REGULARIZING NONLINEAR SCHRÖDINGER EQUATIONS THROUGH PARTIAL OFF-AXIS VARIATIONS

PAOLO ANTONELLI, JACK ARBUNICH, AND CHRISTOF SPARBER

Abstract. We study a class of (focusing) nonlinear Schrödinger-type equations derived recently by Dumas, Lannes and Szeftel within the mathematical description of high intensity laser beams [5]. These equations incorporate the possibility of a (partial) off-axis variation of the group velocity of such laser beams through a second order partial differential operator acting in some, but not necessarily all, spatial directions. We study the well-posedness theory for such models and obtain a regularizing effect, even in the case of only partial off-axis dependence. This provides an answer to an open problem posed in [5].

1. Introduction

Consider the initial value problem for a general (focusing) nonlinear Schrödinger equation (NLS) in $d \geq 1$ spatial dimensions, i.e.

\begin{equation}
\begin{cases}
  i\partial_t u + \Delta u + |u|^{2\sigma} u = 0, \ t \in \mathbb{R}, \ x \in \mathbb{R}^d, \\
  u(0, x) = u_0(x),
\end{cases}
\end{equation}

with $\sigma > 0$, some parameter describing nonlinear effects. The NLS is a canonical model for (weakly) nonlinear wave propagation in dispersive media, cf. [16]. In particular, the cubic case ($\sigma = 1$) is well-studied in the context of nonlinear laser optics, see [6, 16]. The NLS thereby describes diffractive effects which modify the propagation of slowly modulated light rays of geometrical optics over large times. In this context, the variable "$t$" should not be thought of as time, but rather as the main spatial direction of propagation of the ray. Solutions to (1.1) admit several conservation laws. In particular, one finds that

\begin{equation}
\|u(t, \cdot)\|_{L^2}^2 = \|u_0\|_{L^2}^2,
\end{equation}

which corresponds to the conservation of the (total) power, or intensity of the wave train.

From a mathematical point of view, it is well-known that (1.1) is $L^2$–subcritical provided $\sigma < \frac{2}{d}$. In this regime, one can use the dispersive properties of the NLS to obtain global solutions $u \in C(\mathbb{R}_t; L^2(\mathbb{R}))$, satisfying (1.1) in the sense of Duhamel’s integral representation, see e.g. [2]. For $\frac{2}{d} \leq \sigma < \frac{2}{d-2}$, one usually seeks solutions $u(t, \cdot) \in H^1(\mathbb{R}^d)$, in particular this includes the cubic case in dimensions $d = 2$ and 3. However such a solution may not exist for all times $t \in \mathbb{R}$, due to the possibility of finite-time blow-up. In this case

\[
\lim_{t \to T^-} \|\nabla u\|_{L^2} = +\infty,
\]

for some $T < \infty$, depending on the initial data. A rather complete description of this phenomenon is available in the $L^2$–critical case $\sigma = \frac{2}{d}$. In particular, it is

\begin{thebibliography}{99}

\bibitem{1} Dumas, Lannes and Szeftel.

\bibitem{2} 2000 Mathematics Subject Classification. 35Q41, 35C20.

\bibitem{3} Key words and phrases. Nonlinear Schrödinger equation, partial off-axis variation, Strichartz estimates, dispersion, finite-time blow-up, BBM equation.

\bibitem{4} This publication is based on work supported by the NSF through grant no. DMS 1348092.

\end{thebibliography}
known that global solutions exist for intensities \( \|u_0\|_{L^2} < \|Q\|_{L^2} \), where \( Q \) denotes the (stationary) ground state solution associated to (1.1). Above this threshold finite time blow-up appears and has been analyzed in a series of works, see [12, 13] and the references therein.

From the point of view of laser physics, blow-up is usually referred to as optical collapse. However, it is known from physics experiments that higher order effects, neglected in the derivation of (1.1), can arrest such a collapse and instead yield a process called filamentation. The latter corresponds to a complicated interplay between diffraction, self-focusing, and defocusing mechanisms present at high intensities which allow the beam to propagate beyond the theoretical predicted blow-up point, see [6].

In their recent mathematical study [5], Dumas, Lannes and Szeftel derive several new variants of the NLS from the underlying Maxwell’s equations of electromagnetism, in an effort to incorporate additional physical effects not present in (1.1). One of the new NLS type models derived in [5] allows for the possibility of an off-axis variation of the group velocity. It takes into account the fact that self-focusing pulses usually become asymmetric due to variations of the group velocity within off-axis rays, a phenomenon referred to as space-time focusing. Focusing pulses usually become asymmetric due to variations of the group velocity. One of the new NLS type models derived in [5] allows for the possibility of an off-axis dependence, for which we will recover (as we shall see below) the (stationary) ground state solution associated to (1.1). Above this threshold \( y \equiv \) is related to the celebrated Mahoney equation

\[ L = \text{the usual} \]

situation with no off-axis variation, for which we will recover (as we shall see below) the well-posedness theory for NLS.

To this end, the simplest mathematical model is given by

(1.3) \[ i\varepsilon \partial_t u + \Delta u + |u|^2 u = 0, \]

where \( P_\varepsilon \equiv P_\varepsilon(\nabla) \) is a linear, second order, self-adjoint operator such that

\[ \langle P_\varepsilon u, u \rangle_{L^2} \geq \|u\|_{L^2}^2 + \varepsilon^2 \sum_{j=1}^k \|\omega_j \cdot \nabla u\|_{L^2}^2. \]

Here, \( \langle \cdot, \cdot \rangle_{L^2} \) denotes the usual \( L^2 \)-inner product, \( 0 < \varepsilon \leq 1 \) is a small (dimensionless) parameter, and \( \{\omega_j\}_{j=1}^k \in \mathbb{R}^d \), with \( k \leq d \), are some given (linearly independent) vectors representing the off-axis directions. The case \( k = d \) thereby corresponds to a full off-axis dependence of the group velocity, whereas \( k < d \) is referred to as partial off-axis dependence. In the former case, the authors of [5] have shown that solutions \( u(t, \cdot) \in H^1(\mathbb{R}^d) \) to (1.3) exists for all \( t \in \mathbb{R} \), and hence no finite-time blow-up occurs. The situation involving only a partial off-axis dependence, however, is much more involved and it is an open problem posed in [5] to study the regularizing effect associated to this case.

In this work, we shall do so and thus provide an answer to the problem posed in [5]. To this end, we consider the following Cauchy problem

(1.4) \[
\begin{cases}
  i\varepsilon \partial_t u + \Delta u + |u|^2 u = 0, & t \in \mathbb{R}, \ x \in \mathbb{R}^d, \\
  u(0, x) = u_0(x),
\end{cases}
\]

where \( \sigma > 0 \). From now on, we shall split the spatial coordinates into \( x = (x, y) \in \mathbb{R}^{d-k} \times \mathbb{R}^k \) for \( k \leq d \), with the understanding that if \( k = d \), we again identify \( y \equiv x \in \mathbb{R}^d \). In addition, we choose w.l.o.g. \( \omega_j \) to be the \( j \)-th standard basis vectors in \( \mathbb{R}^k \). Explicitly, we then have

(1.5) \[ P_\varepsilon = 1 - \varepsilon^2 \Delta_y = 1 - \varepsilon^2 \sum_{j=1}^k \frac{\partial^2}{\partial y_j^2} \]

\[ 0 \leq k \leq d. \]

With the usual summation convention, the case \( k = 0 \) thereby corresponds to the situation with no off-axis variation, for which we will recover (as we shall see below) the usual \( L^2 \)-well-posed theory for NLS.

Mathematically, (1.3) is related to (1.1), in the same way the Benjamin-Bona-Mahoney equation is related to the celebrated Korteweg-de Vries equation for shallow, uni-directional water waves in \( d = 1 \), see [1]. The difference, when compared
to our case, is that we are not confined to work in only one spatial dimension, and therefore can allow for a partial regularization in \( k < d \) directions (a possibility which seems to have not been considered for BBM-type equations in higher dimensions, see [3]).

When comparing (1.4) to (1.1), one checks that, at least formally, both equations are Hamiltonian systems which conserve the same energy functional, i.e.,

\[
E(t) = \frac{1}{2} \| \nabla u(t, \cdot) \|_{L^2}^2 - \frac{1}{2(\sigma + 1)} \| u(t, \cdot) \|_{L^{2\sigma+2}}^{2\sigma+2} = E(0).
\]

However, instead of the usual \( L^2 \)-conservation law (1.2), one finds

\[
\| P_h^1/2 u(t, \cdot) \|_{L^2}^2 = \| P_h^1/2 u_0 \|_{L^2}^2,
\]

in the case of (1.4). Here, and in the following, \( P_h^1/2 \) is the pseudo-differential operator corresponding to the Fourier symbol

\[
\hat{P}_h^1/2(\eta) = (1 + h^2|\eta|^2)^{1/2}, \quad \text{for } \eta \in \mathbb{R}^k.
\]

The identity (1.7) corresponds to a conservation law for (the square of) the mixed \( L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k)) \)-norm of \( u \), whenever \( \varepsilon > 0 \). In order to understand the influence of partial off-axis variations, it is therefore natural to set up a well-posedness theory in this mixed Sobolev-type space.

With this in mind, we can now state the main results of this work.

**Theorem 1.1** (Partial off-axis variation: subcritical case). Let \( d > k \geq 0 \) and

- either \( k \leq 2 \) and \( 0 \leq \sigma < \frac{2}{d-k} \),
- or \( k > 2 \) and \( 0 \leq \sigma \leq \frac{2}{d-k} \).

Then for any \( u_0 \in L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k)) \) there exists a unique global solution \( u \in C(\mathbb{R}; L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k))) \) to (1.4), depending continuously on the initial data and satisfying the conservation law (1.7) for all \( t \in \mathbb{R} \).

In the result above, we have to exclude the choice \( k \leq 2 \) and \( \sigma = \frac{2}{d-k} \), which corresponds to a critical case that needs to be dealt with separately (see below).

Regardless of that, we see that as soon as \( k > 2 \), i.e., as soon as some partial off-axis variation is present, we can allow for \( L^2 \)-supercritical powers \( \sigma > \frac{2}{d-k} \) and still retain global in-time solutions \( u \). In other words, no finite time blow-up appears in the case of partial off-axis variations, and we can even allow for initial data \( u_0 \) in a space slightly larger than \( H^1(\mathbb{R}^d) \). This answers the question posed in [5].

We now turn to the case of partial off-axis dispersion with critical nonlinearity, for which we can prove an analog of the well-posedness results given in [3]. Note that for \( k = 0 \) (no off-axis variation) we recover the usual \( L^2 \)-critical case \( \sigma = \frac{2}{d} \).

**Theorem 1.2** (Partial off-axis variation: critical case). Let \( 0 \leq k \leq 2 \), and \( \sigma = \frac{2}{d-k} \). Then for any \( u_0 \in L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k)) \) there exists \( 0 < T_{\text{max}}, T_{\text{min}} \leq \infty \) and a unique maximal solution \( u \in C((-T_{\text{min}}, T_{\text{max}}); L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k))) \), satisfying (1.7), for all \( t \in (-T_{\text{min}}, T_{\text{max}}) \). In addition, we have the following blow-up alternative: \( T_{\text{max}} < \infty \), if and only if

\[
\| u \|_{L_t^{2(d-k+2)} \times x^{\frac{2}{d-k} + \frac{2}{d-k}}} \times \{ (0, T_{\text{max}}) \times \mathbb{R}^d \} \to \infty,
\]

and analogously for \( T_{\text{min}} \). Finally, if the \( L_h^2(\mathbb{R}^{d-k}; H_{\text{adm}}^1(\mathbb{R}^k)) \)-norm of the initial datum is sufficiently small, then the solution \( u \) exists for all \( t \in \mathbb{R} \).

For completeness, we shall also state a result in the case of full off-axis variation. Note when \( k = d \), the mixed Sobolev space above simply becomes \( H^1(\mathbb{R}^d) \).
Indeed, we can only expect “nice” dispersive properties in the spatial directions $u$ and artifact of our change of unknown $\Phi$. Morally speaking, the action of $P_\epsilon$ (1.11) allows us to gain a derivative in $y \in \mathbb{R}^k$. However, we also note that the linear semi-group
\begin{equation}
S_\epsilon(t) = e^{itP_\epsilon^{-1}\Delta}
\end{equation}
is no longer dispersive in the same way as the usual Schrödinger group $S_0(t) = e^{it\Delta}$. Indeed, we can only expect “nice” dispersive properties in the spatial directions $x \in \mathbb{R}^{d-k}$, where $P_\epsilon$ does not act, which will play an important role in the derivation of suitable Strichartz estimates (see below). Note that this issue is not simply an artifact of our change of unknown $u \mapsto v$, since $S_\epsilon(t)$ also describes the dispersive properties of (the linear part of) the original equation for $u$, as can be seen by applying $P_\epsilon^{-1}$ to the first line of (1.10). This issue has already been noticed in [5], but the change of unknown $u \mapsto v$, which allows us to treat the partial off-axis variation is a novel idea of the present paper.

We also want to mention that the sign of the nonlinearity (which is focusing) does not play a role in the proofs given below, and hence all of our results also remain true in the defocusing case. Since the latter does not allow for the possibility of finite-time blow-up, we decided not to take it into account in our exposition.

This paper is now organized as follows: In the next section we shall introduce some notation and definitions. Then in Section 3 we shall study the dispersive properties of $S_\epsilon(t)$ and derive appropriate Strichartz estimates in the case of partial off-axis dispersion. These will then be used in Section 4 to prove global well-posedness of (1.10) in the subcritical case. The critical case, and the case of full off-axis dispersion will be treated in Section 5.
2. Basic notations and definitions

As mentioned in the Introduction, we shall denote \( x = (x, y) \in \mathbb{R}^{d-k} \times \mathbb{R}^k \) with the understanding that if, either \( k = 0 \) (no off-axis variation), or if \( k = d \) (full off-axis variation), the variable \( y \) does not appear. We will often use mixed Lebesgue spaces such as \( L^p(x) \) and \( L^q(y) \), whereas the partial Fourier transform with respect to the \( x \)-variable only, will be denoted by

\[
\mathcal{F}(f)(x, \eta) \equiv \hat{f}(x, \eta) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x, y) e^{-i(x \cdot \eta)} \, dy.
\]

Analogously, we denote the partial Fourier transform in \( y \) by \( \mathcal{F}_{y \to \eta} \).

By recalling the (family of) differential operators \( P_\varepsilon = 1 - \varepsilon^2 \Delta_y \), defined in [1.5] with \( 0 < \varepsilon \ll 1 \), we shall introduce the class of mixed Sobolev-type spaces \( L^q_x(\mathbb{R}^{d-k}; H^s_y(\mathbb{R}^k)) \) of order \( s \in \mathbb{R} \), via the following norm

\[
\|f\|_{L^q_x H^s_y} := \|P^{s/2}_1 f\|_{L^q_x L^s_y} = \|(1 + |y|^2)^{s/2} \hat{f}\|_{L^q_x L^s_y}.
\]

Obviously, the Fourier symbol corresponding to \( P^{1/2}_1 \) is nothing but the well-known Japanese bracket \( \langle \eta \rangle = (1 + |\eta|^2)^{1/2} \). However, in order make the dependence of certain estimates on the small parameter \( 0 < \varepsilon \ll 1 \) explicit, we shall often make use of the following set of basic inequalities: Let \( s \geq 0 \), then

\[
\varepsilon^s \|f\|_{H^s} \leq \|P^{s/2}_1 f\|_{L^2} \leq \|f\|_{H^s},
\]

as well as

\[
\|f\|_{H^{-s}} \leq \|P^{-s/2}_\varepsilon f\|_{L^2} \leq \varepsilon^{-s} \|f\|_{H^{-s}}.
\]

Moreover, we shall write \( a \lesssim b \) whenever there exists a universal constant \( C > 0 \), independent of \( \varepsilon \), such that \( a \leq Cb \). In general this constant \( C \) may change from inequality to inequality.

Furthermore, we will also make use of the mixed space-time spaces \( L^q_t L^s_x H^s_y \), which we shall equip with the norm

\[
\|F\|_{L^q_t L^s_x H^s_y} := \left( \int \|F(t)\|_{L^q_x L^s_y}^2 \, dt \right)^{1/2}.
\]

Associated with these spaces is the following notion of Strichartz admissibility.

**Definition 2.1.** Let \( d > k \geq 0 \) be given. We say that the pair \((q, r)\) is *admissible* if \( 2 \leq r \leq \infty, 2 \leq q \leq \infty \), and

\[
\frac{2}{q} = (d-k) \left( \frac{1}{2} - \frac{1}{r} \right) =: \delta(r)
\]

where we omit the endpoint case \((q, r) \neq (2, \frac{2(d-k)}{(d-k)-2}) \) for \( d-k \geq 2 \).

Clearly, if \( k = 0 \), this is just the usual admissibility condition for non-endpoint Strichartz pairs corresponding to the Schrödinger group \( S_0(t) = e^{it\Delta} \) acting on \( \mathbb{R}^d \).
In this section, we shall derive Strichartz estimates associated to \( S(t) = e^{itP_x^{-1} \Delta} \) in the case of partial off-axis variation, i.e. \( d > k \). To this end we first derive a set of basic dispersion estimates.

3.1. Dispersion estimate for \( S(t) \). Recall the notation \( \delta(r) \geq 0 \) introduced in Definition 2. Then we have the following:

**Proposition 3.1.** Let \( r \in [2, \infty] \), and \( t \neq 0 \). Then, for any \( \varepsilon > 0 \), the group of \( L^2 \)-unitary operators \( S(t) = e^{itP_x^{-1} \Delta} \) continuously maps

\[
L^r \rightarrow L^r, \quad \frac{1}{r} = 1,
\]

and it holds

\[
\| S(t) \phi \|_{L^r} \leq |4\pi t|^{-\delta(r)} \| \phi \|_{L^r}.
\]

**Proof.** We shall prove (3.1) using a duality argument and the density of the Schwartz space \( \mathcal{S}(\mathbb{R}^d) \) of smooth, rapidly decaying functions. Namely, we shall show that for \( f, g \in \mathcal{S}(\mathbb{R}^d) \):

\[
|\langle S(t)f, g \rangle_{L^2} | \leq |4\pi t|^{-\delta(r)} \| \mathcal{F} \phi \|_{L^r} \| \phi \|_{L^r}.
\]

The case \( r = 2 \) when \( \delta(r) = 0 \), then follows directly from the fact that \( S(t) \) is unitary. By interpolation, it therefore suffices to only prove the case \( r = \infty \) where \( \delta(r) = \frac{d-k}{2} \), i.e.

\[
|\langle S(t)f, g \rangle_{L^2} | \leq |4\pi t|^{-\frac{d-k}{2}} \| \mathcal{F} \phi \|_{L^r} \| \phi \|_{L^r}.
\]

To this end, we use Plancherel’s identity to write

\[
\langle S(t)f, g \rangle_{L^2} = \langle S(t)f, \hat{g} \rangle_{L^2} = \int_{\mathbb{R}^{d-k} \times \mathbb{R}^k} e^{-\frac{i |\xi|^2}{4\pi t |\eta|^2}} \hat{f}(\xi, \eta) \overline{\hat{g}(\xi, \eta)} \ d\xi \ d\eta
\]

We first compute the inner integral by using inverse partial Fourier inversion in \( \xi \) on \( \hat{f} \) to give

\[
\int_{\mathbb{R}^{d-k}} e^{-\frac{i |\xi|^2}{4\pi t |\eta|^2}} \hat{f}(\xi, \eta) \overline{\hat{g}(\xi, \eta)} \ d\xi
\]

\[
= \frac{1}{(2\pi)^{\frac{d-k}{2}}} \int_{\mathbb{R}^{d-k}} e^{-\frac{i |\xi|^2}{4\pi t |\eta|^2}} \hat{g}(\xi, \eta) \int_{\mathbb{R}^{d-k}} e^{ix \xi} \hat{f}(x, \eta) \ dx \ d\xi
\]

\[
= \int_{\mathbb{R}^{d-k}} \hat{f}(x, \eta) \left( \frac{1}{(2\pi)^{\frac{d-k}{2}}} \int_{\mathbb{R}^{d-k}} e^{ix \xi} e^{-\frac{i |\xi|^2}{4\pi t |\eta|^2}} \hat{g}(\xi, \eta) \ d\xi \right) \ dx
\]

\[
= \int_{\mathbb{R}^{d-k}} \hat{f}(x, \eta) \mathcal{F}^{-1}\left(e^{-\frac{i |\eta|^2}{4\pi t}} g(\cdot, \eta)\right)(x) \ dx,
\]

where the change in the order of integration follows by Fubini’s Theorem. We now recall that, for \( a \in \mathbb{R} \),

\[
\mathcal{F}^{-1}\left(e^{-\frac{i |\eta|^2}{4\pi t}} g(\cdot, \eta)\right)(z) = \left( \frac{a}{2\pi} \right)^{\frac{d-k}{4}} e^{\frac{|a| |z|^2}{4\pi t}}.
\]
By setting $a = 1 + \varepsilon^2 |\eta|^2$ it then follows that the latter term in the integrand of (3.3) becomes
\[
\mathcal{F}_{\xi \to x}^{-1} \left( e^{-\frac{\varepsilon |\eta|^2}{1 + \varepsilon^2 |\eta|^2}} \overline{g(\cdot, \eta)} \right)(x) = \frac{1}{(2\pi)^{\frac{d-2}{2}}} \mathcal{F}_{\xi \to x}^{-1} \left( e^{-\frac{\varepsilon |\eta|^2}{1 + \varepsilon^2 |\eta|^2}} \ast \overline{g(\cdot, \eta)} \right)(x)
\]
(3.4)
\[
= \frac{1 + \varepsilon^2 |\eta|^2}{4\pi t} \int_{\mathbb{R}^{d-k}} e^{\frac{t (1 + \varepsilon^2 |\eta|^2)}{4\pi t} |\xi - \eta|^2} \overline{g(z, \eta)} \, dz.
\]

Now it is clear by (3.3) and (3.4) that
\[
(4\pi t)^{-\frac{d-k}{2}} \langle S_\varepsilon(t)f, g \rangle_{L^2} \leq \int_{\mathbb{R}^{d-k}} \left( \int_{\mathbb{R}^d} (1 + \varepsilon^2 |\eta|^2)^{-\frac{d-k}{2}} |\hat{f}(x, \eta)||\overline{\hat{g}}(z, \eta)| \, dz \, d\eta \right) dx.
\]

Hence it follows after bringing the absolute value inside the integrals and changing the order of integration that
\[
|\langle S_\varepsilon(t)f, g \rangle_{L^2}| \leq |4\pi t|^{-\frac{d-k}{2}} \int_{\mathbb{R}^{d-k}} \left( \int_{\mathbb{R}^d} (1 + \varepsilon^2 |\eta|^2)^{-\frac{d-k}{2}} |\hat{f}(x, \eta)||\overline{\hat{g}}(z, \eta)| \, dz \right) dx d\eta.
\]

A Cauchy Schwartz inequality in $\eta$, followed by Plancherel’s identity then gives
\[
|\langle S_\varepsilon(t)f, g \rangle_{L^2}| \leq |4\pi t|^{-\frac{d-k}{2}} \int_{\mathbb{R}^{d-k}} \left( \|P_\varepsilon^{(d-k)} \hat{f}(x, \cdot)\|_{L^2} \|\overline{\hat{g}}(z, \cdot)\|_{L^2} \right) dx d\eta
\]
\[
\leq |4\pi t|^{-\frac{d-k}{2}} \|P_\varepsilon^{(d-k)} f\|_{L^1_x \mathcal{S}^0_y} \|g\|_{L^1_x \mathcal{S}'^0_y}.
\]

As discussed before, this is sufficient to conclude the desired result. \hfill \Box

**Remark 3.2.** Note that, as $\varepsilon \to 0$, the estimate (3.4) converges to
\[
\|S_0(t)\varphi\|_{L^2_\varepsilon} \leq |4\pi t|^{-(d-k)(\frac{1}{2} - \frac{\delta}{2})} \|\varphi\|_{L^1_\varepsilon \mathcal{S}'^0_\varepsilon},
\]
which is similar to the usual dispersion estimate for the Schrödinger group in dimension $d - k \in \mathbb{N}$, and again reflects the fact that we don’t have dispersion in the $y$-coordinates.

For the following, it is useful to replace $\varphi$ by $P_\varepsilon^{-\frac{d-k}{2}} \varphi$ in (3.4) to obtain a more symmetric version of the dispersion estimate, i.e.
\[
\|S_\varepsilon(t)P_\varepsilon^{-\frac{d-k}{2}} \varphi\|_{L^1_\varepsilon \mathcal{S}'^0_\varepsilon} \leq |4\pi t|^{-\frac{\delta}{2}} \|P_\varepsilon^{-\frac{d-k}{2}} \varphi\|_{L^1_\varepsilon \mathcal{S}'^0_\varepsilon}.
\]

In view of (2.1) and (2.2) this implies that
\[
(3.5) \quad \|S_\varepsilon(t)\varphi\|_{L^1_\varepsilon H^{-\delta(r)}_y} \leq |4\pi t|^{-\frac{\delta}{2}} \|\varphi\|_{L^1_\varepsilon H^{\delta(r)}_y}.
\]

This estimate has the advantage that it uses regular Sobolev-spaces $H^s$, independent of $\varepsilon$, to measure the regularity in $y$ (instead of employing the operator $P_\varepsilon$). The price to pay is that (3.5) no longer converges to the classical dispersion estimate in the limit $\varepsilon \to 0$ (except in the case $r = 2$ for which $\delta(r) = 0$). But since in this work we are not concerned with the limit $\varepsilon \to 0$, we shall ignore this issue in the following and base our Strichartz estimates on (3.3), instead of (3.1).
3.2. Strichartz estimates. Exploiting the dispersion estimate (3.5), we shall now prove space-time Strichartz estimates associated to $S_{\varepsilon}(t)$. These estimates also follow from abstract arguments as in [10]. For the sake of concreteness and due to our somewhat unusual function spaces, we shall give their proof in the non-endpoint case.

Remark 3.3. The case of endpoint Strichartz estimates, i.e., $(q, r) = \left(\frac{2(d-k)}{d-2k}, \frac{d}{d-2k}\right)$ for $d - k \geq 2$, in principle could also be dealt with as in [10], but since we never make use of it in our analysis, we shall not pursue this issue any further.

Proposition 3.4 (Strichartz Estimates). Let $S_{\varepsilon}(t) = e^{itP_{v}^{-1}\Delta}$, and $(q, r), (\gamma, \rho)$ be two arbitrary admissible Strichartz pairs with $0 < \delta(r), \delta(\rho) < 1$. Then there exist constants $C_1, C_2 > 0$, independent of $\varepsilon$, such that

\[
\|S_{\varepsilon}(t)f\|_{L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}} \leq C_1 \|f\|_{L^2},
\]

as well as

\[
\left\|\int_0^t S_{\varepsilon}(t-s)F(s)\,ds\right\|_{L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}} \leq C_2 \|F\|_{L_t^q L_x^r H_{\varepsilon}^\delta(\rho)}.
\]

Proof. We start by first noticing, that (3.6) is equivalent to saying that the map $f \mapsto S_{\varepsilon}(t)f$ is bounded as an operator $L^2 \to L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}$. Let us define the operator $T_{\varepsilon} : L_t^q L_x^r H_{\varepsilon}^\delta(\rho) \to L^2$ by

\[T_{\varepsilon}F = \int_\mathbb{R} S_{\varepsilon}(-s)F(s)\,ds,
\]

and note that its formal adjoint $T_{\varepsilon}^*$ is given by $S_{\varepsilon}(t)f$. Next, we shall show that

\[T_{\varepsilon}^* T_{\varepsilon} = \int_\mathbb{R} S_{\varepsilon}(t-s)F(s)\,ds
\]

is bounded as an operator from $L_t^q L_x^r H_{\varepsilon}^\delta(\rho) \to L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}$. By the generalized Minkowski’s inequality we have

\[\left\|\int_\mathbb{R} S_{\varepsilon}(t-s)F(s)\,ds\right\|_{L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}} \leq \left\|\int_\mathbb{R} S_{\varepsilon}(t-s)F(s)\,ds\right\|_{L_t^q H_{\varepsilon}^{-\delta(r)}} ds\right\|_{L_t^q} ,
\]

and applying the dispersion estimate (3.5), it follows that

\[\|S_{\varepsilon}(t-s)F(s)\|_{L_t^q H_{\varepsilon}^{-\delta(r)}} \leq |4\pi(t-s)|^{-\delta(r)}\|F(s)\|_{L_t^q H_{\varepsilon}^\delta(\rho)}.
\]

Hence recalling that $\delta(r) = \frac{2}{q} < 1$, we see it is then possible to apply the Hardy-Littlewood-Sobolev inequality in order to obtain

\[\left\|\int_\mathbb{R} S_{\varepsilon}(t-s)F(s)\,ds\right\|_{L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}} \leq \left\|\int_\mathbb{R} |t-s|^{-\delta(r)}\|F(s)\|_{L_t^q H_{\varepsilon}^\delta(\rho)}\,ds\right\|_{L_t^q} \leq C \|F\|_{L_t^q L_x^r H_{\varepsilon}^\delta(\rho)}.
\]

We thus have proven that the operator $T_{\varepsilon}^* T_{\varepsilon} : L_t^q L_x^r H_{\varepsilon}^\delta(\rho) \to L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}$ is bounded. A standard functional analysis result for operators on Banach spaces states that

\[
\|T_{\varepsilon}^* T_{\varepsilon}\|_{L_t^q(L_x^r L_x^r): L^2} = \|T_{\varepsilon}^*\|_{L(L^2; L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}; L^2)} \|T_{\varepsilon} T_{\varepsilon}^*\|_{L(L_t^q L_x^r H_{\varepsilon}^\delta(\rho); L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}; L^2)}.
\]

This consequently implies that both

\[T_{\varepsilon} : L_t^q L_x^r H_{\varepsilon}^\delta(\rho) \to L^2 \quad \text{and} \quad T_{\varepsilon}^* : L^2 \to L_t^q L_x^r H_{\varepsilon}^{-\delta(r)}
\]

are bounded. Furthermore, we note that this holds for any non-endpoint admissible pair $(q, r)$. 
Now, choose any arbitrary (non-endpoint) admissible pairs \((\gamma, \rho)\) and \((q, r)\) such that
\[
T_\varepsilon : L^\gamma_x L^\rho_t H^\delta_y \to L^2 \quad \text{and} \quad T_\varepsilon^* : L^2 \to L^\gamma_x L^\rho_t H^{-\delta}(y).
\]
By combining the estimates for the operators \(T_\varepsilon, T_\varepsilon^*\), we then infer that
\[
T_\varepsilon^* T_\varepsilon : L^\gamma_x L^\rho_t H^\delta_y \to L^\gamma_x L^\rho_t H^{-\delta}(y)
\]
is bounded, i.e.
\[
\left\| \int \mathcal{S}_\varepsilon(t-s)F(s) \, ds \right\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)} \leq C \left\| F \right\|_{L^\gamma_x L^\rho_t H^\delta_y},
\]
for any arbitrary \((q, r), (\gamma, \rho)\). We can then invoke Theorem 1.2 from the paper \cite{1} by Christ and Kiselev, to conclude the retarded estimate
\[
\left\| \int_s^t \mathcal{S}_\varepsilon(t-s)F(s) \, ds \right\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)} \leq C \left\| F \right\|_{L^\gamma_x L^\rho_t H^\delta_y}.
\]
In summary, this proves the desired result. \(\square\)

4. The Cauchy problem for partial off-axis variation in the subcritical case

4.1. Well-Posedness in terms of \(v\). In this section we shall give the proof of Theorem 1.1 by proving a global \(L^2\)-based well-posedness result for \(\text{(1.10)}\) with subcritical nonlinearities. To this end, we rewrite \(\text{(1.10)}\) using Duhamel’s formulation, i.e.

\[
(4.1) \quad v(t) = \mathcal{S}_\varepsilon(t)v_0 + \int_0^t \mathcal{S}_\varepsilon(t-s)P_\varepsilon^{-1/2}\left(|P_\varepsilon^{-1/2}v|^2P_\varepsilon^{-1/2}v\right)(s) \, ds =: \Phi(v)(t).
\]

For simplicity, we shall write
\[
\Phi(v)(t) = \mathcal{S}_\varepsilon(t)v_0 + \mathcal{N}(v)(t),
\]
and denote
\[
(4.2) \quad \mathcal{N}(v)(t) := i \int_0^t \mathcal{S}_\varepsilon(t-s)P_\varepsilon^{-1/2}g(P_\varepsilon^{-1/2}v(s)) \, ds,
\]
where \(g(z) = |z|^{2\sigma}z\) with \(\sigma > 0\).

Of course, the basic idea is to prove that \(v \mapsto \Phi(v)\) is a contraction mapping in a suitable Banach space. To this end, the following lemma is key.

Lemma 4.1. Let \(d-k > 0\). Fix a \(T > 0\) and choose the admissible pair \((\gamma, \rho) = \left(\frac{4(\sigma+1)}{q-d-k}, \frac{2(\sigma+1)}{2}\right)\). Then, the following inequality holds

\[
\|\mathcal{N}(v) - \mathcal{N}(v')\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)} \lesssim \varepsilon^{-2(\sigma+1)}T^{1 - \frac{(d-k)\sigma}{2}} \left(\|v\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)}^{2\sigma} + \|v'\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)}^{2\sigma}\right) \|v - v'\|_{L^\gamma_x L^\rho_t H^{-\delta}(y)}\],
\]
whenever \(0 < \sigma \leq \frac{2}{(d-2)_+}\).

Proof. We first note that for our pair \((\gamma, \rho)\) to be non-endpoint admissible for \((d-k) \geq 2\), we require that \(\gamma > 2\), which in turn implies that \(\sigma < \frac{2}{((d-k)-2)_+}\). This condition, however, will always be fulfilled, since

\[
\sigma \leq \frac{2}{(d-2)_+} < \frac{2}{((d-k)-2)_+}.
\]

As a consequence, we also have that \(\delta(\rho) = \frac{(d-k)\sigma}{2(\sigma+1)} < 1\).
Now, as a first step we apply the Strichartz estimate \((3.7)\) and note that
\[
\|N(v)(t) - N(v')(t)\|_{L_t^\gamma L_x^2 H_y^{-(\delta)(\rho)}} \\
\leq C_2 \|P_{\varepsilon}^{-1/2} (g(P_{\varepsilon}^{-1/2} v) - g(P_{\varepsilon}^{-1/2} v'))\|_{L_t^\gamma L_x^2 H_y^{(\rho)}} \\
\leq \varepsilon^{-1} C_2 \|g(P_{\varepsilon}^{-1/2} v) - g(P_{\varepsilon}^{-1/2} v')\|_{L_t^\gamma L_x^2 H_y^{-(1-\delta)(\rho)}}.
\]
where we have also used the scaling \((2.2)\) to obtain the factor \(\varepsilon^{-1}\). Next, by a Sobolev imbedding we have that \(H^s(\mathbb{R}^k) \hookrightarrow L^{\rho'}(\mathbb{R}^k)\) where
\[
s = k \left(1 - \frac{1}{2} \frac{1}{2(\sigma + 1)}\right) = \frac{k \sigma}{2(\sigma + 1)} \in (0, \frac{k}{2}).
\]
In turn, this also implies the dual imbedding \(L^{\rho'}(\mathbb{R}^k) \subset H^{-s}(\mathbb{R}^k)\). Now, if we impose that
\[
1 \geq s + \delta(\rho) = \frac{d \sigma}{2(\sigma + 1)},
\]
which is so whenever \(\sigma \leq \frac{2}{d - 2}\), then \(H^{-(\delta)(\rho)}(\mathbb{R}^k) \subset H^{-(1-\delta)(\rho)}(\mathbb{R}^k)\). Together these allow us to estimate
\[
\|g(P_{\varepsilon}^{-1/2} v) - g(P_{\varepsilon}^{-1/2} v')\|_{H_y^{-(1-\delta)(\rho)}} \leq \|g(P_{\varepsilon}^{-1/2} v) - g(P_{\varepsilon}^{-1/2} v')\|_{H_y^{-(\delta)(\rho)}} \\
\leq C_2 \|\|P_{\varepsilon}^{-1/2} v\|_{H_y^{-(\delta)(\rho)}} + \|P_{\varepsilon}^{-1/2} v'\|_{H_y^{-(\delta)(\rho)}}\|_{L_t^\gamma L_x^2} (v - v')\|_{L_y^{\rho'}} = (*),
\]
where we have also used that for all \(z, w \in \mathbb{C}\),
\[
|g(z) - g(w)| \leq C_\sigma (z^{2\sigma} + |w|^{2\sigma})|z - w|.
\]
Now, recall that \(\rho = 2(\sigma + 1)\) and hence \(\rho' = \frac{d \sigma}{2(\sigma + 1)}\). Thus, by first applying Hölder’s inequality and using \((2.2)\), we obtain
\[
(*) \lesssim (\|P_{\varepsilon}^{-1/2} v\|_{L_y^{\rho'}} + \|P_{\varepsilon}^{-1/2} v'\|_{L_y^{\rho'}})\|P_{\varepsilon}^{-1/2} (v - v')\|_{L_y^{\rho'}} \\
\lesssim \varepsilon^{-(2\sigma + 1)} (\|v\|_{H_y^{-(\delta)(\rho)}} + \|v'\|_{H_y^{-(\delta)(\rho)}})\|v - v'\|_{H_y^{-(\delta)(\rho)}} \\
\lesssim \varepsilon^{-(2\sigma + 1)} (\|v\|_{H_y^{-(\delta)(\rho)}} + \|v'\|_{H_y^{-(\delta)(\rho)}})\|v - v'\|_{H_y^{-(\delta)(\rho)}},
\]
where the last inequality follows since \(H^{-(\delta)(\rho)}(\mathbb{R}^k) \subset H^{-(1-\delta)(\rho)}(\mathbb{R}^k)\), by the same arguments as before.

By employing Hölder’s inequality in \(x\) we consequently infer
\[
\|g(P_{\varepsilon}^{-1/2} v) - g(P_{\varepsilon}^{-1/2} v')\|_{L_y^{\rho'} H_x^{-(1-\delta)(\rho)}} \\
\lesssim \varepsilon^{-(2\sigma + 1)} (\|v\|_{L_y^{\rho'} H_x^{-(\delta)(\rho)}} + \|v'\|_{L_y^{\rho'} H_x^{-(\delta)(\rho)}})\|v - v'\|_{L_y^{\rho'} H_x^{-(\delta)(\rho)}}.
\]
From here, we compute that
\[
\frac{1}{\gamma} = 1 - \frac{(d - k)\sigma}{2} + \frac{2\sigma}{\gamma} + 1.
\]
Thus, taking the \(L_t^\gamma\) norm in \(t\) and applying Hölder’s inequality yields the result of the lemma. \(\Box\)

Using Lemma \((4.4)\) we are now able to prove global well-posedness for \((1.10)\) in the subcritical case. In doing so, we will require a positive exponent \(\alpha \equiv 1 - \frac{(d - k)\sigma}{2}\) of \(T\) in the estimate obtained in Lemma \((4.4)\) i.e., we require \(\sigma < \frac{2}{d - k}\). Since Lemma \((4.4)\) holds for \(\sigma \leq \frac{2}{d - 2}\), we need to distinguish the cases \(k \leq 2\) and \(k > 2\) in the following: One notices immediately that for \(k \leq 2\), we have that \(\frac{2\sigma}{\gamma - k} \leq \frac{2}{d - 2}\), which in turn implies that, in this case, we require the stronger assumption \(\sigma < \frac{2}{d - k}\) to ensure \(\alpha > 0\).
Proposition 4.2. Let $M, T > 0$, to be determined later on, and set

$$\|v\|_{L^2} = \|v_0\|_{L^2}, \quad \forall t \in \mathbb{R}.$$ 

We now set

$$\gamma = \frac{2}{d - k} \sigma = \frac{2}{d - 2}.$$ 

and hence no new restriction arises. We also note that for $k > 2$, the exponent of $T^\alpha$ is positive and is $L^2$-subcritical in the sense that when $\sigma = \frac{2}{d - 2}$ then

$$\alpha = 1 - \frac{(d - 2)\sigma}{2} = \frac{k - 2}{d - 2} > 0.$$

With this in mind, we can now prove the following result.

**Proposition 4.2.** Let $d > k \geq 0$ and

- either $k \leq 2$ and $0 \leq \sigma < \frac{1}{d - k}$,
- or $k > 2$ and $0 \leq \sigma < \frac{2}{d - 2}$.

Then for any $v_0 \in L^2(\mathbb{R}^d)$, there exists a unique global solution

$$v \in C(\mathbb{R}; L^2(\mathbb{R}^d)) \cap \mathcal{V} \cap L^\infty(\mathbb{R}; L^2(\mathbb{R}^d; H^{-q}_y(\mathbb{R}^k))),$$

for any (non-endpoint) admissible pair $(q, r)$. Moreover, $v$ depends continuously on the initial data and satisfies

$$\|v(t, \cdot)\|_{L^2} = \|v_0\|_{L^2}, \quad \forall t \in \mathbb{R}.$$ 

By identifying $v = P^{1/2}_\xi u$, this directly yields a global in-time solution $u \in C(\mathbb{R}; L^2(\mathbb{R}^{d-k}; H^1_r(\mathbb{R}^k)))$ to (1.3) and thus proves Theorem 1.2.

**Proof.** We shall prove Proposition 4.2 in several steps:

Step 1 (Existence): Fix the admissible pair $(\gamma, \rho) = \left(\frac{4(\sigma + 1)}{(d-k)\sigma}, 2(\sigma + 1)\right)$. Let $M, T > 0$, to be determined later on, and set

$$X_{T,M} = \{v \in L^\infty([0, T); L^2(\mathbb{R}^d)) \cap L^\infty([0, T); L^\infty(\mathbb{R}^d; H^{-q}_y(\mathbb{R}^k))) :$$

$$\|v\|_{L^\infty L^2} + \|v\|_{L^1 L^\infty H^{-q}_y} \leq M\}.$$ 

We note that $X_{T,M}$ is a complete metric space equipped with the distance

$$d(v, w) = \|v - w\|_{L^\infty L^2} + \|v - w\|_{L^1 L^\infty H^{-q}_y}.$$ 

Let $v \in X_{T,M}$. Then the Strichartz estimates obtained in Proposition 3.3 together with Lemma 3.4 imply that for any admissible $(q, r)$

$$\|\Phi(v)\|_{L^2 L^r H^{-q}_y} \leq \|S(\tau)v_0\|_{L^1 L^r H^{-q}_y} + \|N(v)\|_{L^1 L^r H^{-q}_y} \leq C_{\sigma, \epsilon} \left(\|v_0\|_{L^2} + T^{1 - \frac{(d-k)\sigma}{2}} M^{2\sigma + 1}\right),$$

as well as

$$\|\Phi(v)(t)\|_{L^2 L^2} \leq \|v_0\|_{L^2} + \|P^\perp P^\perp g(P^\perp v)\|_{L^1 L^r H^{q}_y} \leq C_{\sigma, \epsilon} \left(\|v_0\|_{L^2} + T^{1 - \frac{(d-k)\sigma}{2}} M^{2\sigma + 1}\right).$$

Together, these yield

$$d(\Phi(v), \Phi(v)) \leq 2C_{\sigma, \epsilon} \left(\|v_0\|_{L^2} + T^{1 - \frac{(d-k)\sigma}{2}} M^{2\sigma + 1}\right).$$

We now set

$$M = 4C_{\sigma, \epsilon} \|v_0\|_{L^2}$$

and choose $T > 0$ such that

$$2C_{\sigma, \epsilon} T^{1 - \frac{(d-k)\sigma}{2}} M^{2\sigma + 1} \leq \frac{M}{2}.$$
Then it follows that $\Phi(v) \in X_{T,M}$ for all $v \in X_{T,M}$ so that $\Phi(X_{T,M}) \subset X_{T,M}$. In addition, continuity of $\Phi$ follows from Lemma 4.11. Now, let $v, w \in X_{T,M}$ then

$$\|N(v) - N(w)\|_{L^1_t L_x^2 H^{s(e)}_y} \leq 2C_{\sigma, \varepsilon} M^{2\sigma} T^{1 - \frac{d - 1}{2}} \|d(v, w)\|,$$

which implies that

$$d(\Phi(v), \Phi(w)) \leq \frac{1}{2} d(v, w), \quad \text{for all } v, w \in X_{T,M}.$$

Thus $\Phi$ is a contraction map on $X_{T,M}$ and Banach’s fixed point theorem yields the existence of a unique fixed point $v \in X_{T,M}$, solving (4.1).

Step 2 (Uniqueness): For the same choice $(\gamma, \rho)$ as before, let $v, w \in X_{T,M}$ such that $v_0 = w_0$. Then, in view of (4.1), we have

$$v(t) - w(t) = N(v)(t) - N(w)(t),$$

and thus (4.3) implies that for a sufficiently small time $T > 0$, we have

$$d(v, w) = \|v - w\|_{L^\infty_t L^2_x} + \|v - w\|_{L^1_t L^2_x H^{s(e)}_y}$$

$$= \|N(v) - N(w)\|_{L^\infty_t L^2_x} + \|N(v) - N(w)\|_{L^1_t L^2_x H^{s(e)}_y}$$

$$\leq \frac{1}{2} \|v - w\|_{L^1_t L^2_x H^{s(e)}_y} \leq \frac{1}{2} d(v, w),$$

i.e., $v = w$ on $(0, T)$. Moreover this holds for any admissible pair $(q, r)$ since

$$\|v - w\|_{L^1_t L^2_x H^{s(e)}_y} \leq \frac{1}{2} \|v - w\|_{L^1_t L^2_x H^{s(e)}_y} = 0.$$

Hence we obtain a unique solution $u \in X_{T,M}$ up to some positive (small) time $T = T(\|v_0\|_{L^2}) > 0$. In addition, the solution depends continuously on the initial data, as can be seen by taking two solutions $v, \tilde{v}$ on some common time-interval $0 \leq t \leq T^* = \min\{T, \hat{T}\}$. Then by what was done before we have

$$d(v, \tilde{v}) \leq \|v_0 - \tilde{v}_0\|_{L^2} + C_{T^*, \sigma, \varepsilon} d(v, \tilde{v}),$$

where $C_{T^*, \sigma, \varepsilon} < 1$ if $T^*$ is chosen sufficiently small.

Step 3 (Global existence). In order to show that the solution obtained in Step 1 indeed exists for all times $t \in \mathbb{R}$, let

$$T_{\text{max}} = \sup\{T > 0 : \text{there exists a solution } v(t, \cdot) \text{ on } [0, T)\}.$$

We claim that

if $T_{\text{max}} < +\infty$ then $\lim_{t \to T_{\text{max}}} \|v(t)\|_{L^2} = +\infty$.

Suppose, by contradiction, that $T_{\text{max}} < \infty$ and that there exists a sequence $t_j \to T_{\text{max}}$ such that $\|v(t_j)\|_{L^2} \leq M$. Now for some integer $J$ such that $t_J$ is close to $T_{\text{max}}$ it follows by continuity that $\|v(t_j)\|_{L^2} \leq M$. Now by Step 1, with initial data $v(t_j)$ we extend our solution to the interval $[t_j, t_J + T]$ where we now choose $t_J$ such that

$$t_J + T > T_{\text{max}}$$

gives a contradiction to the definition of $T_{\text{max}}$. Thus $\lim_{t \to T_{\text{max}}} \|v(t)\|_{L^2} = \infty$, if $T_{\text{max}} < \infty$.

Next, let $I \subset [0, T_{\text{max}})$ and assume that $\{w_n(t, \cdot)\}_{n \in \mathbb{N}}$ for $t \in I$ is a sequence of sufficiently smooth and decaying solutions to (4.10) such that $w_n(t, \cdot) \to v(t, \cdot)$ in
L^2. We can then perform the following computation

\[
\frac{d}{dt} \|v_n(t, \cdot)\|_{L^2}^2 = 2 \text{Re} \langle v_n(t, \cdot), \partial_t v_n(t, \cdot) \rangle
\]

which implies that the L^2-norm is constant

\[\|v_n(t, \cdot)\|_{L^2} = \|v_n(0, \cdot)\|_{L^2} \quad \forall t \in I \subset [0, T_{\text{max}}].\]

The fact that v depends continuously on the initial data, then allows us to pass to the limit and obtain

\[\|v(t, \cdot)\|_{L^2} = \|v_0\|_{L^2} \quad \forall t \in I \subset [0, T_{\text{max}}].\]

Now replace v₀ by v(t₀) for some t₀ ∈ (0, T_{\text{max}}). Then \|v(t)\|_{L^2} is locally constant and hence constant. The L^2-conservation law consequently allows us to reapply Step 1 as many times as we wish, preserving the length of the maximal interval in each iteration, and yielding a global in time solution such that T_{\text{max}} = +∞. Since the equation is time-reversible modulo complex conjugation, this yields a global solution for all t ∈ ℝ.

\[\square\]

4.2. Higher order regularity. In this subsection, we are going to prove that the global in-time L^2-solution obtained in Proposition 4.2 enjoys persistence of regularity. Namely, if the initial datum v₀ ∈ H¹, then the corresponding solution v(t, ·) remains in H¹ for all times t ∈ ℝ. We will prove this property by exploiting the Strichartz estimates stated in Proposition 3.4 and the global well-posedness result in L². Similar arguments can be used to obtain a solution v(t, ·) ∈ H^s, s ≥ 1, provided the nonlinearity is sufficiently smooth.

To this end, we first note that the estimate in Lemma 4.1 together with the fact that g(z) = |z|^{2σ+2} is locally Lipschitz implies the following:

**Lemma 4.3.** Let d − k > 0, T > 0, (γ, ρ) = \(\left(\frac{4(\sigma + 1)}{(d-k)\sigma}, 2(\sigma + 1)\right)\), and moreover let us assume that \(\nabla v \in L^1_t L^2_x H^\rho_w \delta(\rho)\). Then we have

\[
\|\nabla \mathcal{N}(v)\|_{L^1_t L^2_x H^\rho_w \delta(\rho)} \lesssim e^{-2(\sigma+1)T} T^{1 - \frac{(d-k)\sigma}{4}} \|v\|_{L^2_t L^2_x H^\rho_w \delta(\rho)} \|\nabla v\|_{L^1_t L^2_x H^\rho_w \delta(\rho)}.
\]

From this we infer:

**Proposition 4.4.** Let v ∈ C(\[\mathbb{R}^d; L^2(\mathbb{R}^d)\] ∩ L^1_{loc}(\[\mathbb{R}^d; L^\infty(\mathbb{R}^d); L^\infty(\mathbb{R}^{d-k}; H^\rho_w(\mathbb{R}^k))\])) be the solution obtained in Proposition 4.2 with initial data v₀ ∈ L^2(\mathbb{R}^d). If, in addition, v₀ ∈ H^s(\mathbb{R}^d), then v ∈ C(\[\mathbb{R}^d; H^s(\mathbb{R}^d)\]).

**Proof.** Let us fix 0 < T < ∞, we are going to show that

\[
\|\nabla v\|_{L^2_t L^\infty_x(\mathbb{R}^d)} \leq C(T, \|v_0\|_{L^2}).
\]

From the global well-posedness result in Proposition 4.2, we know that

\[
\|v\|_{L^1_t L^2_x H^\rho_w \delta(\rho)(\mathbb{R}^d)} \leq C(T, \|v_0\|_{L^2}) =: CT,
\]

where (γ, ρ) = \(\left(\frac{4(\sigma+1)}{(d-k)\sigma}, 2(\sigma+1)\right)\) is the admissible pair used in Lemma 4.1. Let \(\eta > 0\) be a small parameter to be chosen later on. We then divide \[0, T\] into \(N = N(\eta, CT)\) subintervals, i.e., \([0, T] = \bigcup_{j=1}^N I_j\), where \(I_j = [t_{j-1}, t_j]\) and \(0 = t_0 < t_1 < \ldots < t_N = T\), such that

\[
\|v\|_{L^1_{t_j} L^2_x H^\rho_w \delta(\rho)(I_j \times \mathbb{R}^d)} \leq \eta, \quad j = 1, \ldots, N.
\]
Next, we apply the Strichartz estimates established in Proposition 3.4 to the gradient of the integral formula (4.1), to obtain
\[
\|\nabla v\|_{L^1_t L^2_x H^{-\delta(\rho)}_y (I_1 \times \mathbb{R}^d)} \leq C_1 \|\nabla v_0\|_{L^2} + \|\nabla N(v)\|_{L^1_t L^2_x H^{-\delta(\rho)}_y (I_1 \times \mathbb{R}^d)} \leq C_0 \left( \|\nabla v_0\|_{L^2} + \varepsilon^{-2(\sigma+1)} T^{-1 - \frac{(d-k)\sigma}{2}} \eta^{2\sigma} \|\nabla v\|_{L^1_t L^2_x H^{-\delta(\rho)}_y (I_1 \times \mathbb{R}^d)} \right),
\]
where in the last inequality we have used (4.4) and (4.5). (In addition, \( C_0 = \max(C_1, C_2) \), the two constants appearing in the Strichartz estimates.) Now choose \( \eta = \eta(T) \) such that
\[
C_0 \varepsilon^{-2(\sigma+1)} T^{-1 - \frac{(d-k)\sigma}{2}} \eta^{2\sigma} \leq \frac{1}{2},
\]
then we infer that
\[
\|\nabla v\|_{L^1_t L^2_x H^{-\delta(\rho)}_y (I_1 \times \mathbb{R}^d)} \leq 2C_0 \|\nabla v_0\|_{L^2}.
\]
By again using Strichartz estimates we obtain
\[
\|\nabla v\|_{L^\infty_t L^2_x \omega (I_1 \times \mathbb{R}^d)} \leq \|\nabla v_0\|_{L^2} + C_0 \varepsilon^{-2(\sigma+1)} T^{-1 - \frac{(d-k)\sigma}{2}} \eta^{2\sigma} \|\nabla v\|_{L^1_t L^2_x H^{-\delta(\rho)}_y (I_1 \times \mathbb{R}^d)} \leq (1 + C_0) \|\nabla v_0\|_{L^2}.
\]
In particular, we have
\[
\|\nabla v(t_i)\|_{L^2} \leq (1 + C_0) \|\nabla v_0\|_{L^2}.
\]
Using this, we can now iterate the argument on each subinterval \( I_j, j = 1, \ldots, N \), to finally obtain
\[
\|\nabla v\|_{L^\infty_t L^2_x \omega ((0, T) \times \mathbb{R}^d)} \leq C(T, \|\nabla v_0\|_{L^2}).
\]
Together with the conservation law for the \( L^2 \)-norm of \( v \), this yields the desired result. \( \square \)

**Remark 4.5.** Notice that we cannot obtain uniform in-time bounds on the \( H^1 \)-norm of \( v \) by invoking the conserved energy (1.6). Indeed the energy functional, written in terms of \( v \), reads
\[
E(t) = \frac{1}{2} \|P^{1/2} v\|_{L^2}^2 - \frac{1}{2(\sigma + 1)} \|P^{1/2} v\|_{L^{2\sigma+2}}^{2\sigma+2},
\]
which cannot provide a uniform bound on the full gradient of \( v \).

The proposition above yields a solution \( u \) to (1.3), such that \( v(t, \cdot) = P^{1/2}_u u(t, \cdot) \in H^1(\mathbb{R}^d) \) globally in-time. In particular, since
\[
\|u(t, \cdot)\|_{H^1} \leq \|P^{1/2} u(t, \cdot)\|_{H^1},
\]
we infer \( u(t, \cdot) \in H^1(\mathbb{R}^d) \) for all \( t \in \mathbb{R} \), provided \( P^{1/2}_u u_0 \in H^1 \). This shows that for a restricted class of initial data, the solution \( u \) exhibits a sufficient amount of regularity to rule out the possibility of finite time blow-up in the usual sense.

5. The critical case and the case of full off-axis dispersion

In this section, we shall treat the two “extreme” cases and consequently prove Theorems 1.2 and 1.3.
5.1. Partial off-axis dispersion with critical nonlinearity. In the case of partial off-axis dispersion with critical nonlinearity, i.e., \( \sigma = \frac{2}{d-k} \) and \( 0 \leq k \leq 2 \), we see that the estimate obtained in Lemma 4.2 no longer yields a positive power of \( T \). Hence the fixed point argument employed in the subcritical case breaks down. In order to overcome this obstacle, we shall employ the same type of arguments as in [3].

To this end, we first note that a particular admissible pair \((q, r)\), is obtained for
\[
q = r = \frac{2(d - k + 2)}{d - k},
\]
and introduce the following mixed space for any \( I \subset \mathbb{R}_t \):
\[
W(I) = L_t^{\frac{2(d - k + 2)}{d - k}}(I; L_x^{\frac{2(d - k + 2)}{d - k}} H_y \frac{d - k}{d + 2 + \varepsilon}(\mathbb{R}^d)).
\]
Then, we have the following local well-posedness result for \( v \), which directly yields Theorem 1.2 for \( u \) via \( v = F_{\varepsilon}^{1/2} u \).

**Proposition 5.1.** Let \( d - k > 0 \) with \( k \leq 2 \), and \( \sigma = \frac{2}{d-k} \). Then for any \( v_0 \in L^2(\mathbb{R}^d) \), there exists \( 0 < T_{\max}, T_{\min} < \infty \) and a unique maximal solution \( v \in C([-T_{\min}, T_{\max}); L^2(\mathbb{R}^d)] \cap W(I) \) to (1.1), where \( I \) denotes any closed interval \( I \subset (-T_{\min}, T_{\max}) \). Furthermore, \( T_{\max} < \infty \) if and only if
\[
\|v\|_{W(0,T_{\max})} = \infty,
\]
and analogously for \( T_{\min} \). Finally, if \( \|v_0\|_{L^2} \) is sufficiently small, then the solution is global.

Note that here the maximal existence time does not only depend on the size of the initial datum, but rather on the whole profile of the solution, or more precisely on the \( W(I) \)-norm of \( v \).

**Proof.** We shall only give a sketch of the proof for \( t \geq 0 \), since our arguments follow along the same lines as those in [3, Section 3].

Firstly, given a \( T > 0 \), we claim that by choosing \( \delta > 0 \) sufficiently small, such that
\[
\|S_{\varepsilon}(t)v_0\|_{W([0,T])} < \delta,
\]
we obtain a unique solution \( v \in C([0,T]; L^2(\mathbb{R}^d)) \cap W(I) \) to (1.1). Indeed, under assumption (5.2), we shall consider equation (4.1) in the space
\[
Z_{T,\delta} = \{ v \in W([0,T]) \text{ s.t. } \|v\|_{W([0,T])} < 2\delta \}.
\]
By using Lemma 4.1 with \( q = r = \frac{2(d - k + 2)}{d - k} \) and (5.2), we obtain
\[
\|v\|_{W([0,T])} \leq \delta + \|v\|_{W([0,T])}^{\frac{d+k}{d+k-2}},
\]
and hence our claim follows by taking \( \delta > 0 \) sufficiently small. A similar estimate then also gives \( v \in L^q_t L^r_x H_y^{-\frac{d-k}{d+k}}([0,T] \times \mathbb{R}^d) \) for any admissible pair \((q, r)\). Furthermore, since the solution \( v \) satisfies the integral equation (4.1), we also infer that \( v \in C([0,T]; L^2(\mathbb{R}^d)) \).

Moreover, from our Strichartz estimates in Proposition 3.3 we have that
\[
\|S_{\varepsilon}(t)v_0\|_{W(\mathbb{R})} \lesssim \|v_0\|_{L^2},
\]
and hence \( \|S_{\varepsilon}(t)v_0\|_{W([0,T])} \to 0 \) as \( T \to 0 \). Consequently, for \( T > 0 \) small enough, assumption (5.2) is satisfied, yielding a local in-time solution \( v(t, \cdot) \) for \( t \in [0, T] \). We can then iterate this argument to find a maximal existence time \( 0 < T_{\max} \leq \infty \) for which the solution exists for every admissible pair \((q, r)\).
Lastly, we shall prove the blow-up alternative (5.1) by contradiction. Namely, let $T_{\text{max}} < \infty$ and let us assume that $\|v\|_{W(0,T_{\text{max}})} < \infty$. Let $t \in [0, T_{\text{max}})$, then for any $s \in [0, T_{\text{max}} - t)$ we write in view of (4.4) that

$$S_t(s)v(t) = v(t + s) - \mathcal{N}(v(t + \cdot))(s).$$

Applying again Lemma 4.1 we can estimate

$$\|S_t(\cdot)v(t)\|_{W(0, T_{\text{max}} - t)} \leq \|v\|_{W(t, T_{\text{max}})} + C\|v\|_{W(t, T_{\text{max}})},$$

and thus, for $t$ sufficiently close to $T_{\text{max}}$, we have

$$\|S_t(\cdot)v(t)\|_{W(0, T_{\text{max}} - t)} < \delta.$$  

This implies we can extend the solution after the time $T_{\text{max}}$, contradicting its maximality. Finally, standard arguments imply the $L^2$-conservation law and the continuity of the solution with respect to the initial data $v_0$. □

5.2. The case of full off-axis dispersion. We finally turn to the case of full off-axis dispersion, i.e. $d = k$. It is clear from our admissibility condition in Definition 2.4 that in this case, we cannot expect to have any Strichartz estimates. We thus have to resort to a more basic fixed point argument to prove the following result:

**Lemma 5.2.** Let $d = k \geq 1$ and $\sigma \leq \frac{2}{(d-2)^+}$. Then, for any $v_0 \in L^2(\mathbb{R}^d)$, there exists a unique global solution $v \in C([0, T_{M}^2])$, depending continuously on the initial data and satisfying

$$\|v(t, \cdot)\|_{L^2}^2 = \|v_0\|_{L^2}^2, \quad \forall \, t \in \mathbb{R}.$$  

**Proof.** To prove this result it suffices to show that $v \mapsto \Phi(v)$ is a contraction on

$$Y_{T_0, M} = \{v \in L^\infty([0, T]; L^2(\mathbb{R}^d)) : \|v\|_{L^\infty T_{M}^2} \leq M\}.$$  

Let $v, v' \in Y_{T_0, M}$, and recall that $S_t(t)$ is unitary on $L^2$. Using Minkowski’s inequality and the scaling argument (4.2) then yields

$$\|\mathcal{N}(v)(t) - \mathcal{N}(v')(t)\|_{L^2} \leq \varepsilon^{-1} \int_0^t \|g(P_{\varepsilon^2}^{-1/2}v) - g(P_{\varepsilon^2}^{-1/2}v')\|_{H^{-\frac{\sigma}{2(\sigma+1)}}}}(s) \, ds,$$

provided $\frac{\sigma}{2(\sigma+1)} \leq 1$, i.e., $\sigma \leq \frac{2}{(d-2)^+}$. By a similar embedding strategy as in Lemma 4.1 one finds

$$\|g(P_{\varepsilon^2}^{-1/2}v) - g(P_{\varepsilon^2}^{-1/2}v')\|_{H^{-\frac{\sigma}{2(\sigma+1)}}}} \leq (\|P_{\varepsilon^2}^{-1/2}v\|_{L^2}^{2\sigma} + \|P_{\varepsilon^2}^{-1/2}v'\|_{L^2}^{2\sigma}) \|P_{\varepsilon^2}^{-1/2}(v - v')\|_{L^2} \leq \varepsilon^{-(2\sigma+1)}(\|v\|_{L^2}^{2\sigma} + \|v'\|_{L^2}^{2\sigma}) \|v - v'\|_{L^2},$$

which consequently implies that

$$\|\mathcal{N}(v) - \mathcal{N}(v')\|_{L^\infty T_{M}^2} \leq \varepsilon^{-(2\sigma+1)}(\|v\|_{L^\infty T_{M}^2}^{2\sigma} + \|v'\|_{L^\infty T_{M}^2}^{2\sigma}) \|v - v'\|_{L^\infty T_{M}^2}.$$  

Choosing $T > 0$ sufficiently small, Banach’s fixed point theorem directly yields a local in-time solution $v \in C([0, T], L^2(\mathbb{R}^d))$. Analogously to the proof of Proposition 4.2 we may extend the local solution $v$ for all $t \in \mathbb{R}$ using the $L^2$-conservation law. □

This consequently yields Theorem 1.3 for $u$, since in the case of full-off axis dispersion $v = P_{\varepsilon^2}^{1/2} u \in L^2(\mathbb{R}^d)$ implies $u \in H^1(\mathbb{R}^d)$, for any $\varepsilon > 0$. In addition, the $L^2$-conservation for $v$ directly yields (1.2), whereas (1.0) is a standard computation, and valid for any $H^1$-solution $u$.

**Remark 5.3.** Note that (1.7) also implies a uniform in-time bound on the $H^1$-norm of $u(t, \cdot)$, for any $\varepsilon > 0$. In turn, this means that both, the kinetic and the nonlinear potential energy remain uniformly bounded for all $t \in \mathbb{R}$.  

REFERENCES

1. T. B. Benjamin, J. L. Bona, and J. J. Mahony, *Model equations for long waves in nonlinear dispersive systems*. Phil. Trans. Royal Soc. London. Series A, Math. Phys. Sci. 272 (1972), 47–78.

2. T. Cazenave, *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics vol. 10, American Mathematical Society, 2003.

3. T. Cazenave, F. Weissler, *Some remarks on the nonlinear Schrödinger equation in the critical case*. In: Lecture Notes Math. 1394, pp. 18–29, Springer, 1989.

4. M. Christ and A. Kiselev, *Maximal Functions Associated to Filtrations*. J. Funct. Anal. 179 (2001), 409–425.

5. E. Dumas, D. Lannes, and J. Szeftel, *Variants of the focusing NLS equation. Derivation, justification and open problems related to filamentation*. In: CRM Series in Mathematical Physics, pp. 19–75, Springer, 2016.

6. G. Fibich, *The nonlinear Schrödinger equation; Singular solutions and optical collapse*. Appl. Math. Sciences vol. 192, Springer Verlag, 2015.

7. J. Ginibre and G. Velo, *On a class of of nonlinear Schrödinger equations. I. The Cauchy problem, general case*. J. Funct. Anal. 32 (1979), no. 1, 1–32.

8. J. Ginibre and G. Velo, *Generalized Strichartz estimates for the wave equation*. J. Funct. Anal. 133 (1995), no. 1, 50–68.

9. J. A. Goldstein and B. J. Wichnoski, *On the Benjamin Bona Mahony equation in higher dimensions*. Nonlin. Anal. 4 (1980), no. 4, 665–675.

10. M. Keel and T. Tao, *Endpoint Strichartz Estimates*. Amer. J. Math. 120 (1998), 955–980.

11. C. Kenig and F. Merle, *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case*. Invent. math. 166 (2006), no. 3, 645–675.

12. F. Merle and P. Raphael, *On universality of blow up profile for \( L^2 \) critical nonlinear Schrödinger equation*. Invent. Math. 156 (2004), 565–672.

13. F. Merle and P. Raphael, *Sharp lower bound on the blow up rate for critical nonlinear Schrödinger equation*. J. Amer. Math. Soc. 19 (2006), no. 1, 37–90.

14. F. Merle and P. Raphael, *Profiles and quantization of the blow up mass for critical nonlinear Schrödinger equation*. Comm. Math. Phys. 253 (2005), no. 3, 675–704.

15. J. E. Rothenberg, *Space-time focusing: breakdown of the slowly varying envelope approximation in the self-focusing of femtosecond pulses*. Optics Lett. 17 (1992), 1340–1342.

16. C. Sulem and P.-L. Sulem, *The nonlinear Schrödinger equation, Self-focusing and wave collapse*. Springer-Verlag, 1999.

(P. Antonelli) Gran Sasso Science Institute, viale F. Crispi,7,67100 L’Aquila, Italy
E-mail address: paolo.antonelli@gssi.infn.it

(J. Arbunich) Department of Mathematics, Statistics, and Computer Science, M/C 249, University of Illinois at Chicago, 851 S. Morgan Street, Chicago, IL 60607, USA
E-mail address: jarbun2@uic.edu

(C. Sparber) Department of Mathematics, Statistics, and Computer Science, M/C 249, University of Illinois at Chicago, 851 S. Morgan Street, Chicago, IL 60607, USA
E-mail address: sparber@uic.edu