GLOBAL UNIFORM IN N ESTIMATES FOR SOLUTIONS OF A SYSTEM OF HARTREE-FOCK-BOGOLIUBOV TYPE IN THE CASE $\beta < 1$

J. CHONG, X. DONG, M. GRILLAKIS, M. MACHEDON, AND Z. ZHAO

Abstract. We extend the results of the 2019 paper by the third and fourth author globally in time. More precisely, we prove uniform in N estimates for the solutions ϕ, Λ and Γ of a coupled system of Hartree–Fock–Bogoliubov type with interaction potential $V_N(x-y) = N^{3\beta}v(N^{\beta}V_N((x-y))$ with $\beta < 1$. The potential satisfies some technical conditions, but is not small. The initial conditions have finite energy and the "pair correlation" part satisfies a smallness condition, but are otherwise general functions in suitable Sobolev spaces, and the expected correlations in Λ develop dynamically in time. The estimates are expected to improve the Fock space bounds from the 2021 paper of the first and fifth author. This will be addressed in a different paper.

1. Introduction

The general motivation for this paper is the evolution of N Bosons under a mean-field Hamiltonian

$$-\sum_{k=1}^{N} \Delta_{x_i} + \frac{1}{2N}\sum_{k\neq l} V_N(x_k - x_l)$$

where $x_i \in \mathbb{R}^3$, N is large and

$$V_N(x) = N^{3\beta}v(N^{\beta}x)$$

and the potential v is discussed below. (The notation $v_M(x)$ will also be used in sections 2-5 , with a different meaning.) The initial conditions are (exactly or approximately) a tensor product $\phi \otimes \cdots \otimes \phi$.

The exact evolution of the system is approximated by a construction involving just two functions: the condensate $\phi(t, x)$ and a pair

Date: April 5, 2022.
2020 *Mathematics Subject Classification.* 35Q55, 81.
Key words and phrases. Hartree-Fock-Bogoliubov.

excitation function $k(t, x, y)$, and it is

$$\psi_{approx} := e^{-\sqrt{N}\mathcal{A}(\phi(t)} e^{-\mathcal{B}(k(t))}\Omega \tag{1}$$

where

$$\mathcal{A}(\phi) := \int dx \left\{\bar{\phi}(x)a_x - \phi(x)a_x^*\right\} \tag{2}$$

and $e^{-\sqrt{N}\mathcal{A}(\phi)}$ is a unitary operator on Fock space, the Weyl operator. and

$$\mathcal{B}(k) := \frac{1}{2}\int dxdy \left\{\bar{k}(t,x,y)a_x a_y - k(t,x,y)a_x^* a_y^*\right\} \ . \tag{3}$$

The unitary operator $e^{\mathcal{B}(k)}$ is the representation of an (infinite dimensional) real symplectic matrix. Also, Ω is the vacuum. See for instance [10] for background on this construction.

In order for ψ_{approx} to be an approximation to the exact evolution, ϕ and k must satisfy certain PDEs. In the math literature, they were introduced in [10] and independently and in a different context in [1]. They were studied in [11], [12], [5], as well as [2].

To write down the equations it is convenient to consider a self-adjoint kernel

$$\Gamma(t,x,y) = \bar{\phi}(t,x)\phi(t,y) + \frac{1}{N}\left(\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)\right)(t,x,y) := \Gamma_c + \Gamma_p$$

and a symmetric kernel

$$\Lambda(t,x,y) = \phi(t,x)\phi(t,y) + \frac{1}{2N}\mathrm{sh}(2k)(t,x,y) := \Lambda_c + \Lambda_p$$

where

$$\mathrm{sh}(k) := k + \frac{1}{3!}k \circ \bar{k} \circ k + \dots \ ,$$
$$\mathrm{ch}(k) := \delta(x-y) + \frac{1}{2!}\bar{k} \circ k + \dots$$

The functions Λ and Γ have the conceptual meaning of reduced density matrices. Here, $(u \circ v)(x,y) = \int u(x,z)v(z,y)dz$. There are several equivalent ways of expressing the equations. In this section we give a compact, matrix formulation.

For the current paper 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, define

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

where

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

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

Finally, let

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

where I is the identity operator.

The evolution equations for Ω and Ψ (for $t > 0$, with initial conditions at $t = 0$) are

$$\frac{1}{i}\partial_t\Phi - [\Delta_x\delta(x-y)S_3, \Phi]$$
$$= -[(V_N * \rho(t,x))\delta(x-y)S_3, \Phi] - [V_N\Psi^*, \Phi] \quad (4)$$
$$\frac{1}{i}\partial_t\Psi - [\Delta_x\delta(x-y)S_3, \Psi] \quad (5)$$
$$= -[(V_N * \rho)\delta(x-y)S_3, \Psi] - \frac{1}{2N}[S_3, V_N\Psi] - [V_N\Omega^*, \Psi] - \frac{1}{2N}[S_3, V_N\Phi]$$

In addition, the condensate ϕ satisfies

$$\left\{\frac{1}{i}\partial_t - \Delta_{x_1}\right\}\phi(x_1)$$
$$= -\int dy\,\{v_N(x_1-y)\Gamma(y,y)\}\,\phi(x_1)$$
$$-\int dy\big\{v_N(x_1-y)\Gamma_p(y,x_1)\phi(y)$$
$$+\int dy\big\{v_N(x_1-y)\Lambda_p(x_1,y)\big\}\overline{\phi}(y)$$

Here $A^*(x,y) = \bar{A}(y,x)$, $[A,B] = A \circ B - B \circ A$ and V_N acts by pointwise multiplication by $V_N(x-y)$. We will write down these equations in scalar form later, see (67)-(70)). Also, we will write down a simplified model at the end of the introduction.

The arguments of this paper will involve non-local fractional time derivatives, so the values of the solutions at negative times also matter. It is convenient to replace (4), (5) by

$$\frac{1}{i}\partial_t \Phi - [\Delta_x \delta(x-y) S_3, \Phi] = h(t) \times RHS(4) \tag{6}$$

$$\frac{1}{i}\partial_t \Psi - [\Delta_x \delta(x-y) S_3, \Psi] = h(t) \times RHS(5) \tag{7}$$

where $h(t)$ is the characteristic function of $[0, \infty)$ and $RHS(4)$ stands for the right hand side of equation (4). This is the usual solution one gets from Duhamel's formula, and solutions to (4), (5) agree with solutions to (6), (7) fot $t > 0$, provided they have the same initial conditions.

Next, we review the conserved quantities of these equations. See [10] for details. The first conserved quantity is the total number of particles (normalized by division by N):

$$\mathrm{tr}\left\{\Gamma(t)\right\} = \|\phi(t,\cdot)\|^2_{L^2(dx)} + \frac{1}{N}\|\mathrm{sh}(k)(t,\cdot,\cdot)\|^2_{L^2(dxdy)} = 1\ . \tag{8}$$

The second conserved quantity is the energy per particle

$$\begin{aligned} E(t) := \mathrm{tr}\left\{\nabla_{x_1}\cdot\nabla_{x_2}\Gamma(t)\right\} + \frac{1}{2}\int dx_1 dx_2 \left\{V_N(x_1-x_2)\left|\Lambda(t,x_1,x_2)\right|^2\right\} \\ + \frac{1}{2}\int dx_1 dx_2 \left\{V_N(x_1-x_2)\left(\left|\Gamma(t,x_1,x_2)\right|^2 + \Gamma(t,x_1,x_1)\Gamma(t,x_2,x_2)\right)\right\} \\ - \int dx_1 dx_2 \left\{V_N(x_1-x_2)|\phi(t,x_1)|^2|\phi(t,x_2|^2\right\}\ . \end{aligned} \tag{9}$$

The above holds for any Schwartz potential v. In addition, in order to use the estimates of [5], we assume

$$\begin{aligned} &v \text{ is spherically symmetric and} \\ &v \geq 0,\ v \in C_0^\infty,\ \frac{\partial v}{\partial r}(r) \leq 0. \end{aligned} \tag{10}$$

For the initial conditions, we assume there exist a constant C (independent of N) and $\alpha > \frac{1}{2}$ such that

$$\begin{aligned}
&\mathrm{tr}\{\Gamma(0)\} \le C \\
&E(0) \le C \qquad (11)\\
&\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Gamma(0,x,y)\|_{L^2} \le C \text{ (this follows from the previous condition,}\\
&\text{ and will be preserved by the time evolution)}\\
&\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Lambda(0,x,y)\|_{L^2} \le C \text{ (this will also be preserved by the time evolution}\\
&\text{for all } \tfrac{1}{2} \le \alpha \le \alpha_0 \text{ for some } \alpha_0 > \tfrac{1}{2}, \text{ as was shown in [5])}\\
&\||\nabla_x||\nabla_y|\Lambda(0,x,y)\|_{L^2} \le CN.
\end{aligned}$$

The data is also assumed to be in a high H^s space (but not uniformly in N).

In addition, there will be a smallness assumption on the initial conditions for the "pair" components of Γ and Λ. Under the above assumptions, the arguments of [5] imply that for all $\alpha > \frac{1}{2}$, sufficiently close to $\alpha > \frac{1}{2}$, there exists $\epsilon_3 > 0$ such that

$$\int \left||\nabla_x|^{\alpha}|\nabla_y|^{\alpha}\Lambda_p(t,x,y)\right|^2 dxdy \le \frac{C}{N^{\epsilon_3}} \qquad (12)$$

uniformly in t and N. This follows by interpolating between (14) in [5] and Theorem 1.2 in that paper.

In addition, we assume

$$\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Gamma_p(0,\cdot,\cdot)\right\|_{L^2} \le \frac{1}{N^{\epsilon_3}} \qquad (13)$$

α is a number slightly bigger that $\frac{1}{2}$, to be chosen later.

Also, from Proposition 3.4 in [5] we have a Morawetz type estimate

$$||\Gamma(t,x,x)||_{L^2_{t,x}} \lesssim 1 \qquad (14)$$

while from conservation of energy and the trace theorem,

$$||\Gamma(t,x,x)||_{L^\infty(dt)L^2(dx)} \lesssim 1. \qquad (15)$$

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

$$\begin{aligned}
&\|\Lambda\|_{\mathcal{S}_{x,y}} \qquad (16)\\
&= \sup_{p,q \text{ admissible},} \|\Lambda\|_{L^p(dt)L^q(dx)L^2(dy)}\\
&\quad + \sup_{p,q \text{ admissible},} \|\Lambda\|_{L^p(dt)L^q(dy)L^2(dx)}.
\end{aligned}$$

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

$$\|\Lambda\|_{\mathcal{S}_{x,y}} \sim \|\Lambda\|_{L^2(dt)L^6(dx)L^2(dy)} + \|\Lambda\|_{L^2(dt)L^6(dy)L^2(dx)} + \|\Lambda\|_{L^\infty(dt)L^2(dxdy)}.$$

The main result of this paper is

Theorem 1.1. *Let $\Lambda = \Lambda_p + \Lambda_c$, $\Gamma = \Gamma_p + \Gamma_c$ be solutions of* (6), (7) *(or, equivalently,* (67)-(70)*), where the potential satisfies* (10), *and the initial conditions satisfy* (11) *and* (13). *Then we have the a priori estimates*

$$\|\Gamma\|_{L^8(dt)L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \leq C \tag{17}$$

$$\|\nabla_{x+y}\Gamma\|_{L^8(dt)L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \leq C \tag{18}$$

and thus also

$$\|\langle\nabla_{x+y}\rangle^\alpha\Gamma\|_{L^8(dt)L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \leq C. \tag{19}$$

(In what follows, $\alpha > \frac{1}{2}$, close to $\frac{1}{2}$, will be fixed.)

Also, there exists N_0, and $\alpha > \frac{1}{2}$ and C independent of N such that

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda\|_{\mathcal{S}_{x,y}} + \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Gamma\|_{\mathcal{S}_{x,y}} \tag{20}$$

$$+ \|\langle\nabla_{x+y}\rangle^\alpha\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \tag{21}$$

$$+ \||\partial_t|^{\frac{1}{4}}\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \tag{22}$$

$$+ \sup_{x-y}\||\nabla_{x+y}|^\alpha\Gamma\|_{L^2(dtd(x+y))} \leq C \tag{23}$$

for all $N \geq N_0$. In addition, if

$$\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha|\nabla_{x+y}|^j\Lambda(0,\cdot)\right\|_{L^2} + \left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha|\nabla_{x+y}|^j\Gamma(0,\cdot)\right\|_{L^2} \leq C$$

for all $j = 1, \cdots, j_0$, then also

$$\begin{aligned}
&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha|\nabla_{x+y}|^j\Lambda\|_{\mathcal{S}_{x,y}} + \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha|\nabla_{x+y}|^j\Gamma\|_{\mathcal{S}_{x,y}}\\
&+ \|\langle\nabla_{x+y}\rangle^{\alpha+j}\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+ \||\partial_t|^{\frac{1}{4}}|\nabla_{x+y}|^j\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+ \sup_{x-y}\||\nabla_{x+y}|^{\alpha+j}\Gamma\|_{L^2(dtd(x+y))} \leq C.
\end{aligned}$$

The above hold for both the condensate and pair functions.

Remark 1.2. We don't know if $\sup_{x-y}\||\nabla_{x+y}|^\alpha\Gamma\|_{L^2(dtd(x+y))}$ can be replaced by the stronger norm $\||\nabla_{x+y}|^\alpha\Gamma\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$.

For a proof of (17) and (18) see Lemma 6.2. These estimates depend on initial conditions of trace class (or in a Schatten space) and could not be true, even for a linear equation, with just $\dot{H}^s$ initial conditions. See [9], [8]

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

Theorem 1.3. *Let Λ, Γ, ϕ be solutions of* (6), (7), *where the potential satisfies* (10) *and the initial conditions satisfy* (11) *and* (13). *Assume also that*

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

Then, for all $N \geq N_0$ (as in Theorem 1.1)

$$\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}} + \|\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k)\|_{\mathcal{S}_{x,y}} \leq C \log N.$$

Also, assume that for all $j = 1, \cdots, j_0$ we have

$$\||\nabla_{x+y}|^j \mathrm{sh}(2k)(0,\cdot,\cdot)\|_{L^2} + \||\nabla_{x+y}|^j (\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k))(0,\cdot,\cdot))\|_{L^2} \leq C.$$

Then also

$$\|||\nabla_{x+y}|^j \mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}} + \||\nabla_{x+y}|^j (\overline{\mathrm{sh}(k)} \circ \mathrm{sh}(k))\|_{\mathcal{S}_{x,y}} \leq C \log N.$$

Remark 1.4. The above estimates also imply some estimates for $\mathrm{sh}(k)$. In particular,

$$\|\mathrm{sh}(k)\|_{L^p(dx)L^2(dy)} \leq C\|\mathrm{sh}(2k)\|_{L^p(dx)L^2(dy)}. \tag{24}$$

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

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

Corollary 1.5. *Under the assumptions of Theorem 1.1, and the additional assumptions $\|\left\langle \nabla \right\rangle^{\alpha+j} \phi(t=0)\|_{L^2} \leq C$ for all $j = 0, 1, \cdots, j_0$, we have*

$$\|\left\langle \nabla \right\rangle^{\alpha+j} \phi\|_{\mathcal{S}} \leq C$$

We expect the above theorems to have immediate applications to proving a global improved Fock space estimate. This will be addressed in a different paper. We expect to be able to prove

$$\|\psi_{exact}(t) - \psi_{approx}(t)\|_{\mathcal{F}} := \|e^{it\mathcal{H}} e^{-\sqrt{N}\mathcal{A}(\phi_0)} e^{-\mathcal{B}(k(0))}\Omega - e^{i\chi(t)} e^{-\sqrt{N}\mathcal{A}(\phi(t))} e^{-\mathcal{B}(k(t))}\Omega\|_{\mathcal{F}}$$
$$\leq \frac{CP(t)}{N^{\frac{1-\beta}{2}}}$$

for a polynomial $P(t)$, and $0 < \beta < 1$. Currently, the best bounds for growth in time for the above construction are of the form $\frac{e^{Ct}}{N^{\frac{1-\beta}{2}}}$. See [6] for the proof and background material.

Finally, we mention the difficulties surrounding equations (4), (5). 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, treating V_N as δ and ignoring constants, the equations become

$$\mathbf{S}\Lambda_c = \Gamma(t,x,x)\Lambda_c(t,x,y)\} + \Lambda_p(t,x,x)\Gamma_c(t,x,y)$$
$$\mathbf{S}_\pm\Gamma_c = \Gamma(t,x,x)\Gamma_c(t,x,y)\} + \bar{\Lambda}_p(t,x,x)\Lambda_c(t,x,y)$$
$$\mathbf{S}\Lambda_p + \frac{V_N}{N}\Lambda_p = \Gamma(t,x,x)\Lambda_p(t,x,y)\} + \Lambda_p(t,x,x)\Gamma_p(t,x,y) - \frac{V_N}{N}\Lambda_c$$
$$+ \Lambda_c(t,x,x)\Gamma_p(t,x,y)$$
$$\mathbf{S}_\pm\Gamma_p = \Gamma(t,x,x)\Gamma_p(t,x,y) + \Lambda_p(t,x,x)\Lambda_p(t,x,y) + \bar{\Lambda}_c(t,x,x)\Lambda_p(t,x,y).$$

Our method for treating the nonlinear terms requires (roughly) Strichartz estimates for $|\nabla_x|^{\frac{1}{2}}|\nabla_y|^{\frac{1}{2}}\Lambda_{p\ or\ c}$, $|\nabla_x|^{\frac{1}{2}}|\nabla_y|^{\frac{1}{2}}\Gamma_{p\ or\ c}$. But if we apply $|\nabla_x|^{\frac{1}{2}}|\nabla_y|^{\frac{1}{2}}$ to the forcing term $\frac{V_N}{N}\Lambda_c$ in the equation for Λ_p, we get a singularity which approached $\delta(x-y)\Lambda_c$ which cannot be treated by standard $X^{-\frac{1}{2}}$ type techniques.

1.6. **Acknowledgment.** The second and third authors thank Daniel Tataru for the suggestion that the resulting singularities mentioned above are sufficiently special that they can be treated by other methods [20]. Also, we thank Xiaoqi Huang for suggesting several improvements to this paper. He is currently working on extending our current results to the case $\beta = 1$.

J. Chong was supported by the NSF through the RTG grant DMS-RTG 1840314. Z. Zhao was partially supported by the NSF grant of China (No. 12101046) and the Beijing Institute of Technology Research Fund Program for Young Scholars.

2. Statement of the main linear estimates

Let $x, y \in \mathbb{R}^3$, recall $\mathbf{S} = \frac{1}{i}\frac{\partial}{\partial t} - \Delta_x - \Delta_y$ and $h(t)$ denotes the Heaviside function. Consider the equation

$$\begin{aligned}
&\mathbf{S}\Lambda(t,x,y) = h(t)\Big(N^{3\beta-1}v(N^{\beta}(x-y)\Lambda(t,x,y) + G(t,x,y) \\
&+ N^{3\beta-1}v(N^{\beta}(x-y)H(t,x,y)\Big) \qquad (25)\\
&\Lambda(0,\cdot) = \Lambda_0
\end{aligned}$$

for $0 < \beta < 1$ with v is Schwartz.

Recall the definition of $\mathcal{S}_{x,y}$ from (16). Also define the full Strichartz norm (including $L^p(dt)L^q(d(x-y))L^2(d(x+y))$)

$$\begin{aligned}
&\|\Lambda\|_{\mathcal{S}} \qquad (26)\\
&= \sup_{p,q\ admissible} \|\Lambda\|_{L^p(dt)L^q(dx)L^2(dy)}\\
&\quad + \sup_{p,q\ admissible} \|\Lambda\|_{L^p(dt)L^q(dy)L^2(dx)}\\
&\quad + \sup_{p,q\ admissible} \|\Lambda\|_{L^p(dt)L^q(d(x-y))L^2(d(x+y))}
\end{aligned}$$

and 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 as above, and define

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

The reason for excluding $p' = 2$ is that we don't know if we can flip x and y in the double end-point case in Theorem 4.1. The reason $p' = 1$ is excluded is the failure of sharp Sobolev estimates in L^1, see for instance the proof of Lemma 4.6.

Finally, define the "collapsing norms"

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

We will also use the stronger norms $\big\|\Lambda\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$. For the reason we don't work only with this stronger norm see the comments regarding (65). For the reason we don't work only with the collapsing norms, see Remark (6.3).

The simplest form of our theorem is

Theorem 2.1. *Let* Λ *satisfy* (25)*, assume* v *is Schwartz. Let* $0 < \beta < 1$ *and* $\alpha > \frac{1}{2}$ *is sufficiently close to* $\frac{1}{2}$ *(so that* (28)-(30) *hold). Then there* *exists* $\epsilon > 0$ *(depending on* $\beta < 1$*) such that, for* N *sufficiently large,*

$$\begin{aligned}
&\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Lambda\big\|_{\mathcal{S}_{x,y}} + \big\|\big\langle\nabla_{x+y}\big\rangle^\alpha\Lambda\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\big\|\big|\partial_t\big|^{\frac14}\Lambda\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha G\big\|_{\mathcal{S}'_r} + N^{-\epsilon}\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha H\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+N^{-\epsilon}\big\|\big|\partial_t\big|^{\frac14}H\big\|_{collapsing} + N^{-\epsilon}\big\|\big\langle\nabla_{x+y}\big\rangle^\alpha H\big\|_{collapsing}\\
&+\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Lambda_0\big\|_{L^2}.
\end{aligned}$$

Remark 2.2. Notice that the LHS involves the stronger norm $L^2(dt)L^\infty(d(x-y))L^2(d(x+y))$, while the RHS has the weaker "collapsing" norm $L^\infty(d(x-y))L^2(dt)L^2(d(x+y))$.

We first reduce the proof to $\hat{v}$ compactly supported. Let $0 < \epsilon << 1-\beta$ to be chosen below. Start with $v \in \mathcal{S}$ and $\hat{\psi} \in C_0^\infty$, supported in a ball of radius $\frac{1}{10}$, and $\hat{\psi} = 1$ on a neighborhood of 0 and define v_{main} and v_{tail} (depending on N) by $\hat{v}_{main} = \hat{v}(\xi)\hat{\psi}\left(\frac{\xi}{N^\epsilon}\right)$ and $\hat{v}_{tail} = \hat{v}(\xi)\left(1-\hat{\psi}\left(\frac{\xi}{N^\epsilon}\right)\right)$. Since $\hat{v}$ is Schwartz, for any p, $|\xi^\alpha D_\xi^\beta \hat{v}_{tail}(\xi)| \le C_{p,\alpha,\beta}N^{-p}$. Thus we also have $|x^\alpha D_x^\beta v_{tail}(x)| \le C_{p,\alpha,\beta}N^{-p}$ (with a different $C_{p,\alpha,\beta}$, of course).

In all calculations that follow, $N^{3\beta-1}v_{tail}(N^\beta(x-y))$ and its derivatives can be treated as error terms.

It is also simpler to change the notation to $M = N^{\beta+\epsilon}$. Then the Fourier transform of $v_{main}(N^\beta x)$ is $\hat{v}(\frac{\xi}{N^\beta})\hat{\psi}\left(\frac{\xi}{N^{\beta+\epsilon}}\right)$ and is supported in $|\xi| < \frac{M}{10}$. Also, define

$$\begin{aligned}
v_M(x) &= N^{3\beta-1}v(N^\beta x)\\
v_M^1(x) &= N^{3\beta-1}v_{main}(N^\beta x)\\
v_M^2(x) &= N^{3\beta-1}v_{tail}(N^\beta x)
\end{aligned}$$

This definition simplifies the notation during the proof of the main linear theorem (up to section 5). When we deal with the nonlinear equations (starting with section 6) we will use the notation $V_N(x) = N^{3\beta}v(N^\beta x)$.

Thus $v_M^1(x)$ is slightly less singular than $M^2v(Mx)$ as $M \to \infty$, and its Fourier transform is supported in $|\xi| < \frac{M}{10}$, while

$$\|\big\langle\nabla\big\rangle^n v_M^2\|_{L^p} \le C_{n,p}N^{-10} \tag{27}$$

for any $1 \le p \le \infty$, $n \ge 0$. The reader willing to assume $\hat{v}$ is compactly supported in a small neighborhood of 0 can take $M = N$.

At this stage we also choose $\alpha > \frac{1}{2}$, a number $1+$ (slightly bigger than 1), a number $\frac{6}{5}+$ (slightly bigger than $\frac{6}{5}+$) and also $\epsilon_0 > 0$ so that

$$\|v_M^1\|_{L^{\frac{3}{2}}} \lesssim M^{-\epsilon_0} \tag{28}$$

$$\| < \nabla >^{\alpha+\delta_0} v_M^1\|_{L^{\frac{6}{5}+}} \lesssim M^{\alpha+\delta_0}\|v_M^1\|_{L^{\frac{6}{5}+}} \lesssim M^{-\epsilon_0}$$

$$\|\sqrt{N}\langle\nabla\rangle^{\alpha} v_M^1\|_{L^{1+}} + \|\langle\nabla\rangle^{2\alpha} v_M^1\|_{L^{1+}} \tag{29}$$

$$\lesssim \sqrt{N} M^{\alpha(1+\delta_0)}\|v_M^1\|_{L^{1+}} + M^{2\alpha(1+\delta_0)}\|v_M^1\|_{L^{1+}} \lesssim M^{-\epsilon_0}. \tag{30}$$

The above are also true for v_M^2, with a bound of M^{-n} on the right hand side, for any n.

All the implicit constants in $\lesssim$ depend on $\beta < 1$, (which determines the numbers α, δ_0, ϵ_0, $1+$, $\frac{6}{5}+$ described above), and the exponents p_1, p_0 defining and $\mathcal{S}'_r$, but are independent of N (for N large).

3. Estimates in rotated coordinates

In order to prove Theorem 2.1 we will need to adapt standard Sobolev, Bernstein, square function and maximal function estimates to rotated coordinates.

The argument is based on the following lemma:

Lemma 3.1. *Let*

$$R = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{31}$$

(where 1 *stands for the* 3×3 *identity matrix) so that* $\|f \circ R\|_{L^p(dx)L^q(dy)} = \|f\|_{L^p(d(x-y))L^q(d(x+y))}$. *Let* K *be a distribution (possibly* l^2 *valued*[1]*) acting in the* x *variable, and denote* $K\delta = K(x)\delta(y)$ *and* $\delta K = \delta(x)K(y)$ *(tensor products). Assume the following estimate holds, for some* $1 \leq p_1, p_2, q \leq \infty$*:*

$$\|(K\delta) * f\|_{L^{p_1}(dx)L^q(dy)} \lesssim \|f\|_{L^{p_2}(dx)L^q(dy)}$$

Then

$$\|(K\delta) * f\|_{L^{p_1}(d(x-y))L^q(d(x+y))} \lesssim \|f\|_{L^{p_2}(d(x-y))L^q(d(x+y))}$$

or, equivalently

$$\| \left((K\delta) * (f \circ R^{-1})\right) \circ R\|_{L^{p_1}(dx)L^q(dy)} \lesssim \|f\|_{L^{p_2}(dx)L^q(dy)} \tag{32}$$

Also,

$$\|(\delta K) * f\|_{L^{p_1}(d(x-y))L^q(d(x+y))} \lesssim \|f\|_{L^{p_2}(d(x-y))L^q(d(x+y))}$$

[1] In this case, $L^{p_1}(dx)L^q(dy)$ is replaced by $L^{p_1}(dx)L^q(dy)l^2$.

or, equivalently

$$\| \left((\delta K) * (f \circ R^{-1}) \right) \circ R \|_{L^{p_1}(dx) L^q(dy)} \lesssim \|f\|_{L^{p_2}(dx) L^q(dy)}. \tag{33}$$

Proof. In order to prove (32) we use a nonsingular lower triangular matrix L_1 such that

$$(RL_1)^{-1} = \begin{pmatrix} 1 & a \\ 0 & b. \end{pmatrix} \tag{34}$$

Using the invariance of $L^{p_1}(dx)L^q(dy)$ under transformations given by lower triangular matrices, (32) is equivalent to

$$\| \left((K\delta) * (f \circ (RL_1)^{-1}) \right) (RL_1(x,y)) \|_{L^{p_1}(dx) L^q(dy)} \lesssim \|f\|_{L^{p_2}(dx) L^q(dy)}$$

but, by direct calculation (see Lemma 10.6 in the appendix),

$$\left((K\delta) * (f \circ (RL_1)^{-1}) \right) (RL_1(x,y)) = ((K\delta) * f)(x,y).$$

In order to prove (33) we use the same argument, based on a nonsingular lower triangular matrix L_2 such that

$$(RL_2)^{-1} = \begin{pmatrix} c & 1 \\ d & 0 \end{pmatrix} \tag{35}$$

and the calculation

$$\left((\delta K) * (f \circ (RL_2)^{-1}) \right) (RL_2(x,y)) = ((K\delta) * f)(x,y).$$

□

A first consequence is the "Sobolev at an angle" estimate

Lemma 3.2. *Let $\alpha > 0$, $1 \leq p, q, \leq \infty$ and assume the Sobolev estimate $\|u\|_{L^p(dx)} \lesssim \| \left\langle \nabla_x \right\rangle^\alpha \|_{L^q(dx)}$ holds. Then*

$$\begin{aligned} &\|\Lambda\|_{L^p(d(x-y))L^2(d(x+y))} \\ &\lesssim \min\{\| \left\langle \nabla_x \right\rangle^\alpha \Lambda\|_{L^q(d(x-y))L^2(d(x+y))} \| \left\langle \nabla_y \right\rangle^\alpha \Lambda\|_{L^q(d(x-y))L^2(d(x+y))}\} \end{aligned} \tag{36}$$

and also

$$\|\Lambda\|_{L^p(dx)L^2(dy)} \lesssim \| \left\langle \nabla_{x+y} \right\rangle^\alpha \Lambda\|_{L^q(dx))L^2(dy)}.$$

Proof. This follows by using K the kernel of $\left\langle \nabla_x \right\rangle^{-\alpha}$.

□

Another consequence is Bernstein's inequality in rotated coordinates.

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 ϕ_k for $k \geq 0$ by $\hat{\phi}_k(\xi) = \hat{\phi}(\frac{\xi}{2^k})$ and denote

$$P_{|\xi| \leq 2^k} f = f * \phi_k$$

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

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

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

and note that $\sum_{k=0}^{l} \hat{\psi}_k(\xi) = \hat{\phi}(\frac{\xi}{2^l})$ and

$$\sum_{k=0}^{l} f * \psi_k \to f \tag{37}$$

in all L^p spaces ($1 \leq p < \infty$).

More generally, sometimes we will denote by $\hat{\psi}_k(\xi) = \hat{\psi}(\frac{\xi}{2^k})$ for $k \geq 1$ and any $\hat{\psi} \in C_0^\infty(\mathbb{R}^3)$, vanishing on a neighborhood of 0. $\phi = \psi_0$ will only be required to have C_0^∞ Fourier transform. In that case (37) will not be true, but the Bernstein and square function estimates listed below still hold.

The classical Bernstein inequalities are

$$\| <\nabla>^\alpha (\phi_k * f)\|_{L^p(dx)} \lesssim 2^{\alpha k}\|\phi_k * f\|_{L^p(dx)}$$
$$\| <\nabla>^\alpha (\psi_k * f)\|_{L^p(dx)} \sim 2^{\alpha k}\|\psi_k * f\|_{L^p(dx)} \ (\text{ if } k \geq 1)$$

($\alpha \geq 0$, $1 \leq p \leq \infty$). See, for instance, [18]. The (elementary) proof immediately implies (for $1 \leq p, q \leq \infty$)

$$\| <\nabla_x>^\alpha ((\phi_k\delta) * f)\|_{L^p(dx)L^q(dy)} \lesssim 2^{\alpha k}\|(\phi_k\delta) * f\|_{L^p(dx)L^q(dy)}$$
$$\| <\nabla_x>^\alpha ((\psi_k\delta) * f)\|_{L^p(dx)L^q(dy)} \sim 2^{\alpha k}\|(\psi_k\delta) * f\|_{L^p(dx)L^q(dy)}.$$

Using Lemma 3.1 we get

Lemma 3.3. *The following estimates hold*

$$\| <\nabla_x>^\alpha ((\phi_k\delta) * f)\|_{L^p(d(x-y))L^q(d(x+y)} \lesssim 2^{\alpha k}\|(\phi_k\delta) * f\|_{L^p(d(x-y))L^q(d(x+y))}$$
$$\| <\nabla_y>^\alpha ((\delta\phi_k) * f)\|_{L^p(d(x-y))L^q(d(x+y)} \lesssim 2^{\alpha k}\|(\delta\phi_k) * f\|_{L^p(d(x-y))L^q(d(x+y))}$$
$$\| <\nabla_x>^\alpha ((\psi_k\delta) * f)\|_{L^p(d(x-y))L^q(d(x+y)} \sim 2^{\alpha k}\|(\psi_k\delta) * f\|_{L^p(d(x-y))L^q(d(x+y))}$$
$$\| <\nabla_y>^\alpha ((\delta\psi_k) * f)\|_{L^p(d(x-y))L^q(d(x+y)} \sim 2^{\alpha k}\|(\delta\psi_k) * f\|_{L^p(d(x-y))L^q(d(x+y))}$$

Finally, we state two square function estimates. For a function depending only on x, the classical estimate is (for $1 < p < \infty$)

$$\| \left(\sum_{k=0}^{\infty} |f * \psi_k|^2 \right)^{\frac{1}{2}} \|_{L^p(dx)} \sim \|f\|_{L^p(dx)}.$$

The proof can be modified to apply to L^2 valued functions and we have

Lemma 3.4. *Let $1 < p < \infty$. Define $\mathcal{F}\left(P_{|\xi-\eta|\sim 2^k} f\right) = \hat{f}(\xi, \eta)\hat{\psi}\left(\frac{\xi-\eta}{2^k}\right)$ for $k \geq 1$ and $\mathcal{F}\left(P_{|\xi-\eta|\sim 2^0} f\right) = \hat{f}(\xi, \eta)\hat{\phi}(\xi - \eta)$. Then the following estimate holds (for functions which also depend on t)*

$$\| \left(\sum_{k=0}^{\infty} |P_{|\xi-\eta|\sim 2^k} f)|^2 \right)^{\frac{1}{2}} \|_{L^p(d(x-y))L^2(d(x+y)dt)} \sim \|f\|_{L^p(d(x-y))L^2(d(x+y)dt)}.$$

The proof is the same as the standard square function estimate.

Also, we have a result for a "double square function" in rotated coordinates:

Lemma 3.5. *Let $1 < p < \infty$. Then the following estimate holds*

$$\| \left(\sum_{k',k''=0}^{\infty} |f * (\psi_{k'}\delta) * (\delta\psi_{k''})|^2 \right)^{\frac{1}{2}} \|_{L^p(d(x-y))L^2(d(x+y))} \sim \|f\|_{L^p(d(x-y))L^2(d(x+y))}.$$

Proof. For $\lesssim$ see Lemma 10.8 in the Appendix. The opposite inequality is a standard duality argument. □

4. Preliminary estimates for solutions to the linear Schrödinger equation

We will use the following Strichartz estimate (proved in Theorem 2.4, 2.5 of [5]). In $6 + 1$ dimensions,

Theorem 4.1. *Let $\mathbf{S}u = f + g$, $u(0, \cdot) = u_0$. Then*

$$\|u\|_{\mathcal{S}} \lesssim \|f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|g\|_{\mathcal{S}'_r} + \|u_0\|_{L^2}.$$

In the applications that follow, u will be Λ (or Λ_p or Λ_c, or suitable fractional derivatives of Λ), f will be $v_M(x - y)\Lambda(x, y)$ (or suitable derivatives) and g will be G (or suitable derivatives).

After our paper [5] was published, we learned about [13] which contains closely related results (proved with different methods).

Remark 4.2. Another way of obtaining Strichartz estimates on the LHS will be given in Propositions 4.7, 4.8 below.

Using Theorem 4.1 and Lemma 3.2, we can get a (non-sharp) collapsing estimate.

Lemma 4.3. *If* $\mathbf{S}u = f + g$, $u(0, \cdot) = u_0$, *and let* $\alpha > \frac{1}{2}$. *Then*

$$\|u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$$
$$\lesssim \min\{\|\langle\nabla_x\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^\alpha g\|_{\mathcal{S}'_r} + \|\langle\nabla_x\rangle^\alpha u_0\|_{L^2},$$
$$\|\langle\nabla_y\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_y\rangle^\alpha g\|_{\mathcal{S}'_r} + \|\langle\nabla_y\rangle^\alpha u_0\|_{L^2}.\}$$

Proof.

$$\|u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim \|\langle\nabla_x\rangle^\alpha u\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}$$
$$\lesssim \|\langle\nabla_x\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}$$
$$+ \|\langle\nabla_x\rangle^\alpha g\|_{\mathcal{S}'_r} + \|\langle\nabla_x\rangle^\alpha u_0\|_{L^2},$$

and, of course, $\langle\nabla_x\rangle^\alpha$ can be replaced by $\langle\nabla_y\rangle^\alpha$.

□

We record that the above implies

Lemma 4.4. *If* $\mathbf{S}u = f + g$, $u(0, \cdot) = u_0$. *Then*

$$\|\langle\nabla_x\rangle^\alpha u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} + \|\langle\nabla_y\rangle^\alpha u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$$
$$\lesssim \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha g\|_{\mathcal{S}'_r}$$
$$+ \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha u_0\|_{L^2}.$$

We will also need

Lemma 4.5. *If* $\mathbf{S}u = f + g$, $u(0, \cdot) = u_0$. *Then*

$$\|\langle\nabla_{x+y}\rangle^\alpha u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$$
$$\lesssim \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha g\|_{\mathcal{S}'_r}$$
$$+ \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha u_0\|_{L^2}.$$

Proof. The estimate for the homogeneous equation follows from standard Strichartz estimates and "Sobolev at an angle", as in the proof for the inhomogeneous estimate. Thus we can assume $u_0 = 0$. Let $u = \sum_{k,k'=0}^\infty u_{k,k'}$ be a "double" Littlewood-Paley decomposition $u = u * (\psi_k\delta) * (\delta\psi_{k'})$ (see Lemma 3.5) so that $\hat{u}_{k,k'}(\xi, \eta)$ is supported in $|\xi| \sim 2^k$, $|\eta| \sim 2^{k'}$ if $k, k' \geq 1$. We only treat the sum over $k \leq k'$ with $k' \geq 1$ (so x corresponds to the "low" frequency), the remaining part being similar. We use the standard procedure of reducing a Strichartz estimate to a frequency localized estimate, but in the context of mixed coordinates. We have

$$
\begin{aligned}
&\|\langle\nabla_{x+y}\rangle^{\alpha}u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim \|\langle\nabla_{x+y}\rangle^{\alpha}\langle\nabla_x\rangle^{\alpha}u\|\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \quad \text{(Lemma 3.2)}\\
&= \|\langle\nabla_{x+y}\rangle^{\alpha}\langle\nabla_x\rangle^{\alpha}\sum_{0\le k\le k'}u_{k,k'}\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\lesssim \left\|\left(\sum_{0\le k\le k'}\left|\langle\nabla_{x+y}\rangle^{\alpha}\langle\nabla_x\rangle^{\alpha}u_{k,k'}\right|^2\right)^{\frac12}\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \quad \text{(square function estimate)}\\
&\lesssim \left(\sum_{0\le k\le k'}\left\|\langle\nabla_{x+y}\rangle^{\alpha}\langle\nabla_x\rangle^{\alpha}u_{k,k'}\right\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\right)^{\frac12} \quad \text{(Minkowski)}\\
&= \left(\sum_{0\le k\le k'}\left\|\langle\nabla_{x+y}\rangle^{\alpha}\langle\nabla_x\rangle^{\alpha}P_{|\xi+\eta|\lesssim 2^{k'}}u_{k,k'}\right\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\right)^{\frac12}\\
&\lesssim \left(\sum_{0\le k\le k'}\left\|2^{k'\alpha}\langle\nabla_x\rangle^{\alpha}u_{k,k'}\right\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\right)^{\frac12} \quad \text{(Plancherel)}\\
&\lesssim \left(\sum_{0\le k\le k'}\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}u_{k,k'}\right\|^2_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\right)^{\frac12} \quad \text{(Lemma 3.3)}\\
&\\
&\lesssim \left(\sum_{0\le k\le k'}\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}f_{k,k'}\right\|^2_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\right)^{\frac12}\\
&+\left(\sum_{0\le k\le k'}\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}g_{k,k'}\right\|^2_{\mathcal{S}'_r}\right)^{\frac12} \quad \text{(Thm. 4.1)}\\
&\lesssim \left\|\left(\sum_{0\le k\le k'}\left|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}f_{k,k'}\right|^2\right)^{\frac12}\right\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&+\left\|\left(\sum_{0\le k\le k'}\left|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}g_{k,k'}\right|^2\right)^{\frac12}\right\|_{\mathcal{S}'_r} \quad \text{(Minkowski)}\\
&\lesssim \|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}f\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}+\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}g\|_{\mathcal{S}'_r} \quad \text{(square function estimate).}
\end{aligned}
$$

□

Lemma 4.6. *Let* $\mathbf{S}u = f + g$, $u(0,\cdot) = u_0$, *and let* $\frac{1}{2} < \alpha$. *Then*

$$\||\partial_t|^{\frac{1}{4}} u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha f\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} \tag{38}$$
$$+ \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha g\|_{\mathcal{S}'_r} + \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha u_0\|_{L^2}.$$

Proof. The estimate for the homogeneous equation follows from Strichartz estimates and the fact that $|\tau| = |\xi|^2 + |\eta|^2$ on the Fourier support of u, so we can assume $u_0 = 0$. Let $u = \sum_{k,k'=0}^\infty u_{k,k'}$ be a "double" Littlewood-Paley decomposition $u = u*(\psi_k\delta)*(\delta\psi_{k'})$ so that $\hat{u}_{k,k'}(\xi,\eta)$ is supported in $|\xi| \sim 2^k$, $|\eta| \sim 2^{k'}$ if $k, k' \geq 1$. The proof of (38) will use two additional numbers α', α'' satisfying $\frac{1}{2} < \alpha' < \alpha'' < \alpha$. Start by fixing $\frac{1}{2} < \alpha'' < \alpha$. We will prove the frequency localized estimate

$$\||\partial_t|^{\frac{1}{4}} u_{k,k'}\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \tag{39}$$
$$\lesssim \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} f_{k,k'}\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} g_{k,k'}\|_{\mathcal{S}'_r}.$$

Summing the pieces will be easy because $\alpha'' < \alpha$. To prove (39), assume, without loss of generality, $1 \leq k \leq k'$.

Now we localize $u_{k,k'}$ in τ. This changes the initial conditions, but in a controlled way. From Theorem 4.1 we have

$$\|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} u_{k,k'}\|_{L^\infty(dt)L^2(dxdy)}$$
$$\lesssim \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} f_{k,k'}\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} g_{k,k'}\|_{\mathcal{S}'_r}.$$

Since $P_{|\tau|<100\,2^{2k'}}$ acts by convolution in time with a function which is in L^1 uniformly in k', we also have

$$\|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} P_{|\tau|<100\,2^{2k'}} u_{k,k'}(0,\cdot,\cdot)\|_{L^2(dxdy)}$$
$$\lesssim \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} f_{k,k'}\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^{\alpha''} \langle\nabla_y\rangle^{\alpha''} g_{k,k'}\|_{\mathcal{S}'_r}.$$

We have

$$\mathbf{S}P_{|\tau|<100\,2^{2k'}} u_{k,k'} = P_{|\tau|<100\,2^{2k'}} f_{k,k'} + P_{|\tau|<100\,2^{2k'}} g_{k,k'}$$

with initial conditions as discussed above. Using Lemma 3.3 we get

$$\begin{aligned}
&\||\partial_t|^{\frac{1}{4}} P_{|\tau|<100\, 2^{2k'}} u_{k,k'}\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \\
&\lesssim 2^{\frac{k'}{2}} \||P_{|\tau|<100\, 2^{2k'}} u_{k,k'}\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \\
&\lesssim \|\langle \nabla_y \rangle^{\frac{1}{2}} P_{|\tau|<100\, 2^{2k'}} u_{k,k'}\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \\
&\lesssim \|\langle \nabla_x \rangle^{\alpha'} \langle \nabla_y \rangle^{\alpha'} f_{k,k'}\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} \\
&+ \|\langle \nabla_x \rangle^{\alpha'} \langle \nabla_y \rangle^{\alpha'} g_{k,k'}\|_{\mathcal{S}'_r}.
\end{aligned}$$

(See Lemma 4.4.)

Next, consider

$$\mathbf{S} P_{|\tau|>100\, 2^{2k'}} u_{k,k'} = P_{|\tau|>100\, 2^{2k'}} f_{k,k'} + P_{|\tau|>100\, 2^{2k'}} g_{k,k'}$$

and 2^k (the frequency of x) is less than $2^{k'}$ (the frequency of y).

Call either function on the RHS $P_{|\tau|>100\, 2^{2k'}} h_{k,k'}$. The point is that $\mathbf{S} \sim \langle \partial_t \rangle$ is an elliptic operator on the Fourier support of $u_{k,k'}$, and, at the level of symbols and on $L^2(dtdxdy)$, $\langle \partial_t \rangle \geq \langle \nabla_y \rangle^2 \geq \langle \nabla_x \rangle^2$.

Let $\frac{1}{2} < \alpha' < \alpha'' < \alpha$, with α' to be chosen later (the choice will be $\alpha' = \frac{\frac{1}{2}+2\alpha''}{3}$). Using Lemma 3.2,

$$\begin{aligned}
&\|\langle \partial_t \rangle^{\frac{1}{4}} P_{|\tau|>100\, 2^{2k'}} u_{k,k'}\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \\
&\lesssim \|\langle \partial_t \rangle^{\frac{1}{4}} \langle \nabla_x \rangle^{3\alpha'} P_{|\tau|>100\, 2^{2k'}} u_{k,k'}\|_{L^2(dtdxdy)} \\
&\lesssim \|\langle \partial_t \rangle^{\frac{1}{4}-1} \langle \nabla_x \rangle^{3\alpha'} \mathbf{S} P_{|\tau|>100\, 2^{2k'}} u_{k,k'}\|_{L^2(dtdxdy)} \\
&= \|\langle \partial_t \rangle^{-\frac{3}{4}} \langle \nabla_x \rangle^{3\alpha'} P_{|\tau|>100\, 2^{2k'}} h_{k,k'}\|_{L^2(dtdxdy)} \\
&= \|\langle \partial_t \rangle^{-\frac{1}{2}} \left(\langle \partial_t \rangle^{-\frac{1}{4}} \langle \nabla_x \rangle^{3\alpha'-2\alpha''} \right) \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} h_{k,k'}\|_{L^2(dtdxdy)} \\
&\lesssim \|\langle \partial_t \rangle^{-\frac{1}{2}} \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} h_{k,k'}\|_{L^2(dtdxdy)}
\end{aligned}$$

since the choice of α' insures $3\alpha' - 2\alpha'' = \frac{1}{2}$ so that $\langle \partial_t \rangle^{-\frac{1}{4}} \langle \nabla_x \rangle^{3\alpha'-2\alpha''}$ is bounded on L^2 on the Fourier support of $P_{|\tau|>100\, 2^{2k'}} h_{k,k'}$.

By applying Sobolev estimates in t or x, y or $x-y$, and recalling $\langle \partial_t \rangle \geq \langle \nabla_y \rangle^2 \geq \langle \nabla_x \rangle^2$, $\langle \partial_t \rangle \geq \langle \nabla_{x-y} \rangle^2$ we get

$$\begin{aligned}
&\| \langle \partial_t \rangle^{-\frac{1}{2}} \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} h_{k,k'} \|_{L^2(dtdxdy)} \\
&\lesssim \min \Big\{ \| \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} h_{k,k'} \|_{\mathcal{S}'_r}, \\
&\| \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} h_{k,k'} \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \Big\}.
\end{aligned}$$

Recalling the definition of $h_{k,k'}$ the above is dominated by

$$\| \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} f_{k,k'} \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} + \| \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} P_{|\tau|>100\, 2^{2k'}} g_{k,k'} \|_{\mathbf{S}'_r}$$

.

To sum the pieces, use $\alpha'' < \alpha$. For instance,

$$\begin{aligned}
&\sum_{k,k'=0}^{\infty} \| \langle \nabla_x \rangle^{\alpha''} \langle \nabla_y \rangle^{\alpha''} f_{k,k'} \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \\
&\lesssim \sum_{k,k'=0}^{\infty} 2^{(\alpha''-\alpha)k} 2^{(\alpha''-\alpha)k'} \| \langle \nabla_x \rangle^{\alpha} \langle \nabla_y \rangle^{\alpha} (f * (\psi_k \delta) * (\delta \psi_{k'})) \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \\
&\lesssim \sup_{k,k'} \| \langle \nabla_x \rangle^{\alpha} \langle \nabla_y \rangle^{\alpha} f * (\psi_k^1 \delta) * (\delta \psi_{k'}) \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \\
&\lesssim \| \langle \nabla_x \rangle^{\alpha} \langle \nabla_y \rangle^{\alpha} f \|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))}.
\end{aligned}$$

□

Next, we prove a frequency localized sharp result.

Proposition 4.7. *Let*

$$\mathbf{S} u_k = f_k$$

with 0 *initial conditions, and assume* u_k *(and thus also* f_k*) is supported, in Fourier space, at* $|\xi - \eta| \sim 2^k$. *Then*

$$\|u_k\|_{\mathcal{S}_{x,y}} \lesssim \left\| |\nabla_{x+y}|^{\frac{1}{2}} f_k \right\|_{L^1(d(x-y)) L^2(dt) L^2(d(x+y))} + \left\| |\partial_t|^{\frac{1}{4}} f_k \right\|_{L^1(d(x-y)) L^2(dt) L^2(d(x+y))}.$$

Proof. As we did earlier, we decompose $u_k = u_k^1 + u_k^2 + u_k^3$, where

$$\mathbf{S} u_k^1 = P_{10|\tau|^{\frac{1}{2}} \geq 2^k} f_k := f_k^1 \text{ with initial conditions } 0$$

$$\mathcal{F} u_k^2 = \frac{\mathcal{F}\left(P_{10|\tau|^{\frac{1}{2}} \leq 2^k} f_k\right)}{\tau + |\xi|^2 + |\eta|^2} \text{ (this no longer has initial conditions 0)}$$

$$\mathbf{S} u_k^3 = 0, \text{ a correction so that } u_k^2 + u_k^3 \text{ has initial conditions } 0.$$

For u_k^1 the argument is based on the Strichartz and Sobolev estimates at fixed frequency:

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

We used Plancherel and the Fourier support of f_k^1.

For u_k^2, the denominator is comparable with $|\xi-\eta|^2 + |\xi+\eta|^2 \geq 2^{2k} \geq 100|\tau|$, and

$$\|u_k^2\|_{L^2(dtdxdy)} \lesssim \||\nabla_{x-y}|^{-\frac{3}{2}} |\nabla_{x+y}|^{-\frac{1}{2}} f_k\|_{L^2(dtdxdy)}$$

and

$$\||\nabla_{x+y}| u_k^2\|_{L^2(dtdxdy)} \lesssim \||\nabla_{x-y}|^{-\frac{3}{2}} |\nabla_{x+y}|^{\frac{1}{2}} f_k\|_{L^2(dtdxdy)}$$
$$\lesssim \||\nabla_{x+y}|^{\frac{1}{2}} f_k\|_{L^1(d(x-y))L^2(dtd(x+y))}.$$

Thus, using Lemma 3.2 (Sobolev estimates at an angle) we have

$$\|u_k^2\|_{L^2(dt)L^6(dx)L^2(dy)} + \|u_k^2\|_{L^2(dt)L^6(dy)L^2(dx)} \lesssim \||\nabla_{x+y}| u_k^2\|_{L^2(dtdxdy)}$$
$$\lesssim \||\nabla_{x+y}|^{\frac{1}{2}} f_k\|_{L^1(d(x-y))L^2(dtd(x+y))}. \tag{40}$$

Similarly, we have

$$\||\partial_t|^{\frac{1}{2}} u_k^2\|_{L^2(dtdxdy)} \lesssim \||\nabla_{x-y}|^{-\frac{3}{2}} |\partial_t|^{\frac{1}{4}} f_k\|_{L^2(dtdxdy)}$$
$$\lesssim \||\partial_t|^{\frac{1}{4}} f_k\|_{L^1(d(x-y))L^2(dtd(x+y))}.$$

Unfortunately, the desired $L^\infty(dt)$ Sobolev estimate is false, so we proceed slightly differently:

$$\|u_k^2\|_{L^\infty(dt)L^2(dxdy)} \lesssim 2^{-\frac{k}{2}} \Big\| \int_{\tau\in[-2^{2k},2^{2k}]} \frac{1}{|\tau|^{\frac{1}{4}}} |\tau|^{\frac{1}{4}} |\xi-\eta|^{-\frac{3}{2}} |\widetilde{f}_k| d\tau \Big\|_{L^2(d\xi d\eta)}$$
$$\lesssim \||\tau|^{\frac{1}{4}} |\xi-\eta|^{-\frac{3}{2}} |\widetilde{f}_k|\|_{L^2(\tau\in[-2^{2k},2^{2k}]d(\xi d\eta))}$$
$$= c\||\nabla_{x-y}|^{-\frac{3}{2}} |\partial_t|^{\frac{1}{4}} f_k\|_{L^2(dtdxdy)} \lesssim \||\partial_t|^{\frac{1}{4}} f_k\|_{L^1(d(x-y))L^2(dtd(x+y))}.$$

Finally, by interpolation,

$$\|u_k^2\|_{\mathcal{S}_{x,y}} \lesssim \|u_k^2\|_{L^2(dt)L^6(dx)L^2(dy)} + \|u_k^2\|_{L^2(dt)L^6(dy)L^2(dx)} + |u_k^2\|_{L^\infty(dt)L^2(dxdy)}$$

and we get the desired result for u_k^2. Since $\|u_k^2(t=0)\|_{L^2} = \|u_k^3(t=0)\|_{L^2}$, the result for u_k^3 is trivial.

□

We also record the following version:

Proposition 4.8. *Let*

$$\mathbf{S}u = N^3 v(N(x-y)\Lambda := f$$

with 0 *initial conditions and* $v \in \mathcal{S}$.

$$\|u\|_{\mathcal{S}_{x,y}} \lesssim \log N \left(\left\| \left\langle \nabla_{x+y} \right\rangle^{\frac{1}{2}} \Lambda \right\|_{collapsing} + \left\| |\partial_t|^{\frac{1}{4}} \Lambda \right\|_{collapsing} \right).$$

The same result holds for the equation

$$\mathbf{S}u = P_{|\xi|<M} P_{|\eta|<M} \left(N^3 v(N(x-y)\Lambda \right)$$

for all $M > 0$*, with implicit constants independent of* M*.*

(see (41) for the definition of $P_{|\xi|<M} P_{|\eta|<M}$).

Proof. First we consider $P_{|\xi-\eta|>N} u$, and in this case we get the result without a log. In that case

$$\|u\|_{\mathcal{S}_{x,y}} \lesssim \|f\|_{L^2(dt) L^{\frac{6}{5}}(d(x-y)) L^2(d(x+y))} \lesssim N^{\frac{1}{2}} \|\Lambda\|_{collapsing}.$$

If we localize to the region where $|\tau|^{\frac{1}{2}} + |\xi+\eta| > \frac{N}{10}$, the result follows from Plancherel. Thus we can assume $\tilde{u}$ is supported in $|\tau|^{\frac{1}{2}} + |\xi+\eta| < \frac{N}{10}$, $|\xi-\eta| > N$ and in particular we can solve $\mathbf{S}$ by dividing by the symbol, as in the previous proof. Let u_1 be the solution (which differs from u by a solution to the homogeneous equation). Then we get

$$\|u_1\|_{L^2(dtdxdy)} \lesssim \frac{1}{N^2} \|f\|_{L^2} \lesssim \frac{1}{\sqrt{N}} \|\Lambda\|_{collapsing}$$

and also

$$\|u_1\|_{L^2(dt)L^6(dx)L^2(dy)} + \|u_1\|_{L^2(dt)L^6(dy)L^2(dx)} \lesssim \||\nabla_{x+y}| u_1\|_{L^2(dtdxdy)}$$
$$\lesssim \frac{1}{\sqrt{N}} \||\nabla_{x+y}| \Lambda\|_{collapsing} \lesssim \left\| |\nabla_{x+y}|^{\frac{1}{2}} \Lambda \right\|_{collapsing}$$

and also, using Cauchy-Schwartz in τ,

$$\|u_1\|_{L^\infty(dt)L^2(dxdy)} \lesssim \left\| \int_{\tau \in [-N^2, N^2]} \frac{1}{|\tau|^{\frac{1}{4}}} |\tau|^{\frac{1}{4}} |\widetilde{u_1}| d\tau \right\|_{L^2(d\xi d\eta)}$$
$$\lesssim \sqrt{N} \| |\tau|^{\frac{1}{4}} |\widetilde{u_1}| \|_{L^2(\tau \in [-N^2, N^2] d(\xi d\eta))} \lesssim \left\| |\partial_t|^{\frac{1}{4}} \Lambda \right\|_{collapsing}.$$

Next, we consider $P_{|\xi-\eta|<N} u = P_{|\xi-\eta|<1} u + \sum P_{|\xi-\eta| \sim 2^i} u$ where the sum has about $\log N$ terms. Proposition 4.7 applies to each term, and the result follows.

□

5. Proof of Theorem 2.1

Recall the "projections" $P_{|\xi|>M}$ and $P_{|\xi|<M}$ by

$$\mathcal{F}\left(P_{|\xi|<M}\Lambda\right)(\xi,\eta) = \hat{\phi}\left(\frac{|\xi|}{M}\right)\mathcal{F}\Lambda(\xi,\eta) \tag{41}$$

and $P_{|\xi|>M} = 1 - P_{|\xi|<M}$. The function $\hat{\phi}$ is a C_0^∞ function (which can change from line to line).

Here $\mathcal{F}$ denotes the Fourier transform. These multipliers are bounded on $L^q(dx)L^2(dy)$, $L^q(dy)L^2(dx)$ and $L^q(d(x-y))L^2(d(x+y))$ for $1 \le q \le \infty$ (uniformly in M). Also, we adopt the convention that P and ϕ may change from line to line.

We have to show: If $h(t)$ is the Heaviside function and

$$\begin{aligned}
&\mathbf{S}\Lambda(t,x,y) = h(t)\Big(N^{3\beta-1}v(N^\beta(x-y))\Lambda(t,x,y) + G(t,x,y)\\
&\quad + N^{3\beta-1}v(N^\beta(x-y))H(t,x,y)\Big)\\
&\Lambda(0,\cdot) = \Lambda_0
\end{aligned}$$

then

$$\begin{aligned}
&\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda\right\|_{\mathcal{S}_{x,y}}\\
&+\left\||\nabla_{x+y}|^\alpha\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\||\partial_t|^{\frac14}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim \left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\right\|_{\mathcal{S}'_r} + N^{-\epsilon}\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+N^{-\epsilon}\left\||\partial_t|^{\frac14}H\right\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&+N^{-\epsilon}\left\||\nabla_{x+y}|^\alpha H\right\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\\
&+\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\right\|_{L^2}
\end{aligned}$$

and $N^{3\beta-1}v(N^\beta(x-y)$ is denoted by $v_M(x-y)$, and M is of the form $N^{1+\epsilon}$ for some small $\epsilon > 0$.

Before starting the proof we remark that we have the following Strichartz estimate (see Theorem 4.1)

$$\begin{aligned}
&\|\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \lesssim \|v_M(x-y)\Lambda\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&\quad+\|v_M(x-y)H\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))} + \|G\|_{\mathcal{S}'_r} + \|\Lambda_0\|_{L^2}.
\end{aligned}$$

Using Hölder's inequality and the fact that $\|v_M\|_{L^{\frac32}} = O(M^{-\epsilon_0})$ as $M \to \infty$ (see (28)) we can treat the potential term as a perturbation and get

$$\|\Lambda\|_{\mathcal{S}} \lesssim \|\Lambda_0\|_{L^2} + \|v_M(x-y)H\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))} + \|G\|_{\mathcal{S}'_r} \quad (42)$$
$$\lesssim \|\Lambda_0\|_{L^2} + M^{-\epsilon_0}\|H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} + \|G\|_{\mathcal{S}'_r}.$$

We can do the same after taking $|\nabla_x|^\alpha$, but we need a suitable Leibniz rule. Using the outline of [7], [15]. [19], it is possible to prove

Theorem 5.1. *Let $\alpha, \beta \geq 0$. Let $f(x,y) = v(x-y)$ with $v \in \mathcal{S}$. Then*

$$\||\nabla_x|^\alpha |\nabla_y|^\beta \left(f(x,y)g(x,y)\right)\|_{L^r(d(x-y))L^2(d(x+y))}$$
$$\lesssim \| < \nabla_x >^{\alpha+\beta} v\|_{L^{p_1}} \|g\|_{L^{q_1}(d(x-y))L^2(d(x+y))}$$
$$+ \| < \nabla_x >^{\alpha} v\|_{L^{p_2}} \| < \nabla_y >^{\beta} g\|_{L^{q_2}(d(x-y))L^2(d(x+y))}$$
$$+ \| < \nabla_y >^{\beta} v\|_{L^{p_3}} \| < \nabla_x >^{\alpha} g\|_{L^{q_3}(d(x-y))L^2(d(x+y))}$$
$$+ \|v\|_{L^{p_4}} \| < \nabla_x >^{\alpha} < \nabla_y >^{\beta} g\|_{L^{q_4}(d(x-y))L^2(d(x+y))}$$

($\frac{1}{r} = \frac{1}{p_i} + \frac{1}{q_i}$, $1 < r, p_i, q_i < \infty$). In addition, if $\hat{v}$ is supported in $|\xi| \leq \frac{M}{10}$ and $\hat{g}(\xi,\eta)$ is supported in $|\xi| > 10M$, then the $|\nabla_x|^\alpha$ derivatives only fall on g and we have

$$\||\nabla_x|^\alpha |\nabla_y|^\beta \left(f(x,y)g(x,y)\right)\|_{L^r(d(x-y))L^2(d(x+y))}$$
$$\lesssim \| < \nabla_y >^{\beta} v\|_{L^{p_3}} \| < \nabla_x >^{\alpha} g\|_{L^{q_3}(d(x-y))L^2(d(x+y))}$$
$$+ \|v\|_{L^{p_4}} \| < \nabla_x >^{\alpha} < \nabla_y >^{\beta} g\|_{L^{q_4}(d(x-y))L^2(d(x+y))}.$$

However, for our purposes it is enough to use the following version, which is easier to prove. We thank Xiaoqi Huang for suggesting this approach.

Theorem 5.2. *Let $\alpha, \beta \geq 0$. Let v_M as above with $\hat{v}_M$ is supported in $|\xi| \leq \frac{M}{10}$, and let $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, $1 < r, p, q < \infty$. Then*

$$\||\nabla_x|^\alpha \left(v_M(x-y)g(x,y)\right)\|_{L^r(d(x-y))L^2(d(x+y))} \quad (43)$$
$$\lesssim M^\alpha \|v_M\|_{L^r} \|g\|_{L^\infty(d(x-y))L^2(d(x+y))} \quad (44)$$
$$+ \|v_M\|_{L^{p_1}} \| < \nabla_x >^{\alpha} g\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \quad (45)$$

and similarly for $|\nabla_y|^\alpha$. Also, if $\frac{1}{r} = \frac{1}{p_i} + \frac{1}{q_i}$, $1 < r, p_i, q_i < \infty$,

$$\| |\nabla_x|^\alpha \nabla_y|^\beta \left(v_M(x-y)) P_{|\eta|>M} g(x,y) \right) \|_{L^r(d(x-y))L^2(d(x+y))} \tag{46}$$
$$\lesssim M^\alpha \|v_M\|_{L^{p_1}} \| |\nabla_y|^\beta g\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \tag{47}$$
$$+ \|v\|_{L^{p_2}} \| < \nabla_x >^\alpha < \nabla_y >^\beta g\|_{L^{q_2}(d(x-y))L^2(d(x+y))} \tag{48}$$

and also

$$\| |\nabla_x|^\alpha \nabla_y|^\beta \left(v_M(x-y)) P_{|\eta|<M} P_{|\eta|<M} g(x,y) \right) \|_{L^r(d(x-y))L^2(d(x+y))}$$
$$\lesssim M^{\alpha+\beta} \|v_M\|_{L^{p_3}} \|g\|_{L^{q_3}(d(x-y))L^2(d(x+y))}. \tag{49}$$

Proof. We use standard Littlewood-Paley operators $P_{|\xi|\sim 2^i M}$, $P_{|\xi|\lesssim 2^i M}$. In the name of simplicity of notation, we allow the implicit constants in $\sim$, $\lesssim$ to be different in different instances. However, the modified projections will be denoted by $\bar{P}$. Thus, for instance, $\bar{P}_{|\xi|\sim 2^i M} P_{|\xi|\sim 2^i M} = P_{|\xi|\sim 2^i M}$. Also, we use the notation $|\nabla|^\alpha P_{|\xi|\sim 2^i M} g = (2^i M)^\alpha \bar{P}_{|\xi|\sim 2^i M} g$. In the first instance, the multiplier is of the form $\psi(\frac{\xi}{M})$ with $\psi \in C_0^\infty(\mathbb{R}^n)$, $\psi(\xi) = 0$ in a neighborhood of 0. In the second case, ψ is replaced by $|\xi|^\alpha \psi(\xi)$ which has the same properties. When $\bar{P}$ is further modified, it is still denoted $\bar{P}$. This is the convention used in [15]. Finally, denote $|\nabla|^\alpha P_{|\xi|\lesssim M} g = M^\alpha \bar{P}_{|\xi|\lesssim M} g$, where the exact definition of $\bar{P}$ is seen on the Fourier transform side. The corresponding multiplier is not smooth near 0, but the corresponding kernel is in L^1. After the modifications described above, the identity $g = \bar{P}_{|\xi|\lesssim M} g + \sum_{i=1}^\infty \bar{P}_{|\xi|\sim 2^i M} g$ is no longer true. Square function estimates have to be used instead, and the square function operators constructed using $\bar{P}_{|\xi|\sim 2^i M}$ have the same mapping properties as those using $P_{|\xi|\sim 2^i M}$.

For (43), decompose $g = P_{|\xi|<M} g + P_{|\xi|>M} g$. Arguing as in the proof of Bernstein's inequality,

$$|\nabla_x|^\alpha \left(v_M(x-y) P_{|\xi|<M} g(x,y) \right) = |\nabla_x|^\alpha \bar{P}_{|\xi|\lesssim M} \left(v_M(x-y) P_{|\xi|<M} g(x,y) \right)$$
$$= M^\alpha \bar{P}_{|\xi|\lesssim M} \left(v_M(x-y) P_{|\xi|<M} g(x,y) \right).$$

Since $\bar{P}_{|\xi|\lesssim M}$ is given by convolution with a kernel which is in $L^1(dx)$ uniformly in M,

$$\|\bar{P}_{|\xi|\lesssim M} \left(v_M(x-y) P_{|\xi|<M} g(x,y) \right) \|_{L^r(d(x-y))L^2(d(x+y))}$$
$$\lesssim \|v_M(x-y) P_{|\xi|<M} g(x,y)\|_{L^r(d(x-y))L^2(d(x+y))}$$
$$\lesssim \|v_M\|_{L^r} \|g\|_{L^\infty(d(x-y))L^2(d(x+y))}$$

For $P_{|\xi|>M} g$, decompose it as $P_{|\xi|>M} g = \sum_{i=1}^\infty P_{|\xi|\sim 2^i M} g$. Then $|\nabla|^\alpha \left(v_M(x-y) P_{|\xi|\sim 2^i M} g \right) = (2^i M)^\alpha \bar{P}_{|\xi|\sim 2^i M} \left(v_M(x-y) P_{|\xi|\sim 2^i M} g \right)$.

Proving the estimate by duality involves using a test function H with $\|H\|_{L^{r'}(dx)L^2(dy)} = 1$ and looking at

$$\int |\nabla_x|^\alpha \big(v_M(x-y)P_{|\xi|>M}g(x,y)\big)h(x,y)dxdy \tag{50}$$

with $h = H \circ R^{-1}$, or $\|h\|_{L^{r'}(d(x-y))L^2(d(x+y))} = 1$ We have

$$
\begin{aligned}
&|(50)| = \left| \int v_M(x-y) \sum (2^i M)^\alpha P_{|\xi|\sim 2^i M} g \bar{P}_{|\xi|\sim 2^i M} h \right| \\
&\leq \int |v_M(x-y)| \left(\sum |(2^i M)^\alpha P_{|\xi|\sim 2^i M} g|^2\right)^{\frac{1}{2}} \left(\sum |\bar{P}_{|\xi|\sim 2^i M} h|^2\right)^{\frac{1}{2}} \\
&\leq \|v_M\|_{L^{p_1}} \left\| \left(\sum |(2^i M)^\alpha P_{|\xi|\sim 2^i M} g|^2\right)^{\frac{1}{2}} \right\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \\
&\times \left\| \left(\sum |\bar{P}_{|\xi|\sim 2^i M} h|^2 \right|\right)^{\frac{1}{2}} \right\|_{L^{r'}(d(x-y))L^2(d(x+y))} \\
&\lesssim \|v_M\|_{L^{p_1}} \left\| \left(\sum ||\nabla_x|^\alpha \bar{P}_{|\xi|\sim 2^i M} g|^2\right)^{\frac{1}{2}} \right\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \\
&\times \left\| \left(\sum |\bar{P}_{|\xi|\sim 2^i M} h|^2 \right|\right)^{\frac{1}{2}} \right\|_{L^{r'}(d(x-y))L^2(d(x+y))} \\
&\lesssim \|v_M\|_{L^{p_1}} \| < \nabla_x >^\alpha g\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \|h\|_{L^{r'}(d(x-y))L^2(d(x+y))}.
\end{aligned}
$$

This uses the square function estimate in rotated coordinates.

For (46), let $\frac{1}{r} = \frac{1}{p_1} + \frac{1}{q_1}$, $1 < r, p_1, q_1 < \infty$. Then, use the same duality argument as in the previous proof. with $\|h\|_{L^{r'}(d(x-y))L^2(d(x+y))} = 1$.

The bound (47) corresponds to

$$
\begin{aligned}
&\| |\nabla_x|^\alpha |\nabla_y|^\beta \left(v_M(x-y)) P_{|\xi|<M} P_{|\eta|>M} g(x,y) \right) \|_{L^r(d(x-y))L^2(d(x+y))} \\
&= \left\| \sum_{i=1}^\infty \bar{P}_{|\eta|\sim 2^i M} \bar{P}_{|\xi|\lesssim M} \left(M^\alpha v_M(x-y)) P_{|\xi|<M} (2^i M)^\beta P_{|\eta|\sim 2^i M} g(x,y) \right) \right\|_{L^r(d(x-y))L^2(d(x+y))} \\
&= \left| \int \sum_{i=1}^\infty \bar{P}_{|\eta|\sim 2^i M} \bar{P}_{|\xi|\lesssim M} \left(M^\alpha v_M(x-y)) P_{|\xi|<M} (2^i M)^\beta P_{|\eta|\sim 2^i M} g(x,y) \right) h \right| \\
&= \left| \int \sum_{i=1}^\infty \left(M^\alpha v_M(x-y)) P_{|\xi|<M} (2^i M)^\beta P_{|\eta|\sim 2^i M} g(x,y) \right) \bar{P}_{|\eta|\sim 2^i M} \bar{P}_{|\xi|\lesssim M} h \right| \\
&\leq \int |M^\alpha v_M(x-y))| \left(\sum_{i=1}^\infty |P_{|\xi|<M} (2^i M)^\beta P_{|\eta|\sim 2^i M} g|^2 \right)^{\frac{1}{2}} \\
&\times \left(\sum_{i=1}^\infty |\bar{P}_{|\eta|\sim 2^i M} \bar{P}_{|\xi|\lesssim M} h|^2 \right)^{\frac{1}{2}} \\
&\leq \|M^\alpha v_M\|_{L^{p_1}} \left\| \left(\sum_{i=1}^\infty |P_{|\xi|<M} |\nabla_y|^\beta \bar{P}_{|\eta|\sim 2^i M} g|^2 \right)^{\frac{1}{2}} \right\|_{L^{q_1}(d(x-y))L^2(d(x+y))} \\
&\left\| \left(\sum_{i=1}^\infty |\bar{P}_{|\eta|\sim 2^i M} \bar{P}_{|\xi|\lesssim M} h|^2 \right)^{\frac{1}{2}} \right\|_{L^{r'}(d(x-y))L^2(d(x+y))}
\end{aligned}
$$

and the last factor is $\lesssim 1$.

The bound (48) corresponds to

$$
\begin{aligned}
&\| |\nabla_x|^\alpha |\nabla_y|^\beta \left(v_M(x-y)) P_{|\xi|>M} P_{|\eta|>M} g(x,y) \right) \|_{L^r(d(x-y))L^2(d(x+y))} \\
&\lesssim \|v_M\|_{L^{p_2}} \| |\nabla_x|^\alpha |\nabla_y|^\beta g \|_{L^{q_2}(d(x-y))L^2(d(x+y))}.
\end{aligned}
$$

To prove this, do a double Littlewood-Paley decomposition $P_{|\xi|>M} P_{|\eta|>M} g = \sum_{i=1}^\infty \sum_{j=1}^\infty P_{|\eta|\sim 2^i M} P_{|\eta|\sim 2^j M} g$ and proceed as before, using "double square function estimates in rotated coordinates", Lemma 10.8. Finally, the bound (49) follows from Bernstein's inequality.

□

Continuing with the comments preceding the proof of Theorem 2.1,

$$\begin{aligned}
&\||\nabla_x|^\alpha \Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&\lesssim M^\alpha \|v_M\|_{L^{\frac{6}{5}}}\|\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+ M^\alpha \|v_M\|_{L^{\frac{6}{5}}}\|H\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+ \|v_M\|_{L^{\frac{3}{2}}}\|\langle\nabla_x\rangle^\alpha \Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+ \|v_M\|_{L^{\frac{3}{2}}}\|\langle\nabla_x\rangle^\alpha H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+ \|\langle\nabla_x\rangle^\alpha G\|_{\mathcal{S}'_r} + \|\langle\nabla_x\rangle^\alpha \Lambda_0\|_{L^2}.
\end{aligned}$$

Using the "Sobolev at an angle" estimate (Lemma 3.2)

$$\|\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim \|\langle\nabla_x\rangle^\alpha \Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}$$

and similarly for H. Using $\|v_M\|_{L^{\frac{3}{2}}} + M^\alpha\|v_M\|_{L^{\frac{6}{5}}} = O(M^{-\epsilon_0})$ as $M \to \infty$ we can treat the two terms involving the potential and Λ as perturbations and get

$$\begin{aligned}
&\|\langle\nabla_x\rangle^\alpha \Lambda\|_{\mathcal{S}} \lesssim \|\langle\nabla_x\rangle^\alpha \Lambda_0\|_{L^2}\\
&+ M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha H\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^\alpha G\|_{\mathcal{S}'_r}.
\end{aligned} \tag{51}$$

Finally, we can repeat the argument with $|\nabla_x|^\alpha|\nabla_y|^\alpha$. Now we are forced to estimate $M^{2\alpha}\|v_M\|_{L^{\frac{6}{5}}} = O(M)$ and get a sub-optimal estimate

$$\frac{1}{M}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha \Lambda\|_{\mathcal{S}} \lesssim \frac{1}{M}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha \Lambda_0\|_{L^2} \tag{52}$$
$$+ M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} + \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\|_{\mathcal{S}'_r}. \tag{53}$$

This will help control lower order terms.

To continue, we need a frequency 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 $|\xi| > 2$.

Theorem 2.1 follows from the next two more detailed theorems.

Theorem 5.3. *Let Λ satisfy* (25), *and let* $1+$ *denote* $1+\delta_0$ *with* $\delta_0 > 0$ *satisfying* (28), (30). *Then, at high frequencies,*

$$
\begin{aligned}
&\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{\mathcal{S}}+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\eta|>M^{1+}}\Lambda\right\|_{\mathcal{S}}\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\eta|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac14} P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac14} P_{|\eta|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha G\right\|_{\mathcal{S}_r'}+M^{-\epsilon_0}\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Lambda_0\right\|_{L^2}.
\end{aligned}\tag{54}
$$

In addition, the proof will show

$$
\begin{aligned}
&\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Big(v_M(x-y)\big(P_{|\xi|>M^{1+}}\Lambda\big)\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&\lesssim\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha G\right\|_{\mathcal{S}_r'}+M^{-\epsilon_0}\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Lambda_0\right\|_{L^2}.
\end{aligned}
$$

Proof. Roughly speaking, $\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\xi|>M^{1+}}\big(v_M^1(x-y)\Lambda\big)\right\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}$ can be treated as a perturbation because $\left\langle\nabla_x\right\rangle^\alpha$ only falls on Λ. Rigorously, we have

$$
\mathbf{S}P_{|\xi|>M^{1+}}\Lambda=P_{|\xi|>M^{1+}}\Big(v_M^1(x-y)\big(P_{|\xi|>M^{1+}-M}\Lambda\big)\Big)\tag{55}
$$

$$
+P_{|\xi|>M^{1+}}\Big(v_M^2(x-y)\Lambda\Big)+P_{|\xi|>M^{1+}}G+P_{|\xi|>M^{1+}}(v_MH).\tag{56}
$$

We used the fact that $\hat{v}_M^1$ is supported in $|\xi|<\frac{M}{10}$. Next we use the Strichartz estimate of Theorem 4.1 and the collapsing estimate of Lemma 4.5 and Lemma 4.6.

$$\begin{aligned}
&\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha P_{|\xi|>M^{1+}}\Lambda\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\big\|\big|\nabla_{x+y}\big|^\alpha P_{|\xi|>M^{1+}}\Lambda\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\big\|\big|\partial_t\big|^{\frac14} P_{|\xi|>M^{1+}}\Lambda\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}
\end{aligned}$$

$$\begin{aligned}
&\lesssim \|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^1(x-y)\big(P_{|\xi|>M^{1+}-M}\Lambda\big)\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&+\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^2(x-y)\Lambda\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&+\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^1(x-y)\big(P_{|\xi|>M^{1+}-M}H\big)\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&+\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^2(x-y)H\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&+\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha G\big\|_{\mathcal{S}'_r}+\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Lambda_0\big\|_{L^2}.
\end{aligned}$$

For the terms involving v_M^1 we use Theorem 5.2 and the fact $M^{1+}-M>10M$ to conclude that the $\big\langle\nabla_x\big\rangle^\alpha$ derivative only falls on $P_{|\xi|>M^{1+}-M}\Lambda$ and $P_{|\xi|>M^{1+}-M}H$. Then Hölder's inequality, "Sobolev at an angle" (see Lemma 3.2) and our estimates on v_M (see (28)) show

$$\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^1(x-y)\big(P_{|\xi|>M^{1+}-M}\Lambda\big)\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))} \tag{57}$$

$$\lesssim \|v_M\|_{L^{\frac32}}\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\big(P_{|\xi|>M^{1+}-M}\Lambda\big)\Big)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \tag{58}$$

$$+M^\alpha\|v_M\|_{L^{\frac65+}}\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\big(P_{|\xi|>M^{1+}-M}\Lambda\big)\Big)\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} \tag{59}$$

$$\lesssim M^{-\epsilon_0}\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha P_{|\xi|>M^{1+}-M}\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}. \tag{60}$$

Similarly,

$$\begin{aligned}
&\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(v_M^1(x-y)\big(P_{|\xi|>M^{1+}-M}H\big)\Big)\|_{L^2(dt)L^{\frac65}(d(x-y))L^2(d(x+y))}\\
&\lesssim M^{-\epsilon_0}\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha H\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}
\end{aligned}$$

while for the term involving v_M^2 we have

$$\begin{aligned}
&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left(v_M^2(x-y)\Lambda\right)\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\
&\lesssim M^{-10}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}
\end{aligned}$$

(and similarly for H). Thus

$$\begin{aligned}
&\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{\mathcal{S}}\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac{1}{4}}P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim M^{-\epsilon_0}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|>M^{1+}-M}\Lambda\|_{\mathcal{S}}\\
&+M^{-10}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\right\|_{\mathcal{S}_r'}\\
&+M^{-\epsilon_0}\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\right\|_{L^2}.
\end{aligned}$$

The last four terms on the RHS are acceptable, but the first one must be estimated further, by repeating the argument(with the same implicit constants in $\lesssim$) as long as $M^{1+}-kM>10M$, which is essentially $\log\left(\frac{M^{1+}}{M}\right)$ times. At the kth step we get

$$\begin{aligned}
&\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|>M^{1+}-(k-1)M}\Lambda\right\|_{\mathcal{S}}\\
&\lesssim M^{-\epsilon_0}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|>M^{1+}-kM}\Lambda\|_{\mathcal{S}}\\
&+M^{-10}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\right\|_{\mathcal{S}_r'}+M^{-\epsilon_0}\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\right\|_{L^2}.
\end{aligned}$$

Putting together the above k estimates we get

$$\begin{aligned}
&\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{\mathcal{S}}+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\eta|>M^{1+}}\Lambda\right\|_{\mathcal{S}}\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\eta|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac{1}{4}} P_{|\xi|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac{1}{4}} P_{|\eta|>M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim M^{-k\epsilon_0}\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\xi|>M^{1+}-kM}\Lambda\|_{\mathcal{S}}\\
&+M^{-10}\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Lambda\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha G\right\|_{\mathcal{S}_r'}+M^{-\epsilon_0}\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha H\right\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}\\
&+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Lambda_0\right\|_{L^2}.
\end{aligned}$$

Once $k\epsilon_0>1$, we use (53) to complete the proof. □

Now we move to the low frequency part. Here the collapsing norm can be treated perturbatively.

Theorem 5.4. *Let* Λ *satisfy* (25), *and let* $1+$ *denote* $1+\delta_0$ *with* $\delta_0>0$ *satisfying* (28), (30). *Then*

$$\begin{aligned}
&\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{\mathcal{S}_{x,y}} &(61)\\
&+\left\|\left|\nabla_{x+y}\right|^\alpha P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&+\left\|\left|\partial_t\right|^{\frac{1}{4}} P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\
&\lesssim\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha G\right\|_{\mathcal{S}_r'}+M^{-\epsilon_0}\left\|\left|\nabla_{x+y}\right|^\alpha H\right\|_{\textit{collapsing}}\\
&+M^{-\epsilon_0}\left\|\left|\partial_t\right|^{\frac{1}{4}} H\right\|_{\textit{collapsing}}+\left\|\left\langle\nabla_x\right\rangle^\alpha\left\langle\nabla_y\right\rangle^\alpha\Lambda_0\right\|_{L^2}. &(62)
\end{aligned}$$

We first prove the above theorem without including $\mathcal{S}_{x,y}$ on the LHS.

Proof. (excluding the term (61))

$$\begin{aligned}
&\mathbf{S}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\\
&=P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M(x-y)(\Lambda+H)\right)+P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)G\\
&=P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^1(x-y)P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}(\Lambda+H)\right)\\
&+P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^1(x-y)P_{|\xi|\,or\,|\eta|>M^{1+}}(\Lambda+H)\right)\\
&+P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)G\\
&+P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^2(x-y)(\Lambda+H)\right).
\end{aligned}$$

We write

$$\begin{aligned}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda &= \Lambda_1+\Lambda_2+\Lambda_3+\Lambda_4+\Lambda_5\\ &= P_{|\xi|<2M^{1+}}P_{|\eta|<2M^{1+}}\left(\Lambda_1+\Lambda_2+\Lambda_3+\Lambda_4+\Lambda_5\right)\end{aligned}$$

where $\Lambda_1, \cdots \Lambda_5$ are defined by

$$\mathbf{S}\Lambda_1 = P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^1(x-y)P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}(\Lambda+H)\right)$$

$$\mathbf{S}\Lambda_2 = P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^1(x-y)P_{|\xi|\, or\, |\eta|>M^{1+}}P_{|\xi|<10M^{1+}}P_{|\eta|<10M^{1+}}(\Lambda+H)\right)$$

$$\mathbf{S}\Lambda_3 = h(t)P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}G$$

with initial conditions 0, and $\mathbf{S}\Lambda_4 = 0$ with initial conditions $P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda_0$, and finally

$$\mathbf{S}\Lambda_5 = h(t)P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\left(v_M^2(x-y)(\Lambda+H)\right)$$

with initial conditions 0. Putting together the five propositions below, we conclude

$$\begin{aligned}&\left\||\partial_t|^{\frac14}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\ &+\left\||\nabla_{x+y}|^{\alpha}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\ &\lesssim M^{-\epsilon_0}\left(\left\||\partial_t|^{\frac14}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{collapsing}+\left\||\nabla_{x+y}|^{\alpha}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\Lambda\right\|_{collapsing}\right)\\ &+M^{-\epsilon_0}\left(\left\||\partial_t|^{\frac14}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}H\right\|_{collapsing}+\left\||\nabla_{x+y}|^{\alpha}P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}H\right\|_{collapsing}\right)\\ &+\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}G\right\|_{\mathcal{S}'_r}+\left\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Lambda_0\right\|_{L^2}.\end{aligned}$$

The first term on the RHS can be absorbed in the LHS, proving the result. □

We have to prove the following propositions

Proposition 5.5. *Let*

$$\mathbf{S}\Lambda_1 = P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}h(t)\left(v_M^1(x-y)P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}(\Lambda+H)\right)$$

with zero initial conditions. Then

$$\|\Lambda_1\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\left\|P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}(\Lambda+H)\right\|_{collapsing}$$

and similarly

$$\||\nabla_{x+y}|^{\alpha}\Lambda_1\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\left\|P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}|\nabla_{x+y}|^{\alpha}(\Lambda+H)\right\|_{collapsing}$$

$$\||\partial_t|^{\frac14}\Lambda_1\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\left\|P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}|\partial_t|^{\frac14}(\Lambda+H)\right\|_{collapsing}.$$

Since the frequency localization of $\Lambda + H$ plays no role and only the frequency localization of $v_M^1(x-y)\Lambda$ is important, we record a slightly more general result, which implies Proposition 5.5 and will also be used later.

Proposition 5.6. *Let*

$$\mathbf{S}u = P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} h(t) \left(v_M(x-y)\Lambda\right)$$

with zero initial conditions. Then

$$\|u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\|\Lambda\|_{collapsing}$$

and similarly

$$\||\nabla_{x+y}|^\alpha u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\left\||\nabla_{x+y}|^\alpha \Lambda\right\|_{collapsing}$$

$$\||\partial_t|^{\frac{1}{4}} u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\left\||\partial_t|^{\frac{1}{4}} \Lambda\right\|_{collapsing}.$$

Proof. Using Sobolev estimates and Theorem 4.8 we get

$$\|u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim\lesssim M^{-\epsilon_0}\|\Lambda\|_{collapsing}.$$

No modifications are needed to prove

$$\||\nabla_{x+y}|^\alpha u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\||\nabla_{x+y}|^\alpha \Lambda\|_{collapsing}.$$

However, the estimate for $|\partial_t|^{\frac{1}{4}}u$ requires extra care because time derivatives don't preserve initial conditions. To simplify notation, let $F = P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} \left(v_M(x-y)\Lambda\right)$. Let E the fundamental solution of the Schrödinger equation supported in $t \geq 0$, and recall h is the Heaviside function. The usual solution to

$$\mathbf{S}u = F$$

with 0 initial conditions is given (in the region $t > 0$) by

$$u = E * (hF).$$

From the first part of this proof we get

$$\||\partial_t|^{\frac{1}{4}} u\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim M^{-\epsilon_0}\||\partial_t|^{\frac{1}{4}}(hF)\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}.$$

It is easy to show

$$\||\partial_t|^{\frac{1}{4}}(hF)\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))} \lesssim \|(|\partial_t|^{\frac{1}{4}}F)\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}. \tag{63}$$

This can be done by taking Fourier transform in t and using A_2 theory (see [17]), or else the equivalent definition (for $0 < k < 1$):

$$\|u\|_{\dot{H}^k}^2 = \int\int \frac{|u(t)-u(s)|^2}{|t-s|^{1+2k}} dt ds \tag{64}$$

and the generalized Hardy's inequality from [21].

We remark that the corresponding estimate for the stronger norm $\|\Lambda\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$ might not be true. We do not know if

$$\||\partial_t|^{\frac{1}{4}}(hF)\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim \|(|\partial_t|^{\frac{1}{4}}F)\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \tag{65}$$

is true.

□

Proposition 5.7. *Recall* Λ *satisfies* (25) *and*

$$\mathbf{S}\Lambda_2 = P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\left(h(t)v_M^1(x-y)P_{|\xi|\text{ or }|\eta|>M^{1+}}P_{|\xi|<10M^{1+}}P_{|\eta|<10M^{1+}}(\Lambda+H)\right)$$

with zero initial conditions. Then

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_2\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_2\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\ &+ \||\partial_t|^{\frac{1}{4}}\Lambda_2\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\ &\lesssim M^{-\epsilon_0}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\|_{\mathcal{S}'_r}\\ &+ M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))} + M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\big\|_{L^2}.\end{aligned}$$

Proof. Consider (by slight abuse of notation) just one of the contributions to Λ_2 with $P_{|\xi|>M^{1+}}$.

$$\mathbf{S}\Lambda_2 = P_{|\xi|<M^{1+}}P_{|\eta|<M^{1+}}\left(v_M^1(x-y)P_{M^{1+}<|\xi|<10M^{1+}}P_{|\eta|<10M^{1+}}(\Lambda+H)\right). \tag{66}$$

Using Theorem 4.1 and Lemmas 4.5, 4.6 we have

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_2\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_2\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\ &+ \||\partial_t|^{\frac{1}{4}}\Lambda_2\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} \lesssim \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha RHS\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}.\end{aligned}$$

Here we can estimate the RHS directly (using the fractional Leibniz rule and the fact that $\langle\nabla_x\rangle^\alpha$ only falls on $\Lambda+H$)

$$\begin{aligned}&\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left(v_M^1(x-y)P_{|\xi|>M^{1+}}(\Lambda+H)\right)\big\|_{L^2(dt)L^{\frac{6}{5}}(d(x-y))L^2(d(x+y))}\\ &\lesssim M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|>M^{1+}}(\Lambda+H)\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned}$$

This is in the high frequency range, and using Theorem 5.3, the above is

$$\begin{aligned}&\lesssim M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\big\|_{\mathcal{S}'_r} + M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\big\|_{L^2}\\ &+ M^{-\epsilon_0}\big\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha H\big\|_{L^2(dt)L^6(d(x-y))L^2(d(x+y))}.\end{aligned}$$

□

Proposition 5.8. *Recall*

$$\mathbf{S}\Lambda_3 = P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} G$$

with zero initial conditions. Then

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_3\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_3\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&+\||\partial_t|^{\frac{1}{4}}\Lambda_3\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&\lesssim \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\|_{\mathcal{S}_r'}.\end{aligned}$$

Proof. This follows immediately from Theorem 4.1 and Lemmas 4.5, 4.6. □

Proposition 5.9. *Recall*

$$\mathbf{S}\Lambda_4 = 0$$

with initial conditions Λ_0. *Then, for any* $\alpha > \frac{1}{2}$,

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_4\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_4\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} + \||\partial_t|^{\frac{1}{4}}\Lambda_4\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&\lesssim \left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\right\|_{L^2}.\end{aligned}$$

Proof. This follows from Strichartz estimates, see for instance the proof of Lemma 5.3 in [11]. □

Proposition 5.10. *Recall* Λ *satisfies* (25) *and*

$$\mathbf{S}\Lambda_5 = P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} \left(v_M^2(x-y)(\Lambda+H)\right)$$

with initial conditions 0. *Then*

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_5\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_5\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&+\||\partial_t|^{\frac{1}{4}}\Lambda_5\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&\lesssim M^{-9}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\|_{\mathcal{S}_r'} + M^{-9}\left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\right\|_{\mathcal{S}_r'}\end{aligned}$$

Proof. Using Proposition 5.8, we have

$$\begin{aligned}&\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_5\|_{\mathcal{S}} + \|\langle\nabla_{x+y}\rangle^\alpha\Lambda_5\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))} + \||\partial_t|^{\frac{1}{4}}\Lambda_5\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}\\&\lesssim \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}}\left(v_M^2(x-y)\Lambda\right)\|_{\mathcal{S}_r'}.\end{aligned}$$

Here we use the Leibniz rule and (51), (53) and the smallness of v_M^2 to conclude the above is

$$\lesssim M^{-9}\left(\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha G\|_{\mathcal{S}_r'} + \left\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_0\right\|_{L^2}\right).$$

□

To finish the proof the Theorem 2.1, we also need

Theorem 5.11. *Let*

$$\mathbf{S}\Lambda_1 = P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} \left(v_M^1(x-y) P_{|\xi|<M^{1+}} P_{|\eta|<M^{1+}} \Lambda \right)$$

with 0 *initial conditions. Then*

$$\begin{aligned} &\left\| |\nabla_x|^\alpha |\nabla_y|^\alpha P_{|\xi|<2M^{1+}} P_{|\eta|<2M^{1+}} \Lambda_1 \right\|_{\mathcal{S}_{x,y}} \\ &\lesssim \left\| \langle \nabla_{x+y} \rangle^\alpha \Lambda \right\|_{collapsing} + \left\| \langle \partial_t \rangle^{\frac{1}{4}} \Lambda \right\|_{collapsing}. \end{aligned}$$

Proof. This follows from Proposition 4.8 and Bernstein's inequality. □

6. Estimates for the nonlinear equation, step 1

Recall the notation

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

From now on, $V_N(x) = N^{3\beta} v(N^\beta x)$. As in [5] we assume the potential satisfies (10) and the initial conditions satisfy (11) and (13), and also (12) holds.

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 4 relevant equations are

$$\begin{aligned} &\mathbf{S}\Lambda_p + \{V_N * \rho, \Lambda_p\} + \frac{V_N}{N}\Lambda_p && (67) \\ &+ \left((V_N \bar{\Gamma}_p) \circ \Lambda_p + (V_N \Lambda_p) \circ \Gamma_p \right)_{symm} \\ &+ \left((V_N \bar{\Gamma}_c) \circ \Lambda_p + (V_N \Lambda_c) \circ \Gamma_p \right)_{symm} = -\frac{V_N}{N}\Lambda_c \\ &\mathbf{S}_\pm \Gamma_p + [V_N * \rho, \Gamma_p] + \left((V_N \Gamma_p) \circ \Gamma_p + (V_N \bar{\Lambda}_p) \circ \Lambda_p \right)_{skew} \\ &+ \left((V_N \Gamma_c) \circ \Gamma_p + (V_N \bar{\Lambda}_c) \circ \Lambda_p \right)_{skew} = 0 && (68) \\ &\mathbf{S}\Lambda_c + \{V_N * \rho, \Lambda_c\} + \left((V_N \bar{\Gamma}_p) \circ \Lambda_c + (V_N \Lambda_p) \circ \Gamma_c \right)_{symm} = 0 && (69) \\ &\mathbf{S}_\pm \Gamma_c + [V_N * \rho, \Gamma_c] + \left((V_N \Gamma_p) \circ \Gamma_c + (V_N \bar{\Lambda}_p) \circ \Lambda_c \right)_{skew} = 0. && (70) \end{aligned}$$

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

The norms used for Λ_p and Λ_c are called $\mathcal{N}(\Lambda)$ and are

$$\begin{aligned} &\|\Lambda\|_{\mathcal{N}(\Lambda)} = \left\| \langle \nabla_x \rangle^\alpha \langle \nabla_y \rangle^\alpha \Lambda \right\|_{\mathcal{S}_{x,y}} \\ &+ \left\| |\partial_t|^{\frac{1}{4}} \Lambda \right\|_{L^2(dt) L^\infty(d(x-y)) L^2(d(x+y))} + \left\| \langle \nabla_{x+y} \rangle^\alpha \Lambda \right\|_{L^2(dt) L^\infty(d(x-y)) L^2(d(x+y))}. \end{aligned}$$

Because of the quarter time derivative, these norms cannot be localized in an obvious way, and we will devise a way to get around that. The norms for Γ, Γ_p, Γ_c and also Λ_c are

$$\begin{aligned}\|F\|_{\mathcal{N}^1} = &\sup_{p,q \text{ admissible}} \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha F\|_{L^p(dt)L^q(dx)L^2(dy)}\\ &+ \sup_{p,q \text{ admissible}} \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha F\|_{L^p(dt)L^q(dy)L^2(dx)}\\ &+ \left\||\nabla_{x+y}|^\alpha F\right\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}.\end{aligned}$$

The estimates for the linear part of the Λ equations have been studied in the previous sections. For the Γ equation we will only use older, standard estimates

Proposition 6.1. *Let*

$$\begin{aligned}\mathbf{S}_\pm \Gamma &= F\\ \Gamma(0,\cdot) &= \Gamma_0\end{aligned}$$

Then

$$\begin{aligned}&\sup_{p,q \text{ admissible}} \|\Gamma\|_{L^p(dt)L^q(dx)L^2(dy)}\\ &\quad + \sup_{p,q \text{ admissible}} \|\Gamma\|_{L^p(dt)L^q(dy)L^2(dx)}\\ &\lesssim \|\Gamma_0\|_{L^2} + \inf_{p,q \text{ admissible}, p\geq p_0>2} \{\|F\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}, \|F\|_{L^{p'}(dt)L^{q'}(dy)L^2(dx)}\}\\ &\sup_z \||\nabla_x|^\alpha \Gamma(t, x+z, x)\|_{L^2(dtdx)}\\ &\lesssim \|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha \Gamma_0\|_{L^2}\\ &+ \inf_{p,q \text{ admissible}, p\geq p_0>2} \{\|\langle\nabla_x\rangle^\alpha \langle\nabla_y\rangle^\alpha F\|_{L^{p'}(dt)L^{q'}(dx)L^2(dy)}, \|F\|_{L^{p'}(dt)L^{q'}(dy)L^2(dx)}\}.\end{aligned}$$

Proof. For a proof of the homogeneous estimate, see Lemmas 5.1, 5.3 in [11], and also [4]. The inhomogeneous estimate follows from the Christ-Kiselev lemma. Let $T_1 = e^{it\Delta_\pm}$, 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 z and let $T_2 : H^\alpha(dx)H^\alpha(dy) \to L^2(dt)\dot{H}^\alpha(dx)$ be the operator $f \to \left(e^{it\Delta_\pm} f\right)(t, x, x+z)$. Then the inhomogeneous estimate follows by applying the Christ-Kiselev lemma to $T_2T_1^*$.

□

The first step in the analysis of the nonlinear equations uses a priori estimates for $\Gamma(t, x, x)$, see (14) and (15).

Lemma 6.2. *Under the assumptions of Theorem 1.1 we have*

$$\|\Gamma\|_{L^8(dt))L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \lesssim 1 \tag{71}$$

$$\|\nabla_{x+y}\Gamma\|_{L^8(dt))L^\infty(d(x-y))L^{\frac{4}{3}}(d(x+y))} \lesssim 1 \tag{72}$$

and thus, for any $\epsilon_2 > 0$ *there exist* $n = n(\epsilon_2)$ *intervals* $[T_i, T_{i+1}]$ *covering* $[0, \infty)$ *such that*

$$\sup_z \|\langle\nabla_x\rangle^\alpha \Gamma(t, x+z, x)\|_{L^8([T_j,T_{j+1}])L^{\frac{4}{3}}(dx)} \leq \epsilon_2. \tag{73}$$

The idea of proving estimates for NLS by using such a localization in time goes back to Bourgain [3].

Remark 6.3. Notice that the "collapsing estimate" $\sup_z \|\langle\nabla_x\rangle^\alpha \Gamma(t, x+z, x)\|_{L^2(dtdx)} \lesssim 1$ does not imply there exist $n = n(\epsilon_2)$ intervals $[T_i, T_{i+1}]$ covering $[0, \infty)$ such that

$$\sup_z \|\langle\nabla_x\rangle^\alpha \Gamma(t, x+z, x)\|_{L^2([T_j,T_{j+1}])L^2(dx)} \leq \epsilon_2.$$

Proof. We have a pointwise estimate

$$||\Gamma(t, x+z, x-z)| \leq |\Gamma(t, x+z, x+z)|^{\frac{1}{2}} |\Gamma(t, x-z, x-z)|^{\frac{1}{2}}$$

and also

$$||\nabla_x \Gamma(t, x+z, x-z)|$$
$$\leq |E_1(t, x+z, x+z)|^{\frac{1}{2}} |\Gamma(t, x-z, x-z)|^{\frac{1}{2}} + |E_1(t, x-z, x-z)|^{\frac{1}{2}} |\Gamma(t, x+z, x+z)|^{\frac{1}{2}}$$

where $E_1(t, x) = \nabla_x \cdot \nabla_y \Gamma(t, x, y)\big|_{x=y}$ is the kinetic energy density, with $\int |E_1(t, x)| + \Gamma(t, x, x) dx$ uniformly bounded in time and $\|\Gamma(t, x, x)\|_{L^4(dt)L^2(dx)} \lesssim 1$. (71) and (72) follow by applying Hölder's inequality, and these imply (73). □

The above estimates hold for Γ_c and Γ_p separately. Since $\Gamma_c = \bar{\phi} \otimes \phi$ and $\Lambda_c = \phi \otimes \phi$, they also hold for Λ_c. No such a priori estimates are available for Λ_p. However, we assume initial conditions for Λ_p are small. Also, the forcing term in the equation for Λ_p is $\frac{V_N}{2N}\Lambda_c$, and this is small in suitable norms, thus $\|\Lambda_p\|_{\mathcal{N}(\Lambda)}$ will stay small. Here are the details:

In order to use the smallness of the above quantities, we have to localize our estimates to these intervals. However, the necessary norms involve a non-local term, so we have to proceed carefully. Also, we will use a continuity argument, so the right end of the interval must be a variable $T \in [T_i, T_{i+1}]$. Define $\Lambda_c^{i,T}$ $\Lambda_p^{i,T}$, $\Gamma_c^{i,T}$, $\Gamma_p^{i,T}$ be the solution to the standard equations with the RHS multiplied by $\chi_{[T_i,T]}$:

$$\mathbf{S}\Lambda_p^{i,T} + \frac{V_N}{N}\Lambda_p^{i,T} \tag{74}$$
$$= \chi_{[T_i,T]}\Big(- \{V_N * \rho, \Lambda_p\} - \big((V_N\bar{\Gamma}_p)\circ\Lambda_p + (V_N\Lambda_p)\circ\Gamma_p\big)_{symm}$$
$$- \big((V_N\bar{\Gamma}_c)\circ\Lambda_p + (V_N\Lambda_c)\circ\Gamma_p\big)_{symm} - \frac{V_N}{N}\Lambda_c\Big)$$
$$\mathbf{S}_{\pm}\Gamma_p^{i,T} = \chi_{[T_i,T]}\Big(- [V_N * \rho, \Gamma_p] - \big((V_N\Gamma_p)\circ\Gamma_p - (V_N\bar{\Lambda}_p)\circ\Lambda_p\big)_{skew}$$
$$- \big((V_N\Gamma_c)\circ\Gamma_p - (V_N\bar{\Lambda}_c)\circ\Lambda_p\big)_{skew}\Big) \tag{75}$$
$$\mathbf{S}\Lambda_c^{i,T} = \chi_{[T_i,T]}\Big(- \{V_N * \rho, \Lambda_c\} - \big((V_N\bar{\Gamma}_p)\circ\Lambda_c - (V_N\Lambda_p)\circ\Gamma_c\big)_{symm}\Big) \tag{76}$$
$$\mathbf{S}_{\pm}\Gamma_c^{i,T} = \chi_{[T_i,T]}\Big(- [V_N * \rho, \Gamma_c] - \big((V_N\Gamma_p)\circ\Gamma_c - (V_N\bar{\Lambda}_p)\circ\Lambda_c\big)_{skew}\Big) \tag{77}$$

with $\Lambda_c^{i,T}(T_i,\cdot) = \Lambda_c^{i-1,T_i}(T_i,\cdot) = \Lambda_c(T_i,\cdot)$, and similarly for the other three functions. Also, Then $\Lambda_c^{i,T} = \Lambda_c$, etc. in $[T_i,T]$ (but not outside this interval), and similarly for the other three functions. Also, Λ_c^{i,T_i}, etc. satisfies a homogeneous linear equation.

We continue by estimating the four functions on the LHS. The most difficult one is $\Lambda_p^{i,T}$.

Later it will be convenient to have norms which can be made small on small time intervals, so we introduce the restricted Strichartz norms

$$\|F\|_{\mathcal{S}^r_{x,y}} = \|F\|_{L^2(dt)L^6(dx)L^2(dy)} + \|F\|_{L^2(dt)L^6(dy)L^2(dx)}$$
$$+ \|F\|_{L^4(dt)L^3(dx)L^2(dy)} + \|F\|_{L^4(dt)L^3(dx)L^2(dy)}.$$

Notice this is not dual to $\mathcal{S}'_r$, the restrictions on the exponents are different.

Theorem 6.4. *Let $[T_j,T]$, $T_i \le T \le T_{i+1}$, and ϵ_2 be as in* (73)*. There exists a universal constant C such that*

$$\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)} \le C\big\|\langle\nabla_x\rangle^{\alpha}\langle\nabla_y\rangle^{\alpha}\Lambda_p^{i,T}(T_j,\cdot)\big\|_{L^2} + C\epsilon_2\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}$$
$$+ C\epsilon_2\|\Gamma_p^{i,T}\|_{\mathcal{N}^1} + C\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}\|\Gamma_p^{i,T}\|_{\mathcal{S}^r_{x,y}}$$
$$+ CN^{-\epsilon_1}\|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)}.$$

Remark 6.5. Of course $\|\Gamma_p^{i,T}\|_{\mathcal{S}^r_{x,y}} \leq \|\Gamma_p^{i,T}\|_{\mathcal{N}^1}$, but $\|\Gamma_p^{i,T}\|_{\mathcal{S}^r_{x,y}}$ can also be made small on small time intervals. This will be useful when estimating higher order derivatives.

The proof is based on Theorem 2.1: there exists a constant C and $\epsilon_1 > 0$ such that

$$\begin{aligned}
&\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)} \\
&\leq C\bigg(\big\| \big\langle \nabla_x \big\rangle^\alpha \big\langle \nabla_y \big\rangle^\alpha \chi_{[T_i,T]} \big(\{V_N * \rho, \Lambda_p\} + \big((V_N \bar{\Gamma}_p) \circ \Lambda_p + (V_N \Lambda_p) \circ \Gamma_p \big)_{symm} \\
&+ \big((V_N \bar{\Gamma}_c) \circ \Lambda_p + (V_N \Lambda_c) \circ \Gamma_p \big)_{symm} \big) \big\|_{\mathcal{S}'_r} \\
&+ N^{-\epsilon_1} \big\| \big\langle \nabla_x \big\rangle^\alpha \big\langle \nabla_y \big\rangle^\alpha \chi_{[T_i,T]} \Lambda_c \big\|_{L^2(dt) L^6(d(x-y)) L^2(d(x+y))} \\
&+ N^{-\epsilon_1} \big\| \big|\partial_t\big|^{\frac{1}{4}} \chi_{[T_i,T]} \Lambda_c \big\|_{L^\infty(d(x-y)) L^2(dt) L^2(d(x+y))} \qquad (78) \\
&+ N^{-\epsilon_1} \big\| \big|\nabla_{x+y}\big|^\alpha \chi_{[T_i,T]} \Lambda_c \big\|_{L^\infty(d(x-y)) L^2(dt) L^2(d(x+y))} \\
&+ \big\| \big\langle \nabla_x \big\rangle^\alpha \big\langle \nabla_y \big\rangle^\alpha \Lambda_p(T_i) \big\|_{L^2}. \bigg)
\end{aligned}$$

For all terms other than (78), the superscript i, T can be trivially added to Λ, Γ on the RHS. In (78),

$$\begin{aligned}
&\big\| \big|\partial_t\big|^{\frac{1}{4}} \big(\chi_{[T_i,T]} \Lambda_c \big) \big\|_{L^\infty(d(x-y)) L^2(\mathbb{R}) L^2(d(x+y))} \\
&= \big\| \big|\partial_t\big|^{\frac{1}{4}} \big(\chi_{[T_i,T]} \Lambda_c^{i,T} \big) \big\|_{L^\infty(d(x-y)) L^2(\mathbb{R}) L^2(d(x+y))} \\
&\lesssim \big\| \big|\partial_t\big|^{\frac{1}{4}} \Lambda_c^{i,T} \big\|_{L^\infty(d(x-y)) L^2(\mathbb{R}) L^2(d(x+y))} \\
&\text{(as explained for (63))} \\
&\leq \big\| \big|\partial_t\big|^{\frac{1}{4}} \Lambda_c^{i,T} \big\|_{L^2(dt) L^\infty(d(x-y)) L^2(d(x+y))} \\
&\leq \|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)}.
\end{aligned}$$

In the lemmas that follow, we estimate the norms of the nonlinear terms in suitable dual Strichartz norms, using the bound (73) whenever possible.

Lemma 6.6. *Let $[T_i, T]$ be as above. There exists a universal constant C such that*

$$\begin{aligned}
&\big\| \big\langle \nabla_x \big\rangle^\alpha \big\langle \nabla_y \big\rangle^\alpha \bigg(\{V_N * \rho, \Lambda_p^{i,T}\} + V_N \bar{\Gamma}^{i,T} \circ \Lambda_p^{i,T} \bigg) \big\|_{L^{\frac{8}{5}}([T_i,T]) L^{\frac{4}{3}}(dx) L^2(dy)} \\
&\leq C\epsilon_2 \big\| \big\langle \nabla_x \big\rangle^\alpha \big\langle \nabla_y \big\rangle^\alpha \Lambda_p^{i,T} \big\|_{\mathcal{S}^r_{x,y}}.
\end{aligned}$$

The result depends on the a priori bounds for Γ*, but is true with* $\Lambda_p^{i,T}$ *replaced with any other function).*

Proof. The proof is easily reduced to estimating

$$\sup_z \|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\big(\Gamma(t,x,x+z)\Lambda_p^{i,T}(t,x+z,y)\big)\|_{L^{\frac{8}{5}}([T_i,T])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq C\epsilon_2\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_p^{i,T}\|_{\mathcal{S}_{x,y}}.$$

Using the fractional Leibniz rule (see the proof of Theorem 5.1) we have the following estimate, uniformly in z:

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\big(\Gamma(t,x,x+z)\Lambda_p^{i,T}(t,x+z,y)\big)\|_{L^{\frac{8}{5}}([T_i,T])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq C\|\langle\nabla_x\rangle^\alpha\Gamma(t,x,x+z)\|_{L^8([T_i,T])L^{\frac{4}{3}}(dx)}\|\langle\nabla_y\rangle^\alpha\Lambda_p^{i,T}\|_{L^2(dt)L^\infty(dx)L^2(dy)}$$
$$+\|\Gamma(t,x,x+z)\|_{L^8([T_i,T])L^{\frac{12}{7}}(dx)}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_p^{i,T}\|_{L^2(dt)L^6(dx)L^2(dy)}$$
$$\leq C\|\langle\nabla_x\rangle^\alpha\Gamma(t,x,x+z)\|_{L^8([T_i,T])L^{\frac{4}{3}}(dx)}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_p^{i,T}\|_{L^2(dt)L^6(dx)L^2(dy)}$$
$$\leq C_2\epsilon_2\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Lambda_p^{i,T}\|_{L^2(dt)L^6(dx)L^2(dy)}.$$

□

Since Λ_c satisfies the same a priori estimates (based on interaction Morawetz and conservation of energy) as Γ, by the exact same argument we get

Lemma 6.7. *Let* $[T_i,T]$ *be as above. There exists a universal constant* C *such that*

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left((V_N\Lambda_c)\circ\Gamma_p^{i,T}\right)\|_{L^{\frac{8}{5}}([T_i,T])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq C\epsilon_2\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Gamma_p^{i,T}\|_{\mathcal{S}_{x,y}}.$$

The result depends on the a priori bounds for Λ_c*, but is true for any function* $\Gamma_p^{i,T}=\Gamma_p^{i,T}(t,x,y)$*.*

We continue estimating nonlinear terms.

Lemma 6.8. *There exists a universal constant* C *such that*

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left((V_N\Lambda_p^{i,T})\circ\Gamma_p^{i,T}\right)\|_{L^{\frac{4}{3}}(dt)L^{\frac{3}{2}}(dx)L^2(dy)}$$
$$\leq C\|\langle\nabla_{x+y}\rangle^\alpha\Lambda_p^{i,T}\|_{L^\infty(d(x-y))L^2(dt)L^2(dx)}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Gamma_p^{i,T}(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}.$$

This result can be localized to any time interval $[T_i,T_{i+1}]$ *and is true for any two functions, not just* $\Lambda_p^{i,T}$ *and* $\Gamma_p^{i,T}$*.*

Proof. It suffices to estimate

$$\sup_z\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left(\Lambda_p^{i,T}(t,x,x+z)\Gamma_p^{i,T}(t,x+z,y)\right)\|_{L^{\frac{4}{3}}(dt)L^{\frac{3}{2}}(dx)L^2(dy)}.$$

The following holds, uniformly in z:

$$\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\left(\Lambda_p^{i,T}(t,x,x+z)\Gamma_p^{i,T}(t,x+z,y)\right)\|_{L^{\frac{4}{3}}(dt)L^{\frac{3}{2}}(dx)L^2(dy)}$$
$$\lesssim \|\big\langle \nabla_x\big\rangle^{\alpha}\Lambda_p^{i,T}(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p^{i,T}(t,x,y)\|_{L^4(dt)L^6(dx)L^2(dy)}$$
$$+\|\Lambda_p^{i,T}(t,x,x+z)\|_{L^2(dt)L^3(dx)}\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p^{i,T}(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}$$
$$\lesssim \|\big\langle \nabla_x\big\rangle^{\alpha}\Lambda_p^{i,T}(t,x,x+z)\|_{L^2(dt)L^2(dx)}\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p^{i,T}(t,x,y)\|_{L^4(dt)L^3(dx)L^2(dy)}.$$

□

We continue with estimates for $\|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)}$. This is an easy version of the previous theorem. Using the previous lemmas and the trivial version of Theorem 2.1 (without the potential term) we get

Theorem 6.9. *Let $[T_i, T]$ be as above. There exists a universal constant C such that*

$$\|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)} \le C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Lambda_c^{i,T}(T_i,\cdot)\big\|_{L^2} + C\epsilon_2\|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)}$$
$$+ C\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}\|\Gamma_c^{i,T}\|_{\mathcal{S}^{r}_{x,y}}.$$

Using Strichartz estimates for $\mathbf{S}_{\pm}$, and Lemmas 6.6-6.8 we get

Theorem 6.10. *Let $[T_i, T]$ be as above. There exists a universal constant C such that*

$$\|\Gamma_c^{i,T}\|_{\mathcal{N}^1} \le C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_c^{i,T}(T_i,\cdot)\big\|_{L^2} + C\epsilon_2\|\Gamma_c^{i,T}\|_{\mathcal{N}^1}$$
$$+ C\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}\|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)}$$
$$\|\Gamma_p\|_{\mathcal{N}^1} \le C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p^{i,T}(T_i,\cdot)\big\|_{L^2} + C\epsilon_2\|\Gamma_p^{i,T}\|_{\mathcal{N}^1}$$
$$+ C\epsilon_2\|\Lambda_p^{i,T}\|_{\mathcal{N}^1} + C\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}\|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)}.$$

At this stage we take $C\epsilon_2 < \frac{1}{2}$. This determines the number of intervals in the list T_i, T_{i+1}. Call that number n, and notice it is independent of N. Also, for $T \in [T_i, T_{i+1}]$ denote

$$X_i(T) = \|\Lambda_c^{i,T}\|_{\mathcal{N}(\Lambda)} + \|\Gamma_c^{i,T}\|_{\mathcal{N}^1}$$
$$Y_i(T) = \|\Lambda_p^{i,T}\|_{\mathcal{N}(\Lambda)} + \|\Gamma_p^{i,T}\|_{\mathcal{N}^1}.$$

Since we trivially have bounds on $\|\Lambda_{p\ or\ c}\|_{L^\infty[0,T]H^s(dxdy)}$ and $\|\Gamma_{p\ or\ c}\|_{L^\infty[0,T]H^s(dxdy)}$ (for any s) which can grow with T and N, then we do know X_i, Y_i are continuous.

We have established the following estimate for $T_i \le T \le T_{i+1}$:

Corollary 6.11. *The functions X_i, Y_i are continuous and satisfy*

$$X_i(T) \le C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Lambda_c(T_i,\cdot)\big\|_{L^2} + C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_c(T_i,\cdot)\big\|_{L^2} + CX_i(T)Y_i(T)$$
$$Y_i(T) \le C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Lambda_p(T_i,\cdot)\big\|_{L^2} + C\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p(T_i,\cdot)\big\|_{L^2}$$
$$+ CY_i^2(T) + CN^{-\epsilon_1}X_i(T).$$

Now we can state and prove the main theorem of this section.

Theorem 6.12. *Assume* Λ, Γ *and* ϕ *are smooth solutions to the HFB system, with finite energy per particle, uniformly in* N *(see* (9)*). In particular,*

$$\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Lambda_c(0,\cdot)\big\|_{L^2}+\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_c(0,\cdot)\big\|_{L^2}\le C.$$

Assume, in addition,

$$\big\|\big\langle \nabla_x\big\rangle^{\alpha}\big\langle \nabla_y\big\rangle^{\alpha}\Gamma_p(0,\cdot)\big\|_{L^2}\le \frac{C}{N^{\epsilon_3}}$$

(ϵ_3 *as in* (12)*; the corresponding estimate for* Λ_p *holds globally in time). Let* $\epsilon>0$*. There exists* N_0 *such that, if* $N\ge N_0$*, then*

$$\|\Lambda_c\|_{\mathcal{N}(\Lambda)}+\|\Gamma_c\|_{\mathcal{N}^1}\lesssim 1$$
$$\|\Lambda_p\|_{\mathcal{N}(\Lambda)}+\|\Gamma_p\|_{\mathcal{N}^1}\le \epsilon.$$

Proof. Starting at $T_1=0$, we have $X_1(0)\le C$ and $Y_1(0)\le CN^{-\epsilon_3}$, and

$$X_1(T)\le C+CX_1(T)Y_1(T)$$
$$Y_1(T)\le CN^{-\epsilon_3}+CY_1(T)^2+CN^{-\epsilon_1}X_1(T).$$

In the second line, either $Y_1(T)\le 2CY_1(T)^2$, but by continuity we rule this out, or else $Y_1(t)\le 2(CN^{-\epsilon_3}+CN^{-\epsilon_1}X_1(t))$, and if we plug this in the first line we get

$$X_1(T)\le C+C\big(N^{-\epsilon_3}+N^{-\epsilon_1}X_1(T)\big)X_1(T).$$

If N is sufficiently large, we get $X_1(T)\le 2C$. We continue to the next interval, $[T_2,T_3]$. The argument is the same, the initial conditions for X_2 have the same bound, but the initial conditions for Y_2 have grown:

$$Y_2(T_2)\le CN^{-\epsilon_3}+CN^{-\epsilon_1}.$$

We can repeat the argument as long as $Y_i(T_i)$ is sufficiently small. This will be the case if N is sufficiently large, because n, the number of intervals, is independent of N. □

7. Higher order derivatives

Next, we refine the argument to include $x+y$ derivatives. This section uses additional smallness results. Denote $\mathcal{S}^r_{x,y}[T_1,T_2]$ the standard Strichartz norms subject to the restriction $2\le p\le p_1<\infty$ for some large $p_1<\infty$ and $t\in[T_1,T_2]$.

Lemma 7.1. *Under the assumptions of Theorem 6.12, given* $\epsilon > 0$*, we can divide* $[0, \infty)$ *into finitely many intervals (depending only on* ϵ *and the above implicit bounds, as well as those used for* (73)*) such that*

$$\begin{aligned}
&\left\| \left\langle \nabla_{x+y} \right\rangle^{\alpha} \Gamma_{p\, and\, c} \right\|_{L^\infty(d(x-y)) L^8([T_j, T_{j+1}]) L^{\frac{4}{3}}(d(x+y))} \leq \epsilon \\
&\left\| \left\langle \nabla_{x+y} \right\rangle^{\alpha} \Lambda_{p\, and\, c} \right\|_{L^2([T_j, T_{j+1}]) L^\infty(d(x-y)) L^2(d(x+y))} \leq \epsilon \\
&\left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \Gamma_{p\, and\, c} \right\|_{\mathcal{S}^r_{x,y}[T_j, T_{j+1}]} \leq \epsilon \\
&\left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \Lambda_{p\, and\, c} \right\|_{\mathcal{S}^r_{x,y}[T_j, T_{j+1}]} \leq \epsilon.
\end{aligned}$$

Proof. The first estimate is (73). Using Theorem 6.12 we know the following quantities are bounded:

$$\begin{aligned}
&\left\| \left\langle \nabla_{x+y} \right\rangle^{\alpha} \Lambda_{p\, and\, c} \right\|_{L^2(dt) L^\infty(d(x-y)) L^2(d(x+y))} \lesssim 1 \\
&\left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \Gamma_{p\, and\, c} \right\|_{\mathcal{S}_{x,y}} \lesssim 1 \\
&\left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \Lambda_{p\, and\, c} \right\|_{\mathcal{S}_{x,y}} \lesssim 1
\end{aligned}$$

and the result follows. □

Remark 7.2. However, it is not clear we can insure

$$\left\| |\nabla_{x+y}|^{\alpha} \Gamma_{p\, and\, c} \right\|_{L^\infty(d(x-y)) L^2(t \in [T_1, T_2]) L^2(d(x+y))} \leq \epsilon$$

which is why we use $L^\infty(d(x-y)) L^8(t \in [T_1, T_2]) L^{\frac{4}{3}}(d(x+y))$ for which we have an a priori estimate, and which has the same scaling.

Theorem 7.3. *Under the assumption of Theorem 6.12, if we also have*

$$\left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \nabla^j_{x+y} \Lambda_{p\, and\, c}(0, \cdot) \right\|_{L^2} + \left\| \left\langle \nabla_{x} \right\rangle^{\alpha} \left\langle \nabla_{y} \right\rangle^{\alpha} \nabla^j_{x+y} \Gamma_{p\, and\, c}(0, \cdot) \right\|_{L^2} \lesssim 1$$

for $j = 1, \cdots, j_0$*, then*

$$\|\nabla^j_{x+y} \Lambda_c\|_{\mathcal{N}(\Lambda)} + \|\nabla^j_{x+y} \Gamma_c\|_{\mathcal{N}^1} \lesssim 1 \tag{79}$$

$$\|\nabla^j_{x+y} \Lambda_p\|_{\mathcal{N}(\Lambda)} + \|\nabla^j_{x+y} \Gamma_p\|_{\mathcal{N}^1} \lesssim 1. \tag{80}$$

Proof. (Sketch) At this stage, we don't have to distinguish between Λ_p and Λ_c, or Γ_p and Γ_c and we work directly with $\Lambda = \Lambda_p + \Lambda_c$, $\Gamma = \Gamma_p + \Gamma_c$. Schematically, the equations are

$$\begin{aligned}
&\mathbf{S}\Lambda + \frac{V_N}{N}\Lambda = V_N \Lambda \circ \Gamma + V_N \Gamma \circ \Lambda \\
&\mathbf{S}_{\pm}\Gamma = V_N \Lambda \circ \Gamma + V_N \Lambda \circ \Lambda.
\end{aligned}$$

Apply ∇_{x+y} and localize the RHS to $[T_i, T_{i+1}]$:

$$\mathbf{S}\nabla_{x+y}\Lambda^i + \frac{V_N}{N}\nabla_{x+y}\Lambda^i$$
$$= \chi_{[T_i,T_{i+1}]}\Big(V_N\nabla_{x+y}\Lambda\circ\Gamma + V_N\nabla_{x+y}\Gamma\circ\Lambda + V_N\Lambda\circ\nabla_{x+y}\Gamma + V_N\Gamma\circ\nabla_{x+y}\Lambda\Big)$$
$$:= RHS(1) \tag{81}$$
$$\mathbf{S}_{\pm}\nabla_{x+y}\Gamma^i$$
$$= \chi_{[T_i,T_{i+1}]}\Big(V_N\nabla_{x+y}\Gamma\circ\Gamma + V_N\nabla_{x+y}\Lambda\circ\Lambda + V_N\Lambda\circ\nabla_{x+y}\Gamma\circ\Gamma + V_N\Lambda\circ\nabla_{x+y}\Lambda\Big)$$
$$:= RHS(2) \tag{82}$$

with initial conditions at T_i, so $\nabla_{x+y}\Lambda^i = \nabla_{x+y}\Lambda$ and $\nabla_{x+y}\Gamma^i = \nabla_{x+y}\Gamma$ in $[T_i, T_{i+1}]$. By slight abuse of notation, $\nabla_{x+y}F$ is the function $(x, y) \to \nabla_{x+y}F(x, y)$. Now we can use and use Theorems 2.1 and Theorem 6.1:

$$\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\nabla_{x+y}\Lambda^i\big\|_{\mathcal{S}_{x,y}} + \big\|\big\langle\nabla_{x+y}\big\rangle^\alpha\nabla_{x+y}\Lambda^i\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$$
$$+ \big\||\partial_t|^{\frac{1}{4}}\nabla_{x+y}\Lambda^i\big\|_{L^2(dt)L^\infty(d(x-y))L^2(d(x+y))}$$
$$+ \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\nabla_{x+y}\Gamma^i\big\|_{\mathcal{S}_{x,y}} + \big\||\nabla_{x+y}|^\alpha\nabla_{x+y}\Gamma^i\big\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}$$
$$\lesssim \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha RHS(1)\big\|_{\mathcal{S}'_r} + \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha RHS(2)\big\|_{\mathcal{S}'_r}$$
$$+ \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\nabla_{x+y}\Lambda(t = T_i)\big\|_{L^2}$$
$$+ \big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\nabla_{x+y}\Gamma(t = T_i)\big\|_{L^2}.$$

We have to estimate $\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha$ applied to the 8 terms on the RHS. We will show they are all $\leq C\epsilon LHS$, so if $C\epsilon \leq \frac{1}{2}$, the theorem is proved. In the above compositions, the estimates are the same regardless whether the second term is Λ or Γ, so we call the second term B. The estimates are the same estimates as those of Lemmas 6.6 and 6.8, but now we can also use the estimates of Lemma 7.1. If ∇_{x+y} falls on B,

$$\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\Big(V_N\Gamma^i\circ\nabla_{x+y}B^i\Big)\big\|_{L^{\frac{8}{5}}([T_i,T_{i+1}])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq \epsilon\big\|\big\langle\nabla_x\big\rangle^\alpha\big\langle\nabla_y\big\rangle^\alpha\nabla_{x+y}B^i\big\|_{\mathcal{S}^r_{x,y}}$$

and

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\left((V_N\Lambda^i)\circ\nabla_{x+y}B^i\right)\|_{L^{\frac{4}{3}}([T_i,T_{i+1}])L^{\frac{3}{2}}(dx)L^2(dy)}$$
$$\leq\|\langle\nabla_{x+y}\rangle^\alpha\Lambda^i\|_{L^\infty(d(x-y))L^2([T_i,T_{i+1}])L^2(d(x+y))}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\nabla_{x+y}B^i\|_{L^4(dt)L^3(dx)L^2(dy)}$$
$$\leq\epsilon\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\nabla_{x+y}B^i\|_{L^4(dt)L^3(dx)L^2(dy)}.$$

If ∇_{x+y} falls on $V_N\Gamma$,

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Bigg((V_N\nabla_{x+y}\Gamma^i)\circ B^i\Bigg)\|_{L^{\frac{4}{3}}([T_i,T_{i+1}])L^{\frac{3}{2}}(dx)L^2(dy)}$$
$$\leq\|\nabla_{x+y}\Gamma^i\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha B^i\|_{L^4([T_i,T_{i+1}])L^3(dx)L^2(dy)}$$
$$\leq\epsilon\|\nabla_{x+y}\Gamma^i\|_{L^\infty(d(x-y))L^2([T_j,T_{j+1}])L^2(d(x+y))}$$

and similarly

$$\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha\Bigg((V_N\nabla_{x+y}\Lambda^i)\circ B^i\Bigg)\|_{L^{\frac{4}{3}}([T_i,T_{i+1}])L^{\frac{3}{2}}(dx)L^2(dy)}$$
$$\leq\|\nabla_{x+y}\Lambda^i\|_{L^\infty(d(x-y))L^2([T_j,T_{j+1}])L^2(d(x+y))}\|\langle\nabla_x\rangle^\alpha\langle\nabla_y\rangle^\alpha B^i\|_{L^4([T_i,T_{i+1}])L^3(dx)L^2(dy)}$$
$$\leq\epsilon\|\nabla_{x+y}\Lambda^i\|_{L^\infty(d(x-y))L^2(dt)L^2(d(x+y))}.$$

The proof for ∇^j_{x+y} is the same.

□

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

Proof. (of Theorem 1.3). The equations for $\mathrm{sh}(2k)=N\Lambda_p$ and $p_2=N\Gamma_p$ are

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

Let $\epsilon>0$. As in the previous proofs, use (73) and divide $[0,\infty)$ into finitely many intervals $[T_i,T_{i+1}]$ so that

$$\sup_z\|\Gamma(t,x+z,x)\|_{L^8([T_i,T_{i+1}])L^{\frac{12}{7}}(dx)}<\epsilon$$

and

$$\sup_z\|\Lambda(t,x+z,x)\|_{L^2([T_i,T_{i+1}])L^3(dx)}<\epsilon$$

and estimate the dual Strichartz norms

$$\|(V_N*\rho(t,x))\mathrm{sh}(2k)(t,x,y)\|_{L^{\frac{8}{5}}([T_i,T_{i+1}])L^{\frac{4}{3}}(dx)L^2(dy)}+\|(V_N\Gamma^T)\circ\mathrm{sh}(2k)\|_{L^{\frac{8}{5}}([T_i,T_{i+1}])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq C\sup_z\|\Gamma(t,x+z,x)\|_{L^8([T_i,T_{i+1}])L^{\frac{12}{7}}(dx)}\|\mathrm{sh}(2k)\|_{L^2([T_i,T_{i+1}])L^6(dx)L^2(dy)}$$
$$\leq C\epsilon\|\mathrm{sh}(2k)\|_{L^2([T_i,T_{i+1}])L^6(dx)L^2(dy)}$$

and similarly

$$\|(V_N * \rho(t,x))p_2(t,x,y)\|_{L^{\frac{8}{5}}([T_i,T_{i+1}])L^{\frac{4}{3}}(dx)L^2(dy)} + \|(V_N\Gamma)\circ p_2\|_{L^{\frac{8}{5}}([T_i,T_{i+1}])L^{\frac{4}{3}}(dx)L^2(dy)}$$
$$\leq C\epsilon\|p_2\|_{L^2([T_i,T_{i+1}])L^6(dx)L^2(dy)}.$$

Also,

$$\begin{aligned}
&\|(V_N\Lambda)\circ p_2\|_{L^{\frac{8}{5}}(dt)L^{\frac{4}{3}}(dx)L^2(dy)}\\
&\leq C\sup_z \|\Lambda(t,x+z,x)\|_{L^2([T_i,T_{i+1}])L^3(dx)}\|p_2\|_{L^4([T_i,T_{i+1}])L^3(dx)L^2(dy)}\\
&\leq \epsilon\|p_2\|_{L^4([T_i,T_{i+1}])L^3(dx)L^2(dy)}.
\end{aligned}$$

We get, using the estimates of Theorem 1.1 as well as proposition 4.8 and standard Strichartz estimates,

$$\begin{aligned}
&\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}[T_i,T_{i+1}]} + \|p_2\|_{\mathcal{S}_{x,y}[T_i,T_{i+1}]}\\
&\leq C((\|\mathrm{sh}(2k)(t=T_i)\|_{L^2} + \|p_2(t=T_i)\|_{L^2})\\
&+ C\epsilon\big((\|\mathrm{sh}(2k)\|_{\mathcal{S}_{x,y}[T_i,T_{i+1}]} + \|p_2\|_{\mathcal{S}_{x,y}[T_i,T_{i+1}]}\big) + C\log N.
\end{aligned}$$

Thus, if $C\epsilon \leq \frac{1}{2}$ we get the desired result on each interval $[T_i, T_{i+1}]$. The number of such intervals is finite (bounded by universal constants), and the result follows.

The proof for ∇^j_{x+y} is similar.

□

9. 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:

$$\begin{aligned}
&\left\{\frac{1}{i}\partial_t - \Delta_{x_1}\right\}\phi(x_1)\\
&= -\int dy\,\{v_N(x_1-y)\Gamma(y,y)\}\,\phi(x_1) &(83)\\
&- \int dy\big\{v_N(x_1-y)\Gamma_p(y,x_1)\phi(y) &(84)\\
&+ \int dy\big\{v_N(x_1-y)\Lambda_p(x_1,y)\big\}\overline{\phi}(y). &(85)
\end{aligned}$$

Define the standard Strichartz spaces

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

We prove the estimates of Corollary 1.5.

Proof. Using (19) and (21) we split $[0, \infty)$ into finitely many intervals so that

$$\|\big\langle \nabla_{x+y} \big\rangle^{\alpha} \Gamma\|_{L^8(dt)L^{\infty}(d(x-y))L^{\frac{4}{3}}(d(x+y))} \leq \epsilon$$

$$\|\big\langle \nabla_{x+y} \big\rangle^{\alpha} \Lambda\|_{L^2(dt)L^{\infty}(d(x-y))L^2(d(x+y))} \leq \epsilon$$

Using the fractional Leibniz rule, we easily estimate the RHS of the equation for ϕ:

$$\|\big\langle \nabla \big\rangle^{\alpha} (83)\|_{L^{\frac{8}{5}}([T_i.T_{i+1}])L^{\frac{4}{3}}(dx)} \leq C\epsilon \|\big\langle \nabla \big\rangle^{\alpha} \phi\|_{L^2[T_i.T_{i+1})L^6(dx)}$$

$$\|\big\langle \nabla \big\rangle^{\alpha} (84)\|_{L^{\frac{8}{5}}[T_i.T_{i+1})L^{\frac{4}{3}}(dx)} \leq C\epsilon \|\big\langle \nabla \big\rangle^{\alpha} \phi\|_{L^2[T_i.T_{i+1})L^6(dx)}$$

$$\|\big\langle \nabla \big\rangle^{\alpha} (85)\|_{L^1[T_i.T_{i+1})L^2(dx)} \leq C\epsilon \|\big\langle \nabla \big\rangle^{\alpha} \phi\|_{L^2[T_i.T_{i+1})L^6(dx)}$$

thus

$$\|\big\langle \nabla \big\rangle^{\alpha} \phi\|_{\mathcal{S}[T_i,T_{i+1}]} \leq C \|\big\langle \nabla \big\rangle^{\alpha} \phi(t = T_i)\|_{L^2} + 3C\epsilon \|\big\langle \nabla \big\rangle^{\alpha} \phi\|_{L^2[T_i.T_{i+1})L^6(dx)}.$$

By taking $3\epsilon < \frac{1}{2}$, the result for $j = 0$ follows. Next, differentiate the equation and use the same splitting. If the derivative falls on ϕ, the argument is the same. If the derivative falls on Γ, use

$$\sup_{x-y} \| |\nabla_{x+y}|^{\alpha+1} \Gamma\|_{L^2(dtd(x+y))} \lesssim 1.$$

While we don't know if this term can be made small by localizing to time intervals, such a term is coupled with ϕ without extra derivatives, which has been estimated already. For instance,

$$\int dz |v_N(z)| \| \left(|\nabla_x|^{\alpha+1} \Gamma_p(x, x-z) \right) \phi(x-z)\|_{L^1[T_i.T_{i+1})]L^2(dx)}$$

$$+ \int dz |v_N(z)| \left(|\nabla_x| \Gamma_p(x, x-z) \right) \big\langle \nabla \big\rangle^{\alpha} \phi(x-z)\|_{L^1[T_i.T_{i+1})]L^2(dx)}$$

$$\lesssim \sup_{x-y} \| |\nabla_{x+y}|^{\alpha+1} \Gamma_p\|_{L^2L^2} \| \| |\nabla|^{\alpha} \phi\|_{L^2(dt)L^6(dx)} \leq C$$

thus we get

$$\|\big\langle \nabla \big\rangle^{\alpha+1} \phi\|_{\mathcal{S}[T_i,T_{i+1}]} \leq C \|\big\langle \nabla \big\rangle^{\alpha+1} \phi(t = T_i)\|_{L^2} + 3C\epsilon \|\big\langle \nabla \big\rangle^{\alpha+1} \phi\|_{L^2[T_i.T_{i+1})L^6(dx)} + C$$

which proves the result. The case of higher j is similar. □

10. Proof the square function estimates

10.1. The double square function in standard coordinates. This subsection covers well-known results, and is included for the reader's convenience. Let ψ_k ($k \geq 1$) be any functions satisfying $\hat{\psi}_k = \hat{\psi}(\frac{\cdot}{2^k})$ with $\hat{\psi} \in C_0^{\infty}$ vanishing in a neighborhood of 0. Let $\hat{\psi}_0 \in C_0^{\infty}$.

Define $\overrightarrow{K}(x)$ be the infinite column vector $\overrightarrow{\psi(x)} = (\psi_k(x))_{k\geq 0}$. This satisfies the standard estimates for a Calderón-Zygmund operator: in particular (in three space dimensions) $|\overrightarrow{\psi(x)}| \lesssim \frac{1}{|x|^3}$, $|\overrightarrow{\psi(x+y)} - \overrightarrow{\psi(x)}| \lesssim \frac{|y|}{|x|^4}$ if $|x| > 2|y|$. Convolution with $\overrightarrow{K}(x)$ is bounded from L^2 to $L^2 l^2$ by orthogonality.

Denote

$$\overrightarrow{K}_1 f(x,y) = \int \overrightarrow{K}(x') f(x-x', y) dx'$$
$$\overrightarrow{K}_2 f(x,y) = \int \overrightarrow{K}(y') f(x, y-y') dy'.$$

We review the following known results

Lemma 10.2. *Let $1 < p, q < \infty$. Then*

$$\|\overrightarrow{K}_1 f\|_{L^p(dx)L^q(dy)l^2} \lesssim \|f\|_{L^p(dx)L^q(dy)}$$
$$\|\overrightarrow{K}_2 f\|_{L^p(dx)L^q(dy)l^2} \lesssim \|f\|_{L^p(dx)L^q(dy)}.$$

Remark 10.3. The above inequalities are just "linear" formulation of the square function estimate

$$\|S_1 f\|_{L^p(dx)L^q(dy)} = \| \left(\sum_{k=0}^{\infty} |f * (\psi_k \delta)|^2 \right)^{\frac{1}{2}} \|_{L^p(dx)L^q(dy)} \lesssim \|f\|_{L^p(dx)L^q(dy)} \tag{86}$$

$$\|S_2 f\|_{L^p(dx)L^q(dy)} = \| \left(\sum_{k'=0}^{\infty} |f * (\delta \psi_{k'})|^2 \right)^{\frac{1}{2}} \|_{L^p(dx)L^q(dy)} \lesssim \|f\|_{L^p(dx)L^q(dy)}.$$

Proof. The estimate for $\overrightarrow{K}_2$ (or S_2) follows right away from the standard square function estimate in y (for fixed x), followed by L^p in x.

The operator $\overrightarrow{K}_1$ is a Calderón-Zygmund operator. It is bounded from $L^p(dx)L^q(dy)$ to $L^p(dx)L^q(dy)l^2$. See [19], Theorem 2.1 and Corollary 2.3. For the main case we need, $q = 2$, this also follows from Section 5, Chapter 2 in [16]

□

In fact we can do more: Let $\overrightarrow{K}$ be the infinite column kernel $\overrightarrow{\psi(x)}$ as before, but now we multiply it with l^2 vectors $\overrightarrow{f_j}$:

$$\overrightarrow{\psi} \otimes \overrightarrow{f} = (\psi_k f_j)_{j,k}.$$

It maps $l^2 \to l^2 l^2$ with norm given by $|\overrightarrow{\psi}|_{l^2}$, and we repeat the argument, and get

Lemma 10.4. *Define*

$$\overrightarrow{\overrightarrow{K}}_1(\overrightarrow{f})(x,y) = \int \overrightarrow{K}(x') \otimes \overrightarrow{f}(x-x',y)dx' \tag{87}$$

Then, if $1 < p, q < \infty$,

$$\|\overrightarrow{\overrightarrow{K}}_1 \overrightarrow{f}\|_{L^p(dx)L^q(dy)l^2l^2} \lesssim \|\overrightarrow{f}\|_{L^p(dx)L^q(dy)l^2} \tag{88}$$

and thus

$$\|\overrightarrow{\overrightarrow{K}}_1(\overrightarrow{K}_2 f)\|_{L^p(dx)L^q(dy)l^2l^2} \lesssim \|\overrightarrow{f}\|_{L^p(dx)L^q(dy)l^2} \tag{89}$$

or, equivalently

$$\|S_1S_2f\|_{L^p(dx)L^q(dy)} = \|\left(\sum_{k',k''=0}^{\infty} |f * (\psi_{k'}\delta) * (\delta\psi_{k''})|^2\right)^{\frac{1}{2}}\|_{L^p(dx)L^q(dy)} \lesssim \|f\|_{L^p(dx)L^q(dy)}.$$

The result is also true, uniformly in M*, if the dyadic intervals* 2^i *defining the square functions are replaced by* 2^iM.

10.5. **The double square function in rotated coordinates.** Recall

$$R = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

and L_1, L_2 non-singular matrices satisfying satisfying

$$(RL_1)^{-1} = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$$

$$(RL_2)^{-1} = \begin{pmatrix} c & 1 \\ d & 0 \end{pmatrix}.$$

Lemma 10.6. *Let* K*,* f *be functions or distributions on* $\mathbb{R}^3$ *so* $K * f$ *is defined, and let* $K\delta$ *be* $K(x)\delta(y)$*, and* δK $\delta(x)K(y)$*. Then*

$$\left((K\delta) * (f \circ (RL_1)^{-1})\right)(RL_1(x,y)) = ((K\delta) * f)(x,y) \tag{90}$$

$$\left((\delta K) * (f \circ (RL_2)^{-1})\right)(RL_2(x,y)) = ((K\delta) * f)(x,y) \tag{91}$$

Proof. We have

$$\int K(x')\delta(y')f((x,y) - (RL_1)^{-1}(x',y'))$$

$$= \int K(x')\delta(y')f(x-x'-ay', y-by')dx'dy'$$

$$= \int K(x')f(x-x',y)dx'$$

and

$$\int \delta(x')K(y')f((x,y)-(RL_2)^{-1}(x',y'))$$
$$=\int \delta(x')K(y')f(x-cx'-y',y-x')dx'dy'$$
$$=\int K(y')f(x-y',y)dy'.$$

□

This calculation also works if K is an l^2 valued function, such as $\overrightarrow{K}$. Using this remark, we obtain

Lemma 10.7. *Let* $\overrightarrow{K}_1$, $\overrightarrow{K}_2$ *as above. Let* R *be the linear transformation defined by* (31)
Then

$$\|\overrightarrow{K}_{1,2}(f\circ R^{-1})(R(x,y))\|_{L^p(dx)L^q(dy)l^2} \lesssim \|f\|_{L^p(dx)L^q(dy)}. \tag{92}$$

Proof. The proof uses only the invariance of $L^p(dx)L^q(dy)$ under lower triangular invertible matrices: $\|f\|_{L^p(dx)L^q(dy)} = c\|f\circ L_1\|_{L^p(dx)L^q(dy)}$ and (92) is equivalent to

$$\|\overrightarrow{K}_1(f\circ(RL_1)^{-1})(RL_1(x,y))\|_{L^p(dxL^q(dy))} \lesssim \|f\|_{L^p(dx)L^q(dy)}.$$

For $\overrightarrow{K}_1$ we are convolving with $\overrightarrow{K}(x)\delta(y)$. Using (90) we have

$$\overrightarrow{K}_1(f\circ(RL_1)^{-1})(RL_1(x,y)) = \overrightarrow{K}_1(f)(x,y)$$

and we already know from Lemma 10.2 this is bounded on $L^p(dx)L^q(dy)$.

The argument for $\overrightarrow{K}_2$ is similar, but uses L_2:

$$\overrightarrow{K}_2(f\circ(RL_2)^{-1})(RL_2(x,y))$$
$$=\int \overrightarrow{K}(x')f(x-x',y)dx' = \overrightarrow{K}_1(f)(x,y)$$

which is bounded on $L^p(dx)L^q(dy)$, as in the previous case.

□

Next, recall $\overrightarrow{\overrightarrow{K}}_1$ defined by (87). The same proof as before gives

Lemma 10.8. *Let* $1 < p,q < \infty$. *Then*

$$\|\overrightarrow{\overrightarrow{K}}_{1,2}(\overrightarrow{f}\circ R^{-1})(R(x,y))\|_{L^p(dx)L^q(dy)l^2l^2} \lesssim \|\overrightarrow{f}\|_{L^p(dx)L^q(dy)l^2} \tag{93}$$

and, as a corollary,

$$\|\overrightarrow{\overrightarrow{K}}_1(\overrightarrow{K}_2(f \circ R^{-1})(R(x,y))\|_{L^p(dx)L^q(dy)l^2l^2} \lesssim \|\overrightarrow{f}\|_{L^p(dx)L^q(dy)l^2} \tag{94}$$

(95)

or, equivalently,

$$\|(S_1 S_2 f) \circ R\|_{L^p(dx)L^q(dy)}$$

$$= \| \left(\sum_{k',k''=0}^{\infty} |f * (\psi_{k'}\delta) * (\delta\psi_{k''}) \circ R|^2 \right)^{\frac{1}{2}} \|_{L^p(dx)L^q(dy)} \lesssim \|f \circ R\|_{L^p(dx)L^q(dy)}.$$

In other words, we have the double square function estimate in $x-y$, $x+y$ coordinates.

References

[1] V. Bach, S. Breteaux, T. Chen, J. Fröhlich, I. M. Sigal: The time-dependent Hartree-Fock-Bogoliubov equation for Bosons, arXiv:1602.05171.

[2] N. Benedikter, J. Sok, and J. P. Solovej, *The Dirac–Frenkel principle for reduced density matrices, and the Bogoliubov–de Gennes equations*, Annales Henri Poincaré **19** (2018), no. 4, 1167–1214.

[3] J. Bourgain, *Scattering in the energy space and below for 3D NLS*, Journal d'Analyse Mathématique **75** (1998), no. 1, 267–297.

[4] Chen, T., Hong, Y., Pavlovic, N. : Global well posedness for the NLS system for infinitely many Fermions, Arch. Rat. Mechanics and Analysis, **224**, 91–123 (2017).

[5] J. Chong, M. Grillakis, M. Machedon, Z. Zhao, *Global estimates for the Hartree-Fock-Bogoliubov equations* Communications in Partial Differential Equations, 46:10, pp 2015-2055, DOI: 10.1080/03605302.2021.1920615 (2021).

[6] J. Chong, Z. Zhao, *Dynamical Hartree–Fock–Bogoliubov Approximation of Interacting Bosons.* Ann. Henri Poincaré (2021). https://doi.org/10.1007/s00023-021-01100-w.

[7] F. M. Christ, M. I Weinstein, *Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation.* J. Funct. Anal. 100 (1991), no. 1, 87–109.

[8] X. Du, M. Machedon, *Counterexamples to L^p collapsing estimates*, Illinois Journal of Mathematics, Vol. 65, No. 1 (2021), 191 - 200.

[9] R. L. Frank and J. Sabin, *Restriction theorems for orthonormal functions, Strichartz inequalities, and uniform Sobolev estimates*, Amer. J. Math. 139(6), 1649-1691(2017).

[10] 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** (2013), no. 1, 91–111.

[11] M. Grillakis, M. Machedon: Pair excitations and the mean field approximation of interacting bosons, II, Communications in PDE, Vol 42, No 1, 24–67 (2017).

[12] ______, *Uniform in N estimates for a Bosonic system of Hartree–Fock–Bogoliubov type*, Communications in Partial Differential Equations **44** (2019), no. 12, 1431–1465.

[13] Y. Hong *Strichartz estimates for N-body Schrödinger operators with small potential interactions.* Discrete and Continuous Dynamical Systems, 2017, 37 (10) : 5355-5365. doi: 10.3934/dcds.2017233.

[14] M. Keel and T. Tao, *Endpoint strichartz estimates*, American Journal of Mathematics **120** (1998), no. 5, 955–980.

[15] C. Muscalu , J. Pipher , T. Tao, C. Thiele, *Bi-parameter paraproducts*, Acta Math., 193 (2004), 269-296.

[16] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press.

[17] E. M. Stein: *Harmonic Analysis, Real Variable Methods Orthogonality and Oscillatory Integrals*, Princeton University Press, (1993).

[18] T. Tao *Nonlinear Dispersive Equations: Local and Global Analysis*, vol. 106, American Mathematical Soc., 2006.

[19] R. Torres, E. Ward *Leibniz's Rule, Sampling and Wavelets on Mixed Lebesgue Spaces* Journal of Fourier Analysis and Applications 21(5), 2015, Pages 1053–1076 DOI 10.1007/s00041-015-9397-y.

[20] D. Tataru, personal communication, December 2017.

[21] Junyong Zhang, Jiqiang Zheng *Scattering theory for nonlinear Schrödinger equations with inverse-square potential, Journal of Functional Analysis 267 (2014) 2907–2932.*

University of Texas at Austin
Email address: jwchong@math.utexas.edu

University of Maryland, College Park
Email address: away@umd.edu

University of Maryland, College Park
Email address: mng@math.umd.edu

University of Maryland, College Park
Email address: mxm@math.umd.edu

Department of Mathematics and Statistics, Beijing Institute of Technology, Beijing, China

MIIT Key Laboratory of Mathematical Theory and Computation in Information Security, Beijing, China
Email address: zzh@bit.edu.cn