\Lambda(t,x,y).$$

Thus, by Proposition 8.2

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}}\Lambda_2\|_{collapsing}\\
\lesssim&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}N^2v\big(N(x-y)\big)P_{|\xi|\geq 19N}P_{|\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\
\lesssim&\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\xi|\geq 19N}P_{|\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))},
\end{aligned}
$$

where in the last inequality we used the fact that when $|\xi|\geq 19N$, $|\eta|\geq 19N$, $\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}$ essentially only fall on Λ, as well as Hölder's inequality. Recall that when $|\xi+\eta|>\frac{N}{10}$, the last line appears as part of the norm for Λ in the proof of Theorem 2.1 (see (4.15)), thus it is bounded by $\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r}$ for Λ satisfying (8.5).

Case 2: $|\xi+\eta|<\frac{N}{10}$.

In this case,

$$
\mathbf{S}\,P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 20N}\Lambda_2\sim -N^2v\big(N(x-y)\big)P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y).
$$

Thus, by Proposition 8.2

$$
\begin{aligned}
&\||\partial_t|^{\frac{1}{4}}\Lambda_2\|_{collapsing}\\
\lesssim&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}N^2v\big(N(x-y)\big)P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\
&+\||\partial_t|^{\frac{1}{4}}N^{\frac{3}{2}}v\big(N(x-y)\big)P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dtdxdy)}\\
\lesssim&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+N^{\frac{1}{2}}\||\partial_t|^{\frac{1}{4}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
\lesssim&\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi+\eta|<\frac{N}{10}}P_{|\xi-\eta|\geq 19N}\Lambda(t,x,y)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}
\end{aligned}
$$

where in the second inequality we used the fact that when $|\xi-\eta|\geq 19N$, $\langle\nabla_{x-y}\rangle^{\frac{1}{2}}$ essentially only fall on Λ, as well as Hölder's inequality. And in the last inequality we used Bernstein's inequality and the fact that $|\xi-\eta|\geq 20N$.

Recall that when $|\xi+\eta|<\frac{N}{10}$, the last line appears as part of the norm for Λ in the proof of Theorem 2.1 (see (4.40)), thus it is bounded by $\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}G\|_{\mathcal{S}'_r}$ for Λ satisfying (8.5).

Thus it remains to prove Proposition 8.2, to see this, we shall follow the ideas in the proof of Lemma 3.9.

Proof of Proposition 8.2:

Case 1: If $|\tau|^{\frac{1}{2}}>2(|\xi|+|\eta|)$.

Write $u = u^1 + u^2$, where

$$\begin{aligned}\mathcal{F}\,u^1 &= \frac{\mathcal{F}f}{\tau + |\xi|^2 + |\eta|^2}, \text{ this no longer has initial conditions } 0 \\ \mathbf{S}\,u^2 &= 0, \quad \text{a correction so that } u^1 + u^2 \text{ has initial condition } 0.\end{aligned} \tag{8.11}$$

In this case, it suffices to control u_1 since u_2 is only supported where $|\tau| = |\xi|^2 + |\eta|^2$. The goodness about u_1 is that it has the same Fourier support with f. The strategy is based on

$$\begin{aligned}\||\partial_t|^{\frac{1}{4}} u_1\|_{L^\infty(d(x-y))L^2(d(x+y)dt)} &= \|\tau^{\frac{1}{4}} \frac{\mathcal{F}f}{\tau + |\xi|^2 + |\eta|^2}\|_{L^\infty(d(x-y))L^2(d\tau d(\xi+\eta))} \\ &\lesssim \| \int |\tau|^{\frac{1}{4}} \Big| \frac{\mathcal{F}\Big(f\Big)}{\tau + |\xi|^2 + |\eta|^2}\Big| d(\xi - \eta)\|_{L^2(d\tau d(\xi+\eta))}.\end{aligned} \tag{8.12}$$

By Cauchy-Schwarz, we have

$$\begin{aligned}\text{RHS}(8.12) &\lesssim \| \frac{|\tau|^{\frac{1}{4}}}{|\tau|} \int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} |\mathcal{F}f| d(\xi - \eta)\|_{L^2(d\tau d(\xi+\eta))} \\ &\lesssim A\Big\||\nabla_y|^{\frac{1}{2}} |\nabla_x|^{-\frac{1}{2}} f\Big\|_{L^2(dtd(x-y)d(x+y))}\end{aligned} \tag{8.13}$$

where

$$A = \sup_{\tau, \xi+\eta} \frac{|\tau|^{\frac{1}{4}}}{|\tau|} \left(\int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} \frac{|\xi|}{|\eta|} d(\xi - \eta)\right)^{\frac{1}{2}}. \tag{8.14}$$

Changing variables, this is something like

$$A = \sup_{\tau, |u|<|\tau|^{\frac{1}{2}}} \frac{|\tau|^{\frac{1}{4}}}{|\tau|} \left(\int_{|v|<|\tau|^{\frac{1}{2}}} \frac{|u+v|}{|u-v|} dv\right)^{\frac{1}{2}}.$$

After a change of variables this is reduced to $\tau = 1$, and A is bounded. By Sobolev's estimate at an angle, we have

$$\Big\||\nabla_y|^{\frac{1}{2}} |\nabla_x|^{-\frac{1}{2}} f\Big\|_{L^2(dtd(x-y)d(x+y))} \lesssim \||\nabla_x|^{\frac{1}{2}} |\nabla_y|^{\frac{1}{2}} f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.$$

Similarly, by Cauchy-Schwarz, we have

$$\begin{aligned}\text{RHS}(8.12) &\lesssim \| \frac{|\tau|^{\frac{1}{4}}}{|\tau|} \int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} |\mathcal{F}f| d(\xi - \eta)\|_{L^2(d\tau d(\xi+\eta))} \\ &\lesssim A\Big\||\partial_t|^{\frac{1}{4}} |\nabla_{x-y}|^{-\frac{1}{2}} f\Big\|_{L^2(dt)d(x-y)d(x+y)}\end{aligned} \tag{8.15}$$

where

$$A = \sup_{\tau, \xi+\eta} \frac{1}{|\tau|} \left(\int_{|\xi-\eta|<|\tau|^{\frac{1}{2}}} |\xi - \eta|\, d(\xi - \eta)\right)^{\frac{1}{2}}, \tag{8.16}$$

which is bounded. By Sobolev's inequality, we have

$$\big\||\partial_t|^{\frac{1}{4}}|\nabla_{x-y}|^{-\frac{1}{2}}f\big\|_{L^2(dt)d(x-y)d(x+y)} \lesssim \||\partial_t|^{\frac{1}{4}}|\nabla_{x-y}|^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.$$

Case 2: $|\xi| + |\eta| > 2|\tau|^{\frac{1}{2}}$

In this case, since we are assuming that $|\tau|^{\frac{1}{2}} > 10(1 + |\xi + \eta|)$. we have $|\xi - \eta| > |\xi + \eta|$, and also $|\xi - \eta| > |\tau|^{\frac{1}{2}}$. As before, it suffices to bound the right hand side of (8.12). By Cauchy-Schwarz

$$\begin{aligned}\text{RHS}(8.12) &\lesssim \| \int_{2|\xi-\eta|>|\xi+\eta|+|\tau|^{\frac{1}{2}}} |\frac{|\tau|^{\frac{1}{4}}|\mathcal{F}f|}{|\xi-\eta|^2} d(\xi-\eta)\|_{L^2(d\tau d(\xi+\eta))}\\ &\lesssim A\big\||\nabla_y|^{\frac{1}{2}}|\nabla_x|^{-\frac{1}{2}}f\big\|_{L^2(dtd(x-y)d(x+y))}\\ &\lesssim A\||\nabla_x|^{\frac{1}{2}}|\nabla_y|^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\end{aligned} \tag{8.17}$$

where we used Sobolev's estimate at an angle in the last inequality. In this case,

$$A^2 = \sup_{\xi+\eta,\tau} \int_{2|\xi-\eta|>|\xi+\eta|+|\tau|^{\frac{1}{2}}} \frac{|\tau|^{\frac{1}{2}}}{|\xi-\eta|^4}\frac{|\xi|}{|\eta|} d(\xi-\eta). \tag{8.18}$$

Again we scale to $|\tau|^{\frac{1}{2}} + |\xi + \eta| = 1$ and have to estimate

$$\int_{|v|>1} \frac{1}{|v|^4}\frac{|u+v|}{|u-v|}dv.$$

This is bounded uniformly in $|u| < 1$.

Similarly, since we are assuming $|\xi - \eta| \geq 20N$, by Cauchy-Schwarz

$$\begin{aligned}\text{RHS}(8.12) &\lesssim \| \int_{|\xi-\eta|\geq 20N} \frac{|\tau|^{\frac{1}{4}}|\mathcal{F}f|}{|\xi-\eta|^2} d(\xi-\eta)\|_{L^2(d\tau d(\xi+\eta))}\\ &\lesssim N^{-\frac{1}{2}}\big\||\partial_t|^{\frac{1}{4}}f\big\|_{L^2(dtd(x-y)d(x+y))}.\end{aligned}$$

Case 4 :$\frac{1}{2}(|\xi| + |\eta|) < |\tau|^{\frac{1}{2}} < 2(|\xi| + |\eta|)$, In this case, since we are assuming that $|\tau|^{\frac{1}{2}} > 10(1 + |\xi + \eta|)$, this implies $|\tau|^{\frac{1}{2}} \sim |\xi| \sim |\eta| \sim |\xi - \eta|$.

We shall use the decomposition $u = \sum_{k=0}^{\infty} P_{|\tau|\sim 2^k}u$, and the square function estimate

$$\begin{aligned}&\||\partial_t|^{\frac{1}{4}}u\|_{L^\infty(d(x-y))L^2(d(x+y)dt)}\\ &\qquad\sim \|\Big(\sum_{k=0}^{\infty}|P_{|\tau|\sim 2^k}|\partial_t|^{\frac{1}{4}}u|^2\Big)^{\frac{1}{2}}\|_{L^\infty(d(x-y))L^2(d(x+y)dt)}\\ &\qquad\lesssim \Big(\sum_{k=0}^{\infty}2^{2k}\|P_{|\tau|\sim 2^k}u\|^2_{L^\infty(d(x-y))L^2(d(x+y)dt)}\Big)^{\frac{1}{2}}.\end{aligned} \tag{8.19}$$

For each fixed dyadic piece $P_{|\tau|\sim 2^k}u$, in the current case, we have

$$P_{|\tau|\sim 2^k}u = P_{|\tau|\sim 2^k}P_{|\xi|\sim 2^{k/2}}P_{|\eta|\sim 2^{k/2}}u,$$

which implies

$$\begin{aligned}\||P_{|\tau|\sim 2^k}u\|_{collapsing} &\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}P_{|\xi|\sim 2^{k/2}}P_{|\eta|\sim 2^{k/2}}u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\lesssim 2^{-k/2}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\eta|\sim 2^{k/2}}u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\end{aligned}$$

where we used Bernstein's inequality in rotated coordinates twice, see e.g., Lemma 3.1 in [10]. Thus,

$$\begin{aligned}&\||\partial_t|^{\frac{1}{4}}u\|_{L^\infty(d(x-y))L^2(d(x+y)dt)}\\ &\lesssim \Big(\sum_{k=0}^{\infty}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\eta|\sim 2^{k/2}}u\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\Big)^{\frac{1}{2}}\\ &\lesssim \left(\sum_{k=0}^{\infty}\|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}P_{|\eta|\sim 2^{k/2}}f\|^2_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\right)^{\frac{1}{2}}\\ &\lesssim \|\langle\nabla_x\rangle^{\frac{1}{2}}\langle\nabla_y\rangle^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\end{aligned} \tag{8.20}$$

where we used Strichartz (Theorem 3.1) in the second line, and square function estimates in y the the last line.

Similarly, for each fixed dyadic piece $P_{|\tau|\sim 2^k}|\partial_t|^{\frac{1}{4}}u$, in this case since $|\xi-\eta|\sim 2^{\frac{k}{2}}$, by Bernstein's inequality, we have

$$\begin{aligned}&\||P_{|\tau|\sim 2^k}|\partial_t|^{\frac{1}{4}}u\|_{collapsing}\\ &\quad\lesssim \||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\sim 2^{\frac{k}{2}}}u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\ &\quad\lesssim \||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\sim 2^{\frac{k}{2}}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))},\end{aligned} \tag{8.21}$$

where in the last inequality we used Theorem 3.5. Thus,

$$\begin{aligned}&\||\partial_t|^{\frac{1}{4}}u\|_{L^\infty(d(x-y))L^2(d(x+y)dt)}\\ &\lesssim \left(\sum_{k=0}^{\infty}\||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}P_{|\xi-\eta|\sim 2^{\frac{k}{2}}}f\|^2_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\right)^{\frac{1}{2}}\\ &\lesssim \||\partial_t|^{\frac{1}{4}}\langle\nabla_{x-y}\rangle^{\frac{1}{2}}f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\end{aligned} \tag{8.22}$$

where we used square function estimates in $x-y$ the the last line. □

References

[1] V. Bach, S. Breteaux, T. Chen, J. Fröhlich, and I. M. Sigal. The time-dependent Hartree–Fock–Bogoliubov equations for bosons. *Journal of Evolution Equations*, 22(2):46, 2022.

[2] N. Benedikter, G. de Oliveira, and B. Schlein. Quantitative derivation of the Gross–Pitaevskii equation. *Communications on Pure and Applied Mathematics*, 68(8):1399–1482, 2015.

[3] N. Benedikter, M. Porta, B. Schlein, et al. *Effective evolution equations from quantum dynamics*. Springer, 2016.

[4] N. Benedikter, J. Sok, and J. P. Solovej. The Dirac–Frenkel principle for reduced density matrices, and the Bogoliubov–de Gennes equations. In *Annales Henri Poincaré*, volume 19, pages 1167–1214. Springer, 2018.

[5] C. Boccato, C. Brennecke, S. Cenatiempo, and B. Schlein. Complete Bose–Einstein condensation in the Gross–Pitaevskii regime. *Communications in Mathematical Physics*, 359:975–1026, 2018.

[6] C. Boccato, C. Brennecke, S. Cenatiempo, and B. Schlein. Bogoliubov theory in the Gross–Pitaevskii limit. *Acta Mathematica*, 222(2):219–335, 2019.

[7] N. Bogoliubov. On the theory of superfluidity. *J. Phys*, 11(1):23, 1947.

[8] J.-M. Bouclet and H. Mizutani. Uniform resolvent and Strichartz estimates for Schrödinger equations with critical singularities. *Transactions of the American Mathematical Society*, 370(10):7293–7333, 2018.

[9] T. Cazenave and F. B. Weissler. The Cauchy problem for the critical nonlinear Schrödinger equation in H^s. *Nonlinear Analysis: Theory, Methods & Applications*, 14(10):807–836, 1990.

[10] J. Chong, X. Dong, M. Grillakis, M. Machedon, and Z. Zhao. Global uniform in N estimates for solutions of a system of Hartree-Fock-Bogolyubov type in the case $\beta < 1$. *in press, Peking Math J*, 2024.

[11] J. Chong, M. Grillakis, M. Machedon, and Z. Zhao. Global estimates for the Hartree-Fock-Bogolyubov equations. *Communications in Partial Differential Equations*, 46(10):2015–2055, 2021.

[12] J. Chong and Z. Zhao. Dynamical Hartree-Fock-Bogolyubov approximation of interacting bosons. In *Annales Henri Poincaré*, pages 1–59. Springer, 2020.

[13] A. Elgart, L. Erdös, B. Schlein, and H.-T. Yau. Gross–Pitaevskii equation as the mean field limit of weakly coupled bosons. *Archive for rational mechanics and analysis*, 179(2):265–283, 2006.

[14] L. Erdös, B. Schlein, and H.-T. Yau. Rigorous derivation of the Gross–Pitaevskii equation. *Physical review letters*, 98(4):040404, 2007.

[15] L. Erdös, B. Schlein, and H.-T. Yau. Rigorous derivation of the Gross–Pitaevskii equation with a large interaction potential. *Journal of the American Mathematical Society*, 22(4):1099–1156, 2009.

[16] L. Erdös, B. Schlein, and H.-T. Yau. Derivation of the Gross–Pitaevskii equation for the dynamics of Bose-Einstein condensate. *Annals of Mathematics*, pages 291–370, 2010.

[17] V. Georgiev and A. Ivanov. Existence and mapping properties of the wave operator for the Schrödinger equation with singular potential. *Proceedings of the American Mathematical Society*, 133(7):1993–2003, 2005.

[18] J. Ginibre and G. Velo. The classical field limit of scattering theory for non-relativistic many-boson systems. i. *Communications in Mathematical Physics*, 66:37–76, 1979.

[19] M. Grillakis and M. Machedon. Beyond mean field: on the role of pair excitations in the evolution of condensates. *Journal of fixed point theory and applications*, 14(1):91–111, 2013.

[20] M. Grillakis and M. Machedon. Pair excitations and the mean field approximation of interacting bosons, ii. *Communications in Partial Differential Equations*, 42(1):24–67, 2017.

[21] M. Grillakis and M. Machedon. Uniform in N estimates for a bosonic system of hartree–fock–bogoliubov type. *Communications in Partial Differential Equations*, 44(12):1431–1465, 2019.

[22] M. Grillakis, M. Machedon, and D. Margetis. Second-order corrections to mean field evolution of weakly interacting bosons. i. *Communications in Mathematical Physics*, 294(1):273–301, 2010.

[23] M. Grillakis, M. Machedon, and D. Margetis. Second-order corrections to mean field evolution of weakly interacting bosons. ii. *Advances in Mathematics*, 228(3):1788–1815, 2011.
[24] K. Hepp. The classical limit for quantum mechanical correlation functions. *Communications in Mathematical Physics*, 35:265–277, 1974.
[25] X. Huang, J. Chong, M. Machedon, and Z. Zhao. Fock space estimates. *In preparation.*
[26] X. Huang and C. D. Sogge. Uniform Sobolev estimates in R^n involving singular potentials. *The Journal of Geometric Analysis*, 32(2):1–36, 2022.
[27] J.-L. Journé, A. Soffer, and C. D. Sogge. Decay estimates for Schrödinger operators. *Communications on Pure and Applied mathematics*, 44(5):573–604, 1991.
[28] E. Kuz. Exact evolution versus mean field with second-order correction for bosons interacting via short-range two-body potential. *Differential Integral Equations*, 30:587–630, 2017.
[29] T. D. Lee, K. Huang, and C. N. Yang. Eigenvalues and eigenfunctions of a Bose system of hard spheres and its low-temperature properties. *Phys. Rev.*, 106:1135–1145, 1957.
[30] E. H. Lieb, R. Seiringer, J. P. Solovej, and J. Yngvason. *The mathematics of the Bose gas and its condensation*, volume 34. Springer Science & Business Media, 2005.
[31] C. S. Morawetz. Time decay for the nonlinear Klein-Gordon equation. *Proceedings of the Royal Society of London. Series A. Mathematical and physical sciences*, 306(1486):291–296, 1968.
[32] P. Pickl. Derivation of the time dependent Gross–Pitaevskii equation with external fields. *Reviews in Mathematical Physics*, 27(01):1550003, 2015.
[33] I. Rodnianski and W. Schlag. Time decay for solutions of Schrödinger equations with rough and time dependent potentials. *Inventiones mathematicae*, 155(3):451–513, 2004.
[34] I. Rodnianski and B. Schlein. Quantum fluctuations and rate of convergence towards mean field dynamics. *Communications in Mathematical Physics*, 291(1):31–61, 2009.
[35] T. T. Wu. Some nonequilibrium properties of a Bose system of hard spheres at extremely low temperatures. *Journal of Mathematical Physics*, 2(1):105–123, 1961.
[36] K. Yajima. The $\mathrm{W}^{k,p}$-continuity of wave operators for Schrödinger operators. *Journal of the Mathematical Society of Japan*, 47(3):551–581, 1995.
[37] J. Zhang and J. Zheng. Scattering theory for nonlinear Schrödinger equations with inverse-square potential. *Journal of Functional Analysis*, 267(8):2907–2932, 2014.

(X.H.) Department of Mathematics, Louisiana State University, Baton Rouge. LA 70808

Email address: `xhuang49@lsu.edu`