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 [7]. 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 investigate the initial value problem for such models and obtain global well-posedness in L^2-supercritical situations, even in the case of only partial off-axis dependence. This provides an answer to an open problem posed in [7].

1. Introduction

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

$$\begin{cases} i\partial_t u + \Delta u + |u|^{2\sigma} u = 0, \ t \in \mathbb{R}, \ \mathbf{x} \in \mathbb{R}^d, \\ u(0,\mathbf{x}) = u_0(\mathbf{x}), \end{cases} \tag{1.1}$$

with $\sigma > 0$, some parameter describing nonlinear effects. The NLS is a canonical model for (weakly) nonlinear wave propagation in dispersive media, cf. [20]. In particular, the *cubic* case ($\sigma = 1$) is well-studied in the context of nonlinear laser optics, see [8, 20]. 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

$$\|u(t,\cdot)\|_{L^2}^2 = \|u_0\|_{L^2}^2, \tag{1.2}$$

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 $u \in C(\mathbb{R}_t; L^2(\mathbb{R}^d))$, satisfying (1.1) in the sense of Duhamel's integral representation, see e.g. [4]. For $\frac{2}{d} \leqslant \sigma < \frac{2}{(d-2)_+}$ 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

$$\lim_{t \to T_-} \|\nabla u(t,\cdot)\|_{L^2} = +\infty$$

Date: January 17, 2019.

2000 *Mathematics Subject Classification.* 35Q41, 35C20.

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

This publication is based on work supported by the NSF through grant no. DMS 1348092. The authors thank the anonymous referee for many helpful comments on improving our manuscript.

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 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 [13, 14, 15] 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 [8].

In their recent mathematical study [7], Dumas, Lannes and Szeftel derive several new variants of the NLS from the underlying Maxwell 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 [7] 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* in the optics literature, cf. [17]. To this end, the simplest mathematical model is given by

$$iP_\varepsilon \partial_t u + \Delta u + |u|^2 u = 0, \tag{1.3}$$

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

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

Here, $\langle \cdot, \cdot \rangle_{L^2}$ denotes the usual $L^2(\mathbb{R}^d)$ inner product, $0 < \varepsilon \leqslant 1$ is a small (dimensionless) parameter, and $\{\omega_j\}_{j=1}^k \in \mathbb{R}^d$, with $k \leqslant d$, 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 [7] have shown that solutions $u(t, \cdot) \in H^1(\mathbb{R}^d)$ to (1.3) exist 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 [7] to prove global well-posedness in this case.

In this work, we shall do so and thus provide an answer to the problem posed in [7]. To this end, we consider the following Cauchy problem:

$$\begin{cases} iP_\varepsilon \partial_t u + \Delta u + |u|^{2\sigma} u = 0, \ t \in \mathbb{R}, \ \mathbf{x} \in \mathbb{R}^d, \\ u(0, \mathbf{x}) = u_0(\mathbf{x}), \end{cases} \tag{1.4}$$

where $\sigma > 0$. From now on, we shall split the spatial coordinates into $\mathbf{x} = (x, y) \in \mathbb{R}^{d-k} \times \mathbb{R}^k$ for $k \leqslant d$, with the understanding that if $k = d$, we again identify $y \equiv \mathbf{x} \in \mathbb{R}^d$. In addition, we choose without loss of generality ω_j to be the jth standard basis vectors in $\mathbb{R}^k$. Explicitly, we then have

$$P_\varepsilon = 1 - \varepsilon^2 \Delta_y = 1 - \varepsilon^2 \sum_{j=1}^{k} \frac{\partial^2}{\partial y_j^2}, \quad 0 \leqslant k \leqslant d. \tag{1.5}$$

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-posedness theory for NLS.

Mathematically, (1.4) 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, unidirectional water waves in $d = 1$, see [2]. 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 [10]).

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

$$E(t) = \frac{1}{2}\|\nabla u(t,\cdot)\|_{L^2}^2 - \frac{1}{2(\sigma+1)}\|u(t,\cdot)\|_{L^{2\sigma+2}}^{2\sigma+2} = E(0). \tag{1.6}$$

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

$$\|P_\varepsilon^{1/2} u(t,\cdot)\|_{L^2}^2 = \|P_\varepsilon^{1/2} u_0\|_{L^2}^2 \tag{1.7}$$

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

$$\widehat{P}_\varepsilon^{1/2}(\eta) = (1 + \varepsilon^2|\eta|^2)^{1/2} \quad \text{for } \eta \in \mathbb{R}^k. \tag{1.8}$$

The identity (1.7) corresponds to a conservation law for (the square of) the mixed $L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k))$-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 $d > k \geqslant 0$ and*

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

Then for any $u_0 \in L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k))$ there exists a unique global-in-time solution $u \in C(\mathbb{R}_t; L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k)))$ 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 $k \leqslant 2$ 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 $k > 0$, i.e., as soon as some partial off-axis variation is present, we can allow for L^2-*supercritical* powers $\sigma > \frac{2}{d}$ 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 $H^1(\mathbb{R}^d)$.

We now turn to the case of partial off-axis dispersion with critical nonlinearity, for which we can prove an analogue of the well-posedness results given in [5]. 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 $0 \leqslant k \leqslant 2$, and $\sigma = \frac{2}{d-k}$. Then for any $u_0 \in L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k))$ there exist times $0 < T_{\max}, T_{\min} \leqslant \infty$ and a unique maximal solution $u \in C((-T_{\min}, T_{\max}); L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k)))$, satisfying* (1.7) *for all $t \in (-T_{\min}, T_{\max})$. In addition, we have the following blow-up alternative: $T_{\max} < \infty$ if and only if*

$$\|u\|_{L^{\frac{2(d-k+2)}{d-k}}\left([0,T_{\max})\times\mathbb{R}_x^{d-k}; H^{\frac{2}{d-k+2}}(\mathbb{R}_y^k)\right)} = \infty,$$

and analogously for $T_{\min}$. Finally, if the $L^2(\mathbb{R}_x^{d-k}; H^1(\mathbb{R}_y^k))$-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 $H^1(\mathbb{R}^d)$.

Theorem 1.3 (Full off-axis variation)**.** *Let $k = d$ and $0 \leqslant \sigma \leqslant \frac{2}{(d-2)_+}$. Then for any $u_0 \in H^1(\mathbb{R}^d)$ there exists a unique global-in-time solution $u \in C(\mathbb{R}_t; H^1(\mathbb{R}^d))$ to* (1.4)*, depending continuously on the initial data and satisfying the conservation laws* (1.6) *and* (1.7) *for all $t \in \mathbb{R}$.*

This is a slight generalization of the result given in [7], where only the cubic case is treated. Note that we can allow for $\sigma = \frac{2}{(d-2)_+}$, i.e., the H^1-critical power, in contrast to the usual theory of NLS without off-axis variation, cf. [12].

In order to prove all of these theorems, we shall employ the following change of unknown

$$v(t, \mathbf{x}) := P_\varepsilon^{1/2} u(t, \mathbf{x}), \tag{1.9}$$

and rewrite the Cauchy problem (1.4) in the form

$$\begin{cases} i\partial_t v + P_\varepsilon^{-1}\Delta v + P_\varepsilon^{-1/2}\big(|P_\varepsilon^{-1/2} v|^{2\sigma} P_\varepsilon^{-1/2} v\big) = 0, \ t \in \mathbb{R}, \ \mathbf{x} \in \mathbb{R}^d, \\ v(0, \mathbf{x}) = P_\varepsilon^{1/2} u_0(\mathbf{x}) \equiv v_0(\mathbf{x}). \end{cases} \tag{1.10}$$

Instead of (1.7), this new equation conserves

$$\|v(t, \cdot)\|_{L^2}^2 = \|P_\varepsilon^{1/2} u(t, \cdot)\|_{L^2}^2 = \|P_\varepsilon^{1/2} u_0\|_{L^2}^2 = \|v_0\|_{L^2}^2,$$

i.e., the usual L^2-conservation law. We therefore aim to set-up an L^2-based well-posedness theory for (1.10), written in Duhamel's form, i.e.

$$v(t) = e^{itP_\varepsilon^{-1}\Delta} v_0 + i \int_0^t e^{i(t-s)P_\varepsilon^{-1}\Delta} P_\varepsilon^{-1/2}(|P_\varepsilon^{-1/2} v|^2 P_\varepsilon^{-1/2} v)(s)\, ds.$$

The advantage of working with v instead of u lies in the fact that it allows us to exploit the regularizing properties of the operator $P_\varepsilon^{-1/2}$ acting on the nonlinearity. Roughly speaking, the action of $P_\varepsilon^{-1/2}$ allows us to gain a derivative in $y \in \mathbb{R}^k$. However, we also note that the linear semi-group

$$S_\varepsilon(t) = e^{itP_\varepsilon^{-1}\Delta} \tag{1.11}$$

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_ε does not act, which will play an important role in the derivation of suitable Strichartz estimates (see below). It has been proved in [3] that in the case of full off-axis dependence, $S_\varepsilon(t)$ does not admit any Strichartz estimates. Note that this issue is not simply an artifact of our change of unknown $u \mapsto v$, since $S_\varepsilon(t)$ also describes the dispersive properties of (the linear part of) the original equation for u, as can be seen by applying P_ε^{-1} to the first line of (1.4). This issue has already been noticed in [7], 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.

This paper is organized as follows: In the next section we shall introduce some notations and definitions. Then in Section 3, we shall study the dispersive properties of $S_\varepsilon(t)$ 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 $\mathbf{x} = (x, y) \in \mathbb{R}^{d-k} \times \mathbb{R}^k$ 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 $L^p(\mathbb{R}^{d-k}_x; L^q(\mathbb{R}^k_y))$, which will be shortly denoted by $L^p_x L^q_y$. These spaces are equipped with the following norms:

$$\|f\|_{L^p_x L^q_y} := \left(\int_{\mathbb{R}^{d-k}} \left(\int_{\mathbb{R}^k} |f(x,y)|^q \, dy \right)^{\frac{p}{q}} dx \right)^{\frac{1}{p}}.$$

We denote the usual Fourier transform of a function $f = f(x, y)$ as

$$(\mathcal{F}f)(\xi, \eta) \equiv \widehat{f}(\xi, \eta) = \frac{1}{(2\pi)^{d/2}} \iint_{\mathbb{R}^d} f(x,y) e^{-i(x\cdot\xi + y\cdot\eta)} \, dx \, dy,$$

whereas the partial Fourier transform with respect to the y-variable only will be denoted by

$$(\mathcal{F}_{y\to\eta} f)(x, \eta) \equiv \tilde{f}(x, \eta) = \frac{1}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} f(x,y) e^{-iy\cdot\eta} \, dy.$$

Analogously, we denote the partial Fourier transform in x by $\mathcal{F}_{x\to\xi}$.

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

$$\|f\|_{L^p_x H^s_y} := \left\| P_1^{s/2} f \right\|_{L^p_x L^2_y} \equiv \|(1 + |\eta|^2)^{s/2} \tilde{f}\|_{L^p_x L^2_\eta}.$$

Obviously, the Fourier symbol corresponding to $P_1^{1/2}$ is nothing but the well-known Japanese bracket $\langle \eta \rangle = (1 + |\eta|^2)^{1/2}$ used in the definition of H^s. Incorporating the small parameter $0 < \varepsilon \leqslant 1$ comes at the expense of some (possibly) ε-dependent constants: Indeed, for $s \geqslant 0$, we have

$$\varepsilon^s \|f\|_{H^s} \leqslant \|P_\varepsilon^{s/2} f\|_{L^2} \leqslant \|f\|_{H^s}, \tag{2.1}$$

as well as

$$\|f\|_{H^{-s}} \leqslant \|P_\varepsilon^{-s/2} f\|_{L^2} \leqslant \varepsilon^{-s} \|f\|_{H^{-s}}. \tag{2.2}$$

From now on, we shall write $a \lesssim b$ whenever there exists a universal constant $C > 0$, independent of ε, such that $a \leqslant Cb$. In general this constant C may change from inequality to inequality.

Furthermore, for any time interval $I \subset \mathbb{R}$ we will also make use of the mixed space-time spaces $L^q(I_t, L^p(\mathbb{R}^{d-k}_x; H^s(\mathbb{R}^k_y)))$ briefly denoted by $L^q_t L^p_x H^s_y(I)$, or simply $L^q_t L^p_x H^s_y$, whenever the time interval is clear. These spaces are equipped with the norm

$$\|F\|_{L^q_t L^p_x H^s_y} := \left(\int_I \|F(t)\|^q_{L^p_x H^s_y} \, dt \right)^{\frac{1}{q}}.$$

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

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

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

where we omit the endpoint case, i.e., $(q, r) \neq (2, \frac{2(d-k)}{(d-k-2)_+})$ for $d - k \geqslant 2$.

Clearly, if $k = 0$, this is just the usual admissibility condition for nonendpoint Strichartz pairs corresponding to the Schrödinger group $S_0(t) = e^{it\Delta}$ acting on $\mathbb{R}^d$.

3. Dispersive properties with partial off-axis variation

In this section, we shall derive Strichartz estimates associated to $S_\varepsilon(t) = e^{itP_\varepsilon^{-1}\Delta}$ in the case of partial off-axis variation, i.e. $d > k$. To this end we first derive a set of basic dispersion estimates associated to this linear propagator.

3.1. Dispersion estimate for $\mathbf{S}_\varepsilon(\mathbf{t})$.

Recall the notation $\delta(r) \geqslant 0$ introduced in Definition 2.1. 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 $S_\varepsilon(t) = e^{itP_\varepsilon^{-1}\Delta}$ continuously maps*

$$L^{r'}(\mathbb{R}^{d-k}_x; H^{\delta(r)}(\mathbb{R}^k_y)) \to L^r(\mathbb{R}^{d-k}_x; H^{-\delta(r)}(\mathbb{R}^k_y)), \quad \textit{for} \ \ \frac{1}{r} + \frac{1}{r'} = 1,$$

and it holds that

$$\|S_\varepsilon(t) f\|_{L^r_x H^{-\delta(r)}_y} \leqslant |4\pi t|^{-\delta(r)} \|f\|_{L^{r'}_x H^{\delta(r)}_y}. \tag{3.1}$$

Proof. The estimate (3.1) will in itself be a consequence of the following inequality, which is more directly linked to the explicit form of our propagator $S_\varepsilon(t) = e^{itP_\varepsilon^{-1}\Delta}$:

$$\|S_\varepsilon(t) f\|_{L^r_x L^2_y} \leqslant |4\pi t|^{-\delta(r)} \|P_\varepsilon^{\delta(r)} f\|_{L^{r'}_x L^2_y}. \tag{3.2}$$

Indeed, if we replace f by $P_\varepsilon^{-\frac{\delta(r)}{2}} f$ in (3.2) and keep in mind the basic estimates (2.2) and (2.1), we obtain (3.1) through the string of inequalities

$$\begin{aligned} \|S_\varepsilon(t) f\|_{L^r_x H^{-\delta(r)}_y} &\leqslant \|S_\varepsilon(t) P_\varepsilon^{-\frac{\delta(r)}{2}} f\|_{L^r_x L^2_y} \leqslant |4\pi t|^{-\delta(r)} \|P_\varepsilon^{\frac{\delta(r)}{2}} f\|_{L^{r'}_x L^2_y} \\ &\leqslant |4\pi t|^{-\delta(r)} \|f\|_{L^{r'}_x H^{\delta(r)}_y}, \end{aligned}$$

which also ensures the continuity of $S_\varepsilon(t)$. We also point out that there are no ε-dependent constants involved in any of these inequalities.

In order to prove (3.2), we first note that by density, it is enough to show this for $f \in \mathcal{S}(\mathbb{R}^d)$, the space of smooth and rapidly decaying functions. Moreover, we shall argue by duality and rather prove that for $f, g \in \mathcal{S}(\mathbb{R}^d)$,

$$|\langle S_\varepsilon(t) f, g\rangle_{L^2}| \leqslant |4\pi t|^{-\delta(r)} \|P_\varepsilon^{\delta(r)} f\|_{L^{r'}_x L^2_y} \|g\|_{L^{r'}_x L^2_y}. \tag{3.3}$$

In the trivial case $r = 2$, $\delta(r) = 0$, this estimate directly follows by Cauchy–Schwarz and the fact that $S_\varepsilon(t)$ is unitary on L^2:

$$|\langle S_\varepsilon(t) f, g\rangle_{L^2}| \leqslant \|S_\varepsilon(t) f\|_{L^2} \|g\|_{L^2} = \|f\|_{L^2} \|g\|_{L^2}. \tag{3.4}$$

Next, we treat the case $r = \infty$, $\delta(r) = \frac{d-k}{2}$, i.e., we want to show that for $f, g \in \mathcal{S}(\mathbb{R}^d)$ it holds that

$$|\langle S_\varepsilon(t) f, g\rangle_{L^2}| \leqslant |4\pi t|^{-\frac{d-k}{2}} \|P_\varepsilon^{\frac{d-k}{2}} f\|_{L^1_x L^2_y} \|g\|_{L^1_x L^2_y}. \tag{3.5}$$

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

$$\begin{aligned} \langle S_\varepsilon(t) f, g\rangle_{L^2} &= \big\langle \, \widehat{(S_\varepsilon(t) f)}, \widehat{g} \, \big\rangle_{L^2} = \iint\limits_{\mathbb{R}^{d-k}\times\mathbb{R}^k} e^{-\frac{i(|\eta|^2+|\xi|^2)t}{1+\varepsilon^2|\eta|^2}} \widehat{f}(\xi,\eta)\overline{\widehat{g}(\xi,\eta)}\, d\xi\, d\eta \\ &= \int\limits_{\mathbb{R}^k} e^{-\frac{i|\eta|^2 t}{1+\varepsilon^2|\eta|^2}} \Big(\int\limits_{\mathbb{R}^{d-k}} e^{-\frac{i|\xi|^2 t}{1+\varepsilon^2|\eta|^2}} \widehat{f}(\xi,\eta)\overline{\widehat{g}(\xi,\eta)}\, d\xi \Big) d\eta. \end{aligned}$$

Here, we first compute the inner integral by writing out the partial Fourier transform in ξ on $\widehat{g}$ to obtain

$$
\begin{aligned}
&\int\limits_{\mathbb{R}^{d-k}} e^{-\frac{i|\xi|^2 t}{1+\varepsilon^2|\eta|^2}}\widehat{f}(\xi,\eta)\overline{\widehat{g}(\xi,\eta)}\,d\xi = \\
&\qquad = \frac{1}{(2\pi)^{\frac{d-k}{2}}}\int\limits_{\mathbb{R}^{d-k}} e^{-\frac{i|\xi|^2 t}{1+\varepsilon^2|\eta|^2}}\widehat{f}(\xi,\eta)\int\limits_{\mathbb{R}^{d-k}} e^{ix\cdot\xi}\overline{\tilde{g}(x,\eta)}\,dxd\xi \\
&\qquad = \int\limits_{\mathbb{R}^{d-k}} \overline{\tilde{g}(x,\eta)}\Big(\frac{1}{(2\pi)^{\frac{d-k}{2}}}\int\limits_{\mathbb{R}^{d-k}} e^{ix\cdot\xi}e^{-\frac{i|\xi|^2 t}{1+\varepsilon^2|\eta|^2}}\widehat{f}(\xi,\eta)\,d\xi\Big)dx \\
&\qquad = \int\limits_{\mathbb{R}^{d-k}} \overline{\tilde{g}(x,\eta)}\mathcal{F}^{-1}_{\xi\to x}\Big(e^{-\frac{i|\cdot|^2 t}{1+\varepsilon^2|\eta|^2}}\widehat{f}(\cdot,\eta)\Big)(x)dx,
\end{aligned}
\tag{3.6}
$$

where we have used Fubini's theorem to change the order of integration. We now recall that for $a\in\mathbb{R}$,

$$
\mathcal{F}^{-1}_{\xi\to x}\Big(e^{-\frac{i|\cdot|^2 t}{a}}\Big)(z) = \Big(\frac{a}{2it}\Big)^{\frac{d-k}{2}} e^{\frac{ia|z|^2}{4t}}.
$$

By setting $a = 1+\varepsilon^2|\eta|^2$, we can express the integrand in the last line of (3.6) as

$$
\begin{aligned}
\mathcal{F}^{-1}_{\xi\to x}\Big(e^{-\frac{i|\cdot|^2 t}{1+\varepsilon^2|\eta|^2}}\widehat{f}(\cdot,\eta)\Big)(x) &= \frac{1}{(2\pi)^{\frac{d-k}{2}}}\Big(\mathcal{F}^{-1}_{\xi\to x}\Big(e^{-\frac{i|\cdot|^2 t}{1+\varepsilon^2|\eta|^2}}\Big)*\tilde{f}(\cdot,\eta)\Big)(x) \\
&= \Big(\frac{1+\varepsilon^2|\eta|^2}{4\pi it}\Big)^{\frac{d-k}{2}}\int\limits_{\mathbb{R}^{d-k}} e^{\frac{i(1+\varepsilon^2|\eta|^2)|x-z|^2}{4t}}\tilde{f}(z,\eta)\,dz.
\end{aligned}
\tag{3.7}
$$

Now it is clear by (3.6) and (3.7) that

$$
\begin{aligned}
&(4\pi it)^{\frac{d-k}{2}}\langle S_\varepsilon(t)f,g\rangle_{L^2} \\
&= \iiint\limits_{\mathbb{R}^k\times(\mathbb{R}^{d-k})^2}(1+\varepsilon^2|\eta|^2)^{\frac{d-k}{2}}e^{-\frac{i|\eta|^2 t}{1+\varepsilon^2|\eta|^2}}e^{\frac{i(1+\varepsilon^2|\eta|^2)|x-z|^2}{4t}}\tilde{f}(z,\eta)\overline{\tilde{g}(x,\eta)}\,dz\,dx\,d\eta.
\end{aligned}
$$

This implies the following estimate:

$$
|\langle S_\varepsilon(t)f,g\rangle_{L^2}| \leqslant |4\pi t|^{-\frac{d-k}{2}}\int\limits_{(\mathbb{R}^{d-k})^2}\Big(\int\limits_{\mathbb{R}^k}(1+\varepsilon^2|\eta|^2)^{\frac{d-k}{2}}|\tilde{f}(z,\eta)||\tilde{g}(x,\eta)|\,d\eta\Big)dx\,dz.
$$

A Cauchy–Schwarz inequality in η, followed by Plancherel's identity, then gives

$$
\begin{aligned}
|\langle S_\varepsilon(t)f,g\rangle_{L^2}| &\leqslant |4\pi t|^{-\frac{d-k}{2}}\iint\limits_{(\mathbb{R}^{d-k})^2}\Big(\|\widehat{P}_\varepsilon^{\frac{d-k}{2}}\tilde{f}(z,\cdot)\|_{L^2_\eta}\|\tilde{g}(x,\cdot)\|_{L^2_\eta}\Big)dx\,dz \\
&\leqslant |4\pi t|^{-\frac{d-k}{2}}\|P_\varepsilon^{\frac{d-k}{2}}f\|_{L^1_xL^2_y}\|g\|_{L^1_xL^2_y},
\end{aligned}
$$

which is the desired estimate (3.5). Notice that by replacing $f\mapsto |4\pi t|^{\frac{(d-k)}{2}}P_\varepsilon^{-\frac{d-k}{2}}f$ in (3.5), this yields that the operator

$$
|4\pi t|^{\frac{d-k}{2}}S_\varepsilon(t)P_\varepsilon^{-\frac{d-k}{2}}\;:\;L^1_xL^2_y\to L^\infty_xL^2_y\quad\text{is bounded,}
\tag{3.8}
$$

with norm

$$
\||4\pi t|^{\frac{d-k}{2}}S_\varepsilon(t)P_\varepsilon^{-\frac{d-k}{2}}\|\leqslant 1.
$$

We have thus proved (3.2) in the two endpoint cases $r=2$ and $r=\infty$. The intermediate cases of (3.2) then follow by Stein's interpolation theorem [18, 19].

To this end, we consider, for any $z \in \Omega := \{0 \leqslant \mathrm{Re} z \leqslant 1\} \subset \mathbb{C}$, the family of interpolating operators T_z given by

$$\mathcal{F}(T_z f)(\xi, \eta) = |4\pi t|^{\frac{(d-k)z}{2}} (1+\varepsilon^2|\eta|^2)^{-\frac{d-k}{2}z} e^{-it(1+\varepsilon^2|\eta|^2)^{-1}(|\xi|^2+|\eta|^2)} \widehat{f}(\xi, \eta).$$

Clearly, for $z = 0$, this is nothing but the Fourier transform of $S_\varepsilon(t)$, which we know to be bounded $L^2 \to L^2$ in view of (3.4). For $z = 1$, we obtain the second endpoint case given by (3.8). In addition, it is straightforward to check that $\{T_z\}_{z\in\Omega}$ is an admissible family of linear operators satisfying the hypotheses of Theorem V.4.1 in [19]. The theorem then requires us to bound T_z at the edges of the strip Ω:

For $\mu \in \mathbb{R}$, the following estimate for $z = 0 + i\mu$ uses (3.4) and Plancherel in y, to give

$$\begin{aligned} |\langle T_{0+i\mu} f, g\rangle_{L^2}| &= |\langle \tilde{S}_\varepsilon(t)\big((|4\pi t|^{-1}\widehat{P}_\varepsilon)^{\frac{-i(d-k)\mu}{2}} \tilde{f}\big), \tilde{g}\rangle_{L^2}| \\ &= \|e^{\frac{-i(d-k)\mu}{2}\ln(|4\pi t|^{-1}\widehat{P}_\varepsilon)} \tilde{f}\|_{L^2_x L^2_\eta} \|g\|_{L^2} = \|f\|_{L^2_x L^2_y} \|g\|_{L^2_x L^2_y}. \end{aligned}$$

The estimate for $z = 1 + i\mu$ follows similarly, but now using (3.5), so that

$$\begin{aligned} |\langle T_{1+i\mu} f, g\rangle_{L^2}| &= |4\pi t|^{\frac{(d-k)}{2}} |\langle \tilde{S}_\varepsilon(t)\big((|4\pi t|^{-1}\widehat{P}_\varepsilon)^{-i\frac{(d-k)\mu}{2}} \widehat{P}_\varepsilon^{-\frac{(d-k)}{2}} \tilde{f}\big), \tilde{g}\rangle_{L^2}| \\ &\leqslant \|\widehat{P}_\varepsilon^{\frac{(d-k)}{2}} e^{\frac{-i(d-k)\mu}{2}\ln(|4\pi t|^{-1}\widehat{P}_\varepsilon)} \widehat{P}_\varepsilon^{-\frac{(d-k)}{2}} \tilde{f}\|_{L^1_x L^2_\eta} \|\tilde{g}\|_{L^1_x L^2_\eta} \\ &\leqslant \|f\|_{L^1_x L^2_y} \|g\|_{L^1_x L^2_y}. \end{aligned}$$

Noting that the constants produce no growth in $z \in \mathbb{C}$, then the quoted version of Stein interpolation in [19] implies for $0 \leqslant \theta = 1 - \frac{2}{r} \leqslant 1$ and $r \in [2, \infty]$ the following estimate

$$|4\pi t|^{\delta(r)} \|P_\varepsilon^{-\delta(r)} f\|_{L^r_x L^2_y} = \|T_\theta f\|_{L^r_x L^2_y} \leqslant \|f\|_{L^{r'}_x L^2_y},$$

which by replacing f by $P_\varepsilon^{\delta(r)} f$ and dividing the above inequality by $|4\pi t|^{\delta(r)}$ gives (3.3). Again, we note that there are no ε-dependent constants arising from this interpolation step. Moreover, since the proof of this theorem exploits a density argument using simple functions, the result directly applies also to the mixed spaces $L^r_x L^2_y$ under consideration. □

Remark 3.2. Note that, as $\varepsilon \to 0$, the estimate (3.2) converges to

$$\big\|S_0(t) f\big\|_{L^r_x L^2_y} \leqslant |4\pi t|^{-(d-k)\left(\frac{1}{2}-\frac{1}{r}\right)} \big\|f\big\|_{L^{r'}_x L^2_y},$$

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 obtain dispersion in the y-coordinates when $\varepsilon > 0$. Deriving estimate (3.1) from (3.2) has the advantage that we can use standard Sobolev spaces H^s, independent of ε, to measure the regularity in y (instead of employing the operator P_ε). The price to pay is that (3.1) 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 (3.1).

3.2. **Strichartz estimates.** Exploiting the dispersion estimate (3.1), we shall now prove space-time Strichartz estimates associated to $S_\varepsilon(t)$. These estimates also follow from abstract arguments as in [1, 9, 11]. For the sake of concreteness and due to our somewhat unusual function spaces, we shall give their proof in the nonendpoint case.

Remark 3.3. The case of endpoint Strichartz estimates, i.e., $(q, r) = \big(2, \frac{2(d-k)}{(d-k-2)_+}\big)$ for $d - k \geqslant 2$, in principle could also be dealt with as in [11], 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 $S_\varepsilon(t) = e^{itP_\varepsilon^{-1}\Delta}$ and $(q,r),(\gamma,\rho)$ be two arbitrary admissible Strichartz pairs with $0<\delta(r),\delta(\rho)<1$. Then for any time interval I, there exist constants $C_1, C_2 > 0$, independent of ε and I, such that*

$$\|S_\varepsilon(\cdot)f\|_{L_t^q L_x^r H_y^{-\delta(r)}} \leqslant C_1 \|f\|_{L^2}, \tag{3.9}$$

as well as

$$\left\| \int_0^t S_\varepsilon(\cdot - s)F(s)\, ds \right\|_{L_t^q L_x^r H_y^{-\delta(r)}} \leqslant C_2 \|F\|_{L_t^{\gamma'} L_x^{\rho'} H_y^{\delta(\rho)}}. \tag{3.10}$$

Proof. We start by first noticing that (3.9) is equivalent to saying that the map $f \mapsto S_\varepsilon(t)f$ is bounded as an operator $L^2 \to L_t^q L_x^r H_y^{-\delta(r)}$. Let us define the operator $T_\varepsilon : L_t^{q'} L_x^{r'} H_y^{\delta(r)} \to L^2$ by

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

and note that its formal adjoint T_ε^* is the map $f \mapsto S_\varepsilon(t)f$. Next, we shall show that

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

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

$$\left\| \int_{\mathbb{R}} S_\varepsilon(\cdot - s)F(s)\, ds \right\|_{L_t^q L_x^r H_y^{-\delta(r)}} \leqslant \left\| \int_{\mathbb{R}} \|S_\varepsilon(\cdot - s)F(s)\|_{L_x^r H_y^{-\delta(r)}}\, ds \right\|_{L_t^q},$$

and applying the dispersion estimate (3.1), it follows that

$$\|S_\varepsilon(t-s)F(s)\|_{L_x^r H_y^{-\delta(r)}} \leqslant |4\pi(t-s)|^{-\delta(r)} \|F(s)\|_{L_x^{r'} H_y^{\delta(r)}}.$$

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

$$\begin{aligned} \left\| \int_{\mathbb{R}} S_\varepsilon(\cdot - s)F(s)\, ds \right\|_{L_t^q L_x^r H_y^{-\delta(r)}} &\leqslant \left\| \int_{\mathbb{R}} |4\pi(\cdot - s)|^{-\delta(r)} \|F(s)\|_{L_x^{r'} H_y^{\delta(r)}}\, ds \right\|_{L_t^q} \\ &\leqslant C\|F\|_{L_t^{q'} L_x^{r'} H_y^{\delta(r)}}. \end{aligned}$$

We thus have proven that the operator $T_\varepsilon^* T_\varepsilon : L_t^{q'} L_x^{r'} H_y^{\delta(r)} \to L_t^q L_x^r H_y^{-\delta(r)}$ is bounded. A standard functional analysis result for operators on Banach spaces (see, e.g., [1]) states that

$$\|T_\varepsilon\|^2_{\mathcal{L}(L_t^{q'} L_x^{r'} H_y^{\delta(r)}; L^2)} = \|T_\varepsilon^*\|^2_{\mathcal{L}(L^2 : L_t^q L_x^r H_y^{-\delta(r)})} = \|T_\varepsilon^* T_\varepsilon\|_{\mathcal{L}(L_t^{q'} L_x^{r'} H_y^{\delta(r)}; L_t^q L_x^r H_y^{-\delta(r)})}.$$

This consequently implies that both

$$T_\varepsilon : L_t^{q'} L_x^{r'} H_y^{\delta(r)} \to L^2 \quad \text{and} \quad T_\varepsilon^* : L^2 \to L_t^q L_x^r H_y^{-\delta(r)}$$

are bounded with norms independent of ε. In particular, (3.9) is proved. Furthermore, we note that this holds for any nonendpoint admissible pair (q,r).

Now, choose any arbitrary (nonendpoint) admissible pairs (γ,ρ) and (q,r) such that

$$T_\varepsilon : L_t^{\gamma'} L_x^{\rho'} H_y^{\delta(\rho)} \to L^2 \quad \text{and} \quad T_\varepsilon^* : L^2 \to L_t^q L_x^r H_y^{-\delta(r)}.$$

By combining the estimates for the operators T_ε, T_ε^*, we then infer that

$$T_\varepsilon^* T_\varepsilon : L_t^{\gamma'} L_x^{\rho'} H_y^{\delta(\rho)} \to L_t^q L_x^r H_y^{-\delta(r)}$$

is bounded, i.e.,

$$\left\| \int_{\mathbb{R}} S_\varepsilon(\cdot - s)F(s)\, ds \right\|_{L_t^q L_x^r H_y^{-\delta(r)}} \leqslant C\|F\|_{L_t^{\gamma'} L_x^{\rho'} H_y^{\delta(\rho)}},$$

for any arbitrary $(q,r),(\gamma,\rho)$. We can then invoke Theorem 1.2 from the paper [6] by Christ and Kiselev to conclude the retarded estimate

$$\left\| \int_{s<t} S_\varepsilon(\cdot - s)F(s)\, ds \right\|_{L^q_t L^r_x H^{-\delta(r)}_y} \leqslant C\|F\|_{L^{\gamma'}_t L^{\rho'}_x H^{\delta(\rho)}_y}.$$

In summary, this proves the desired result. □

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

In this section we shall give the proof of Theorem 1.1 by proving a global L^2-based well-posedness result for (1.10) with subcritical nonlinearities. In a second step we shall establish the additional H^1-regularity of the solution.

4.1. **Well-posedness in terms of v.** We rewrite (1.10) using Duhamel's formulation, i.e.,

$$v(t) = S_\varepsilon(t)v_0 + i\int_0^t S_\varepsilon(t-s)P_\varepsilon^{-1/2}(|P_\varepsilon^{-1/2}v|^{2\sigma}P_\varepsilon^{-1/2}v)(s)\, ds =: \Phi(v)(t). \tag{4.1}$$

For the sake of brevity, we shall also write

$$\Phi(v)(t) = S_\varepsilon(t)v_0 + \mathcal{N}(v)(t)$$

and denote

$$\mathcal{N}(v)(t) := i\int_0^t S_\varepsilon(t-s)P_\varepsilon^{-1/2}g(P_\varepsilon^{-1/2}v(s))\, ds, \tag{4.2}$$

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 $T>0$ and choose the admissible pair*

$$(\gamma,\rho) = \left(\frac{4(\sigma+1)}{(d-k)\sigma}, 2(\sigma+1)\right).$$

Then, in the space-time slab $\mathbb{R}^d \times [0,T]$ the inequality

$$\begin{aligned}&\|\mathcal{N}(v) - \mathcal{N}(v')\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y}\\ &\lesssim \varepsilon^{-2(\sigma+1)}T^{1-\frac{(d-k)\sigma}{2}}\left(\|v\|^{2\sigma}_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} + \|v'\|^{2\sigma}_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y}\right)\|v-v'\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y},\end{aligned}$$

holds, provided $0<\sigma\leqslant \frac{2}{(d-2)_+}$.

The case $k=0$ is classical and thus we will only give the proof for $d>k>0$.

Proof. We first note that for our pair (γ,ρ) to be non-endpoint admissible for $d-k\geqslant 2$, we require that $\gamma>2$, which in turn is equivalent to $\sigma < \frac{2}{(d-k-2)_+}$. However, this condition will always be fulfilled since

$$\sigma \leqslant \frac{2}{(d-2)_+} < \frac{2}{(d-k-2)_+}.$$

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 (3.10) and note that

$$\begin{aligned}\|\mathcal{N}(v) - \mathcal{N}(v')\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} &\leqslant C_2\|P_\varepsilon^{-1/2}(g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{-1/2}v'))\|_{L^{\gamma'}_t L^{\rho'}_x H^{\delta(\rho)}_y}\\ &\leqslant \varepsilon^{-1}C_2\|g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{-1/2}v')\|_{L^{\gamma'}_t L^{\rho'}_x H^{-(1-\delta(\rho))}_y},\end{aligned}$$

where we have also used the scaling (2.2) to obtain the factor ε^{-1}. Next, by a Sobolev embedding we have that $H^s(\mathbb{R}^k) \hookrightarrow L^\rho(\mathbb{R}^k)$, where

$$s = k\left(\frac{1}{2} - \frac{1}{2(\sigma+1)}\right) = \frac{k\sigma}{2(\sigma+1)} \in \Big(0, \frac{k}{2}\Big).$$

In turn, this also implies the dual embedding $L^{\rho'}(\mathbb{R}^k) \hookrightarrow H^{-s}(\mathbb{R}^k)$. Now, if we impose that

$$1 \geqslant s + \delta(\rho) = \frac{d\sigma}{2(\sigma+1)},$$

which is so whenever $\sigma \leqslant \frac{2}{(d-2)_+}$, then $H^{-s}(\mathbb{R}^k) \hookrightarrow H^{-(1-\delta(\rho))}(\mathbb{R}^k)$. Together these allow us to estimate

$$\begin{aligned}\|g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{-1/2}v')\|_{H_y^{-(1-\delta(\rho))}} &\leqslant \|g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{-1/2}v')\|_{H_y^{-s}} \\ &\leqslant C_\sigma \|(|P_\varepsilon^{-1/2}v|^{2\sigma} + |P_\varepsilon^{-1/2}v'|^{2\sigma})P_\varepsilon^{-1/2}(v-v')\|_{L_y^{\rho'}} = (*),\end{aligned}$$

where we have also used that for all $z, w \in \mathbb{C}$,

$$|g(z) - g(w)| \leqslant C_\sigma(|z|^{2\sigma} + |w|^{2\sigma})|z-w|.$$

Now, recall that $\rho = 2(\sigma+1)$ and hence $\frac{1}{\rho'} = \frac{2\sigma}{\rho} + \frac{1}{\rho}$. Thus, by first applying Hölder's inequality and using (2.2), we obtain

$$\begin{aligned}(*) &\lesssim (\|P_\varepsilon^{-1/2}v\|_{L_y^\rho}^{2\sigma} + \|P_\varepsilon^{-1/2}v'\|_{L_y^\rho}^{2\sigma})\|P_\varepsilon^{-1/2}(v-v')\|_{L_y^\rho} \\ &\lesssim \varepsilon^{-(2\sigma+1)}(\|v\|_{H_y^{-(1-s)}}^{2\sigma} + \|v'\|_{H_y^{-(1-s)}}^{2\sigma})\|v-v'\|_{H_y^{-(1-s)}} \\ &\lesssim \varepsilon^{-(2\sigma+1)}(\|v\|_{H_y^{-\delta(\rho)}}^{2\sigma} + \|v'\|_{H_y^{-\delta(\rho)}}^{2\sigma})\|v-v'\|_{H_y^{-\delta(\rho)}},\end{aligned}$$

where the last inequality follows from $H^{-\delta(\rho)}(\mathbb{R}^k) \hookrightarrow H^{-(1-s)}(\mathbb{R}^k)$, by the same arguments as before. Employing Hölder's inequality once more in x, we consequently infer

$$\begin{aligned}&\|g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{-1/2}v')\|_{L_x^{\rho'}H_y^{-(1-\delta(\rho))}} \\ &\qquad \lesssim \varepsilon^{-(2\sigma+1)}(\|v\|_{L_x^\rho H_y^{-\delta(\rho)}}^{2\sigma} + \|v'\|_{L_x^\rho H_y^{-\delta(\rho)}}^{2\sigma})\|v-v'\|_{L_x^\rho H_y^{-\delta(\rho)}}.\end{aligned}$$

From here, we compute that

$$\frac{1}{\gamma'} = 1 - \frac{(d-k)\sigma}{2} + \frac{2\sigma}{\gamma} + \frac{1}{\gamma}\,.$$

Thus, taking the $L^{\gamma'}$ norm in t and applying Hölder's inequality yields the result of the lemma. □

Using Lemma 4.1, 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 4.1, i.e., we require $\sigma < \frac{2}{d-k}$. Since Lemma 4.1 holds for $\sigma \leqslant \frac{2}{(d-2)_+}$, we need to distinguish the cases $k \leqslant 2$ and $k > 2$ in the following.

One notices immediately that for $k \leqslant 2$, we have that $\frac{2}{d-k} \leqslant \frac{2}{(d-2)_+}$, which in turn implies that, in this case, we require the stronger assumption $\sigma < \frac{2}{d-k}$ to ensure $\alpha > 0$. However, for $k > 2$ (and thus $d > 3$), it holds that

$$\frac{2}{d-2} < \frac{2}{d-k} < \frac{2}{(d-k-2)_+},$$

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

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

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

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

- *either $k \leqslant 2$ and $0 \leqslant \sigma < \frac{2}{d-k}$*
- *or $k > 2$ and $0 \leqslant \sigma \leqslant \frac{2}{d-2}$.*

Then for any $v_0 \in L^2(\mathbb{R}^d)$, there exists a unique global solution to (1.10)

$$v \in C(\mathbb{R}_t, L^2(\mathbb{R}^d)) \cap L^q_{\text{loc}}(\mathbb{R}_t; L^r(\mathbb{R}^{d-k}_x; H^{-\delta(r)}(\mathbb{R}^k_y)))$$

for any (nonendpoint) 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 $v = P_\varepsilon^{1/2} u$, this directly yields a global-in-time solution $u \in C(\mathbb{R}; L^2(\mathbb{R}^{d-k}_x; H^1(\mathbb{R}^k_y)))$ to (1.4) and thus proves Theorem 1.1. Note that here continuous dependence on the initial data precisely means that for $T > 0$ the map $v_0 \mapsto v|_{[-T,T]}$ is continuous as a map

$$L^2(\mathbb{R}^d) \to C([-T, T], L^2(\mathbb{R}^d)) \cap L^q([-T, T]; L^r(\mathbb{R}^{d-k}_x; H^{-\delta(r)}(\mathbb{R}^k_y))).$$

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 and denote $I = [0, T]$, and set

$$X_{T,M} = \{v \in L^\infty_t L^2(I) \cap L^q_t L^r_x H^{-\delta(r)}_y(I) : \|v\|_{L^\infty_t L^2} + \|v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} \leqslant M\}.$$

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

$$d(v, w) = \|v - w\|_{L^\infty_t L^2_{x,y}} + \|v - w\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y}.$$

Let $v \in X_{T,M}$. Then the Strichartz estimates obtained in Proposition 3.4 together with Lemma 4.1 imply that

$$\begin{aligned}\|\Phi(v)\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} &\leqslant \|S_\varepsilon(\cdot)v_0\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} + \|\mathcal{N}(v)\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} \\ &\leqslant C_{\sigma,\varepsilon}\Big(\|v_0\|_{L^2} + T^{1-\frac{(d-k)\sigma}{2}} M^{2\sigma+1}\Big),\end{aligned}$$

as well as

$$\begin{aligned}\|\Phi(v)\|_{L^\infty_t L^2} &\leqslant \|v_0\|_{L^2} + C_2\|P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2} v)\|_{L^{\gamma'}_t L^{\rho'}_x H^{\delta(\rho)}_y} \\ &\leqslant C_{\sigma,\varepsilon}\Big(\|v_0\|_{L^2} + T^{1-\frac{(d-k)\sigma}{2}} M^{2\sigma+1}\Big).\end{aligned}$$

Together, these yield

$$\|\Phi(v)\|_{L^\infty_t L^2} + \|\Phi(v)\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y} \leqslant 2C_{\sigma,\varepsilon}\Big(\|v_0\|_{L^2} + T^{1-\frac{(d-k)\sigma}{2}} M^{2\sigma+1}\Big).$$

We now choose M such that

$$3M = 8C_{\sigma,\varepsilon}\|v_0\|_{L^2}$$

and choose $T > 0$ such that

$$2C_{\sigma,\varepsilon} T^{1-\frac{(d-k)\sigma}{2}} M^{2\sigma+1} \leqslant \frac{M}{4}. \tag{4.3}$$

Then it follows that $\Phi(v) \in X_{T,M}$ for all $v \in X_{T,M}$ so that $\Phi(X_{T,M}) \subset X_{T,M}$. Now, let $v, w \in X_{T,M}$. Then by Lemma (4.1) and using (4.3) we have

$$\|\mathcal{N}(v) - \mathcal{N}(w)\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}} \leqslant 2C_{\sigma,\varepsilon} M^{2\sigma} T^{1-\frac{(d-k)\sigma}{2}} \|v - w\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}} \leqslant \frac{1}{4}\|v - w\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}}, \tag{4.4}$$

which together with the same estimate for the $L_t^\infty H^1$-norm gives

$$d(\Phi(v), \Phi(w)) \leqslant \frac{1}{2} d(v, w), \quad \forall v, w \in X_{T,M}.$$

Thus Φ is a contraction map on $X_{T,M}$ and Banach's fixed point theorem yields the existence of a unique fixed point $v \in X_{T,M}$. Furthermore, since the solution v satisfies the integral equation (4.1), we infer continuity in time, i.e., $v \in C(I; L^2(\mathbb{R}^d))$.

Moreover, if $v \in X_{T,M}$, then $v \in L_t^q L_x^r H_y^{-\delta(r)}(I)$ for any admissible pair (q, r), since by our Strichartz estimates

$$\|v\|_{L_t^q L_x^r H_y^{-\delta(r)}} \equiv \|\Phi(v)\|_{L_t^q L_x^r H_y^{-\delta(r)}} \leqslant C_1\|v_0\|_{L^2} + C_2\|P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2} v)\|_{L_t^{\gamma'} L_x^{\rho'} H_y^{\delta(\rho)}},$$

which is estimated as in the proof of Lemma 4.1.

Step 2 (Uniqueness): Let $I = [0, T]$ and $v, w \in C(I; L^2) \cap L_t^q L_x^r H_y^{-\delta(r)}(I)$ be two solutions to (4.1) with $\varphi = v_0 = w_0$. Then as in Step 1, we have $v, w \in X_{T,M}$ with $3M = 8C_{\sigma,\varepsilon}\|\varphi\|_{L^2}$ and T given by (4.3). Since the difference of v and w is given by

$$(v - w)(t) = \mathcal{N}(v)(t) - \mathcal{N}(w)(t),$$

then we can apply (4.4) from Step 1 on the interval I to obtain

$$\|v - w\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}(I)} \leq \frac{1}{4}\|v - w\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}(I)}.$$

From this we conclude (local) uniqueness

$$\|v - w\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}(I)} = 0,$$

i.e., $v = w$ on $I = [0, T]$.

In addition, the solution depends continuously on the initial data, as can be seen by taking two solutions $v, \tilde{v}$ on a common time interval $I_c = \min\{I, \tilde{I}\}$. Then by what was done above, we have that $v, \tilde{v} \in X_{T,M}$ with $3M = 8\max\{\|v_0\|_{L^2}, \|\tilde{v}_0\|_{L^2}\}$ and $T = |I_c|$ satisfying (4.3) so that

$$d(v, \tilde{v}) \leqslant \|v_0 - \tilde{v}_0\|_{L^2} + \frac{1}{2} d(v, \tilde{v}),$$

which proves the continuous dependence on the initial data, after extending the argument to the interval I_c.

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

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

We claim that

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

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

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

This gives a contradiction to the definition of $T_{\max}$.

Next, we shall prove that the L^2-norm of v is conserved along the time-evolution. To this end, we adapt an elegant argument given in [16], which has the advantage that it does not require an approximation procedure using a sequence of sufficiently smooth solutions (as is classically done, see, e.g., [4]). First note that by Step 1 we have $v \in C([0,T]; L^2(\mathbb{R}^d))$ for any $T < T_{\max}$. We then rewrite Duhamel's formula (4.1), using the continuity of the semigroup S_ε to propagate backward in time

$$S_\varepsilon(-t)v(t) = v_0 + S_\varepsilon(-t)\mathcal{N}(v)(t). \tag{4.5}$$

The fact that $S_\varepsilon(\cdot)$ is unitary in L^2 implies $\|v(t)\|_{L^2} = \|S_\varepsilon(-t)v(t)\|_{L^2}$. The latter can be expressed using the above identity to obtain

$$\begin{aligned}\|v(t)\|_{L^2}^2 &= \|v_0\|_{L^2}^2 + 2\mathrm{Re}\left\langle S_\varepsilon(-t)\mathcal{N}(v)(t), v_0\right\rangle_{L^2} + \|S_\varepsilon(-t)\mathcal{N}(v)(t)\|_{L^2}^2 \\ &=: \|v_0\|_{L^2}^2 + \mathcal{I}_1 + \mathcal{I}_2.\end{aligned}$$

We want to show that $\mathcal{I}_1 + \mathcal{I}_2 = 0$. In view of (4.2) we can rewrite

$$\begin{aligned}\mathcal{I}_1 &= -2\mathrm{Im}\left\langle \int_0^t S_\varepsilon(-s)P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2}v)(s)\, ds, v_0\right\rangle_{L^2} \\ &= -2\mathrm{Im}\int_0^t \left\langle P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2}v)(s), S_\varepsilon(s)v_0\right\rangle_{L^2} ds.\end{aligned}$$

By duality in y and Hölder's inequality in both x and t we find that this quantity is indeed finite

$$|\mathcal{I}_1| \leqslant 2\|P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2}v)\|_{L_t^{\gamma'}L_x^{\rho'}H_y^{\delta(\rho)}} \|S_\varepsilon(s)v_0\|_{L_t^{\gamma}L_x^{\rho}H_y^{-\delta(\rho)}} < \infty.$$

Denoting for simplicity $G_\varepsilon(\cdot) = P_\varepsilon^{-1/2} g(P_\varepsilon^{-1/2}v)(\cdot)$, we perform the following computation:

$$\begin{aligned}\mathcal{I}_2 &\equiv \left\langle \int_0^t S_\varepsilon(-s)G_\varepsilon(s)\, ds, \int_0^t S_\varepsilon(-s')G_\varepsilon(s')\, ds'\right\rangle_{L^2} \\ &= \int_0^t \left\langle S_\varepsilon(-s)G_\varepsilon(s), \Big(\int_0^s + \int_s^t\Big) S_\varepsilon(-s')G_\varepsilon(s')\, ds'\right\rangle_{L^2} ds \\ &= \int_0^t \left\langle G_\varepsilon(s), \int_0^s S_\varepsilon(s-s')G_\varepsilon(s')\, ds'\right\rangle_{L^2} ds \\ &\qquad + \int_0^t \int_s^t \left\langle S_\varepsilon(s'-s)G_\varepsilon(s), G_\varepsilon(s')\right\rangle_{L^2} ds'\, ds \\ &= \int_0^t \left\langle G_\varepsilon(s), -i\mathcal{N}(v)(s)\right\rangle_{L^2} ds + \int_0^t \left\langle \int_0^{s'} S_\varepsilon(s'-s)G_\varepsilon(s)\, ds, G_\varepsilon(s')\right\rangle_{L^2} ds' \\ &= 2\mathrm{Re}\int_0^t \left\langle G_\varepsilon(s), -i\mathcal{N}(v)(s)\right\rangle_{L^2} ds.\end{aligned}$$

Using the integral formulation (4.5), we can express $-i\mathcal{N}(v)(s)$ and write

$$\mathcal{I}_2 = 2\mathrm{Re}\Big(\int_0^t \left\langle G_\varepsilon(s), iS_\varepsilon(s)v_0\right\rangle_{L^2} ds + \int_0^t \left\langle G_\varepsilon(s), -iv(s)\right\rangle_{L^2} ds\Big). \tag{4.6}$$

Here we note that the particular form of our nonlinearity implies

$$\mathrm{Re}\left\langle G_\varepsilon(\cdot), -iv(\cdot)\right\rangle_{L^2} = \mathrm{Im}\left\langle g(P_\varepsilon^{-1/2}v)(\cdot), P_\varepsilon^{-1/2}v(\cdot)\right\rangle_{L^2} = \mathrm{Im}\,\|P_\varepsilon^{-1/2}v(\cdot)\|_{L^{2\sigma+2}}^{2\sigma+2} = 0,$$

and thus the second term on the right-hand side of (4.6) simply vanishes. In summary, we find

$$\mathcal{I}_2 = 2\mathrm{Re}\int_0^t \left\langle G_\varepsilon(s), iS_\varepsilon(s)v_0\right\rangle_{L^2} ds = 2\mathrm{Im}\int_0^t \left\langle S_\varepsilon(-s)G_\varepsilon(s)\, ds, v_0\right\rangle_{L^2} \equiv -\mathcal{I}_1,$$

which proves that

$$\|v(t)\|_{L^2} = \|v_0\|_{L^2} \quad \forall t \in [0,T].$$

This conservation law allows us to reapply Step 1 as many times as we wish, thereby preserving the length of the maximal interval in each iteration, and yielding $T_{\max} = +\infty$. Since the equation is time-reversible modulo complex conjugation, this yields a global solution for all $t \in \mathbb{R}$. □

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 $v_0 \in H^1$, then the corresponding solution $v(t,\cdot)$ remains in H^1 for all times $t \in \mathbb{R}$. We will prove this property by exploiting the Strichartz estimates stated in Proposition 3.4 and the global well-posedness result in L^2. Similar arguments can be used to obtain a solution $v(t,\cdot) \in H^s$, $s \geqslant 1$, provided the nonlinearity is sufficiently smooth.

Proposition 4.3. *Let $v \in C(\mathbb{R}_t, L^2(\mathbb{R}^d)) \cap L^q_{\rm loc}(\mathbb{R}_t; L^r(\mathbb{R}^{d-k}_x; H^{-\delta(r)}(\mathbb{R}^k_y)))$ be the solution obtained in Proposition 4.2 with initial data $v_0 \in L^2(\mathbb{R}^d)$. If, in addition, $v_0 \in H^1(\mathbb{R}^d)$, then $v \in C(\mathbb{R}_t; H^1(\mathbb{R}^d))$.*

Proof. Let us fix a $0 < T < \infty$. We are going to show that

$$\|\nabla v\|_{L^\infty_t L^2([0,T])} \leqslant K(T, \|\nabla v_0\|_{L^2}). \tag{4.7}$$

Having in mind the conservation property of the L^2-norm of v, this estimate is sufficient to conclude the desired result.

To obtain (4.7), we first recall from Proposition 4.2 that

$$\|v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y([0,T])} \leqslant C(T, \|v_0\|_{L^2}) =: C_T,$$

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

$$\|v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)} \leqslant \lambda, \quad j = 1, \ldots, N. \tag{4.8}$$

First we estimate the gradient of (4.2) by a similar strategy as in Lemma 4.1 with $v' = 0$. By applying the Strichartz estimate (3.10) and the appropriate embeddings in y gives

$$\|\nabla \mathcal{N}(v)\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)} \leqslant \varepsilon^{-1} C_2 \|\nabla g(P_\varepsilon^{1/2} v)\|_{L^{\gamma'}_t L^{\rho'}_x L^{\rho'}_y}.$$

Since the nonlinearity is smooth, this allows us to estimate in y as follows:

$$\begin{aligned}\|\nabla g(P_\varepsilon^{1/2} v)\|_{L^{\rho'}_y} &\leqslant (2\sigma+1)\|P_\varepsilon^{-1/2} v\|^{2\sigma}_{L^\rho_y} \|P_\varepsilon^{-1/2}\nabla v\|_{L^\rho_y} \\ &\lesssim \varepsilon^{-(2\sigma+1)} \|v\|^{2\sigma}_{H^{-\delta(\rho)}_y} \|\nabla v\|_{H^{-\delta(\rho)}_y}.\end{aligned}$$

Combining this with a Hölder estimate in x and t, similarly as in Lemma 4.1 above, we obtain

$$\|\nabla \mathcal{N}(v)\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)} \lesssim \varepsilon^{-2(\sigma+1)} |I_j|^{1-\frac{(d-k)\sigma}{2}} \lambda^{2\sigma} \|\nabla v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)}.$$

Hence on each subinterval I_j we have that

$$\begin{aligned}&\|\nabla v\|_{L^\infty_t L^2(I_j)} + \|\nabla v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)} \\ &\qquad \leqslant C_\varepsilon \Big(\|\nabla v_{j-1}\|_{L^2} + |I_j|^{1-\frac{(d-k)\sigma}{2}} \lambda^{2\sigma} \|\nabla v\|_{L^\gamma_t L^\rho_x H^{-\delta(\rho)}_y(I_j)}\Big)\end{aligned}$$

for $j = 1, \dots, N$ where we write ∇v_{j-1} to denote $\nabla v(t_{j-1})$. Now choose $\lambda = \lambda(C_\varepsilon, T)$ such that

$$C_\varepsilon T^{1-\frac{(d-k)\sigma}{2}} \lambda^{2\sigma} < 1.$$

Since $|I_j| \leqslant T$ we infer the estimate

$$\|\nabla v\|_{L_t^\infty L^2(I_j)} + \|\nabla v\|_{L_t^\gamma L_x^\rho H_y^{-\delta(\rho)}(I_j)} \leqslant K_j^\varepsilon \|\nabla v_{j-1}\|_{L^2},$$

for some constant K_j^ε which depends on ε. In particular, for $j = 1, \dots, N$ we have

$$\|\nabla v_j\|_{L^2} \leqslant K_j^\varepsilon \|\nabla v_{j-1}\|_{L^2}.$$

Using this, we iterate the argument on each subinterval I_j, $j = 1, \dots, N$, to obtain the desired estimate (4.7). □

Remark 4.4. Notice that we cannot obtain uniform-in-time bounds on the H^1-norm of v by invoking the energy (1.6). Indeed the energy functional, written in terms of v, reads

$$E(t) = \frac{1}{2}\|P_\varepsilon^{-1/2} \nabla v\|_{L^2}^2 - \frac{1}{2(\sigma+1)} \|P_\varepsilon^{-1/2} v\|_{L^{2\sigma+2}}^{2\sigma+2},$$

which cannot provide a uniform bound on the full gradient of v.

The proposition above yields a solution u to (1.4) such that $v(t,\cdot) = P_\varepsilon^{1/2} u(t,\cdot) \in H^1(\mathbb{R}^d)$ globally in time. In particular, since

$$\|u(t,\cdot)\|_{H^1} \leqslant \|P_\varepsilon^{1/2} u(t,\cdot)\|_{H^1},$$

we infer $u(t,\cdot) \in H^1(\mathbb{R}^d)$ for all $t \in \mathbb{R}$, provided $P_\varepsilon^{1/2} u_0 \in H^1$. 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 $0 \leqslant k \leqslant 2$, we see that the estimate obtained in Lemma 4.1 no longer yields a positive power α 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 [5].

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 $I \subset \mathbb{R}_t$:

$$W(I) = L^{\frac{2(d-k+2)}{d-k}}\Big(I \times \mathbb{R}_x^{d-k}; H^{-\frac{d-k}{d-k+2}}(\mathbb{R}_y^k)\Big).$$

Then, we have the following local well-posedness result for v, which directly yields Theorem 1.2 for u via $v = P_\varepsilon^{1/2} u$.

Proposition 5.1. *Let $d-k > 0$ with $k \leqslant 2$, and $\sigma = \frac{2}{d-k}$. Then for any $v_0 \in L^2(\mathbb{R}^d)$, there exist times $0 < T_{\max}, T_{\min} \leqslant \infty$ and a unique maximal solution $v \in C((-T_{\min}, T_{\max}); L^2(\mathbb{R}^d)) \cap W(I)$ to (1.10), where I denotes any closed time interval $I \subset (-T_{\min}, T_{\max})$. Furthermore, $T_{\max} < \infty$ if and only if*

$$\|v\|_{W([0,T_{\max}))} = \infty, \tag{5.1}$$

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 depends not only 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 $t \geqslant 0$, since our arguments follow along the same lines as those in [5, Section 3]; see also [4, Chapter 4.7].

Firstly, given a $T > 0$, we claim that by choosing $\delta > 0$ sufficiently small and such that

$$\|S_\varepsilon(\cdot)v_0\|_{W([0,T])} < \delta, \tag{5.2}$$

we obtain a unique solution $v \in C([0,T]; L^2(\mathbb{R}^d)) \cap W([0,T])$ to (1.10). Indeed, under assumption (5.2), the operator $v \mapsto \Phi(v)$, defined by (4.1) with $\sigma = \frac{2}{d-k}$, admits a unique fixed point in

$$Z_{T,\delta} = \{v \in W([0,T]) \text{ s.t. } \|v\|_{W([0,T])} < 2\delta\}.$$

As in Proposition 4.2, by means of the Strichartz estimates one can then show that $v \in L^q_t L^r_x H^{-\delta(r)}_y(0,T)$ for every admissible pair (q,r). Moreover, since the solution v satisfies the integral equation (4.1), we also infer $v \in C([0,T]; L^2(\mathbb{R}^d))$.

To see that $\Phi(v)$ has a fixed point, we use (4.4) with $\gamma = \rho = \frac{2(d-k+2)}{d-k}$ and (5.2), to obtain

$$\|\Phi(v)\|_{W([0,T])} \leqslant \delta + C_\varepsilon \|v\|^{\frac{4+d-k}{d-k}}_{W([0,T])}.$$

Since $\frac{4+d-k}{d-k} > 1$, choosing δ small enough guarantees that $\Phi : Z_{T,\delta} \to Z_{T,\delta}$. Next, Lemma 4.1 implies the estimate

$$\|\Phi(v) - \Phi(w)\|_{W([0,T])} \leqslant C_\varepsilon \left(\|v\|^{\frac{4}{d-k}}_{W([0,T])} + \|w\|^{\frac{4}{d-k}}_{W([0,T])} \right) \|v - w\|_{W([0,T])}, \tag{5.3}$$

where C_ε is independent of T. Here, the fact that $\frac{4}{d-k} > 0$ and $\delta > 0$ is sufficiently small (independent of v_0 and T) implies that $v \mapsto \Phi(v)$ is a contraction on $Z_{T,\delta}$. That this choice of $\delta > 0$ is always possible follows from our Strichartz estimate and from

$$\|S_\varepsilon(t)v_0\|_{W([0,T])} \xrightarrow{T \to 0} 0. \tag{5.4}$$

Consequently, for $T > 0$ small enough, assumption (5.2) is satisfied, yielding a unique local-in-time solution $v(t,\cdot)$ for $t \in [0,T]$.

By a similar argument as in Proposition 4.2 (see also [4, 5]), one can prove uniqueness by letting $v = \Phi(v), w = \Phi(w) \in W([0,T])$ and having in mind that

$$\left(\|v\|^{\frac{4}{d-k}}_{W([0,T])} + \|w\|^{\frac{4}{d-k}}_{W([0,T])} \right) \xrightarrow{T \to 0} 0.$$

From (5.3), we thus conclude that $v = w$ for $T > 0$ sufficiently small. We can then iterate this argument to find a maximal existence time $0 < T_{\max} \leqslant \infty$ for which the unique solution exists for every admissible pair (q,r).

Next, we shall prove the blow-up alternative (5.1) by contradiction. Namely, let $T_{\max} < \infty$ and let us assume that $\|v\|_{W([0,T_{\max}))} < \infty$. Let $t \in [0, T_{\max})$, then for any $s \in [0, T_{\max} - t)$ we write in view of (4.1) that

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

Applying again Lemma 4.1 we can estimate

$$\|S_\varepsilon(\cdot)v(t)\|_{W([0,T_{max}-t))} \leqslant \|v\|_{W([t,T_{max}))} + C_\varepsilon \|v\|^{\frac{4+d-k}{d-k}}_{W([t,T_{max}))}$$

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

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

This implies we can extend the solution after the time $T_{\max}$, contradicting its maximality.

Finally, in order to conclude global existence of small solutions, we note that, by a global-in-time Strichartz estimate,

$$\|S_\varepsilon(\cdot)v_0\|_{W(\mathbb{R})} \leqslant C_1\|v_0\|_{L^2}.$$

This implies that if $\|v_0\|_{L^2}$ is small enough depending on $\delta > 0$, we have

$$\|S_\varepsilon(\cdot)v_0\|_{W(\mathbb{R})} < \delta.$$

Hence, assumption (5.2) is satisfied for all $T \in \mathbb{R}$ and the same continuity argument as before allows one to repeat the contraction argument with $T = \pm\infty$, cf. [4, Remark 4.7.5]. In summary, this yields a unique global solution $v(t,\cdot) \in L^2(\mathbb{R}^d)$ for sufficiently small initial data. ☐

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 2.1, that in this case, we cannot expect to have any Strichartz estimates (see also [3] for more details). We thus have to resort to a more basic fixed point argument to prove the following result.

Lemma 5.2. *Let $d = k \geqslant 1$ and $\sigma \leqslant \frac{2}{(d-2)_+}$. Then, for any $v_0 \in L^2(\mathbb{R}^d)$, there exists a unique global solution $v \in C(\mathbb{R}_t, L^2(\mathbb{R}^d))$ to* (1.10)*, 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

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

Let $v, v' \in Y_{T,M}$, and recall that $S_\varepsilon(t)$ is unitary on L^2. Using Minkowski's inequality and the scaling argument (2.2) then yields

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

provided $\frac{d\sigma}{2(\sigma+1)} \leqslant 1$, i.e., $\sigma \leqslant \frac{2}{(d-2)_+}$.

By a similar embedding strategy as in Lemma 4.1 one finds

$$\begin{aligned}\|g(P_\varepsilon^{-1/2}v) - g(P_\varepsilon^{1/2}v')\|_{H^{-\frac{d\sigma}{2(\sigma+1)}}} &\leqslant (\|P_\varepsilon^{-1/2}v\|_{L^\rho}^{2\sigma} + \|P_\varepsilon^{-1/2}v'\|_{L^\rho}^{2\sigma})\|P_\varepsilon^{-1/2}(v-v')\|_{L^\rho}\\ &\leqslant \varepsilon^{-(2\sigma+1)}(\|v\|_{L^2}^{2\sigma} + \|v'\|_{L^2}^{2\sigma})\|v-v'\|_{L^2},\end{aligned}$$

which consequently implies that

$$\|\mathcal{N}(v) - \mathcal{N}(v')\|_{L^\infty_t L^2} \leqslant \varepsilon^{-(2\sigma+1)}T(\|v\|_{L^\infty_t L^2}^{2\sigma} + \|v'\|_{L^\infty_t L^2}^{2\sigma})\|v-v'\|_{L^\infty_t L^2}.$$

Choosing $T > 0$ sufficiently small, Banach's fixed point theorem directly yields a local-in-time solution $v \in C([0,T], L^2(\mathbb{R}^d))$. The conservation property of the L^2-norm of v can then be shown analogously as in the proof of Proposition 4.2. This consequently allows us to extend the local solution v for all $t \in \mathbb{R}$. ☐

This directly yields Theorem 1.3 for u, since in the case of full-off axis dispersion $v = P_\varepsilon^{1/2}u \in L^2(\mathbb{R}^d)$ implies $u \in H^1(\mathbb{R}^d)$ for any $\varepsilon > 0$. In addition, the L^2-conservation for v directly yields (1.7), whereas (1.6) is a standard computation, and valid for any H^1-solution u. Finally, it is straightforward to extend the solution to $v(t,\cdot) \in H^s(\mathbb{R}^d)$ for any $s > 0$ provided the initial data satisfies $v_0 \in H^s(\mathbb{R}^d)$.

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. H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier Analysis and Nonlinear Partial Differential Equations.* Grundlehren der Mathematischen Wissenschaften Vol. 343, Springer Verlag, New York, 2011.
2. 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.
3. R. Carles, *On Schrödinger equations with modified dispersion.* Dyn. Partial Differ. Equ. **8** (2011), no. 3, 173–184.
4. T. Cazenave, *Semilinear Schrödinger equations.* Courant Lecture Notes in Mathematics vol. 10, American Mathematical Society, 2003.
5. 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.
6. M. Christ and A. Kiselev, *Maximal Functions Associated to Filtrations.* J. Funct. Anal. **179** (2001), 409–425.
7. 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.
8. G. Fibich, *The nonlinear Schrödinger equation; Singular solutions and optical collapse*. Appl. Math. Sciences vol. 192, Springer Verlag, 2015.
9. J. Ginibre and G. Velo, *Smoothing properties and retarded estimates for some dispersive evolution equations*, Comm. Math. Phys. **144** (1992), no. 1, 163–188.
10. J. A. Goldstein and B.J. Wichnoski, *On the Benjamin-Bona-Mahony equation in higher dimensions*, Nonlin. Anal. **4** (1980), no. 4, 665–675.
11. M. Keel and T. Tao, *Endpoint Strichartz Estimates.* Amer. J. Math. **120** (1998), 955–980.
12. 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.
13. F. Merle, and P. Raphael, *On universality of blow up profile for L^2 critical nonlinear Schrödinger equation*. Invent. Math. **156** (2004), 565–672.
14. 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.
15. 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.
16. T. Ozawa, *Remarks on proofs of conservation laws for nonlinear Schrödinger equations*, Calc. Var. Partial Differ. Equ. 25, No. 3, 403-408 (2006).
17. 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.
18. E. M. Stein, *Interpolation of linear operators.* Trans. Amer. Math. Soc. **83** (1956), 482–492.
19. E. M. Stein and G. Weiss, *Introduction to Fourier analysis on euclidean spaces*, Princeton University Press, 1971.
20. 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'Aquilla, Italy
Email 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
Email 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
Email address: `sparber@uic.edu`