



Dynamics of the nonlinear Klein–Gordon equation in the nonrelativistic limit

S. Pasquali¹

Received: 3 July 2018 / Accepted: 27 October 2018 / Published online: 3 November 2018
© Fondazione Annali di Matematica Pura ed Applicata and Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

We study the nonlinear Klein–Gordon (NLKG) equation on a manifold M in the nonrelativistic limit, namely as the speed of light c tends to infinity. In particular, we consider a higher-order normalized approximation of NLKG (which corresponds to the NLS at order $r = 1$) and prove that when M is a smooth compact manifold or \mathbb{R}^d , the solution of the approximating equation approximates the solution of the NLKG locally uniformly in time. When $M = \mathbb{R}^d$, $d \geq 2$, we also prove that for $r \geq 2$ small radiation solutions of the order- r normalized equation approximate solutions of the nonlinear NLKG up to times of order $\mathcal{O}(c^{2(r-1)})$. We also prove a global existence result uniform with respect to c for the NLKG equation on \mathbb{R}^3 with cubic nonlinearity for small initial data and Strichartz estimates for the Klein–Gordon equation with potential on \mathbb{R}^3 .

Keywords Nonrelativistic limit · Nonlinear Klein–Gordon equation · Birkhoff normal form · Long-time behavior

Mathematics Subject Classification 37K55 · 70H08 · 70K45 · 81Q05

1 Introduction

In this paper the nonlinear Klein–Gordon (NLKG) equation in the nonrelativistic limit, namely as the speed of light c tends to infinity, is studied. Formal computations going back to the first half of the last century suggest that, up to corrections of order $\mathcal{O}(c^{-2})$, the system should be described by the nonlinear Schrödinger (NLS) equation. Subsequent mathematical results have shown that the NLS describes the dynamics over timescales of order $\mathcal{O}(1)$.

The nonrelativistic limit for the Klein–Gordon equation on \mathbb{R}^d has been extensively studied over more than 30 years, and essentially all the known results only show convergence of the

This research was supported by by ERC Grant 306414 HamPDEs under FP7.

✉ S. Pasquali
spasquali@mat.uniroma3.it

¹ Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Largo S. Leonardo Murialdo 1, 00146 Rome, Italy

solutions of NLKG to the solutions of the approximate equation for times of order $\mathcal{O}(1)$. The typical statement ensures convergence locally uniformly in time. In a first series of results (see [35,42,57]) it was shown that, if the initial data are in a certain smoothness class, then the solutions converge in a weaker topology to the solutions of the approximating equation. These are informally called “results with loss of smoothness.” Although in this paper a longer time convergence is proved, our results also fill in this group.

Some other results, essentially due to Machihara, Masmoudi, Nakanishi and Ozawa, ensure convergence without loss of regularity in the energy space, again over timescales of order $\mathcal{O}(1)$ (see [36,38,44]).

Concerning radiation solutions there is a remarkable result (see [43]) by Nakanishi, who considered the complex NLKG in the defocusing case, in which it is known that all solutions scatter (and thus the scattering operator exists), and proved that the scattering operator of the NLKG equation converges to the scattering operator of the NLS. It is important to remark that this result is not contained in the one proved here and does not contain it.

Recently Lu and Zhang in [34] proved a result which concerns the NLKG with a quadratic nonlinearity. Here the problem is that the typical scale over which the standard approach allows to control the dynamics is $\mathcal{O}(c^{-1})$, while the dynamics of the approximating equation takes place over timescales of order $\mathcal{O}(1)$. In that work the authors are able to use a normal form transformation (in a spirit quite different from ours) in order to extend the time of validity of the approximation over the $\mathcal{O}(1)$ timescale. We did not try to reproduce or extend that result.

In this paper we prove two kinds of results for the dynamics of NLKG: a global existence result (see Theorem 1) which is uniform for sufficiently large values of $c > 0$, and approximation results (see Theorems 2 and 3) that allow to approximate solutions of NLKG by solutions of suitable higher-order NLS equations. Approximation results are different in the case where the equation lives on \mathbb{R}^d or in a compact manifold: When M is a smooth compact manifold or \mathbb{R}^d the solution of NLS approximates the solution of the original equation locally uniformly in time; when $M = \mathbb{R}^d$, $d \geq 2$, it is possible to prove that for $r > 1$ solutions of the order- r normalized equation approximate solutions of the NLKG equation up to times of order $\mathcal{O}(c^{2(r-1)})$.

The present paper can be thought as an example in which techniques from canonical perturbation theory are used together with results from the theory of dispersive equations in order to understand the singular limit of Hamiltonian PDEs. In this context, the nonrelativistic limit of the NLKG is a relevant example.

The issue of nonrelativistic limit has been studied also in the more general Maxwell–Klein–Gordon system [10,39], in the Klein–Gordon–Zakharov system [40,41], in the Hartree equation [17] and in the pseudo-relativistic NLS [18]. However, all these results proved the convergence of the solutions of the limiting system in the energy space ([17] studied also the convergence in H^k), *locally uniformly in time*; no information could be obtained about the convergence of solutions for longer (in the case of NLKG, which means c -dependent) timescales. On the other hand, in the recent [27], which studies the nonrelativistic limit of the Vlasov–Maxwell system, the authors were able to prove a stability result for solutions which lie in a neighborhood of stable equilibria of the system; this result is valid for times which are polynomial in terms of the inverse of the speed of light, and does not exhibit loss of smoothness.

Other examples of singular perturbation problems that have been studied either with canonical perturbation theory or with multiscale analysis are the problem of the continuous approximation of lattice dynamics (see, e.g., [6,51]) and the semiclassical analysis of Schrödinger operators (see, e.g., [1,46]). In the framework of lattice dynamics, the timescale

covered by all known results is that typical of averaging theorems, which corresponds to our $\mathcal{O}(1)$ timescale. The methods developed in the present paper should allow to extend the time of validity of those results.

The paper is organized as follows. In Sect. 2 we state the results of the paper, together with examples and comments. In Sect. 3 we show Strichartz estimates for the linear KG equation and for the KG equation with potential, as well as a global existence result uniform with respect to c for the cubic NLKG equation on \mathbb{R}^3 . In Sect. 4 we state the main abstract result of the paper. In Sect. 5 we present the proof of the abstract result, which is based on a Galerkin cutoff technique, along with remarks and variant of the result. Next, in Sect. 6 we apply the abstract theorem to the NLKG equation, making explicit computations of the normal form at the first and at the second step. In Sect. 7 we deduce a result about the approximation of solutions locally uniformly in time. In Sect. 8 we study the properties of the normalized equation, namely its dispersive properties in the linear case and its well-posedness for solutions with small initial data in the nonlinear case. In Sect. 9 we discuss the approximation for longer timescales: In particular, to deduce the latter we exploit some dispersive properties of the KG equation reported in Sect. 3. Finally, in “Appendix A” we report all technical lemmata used in Birkhoff normal form estimates (the approach is essentially the same as in [2]), and in “Appendix B” we prove some interpolation theory results for relativistic Sobolev spaces, and we exploit them to deduce Strichartz estimates for the KG equation with potential.

2 Statement of the main results

The NLKG equation describes the motion of a spinless particle with mass $m > 0$. Consider first the real NLKG

$$\frac{\hbar^2}{2mc^2}u_{tt} - \frac{\hbar^2}{2m}\Delta u + \frac{mc^2}{2}u + \lambda|u|^{2(l-1)}u = 0, \tag{1}$$

where $c > 0$ is the speed of light, $\hbar > 0$ is the Planck constant, $\lambda \in \mathbb{R}, l \geq 2, c > 0$.

In the following $m = 1, \hbar = 1$. As anticipated above, one is interested in the behavior of solutions as $c \rightarrow \infty$.

First it is convenient to reduce Eq. (1) to a first-order system, by making the following symplectic change variables

$$\psi := \frac{1}{\sqrt{2}} \left[\left(\frac{\langle \nabla \rangle_c}{c} \right)^{1/2} u - i \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} v \right], \quad v = u_t/c^2,$$

where

$$\langle \nabla \rangle_c := (c^2 - \Delta)^{1/2}, \tag{2}$$

which reduces (1) to the form

$$-i\psi_t = c\langle \nabla \rangle_c \psi + \frac{\lambda}{2^l} \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\psi + \bar{\psi}) \right]^{2l-1}, \tag{3}$$

which is Hamiltonian with Hamiltonian function given by

$$H(\bar{\psi}, \psi) = \langle \bar{\psi}, c\langle \nabla \rangle_c \psi \rangle + \frac{\lambda}{2^l} \int \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \frac{\psi + \bar{\psi}}{\sqrt{2}} \right]^{2l} dx. \tag{4}$$

To state our first result, introduce for any $k \in \mathbb{R}$ and for any $1 < p < \infty$ the following relativistic Sobolev spaces

$$\mathcal{W}_c^{k,p}(\mathbb{R}^3) := \left\{ u \in L^p : \|u\|_{\mathcal{W}_c^{k,p}} := \|c^{-k} \langle \nabla \rangle_c^k u\|_{L^p} < +\infty \right\}, \tag{5}$$

$$\mathcal{H}_c^k(\mathbb{R}^3) := \left\{ u \in L^2 : \|u\|_{\mathcal{H}_c^k} := \|c^{-k} \langle \nabla \rangle_c^k u\|_{L^2} < +\infty \right\}, \tag{6}$$

and remark that the energy space is $\mathcal{H}_c^{1/2}$. Remark that for finite $c > 0$ such spaces coincide with the standard Sobolev spaces, while for $c = \infty$ they are equivalent to the Lebesgue spaces L^p .

In the following the notation $a \lesssim b$ is used to mean: there exists a positive constant K that does not depend on c such that $a \leq Kb$.

We begin with a global existence result for the NLKG (3) in the cubic case, $l = 2$, for small initial data.

Theorem 1 Consider Eq. (3) with $l = 2$ on \mathbb{R}^3 .

There exist $\varepsilon_* > 0$ and $c_* > 0$ such that for any $c > c_*$, if the norm of the initial datum ψ_0 fulfills

$$\|\psi_0\|_{\mathcal{H}_c^{1/2}} \leq \varepsilon_*, \tag{7}$$

then the corresponding solution $\psi(t)$ of (3) exists globally in time:

$$\|\psi(t)\|_{L_t^\infty \mathcal{H}_c^{1/2}} \lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}}. \tag{8}$$

We remark that the constant involved in the estimate (8) does not depend on c .

Remark 1 For finite c this is the standard result for small amplitude solution, while for $c = \infty$ it becomes the standard result for the NLS: Thus Theorem 1 interpolates between these apparently completely different situations. Remark that the lack of a priori estimates for the solutions of NLKG in the limit $c \rightarrow \infty$ was the main obstruction in order to obtain global existence results uniform in c in standard Sobolev spaces.

One is now interested in discussing the approximation of the solutions of NLKG with NLS-type equations. Before giving the result we describe the general strategy we use to get them.

Remark that Eq. (1) is Hamiltonian with Hamiltonian function (4). If one divides the Hamiltonian by a factor c^2 (which corresponds to a rescaling of time) and expands in powers of c^{-2} it takes the form

$$\langle \psi, \bar{\psi} \rangle + \frac{1}{c^2} P_c(\psi, \bar{\psi}) \tag{9}$$

with a suitable function P_c . One can notice that this Hamiltonian is a perturbation of $h_0 := \langle \psi, \bar{\psi} \rangle$, which is the generator of the standard gauge transform and which in particular admits a flow that is periodic in time. Thus the idea is to exploit canonical perturbation theory in order to conjugate such a Hamiltonian system to a system in normal form, up to remainders of order $\mathcal{O}(c^{-2r})$, for any given $r \geq 1$.

The problem is that the perturbation P_c has a vector field which is small only as an operator extracting derivatives. One can Taylor expand P_c and its vector field, but the number of derivatives extracted at each order increases. This situation is typical in singular perturbation problems. Problems of this kind have already been studied with canonical perturbation theory, but the price to pay to get a normal form is that the remainder of the perturbation turns out to be an operator that extracts a large number of derivatives.

In Sect. 6 the normal form equation is explicitly computed in the case $r = 2$:

$$\begin{aligned}
 -i\psi_t &= c^2\psi - \frac{1}{2}\Delta\psi + \frac{3}{4}\lambda|\psi|^2\psi \\
 &+ \frac{1}{c^2}\left[\frac{51}{8}\lambda^2|\psi|^4\psi + \frac{3}{16}\lambda(2|\psi|^2\Delta\psi + \psi^2\Delta\bar{\psi} + \Delta(|\psi|^2\bar{\psi})) - \frac{1}{8}\Delta^2\psi\right], \tag{10}
 \end{aligned}$$

namely a singular perturbation of a gauge-transformed NLS equation. If one, after a gauge transformation, only considers the first-order terms, one has the NLS, for which radiation solution exist (for example, in the defocusing case all solutions are of radiation type). For higher-order NLS there are very few results (see, for example, [37]).

The standard way to exploit such a ‘‘singular’’ normal form is to use it just to construct some approximate solution of the original system, and then to apply Gronwall lemma in order to estimate the difference with a true solution with the same initial datum (see, for example, [4]).

This strategy works also here, but it only leads to a control of the solutions over times of order $\mathcal{O}(c^2)$. When scaled back to the physical time, this allows to justify the approximation of the solutions of NLKG by solutions of the NLS over timescales of order $\mathcal{O}(1)$, on any manifold admitting a Littlewood–Paley decomposition (such as Riemannian smooth compact manifolds, or \mathbb{R}^d ; see the introduction of [12] for the construction of Littlewood–Paley decomposition on manifolds).

Theorem 2 *Let M be a manifold which admits a Littlewood–Paley decomposition, and consider Eq. (3) on M .*

Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$, $1 < p < +\infty$. Then $\exists k_0 = k_0(r) > 0$ with the following properties: For any $k \geq k_1$ there exists $c_{l,r,k,p,R} \gg 1$ such that for any $c > c_{l,r,k,p,R}$, if

$$\|\psi_0\|_{k+k_0,p} \leq R$$

and there exists $T = T_{r,k,p} > 0$ such that the solution ψ_r of the equation in normal form up to order r (98) with the initial datum ψ_0 satisfies

$$\|\psi_r(t)\|_{k+k_0,p} \leq 2R, \quad \text{for } 0 \leq t \leq T,$$

then

$$\|\psi(t) - \psi_r(t)\|_{k,p} \lesssim \frac{1}{c^2}, \quad \text{for } 0 \leq t \leq T. \tag{11}$$

where $\psi(t)$ is the solution of (3) with the initial datum ψ_0 .

A similar result has been obtained for the case $M = \mathbb{T}^d$ by Faou and Schratz, who aimed to construct numerical schemes which are robust in the nonrelativistic limit (see [23]; see also [7,8] and to [9] for the numerical analysis of the nonrelativistic limit of the NLKG).

The idea one uses here in order to improve the timescale of the result is that of substituting Gronwall lemma with a more sophisticated tool, namely dispersive estimates and the retarded Strichartz estimate. This can be done, provided one can prove a dispersive or a Strichartz estimate for the linearization of Eq. (3) on the approximate solution, uniformly in c .

In order to state our approximation result for the linear case, we consider the approximate equation given by the Hamilton equations of the normal form truncated at order $\mathcal{O}(c^{-2r})$, and let ψ_r be a solution of such a linearized normal form equation.

Theorem 3 *Fix $r \geq 1$ and $k_1 \gg 1$. Then $\exists k_0 = k_0(r) > 0$ such that for any $k \geq k_1$, if we denote by ψ_r the solution of the linearized normal equation (105) with the initial datum*

$\psi_0 \in H^{k+k_0}$ and by ψ the solution of the linear KG equation (12) with the same initial datum, then there exists $c^* := c^*(r, k) > 0$ such that for any $c > c^*$

$$\sup_{t \in [0, T]} \|\psi(t) - \psi_r(t)\|_{H_x^k} \lesssim \frac{1}{c^2}, \quad T \lesssim c^{2(r-1)}.$$

This result has been proved in the case $r = 1$ in Appendix A of [14].

Next we consider the approximation of small radiation solutions of the NLKG equation.

Theorem 4 Consider (3) on \mathbb{R}^d , $d \geq 2$. Let $r > 1$, and fix $k_1 \gg 1$. Assume that $l \geq 2$ and $r < \frac{d}{2}(l - 1)$. Then $\exists k_0 = k_0(r) > 0$ such that for any $k \geq k_1$ and for any $\sigma > 0$ the following holds: Consider the solution ψ_r of the normalized equation (98), with the initial datum $\psi_{r,0} \in H^{k+k_0+\sigma+d/2}$. Then there exist $\alpha^* := \alpha^*(d, l, r) > 0$ and there exists $c^* := c^*(r, k) > 1$, such that for any $\alpha > \alpha^*$ and for any $c > c^*$, if $\psi_{r,0}$ satisfies

$$\|\psi_{r,0}\|_{H^{k+k_0+\sigma+d/2}} \lesssim c^{-\alpha},$$

then

$$\sup_{t \in [0, T]} \|\psi(t) - \psi_r(t)\|_{H_x^k} \lesssim \frac{1}{c^2}, \quad T \lesssim c^{2(r-1)},$$

where $\psi(t)$ is the solution of (3) with the initial datum $\psi_{r,0}$.

Remark 2 The assumption of existence of ψ_r up to times of order $\mathcal{O}(c^{2(r-1)})$ is actually a delicate matter. Equation (10), for example, is a quasilinear perturbation of a fourth-order Schrödinger equation (4NLS). Even if we restrict to the case $r = 2$, the issues of global well-posedness and scattering for solutions with large initial data for Eq. (10) have not been solved. For solutions with small initial data, on the other hand, there are some papers dealing with the local well-posedness of 4NLS (see, for example, [28]) and with global well-posedness and scattering of 4NLS (see [50]). In Sec. 8.2 we prove the local well-posedness for times of order $\mathcal{O}(c^{2(r-1)})$ for solutions of the order- r normalized equation with small initial data under the assumptions that $l \geq 2$ and $r < \frac{d}{2}(l - 1)$.

Remark 3 Just to be explicit, we make some examples of Theorem 4. For $M = \mathbb{R}^2$ and a nonlinearity of order $2l$, we can justify the approximation of small radiation solutions up to times of order $\mathcal{O}(c^{2(r-1)})$, for $r < l - 1$. For $M = \mathbb{R}^3$ and a nonlinearity of order $2l$, we can justify the approximation of small radiation solutions up to times of order $\mathcal{O}(c^{2(r-1)})$, for $r < \frac{3}{2}(l - 1)$.

There are some equations, namely the ones in which $\frac{d}{2}(l - 1) \leq 2$, in which we cannot justify the approximation over long timescales (we mention, for example, the cubic NLKG in 2, 3 and 4 dimensions, or the quintic NLKG in 2 dimensions).

There are other well-known solutions of NLS which would be interesting to study; indeed, it is well known that in the case of mixed-type nonlinearity

$$i\psi_t = -\Delta\psi - (|\psi|^2 - |\psi|^4)\psi,$$

such an equation admits linearly stable solitary wave solutions; it can also be proved that the standing waves of NLS can be modified in order to obtain standing wave solutions of the normal form of order r , for any r . It would be of clear interest to prove that true solutions starting close to such standing wave remain close to them for long times (remark that the NLKG does not admit stable standing wave solutions, see [45]); in order to get such a result,

one should prove a Strichartz estimate for NLKG close to the approximate solution and uniformly in c .

Before closing the subsection, a few technical comments are as follows: The first one is that here we develop normal form in the framework of the spaces $W^{k,p}$, while known results in Galerkin averaging theory only allow to deal with the spaces H^k . This is due to the fact that the Fourier analysis is used in order to approximate the derivatives operators with bounded operators. Thus the first technical step needed in order to be able to exploit dispersion is to reformulate Galerkin averaging theory in terms of dyadic decompositions. This is done in Theorem 7.

Second, the condition on r in Theorem 4 depends on the assumption in which we were able to prove a well-posedness result for the normalized equation, which in turn depends on the approach presented recently in [50]; we do not exclude that this technical condition could be improved.

3 Dispersive properties of the Klein–Gordon equation

We briefly recall some classical notion of Fourier analysis on \mathbb{R}^d . Recall the definition of the space of Schwartz (or rapidly decreasing) functions,

$$\mathcal{S} := \left\{ f \in C^\infty(\mathbb{R}^d, \mathbb{R}) \mid \sup_{x \in \mathbb{R}^d} (1 + |x|^2)^{\alpha/2} |\partial^\beta f(x)| < +\infty, \quad \forall \alpha \in \mathbb{N}^d, \forall \beta \in \mathbb{N}^d \right\}.$$

In the following $\langle x \rangle := (1 + |x|^2)^{1/2}$.

Now, for any $f \in \mathcal{S}$ the Fourier transform of f , $\hat{f} : \mathbb{R}^d \rightarrow \mathbb{R}$, is defined by the following formula

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-i\langle x, \xi \rangle} dx, \quad \forall \xi \in \mathbb{R}^d,$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in \mathbb{R}^d .

At the beginning we obtain Strichartz estimates for the linear equation

$$-i \psi_t = c \langle \nabla \rangle_c \psi, \quad x \in \mathbb{R}^d. \tag{12}$$

Proposition 1 *Let $d \geq 2$. For any Schrödinger-admissible couples (p, q) and (r, s) , namely such that*

$$\begin{aligned} 2 &\leq p, r \leq \infty, \\ 2 &\leq q, s \leq \frac{2d}{d-2}, \\ \frac{2}{p} + \frac{d}{q} &= \frac{d}{2}, \quad \frac{2}{r} + \frac{d}{s} = \frac{d}{2}, \\ (p, q, d), (r, s, d) &\neq (2, +\infty, 2), \end{aligned}$$

one has

$$\left\| \langle \nabla \rangle_c^{\frac{1}{q} - \frac{1}{p}} e^{it c \langle \nabla \rangle_c} \psi_0 \right\|_{L_t^p L_x^q} \lesssim c^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} \|\langle \nabla \rangle_c^{1/2} \psi_0\|_{L^2}, \tag{13}$$

$$\left\| \langle \nabla \rangle_c^{\frac{1}{q} - \frac{1}{p}} \int_0^t e^{i(t-s) c \langle \nabla \rangle_c} F(s) ds \right\|_{L_t^p L_x^q} \lesssim c^{\frac{1}{q} - \frac{1}{p} + \frac{1}{s} - \frac{1}{r} - 1} \left\| \langle \nabla \rangle_c^{-\frac{1}{s} + 1} F \right\|_{L_t^r L_x^s}. \tag{14}$$

Remark 4 By choosing $p = +\infty$ and $q = 2$, we get the following a priori estimate for finite energy solutions of (12),

$$\left\| c^{1/2} \langle \nabla \rangle_c^{1/2} e^{it c \langle \nabla \rangle_c} \psi_0 \right\|_{L_t^\infty L_x^2} \lesssim \left\| c^{1/2} \langle \nabla \rangle_c^{1/2} \psi_0 \right\|_{L^2}.$$

We also point out that, since the operators $\langle \nabla \rangle$ and $\langle \nabla \rangle_c$ commute, the above estimates in the spaces $L_t^p L_x^q$ extend to estimates in $L_t^p W_x^{k,q}$ for any $k \geq 0$.

Proof We recall a result reported by D’Ancona–Fanelli in [21] for the operator $\langle \nabla \rangle := \langle \nabla \rangle_1$.

Lemma 1 For all (p, q) Schrödinger-admissible exponents

$$\| e^{i\tau \langle \nabla \rangle} \phi_0 \|_{L_\tau^p W_y^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}, q}} = \left\| \langle \nabla \rangle^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} e^{it \langle \nabla \rangle} \phi_0 \right\|_{L_\tau^p L_y^q} \leq \| \phi_0 \|_{L_y^2}.$$

Now, the solution of Eq.(12) satisfies $\hat{\psi}(t, \xi) = e^{ic(\xi)ct} \hat{\psi}_0(\xi)$. We then define $\eta := \xi/c$, in order to have that

$$\hat{\phi}(c^2t, \eta) := \hat{\psi}(t, c\eta) = \hat{\psi}(t, \xi),$$

and in particular that $\hat{\phi}_0(\eta) = \hat{\psi}_0(\xi)$.

Since

$$\langle \xi \rangle_c = \sqrt{c^2 + |\xi|^2} = c\sqrt{1 + |\xi|^2/c^2}, \tag{15}$$

we get

$$\begin{aligned} \hat{\phi}(t, \eta) &= e^{it c^2 \langle \xi/c \rangle} \hat{\phi}_0(\xi/c) \\ &= e^{it c^2 \langle \eta \rangle} \hat{\phi}_0(\eta) \\ &= e^{i\tau \langle \eta \rangle} \hat{\phi}_0(\eta) \end{aligned}$$

if we set $\tau := c^2t$. Now, by setting $y := cx$ a simple scaling argument leads to

$$\| e^{i\tau \langle \nabla \rangle} \phi_0 \|_{L_\tau^p L_y^q} \lesssim \left\| \langle \nabla \rangle^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} \phi_0 \right\|_{L^2} = \left\| \langle \eta \rangle^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} \hat{\phi}_0 \right\|_{L^2}$$

and since

$$\begin{aligned} \| \langle \eta \rangle^k \hat{\phi}_0 \|_{L^2}^2 &= \int_{\mathbb{R}^d} \langle \eta \rangle^{2k} |\hat{\phi}_0(\eta)|^2 d\eta \\ &= \int_{\mathbb{R}^d} \left\langle \frac{\xi}{c} \right\rangle^{2k} |\hat{\phi}_0(\eta/c)|^2 \frac{d\xi}{c^d} = \frac{1}{c^{2k+d}} \int_{\mathbb{R}^d} \langle \xi \rangle_c^{2k} |\hat{\psi}_0(\xi)|^2 d\xi, \end{aligned}$$

we get

$$\| \langle \eta \rangle^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} \hat{\phi}_0 \|_{L^2} = \frac{1}{c^{\frac{d}{2} - \frac{1}{q} + \frac{1}{p} + \frac{1}{2}}} \left\| \langle \nabla \rangle_c^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} \psi_0 \right\|_{L^2}, \tag{16}$$

while on the other hand

$$\begin{aligned} \psi(t, x) &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle \xi, x \rangle} \hat{\psi}(t, \xi) d\xi = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle \eta, cx \rangle} \hat{\psi}(t, c\eta) c^d d\eta \\ &= (2\pi)^{-d/2} c^d \int_{\mathbb{R}^d} e^{i\langle \eta, cx \rangle} \hat{\phi}(c^2t, \eta) d\eta = c^d \phi(c^2t, cx), \end{aligned}$$

yields

$$\|\psi\|_{L_t^p L_x^q} = c^{d-d/q-2/p} \|\phi\|_{L_t^p L_x^q}. \tag{17}$$

Hence we can deduce (13); via a scaling argument, we can also deduce (14). \square

One important application of the Strichartz estimates for the free Klein–Gordon equation is Theorem 1, namely a global existence result *uniform with respect to c* for the NLKG equation (3) on \mathbb{R}^3 with cubic nonlinearity ($l = 2$), for small initial data.

Proof (Theorem 1) It just suffices to apply Duhamel formula,

$$\psi(t) = e^{it\Delta_c} \psi_0 + i \frac{\lambda}{2^2} \int_0^t e^{i(t-s)\Delta_c} \left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} \left[\left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right]^3,$$

and Proposition 1 with $p = +\infty$ and $q = 2$, in order to get that

$$\|\psi(t)\|_{L_t^\infty \mathcal{H}_c^{1/2}} \lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + c^{1/s-1/r} \left\| \nabla_c^{1/r-1/s} \left[\left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right]^3 \right\|_{L_t^{r'} L_x^{s'}},$$

but by choosing $r = +\infty$ and by exploiting Hölder inequality and Sobolev embedding we get

$$\begin{aligned} \|\psi(t)\|_{L_t^\infty \mathcal{H}_c^{1/2}} &\lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right]^3 \right\|_{L_t^1 L_x^2} \\ &\lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right]^2 \right\|_{L_t^1 L_x^3} \left\| \left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right\|_{L_t^\infty L_x^6} \\ &\lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + \left\| \left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right\|_{L_t^2 L_x^6}^2 \left\| \left(\frac{c}{\langle \nabla \rangle_c}\right)^{1/2} (\psi + \bar{\psi}) \right\|_{L_t^\infty L_x^6} \\ &\lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + \|\psi\|_{L_t^2 \mathcal{W}_c^{-1/2,6}}^2 \|\psi\|_{L_t^\infty \mathcal{W}_c^{-1/2,6}} \\ &\lesssim \|\psi_0\|_{\mathcal{H}_c^{1/2}} + \|\psi\|_{L_t^2 \mathcal{W}_c^{-1/3,6}}^2 \|\psi\|_{L_t^\infty \mathcal{H}_c^{1/2}}, \end{aligned}$$

and one can conclude by a standard continuation argument. \square

We also give a formulation of the Kato–Ponce inequality for the relativistic Sobolev spaces.

Proposition 2 Let $f, g \in \mathcal{S}(\mathbb{R}^d)$, and let $c > 0, 1 < r < \infty$ and $k \geq 0$. Then

$$\|f g\|_{\mathcal{W}_c^{k,r}} \lesssim \|f\|_{\mathcal{W}_c^{k,r_1}} \|g\|_{L^{r_2}} + \|f\|_{L^{r_3}} \|g\|_{\mathcal{W}_c^{k,r_4}}, \tag{18}$$

with

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r_3} + \frac{1}{r_4}, \quad 1 < r_1, r_4 < +\infty.$$

Remark 5 For $c = 1$ Eq. (18) reduces to the classical Kato–Ponce inequality.

Proof We follow an argument by Cordero and Zucco (see Theorem 2.3 in [19]).

We introduce the dilation operator $S_c(f)(x) := f(x/c)$, for any $c > 0$.

Then we apply the classical Kato–Ponce inequality to the rescaled product $S_c(fg) = S_c(f) S_c(g)$,

$$\|S_c(fg)\|_{W^{k,r}} \lesssim \|S_c(f)\|_{W^{k,r_1}} \|S_c(g)\|_{L^{r_2}} + \|S_c(f)\|_{L^{r_3}} \|S_c(g)\|_{W^{k,r_4}}, \tag{19}$$

where

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r_3} + \frac{1}{r_4}, \quad 1 < r_1, r_4 < +\infty.$$

Now, combining the commutativity property

$$\langle \nabla \rangle^k S_c(f)(x) = c^{-k} S_c(\langle \nabla \rangle_c^k f)(x),$$

with the equality $\|S_c(f)\|_{L^r} = c^{-d/r} \|f\|_{L^r}$, we can rewrite (19) as

$$\|\langle \nabla \rangle^k (fg)\|_{L^r} \lesssim \|\langle \nabla \rangle^k f\|_{L^{r_1}} \|g\|_{L^{r_2}} + \|f\|_{L^{r_3}} \|\langle \nabla \rangle^k g\|_{L^{r_4}},$$

and this leads to the thesis. □

We conclude with another dispersive result, which could be interesting in itself: by exploiting the boundedness of the wave operators for the Schrödinger equation, we can deduce Strichartz estimates for the KG equation with potential.

Theorem 5 *Let $c \geq 1$, and consider the operator*

$$\mathcal{H}(x) := c(c^2 - \Delta + V(x))^{1/2} = \mathcal{H}_0(1 + \langle \nabla \rangle_c^{-2} V)^{1/2}, \tag{20}$$

where $V \in C(\mathbb{R}^3, \mathbb{R})$ is a potential such that

$$|V(x)| + |\nabla V(x)| \lesssim \langle x \rangle^{-\beta}, \quad x \in \mathbb{R}^3,$$

for some $\beta > 5$, and that 0 is neither an eigenvalue nor a resonance for the operator $-\Delta + V(x)$. Let (p, q) be a Schrödinger-admissible couple, and assume that $\psi_0 \in \langle \nabla \rangle_c^{-1/2} L^2$ is orthogonal to the bound states of $-\Delta + V(x)$. Then

$$\|\langle \nabla \rangle_c^{\frac{1}{q} - \frac{1}{p}} e^{it\mathcal{H}(x)} \psi_0\|_{L_t^p L_x^q} \lesssim c^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} \|\langle \nabla \rangle_c^{1/2} \psi_0\|_{L^2}. \tag{21}$$

In order to prove Theorem 5 we recall Yajima’s result on wave operators [60] (where we denote by $P_c(-\Delta + V)$ the projection onto the continuous spectrum of the operator $-\Delta + V$).

Theorem 6 *Assume that*

- 0 is neither an eigenvalue nor a resonance for $-\Delta + V$;
- $|\partial^\alpha V(x)| \lesssim \langle x \rangle^{-\beta}$ for $|\alpha| \leq k$, for some $\beta > 5$.

Consider the strong limits

$$\mathcal{W}_\pm := \lim_{t \rightarrow \pm\infty} e^{it(-\Delta + V)} e^{it\Delta}, \quad \mathcal{Z}_\pm := \lim_{t \rightarrow \pm\infty} e^{-it\Delta} e^{it(\Delta - V)} P_c(-\Delta + V).$$

Then $\mathcal{W}_\pm : L^2 \rightarrow P_c(-\Delta + V)L^2$ are isomorphic isometries which extend into isomorphisms $\mathcal{W}_\pm : W^{k,p} \rightarrow P_c(-\Delta + V)W^{k,p}$ for all $p \in [1, +\infty]$, with inverses \mathcal{Z}_\pm . Furthermore, for any Borel function $f(\cdot)$ we have

$$f(-\Delta + V)P_c(-\Delta + V) = \mathcal{W}_\pm f(-\Delta) \mathcal{Z}_\pm, \quad f(-\Delta) = \mathcal{Z}_\pm f(-\Delta + V) P_c(-\Delta + V) \mathcal{W}_\pm. \tag{22}$$

Now, in the case $c = 1$ one can derive Strichartz estimates for $\mathcal{H}(x)$ from the Strichartz estimates for the free KG equation, just by applying the aforementioned theorem by Yajima in the case $k = 1$ (since $1/p - 1/q + 1/2 \in [0, 5/6]$ for all Schrödinger-admissible couples (p, q)). This was already proved in [5] (see Lemma 6.3). In the general case, this follows from an interpolation theory argument, and we defer it to Appendix B.

4 Galerkin averaging method

Consider the scale of Banach spaces $W^{k,p}(M, \mathbb{C}^n \times \mathbb{C}^n) \ni (\psi, \bar{\psi})$ ($k \geq 1, 1 < p < +\infty, n \in \mathbb{N}_0$) endowed by the standard symplectic form. Having fixed k and p , and $U_{k,p} \subset W^{k,p}$ open, we define the gradient of $H \in C^\infty(U_{k,p}, \mathbb{R})$ w.r.t. $\bar{\psi}$ as the unique function s.t.

$$\langle \nabla_{\bar{\psi}} H, \bar{h} \rangle = d_{\bar{\psi}} H \bar{h}, \quad \forall h \in W^{k,p},$$

so that the Hamiltonian vector field of a Hamiltonian function H is given by

$$X_H(\psi, \bar{\psi}) = \left(i \nabla_{\bar{\psi}} H, -i \nabla_{\psi} H \right).$$

The open ball of radius R and center 0 in $W^{k,p}$ will be denoted by $B_{k,p}(R)$.

Now, we call an *admissible family of cutoff (pseudo-differential) operators* a sequence $(\pi_j(D))_{j \geq 0}$, where $\pi_j(D) : W^{k,p} \rightarrow W^{k,p}$ for any $j \geq 0$, such that

- for any $j \geq 0$ and for any $f \in W^{k,p}$

$$f = \sum_{j \geq 0} \pi_j(D) f;$$

- for any $j \geq 0$ $\pi_j(D)$ can be extended to a self-adjoint operator on L^2 , and there exist constants $K_1, K_2 > 0$ such that

$$K_1 \left(\sum_{j \geq 0} \|\pi_j(D) f\|_{L^2}^2 \right)^{1/2} \leq \|f\|_{L^2} \leq K_2 \left(\sum_{j \geq 0} \|\pi_j(D) f\|_{L^2}^2 \right)^{1/2};$$

- for any $j \geq 0$, if we denote by $\Pi_j(D) := \sum_{l=0}^j \pi_l(D)$, there exist positive constants K' (possibly depending on k and p) such that

$$\|\Pi_j f\|_{k,p} \leq K' \|f\|_{k,p} \quad \forall f \in W^{k,p};$$

- there exist positive constants K_1'', K_2'' (possibly depending on k and p) and an increasing and unbounded sequence $(K_j)_{j \in \mathbb{N}} \subset \mathbb{R}_+$ such that

$$K_1'' \|f\|_{W^{k,p}} \leq \left\| \left[\sum_{j \in \mathbb{N}} K_j^{2k} |\pi_j(D) f|^2 \right]^{1/2} \right\|_{L^p} \leq K_2'' \|f\|_{W^{k,p}}. \tag{23}$$

Remark 6 Let $k \geq 0, M$ be either \mathbb{R}^d or the d -dimensional torus \mathbb{T}^d , and consider the Sobolev space $H^k = H^k(M)$. One can readily check that Fourier projection operators on H^k

$$\pi_j \psi(x) := (2\pi)^{-d/2} \int_{j-1 \leq |k| \leq j} \hat{\psi}(k) e^{ik \cdot x} dk, \quad j \geq 1$$

form an admissible family of cutoff operators. In this case we have

$$\Pi_N \psi(x) := (2\pi)^{-d/2} \int_{|k| \leq N} \hat{\psi}(k) e^{ik \cdot x} dk, \quad N \geq 0,$$

and the constants $(K_j)_{j \in \mathbb{N}}$ in (23) are given by $K_j := j$.

Remark 7 Let $k \geq 0, 1 < p < +\infty$; we now introduce the Littlewood–Paley decomposition on the Sobolev space $W^{k,p} = W^{k,p}(\mathbb{R}^d)$ (see [56], Ch. 13.5).

In order to do this, define the cutoff operators in $W^{k,p}$ in the following way: Start with a smooth, radial nonnegative function $\phi_0 : \mathbb{R}^d \rightarrow \mathbb{R}$ such that $\phi_0(\xi) = 1$ for $|\xi| \leq 1/2$, and $\phi_0(\xi) = 0$ for $|\xi| \geq 1$; then, define $\phi_1(\xi) := \phi_0(\xi/2) - \phi_0(\xi)$, and set

$$\phi_j(\xi) := \phi_1(2^{1-j}\xi), \quad j \geq 2. \tag{24}$$

Then $(\phi_j)_{j \geq 0}$ is a partition of unity,

$$\sum_{j \geq 0} \phi_j(\xi) = 1.$$

Now, for each $j \in \mathbb{N}$ and each $f \in W^{k,2}$, we can define $\phi_j(D)f$ by

$$\mathcal{F}(\phi_j(D)f)(\xi) := \phi_j(\xi) \hat{f}(\xi).$$

It is well known that for $p \in (1, +\infty)$ the map $\Phi : L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d, l^2)$,

$$\Phi(f) := (\phi_j(D)f)_{j \in \mathbb{N}},$$

maps $L^p(\mathbb{R}^d)$ isomorphically onto a closed subspace of $L^p(\mathbb{R}^d, l^2)$, and we have compatibility of norms ([56], Ch. 13.5, (5.45)–(5.46)),

$$K'_p \|f\|_{L^p} \leq \|\Phi(f)\|_{L^p(\mathbb{R}^d, l^2)} := \left\| \left[\sum_{j \in \mathbb{N}} |\phi_j(D)f|^2 \right]^{1/2} \right\|_{L^p} \leq K_p \|f\|_{L^p},$$

and similarly for the $W^{k,p}$ -norm, i.e., for any $k > 0$ and $p \in (1, +\infty)$

$$K'_{k,p} \|f\|_{W^{k,p}} \leq \left\| \left[\sum_{j \in \mathbb{N}} 2^{2jk} |\phi_j(D)f|^2 \right]^{1/2} \right\|_{L^p} \leq K_{k,p} \|f\|_{W^{k,p}}. \tag{25}$$

We then define the cutoff operator Π_N by

$$\Pi_N \psi := \sum_{j \leq N} \phi_j(D)\psi. \tag{26}$$

Hence, according to the above definition, the sequence $(\phi_j(D))_{j \geq 0}$ is an admissible family of cutoff operators.

We point out that the Littlewood–Paley decomposition, along with equality (25), can be extended to compact manifolds (see [13]), as well as to some particular noncompact manifolds (see [12]).

Now we consider a Hamiltonian system of the form

$$H = h_0 + \varepsilon h + \varepsilon F, \tag{27}$$

where $\varepsilon > 0$ is a parameter. We fix an admissible family of cutoff operators $(\pi_j(D))_{j \geq 0}$ on $W^{k,p}(\mathbb{R}^d)$. We assume that

PER h_0 generates a linear periodic flow Φ^t with period 2π ,

$$\Phi^{t+2\pi} = \Phi^t \quad \forall t.$$

We also assume that Φ^t is analytic from $W^{k,p}$ to itself for any $k \geq 1$, and for any $p \in (1, +\infty)$;

INV for any $k \geq 1$, for any $p \in (1, +\infty)$, Φ^t leaves invariant the space $\Pi_j W^{k,p}$ for any $j \geq 0$. Furthermore, for any $j \geq 0$

$$\pi_j(D) \circ \Phi^t = \Phi^t \circ \pi_j(D);$$

NF h is in normal form, namely

$$h \circ \Phi^t = h.$$

Next we assume that both the Hamiltonian and the vector field of both h and F admit an asymptotic expansion in ε of the form

$$h \sim \sum_{j \geq 1} \varepsilon^{j-1} h_j, \quad F \sim \sum_{j \geq 1} \varepsilon^{j-1} F_j, \tag{28}$$

$$X_h \sim \sum_{j \geq 1} \varepsilon^{j-1} X_{h_j}, \quad X_F \sim \sum_{j \geq 1} \varepsilon^{j-1} X_{F_j}, \tag{29}$$

and that the following properties are satisfied

HVF There exists $R^* > 0$ such that for any $j \geq 1$

- X_{h_j} is analytic from $B_{k+2j,p}(R^*)$ to $W^{k,p}$;
- X_{F_j} is analytic from $B_{k+2(j-1),p}(R^*)$ to $W^{k,p}$.

Moreover, for any $r \geq 1$ we have that

- $X_{h - \sum_{j=1}^r \varepsilon^{j-1} h_j}$ is analytic from $B_{k+2(r+1),p}(R^*)$ to $W^{k,p}$;
- $X_{F - \sum_{j=1}^r \varepsilon^{j-1} F_j}$ is analytic from $B_{k+2r,p}(R^*)$ to $W^{k,p}$.

The main result of this section is the following theorem.

Theorem 7 Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$, $1 < p < +\infty$. Consider (27), and assume PER, INV (with respect to the Littlewood–Paley decomposition), NF and HVF. Then $\exists k_0 = k_0(r) > 0$ with the following properties: For any $k \geq k_1$ there exists $\varepsilon_{r,k,p} \ll 1$ such that for any $\varepsilon < \varepsilon_{r,k,p}$ there exists $\mathcal{T}_\varepsilon^{(r)} : B_{k,p}(R) \rightarrow B_{k,p}(2R)$ analytic canonical transformation such that

$$H_r := H \circ \mathcal{T}_\varepsilon^{(r)} = h_0 + \sum_{j=1}^r \varepsilon^j \mathcal{Z}_j + \varepsilon^{r+1} \mathcal{R}^{(r)},$$

where \mathcal{Z}_j are in normal form, namely

$$\{\mathcal{Z}_j, h_0\} = 0, \tag{30}$$

and

$$\sup_{B_{k+k_0,p}(R)} \|X_{\mathcal{Z}_j}\|_{W^{k,p}} \leq C_{k,p},$$

$$\sup_{B_{k+k_0,p}(R)} \|X_{\mathcal{Z}^{(r)}}\|_{W^{k,p}} \leq C_{k,p}, \tag{31}$$

$$\sup_{B_{k,p}(R)} \|\mathcal{Z}_\varepsilon^{(r)} - id\|_{W^{k,p}} \leq C_{k,p} \varepsilon. \tag{32}$$

In particular, we have that

$$\mathcal{Z}_1(\psi, \bar{\psi}) = h_1(\psi, \bar{\psi}) + \langle F_1 \rangle (\psi, \bar{\psi}),$$

where $\langle F_1 \rangle (\psi, \bar{\psi}) := \int_0^{2\pi} F_1 \circ \Phi^t(\psi, \bar{\psi}) \frac{dt}{2\pi}$.

5 Proof of Theorem 7

We first make a Galerkin cutoff through the Littlewood–Paley decomposition (see [56], Ch. 13.5).

In order to do this, fix $N \in \mathbb{N}$, $N \gg 1$, and introduce the cutoff operators Π_N in $W^{k,p}$ by

$$\Pi_N \psi := \sum_{j \leq N} \phi_j(D) \psi,$$

where $\phi_j(D)$ are the operators we introduced in Remark 7.

We notice that by assumption INV the Hamiltonian vector field of h_0 generates a continuous flow Φ^t which leaves $\Pi_N W^{k,p}$ invariant.

Now we set $H = H_{N,r} + \mathcal{R}_{N,r} + \mathcal{R}_r$, where

$$H_{N,r} := h_0 + \varepsilon h_{N,r} + \varepsilon F_{N,r}, \tag{33}$$

$$h_{N,r} := \sum_{j=1}^r \varepsilon^{j-1} h_{j,N}, \quad h_{j,N} := h_j \circ \Pi_N, \tag{34}$$

$$F_{N,r} := \sum_{j=1}^r \varepsilon^{j-1} F_{j,N}, \quad F_{j,N} := F_j \circ \Pi_N, \tag{35}$$

and

$$\mathcal{R}_{N,r} := h_0 + \sum_{j=1}^r \varepsilon^j h_j + \sum_{j=1}^r \varepsilon^j F_j - H_{N,r}, \tag{36}$$

$$\mathcal{R}_r := \varepsilon \left(h - \sum_{j=1}^r \varepsilon^{j-1} h_j \right) + \varepsilon \left(F - \sum_{j=1}^r \varepsilon^{j-1} F_j \right). \tag{37}$$

The system described by the Hamiltonian (33) is the one that we will put in normal form. In the following we will use the notation $a \lesssim b$ to mean: there exists a positive constant K independent of N and R (but dependent on r, k and p), such that $a \leq Kb$.

We exploit the following intermediate results:

Lemma 2 For any $k \geq k_1$ and $p \in (1, +\infty)$ there exists $B_{k,p}(R) \subset W^{k,p}$ s.t. $\forall \sigma > 0, N > 0$

$$\sup_{B_{k+\sigma+2(r+1),p}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{W^{k,p}} \lesssim \frac{\varepsilon}{2^{\sigma(N+1)}}, \tag{38}$$

$$\sup_{B_{k+2(r+1),p}(R)} \|X_{\mathcal{R}_r}(\psi, \bar{\psi})\|_{W^{k,p}} \lesssim \varepsilon^{r+1}. \tag{39}$$

Proof We recall that $\mathcal{R}_{N,r} = h_0 + \sum_{j=1}^r \varepsilon^j h_j + \sum_{j=1}^r \varepsilon^j F_j - H_{N,r}$.

Now, $\|id - \Pi_N\|_{W^{k+\sigma,p} \rightarrow W^{k,p}} \lesssim 2^{-\sigma(N+1)}$, since

$$\begin{aligned} \left\| \sum_{j \geq N+1} \phi_j(D) f \right\|_{W^{k,p}} &\lesssim \left\| \left[\sum_{j \geq N+1} |2^{jk} \phi_j(D) f|^2 \right]^{1/2} \right\|_{L^p} \\ &\lesssim 2^{-\sigma(N+1)} \left\| \left[\sum_{j \geq N+1} |2^{j(k+\sigma)} \phi_j(D) f|^2 \right]^{1/2} \right\|_{L^p} \\ &\lesssim 2^{-\sigma(N+1)} \|f\|_{W^{k+\sigma,p}}, \end{aligned}$$

hence

$$\begin{aligned} &\sup_{\psi \in B_{k+2(r+1)+\sigma,p}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{W^{k,p}} \\ &\lesssim \|dX_{\sum_{j=1}^r \varepsilon^j (h_j + F_j)}\|_{L^\infty(B_{k+2(r+1),p}(R), W^{k,p})} \|id - \Pi_N\|_{L^\infty(B_{k+2(r+1)+\sigma,p}(R), B_{k+2(r+1),p}(R))} \\ &\lesssim \varepsilon 2^{-\sigma(N+1)}. \end{aligned}$$

The estimate of $X_{\mathcal{R}_r}$ follows from the hypothesis HVF. □

Lemma 3 Let $j \geq 1$. Then for any $k \geq k_1 + 2(j - 1)$ and $p \in (1, +\infty)$ there exists $B_{k,p}(R) \subset W^{k,p}$ such that

$$\begin{aligned} \sup_{B_{k,p}(R)} \|X_{h_{j,N}}(\psi, \bar{\psi})\|_{k,p} &\leq K_{j,k,p}^{(h)} 2^{2jN}, \\ \sup_{B_{k,p}(R)} \|X_{F_{j,N}}(\psi, \bar{\psi})\|_{k,p} &\leq K_{j,k,p}^{(F)} 2^{2(j-1)N}, \end{aligned}$$

where

$$\begin{aligned} K_{j,k,p}^{(h)} &:= \sup_{B_{k,p}(R)} \|X_{h_j}(\psi, \bar{\psi})\|_{k-2j,p}, \\ K_{j,k,p}^{(F)} &:= \sup_{B_{k,p}(R)} \|X_{F_j}(\psi, \bar{\psi})\|_{k-2(j-1),p}. \end{aligned}$$

Proof It follows from

$$\sup_{\psi \in B_{k,p}(R)} \left\| \sum_{h \leq N} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi}) \right\|_{W^{k,p}} \lesssim \sup_{\psi \in B_{k,p}(R)} \left\| \left[\sum_{h \leq N} |2^{hk} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi})|^2 \right]^{1/2} \right\|_{L^p} \tag{40}$$

$$\leq 2^{2(j-1)N} \sup_{\psi \in B_{k,p}(R)} \left\| \left[\sum_{h \leq N} |2^{h[k-2(j-1)]} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi})|^2 \right]^{1/2} \right\|_{L^p} \tag{41}$$

$$\lesssim 2^{2(j-1)N} \sup_{\psi \in B_{k,p}(R)} \|X_{F_{j,N}}(\psi, \bar{\psi})\|_{k-2(j-1),p} \tag{42}$$

$$= K_{j,k,p}^{(F)} 2^{2(j-1)N}, \tag{43}$$

and similarly for $X_{h_{j,N}}$. □

Next we have to normalize the system (33). In order to do this we need a slight reformulation of Theorem 4.4 in [2]. Here we report a statement of the result adapted to our context.

Lemma 4 *Let $k \geq k_1 + 2r$, $p \in (1, +\infty)$, $R > 0$, and consider the system (33). Assume that $\varepsilon < 2^{-4Nr}$, and that*

$$\left(K_{k,p}^{(F,r)} + K_{k,p}^{(h,r)} \right) r 2^{2Nr} \varepsilon < 2^{-9} e^{-1} \pi^{-1} R, \tag{44}$$

where

$$K_{k,p}^{(F,r)} := \sup_{1 \leq j \leq r} \sup_{\psi \in B_{k,p}(R)} \|X_{F_j}(\psi, \bar{\psi})\|_{k-2(j-1),p},$$

$$K_{k,p}^{(h,r)} := \sup_{1 \leq j \leq r} \sup_{\psi \in B_{k,p}(R)} \|X_{h_j}(\psi, \bar{\psi})\|_{k-2j,p}.$$

Then there exists an analytic canonical transformation $\mathcal{T}_{\varepsilon,N}^{(r)} : B_{k,p}(R) \rightarrow B_{k,p}(2R)$ such that

$$\sup_{B_{k,p}(R/2)} \|\mathcal{T}_{\varepsilon,N}^{(r)}(\psi, \bar{\psi}) - (\psi, \bar{\psi})\|_{W^{k,p}} \leq 4\pi r K_{k,p}^{(F,r)} 2^{2Nr} \varepsilon,$$

and that puts (33) in normal form up to a small remainder,

$$H_{N,r} \circ \mathcal{T}_{\varepsilon,N}^{(r)} = h_0 + \varepsilon h_{N,r} + \varepsilon Z_N^{(r)} + \varepsilon^{r+1} \mathcal{R}_N^{(r)}, \tag{45}$$

with $Z_N^{(r)}$ is in normal form, namely $\{h_{0,N}, Z_N^{(r)}\} = 0$, and

$$\begin{aligned} \sup_{B_{k,p}(R/2)} \|X_{Z_N^{(r)}}(\psi, \bar{\psi})\|_{k,p} &\leq 4 2^{2Nr} \varepsilon \left(r K_{k,p}^{(F,r)} + r K_{k,p}^{(h,r)} \right) r 2^{2Nr} K_{k,p}^{(F,r)} \\ &= 4r^2 K_{k,p}^{(F,r)} (K_{k,p}^{(F,r)} + K_{k,p}^{(h,r)}) 2^{4Nr} \varepsilon, \end{aligned} \tag{46}$$

$$\sup_{B_{k,p}(R/2)} \|X_{\mathcal{R}_N^{(r)}}(\psi, \bar{\psi})\|_{k,p} \tag{47}$$

$$\leq 2^8 e^{\frac{T}{R}} (K_{k,p}^{(F,r)} + K_{k,p}^{(h,r)}) r 2^{2Nr} \tag{48}$$

$$\times \left[\frac{4T}{R} \left(2^9 3^2 e \frac{T}{R} (K_{k,p}^{(F,r)} + K_{k,p}^{(F,r)}) K_{k,p}^{(F,r)} r^2 2^{4Nr} \varepsilon + 5 K_{k,p}^{(h,r)} r 2^{2Nr} + 5 K_{k,p}^{(F,r)} r 2^{2Nr} \right) r \right]^r \tag{49}$$

The proof of Lemma 4 is postponed to ‘‘Appendix A.’’

Remark 8 In the original notation of Theorem 4.4 in [2] we set

$$\begin{aligned} \mathcal{P} &= W^{k,p}, \\ h_\omega &= h_0, \\ \hat{h} &= \varepsilon h_{N,r}, \\ f &= \varepsilon F_{N,r}, \\ f_1 &= r = g \equiv 0, \\ F &= K_{k,p}^{(F,r)} r 2^{2Nr} \varepsilon, \\ F_0 &= K_{k,p}^{(h,r)} r 2^{2Nr} \varepsilon. \end{aligned}$$

Remark 9 Actually, Lemma 4 would also hold under a weaker smallness assumption on ε : It would be enough that $\varepsilon < 2^{-2N}$, and that

$$\varepsilon \left[K_{k,p}^{(F,r)} \frac{1 - 2^{2Nr} \varepsilon^r}{1 - 2^{2N} \varepsilon} + K_{k,p}^{(h,r)} \frac{2^{2N} (1 - 2^{2Nr} \varepsilon^r)}{1 - 2^{2N} \varepsilon} \right] < 2^{-9} e^{-1} \pi^{-1} R \tag{50}$$

is satisfied. However, condition (50) is less explicit than (44), which allows us to apply directly the scheme of [2]. The disadvantage of the stronger smallness assumption (44) is that it holds for a smaller range of ε , and that at the end of the proof it will force us to choose a larger parameter $\sigma = 4r^2$. By using (50) and by making a more careful analysis, it may be possible to prove Theorem 7 also by choosing $\sigma = 2r$.

Now we conclude with the proof of Theorem 7.

Proof Now consider the transformation $\mathcal{T}_{\varepsilon,N}^{(r)}$ defined by Lemma 4, then

$$(\mathcal{T}_{\varepsilon,N}^{(r)})^* H = h_0 + \sum_{j=1}^r \varepsilon^j h_{j,N} + \varepsilon Z_N^{(r)} + \varepsilon^{r+1} \mathcal{R}_N^{(r)} + \varepsilon^r \mathcal{R}_{Gal}$$

where we recall that

$$\varepsilon^r \mathcal{R}_{Gal} := \left(\mathcal{T}_{\varepsilon,N}^{(r)} \right)^* (\mathcal{R}_{N,r} + \mathcal{R}_r).$$

By exploiting Lemma 4 we can estimate the vector field of $\mathcal{R}_N^{(r)}$, while by using Lemma 2 and (275) we get

$$\sup_{B_{k+\sigma+2(r+1),p}(R/2)} \|X_{\mathcal{R}_{Gal}}(\psi, \bar{\psi})\|_{W^{k,p}} \lesssim \left(\frac{\varepsilon}{2^\sigma(N+1)} + \frac{\varepsilon^{r+1}}{\sigma + 2(r+1)} \right). \tag{51}$$

To get the result choose

$$\begin{aligned} k_0 &= \sigma + 2(r+1), \\ N &= r\sigma^{-1} \log_2(1/\varepsilon) - 1, \\ \sigma &= 4r^2. \end{aligned}$$

□

Remark 10 The compatibility condition $N \geq 1$ and (44) lead to

$$\varepsilon \leq \left[2^{-9} e^{-1} \pi^{-1} R (K_{k,p}^{(F,r)} + K_{k,p}^{(h,r)})^{-1} r^{-1} 2^{-2r} \right]^{\frac{\sigma}{2r}} =: \varepsilon_{r,k,p} \leq 2^{-2\sigma/r} \leq 2^{-8r}.$$

Remark 11 We point out the fact that Theorem 7 holds for the scale of Banach spaces $W^{k,p}(M, \mathbb{C}^n \times \mathbb{C}^n)$, where $k \geq 1, 1 < p < +\infty, n \in \mathbb{N}_0$, and where M is a smooth manifold on which the Littlewood–Paley decomposition can be constructed, for example, a compact manifold (see sect. 2.1 in [13]), \mathbb{R}^d , or a noncompact manifold satisfying some technical assumptions (see [12]).

If we restrict to the case $p = 2$, and we consider M as either \mathbb{R}^d or the d -dimensional torus \mathbb{T}^d , we can prove an analogous result for Hamiltonians $H(\psi, \bar{\psi})$ with $(\psi, \bar{\psi}) \in H^k := W^{k,2}(M, \mathbb{C} \times \mathbb{C})$. In the following we denote by $B_k(R)$ the open ball of radius R and center 0 in H^k . We recall that the Fourier projection operator on H^k is given by

$$\pi_j \psi(x) := (2\pi)^{-d/2} \int_{j-1 \leq |k| \leq j} \hat{\psi}(k) e^{ik \cdot x} dk, \quad j \geq 1.$$

Theorem 8 Fix $r \geq 1, R > 0, k_1 \gg 1$. Consider (27), and assume PER, INV (with respect to Fourier projection operators), NF and HVF. Then $\exists k_0 = k_0(r) > 0$ with the following properties: For any $k \geq k_1$ there exists $\varepsilon_{r,k} \ll 1$ such that for any $\varepsilon < \varepsilon_{r,k}$ there exists $\mathcal{T}_\varepsilon^{(r)} : B_k(R) \rightarrow B_k(2R)$ transformation s.t.

$$H_r := H \circ \mathcal{T}_\varepsilon^{(r)} = h_0 + \sum_{j=1}^r \varepsilon^j \mathcal{Z}_j + \varepsilon^{r+1} \mathcal{R}^{(r)},$$

where \mathcal{Z}_j are in normal form, namely

$$\{\mathcal{Z}_j, h_0\} = 0, \tag{52}$$

and

$$\sup_{B_{k+k_0}(R)} \|X_{\mathcal{R}^{(r)}}\|_{H^k} \leq C_k, \tag{53}$$

$$\sup_{B_k(R)} \|\mathcal{T}_\varepsilon^{(r)} - id\|_{H^k} \leq C_k \varepsilon. \tag{54}$$

In particular, we have that

$$\mathcal{Z}_1(\psi, \bar{\psi}) = h_1(\psi, \bar{\psi}) + \langle F_1 \rangle (\psi, \bar{\psi}),$$

where $\langle F_1 \rangle (\psi, \bar{\psi}) := \int_0^{2\pi} F_1 \circ \Phi^t(\psi, \bar{\psi}) \frac{dt}{2\pi}$.

The only technical difference between the proofs of Theorem 7 and the proof of Theorem 8 is that we exploit the Fourier cutoff operator

$$\Pi_N \psi(x) := \int_{|k| \leq N} \hat{\psi}(k) e^{ik \cdot x} dk,$$

as in [3]. This in turn affects (38), which in this case reads

$$\sup_{B_{k+\sigma+2(r+1)}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{H^k} \lesssim \frac{\varepsilon}{N^\sigma}, \tag{55}$$

and (51), for which we have to choose a bigger cutoff, $N = \varepsilon^{-r\sigma}$.

6 Application to the nonlinear Klein–Gordon equation

6.1 The real nonlinear Klein–Gordon equation

We first consider the Hamiltonian of the real nonlinear Klein–Gordon equation with power-type nonlinearity on a smooth manifold M (M is such that the Littlewood–Paley decomposition is well defined; take, for example, a smooth compact manifold, or \mathbb{R}^d). The Hamiltonian is of the form

$$H(u, v) = \frac{c^2}{2} \langle v, v \rangle + \frac{1}{2} \langle u, \langle \nabla \rangle_c^2 u \rangle + \lambda \int \frac{u^{2l}}{2l}, \tag{56}$$

where $\langle \nabla \rangle_c := (c^2 - \Delta)^{1/2}$, $\lambda \in \mathbb{R}$, $l \geq 2$.

If we introduce the complex-valued variable

$$\psi := \frac{1}{\sqrt{2}} \left[\left(\frac{\langle \nabla \rangle_c}{c} \right)^{1/2} u - i \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} v \right], \tag{57}$$

(the corresponding symplectic 2-form becomes $i d\psi \wedge d\bar{\psi}$), the Hamiltonian (56) in the coordinates $(\psi, \bar{\psi})$ is

$$H(\bar{\psi}, \psi) = \langle \bar{\psi}, c \langle \nabla \rangle_c \psi \rangle + \frac{\lambda}{2l} \int \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \frac{\psi + \bar{\psi}}{\sqrt{2}} \right]^{2l} dx. \tag{58}$$

If we rescale the time by a factor c^2 , the Hamiltonian takes the form (27), with $\varepsilon = \frac{1}{c^2}$, and

$$H(\psi, \bar{\psi}) = h_0(\psi, \bar{\psi}) + \varepsilon h(\psi, \bar{\psi}) + \varepsilon F(\psi, \bar{\psi}), \tag{59}$$

where

$$h_0(\psi, \bar{\psi}) = \langle \bar{\psi}, \psi \rangle, \tag{60}$$

$$h(\psi, \bar{\psi}) = \langle \bar{\psi}, (c \langle \nabla \rangle_c - c^2) \psi \rangle \sim \sum_{j \geq 1} \varepsilon^{j-1} \langle \bar{\psi}, a_j \Delta^j \psi \rangle =: \sum_{j \geq 1} \varepsilon^{j-1} h_j(\psi, \bar{\psi}), \tag{61}$$

$$F(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1}l} \int \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\psi + \bar{\psi}) \right]^{2l} dx \tag{62}$$

$$\sim \frac{\lambda}{2^{l+1}l} \int (\psi + \bar{\psi})^{2l} dx$$

$$+ \varepsilon b_2 \int \left[(\psi + \bar{\psi})^{2l-1} \Delta(\psi + \bar{\psi}) + \dots + (\psi + \bar{\psi}) \Delta \left((\psi + \bar{\psi})^{2l-1} \right) \right] dx$$

$$+ \mathcal{O}(\varepsilon^2)$$

$$=: \sum_{j \geq 1} \varepsilon^{j-1} F_j(\psi, \bar{\psi}), \tag{63}$$

where $(a_j)_{j \geq 1}$ and $(b_j)_{j \geq 1}$ are real coefficients, and $F_j(\psi, \bar{\psi})$ is a polynomial function of the variables ψ and $\bar{\psi}$ (along with their derivatives) and which admits a bounded vector field from a neighborhood of the origin in $W^{k+2(j-1), p}$ to $W^{k, p}$ for any $1 < p < +\infty$.

This description clearly fits the scheme treated in the previous section, and one can easily check that assumptions PER, NF and HVF are satisfied. Therefore, we can apply Theorem 7 to the Hamiltonian (59).

Remark 12 About the normal forms obtained by applying Theorem 7, we remark that in the first step (case $r = 1$ in the statement of the theorem) the homological equation we get is of the form

$$\{\chi_1, h_0\} + F_1 = \langle F_1 \rangle, \tag{64}$$

where $F_1(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1}l} \int (\psi + \bar{\psi})^{2l} dx$. Hence the transformed Hamiltonian is of the form

$$H_1(\psi, \bar{\psi}) = h_0(\psi, \bar{\psi}) + \frac{1}{c^2} \left[-\frac{1}{2} \langle \bar{\psi}, \Delta \psi \rangle + \langle F_1 \rangle(\psi, \bar{\psi}) \right] + \frac{1}{c^4} \mathcal{R}^{(1)}(\psi, \bar{\psi}), \tag{65}$$

where

$$\langle F_1 \rangle(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1}l} \binom{2l}{l} \int |\psi|^{2l} dx. \tag{66}$$

If we neglect the remainder and we derive the corresponding equation of motion for the system, we get

$$-i\psi_t = \psi + \frac{1}{c^2} \left[-\frac{1}{2} \Delta \psi + \frac{\lambda}{2^{l+1}} \binom{2l}{l} |\psi|^{2(l-1)} \psi \right], \tag{67}$$

which is the NLS, and the Hamiltonian which generates the canonical transformation is given by

$$\chi_1(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1}l} \sum_{\substack{j=0, \dots, 2l \\ j \neq l}} \frac{1}{i 2^{l-j}} \binom{2l}{j} \int \psi^{2l-j} \bar{\psi}^j dx. \tag{68}$$

Remark 13 Now we iterate the construction by passing to the case $r = 2$.

If we neglect the remainder of order c^{-6} , we have that

$$\begin{aligned} H \circ \mathcal{F}^{(1)} &= h_0 + \frac{1}{c^2} h_1 + \frac{1}{c^4} \{\chi_1, h_1\} + \frac{1}{c^4} h_2 \\ &\quad + \frac{1}{c^2} \langle F_1 \rangle + \frac{1}{c^4} \{\chi_1, F_1\} + \frac{1}{2c^4} \{\chi_1, \{\chi_1, h_0\}\} + \frac{1}{c^4} F_2 \\ &= h_0 + \frac{1}{c^2} [h_1 + \langle F_1 \rangle] + \frac{1}{c^4} \left[\{\chi_1, h_1\} + h_2 + \{\chi_1, F_1\} + \frac{1}{2} \{\chi_1, \langle F_1 \rangle - F_1\} + F_2 \right], \end{aligned} \tag{69}$$

(70)

where $h_1(\psi, \bar{\psi}) = -\frac{1}{2} \langle \bar{\psi}, \Delta \psi \rangle$, and χ_1 is of the form (68).

Now we compute the terms of order $\frac{1}{c^4}$.

$$\begin{aligned} \{\chi_1, h_1\} &= d\chi_1 X_{h_1} = \frac{\partial \chi_1}{\partial \psi} \cdot i \frac{\partial h_1}{\partial \bar{\psi}} - i \frac{\partial \chi_1}{\partial \bar{\psi}} \frac{\partial h_1}{\partial \psi} \\ &= -\frac{\lambda}{2^{l+3}l} \int \left[\sum_{\substack{j=0, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} (2l-j) \psi^{2l-j-1} \bar{\psi}^j \right] \Delta \psi \, dx \\ &\quad + \frac{\lambda}{2^{l+3}l} \int \left[\sum_{\substack{j=1, \dots, 2l \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} j \psi^{2l-j} \bar{\psi}^{j-1} \right] \Delta \bar{\psi} \, dx \\ &= -\frac{\lambda}{2^{l+3}l} \int \Delta \psi \, \psi^{2l-1} + \Delta \bar{\psi} \, \bar{\psi}^{2l-1} \, dx \end{aligned}$$

$$\begin{aligned}
 & - \frac{\lambda}{2^{l+3}l} \int \sum_{\substack{j=1, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} \int (2l-j) \psi^{2l-j-1} \bar{\psi}^j \Delta \psi \\
 & - j \psi^{2l-j} \bar{\psi}^{j-1} \Delta \bar{\psi} \, dx,
 \end{aligned} \tag{71}$$

and since $j \neq l$ in the sum we have that

$$\langle \{\chi_1, h_1\} \rangle = 0. \tag{72}$$

Next,

$$h_2 = -\frac{1}{8} \langle \bar{\psi}, \Delta^2 \psi \rangle, \tag{73}$$

$\{\chi_1, F_1\}$

$$\begin{aligned}
 & = \frac{\lambda^2}{2^{2l+3}l^2} \int \left[\sum_{\substack{j=0, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} (2l-j) \psi^{2l-j-1} \bar{\psi}^j \right] \left[\sum_{h=1}^{2l} \binom{2l}{h} h \psi^{2l-h} \bar{\psi}^{h-1} \right] dx \\
 & - \frac{\lambda^2}{2^{2l+3}l^2} \int \left[\sum_{\substack{j=1, \dots, 2l \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} j \psi^{l-j} \bar{\psi}^{j-1} \right] \left[\sum_{h=0}^{2l-1} \binom{2l}{h} (2l-h) \psi^{2l-h-1} \bar{\psi}^h \right] dx \\
 & = \frac{\lambda^2}{2^{2l+3}l^2} \sum_{\substack{j, h=1, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} \binom{2l}{h} [(2l-j)h - j(2l-h)] \int \psi^{4l-j-h-1} \bar{\psi}^{j+h-1} dx \\
 & + \frac{\lambda^2}{2^{2l+3}l^2} 2 \int \psi^{2l-1} \left[\sum_{h=1}^{2l} \binom{2l}{h} h \psi^{2l-h} \bar{\psi}^{h-1} \right] dx \\
 & + \frac{\lambda^2}{2^{2l+3}l^2} 2l \int \left[\sum_{\substack{j=0, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} (2l-j) \psi^{2l-j-1} \bar{\psi}^j \right] \bar{\psi}^{2l-1} dx \\
 & + \frac{\lambda^2}{2^{2l+3}l^2} 2 \int \bar{\psi}^{2l-1} \left[\sum_{h=0}^{2l-1} \binom{2l}{h} (2l-h) \psi^{2l-h-1} \bar{\psi}^h \right] dx \\
 & - \frac{\lambda^2}{2^{2l+3}l^2} 2l \int \left[\sum_{\substack{j=1, \dots, 2l \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} j \psi^{2l-j} \bar{\psi}^{j-1} \right] \psi^{2l-1} dx,
 \end{aligned} \tag{74}$$

$$\langle \{\chi_1, F_1\} \rangle = \lambda^2 K(l) \int |\psi|^{2(2l-1)} dx, \tag{75}$$

$$K(l) := \frac{1}{2^{2l+3}l^2} \left\{ \left(\sum_{\substack{j, h=1, \dots, 2l-1 \\ j \neq l \\ j+h=2l}} \frac{1}{l-j} \binom{2l}{j} \binom{2l}{h} [(2l-j)h - j(2l-h)] \right) + 16l \right\}, \tag{76}$$

where $K(l) > 0$ by the conditions on j and h in the sum.

Then,

$$\begin{aligned}
 & \{\chi_1, \langle F_1 \rangle\} \\
 &= \frac{\lambda^2}{2^{2l+3}l^2} \binom{2l}{l} \int \sum_{\substack{j=0, \dots, 2l-1 \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} (2l-j)l \psi^{2l-j-1} \bar{\psi}^j \psi^l \bar{\psi}^{l-1} \, dx \\
 &- \frac{\lambda^2}{2^{2l+3}l^2} \binom{2l}{l} \int \sum_{\substack{j=1, \dots, 2l \\ j \neq l}} \frac{1}{l-j} \binom{2l}{j} jl \psi^{2l-j} \bar{\psi}^{j-1} \psi^{l-1} \bar{\psi}^l \, dx \\
 &= \frac{\lambda^2}{2^{2l+3}l^2} \binom{2l}{l} \left[\binom{2l}{l} 2 \int \psi^{3l-1} \bar{\psi}^{l-1} + \psi^{l-1} \bar{\psi}^{3l-1} \, dx \right. \\
 &\quad \left. + \sum_{\substack{j=1, \dots, 2l-1 \\ j \neq l}} 2l \binom{2l}{j} \int \psi^{3l-j-1} \bar{\psi}^{j+l-1} \, dx \right], \tag{77}
 \end{aligned}$$

and since $j \neq l$ in the sum we have that

$$\langle \{\chi_1, \langle F_1 \rangle\} \rangle = 0. \tag{78}$$

Furthermore,

$$\begin{aligned}
 F_2 &= \frac{\lambda}{2^{l+3}l} 2l \int (\psi + \bar{\psi})^{2l-1} \Delta(\psi + \bar{\psi}) \, dx \\
 &= \frac{\lambda}{2^{l+2}} \sum_{j=0}^{2l-1} \binom{2l-1}{j} \int \psi^{2l-j-1} \bar{\psi}^j (\Delta\psi + \Delta\bar{\psi}) \, dx, \tag{79}
 \end{aligned}$$

$$\begin{aligned}
 \langle F_2 \rangle &= \frac{\lambda}{2^{l+2}} \int \binom{2l-1}{l} \psi^{l-1} \bar{\psi}^l \Delta\psi + \binom{2l-1}{l-1} \psi^l \bar{\psi}^{l-1} \Delta\bar{\psi} \, dx \\
 &= \frac{\lambda}{2^{l+2}} \binom{2l-1}{l} \int |\psi|^{2(l-1)} (\bar{\psi} \Delta\psi + \psi \Delta\bar{\psi}) \, dx \tag{80}
 \end{aligned}$$

Hence, up to a remainder of order $O\left(\frac{1}{c^6}\right)$, we have that

$$\begin{aligned}
 H_2 &= h_0 + \frac{1}{c^2} \int \left[-\frac{1}{2} \langle \bar{\psi}, \Delta\psi \rangle + \frac{\lambda}{2^{l+1}l} \binom{2l}{l} |\psi|^{2l} \right] dx \\
 &\quad + \frac{1}{c^4} \int \left[\lambda^2 K(l) |\psi|^{2(2l-1)} + \frac{\lambda}{2^{l+2}} \binom{2l-1}{l} |\psi|^{2(l-1)} (\bar{\psi} \Delta\psi + \psi \Delta\bar{\psi}) - \frac{1}{8} \langle \bar{\psi}, \Delta^2\psi \rangle \right] dx, \tag{81}
 \end{aligned}$$

which, by neglecting h_0 (that yields only a gauge factor) and by rescaling the time, leads to the following equations of motion

$$\begin{aligned}
 -i\psi_t &= -\frac{1}{2} \Delta\psi + \frac{\lambda}{2^{l+1}} \binom{2l}{l} |\psi|^{2(l-1)} \psi + \frac{1}{c^2} \left[-\frac{1}{8} \Delta^2\psi + \lambda^2 K(l) (2l-1) |\psi|^{4(l-1)} \psi \right] \\
 &\quad + \frac{1}{c^2} \left[\frac{\lambda}{2^{l+2}} \binom{2l-1}{l} (l |\psi|^{2(l-1)} \Delta\psi + (l-1) |\psi|^{2(l-2)} \psi^2 \Delta\bar{\psi} + \Delta(|\psi|^{2(l-1)} \bar{\psi})) \right], \tag{82}
 \end{aligned}$$

which, for example, in the case of a cubic nonlinearity ($l = 2$) reads

$$\begin{aligned}
 -i\psi_t &= -\frac{1}{2}\Delta\psi + \frac{3}{4}\lambda|\psi|^2\psi \\
 &+ \frac{1}{c^2}\left[\frac{51}{8}\lambda^2|\psi|^4\psi + \frac{3}{16}\lambda(2|\psi|^2\Delta\psi + \psi^2\Delta\bar{\psi} + \Delta(|\psi|^2\bar{\psi})) - \frac{1}{8}\Delta^2\psi\right]. \tag{83}
 \end{aligned}$$

To the author’s knowledge, Eq. (83) has never been studied before. It is the nonlinear analogue of a linear higher-order Schrödinger equation that appears in [14, 15] in the context of semirelativistic equations. Indeed, the linearization of Eq. (83) is studied within the framework of relativistic quantum field theory, as an approximation of nonlocal kinetic terms; Carles, Lucha and Moulay studied the well-posedness of these approximations, as well as the convergence of the equations as the order of truncation goes to infinity, in the linear case, also when one takes into account the effects of some time-independent potentials (e.g., bounded potentials, the harmonic oscillator potential and the Coulomb potential).

Remark 14 We point out that the case of the one-dimensional cubic defocusing NLKG is also interesting, since for $\lambda = 1$ the normalized equation at first step is the cubic defocusing NLS, which is known to be integrable by the inverse scattering method. It would be interesting to reach a better understanding of the one-dimensional normalized equation, even in the case $r = 2$.

Even though there is a one-dimensional integrable 4NLS equation related to the dynamics of a vortex filament (see [52] and references therein),

$$\begin{aligned}
 i\psi_t + \psi_{xx} + \frac{1}{2}|\psi|^2\psi - v\left[\psi_{xxxx} + \frac{3}{2}|\psi|^2\psi_{xx} + \frac{3}{2}\psi_x^2\bar{\psi} + \frac{3}{8}|\psi|^4\psi + \frac{1}{2}(|\psi|^2)_{xx}\psi\right] \\
 = 0, \quad v \in \mathbb{R} \tag{84}
 \end{aligned}$$

apparently there is no obvious relation between the above equation and Eq. (83).

6.2 The complex nonlinear Klein–Gordon equation

Now we consider the Hamiltonian of the complex nonlinear Klein–Gordon equation with power-type nonlinearity on a smooth manifold M (take, for example, a smooth compact manifold, or \mathbb{R}^d)

$$H(w, p_w) = \frac{c^2}{2} \langle p_w, p_w \rangle + \frac{1}{2} \langle w, \langle \nabla \rangle_c^2 w \rangle + \lambda \int \frac{|w|^{2l}}{2l}, \tag{85}$$

where $w : \mathbb{R} \times M \rightarrow \mathbb{C}$, $\langle \nabla \rangle_c := (c^2 - \Delta)^{1/2}$, $\lambda \in \mathbb{R}$, $l \geq 2$.

If we rewrite the Hamiltonian in terms of $u := \text{Re}(w)$ and $v := \text{Im}(w)$, we have

$$\begin{aligned}
 H(u, v, p_u, p_v) &= \frac{c^2}{2} (\langle p_u, p_u \rangle + \langle p_v, p_v \rangle) + \frac{1}{2} (|\nabla u|^2 + |\nabla v|^2) \\
 &+ \frac{c^2}{2} (u^2 + v^2) + \lambda \int \frac{(u^2 + v^2)^l}{2l}. \tag{86}
 \end{aligned}$$

We consider by simplicity only the cubic case ($l = 2$), but the argument may be readily generalized to the other power-type nonlinearities.

If we introduce the variables

$$\psi := \frac{1}{\sqrt{2}} \left[\left(\frac{\langle \nabla \rangle_c}{c} \right)^{1/2} u - i \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} p_u \right], \tag{87}$$

$$\phi := \frac{1}{\sqrt{2}} \left[\left(\frac{\langle \nabla \rangle_c}{c} \right)^{1/2} v + i \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} p_v \right], \tag{88}$$

(the corresponding symplectic 2-form becomes $i d\psi \wedge d\bar{\psi} - i d\phi \wedge d\bar{\phi}$), the Hamiltonian (85) in the coordinates $(\psi, \phi, \bar{\psi}, \bar{\phi})$ reads

$$H(\psi, \phi, \bar{\psi}, \bar{\phi}) = \langle \bar{\psi}, c \langle \nabla \rangle_c \psi \rangle + \langle \bar{\phi}, c \langle \nabla \rangle_c \phi \rangle \tag{89}$$

$$+ \frac{\lambda}{16} \int_M \left[\left\langle \psi + \bar{\psi}, \frac{c}{\langle \nabla \rangle_c} (\psi + \bar{\psi}) \right\rangle + \left\langle \phi + \bar{\phi}, \frac{c}{\langle \nabla \rangle_c} (\phi + \bar{\phi}) \right\rangle \right]^2 dx, \tag{90}$$

with corresponding equations of motion

$$\begin{cases} -i\psi_t &= c \langle \nabla \rangle_c \psi + \frac{1}{4} \left[\left\langle \psi + \bar{\psi}, \frac{c}{\langle \nabla \rangle_c} (\psi + \bar{\psi}) \right\rangle + \left\langle \phi + \bar{\phi}, \frac{c}{\langle \nabla \rangle_c} (\phi + \bar{\phi}) \right\rangle \right] \frac{c}{\langle \nabla \rangle_c} (\psi + \bar{\psi}), \\ i\phi_t &= c \langle \nabla \rangle_c \phi + \frac{1}{4} \left[\left\langle \psi + \bar{\psi}, \frac{c}{\langle \nabla \rangle_c} (\psi + \bar{\psi}) \right\rangle + \left\langle \phi + \bar{\phi}, \frac{c}{\langle \nabla \rangle_c} (\phi + \bar{\phi}) \right\rangle \right] \frac{c}{\langle \nabla \rangle_c} (\phi + \bar{\phi}). \end{cases}$$

If we rescale the time by a factor c^2 , the Hamiltonian takes the form (27), with $\varepsilon = \frac{1}{c^2}$, and

$$H(\psi, \phi, \bar{\psi}, \bar{\phi}) = H_0(\psi, \phi, \bar{\psi}, \bar{\phi}) + \varepsilon h(\psi, \phi, \bar{\psi}, \bar{\phi}) + \varepsilon F(\psi, \phi, \bar{\psi}, \bar{\phi}), \tag{91}$$

where

$$H_0(\psi, \phi, \bar{\psi}, \bar{\phi}) = \langle \bar{\psi}, \psi \rangle + \langle \bar{\phi}, \phi \rangle, \tag{92}$$

$$\begin{aligned} h(\psi, \phi, \bar{\psi}, \bar{\phi}) &= \langle \bar{\psi}, (c \langle \nabla \rangle_c - c^2) \psi \rangle - \langle \bar{\phi}, (c \langle \nabla \rangle_c - c^2) \phi \rangle \\ &\sim \sum_{j \geq 1} \varepsilon^{j-1} \left(\langle \bar{\psi}, a_j \Delta^j \psi \rangle + \langle \bar{\phi}, a_j \Delta^j \phi \rangle \right) \\ &=: \sum_{j \geq 1} \varepsilon^{j-1} (h_j(\psi, \phi, \bar{\psi}, \bar{\phi})), \end{aligned} \tag{93}$$

$$\begin{aligned} F(\psi, \phi, \bar{\psi}, \bar{\phi}) &= \frac{\lambda}{16} \int_{\mathbb{T}} \left[\left\langle \psi + \bar{\psi}, \frac{c}{\langle \nabla \rangle_c} (\psi + \bar{\psi}) \right\rangle + \left\langle \phi + \bar{\phi}, \frac{c}{\langle \nabla \rangle_c} (\phi + \bar{\phi}) \right\rangle \right]^2 dx, \\ &\sim \frac{\lambda}{16} \int [|\psi + \bar{\psi}|^2 + |\phi + \bar{\phi}|^2]^2 dx \\ &\quad + \mathcal{O}(\varepsilon) \\ &=: \sum_{j \geq 1} \varepsilon^{j-1} F_j(\psi, \phi, \bar{\psi}, \bar{\phi}), \end{aligned} \tag{94}$$

where $(a_j)_{j \geq 1}$ are real coefficients, and $F_j(\psi, \phi, \bar{\psi}, \bar{\phi})$ is a polynomial function of the variables $\psi, \phi, \bar{\psi}, \bar{\phi}$ (along with their derivatives) and which admits a bounded vector field from a neighborhood of the origin in $W^{k+2(j-1), p}(\mathbb{R}^d, \mathbb{C}^2 \times \mathbb{C}^2)$ to $W^{k, p}(\mathbb{R}^d, \mathbb{C}^2 \times \mathbb{C}^2)$ for any $1 < p < +\infty$.

This description clearly fits the scheme treated in Sect. 4 with $n = 2$, and one can easily check that assumptions PER, NF and HVF are satisfied. Therefore, we can apply Theorem 7 to the Hamiltonian (91).

Remark 15 About the normal forms obtained by applying Theorem 7, we remark that in the first step (case $r = 1$ in the statement of the theorem) the homological equation we get is of the form

$$\{\chi_1, h_0\} + F_1 = \langle F_1 \rangle, \tag{95}$$

where $F_1(\psi, \bar{\psi}) = \frac{\lambda}{16} \int [|\psi + \bar{\psi}|^2 + |\phi + \bar{\phi}|^2]^2 dx$. Hence the transformed Hamiltonian is of the form

$$\begin{aligned} H_1(\psi, \phi, \bar{\psi}, \bar{\phi}) &= h_0(\psi, \phi, \bar{\psi}, \bar{\phi}) + \frac{1}{c^2} \left[-\frac{1}{2} ((\bar{\psi}, \Delta\psi) + (\bar{\phi}, \Delta\phi)) + \langle F_1 \rangle(\psi, \phi, \bar{\psi}, \bar{\phi}) \right] \\ &\quad + \frac{1}{c^4} \mathcal{R}^{(1)}(\psi, \phi, \bar{\psi}, \bar{\phi}), \end{aligned} \tag{96}$$

where

$$\begin{aligned} \langle F_1 \rangle &= \frac{\lambda}{16} [6\psi^2\bar{\psi}^2 + 6\phi^2\bar{\phi}^2 + 8\psi\bar{\psi}\phi\bar{\phi} + 2\psi^2\phi^2 + 2\bar{\psi}^2\bar{\phi}^2] \\ &= \frac{\lambda}{8} [3(|\psi|^2 + |\phi|^2)^2 + 2(\psi\phi - \bar{\psi}\bar{\phi})^2]. \end{aligned}$$

If we neglect the remainder and we derive the corresponding equations of motion for the system, we get

$$\begin{cases} -i\psi_t &= \psi + \frac{1}{c^2} \left\{ -\frac{1}{2}\Delta\psi + \frac{\lambda}{4} [3(|\psi|^2 + |\phi|^2)\psi + 2(\psi\phi + \bar{\psi}\bar{\phi})\bar{\phi}] \right\}, \\ i\phi_t &= \phi + \frac{1}{c^2} \left\{ -\frac{1}{2}\Delta\phi + \frac{\lambda}{4} [3(|\psi|^2 + |\phi|^2)\phi + 2(\psi\phi + \bar{\psi}\bar{\phi})\bar{\psi}] \right\}, \end{cases} \tag{97}$$

which is a system of two coupled NLS equations.

7 Dynamics

Now we want to exploit the result of the previous section in order to deduce some consequences about the dynamics of the NLKG equation (3) in the nonrelativistic limit. Consider the *simplified system*, that is, the Hamiltonian H_r in the notations of Theorem 7, where we neglect the remainder:

$$H_{simp} := h_0 + \varepsilon(h_1 + \langle F_1 \rangle) + \sum_{j=2}^r \varepsilon^j (h_j + Z_j).$$

We recall that in the case of the NLKG the simplified system is actually the NLS (given by $h_0 + \varepsilon(h_1 + \langle F_1 \rangle)$), plus higher-order normalized corrections. Now let ψ_r be a solution of

$$-i\dot{\psi}_r = X_{H_{simp}}(\psi_r), \tag{98}$$

then $\psi_a(t, x) := \mathcal{F}^{(r)}(\psi_r(c^2t, x))$ solves

$$\dot{\psi}_a = ic\langle \nabla \rangle_c \psi_a + \frac{\lambda}{2l} \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \frac{\psi_a + \bar{\psi}_a}{\sqrt{2}} \right]^{2l-1} - \frac{1}{c^{2r}} X_{\mathcal{F}^{(r)*}\mathcal{R}^{(r)}}(\psi_a, \bar{\psi}_a), \tag{99}$$

that is, the NLKG plus a remainder of order c^{-2r} (in the following we will refer to Eq. (99) as *approximate equation*, and to ψ_a as the *approximate solution* of the original NLKG). We

point out that the original NLKG and the approximate equation differ only by a remainder of order c^{-2r} , which is evaluated on the approximate solution. This fact is extremely important: indeed, if one can prove the smoothness of the approximate solution (which often is easier to check than the smoothness of the solution of the original equation), then the contribution of the remainder may be considered small in the nonrelativistic limit. This property is rather general and has been already applied in the framework of normal form theory (see, for example, [4]).

Now let ψ be a solution of the NLKG equation (3) with the initial datum ψ_0 , and let $\delta := \psi - \psi_a$ be the error between the solution of the approximate equation and the original one. One can check that δ fulfills

$$\dot{\delta} = i c \langle \nabla \rangle_c \delta + [P(\psi_a + \delta, \bar{\psi}_a + \bar{\delta}) - P(\psi_a, \bar{\psi}_a)] + \frac{1}{c^{2r}} X_{\mathcal{G}(r) * \mathcal{R}(r)}(\psi_a(t), \bar{\psi}_a(t)),$$

where

$$P(\psi, \bar{\psi}) = \frac{\lambda}{2l} \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \frac{\psi + \bar{\psi}}{\sqrt{2}} \right]^{2l-1}. \tag{100}$$

Thus we get

$$\begin{aligned} \dot{\delta} &= i c \langle \nabla \rangle_c \delta + dP(\psi_a(t))\delta + \mathcal{O}(\delta^2) + \mathcal{O}\left(\frac{1}{c^{2r}}\right); \\ \delta(t) &= e^{itc\langle \nabla \rangle_c} \delta_0 + \int_0^t e^{i(t-s)c\langle \nabla \rangle_c} dP(\psi_a(s))\delta(s) ds + \mathcal{O}(\delta^2) + \mathcal{O}\left(\frac{1}{c^{2r}}\right). \end{aligned} \tag{101}$$

By applying Gronwall inequality to (101) we obtain

Proposition 3 Fix $r \geq 1, R > 0, k_1 \gg 1, 1 < p < +\infty$. Then $\exists k_0 = k_0(r) > 0$ with the following properties: For any $k \geq k_1$ there exists $c_{l,r,k,p,R} \gg 1$ such that for any $c > c_{l,r,k,p,R}$, if we assume that

$$\|\psi_0\|_{k+k_0,p} \leq R$$

and that there exists $T = T_{r,k,p} > 0$ such that the solution of (98) satisfies

$$\|\psi_r(t)\|_{k+k_0,p} \leq 2R, \quad \text{for } 0 \leq t \leq T,$$

then

$$\|\delta(t)\|_{k,p} \leq C_{k,p} c^{-2r}, \quad \text{for } 0 \leq t \leq T. \tag{102}$$

Remark 16 If we restrict to $p = 2$, and to $M = \mathbb{T}^d$, the above result is actually a reformulation of Theorem 3.2 in [23]. We also remark that the time interval $[0, T]$ in which estimate (102) is valid is independent of c .

Remark 17 By exploiting estimate (32) about the canonical transformation, Proposition 3 leads immediately to a proof of Theorem 2.

In order to study the evolution of the error between the approximate solution and the solution of the NLKG over longer (namely, c -dependent) timescales, we observe that the error is described by

$$\dot{\delta}(t) = i c \langle \nabla \rangle_c \delta(t) + dP(\psi_a(t))\delta(t); \tag{103}$$

$$\delta(t) = e^{itc\langle \nabla \rangle_c} \delta_0 + \int_0^t e^{i(t-s)c\langle \nabla \rangle_c} dP(\psi_a(s))\delta(s) ds, \tag{104}$$

up to a remainder which is small, if we assume the smoothness of ψ_a .

Equation (103) in the context of dispersive PDEs is known as *semirelativistic spinless Salpeter equation* with a time-dependent potential. This system was introduced as a first order in time analogue of the KG equation for the Lorentz covariant description of bound states within the framework of relativistic quantum field theory, and, despite the nonlocality of its Hamiltonian, some of its properties have already been studied. (See [55] for a study from a physical point of view; for a more mathematical approach, see [33] and the more recent works [14,15], which are closer to the spirit of our approximation.)

It seems reasonable to estimate the solution of Eq.(103) by studying and by exploiting its dispersive properties, and this will be the aim of the following sections. From now on we will consider only the case $M = \mathbb{R}^d$ for $d \geq 2$.

8 Properties of the normal form equation

8.1 Linear case

Now let $r \geq 1, d \geq 2$. In [14,15] the authors proved that the linearized normal form system, namely the one that corresponds (up to a rescaling of time by a factor c^2) to

$$\begin{aligned}
 -i\dot{\psi}_r &= X_{h_0 + \sum_{j=1}^r \varepsilon^j h_j}(\psi_r), \\
 \psi_r(0) &= \psi_0,
 \end{aligned}
 \tag{105}$$

admits a unique solution in $L^\infty(\mathbb{R})H^{k+k_0}(\mathbb{R}^d)$ (this is a simple application of the properties of the Fourier transform), and by a perturbative argument they also proved the global existence also for the higher order Schrödinger equation with a bounded time-independent potential.

Moreover, by following the arguments of Theorem 4.1 in [31] and Lemma 4.3 in [14] one obtains the following dispersive estimates and local-in-time Strichartz estimates for solutions of the linearized normal form equation (105).

Proposition 4 (Fig. 1) *Let $r \geq 1$ and $d \geq 2$, and denote by $\mathcal{U}_r(t)$ the evolution operator of (105) at the time c^2t ($c \geq 1, t > 0$). Then one has the following local-in-time dispersive estimate*

$$\|\mathcal{U}_r(t)\|_{L^1(\mathbb{R}^d) \rightarrow L^\infty(\mathbb{R}^d)} \lesssim c^{d(1-\frac{1}{r})} |t|^{-d/(2r)}, \quad 0 < |t| \lesssim c^{2(r-1)}.
 \tag{106}$$

On the other hand, $\mathcal{U}_r(t)$ is unitary on $L^2(\mathbb{R}^d)$.

Now introduce the following set of admissible exponent pairs:

$$\Delta_r := \{(p, q) : (1/p, 1/q) \text{ lies in the closed quadrilateral } ABCD\},
 \tag{107}$$

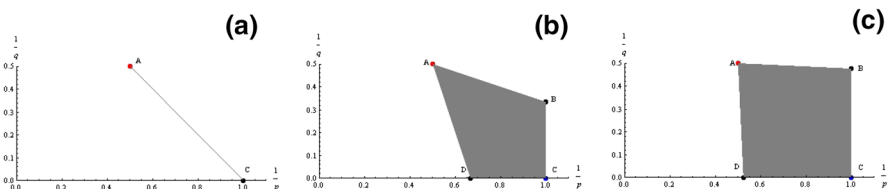


Fig. 1 Set of admissible exponents Δ_r for different values of r : **a** $r = 1$ (this is the Schrödinger case); **b** $r = 2$; **c** $r = 11$

where

$$A = \left(\frac{1}{2}, \frac{1}{2}\right), \quad B = \left(1, \frac{1}{\tau_r}\right), \quad C = (1, 0), \quad D = \left(\frac{1}{\tau_r'}, 0\right),$$

$$\tau_r = \frac{2r-1}{r-1}, \quad \frac{1}{\tau_r} + \frac{1}{\tau_r'} = 1.$$

Then for any $(p, q) \in \Delta_r \setminus \{(2, 2), (1, \tau_r), (\tau_r', \infty)\}$

$$\|\mathcal{U}_r(t)\|_{L^p(\mathbb{R}^d) \rightarrow L^q(\mathbb{R}^d)} \lesssim c^{d\left(1-\frac{1}{r}\right)\left(\frac{1}{p}-\frac{1}{q}\right)} |t|^{-\frac{d}{2r}\left(\frac{1}{q}-\frac{1}{p}\right)}, \quad 0 < |t| \lesssim c^{2(r-1)}, \quad (108)$$

Let $r \geq 1$ and $d \geq 2$: In the following lemma (p, q) is called an order- r admissible pair when $2 \leq p, q \leq +\infty$ for $r \geq 2$ ($2 \leq q \leq 2d/(d-2)$ for $r = 1$), and

$$\frac{2}{p} + \frac{d}{rq} = \frac{d}{2r}. \quad (109)$$

Proposition 5 *Let $r \geq 1$ and $d \geq 2$, and denote by $\mathcal{U}_r(t)$ the evolution operator of (105) at the time c^2t ($c \geq 1, t > 0$). Let (p, q) and (a, b) be order- r admissible pairs, then for any $T \lesssim c^{2(r-1)}$*

$$\|\mathcal{U}_r(t)\phi_0\|_{L^p([0,T]L^q(\mathbb{R}^d))} \lesssim c^{d\left(1-\frac{1}{r}\right)\left(\frac{1}{2}-\frac{1}{q}\right)} \|\phi_0\|_{L^2(\mathbb{R}^d)} = c^{\left(1-\frac{1}{r}\right)\frac{2r}{p}} \|\phi_0\|_{L^2(\mathbb{R}^d)}, \quad (110)$$

$$\left\| \int_0^t \mathcal{U}_r(t-\tau)\phi(\tau)d\tau \right\|_{L^p([0,T]L^q(\mathbb{R}^d))} \lesssim c^{\left(1-\frac{1}{r}\right)2r\left(\frac{1}{p}+\frac{1}{a}\right)} \|\phi\|_{L^{a'}([0,T]L^{b'}(\mathbb{R}^d))}. \quad (111)$$

8.2 Well-posedness of higher-order nonlinear Schrödinger equations with small data

Here we discuss the local well-posedness of

$$-i\psi_t = A_{c,r}\psi + P((\partial_x^\alpha \psi)_{|\alpha| \leq 2(r-1)}, (\partial_x^\alpha \bar{\psi})_{|\alpha| \leq 2(r-1)}), \quad t \in I, \quad x \in \mathbb{R}^d, \quad (112)$$

$$\psi(0, x) = \psi_0(x), \quad (113)$$

where $r \geq 2, I := [0, T], T > 0$,

$$A_{c,r} = c^2 - \sum_{j=1}^r \frac{\Delta^j}{c^{2(j-1)}}, \quad c \geq 1,$$

and P is an analytic function at the origin of the form

$$P(z) = \sum_{m+1 \leq |\beta| < M} a_\beta z^\beta, \quad |a_\beta| \leq K^{|\beta|}, \quad |z| \ll 1, \quad (114)$$

where $M > m \geq 2, m, M \in \mathbb{N}$.

We will exploit this result during the proof of Theorem 4. We will adapt an argument of [50] in order to show the local well-posedness of equation for data with small norm in the so-called modulation spaces.

Modulation spaces $M_{p,q}^s$ ($s \in \mathbb{R}, 0 < p, q < +\infty$) were introduced by Feichtinger, and they can be seen as a variant of Besov spaces, in the sense that they allow to perform a frequency decomposition of operators, and to study their properties with respect to lower and higher frequencies. This spaces were recently used in order to prove global well-posedness and scattering for small data for nonlinear dispersive PDEs, especially in the case of derivative

nonlinearities (see, for example, [50,58,59]). We refer to [49] for a survey about modulation spaces and nonlinear evolution equations.

We define the norm on modulation spaces via the following decomposition: Let $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}$ be a function such that

$$\text{supp}(\sigma) \subset [-3/4, 3/4]^d,$$

and consider a function sequence $(\sigma_k)_{k \in \mathbb{Z}^d}$ satisfying

$$\sigma_k(\cdot) = \sigma(\cdot - k), \tag{115}$$

$$\sum_{k \in \mathbb{Z}^d} \sigma_k(\xi) = 1, \quad \forall \xi \in \mathbb{R}^d. \tag{116}$$

Denote by

$$\mathcal{Y}_d := \{(\sigma_k)_{k \in \mathbb{Z}^d} : (\sigma_k)_{k \in \mathbb{Z}^d} \text{ satisfies (115)–(116)}\}.$$

Let $(\sigma_k)_{k \in \mathbb{Z}^d} \in \mathcal{Y}_d$, and define the frequency-uniform decomposition operators

$$\square_k := \mathcal{F}^{-1} \sigma_k \mathcal{F}, \tag{117}$$

where by \mathcal{F} we denote the Fourier transform on \mathbb{R}^d , then we define the modulation spaces $M_{p,q}^s(\mathbb{R}^d)$ via the following norm,

$$\|f\|_{M_{p,q}^s(\mathbb{R}^d)} := \left(\sum_{k \in \mathbb{Z}^d} \langle k \rangle^{sq} \|\square_k f\|_p^q \right)^{1/q}, \quad s \in \mathbb{R}, 0 < p, q < +\infty. \tag{118}$$

Actually, in our application we will always be interested in the spaces $M_{p,1}^s(\mathbb{R}^d)$ with $s \in \mathbb{R}$ and $p > 1$. We just mention some properties of modulation spaces.

Proposition 6 *Let $s, s_1, s_2 \in \mathbb{R}$ and $1 < p, p_1, p_2 < +\infty$.*

1. $M_{p,1}^s(\mathbb{R}^d)$ is a Banach space;
2. $\mathcal{S}(\mathbb{R}^d) \subset M_{p,1}^s(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$;
3. $\mathcal{S}(\mathbb{R}^d)$ is dense in $M_{p,1}^s(\mathbb{R}^d)$;
4. if $s_2 \leq s_1$ and $p_1 \leq p_2$, then $M_{p_1,1}^{s_1} \subseteq M_{p_2,1}^{s_2}$;
5. $M_{p,1}^0(\mathbb{R}^d) \subseteq L^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$;
6. let $\tau(p) = \max(0, d(1 - 1/p), d/p)$ and $s_1 > s_2 + \tau(p)$, then $W^{s_1,p}(\mathbb{R}^d) \subset M_{p,1}^{s_2}$;
7. let $s_1 \geq s_2$, then $M_{p,1}^{s_1}(\mathbb{R}^d) \subset W^{s_2,p}(\mathbb{R}^d)$.

The last two properties are not trivial and have been proved in [32].

We also introduce other spaces which are often used in this context: the anisotropic Lebesgue space $L_{x_i; (x_j)_{j \neq i}, t}^{p_1, p_2}$,

$$\|f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{p_1, p_2}} := \left\| \|f\|_{L_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d, t}^{p_2}(\mathbb{R}^{d-1} \times I)} \right\|_{L_{x_i}^{p_1}(\mathbb{R})},$$

and, for any Banach space X , the spaces $l_{\square}^{1,s}(X)$ and $l_{\square,i}^{1,s}(X)$,

$$\|f\|_{l_{\square}^{1,s}(X)} := \sum_{k \in \mathbb{Z}^d} \langle k \rangle^s \|\square_k f\|_X, \tag{119}$$

$$\|f\|_{l_{\square,i,c}^{1,s}(X)} := \sum_{k \in \mathbb{Z}^d} \langle k \rangle^s \|\square_k f\|_X, \quad \mathbb{Z}_i^d := \left\{ k \in \mathbb{Z}^d : |k_i| = \max_{1 \leq j \leq d} |k_j|, |k_i| > c \right\}. \tag{120}$$

For simplicity, we write $l_{\square}^1(X) = l_{\square}^{1,0}(X)$ and $M_{p,1}^s = M_{p,1}^s(\mathbb{R}^d)$.

Proposition 7 *Let $d \geq 2, m \geq 2, m > 4r/d$ and $s > 2(r - 1) + 1/m$.*

- (i) *There exist $c_0 > 1$ and $\delta_0 = \delta_0(d, m, r) > 0$ such that for any $c \geq c_0$, for any $\delta > \delta_0$ and for any $\psi_0 \in M_{2,1}^s$ with $\|\psi_0\|_{M_{2,1}^s} \leq c^{-\delta}$ Eq. (112) admits a unique solution $\psi \in C(I, M_{2,1}^s) \cap D$, where $T = T(\|\psi_0\|_{M_{2,1}^s}) = \mathcal{O}(c^{2(r-1)})$, and*

$$\|\psi\|_D = \sum_{\alpha=0}^{2(r-1)} \sum_{i,l=1}^d \|\partial_{x_l}^{\alpha} \psi\|_{l_{\square,i,c}^{1,s-r+1/2}(L_{x_i;(x_j)_{j \neq i,l}}^{\infty,2}) \cap l_{\square}^{1,s}(L_{x_i;(x_j)_{j \neq i,l}}^{m,\infty}) \cap l_{\square}^{1,s+1/m}(L_x^{\infty} L_x^2 \cap L_{t,x}^{2+m})} \lesssim c^{-\delta}. \tag{121}$$

- (ii) *Moreover, if $s \geq s_0(d) := d + 2 + \frac{1}{2}$, then there exists $\delta_1 = \delta_1(d, m, r) > 0$ such that for any $c \geq c_0$, for any $\delta > \delta_1$ and for any $\psi_0 \in M_{2,1}^s$ with $\|\psi_0\|_{M_{2,1}^s} \leq c^{-\delta}$ Eq. (112) admits a unique solution $\psi \in C(I, H^s)$, where $T = T(\|\psi_0\|_{M_{2,1}^s}) = \mathcal{O}(c^{2(r-1)})$, and*

$$\|\psi(t)\|_{H^s} \lesssim c^{-\delta}, \quad |t| \lesssim c^{2(r-1)}. \tag{122}$$

From the above proposition and from the embedding $H^{s+\sigma+d/2} \subset M_{2,1}^s$ for any $\sigma > 0$ we can deduce

Corollary 1 *Let $d \geq 2, l \geq 2, r < \frac{d}{2}(l - 1)$ and $s > 2(r - 1) + \frac{1}{2(l-1)}$. Then there exist $c_0 > 1, \delta_0 = \delta_0(d, l, r) > 0$ and $\delta_1 = \delta_1(d, l, r) > 0$ such that for any $c \geq c_0$, for any $\delta > \max(\delta_0, \delta_1)$, for any $\sigma > 0$ and for any $\psi_0 \in H^{s+\sigma+d/2}$ with $\|\psi_0\|_{H^{s+\sigma+d/2}} \leq c^{-\delta}$ the normal form equation for (56) admits a unique solution $\psi \in C([0, T], H^{s+\sigma+d/2}) \cap D$, where $T = T(\|\psi_0\|_{H^{s+\sigma+d/2}}) = \mathcal{O}(c^{2(r-1)})$, and (121) holds. Furthermore, we have that $\psi \in L^{\infty}(I)H^{s+\sigma+d/2}(\mathbb{R}^d)$, and*

$$\|\psi(t)\|_{H^{s+\sigma+d/2}} \lesssim c^{-\delta}, \quad |t| \lesssim c^{2(r-1)}. \tag{123}$$

Since the nonlinearity in Eq. (112) involves derivatives, this could cause a loss of derivatives as long as we rely only on energy estimates, on dispersive estimates or on Strichartz estimates. In order to overcome such a problem, we study the time decay of the operator $\mathcal{U}_r(t) := e^{itA_{c,r}}$, its local smoothing property, Strichartz estimates with \square_k -decomposition and maximal function estimates in the framework of frequency-uniform localization.

The rest of this subsection is devoted to the proof of Proposition 7. For convenience, we will always use the following function sequence $(\sigma_k)_{k \in \mathbb{Z}^d}$ to define modulation spaces.

Lemma 5 *Let $(\eta_k)_{k \in \mathbb{Z}} \in \mathcal{Y}_1$, and assume that $\text{supp}(\eta_k) \subset [k - 2/3, k + 2/3]$. Consider*

$$\sigma_k(\xi) := \eta_{k_1}(\xi_1) \dots \eta_{k_d}(\xi_d), \quad k = (k_1, \dots, k_d) \in \mathbb{Z}^d, \tag{124}$$

then $(\sigma_k)_{k \in \mathbb{Z}^d} \in \mathcal{Y}_d$.

For convenience, we also write

$$\tilde{\sigma}\sigma_k = \sum_{\|l\|_\infty \leq 1} \sigma_{k+l}, \quad \tilde{\sigma}\square_k = \sum_{\|l\|_\infty \leq 1} \square_{k+l}, \quad k \in \mathbb{Z}^d, \tag{125}$$

and one can check that

$$\tilde{\sigma}\sigma_k\sigma_k = \sigma_k, \quad \tilde{\sigma}\square_k \circ \square_k = \square_k, \quad k \in \mathbb{Z}^d. \tag{126}$$

We also write $\mathcal{A}_r f(t, x) := \int_0^t \mathcal{U}_r(t - \tau) f(\tau, x) d\tau$.

8.2.1 Time decay

Now, the time decay of the operator $\mathcal{U}_r(t)$ is known (see (106)), but now we are interested in its frequency-localized version, and we want to consider lower, medium and higher frequency separately. For simplicity we discuss the case $r = 2$, and we defer to the end of this section a remark about the case $r > 2$. So, consider

$$\mathcal{U}_2(t) = e^{itA_{c,2}} = e^{ic^2t} \mathcal{F}^{-1} e^{it\left(|\xi|^2 - \frac{|\xi|^4}{c^2}\right)} \mathcal{F},$$

and write $\varepsilon = c^{-2}$. It is known that the time decay of $\mathcal{U}_2(t)$ is determined by the critical points of $P_2(|\xi|) = |\xi|^2 - \varepsilon|\xi|^4$. Notice that $P'_2(R) = 4R(\varepsilon^{1/2}R + \frac{1}{\sqrt{2}})(\varepsilon^{1/2}R - \frac{1}{\sqrt{2}})$, the singular points of P_2 are $\xi = 0$ and the points of the sphere $\xi = (2\varepsilon)^{-1/2}$. To handle these points, we exploit Littlewood–Paley decomposition, van der Corput lemma and some properties of the Fourier transform of radial functions.

Indeed, it is known that the Fourier transform of a radial function f is radial,

$$\mathcal{F}f(\xi) = 2\pi \int_0^\infty f(R)R^{d-1}(R|\xi|)^{-(d-2)/2} J_{\frac{d-2}{2}}(R|\xi|)dR,$$

where J_m is the order m Bessel function,

$$J_m(R) = \frac{(R/2)^m}{\Gamma(m + 1/2)\pi^{1/2}} \int_{-1}^1 e^{iRt}(1 - t^2)^{m-1/2}dt, \quad m > -1/2.$$

By following the computations in [50] we obtain that

$$\begin{aligned} \mathcal{F}f(s) &= K_d\pi \int_0^\infty f(R)R^{d-1}e^{-iRs}\bar{h}(Rs)dR \\ &\quad + K_d\pi \int_0^\infty f(R)R^{d-1}e^{iRs}h(Rs)dR, \quad K_d > 0, \end{aligned} \tag{127}$$

$$|h^{(k)}(R)| \leq K_d(1 + R)^{-\frac{d-1}{2}-k}, \quad \forall k \geq 0. \tag{128}$$

Now we make a Littlewood–Paley decomposition of the frequencies: Choose ρ a smooth cutoff function equal to 1 in the unit ball and equal to 0 outside the ball of radius 2, write $\phi_0 = \rho(\cdot) - \rho(2\cdot)$, $\phi_j(\cdot) = \mathcal{F}^{-1}\phi_0(2^{-j}\cdot)\mathcal{F}$, $j \in \mathbb{Z}$, and consider

$$\begin{aligned} \mathcal{U}_2(t)\psi_0 &= \sum_{|j| \leq K} \phi_j(D)\mathcal{U}_2(t)\psi_0 + \sum_{j < -K} \phi_j(D)\mathcal{U}_2(t)\psi_0 + \sum_{j > K} \phi_j(D)\mathcal{U}_2(t)\psi_0 \\ &=: P_=< \mathcal{U}_2(t)\psi_0 + P_< \mathcal{U}_2(t)\psi_0 + P_> \mathcal{U}_2(t)\psi_0, \end{aligned} \tag{129}$$

where

$$K := K(\varepsilon) = 10 - \frac{1}{2} \lceil \log_2 \varepsilon \rceil. \tag{130}$$

Notice that the singular point $R = 0$ is in the support set of $\mathcal{F}(P_{\neq} \mathcal{U}_2(t)\psi_0)$. Roughly speaking, if $j < -K$, the dominant term in $P_2(R)$ is R^2 , while if $j > K$ the dominant term in $P_2(R)$ is εR^4 ; hence, by (106)

$$\|P_{<} \mathcal{U}_2(t)\psi_0\|_{L^\infty} \lesssim |t|^{-d/2} \|\psi_0\|_{L^1}, \tag{131}$$

$$\|P_{>} \mathcal{U}_2(t)\psi_0\|_{L^\infty} \lesssim c^{d/2} |t|^{-d/4} \|\psi_0\|_{L^1}, \quad 0 < |t| \lesssim c^2. \tag{132}$$

The time decay estimate for $P_{\neq} \mathcal{U}_2(t)\psi_0$ is more difficult, since $P_2(R)$ has a singular point in $R = R_1 := (2\varepsilon)^{-1/2}$, which corresponds to the sphere $|\xi| = R_1$ in the support set of $\mathcal{F}(P_{\neq} \mathcal{U}_2(t)\psi_0)$. We notice that also the point that satisfies $P_2''(R) = 0$, $R = (6\varepsilon)^{-1/2}$, corresponds to a sphere $\xi = R_2$ contained in the support set of $\mathcal{F}(P_{\neq} \mathcal{U}_2(t)\psi_0)$; we shall use this fact later.

In order to handle the singular point R_1 , we perform another decomposition around the sphere $|\xi| = R_1$. Denote $\tilde{\sigma}\rho(\cdot) = \rho(2^{-K}\cdot) - \rho(2^{(K+1)}\cdot)$, then $P_{\neq} = \mathcal{F}^{-1} \tilde{\sigma}\rho \mathcal{F}$; write $P_k = \mathcal{F}^{-1} \phi_k(|\xi| - R_1) \mathcal{F}$, we get

$$\sum_{|j| \leq K} \phi_j(D) \mathcal{U}_2(t)\psi_0 = \sum_{k \in \mathbb{Z}} P_{\neq} P_k \mathcal{U}_2(t)\psi_0 \tag{133}$$

By Young’s inequality

$$\|P_{\neq} P_k \mathcal{U}_2(t)\psi_0\|_{L^\infty} \lesssim \|\mathcal{F}^{-1}(\tilde{\sigma}\rho\phi_k(|\xi| - R_1)e^{-itP_2(|\xi|)})\|_{L^\infty} \|\psi_0\|_{L^1}. \tag{134}$$

Moreover,

$$\begin{aligned} & \mathcal{F}^{-1}(\tilde{\sigma}\rho\phi_k(|\xi| - R_1)e^{-itP_2(|\xi|)}) \\ & \stackrel{(127)}{=} K_d \pi \int_0^\infty R^{d-1} \tilde{\sigma}\rho(R)\phi_k(R - R_1)e^{-itP_2(R)-iR|x|} \bar{h}(R|x) dR \\ & \quad + K_d \pi \int_0^\infty R^{d-1} \tilde{\sigma}\rho(R)\phi_k(R - R_1)e^{-itP_2(R)+iR|x|} h(R|x) dR \\ & =: A_k(|x|) + B_k(|x|). \end{aligned}$$

In order to estimate $A_k(s)$ we rewrite it as

$$A_k(s) = K_d \pi \left(\int_{R_1}^\infty + \int_0^{R_1} \right) R^{d-1} \tilde{\sigma}\rho(R)\phi_k(R - R_1)e^{-itP_2(R)-iRs} \bar{h}(Rs) dR \tag{135}$$

$$=: A_k^{(1)}(s) + A_k^{(2)}(s). \tag{136}$$

We begin by estimating $A_k^{(1)}$: Notice that $A_k^{(1)}(s)$ for $k > K + 2$; hence, we can assume that $k \leq K + 2$. By a change of variables we obtain

$$\begin{aligned} A_k^{(1)}(s) & \stackrel{R=R_1+2^k\sigma}{=} 2^k K_d \pi e^{-iR_1s} \int_{1/2}^2 F(\sigma) e^{it2^k\tilde{\sigma}P_2(\sigma)} d\sigma, \\ F(\sigma) & := (R_1 + 2^k\sigma)^{d-1} \tilde{\sigma}\rho(R_1 + 2^k\sigma)\phi_0(\sigma)\bar{h}((R_1 + 2^k\sigma)s), \\ \tilde{\sigma}P_2(\sigma) & := (2^{2k}t)^{-1}(tP_2(R_1 + 2^k\sigma) - 2^k\sigma s). \end{aligned}$$

One can check that

$$|\tilde{\sigma} P_2'(\sigma)| = \left| 4(R_1 + 2^k \sigma)(2R_1 + 2^k \sigma)\sigma\varepsilon - \frac{s}{t2^k} \right|.$$

Let $s \gg 1$; if $s \ll 2^k t/\varepsilon$, then

$$\begin{aligned} |F^{(m)}(\sigma)| &\lesssim 1, \quad \forall m \geq 1, \quad |\tilde{\sigma} P_2'(\sigma)| \lesssim \varepsilon, \quad |\tilde{\sigma} P_2''(\sigma)| \\ &\lesssim \varepsilon^{1/2}, \quad |\tilde{\sigma} P_2'''(\sigma)| \lesssim \varepsilon, \quad |\tilde{\sigma} P_2^{(m)}(\sigma)| \stackrel{\varepsilon \leq 1}{\lesssim} 1, \quad \forall m \geq 4 \end{aligned}$$

while for $s \gg 2^k t/\varepsilon$

$$|F^{(m)}(\sigma)| \lesssim 1, \quad \forall m \geq 1, \quad |\tilde{\sigma} P_2^{(m)}(\sigma)| \stackrel{\varepsilon \leq 1}{\lesssim} 1, \quad \forall m \geq 1.$$

Integrating by parts we get

$$\begin{aligned} A_k^{(1)}(s) &= 2^k (2^{2k} t)^{-N} K_d \pi e^{iR_1 s} \int_{1/2}^2 e^{it2^{2k}\tilde{\sigma} P_2(\sigma)} \frac{d}{d\sigma} \left(\frac{1}{\tilde{\sigma} P_2'(\sigma)} \cdots \frac{d}{d\sigma} \left(\frac{1}{\tilde{\sigma} P_2'(\sigma)} \frac{d}{d\sigma} \left(\frac{F(\sigma)}{\tilde{\sigma} P_2'(\sigma)} \right) \right) \right) d\sigma. \end{aligned}$$

Therefore,

$$|A_k^{(1)}(s)| \lesssim 2^k (2^{2k} t)^{-N}. \tag{137}$$

If $s \sim 2^k t/\varepsilon$, we apply van der Corput lemma,

$$\begin{aligned} |A_k^{(1)}(s)| &\lesssim 2^k (2^{2k} t)^{-1/2} \int_{1/2}^2 |\partial_\sigma F(\sigma)| d\sigma \\ &\stackrel{(128)}{\lesssim} 2^k (2^{2k} t)^{-1/2} s^{-(d-1)/2} \lesssim 2^k (2^{2k} t)^{-d/2} \varepsilon^{(d-1)/2}. \end{aligned}$$

Moreover, we can check that $|A_k^{(1)}(s)| \lesssim 2^k$; hence, for $s \gg 1$

$$|A_k^{(1)}(s)| \stackrel{\varepsilon \leq 1}{\lesssim} 2^k \min(1, (2^{2k} t)^{-d/2}). \tag{138}$$

If $s \lesssim 1$, we rewrite $A_k^{(1)}$ in the following form

$$\begin{aligned} A_k^{(1)}(s) &= 2^k K_d \pi e^{-iR_1 s} \int_{1/2}^2 F_1(\sigma) e^{itP_2(R_1+2^k\sigma)} d\sigma, \\ F_1(\sigma) &:= (R_1 + 2^k \sigma)^{d-1} \tilde{\sigma} \rho(R_1 + 2^k \sigma) \phi_0(\sigma) \bar{h}((R_1 + 2^k \sigma)s) e^{-i2^k \sigma s}. \end{aligned}$$

Again integrating by parts, we obtain

$$|A_k^{(1)}(s)| \lesssim 2^k \min(1, (2^{2k} t)^{-d/2}). \tag{139}$$

Now we estimate $A_k^{(2)}$. We notice that $R_2 \in \text{supp}(\phi_k(R_1 - \cdot))$ if and only if $k \in \{-2, -1\}$; when $k \notin \{-2, -1\}$, one can repeat the above argument and show that

$$|A_k^{(2)}(s)| \lesssim 2^k \min(1, (2^{2k} t)^{-d/2}). \tag{140}$$

Let $k \in \{-2, -1\}$. If $s \ll t$ or $s \gg t$ we have by integration by parts that

$$|A_k^{(2)}(s)| \lesssim \min(1, t^{-N}), \quad \forall N \in \mathbb{N}.$$

On the other hand, if $s \sim t$ we can use van der Corput lemma and obtain

$$|A_k^{(2)}(s)| \lesssim t^{-1/3} s^{-(d-1)/2} \lesssim t^{-\frac{d}{2} + \frac{1}{6}}.$$

Therefore, for $k \in \{-2, -1\}$ we have

$$|A_k^{(2)}(s)| \lesssim \min\left(1, t^{-\frac{d}{2} + \frac{1}{6}}\right). \tag{141}$$

Combining (140) and (141) we can deduce that

$$|A_k^{(2)}(s)| \lesssim 2^k \min\left(1, (2^{2k}t)^{-\frac{d}{2} + \frac{1}{6}}\right). \tag{142}$$

If we sum up all the A_k for $k \leq K + 2$ we finally conclude that for any $d \geq 2$

$$\|P = \mathcal{U}_2(t)\psi_0\|_{L^\infty} \lesssim c \min(|t|^{-d/2}, |t|^{-d/2+1/6}) \|\psi_0\|_{L^1}. \tag{143}$$

Remark 18 In the general case $r > 2$, we have to determine critical points for the polynomial

$$P_r(R) = \sum_{j=1}^r (-1)^{j+1} \varepsilon^{j-1} R^{2j}, \tag{144}$$

namely the roots of the polynomial

$$P'_r(R) = \sum_{j=1}^r (-1)^{j+1} \varepsilon^{j-1} 2j R^{2j-1} = R \left(\sum_{j=1}^r (-1)^{j+1} \varepsilon^{j-1} 2j R^{2(j-1)} \right). \tag{145}$$

Besides the trivial value $R = 0$, which we deal as in the case $r = 2$, one should rely on lower and upper bounds to determine the other (if any) real roots. For a lower bound, we rely on a well-known corollary of Rouché theorem from complex analysis, and we obtain that the other roots satisfy

$$\begin{aligned} R &\geq \frac{2}{\max\left(2, \sum_{j=1}^r 2j\varepsilon^{j-1}\right)} \\ &\geq \frac{2}{\max\left(2, 2r \sum_{j=0}^{r-1} \varepsilon^j\right)} \\ &\stackrel{\varepsilon \leq 1/2}{\geq} \frac{2}{\max(2, 4r\varepsilon)} \stackrel{\varepsilon \ll 1/(2r)}{\geq} 1. \end{aligned}$$

For what concerns an upper bound, we exploit an old result by Fujiwara [24], and we get that the roots satisfy

$$\begin{aligned} R &\leq \max_{1 \leq j \leq r-1} \left(2(r-1) \frac{2j\varepsilon^{j-1}}{2r\varepsilon^{r-1}} \right)^{\frac{1}{2(j-1)}} \\ &\leq 2(r-1) \max_{1 \leq j \leq r-1} \left(\frac{j}{r} \right)^{\frac{1}{2(j-1)}} \varepsilon^{\frac{j-r}{2(j-1)}} \\ &\stackrel{\varepsilon \leq 1}{\leq} K_r \varepsilon^{-1/2} \end{aligned}$$

for some $K_r > 0$.

Hence, in the case $r > 2$, if ε sufficiently small (depending on r), then the polynomial P'_r has critical points (apart from 0) which have modulus between 1 and $\mathcal{O}(\varepsilon^{-1/2})$ (a similar

argument works also for the polynomial P_r''), and this affects the medium-frequency decay of $\mathcal{U}_r(t)$. In any case, we can deal with this problem as in the case $r = 2$, and we get

$$\|P_{<} \mathcal{U}_r(t)\psi_0\|_{L^\infty} \lesssim |t|^{-d/2} \|\psi_0\|_{L^1}, \tag{146}$$

$$\|P_{=} \mathcal{U}_r(t)\psi_0\|_{L^\infty} \lesssim c \min(|t|^{-d/2}, |t|^{-d/2+1/6}) \|\psi_0\|_{L^1}, \tag{147}$$

$$\|P_{>} \mathcal{U}_r(t)\psi_0\|_{L^\infty} \lesssim c^{d/2} |t|^{-\frac{d}{2r}} \|\psi_0\|_{L^1}, \quad 0 < |t| \lesssim c^{2(r-1)}. \tag{148}$$

8.2.2 Smoothing estimates

As already pointed out, one needs smoothing estimates to ensure the well-posedness of Eq. (112) because of the presence of derivatives in the nonlinearity. Again, we first consider the case $r = 2$, and then we mention the results for $r > 2$.

Proposition 8 For any $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ with $|k_i| = |k|_\infty$ and $|k_i| \gtrsim c$

$$\left\| \square_k D_{x_i}^{3/2} \mathcal{A}_2(t)\psi_0 \right\|_{L_{x_i:(x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim c \|\square_k \psi_0\|_{L^2}. \tag{149}$$

Proof It suffices to consider the case $i = 1$. For convenience, we write $\bar{z} = (z_1, \dots, z_d)$. Then,

$$\begin{aligned} \left\| \square_k D_{x_1}^{3/2} \mathcal{A}_2(t)\psi_0 \right\|_{L_{x_1:(x_j)_{j \neq 1}, t}^{\infty, 2}} &= \left\| \int \sigma_k(\xi) |\xi_1|^{3/2} e^{itP_2(|\xi|)} \mathcal{F}(\psi_0)(\xi) e^{ix_1 \xi_1} d\xi \right\|_{L_{x_1}^\infty L_{\xi, t}^2} \\ &\lesssim \left\| \int \eta_{k_1}(\xi_1) |\xi_1|^{3/2} e^{itP_2(|\xi|)} \mathcal{F}(\psi_0)(\xi) e^{ix_1 \xi_1} d\xi \right\|_{L_{x_1}^\infty L_{\xi, t}^2} =: L. \end{aligned}$$

Now, we estimate L : If $k_1 \gtrsim c$, then $\xi_1 > 0$ for $\xi \in \text{supp}(\eta_{k_1})$. Hence, by changing variable, $\theta = P_2(|\xi|)$, we get

$$\begin{aligned} L &\lesssim \left\| \int \eta_{k_1}(\xi_1(\theta)) \xi_1(\theta)^{3/2} e^{i\theta} \mathcal{F}(\psi_0)(\xi(\theta)) e^{ix_1 \xi_1(\theta)} \frac{1}{2} \xi_1^{-1}(\theta) \left(2 \frac{|\xi|^2}{c^2} - 1\right)^{-1} \right\|_{L_{x_1}^\infty L_{\xi, t}^2} \\ &\lesssim \left\| \eta_{k_1}(\xi_1(\theta)) \xi_1(\theta)^{1/2} \mathcal{F}(\psi_0)(\xi(\theta)) \left(2 \frac{|\xi|^2}{c^2} - 1\right)^{-1} \right\|_{L_\theta^2 L_{\xi}^2} \\ &\lesssim \left\| \eta_{k_1}(\xi_1) \xi_1^{1/2} \mathcal{F}(\psi_0)(\xi) \left(2 \frac{|\xi|^2}{c^2} - 1\right)^{-1} \left(2 \frac{|\xi|^2}{c^2} - 1\right)^{1/2} \xi_1^{1/2} \right\|_{L_{\xi}^2} \\ &= \left\| \eta_{k_1}(\xi_1) \xi_1 \mathcal{F}(\psi_0)(\xi) \left(2 \frac{|\xi|^2}{c^2} - 1\right)^{-1/2} \right\|_{L_{\xi}^2} \lesssim c \|\psi_0\|_{L^2}. \end{aligned}$$

The proof for the case $k_1 \lesssim -c$ is similar. □

By duality we have the following

Proposition 9 For any $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ with $|k_i| = |k|_\infty$ and $|k_i| \gtrsim c$

$$\left\| \square_k \partial_{x_i}^2 \mathcal{A}_2 f \right\|_{L_t^\infty L_x^2} \lesssim c \|\square_k D_i^{1/2} f\|_{L_{x_i:(x_j)_{j \neq i}, t}^{1, 2}}. \tag{150}$$

Now consider the inhomogeneous Cauchy problem

$$-i\psi_t = A_{c,2}\psi + f(t, x), \quad \psi(0, x) = 0. \tag{151}$$

Proposition 10 For any $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ with $|k_i| = |k|_\infty$ and $|k_i| \gtrsim c$

$$\|\square_k \partial_{x_i}^2 \psi\|_{L_{x_i:(x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim \|\square_k f\|_{L_{x_i:(x_j)_{j \neq i}, t}^{1, 2}}. \tag{152}$$

Proof It suffices to consider $i = 1$. We write

$$\psi = \mathcal{F}_{\tau, \xi}^{-1} \frac{1}{\tau - c^2 - P_2(|\xi|)} (\mathcal{F}_{t,x} f)(\tau, \xi).$$

We have

$$\partial_{x_i}^2 \psi = \mathcal{F}_{\tau, \xi}^{-1} \frac{\xi_1^2}{P_2(|\xi|) + c^2 - \tau} \mathcal{F}_{t,x} f. \tag{153}$$

We want to show that

$$\left\| \mathcal{F}_{\tau, \xi}^{-1} \frac{\eta_{k_1}(\xi_1) \xi_1^2}{P_2(|\xi|) + c^2 - \tau} \mathcal{F}_{t,x} f \right\|_{L_{x_1}^\infty L_{\xi, t}^2} \lesssim \left\| \mathcal{F}_{\xi_1}^{-1} \eta_{k_1}(\xi_1) \mathcal{F}_{x_1} f \right\|_{L_{x_1}^1 L_{\xi, t}^2},$$

which, by Young’s inequality, is equivalent to show that

$$\sup_{x_1, \tau, \xi_j (j \neq 1)} \left| \mathcal{F}_{\xi_1}^{-1} \frac{\sigma_k(\xi) \xi_1^2}{P_2(|\xi|) + c^2 - \tau} \right| \lesssim 1. \tag{154}$$

We prove (154): First, notice that when $|k_1| = |k|_\infty$, then $|\xi_1| \sim |\xi|_\infty$ for $\xi \in \text{supp}(\sigma_k)$. We split the argument according to the cases $\tau - c^2 > 0$ and $\tau - c^2 \leq 0$. In the case $\tau - c^2 > 0$

$$\sup_{x_1, \tau, \xi_j (j \neq 1)} \left| \mathcal{F}_{\xi_1}^{-1} \frac{\sigma_k(\xi) \xi_1^2}{P_2(|\xi|) + c^2 - \tau} \right| \lesssim \left| \int_{|k_1-3/4}^{k_1+3/4} \frac{c^2}{\xi_1^2} d\xi_1 \right| \stackrel{|k_1| \gtrsim c}{\lesssim} 1.$$

When $\tau - c^2 \leq 0$ we set $\tau_2 := \tau_2(c) = c \left(\sqrt{\frac{5}{4} - \frac{\tau}{c^2}} - \frac{1}{2} \right) > 0$, in order to write

$$P_2(|\xi|) + c^2 - \tau = \left(\frac{|\xi|^2}{c} + \tau_2 \right) \left(-\frac{|\xi|^2}{c} + \tau_2 + c \right).$$

Hence

$$\begin{aligned} \mathcal{F}_{\xi_1}^{-1} \frac{\sigma_k(\xi) \xi_1^2}{P_2(|\xi|) + c^2 - \tau} &= \mathcal{F}_{\xi_1}^{-1} \frac{\sigma_k(\xi) \xi_1^2}{\left(\frac{|\xi|^2}{c} + \tau_2 \right) \left(-\frac{|\xi|^2}{c} + \tau_2 + c \right)} \\ &= \mathcal{F}_{\xi_1}^{-1} \frac{\sigma_k(\xi) \xi_1^2}{\left(\frac{\xi_1^2}{c} + \frac{|\bar{\xi}|^2}{c} + \tau_2 \right) \left(-\frac{\xi_1^2}{c} - \frac{|\bar{\xi}|^2}{c} + \tau_2 + c \right)}. \end{aligned} \tag{155}$$

When $|\bar{\xi}|^2 \geq c(\tau_2 + c)$, we can treat the problem as before.

Next, we consider the case $|\bar{\xi}|^2 < c(\tau_2 + c)$. Let

$$\begin{aligned} A^2 &:= A(\bar{\xi}, \tau, c)^2 = \frac{|\bar{\xi}|^2}{c} + \tau_2, \\ B^2 &:= B(\bar{\xi}, \tau, c)^2 = -\left(\frac{|\bar{\xi}|^2}{c} - \tau_2 - c \right), \end{aligned}$$

then

$$\begin{aligned}
 & \mathcal{F}_{\xi_1}^{-1} \frac{\eta_{k_1}(\xi_1)\xi_1^2}{\left(\frac{|\xi_1|^2}{c} + \tau_2\right)\left(-\frac{|\xi_1|^2}{c} + \tau_2 + c\right)} \\
 &= \mathcal{F}_{\xi_1}^{-1} \frac{\xi_1}{\frac{\xi_1^2}{c} + A^2} \frac{\xi_1}{B^2 - \frac{\xi_1^2}{c}} \eta_{k_1}(\xi_1) \\
 &= \frac{c^{1/2}}{2} \mathcal{F}_{\xi_1}^{-1} \frac{\xi_1}{\frac{\xi_1^2}{c} + A^2} \left(\frac{1}{B - \frac{\xi_1}{c^{1/2}}} - \frac{1}{B + \frac{\xi_1}{c^{1/2}}} \right) \eta_{k_1}(\xi_1) \\
 &=: I + II.
 \end{aligned}$$

We estimate only I , as the argument of II is similar. First we write

$$I = -\frac{c}{2} \mathcal{F}_{\xi_1}^{-1} \frac{\eta_{k_1}(\xi_1)}{\frac{\xi_1^2}{c} + A^2} + \frac{c}{2} \mathcal{F}_{\xi_1}^{-1} \frac{\eta_{k_1}(\xi_1)B}{\left(B - \frac{\xi_1}{c^{1/2}}\right)\left(\frac{\xi_1^2}{c} + A^2\right)} := I_1 + I_2.$$

Since $\mathcal{F}_{\xi_1}^{-1}(1/\xi_1)$ is the function $sgn(\xi_1)$, we have that I_1 is bounded uniformly with respect to c . For I_2 , it suffices to show

$$cB \sup_{x_1} \left| \mathcal{F}_{\xi_1}^{-1} \frac{1}{\left(B - \frac{\xi_1}{c^{1/2}}\right)\left(\frac{\xi_1^2}{c} + A^2\right)} \right| \lesssim 1.$$

Since $|\mathcal{F}(e^{-|\cdot|})(\xi)| \lesssim \frac{1}{1+|\xi|^2}$,

$$\begin{aligned}
 cB \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{\left(B - \frac{\xi_1}{c^{1/2}}\right)\left(\frac{\xi_1^2}{c} + A^2\right)} \right\|_{L^\infty_{x_1}} &\lesssim cB \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{B - \frac{\xi_1}{c^{1/2}}} \right\|_{L^\infty_{x_1}} \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{\frac{\xi_1^2}{c} + A^2} \right\|_{L^1_{x_1}} \\
 &\lesssim \frac{c^2 B}{A^2} \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{B - \frac{\xi_1}{c^{1/2}}} \right\|_{L^\infty_{x_1}} \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{\xi_1^2 A^{-2} + 1} \right\|_{L^1_{x_1}} \\
 &\lesssim B \left\| \mathcal{F}_{\xi_1}^{-1} \frac{1}{B - \frac{\xi_1}{c^{1/2}}} \right\|_{L^\infty_{x_1}} \cdot \frac{c^2}{A^2} \|Ae^{-A|\cdot|}\|_{L^1_{x_1}} \lesssim 1.
 \end{aligned}$$

Finally, we observe that in general the solution ψ of (151) may not vanish at $t = 0$. However, by Parseval’s identity

$$\psi(0, x) = \psi(t, x)|_{t=0} = K \int_I \mathcal{U}_2(s) \mathcal{F}(f)(s, x) ds,$$

for some $K > 0$, and if we combine it with (150), we have that $\square_k \mathcal{U}_2(t) d_{x_1}^2 \psi(0, x) \in L^2$. Hence, by (149)

$$\tilde{\sigma} \psi(t) := \psi(t) - \mathcal{U}_2(t) \psi(0, \cdot) = i \int_I \mathcal{U}_2(t - \tau) f(\tau) d\tau \tag{156}$$

is the solution of (151), and it satisfies (152). □

Lemma 6 For any $\sigma \in \mathbb{R}$ and $k \in \mathbb{Z}^d$ with $|k_i| \geq 4$,

$$\|\square_k D_{x_i}^\sigma \psi\|_{L_{x_i; (x_j)_{j \neq i}, t}^{p_1, p_2}} \lesssim \langle k_i \rangle^\sigma \|\square_k \psi\|_{L_{x_i; (x_j)_{j \neq i}, t}^{p_1, p_2}}. \tag{157}$$

If we replace $D_{x_i}^\sigma$ by $\partial_{x_i}^\sigma$, the above inequality holds for all $k \in \mathbb{Z}^d$.

Proof See the proof of Lemma 3.4 in [58]. One can check that both sides of (157) are equivalent for $|k_i| \geq 4$. □

By combining (152), (150) and (157) we obtain

Proposition 11 For any $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ with $|k_i| = |k|_\infty \gtrsim c$ we have

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim \|\square_k f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}, \tag{158}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_t^\infty L_x^2} \lesssim c \langle |k_i| \rangle^{1/2} \|\square_k f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}. \tag{159}$$

Remark 19 For the case $r > 2$ we replace (149), (150), (152), (158) and (159) with

$$\|\square_k D_{x_i}^{r-1/2} \mathcal{A}_r(t) \psi_0\|_{L_{x_i; (x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim c^{r-1} \|\square_k \psi_0\|_{L^2}, \tag{160}$$

$$\|\square_k \partial_{x_i}^r \mathcal{A}_r f\|_{L_t^\infty L_x^2} \lesssim c^{r-1} \|\square_k D_i^{1/2} f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}, \tag{161}$$

$$\|\square_k \partial_{x_i}^{2(r-1)} \psi\|_{L_{x_i; (x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim \|\square_k f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}, \tag{162}$$

$$\|\square_k \partial_{x_i}^{2(r-1)} \mathcal{A}_r f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{\infty, 2}} \lesssim \|\square_k f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}, \tag{163}$$

$$\|\square_k \partial_{x_i}^{2(r-1)} \mathcal{A}_r f\|_{L_t^\infty L_x^2} \lesssim c^{r-1} \langle |k_i| \rangle^{r-3/2} \|\square_k f\|_{L_{x_i; (x_j)_{j \neq i}, t}^{1, 2}}. \tag{164}$$

Remark 20 We point out the fact that we have worked out smoothing estimates only in the higher frequencies. As in [50], only these smoothing estimates are needed in order to discuss the well-posedness of (112).

8.2.3 Strichartz estimates

By exploiting (110) we can deduce Strichartz estimates for solutions of (112) combined with \square_k -decomposition operators.

Proposition 12 Let $r \geq 1$, $d \geq 2$, $c \geq 1$, $t > 0$. Let (p, q) and (a, b) be order- r admissible pairs. Then for any $0 < T \lesssim c^{2(r-1)}$ and for any $k \in \mathbb{Z}^d$ with $|k| \gtrsim K$ ($K = K(c)$ is defined in (130))

$$\begin{aligned} \|\square_k \mathcal{U}_r(t) \phi_0\|_{L^p([0, T]) L^q(\mathbb{R}^d)} &\lesssim c^d \left(1 - \frac{1}{r}\right) \left(\frac{1}{2} - \frac{1}{q}\right) \|\square_k \phi_0\|_{L^2(\mathbb{R}^d)} \\ &= c \left(1 - \frac{1}{r}\right)^{\frac{2r}{p}} \|\square_k \phi_0\|_{L^2(\mathbb{R}^d)}, \end{aligned} \tag{165}$$

$$\left\| \square_k \int_0^t \mathcal{U}_r(t - \tau) \phi(\tau) d\tau \right\|_{L^p([0, T]) L^q(\mathbb{R}^d)} \lesssim c \left(1 - \frac{1}{r}\right)^{2r} \left(\frac{1}{p} + \frac{1}{a}\right) \|\square_k \phi\|_{L^{a'}([0, T]) L^{b'}(\mathbb{R}^d)}. \tag{166}$$

Furthermore, by (106) we have that

$$\|\square_k \mathcal{U}_r(t)\|_{L^1 \rightarrow L^\infty} \lesssim c^{d(1-\frac{1}{r})} \langle t \rangle^{-d/(2r)}, \quad 0 < |t| \lesssim c^{2(r-1)},$$

and by following closely the argument in Section 5 of [59] we can deduce

Proposition 13 *Let $r \geq 1, d \geq 2, c \geq 1$. Let (p, q) be a Schrödinger-admissible pair; then*

$$\|\mathcal{U}_r(t)\psi_0\|_{L^1_{\square} (L^p_t([0, T])L^q_x)} \lesssim c^{(1-\frac{1}{r})\frac{2r}{p}} \|\psi_0\|_{M^2_{2,1}}, \quad 0 < T \lesssim c^{2(r-1)}, \quad (167)$$

$$\|\mathcal{A}_r f\|_{L^1_{\square} (L^p_t([0, T])L^q_x) \cap L^1_{\square} (L^\infty_t([0, T])L^2_x)} \lesssim c^{(1-\frac{1}{r})\frac{4r}{p}} \|f\|_{L^1_{\square} (L^{p'}([0, T])L^{q'}(\mathbb{R}^d))}. \quad (168)$$

8.2.4 Maximal function estimates

In this subsection we study the maximal function estimates for the semigroup $\mathcal{U}_r(t)$ and the integral operator $\int_0^t \mathcal{U}_r(t - \tau) \cdot d\tau$ in anisotropic Lebesgue spaces. To do this, we will need the time decay properties proved in Sec. 8.2.1. As always, we first prove results for the case $r = 2$, and then we write the modification for the general case.

Lemma 7 *1. Let $q \geq 2, \frac{8}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k| \gtrsim K(c)$, then*

$$\|\square_k \mathcal{U}_2(t)\psi_0\|_{L^{q,\infty}_{x_i:(x_j)_{j \neq i,t}}} \lesssim c^{d/2} \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \quad 0 < |t| \lesssim c^2, \quad \forall i = 1, \dots, d. \quad (169)$$

2. Let $q \geq 2, \frac{4}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k| \lesssim K(c)$, then

$$\|\square_k \mathcal{U}_2(t)\psi_0\|_{L^{q,\infty}_{x_i:(x_j)_{j \neq i,t}}} \lesssim c \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \quad \forall i = 1, \dots, d. \quad (170)$$

Proof Clearly it suffices to show the thesis for $i = 1$; recall that for any $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ we denote $\bar{x} = (x_2, \dots, x_d)$. By a standard TT^* argument, (169) is equivalent to

$$\left\| \int_{\mathbb{R}^d} e^{i\langle x, \xi \rangle} e^{it(c^2 + P_2(|\xi|))} \sigma_k(\xi) d\xi \right\|_{L^{q/2,\infty}_{x_1;\bar{x},t}} \lesssim \langle k \rangle^{2/q}. \quad (171)$$

If $|k| \gtrsim K(c)$, then

$$\|\mathcal{F}^{-1} e^{it(c^2 + P_2(|\xi|))} \sigma_k(\xi)\|_{L^\infty_x} \stackrel{(132)}{\lesssim} c^{d/2} \langle k \rangle^{-d} |t|^{-d/4}, \quad 0 < |t| \lesssim c^2; \quad (172)$$

on the other hand

$$\|\square_k \mathcal{U}_2(t)\mathcal{F}^{-1} \sigma_k\|_{L^\infty_{t,x}} \lesssim \|\square_k \mathcal{U}_2(t)\mathcal{F}^{-1} \sigma_k\|_{L^r_t L^\infty_x} \lesssim 1. \quad (173)$$

If we combine (172) and (173), we obtain

$$|\square_k \mathcal{U}_2(t)\mathcal{F}^{-1} \sigma_k| \lesssim c^{d/2} (1 + \langle k \rangle^4 |t|)^{-d/4}, \quad 0 < |t| \lesssim c^2. \quad (174)$$

Now, if $|x_1| \gtrsim 1 + |t| \langle k \rangle^5$, by integrating by parts we get

$$|\square_k \mathcal{U}_2(t)\mathcal{F}^{-1} \sigma_k| \lesssim c^{d/2} \langle x_1 \rangle^{-2}. \quad (175)$$

If $|x_1| \lesssim 1 + |t| \langle k \rangle^5$, by (174) we can deduce

$$|\square_k \mathcal{U}_2(t)\mathcal{F}^{-1} \sigma_k| \lesssim c^{d/2} (1 + |x_1| \langle k \rangle^{-1})^{-d/4} \quad (176)$$

Combining (175) and (176) we have

$$\sup_{\bar{x}, t} |\square_k \mathcal{W}_2(t) \mathcal{F}^{-1} \sigma_k| \lesssim c^{d/2} \langle x_1 \rangle^{-2} + c^{d/2} (1 + |x_1| \langle k \rangle^{-1})^{-d/4}, \tag{177}$$

from which, by taking the $L_{x_1}^{q/2}$ norm on both sides, we obtain (171). The proof for the case $|k| \lesssim K(c)$ is similar. \square

Lemma 8 *Let $q \geq 2$, $\frac{8}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k_i| \gtrsim K(c)^2$, then*

$$\|\square_k \mathcal{A}_2 f\|_{L_{x_i: (x_j)_{j \neq i, t}}^{q, \infty}} \lesssim c^{d/2} \langle k_i \rangle^{-3/2+1/q} \|\square_k f\|_{L_{x_i: (x_j)_{j \neq i, t}}^{1,2}}, \quad 0 < |t| \lesssim c^2, \quad \forall i = 1, \dots, d. \tag{178}$$

Proof It suffices to prove the case $i = 1$. Recall that the solution of (151) is of the form

$$\psi = \mathcal{F}_{\tau, \xi}^{-1} \frac{1}{c^2 + P_2(|\xi|) - \tau} \mathcal{F}_{t, x} f;$$

hence, its frequency localization can be written as

$$\square_k \psi = \mathcal{F}_{\tau, \xi}^{-1} \frac{1}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t, x} \square_k f)(\tau, \xi).$$

For convenience, we introduce the following regions

$$\begin{aligned} \mathbb{E}_1 &= \{ \tau - c^2 \leq -c^2/4 \}, \\ \mathbb{E}_2 &= \left\{ -c^2/4 \leq \tau - c^2 \leq |\bar{\xi}|^2 \left(-\frac{|\bar{\xi}|^2}{c^2} + 1 \right) \right\}, \\ \mathbb{E}_3 &= \left\{ \tau - c^2 \geq |\bar{\xi}|^2 \left(-\frac{|\bar{\xi}|^2}{c^2} + 1 \right) \right\}, \end{aligned}$$

and we make the following decomposition

$$c^2 + P_2(|\xi|) - \tau = \begin{cases} \left(\frac{|\bar{\xi}|^2}{c} + \tau_2(c, \tau) \right) \left(-\frac{\xi_1}{c^{1/2}} + a \right) \left(\frac{\xi_1}{c^{1/2}} + a \right) & (\bar{\xi}, \tau) \in \mathbb{E}_1, \\ \left(\frac{|\bar{\xi}|^2}{c} + \tau_2(c, \tau) \right) \left(-\frac{|\bar{\xi}|^2}{c} + \tau_2(c, \tau) + c \right) & (\bar{\xi}, \tau) \in \mathbb{E}_2, \\ -\left(\frac{|\bar{\xi}|^2}{c} - \frac{c}{2} \right)^2 + \left(\frac{5}{4}c^2 - \tau \right) & (\bar{\xi}, \tau) \in \mathbb{E}_3, \end{cases} \tag{179}$$

where $a = a(c, \bar{\xi}, \tau) := (\tau_2(c, \tau) - |\bar{\xi}|^2/c + c)^{1/2}$. We denote

$$\square_k \psi_i = \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{E}_i}(\bar{\xi}, \tau)}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t, x} \square_k f)(\tau, \xi), \quad i = 1, 2, 3.$$

First, we estimate $\square_k \psi_1$. Set $\tilde{\sigma} \eta_{k_1}(\xi_1) = \sum_{|l| \leq 10} \eta_{k_1+l}(\xi_1)$. First we notice that

$$\begin{aligned} \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)}{c^2 + P_2(|\xi|) - \tau} &= \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)}{(2\tau_2(c, \tau) + c) \left(\frac{|\bar{\xi}|^2}{c} + \tau_2(c, \tau) \right)} \\ &+ \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)}{2a(2\tau_2(c, \tau) + c)} \left(\frac{1}{-\frac{\xi_1}{c^{1/2}} + a} + \frac{1}{\frac{\xi_1}{c^{1/2}} + a} \right) \\ &=: \sum_{j=1}^3 A_j(c, \xi, \tau). \end{aligned}$$

According to the above decomposition, we can rewrite $\square_k \psi_1$ as

$$\begin{aligned} \square_k \psi_1 &= \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau) \tilde{\sigma} \eta_{k_1}(ac^{1/2})}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t,x} \square_k f)(\tau, \xi) \\ &\quad + \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau) (1 - \tilde{\sigma} \eta_{k_1}(ac^{1/2}))}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t,x} \square_k f)(\tau, \xi) \\ &= \sum_{j=1}^3 \mathcal{F}_{\tau, \xi}^{-1} \chi_{\mathbb{E}_1}(\bar{\xi}, \tau) A_j(\xi, \tau) \tilde{\sigma} \eta_{k_1}(ac^{1/2}) (\mathcal{F}_{t,x} \square_k f)(\tau, \xi) \\ &\quad + \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau) (1 - \tilde{\sigma} \eta_{k_1}(ac^{1/2}))}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t,x} \square_k f)(\tau, \xi) \\ &=: I + II + III + IV. \end{aligned}$$

Case $k_1 \gtrsim K(c)^2$: First, we estimate II. Let $\tilde{\sigma} \sigma_k$ be as in (125), then

$$\begin{aligned} II &= \int_{I \times \mathbb{R}^d} \frac{e^{it\tau + i\langle \bar{x}, \bar{\xi} \rangle} \chi_{\mathbb{E}_1}(\bar{\xi}, \tau) \tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) \tilde{\sigma} \eta_{k_1}(ac^{1/2}) \square_k \widehat{f}(y_1, \cdot)}{2a(2\tau_2(c, \tau) + c)} (\bar{\xi}, \tau) c^{1/2} e^{i(x_1 - y_1)ac^{1/2}} \\ &\quad \times \operatorname{sgn}(x_1 - y_1) d\bar{\xi} dy_1 d\tau. \end{aligned}$$

By changing variable, $\xi_1 = c^{1/2}a(c, \bar{\xi}, \tau)$, and by setting $\tilde{\sigma} \rho_k(\xi) = \tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) \tilde{\sigma} \eta_{k_1}(\xi_1)$, we obtain

$$|II| \lesssim \left| \int dy_1 \operatorname{sgn}(x_1 - y_1) \int e^{it(c^2 + P_2(|\xi|))} e^{i(x_1 - y_1)\xi_1 + i\langle \bar{x}, \bar{\xi} \rangle} \tilde{\sigma} \rho_k(\xi) \square_k \widehat{f}(y_1, \cdot) (c^2 + P_2(|\xi|), \tau) d\xi \right|,$$

and by applying (169) we get

$$\begin{aligned} \|II\|_{L_{x_1; \bar{x}, t}^{q, \infty}} &\lesssim \int dy_1 \left\| \int e^{it(c^2 + P_2(|\xi|))} e^{i(x_1 - y_1)\xi_1 + i\langle \bar{x}, \bar{\xi} \rangle} \tilde{\sigma} \rho_k(\xi) \square_k \widehat{f}(y_1, \cdot) (c^2 + P_2(|\xi|), \tau) d\xi \right\|_{L_{x_1; \bar{x}, t}^{q, \infty}} \\ &\lesssim c^{d/2} \langle k_1 \rangle^{1/q} \int \|\tilde{\sigma} \rho_k(\xi) \square_k \widehat{f}(y_1, \cdot) (c^2 + P_2(|\xi|), \tau)\|_{L_{\xi}^2} dy_1 \\ &\stackrel{(150), (157)}{\lesssim} c c^{d/2} \langle k_1 \rangle^{1/q - 3/2} \|\square_k f\|_{L_{x_1; \bar{x}, t}^{1, 2}}. \end{aligned} \tag{180}$$

Since $k_1 > 0$, III has the same upper bound as in (180).

Now we estimate IV: First notice that

$$\begin{aligned} IV &= \int dy_1 \int e^{it\tau + i\langle \bar{x}, \bar{\xi} \rangle} \tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) \square_k \widehat{f}(y_1, \cdot) (\bar{\xi}, \tau) K(x_1 - y_1, a, \bar{\xi}) d\bar{\xi}, \\ K(x_1, a, \bar{\xi}) &= \chi_{\mathbb{E}_1}(\bar{\xi}, \tau) (1 - \tilde{\sigma} \eta_{k_1}(ac^{1/2})) \int \frac{\sum_{|l| \leq 1} \eta_{k_1 + l}(\xi_1) e^{ix_1 \xi_1}}{c^2 + P_2(|\xi|) - \tau} d\xi_1. \end{aligned}$$

By Young’s inequality for convolutions, Hölder’s inequality and Minkowski’s inequality we have

$$\begin{aligned} \|IV\|_{L_{x_1; \bar{x}, t}^{q, \infty}} &\leq \left\| \int \|\tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) \square_k \widehat{f}(y_1, \cdot) (\bar{\xi}, \tau) K(x_1 - y_1, a, \bar{\xi})\|_{L_{\bar{\xi}, \tau}^1} dy_1 \right\|_{L_{x_1}^q} \\ &\leq \|\square_k f\|_{L_{x_1; \bar{x}, t}^{1, 2}} \|\tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) K(x_1, a, \bar{\xi})\|_{L_{x_1; \bar{x}, \tau}^{q, 2}} \\ &\lesssim \|\square_k f\|_{L_{x_1; \bar{x}, t}^{1, 2}} \|\tilde{\sigma} \sigma_{\bar{k}}(\bar{\xi}) K(x_1, a, \bar{\xi})\|_{L_{\bar{\xi}}^\infty L_\tau^2 L_{x_1}^q}. \end{aligned}$$

Integrating by parts it follows that

$$\begin{aligned} & \|\tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi})K(x_1, a, \xi)\|_{L_{\bar{\xi}}^{\infty}L_{\tau}^2L_{x_1}^q} \\ & \lesssim \sup_{|\bar{\xi}-k|_{\infty}\leq 3} \sum_{j=0}^1 \|\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)(1 - \tilde{\sigma}\eta_{k_1}(ac^{1/2}))\partial_{\xi_1}^j(c^2 + P_2(|\xi|) - \tau)^{-1}\|_{L_{\tau}^2}. \end{aligned} \tag{181}$$

Noticing that $|\xi - ac^{1/2}| \geq c^{1/2} \geq 1$ in the support set of $(1 - \tilde{\sigma}\eta_{k_1}(ac^{1/2}))\chi_{|\xi_1-k_1|\leq 3}\partial_{\xi_1}^j(c^2 + P_2(|\xi|) - \tau)^{-1}$ we can deduce from (179) that there is no singularity if we integrate (181), and this gives

$$\|\tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi})K(x_1, a, \bar{\xi})\|_{L_{\bar{\xi}}^{\infty}L_{\tau}^2L_{x_1}^q} \lesssim c^{1/2}|k_1|^{-3/2}.$$

Now we estimate I : We begin by setting

$$J(x_1, a, \bar{\xi}) = \chi_{\mathbb{E}_1}(\bar{\xi}, \tau)\tilde{\sigma}\eta_{k_1}(ac^{1/2}) \int \frac{\sum_{|l|\leq 1} \eta_{k+l}(\xi_1)e^{ix_1\xi_1}}{(2\tau_2(c, \tau) + c) \left(\frac{|xi|^2}{c} + \tau_2(c, \tau)\right)} d\xi_1. \tag{182}$$

One can check that

$$I = \int dy_1 \int e^{it\tau + i\langle \bar{x}, \bar{\xi} \rangle} \tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi}) \square_k \widehat{f}(y_1, \cdot)(\bar{\xi}, \tau) J(x_1 - y_1, a, \bar{\xi}) d\bar{\xi} d\tau.$$

Similar to the estimate of IV , by Young’s, Hölder’s and Minkowski’s inequalities we obtain

$$\|I\|_{L_{x_1; \bar{x}, t}^{q, \infty}} \lesssim c^{1/2} \|\square_k f\|_{L_{x_1; \bar{x}, t}^{1, 2}} \|\tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi})J(x_1, a, \bar{\xi})\|_{L_{\bar{\xi}}^{\infty}L_{\tau}^2L_{x_1}^q}.$$

By integration by parts we get

$$|J(x_1, a, \xi)| \lesssim \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)\tilde{\sigma}\eta_{k_1}(ac^{1/2})}{(2\tau_2(c, \tau) + c)(1 + |x_1|)} \sum_{j=0}^1 \int_{|\xi_1-k_1|\leq 3} \left| \partial_{\xi_1}^j \left(\frac{|xi|^2}{c} + \tau_2(c, \tau) \right)^{-1} \right| d\xi_1.$$

Therefore,

$$\begin{aligned} & \|\tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi})J(x_1, a, \bar{\xi})\|_{L_{\bar{\xi}}^{\infty}L_{\tau}^2L_{x_1}^q} \\ & \lesssim \sup_{|\bar{\xi}-k|_{\infty}\leq 3} \sum_{j=0}^1 \left\| \frac{\chi_{\mathbb{E}_1}(\bar{\xi}, \tau)\tilde{\sigma}\eta_{k_1}(ac^{1/2})}{(2\tau_2(c, \tau) + c)(1 + |x_1|)} \sum_{j=0}^1 \left| \partial_{\xi_1}^j \left(\frac{|xi|^2}{c} + \tau_2(c, \tau) \right)^{-1} \right| \right\|_{L_{\tau}^2}, \end{aligned} \tag{183}$$

and noticing that $|ac^{1/2} - k_1| \leq 20$ in the support set of $\tilde{\sigma}\eta_{k_1}(ac^{1/2})$, we can deduce that $2\tau_2(c, \tau) + c \gtrsim k_1^2$, and finally we obtain

$$\|\tilde{\sigma}\sigma_{\bar{k}}(\bar{\xi})J(x_1, a, \bar{\xi})\|_{L_{\bar{\xi}}^{\infty}L_{\tau}^2L_{x_1}^q} \lesssim |k_1|^{-2}. \tag{184}$$

The proof for the case $k \lesssim -K(c)^2$ is similar. Furthermore, in the estimate of $\square_k \psi_2$ and $\square_k \psi_3$ we can check that there is no singularity in $(c^2 + P_2(|\xi|) - \tau)^{-1}$ for $|\xi_1| \geq c^{1/2}$ and $(\bar{\xi}, \tau) \in \mathbb{E}_2 \cup \mathbb{E}_3$. Hence one can argue as in (182)–(184) and conclude. \square

In the last lemma we proved that $\square_k \mathcal{A}_2 : L_{x_1, (x_j)_{j \neq 2}, t}^{1, 2} \rightarrow L_{x_1, (x_j)_{j \neq 2}, t}^{2, \infty}$. In the next lemma we show that $\square_k \mathcal{A}_2 : L_{x_2, (x_j)_{j \neq 1}, t}^{1, 2} \rightarrow L_{x_1, (x_j)_{j \neq 2}, t}^{2, \infty}$.

Lemma 9 *Let $q \geq 2$, $\frac{8}{d} < q \leq +\infty$, $k \in \mathbb{Z}^d$ with $|k_i| \gtrsim c$ and $h, i \in \{1, \dots, d\}$ with $h \neq i$, then*

$$\|\square_k \mathcal{A}_2 f\|_{L^{q, \infty}_{x_h; (x_j)_{j \neq h, i}}} \lesssim c^{1+d/2} \langle k_i \rangle^{-3/2+1/q} \|\square_k f\|_{L^{1,2}_{x_i; (x_j)_{j \neq i, i}}}, \quad 0 < |t| \lesssim c^2. \quad (185)$$

Proof It clearly suffices to consider the case $h = 1, i = 2$ and $k_2 \gtrsim c$. The proof goes along the same line of that of (178), and we will only prove in detail the parts that are different. For convenience, we denote $\tilde{\sigma} \xi = (\xi_1, \xi_3, \dots, \xi_d)$. We introduce the following regions

$$\begin{aligned} \mathbb{F}_1 &= \{ \tau - c^2 \leq -c^2/4 \}, \\ \mathbb{F}_2 &= \left\{ -c^2/4 \leq \tau - c^2 \leq |\tilde{\sigma} \xi|^2 \left(-\frac{|\tilde{\sigma} \xi|^2}{c^2} + 1 \right) \right\}, \\ \mathbb{F}_3 &= \left\{ \tau - c^2 \geq |\tilde{\sigma} \xi|^2 \left(-\frac{|\tilde{\sigma} \xi|^2}{c^2} + 1 \right) \right\}, \end{aligned}$$

and we make the following decomposition

$$c^2 + P_2(|\xi|) - \tau = \begin{cases} \left(\frac{|\xi|^2}{c} + \tau_2(c, \tau) \right) \left(-\frac{\xi_2}{c^{1/2}} + a \right) \left(\frac{\xi_2}{c^{1/2}} + a \right) & (\tilde{\sigma} \xi, \tau) \in \mathbb{F}_1, \\ \left(\frac{|\xi|^2}{c} + \tau_2(c, \tau) \right) \left(-\frac{|\xi|^2}{c} + \tau_2(c, \tau) + c \right), & (\tilde{\sigma} \xi, \tau) \in \mathbb{F}_2, \\ -\left(\frac{|\xi|^2}{c} - \frac{c}{2} \right)^2 + \left(\frac{5}{4}c^2 - \tau \right), & (\tilde{\sigma} \xi, \tau) \in \mathbb{F}_3, \end{cases} \quad (186)$$

where $b = b(c, \tilde{\sigma} \xi, \tau) := (\tau_2(c, \tau) - |\tilde{\sigma} \xi|^2/c + c)^{1/2}$, $\tau_2(c, \tau) = c \left(\sqrt{\frac{5}{4} - \frac{\tau}{c^2}} - \frac{1}{2} \right)$. We denote

$$\square_k \tilde{\sigma} \psi_i = \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{F}_i}(\tilde{\sigma} \xi, \tau)}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t, x} \square_k f)(\tau, \xi), \quad i = 1, 2, 3.$$

We estimate $\square_k \tilde{\sigma} \psi_1$, since by definition of the regions \mathbb{F}_i the estimate of the other terms follow more easily, like in the last Lemma.

Set $\tilde{\sigma} \eta_{k_2}(\xi_2) = \sum_{|l| \leq 10} \eta_{k_2+l}(\xi_2)$. First we notice that

$$\begin{aligned} \frac{\chi_{\mathbb{F}_1}(\tilde{\sigma} \xi, \tau)}{c^2 + P_2(|\xi|) - \tau} &= \frac{\chi_{\mathbb{F}_1}(\tilde{\sigma} \xi, \tau)}{(2\tau_2(c, \tau) + c) \left(\frac{|\xi|^2}{c} + \tau_2(c, \tau) \right)} \\ &+ \frac{\chi_{\mathbb{F}_1}(\tilde{\sigma} \xi, \tau)}{2b(2\tau_2(c, \tau) + c)} \left(-\frac{1}{c^{1/2}} + b + \frac{1}{c^{1/2}} + b \right) \\ &=: \sum_{j=1}^3 B_j(c, \xi, \tau). \end{aligned}$$

According to the above decomposition, we can rewrite $\square_k \psi_1$ as

$$\begin{aligned} &\square_k \tilde{\sigma} \psi_1 \\ &= \sum_{j=1}^3 \mathcal{F}_{\tau, \xi}^{-1} \chi_{\mathbb{F}_1}(\tilde{\sigma} \xi, \tau) B_j(\xi, \tau) \tilde{\sigma} \eta_{k_2}(bc^{1/2}) (\mathcal{F}_{t, x} \square_k f)(\tau, \xi) \\ &+ \mathcal{F}_{\tau, \xi}^{-1} \frac{\chi_{\mathbb{F}_1}(\tilde{\sigma} \xi, \tau) (1 - \tilde{\sigma} \eta_{k_2}(bc^{1/2}))}{c^2 + P_2(|\xi|) - \tau} (\mathcal{F}_{t, x} \square_k f)(\tau, \xi) \\ &=: I + II + III + IV. \end{aligned}$$

The estimates of *II* and *III* follow in the same way as for (178) by exchanging the roles of ξ_1 and ξ_2 . Now we estimate *I*: Set

$$m(\xi, \tau) = \frac{\chi_{\mathbb{E}_1}(\tilde{\sigma}\xi, \tau)\tilde{\sigma}\sigma_{\bar{k}}(\tilde{\sigma}\xi)\tilde{\sigma}\eta_{k_2}(bc^{1/2})}{2b(2\tau_2(c, \tau) + c)}, \tag{187}$$

and notice that $2\tau_2(c, \tau) + c \gtrsim k_2^2$ in the support set of m ; hence, for sufficiently large c , we have

$$m(\xi, \tau) \lesssim \frac{\chi_{\mathbb{F}_1}\tilde{\sigma}\sigma_k(\xi)}{k_2^4},$$

and therefore,

$$\|m\|_{L_{\xi_2}^1 L_{\xi_3, \dots, \xi_d, \tau}^2 L_{\xi_1}^{2q/(q-2)}} \lesssim |k_2|^{-2}. \tag{188}$$

Now, since by Young’s, Hölder’s and Minkowski’s inequalities we have

$$\begin{aligned} \|\mathcal{F}_{\xi, \tau}^{-1}m(\xi, \tau)(\mathcal{F}_{t,x}\square_k f)\|_{L_{x_1;\bar{x},t}^{q,\infty}} &\lesssim \|\mathcal{F}_{\xi, \tau}^{-1}m(\xi, \tau)(\mathcal{F}_{t,x}\square_k f)\|_{L_{x_1}^q L_{\tilde{\sigma}\xi, \tau}^1 L_{\xi_1}^{q'}} \\ &\lesssim \|m(\xi, \tau)(\mathcal{F}_{t,x}\square_k f)\|_{L_{\tilde{\sigma}\xi, \tau}^1 L_{\xi_1}^{q'}} \\ &\lesssim \|m\|_{L_{\xi_2}^1 L_{\xi_3, \dots, \xi_d, \tau}^2 L_{\xi_1}^{2q/(q-2)}} \|\mathcal{F}_{t,x}\square_k f\|_{L_{\xi_2}^\infty L_{(\xi_j)_{j\neq 2, \tau}}^2} \\ &\lesssim \|m\|_{L_{\xi_2}^1 L_{\xi_3, \dots, \xi_d, \tau}^2 L_{\xi_1}^{2q/(q-2)}} \|\square_k f\|_{L_{x_2}^1 L_{(x_j)_{j\neq 2, t}}^2}, \end{aligned} \tag{189}$$

we can deduce that

$$\|I\|_{L_{x_1;\bar{x},t}^{q,\infty}} \lesssim |k_2|^{-2} \|\square_k f\|_{L_{x_1}^1 L_{(x_j)_{j\neq 2, t}}^2}. \tag{190}$$

Now we estimate *IV*: Set

$$m_k(\xi, \tau) := \frac{\chi_{\mathbb{F}_3}(\tilde{\sigma}\xi, \tau)\tilde{\sigma}\sigma_k(\xi)(1 - \tilde{\sigma}\eta_{k_2}(b))}{c^2 + P_2(|\xi|) - \tau}, \quad \mathcal{M}_k(f) := \mathcal{F}_{\tau, \xi}^{-1}m_k(\xi, \tau)(\mathcal{F}_{t,x}f), \tag{191}$$

and notice that $\mathcal{M}_k(f)$ is the solution of the inhomogeneous equation

$$-i\psi_t = A_{c,2}\psi - \mathcal{F}_{\tau, \xi}^{-1}m_k(\xi, \tau)(c^2 + P_2(|\xi|) - \tau)(\mathcal{F}_{t,x}f).$$

Applying (159) (recall that $k_2 \gtrsim c$), we have

$$\begin{aligned} \|\mathcal{M}_k(f)\|_{L_{x,t}^\infty} &\lesssim c^{3/2} \|\mathcal{M}_k(f)\|_{L_t^\infty L_x^2} \\ &\lesssim c^{d/2} c |k_2|^{-3/2} \|f\|_{L_{x_2}^1 L_{(x_j)_{j\neq 2, t}}^2} \\ &= c^{\frac{d}{2}+1} |k_2|^{-3/2} \|f\|_{L_{x_2}^1 L_{(x_j)_{j\neq 2, t}}^2}. \end{aligned} \tag{192}$$

Next, for $(\xi, \tau) \in \text{supp}(m_k)$,

$$|c^2 + P_2(|\xi|) - \tau| \stackrel{(\tilde{\sigma}\xi, \tau) \in \mathbb{F}_3}{\gtrsim} c^{-1} \langle k \rangle^2 |k_2|. \tag{193}$$

By the definition of b we have that for $\frac{|\tilde{\sigma}\xi|}{c} \lesssim \frac{1}{2}(c + \tau_2(c, \tau))$

$$|c^2 + P_2(|\xi|) - \tau| \gtrsim (c + \tau_2(c, \tau))^{3/2}, \quad (\xi, \tau) \in \text{supp}(m_k), \tag{194}$$

while for $\frac{|\tilde{\alpha}\xi|}{c} \gtrsim \frac{1}{2}(c + \tau_2(c, \tau))$ we can exploit the fact that $|k|_\infty = |k_2| \gtrsim c$ to obtain again that

$$|c^2 + P_2(|\xi|) - \tau| \gtrsim (c + \tau_2(c, \tau))^{3/2}, \quad (\xi, \tau) \in \text{supp}(m_k), \tag{195}$$

and by combining (191) with (194)–(195) we obtain

$$m_k(\xi, \tau) \lesssim c \frac{\chi_{\tau \geq \frac{3}{4}c^2} \tilde{\sigma} \sigma_k(\xi)}{(|k_2|^2 + c + \tau_2(c, \tau))^{3/2}}, \tag{196}$$

which gives

$$\|m_k\|_{L^1_{\xi_2} L^2_{\xi_3, \dots, \xi_d, \tau} L^\infty_{\xi_1}} \lesssim c|k_2|^{-1}. \tag{197}$$

Therefore, from (191) and (189) we can deduce

$$\|\mathcal{M}_k(f)\|_{L^{2, \infty}_{x_1; \bar{x}, t}} \lesssim c|k_2|^{-1} \|f\|_{L^1_{x_2} L^2_{(x_j)_{j \neq 2, t}}}. \tag{198}$$

For any $q \geq 2$ we obtain by interpolation between (192) and (198)

$$\|\mathcal{M}_k(f)\|_{L^{q, \infty}_{x_1; \bar{x}, t}} \lesssim c^{1+d(\frac{1}{2}-\frac{1}{q})} |k_2|^{-3/2+1/q} \|f\|_{L^1_{x_2} L^2_{(x_j)_{j \neq 2, t}}}, \tag{199}$$

and replacing f by $\square_k f$ in (199), we finally obtain

$$\|IV\|_{L^{q, \infty}_{x_1; \bar{x}, t}} \lesssim c^{1+d(\frac{1}{2}-\frac{1}{q})} |k_2|^{-3/2+1/q} \|f\|_{L^1_{x_2} L^2_{(x_j)_{j \neq 2, t}}}.$$

□

If we collect (178) and (185), we can deduce

Lemma 10 *Let $q \geq 2$, $\frac{8}{d} < q \leq +\infty$, $k \in \mathbb{Z}^d$ with $|k_i| \gtrsim c$ and $h, i \in \{1, \dots, d\}$, then*

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L^{q, \infty}_{x_h; (x_j)_{j \neq h, t}}} \lesssim c^{1+d/2} \langle k_i \rangle^{1/2+1/q} \|\square_k f\|_{L^{1,2}_{x_i; (x_j)_{j \neq i, t}}}, \quad 0 < |t| \lesssim c^2. \tag{200}$$

Remark 21 In the general case $r > 2$ we have

1. Let $q \geq 2$, $\frac{4r}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k| \gtrsim K(c)$, then

$$\begin{aligned} \|\square_k \mathcal{W}_r(t) \psi_0\|_{L^{q, \infty}_{x_i; (x_j)_{j \neq i, t}}} &\lesssim c^{d(1-\frac{1}{r})} \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \\ 0 < |t| &\lesssim c^{2(r-1)}, \quad \forall i = 1, \dots, d. \end{aligned} \tag{201}$$

2. Let $q \geq 2$, $\frac{4}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k| \lesssim K(c)$, then

$$\|\square_k \mathcal{W}_r(t) \psi_0\|_{L^{q, \infty}_{x_i; (x_j)_{j \neq i, t}}} \lesssim c \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \quad \forall i = 1, \dots, d. \tag{202}$$

3. Let $q \geq 2$, $\frac{4r}{d} < q \leq +\infty$ and $k \in \mathbb{Z}^d$ with $|k_i| \gtrsim K(c)^2$ and $i \in \{1, \dots, d\}$, then

$$\|\square_k \mathcal{A}_r f\|_{L^{q, \infty}_{x_i; (x_j)_{j \neq i, t}}} \lesssim c^{r-1+d(1-\frac{1}{r})} \langle k_i \rangle^{-r+1/2+1/q} \|\square_k f\|_{L^{1,2}_{x_i; (x_j)_{j \neq i, t}}}, \quad 0 < |t| \lesssim c^{2(r-1)}. \tag{203}$$

4. Let $q \geq 2, \frac{4r}{d} < q \leq +\infty, k \in \mathbb{Z}^d$ with $|k_i| \gtrsim c$ and $h, i \in \{1, \dots, d\}$ with $h \neq i$, then

$$\begin{aligned} & \|\square_k \mathcal{A}_r f\|_{L^{q,\infty}_{x_h;(x_j)_{j \neq h},t}} \\ & \lesssim c^{r-1+d(1-\frac{1}{r})} \langle k_i \rangle^{-r+1/2+1/q} \|\square_k f\|_{L^{1,2}_{x_i;(x_j)_{j \neq i},t}}, \quad 0 < |t| \lesssim c^{2(r-1)}. \end{aligned} \tag{204}$$

5. Let $q \geq 2, \frac{4r}{d} < q \leq +\infty, k \in \mathbb{Z}^d$ with $|k_i| \gtrsim c$ and $h, i \in \{1, \dots, d\}$, then

$$\begin{aligned} & \|\square_k \partial_{x_i}^{2(r-1)} \mathcal{A}_r f\|_{L^{q,\infty}_{x_h;(x_j)_{j \neq h},t}} \\ & \lesssim c^{r-1+d(1-\frac{1}{r})} \langle k_i \rangle^{r-3/2+1/q} \|\square_k f\|_{L^{1,2}_{x_i;(x_j)_{j \neq i},t}}, \quad 0 < |t| \lesssim c^{2(r-1)}. \end{aligned} \tag{205}$$

8.2.5 Proof of the local well-posedness

In this subsection we use smoothing estimates, Strichartz estimates and maximal function estimates in order to prove Proposition 7. In order to do so, it seems necessary to estimate norms in which partial derivatives and anisotropic Lebesgue spaces have different directions, for example $\|\partial_{x_1}^2 \square_k \mathcal{A} f\|_{L^{2,\infty}_{x_2;(x_j)_{j \neq 2},t}}$ with $|k_\infty| = |k_3|$. As usual, we show results for the case $r = 2$, and then we point out the modifications for the case $r > 2$.

Lemma 11 *Let $i, l, m \in \{1, \dots, d\}, 1 \leq p, q, \leq +\infty$. Assume that $k = (k_1, \dots, k_d)$ with $|k|_\infty = |k_m| \gtrsim c$, then*

$$\|\square_k \partial_{x_l}^2 f\|_{L^{p,q}_{x_i;(x_j)_{j \neq i},t}} \lesssim \|\square_k \partial_{x_m}^2 f\|_{L^{p,q}_{x_i;(x_j)_{j \neq i},t}}. \tag{206}$$

Proof

$$\begin{aligned} \|\square_k \partial_{x_l}^2 f\|_{L^{p,q}_{x_i;(x_j)_{j \neq i},t}} & \lesssim \sum_{|h_l|_\infty, |h_m|_\infty \leq 1} \left\| \mathcal{F}_{\xi_l, \xi_m}^{-1} \left(\left(\frac{\xi_l}{\xi_m} \right)^2 \eta_{k_l+h_l}(\xi_l) \eta_{k_m+l_m}(\xi_m) \right) \right\|_{L^1(\mathbb{R}^2)} \\ & \quad \times \|\square_k \partial_{x_m}^2 f\|_{L^{p,q}_{x_i;(x_j)_{j \neq i},t}} \\ & \lesssim \|\square_k \partial_{x_m}^2 f\|_{L^{p,q}_{x_i;(x_j)_{j \neq i},t}}. \end{aligned}$$

□

Lemma 12 1. *Let (a, b) be order-2 admissible, $i \in \{1, \dots, d\}, q \geq 2, \frac{8}{d} < q < +\infty$ and $k \in \mathbb{Z}^d$ with $|k|_\infty \gtrsim K(c)$, then*

$$\|\square_k \partial_{x_i}^\alpha \mathcal{A}_2 f\|_{L^{q,\infty}_{x_i;(x_j)_{j \neq i},t}} \lesssim c^{\frac{d}{2} + \frac{2}{a}} \langle |k|_\infty \rangle^{\alpha+1/q} \|\square_k f\|_{L_t^q L_x^{b'}}, \quad 0 < |t| \lesssim c^2. \tag{207}$$

2. *Let (a, b) be Schrödinger-admissible, $i \in \{1, \dots, d\}$, then*

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_t^q L_x^b} \lesssim c^{1+4/p} \langle |k|_\infty \rangle^{1/2} \|\square_k f\|_{L^{1,2}_{x_i;(x_j)_{j \neq i},t}}, \quad 0 < |t| \lesssim c^2 \tag{208}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L^\infty_{x_i;(x_j)_{j \neq i},t}} \lesssim c^{1+4/p} \langle |k|_\infty \rangle^{1/2} \|\square_k f\|_{L_t^{q'} L_x^{b'}}, \quad 0 < |t| \lesssim c^2. \tag{209}$$

Proof Denote

$$\mathcal{L}_k(f, \psi) = \int \left(\square_k \int \mathcal{W}_2(t - \tau) f(\tau) d\tau, \psi(t) \right) dt.$$

By duality and the maximal function estimate (169)

$$\begin{aligned} |\mathcal{L}_k(f, \psi)| &\leq \|\square_k f\|_{L_{x_1}^{q'} L_{\bar{x},t}^1} \sum_{|l|_\infty \leq 1} \left\| \square_{k+l} \int \mathcal{W}_2(t - \tau) \psi(\tau) d\tau \right\|_{L_{x_1}^q L_{\bar{x},t}^\infty} \\ &\leq \|\square_k f\|_{L_{x_1}^{q'} L_{\bar{x},t}^1} \sum_{|l|_\infty \leq 1} \int \|\square_{k+l} \mathcal{W}_2(t - \tau) \psi(\tau) d\tau\|_{L_{x_1}^q L_{\bar{x},t}^\infty} \\ &\stackrel{(169)}{\leq} c^{d/2} \langle k \rangle^{1/q} \|\square_k f\|_{L_{x_1}^{q'} L_{\bar{x},t}^1} \|\psi\|_{L_t^1 L_x^2}, \end{aligned}$$

so by duality we obtain

$$\left\| \square_k \int \mathcal{W}_2(t - \tau) f(\tau) d\tau \right\|_{L_t^\infty L_x^2} \lesssim c^{d/2} \langle k \rangle^{1/q} \|\square_k f\|_{L_{x_1}^{q'} L_{\bar{x},t}^1}. \tag{210}$$

Therefore, by duality, Strichartz estimates (166) and (210)

$$\begin{aligned} |\mathcal{L}_k(f, \psi)| &\leq \left\| \square_k \int \mathcal{W}_2(-\tau) f(\tau) d\tau \right\|_{L_x^2} \left\| \square_k \int \mathcal{W}_2(-t) \psi(t) dt \right\|_{L_x^2} \\ &\lesssim c^{d/2} \langle k \rangle^{1/q} \|f\|_{L_{x_1}^{q'} L_{\bar{x},t}^1} c^{(1-1/r)2r/a} \|\square_k \psi\|_{L_t^{a'} L_x^{b'}}, \end{aligned} \tag{211}$$

which implies (207) for $q > 2$ or $a > 2$. In the case $a = q = 2$, (207) can be directly deduced from (169). Furthermore, by (168), (157) and (159) we get

$$\mathcal{L}_k(\partial_{x_1}^2 f, \psi) \lesssim c^{1+4/p} \langle |k|_\infty \rangle^{1/2} \|\square_k f\|_{L_{x_i:(x_j)_{j \neq i},t}^{1,2}} c^{4/p} \|\psi\|_{L_t^{a'} L_x^{b'}}, \tag{212}$$

and we can deduce (208); by exchanging f and ψ , we get (209). □

We now summarize the results we will use in order to prove the local well-posedness of (112): We omit the proof, it follows from the results of the previous subsections, together with (206).

Proposition 14 *Let $d \geq 2$, $8/d \leq p < +\infty$, $2 \leq q < +\infty$, $q > 8/d$, $k \in \mathbb{Z}^d$ with $|k|_\infty = |k_i| \gtrsim c$, $h, i, l \in \{1, \dots, d\}$. Then*

$$\left\| \square_k D_{x_i}^{3/2} \mathcal{W}_2(t) \psi_0 \right\|_{L_{x_i:(x_j)_{j \neq i},t}^{\infty,2}} \lesssim c \|\square_k \psi_0\|_{L^2}, \tag{213}$$

$$\|\square_k \mathcal{W}_2(t) \psi_0\|_{L_{x_i:(x_j)_{j \neq i},t}^{q,\infty}} \lesssim c^{d/2} \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \quad 0 < |t| \lesssim c^2, \tag{214}$$

$$\|\square_k \mathcal{W}_2(t) \phi_0\|_{L_t^\infty L_x^2 \cap L_x^{2+p}} \lesssim c^{\frac{4}{p(p+2)}} \|\square_k \phi_0\|_{L^2}, \quad 0 < |t| \lesssim c^2 \tag{215}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_{x_i:(x_j)_{j \neq i},t}^{\infty,2}} \lesssim \|\square_k f\|_{L_{x_i:(x_j)_{j \neq i},t}^{1,2}}, \tag{216}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_{x_h:(x_j)_{j \neq h},t}^{q,\infty}} \lesssim c^{1+d/2} \langle k_i \rangle^{1/2+1/q} \|\square_k f\|_{L_{x_i:(x_j)_{j \neq i},t}^{1,2}}, \quad 0 < |t| \lesssim c^2, \tag{217}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_{l,t}^\infty L_{\bar{x}}^2 \cap L_{l,x}^{2+p}} \lesssim c^{1+\frac{4}{p(p+2)}} \langle k_i \rangle^{1/2} \|\square_k f\|_{L_{x_i:(x_j)_{j \neq i},t}^{1,2}}, \quad 0 < |t| \lesssim c^2 \tag{218}$$

$$\|\square_k \partial_{x_i}^2 \mathcal{A}_2 f\|_{L_{x_i:(x_j)_{j \neq i},t}^{\infty,2}} \lesssim c^{1+\frac{4}{p(p+2)}} \langle k_i \rangle^{1/2} \|\square_k f\|_{L_{l,x}^{(2+p)/(1+p)}}, \quad 0 < |t| \lesssim c^2, \tag{219}$$

$$\|\square_k \partial_{x_l}^2 \mathcal{A}_2 f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{q,\infty}} \lesssim c^{\frac{d}{2} + \frac{2}{p+2} + \frac{4}{p(p+2)}} \langle k_i \rangle^{2+1/q} \|\square_k f\|_{L_{t,x}^{(2+p)/(1+p)}}, \quad 0 < |t| \lesssim c^2, \tag{220}$$

$$\|\square_k \mathcal{A}_2 f\|_{L_t^\infty L_x^2 \cap L_{t,x}^{2+p}} \lesssim c^{\frac{8}{p(p+2)}} \|\square_k f\|_{L_{t,x}^{(2+p)/(1+p)}}. \tag{221}$$

For the case $r > 2$ we have the following results

Remark 22 1. Let (a, b) be order- r admissible, $i \in \{1, \dots, d\}$, $q \geq 2$, $\frac{4r}{d} < q < +\infty$ and $k \in \mathbb{Z}^d$ with $|k|_\infty \gtrsim K(c)$, then

$$\|\square_k \partial_{x_i}^\alpha \mathcal{A}_r f\|_{L_{x_i;(x_j)_{j \neq i,t}}^{q,\infty}} \lesssim c^{d(1-\frac{1}{r}) + (1-\frac{1}{r})\frac{2r}{a}} \langle |k|_\infty \rangle^{\alpha+1/q} \|\square_k f\|_{L_t^{a'} L_x^{b'}}, \quad 0 < |t| \lesssim c^{2(r-1)}. \tag{222}$$

2. Let (a, b) be Schrödinger-admissible, $i \in \{1, \dots, d\}$, then

$$\|\square_k \partial_{x_i}^{2(r-1)} \mathcal{A}_r f\|_{L_t^a L_x^b} \lesssim c^{r-1+2r/a} \langle |k|_\infty \rangle^{r-3/2} \|\square_k f\|_{L_{x_i;(x_j)_{j \neq i,t}}^{1,2}}, \quad 0 < |t| \lesssim c^{2(r-1)} \tag{223}$$

$$\|\square_k \partial_{x_i}^{2(r-1)} \mathcal{A}_2 f\|_{L_{x_i;(x_j)_{j \neq i,t}}^{\infty,2}} \lesssim c^{r-1+2r/a} \langle |k|_\infty \rangle^{r-3/2} \|\square_k f\|_{L_t^{a'} L_x^{b'}}, \quad 0 < |t| \lesssim c^{2(r-1)}. \tag{224}$$

Proposition 15 Let $d \geq 2$, $4r/d \leq p < +\infty$, $2 \leq q < +\infty$, $q > 4r/d$, $k \in \mathbb{Z}^d$ with $|k|_\infty = |k_i| \gtrsim c$, $h, i, l \in \{1, \dots, d\}$. Then

$$\|\square_k D_{x_i}^{r-1/2} \mathcal{W}_r(t) \psi_0\|_{L_{x_i;(x_j)_{j \neq i,t}}^{\infty,2}} \lesssim c^{r-1} \|\square_k \psi_0\|_{L^2}, \tag{225}$$

$$\|\square_k \mathcal{W}_r(t) \psi_0\|_{L_{x_i;(x_j)_{j \neq i,t}}^{q,\infty}} \lesssim c^{d(1-\frac{1}{r})} \langle k \rangle^{1/q} \|\square_k \psi_0\|_{L^2}, \quad 0 < |t| \lesssim c^{2(r-1)}, \tag{226}$$

$$\|\square_k \mathcal{W}_r(t) \phi_0\|_{L_t^\infty L_x^2 \cap L_x^{2(r-1)+p}} \lesssim c^{\frac{4(r-1)^2}{p(p+2(r-1))}} \|\square_k \phi_0\|_{L^2(\mathbb{R}^d)}, \quad 0 < |t| \lesssim c^{2(r-1)} \tag{227}$$

$$\|\square_k \partial_{x_l}^{2(r-1)} \mathcal{A}_r f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{\infty,2}} \lesssim \|\square_k f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{1,2}}, \tag{228}$$

$$\|\square_k \partial_{x_l}^{2(r-1)} \mathcal{A}_r f\|_{L_{x_h;(x_j)_{j \neq h,t}}^{q,\infty}} \lesssim c^{r-1+d(1-\frac{1}{r})} \langle k_i \rangle^{r-3/2+1/q} \|\square_k f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{1,2}}, \tag{229}$$

$$0 < |t| \lesssim c^{2(r-1)},$$

$$\|\square_k \partial_{x_l}^{2(r-1)} \mathcal{A}_r f\|_{L_t^\infty L_x^2 \cap L_{t,x}^{2(r-1)+p}} \lesssim c^{r-1+\frac{4(r-1)^2}{p(p+2(r-1))}} \langle k_i \rangle^{r-3/2} \|\square_k f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{1,2}}, \tag{230}$$

$$0 < |t| \lesssim c^{2(r-1)}$$

$$\|\square_k \partial_{x_l}^{2(r-1)} \mathcal{A}_r f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{\infty,2}} \lesssim c^{r-1+\frac{4(r-1)^2}{p(p+2(r-1))}} \langle k_i \rangle^{r-3/2} \|\square_k f\|_{L_{t,x}^{\frac{2(r-1)+p}{2r-1+p}}}, \tag{231}$$

$$0 < |t| \lesssim c^{2(r-1)},$$

$$\|\square_k \partial_{x_l}^{2(r-1)} \mathcal{A}_r f\|_{L_{x_l;(x_j)_{j \neq l,t}}^{q,\infty}} \lesssim c^{\frac{d}{2} + \frac{2r}{p+2(r-1)} + \frac{4(r-1)^2}{p(p+2(r-1))}} \langle k_i \rangle^{2(r-1)+1/q} \|\square_k f\|_{L_{t,x}^{\frac{2(r-1)+p}{2r-1+p}}}, \tag{232}$$

$$0 < |t| \lesssim c^{2(r-1)},$$

$$\|\square_k \mathcal{A}_r f\|_{L_t^\infty L_x^2 \cap L_{t,x}^{2(r-1)+p}} \lesssim c^{\frac{8(r-1)^2}{p(p+2(r-1))}} \|\square_k f\|_{L_{t,x}^{\frac{2(r-1)+p}{2r-1+p}}}. \tag{233}$$

For convenience, we state some technical results related to nonlinear mapping estimates. For $i = 1, \dots, d$ and $N \in \mathbb{N}$ we set

$$\mathbb{B}_{i,1}^{(N)} := \left\{ (k^{(1)}, \dots, k^{(N)}) \in (\mathbb{Z}^d)^N : \max (|k_i^{(1)}|, \dots, |k_i^{(N)}|) \gtrsim c \right\},$$

$$\mathbb{B}_{i,2}^{(N)} := \left\{ (k^{(1)}, \dots, k^{(N)}) \in (\mathbb{Z}^d)^N : \max (|k_i^{(1)}|, \dots, |k_i^{(N)}|) \lesssim c \right\}.$$

Lemma 13 *Let $s \geq 0, N \geq 3, i \in \{1, \dots, d\}$, then*

$$\left\| \sum_{\mathbb{B}_{i,1}^{(N)}} \square_{k^{(1)}} \psi_1 \cdots \square_{k^{(N)}} \psi_N \right\|_{l_{\square,i,c}^{1,s}(L_{x_1,(x_j)_{j \neq 2},t}^{1,2})}} \lesssim \sum_{\alpha=1}^N \|\psi_\alpha\|_{\cap_{h=1}^d l_{\square,h,c}^{1,s}(L_{x_h,(x_j)_{j \neq h},t}^{\infty,2})}} \prod_{\substack{\beta=1,\dots,d \\ \beta \neq \alpha}} \|\psi_\beta\|_{\cap_{h=1}^d l_{\square,h}^1(L_{x_h,(x_j)_{j \neq h},t}^{N-1,\infty})}}. \tag{234}$$

Proof See proof of Lemma 3.1 in [50]. □

Lemma 14 *Let $N \geq 1$ and $i \in \{1, \dots, d\}$, and assume that $1 \leq p, q, p_1, q_1, \dots, p_N, q_N \leq +\infty$ satisfy*

$$\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_N}, \quad \frac{1}{q} = \frac{1}{q_1} + \cdots + \frac{1}{q_N},$$

then

$$\left\| \sum_{\mathbb{B}_{i,2}^{(N)}} \square_{k^{(1)}} \psi_1 \cdots \square_{k^{(N)}} \psi_N \right\|_{l_{\square,i,c}^1(L_x^q L_x^p)} \lesssim c^d N^d \sum_{\mathbb{B}_{i,2}^{(N)}} \prod_{i=1}^N \|\square_{k^{(i)}} \psi_i\|_{L_t^{q_i} L_x^{p_i}}. \tag{235}$$

Proof See proof of Lemma 3.3 in [50]. □

Lemma 15 *Let $s \geq 0, N \geq 1$ and $i \in \{1, \dots, d\}$, and assume that $1 \leq p, q, p_1, q_1, \dots, p_N, q_N \leq +\infty$ satisfy*

$$\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_N}, \quad \frac{1}{q} = \frac{1}{q_1} + \cdots + \frac{1}{q_N},$$

then

$$\|\psi_1 \cdots \psi_N\|_{l_{\square}^{1,s}(L_t^p L_x^q)} \lesssim N^d \prod_{i=1}^N \|\psi_i\|_{l_{\square}^{1,s}(L_t^{p_i} L_x^{q_i})}. \tag{236}$$

Proof See proof of Lemma 8.2 in [59]. □

Proof (Proposition 7, part (i), case $r = 2$) Since the nonlinearity contains terms of the form $(\partial_x^\alpha \psi)^\beta$ with $|\alpha| \leq 2, |\beta| \geq m + 1$, we introduce the space

$$D := \left\{ \psi \in \mathcal{S}' : \|\psi\|_D := \sum_{|\alpha| \leq 2} \sum_{l=1}^3 \sum_{i,j=1}^d \rho_l^{(i)} (\partial_{x_j}^\alpha \psi) \lesssim c^{-\delta_0} \right\},$$

where

$$\begin{aligned} \rho_1^{(i)}(\psi) &:= \|\psi\|_{L_{\square, i, c}^{1, s-r+1/2+1/m}(L_{x_i; (\alpha_j)_{j \neq i, t}}^{\infty, 2})}, \\ \rho_2^{(i)}(\psi) &:= \|\psi\|_{L_{\square}^{1, s}(L_{x_i; (\alpha_j)_{j \neq i, t}}^{m, \infty})}, \\ \rho_3^{(i)}(\psi) &:= \|\psi\|_{L_{\square}^{1, s+1/m}(L_t^{\infty} L_{t, x}^2 \cap L_{t, x}^{2+m})}. \end{aligned}$$

and for some $\delta_0 > 0$ that we will choose later.

Since $\|\psi\|_D = \|\bar{\psi}\|_D$, without loss of generality we can assume that the nonlinearity contain only terms of the form

$$\psi^{\beta_0} (\partial_x^{\alpha_1} \psi)^{\beta_1} (\partial_x^{\alpha_2} \psi)^{\beta_2} =: \Psi_1 \dots \Psi_R,$$

where $R := |\beta| = \beta_0 + |\beta_1| + |\beta_2|$, $|\alpha_i| = i$ ($i = 1, 2$).

To prove the first part of Proposition 7 we will show that the map

$$\begin{aligned} \mathcal{F} : D &\rightarrow D, \\ \psi(t) &\mapsto \mathcal{Z}_2(t)\psi_0 + i\mathcal{A}_2 P \left((\partial_x^{\alpha} \psi)_{|\alpha| \leq 2}, (\partial_x^{\alpha} \bar{\psi})_{|\alpha| \leq 2} \right) \end{aligned}$$

is a contraction mapping.

First, we have that by Proposition 14

$$\|\mathcal{Z}_2(t)\psi_0\|_D \lesssim c^{\frac{d}{2} + \frac{4}{m(m+2)}} \|\psi_0\|_{M_{2,1}^{s+3+1/m}}.$$

Now, for the estimate of $\rho_1^{(i)}(\mathcal{A}_2 \partial_{x_j}^{\alpha} F)$ ($i, j = 1, \dots, d$) it suffices to estimate $\rho_1^{(1)}(\mathcal{A}_2 \partial_{x_1}^{\alpha} F)$: Indeed, by (206)

$$\rho_1^{(1)}(\mathcal{A}_2 \partial_{x_2}^{\alpha} F) \lesssim \rho_1^{(1)}(\mathcal{A}_2 \partial_{x_1}^{\alpha} F).$$

Using frequency-uniform decomposition, we write

$$\square_k(\Psi_1 \dots \Psi_R) = \sum_{\mathbb{B}_{1,1}^{(R)}} \square_k(\square_{k(1)} \Psi_1 \dots \square_{k(R)} \Psi_R) + \sum_{\mathbb{B}_{1,2}^{(R)}} \square_k(\square_{k(1)} \Psi_1 \dots \square_{k(R)} \Psi_R).$$

By exploiting (216) and (234) for the first sum and (219) and (236) for the second sum we obtain

$$\begin{aligned} \rho_1^{(1)}(\mathcal{A}_2 \partial_{x_1}^{\alpha} (\Psi_1 \dots \Psi_R)) &\lesssim \left\| \sum_{\mathbb{B}_{1,1}^{(R)}} \square_{k(1)} \Psi_1 \dots \square_{k(R)} \Psi_R \right\|_{L_{\square, 1, c}^{1, s-r+1/2+1/m}(L_{x_1, (x_j)_{j \neq 1, t}}^{1, 2})} \\ &\quad + c^{1 + \frac{4}{R^2-1}} \left\| \sum_{\mathbb{B}_{1,1}^{(R)}} \square_{k(1)} \Psi_1 \dots \square_{k(R)} \Psi_R \right\|_{L_{\square, 1, c}^1(L_{t, x}^{\frac{R+1}{R}})} \\ &\lesssim c^{1 + \frac{4}{R^2-1} + d} R^d \|\psi\|_D^R. \end{aligned}$$

Next, we estimate $\rho_2^{(1)}(\mathcal{A}_2(\Psi_1 \cdots \Psi_R))$ and $\rho_3^{(1)}(\mathcal{A}_2(\Psi_1 \cdots \Psi_R))$. By (221) and (220) we have

$$\begin{aligned} \sum_{j=2}^3 \rho_j^{(1)}(\mathcal{A}_2(\Psi_1 \cdots \Psi_R)) &\lesssim c^{\frac{d}{2} + \frac{2}{m+2} + \frac{8}{m(m+2)}} \|\Psi_1 \cdots \Psi_R\|_{L^{1,s+1/m} \left(L_{t,x}^{\frac{2+m}{1+m}} \right)} \\ &\stackrel{(236)}{\lesssim} c^{\frac{d}{2} + \frac{2}{m+2} + \frac{8}{m(m+2)}} R^d \|\psi\|_D^R. \end{aligned}$$

Then we consider $\rho_2^{(1)}(\mathcal{A}_2 \partial_{x_1}^2(\Psi_1 \cdots \Psi_R))$: We have

$$\begin{aligned} \rho_2^{(1)}(\mathcal{A}_2 \partial_{x_1}^2(\Psi_1 \cdots \Psi_R)) &\lesssim \left(\sum_{\substack{k \in \mathbb{Z}^d \\ |k|_\infty \gtrsim c}} + \sum_{\substack{k \in \mathbb{Z}^d \\ |k|_\infty \lesssim c}} \right) \|\square_k \mathcal{A}_2 \partial_{x_1}^2(\Psi_1 \cdots \Psi_R)\|_{L_{x_1; (\alpha_j)_{j \neq 1, t}}^{m, \infty}} \\ &=: III + IV. \end{aligned}$$

Again by (220) and (236) we obtain

$$\begin{aligned} IV &\lesssim c^{\frac{d}{2} + \frac{2}{m+2} + \frac{4}{m(m+2)}} \|\Psi_1 \cdots \Psi_R\|_{L^{1,s+1/m} \left(L_{t,x}^{\frac{2+m}{1+m}} \right)} \\ &\lesssim c^{\frac{d}{2} + \frac{2}{m+2} + \frac{4}{m(m+2)}} R^d \|\psi\|_D^R. \end{aligned}$$

Furthermore, we have that

$$\begin{aligned} III &\lesssim \left(\sum_{k \in \mathbb{Z}_1^d} + \cdots + \sum_{k \in \mathbb{Z}_d^d} \right) \|\square_k \mathcal{A}_2 \partial_{x_1}^2(\Psi_1 \cdots \Psi_R)\|_{L_{x_1; (\alpha_j)_{j \neq 1, t}}^{m, \infty}} \\ &=: G_1(\psi) + \cdots G_d(\psi). \end{aligned}$$

Using the frequency-uniform decomposition, (217), (234) and (235) we have that

$$G_i(\psi) \lesssim c^{1+3\frac{d}{2}} R^d \|\psi\|_D^R, \quad i = 1, \dots, d,$$

therefore

$$III \lesssim c^{1+3\frac{d}{2}} R^d \|\psi\|_D^R,$$

Finally, we estimate $\rho_3^{(1)}(\mathcal{A}_2 \partial_{x_i}^2(\Psi_1 \cdots \Psi_R))$. It suffices to consider the case $i = 1$: By (168) and (157) we have

$$\|\square_k \mathcal{A}_2 \partial_{x_1}^2 f\|_{L_t^\infty L^2 \cap L_{t,x}^{2+m}} \lesssim c^{\frac{4}{m(m+2)}} \langle k_1 \rangle^2 \|\square_k f\|_{L_{t,x}^{\frac{2+m}{1+m}}},$$

and by (218) and (207) we obtain

$$\rho_3^{(1)}(\mathcal{A}_2 \partial_{x_1}^2(\Psi_1 \cdots \Psi_R)) \lesssim c^{1 + \frac{8}{m(m+2)} + \frac{d}{2}} R^d \|\psi\|_D^R.$$

Collecting all estimates, we have

$$\|\mathcal{F}(\psi)\|_D \lesssim c^{\frac{d}{2} + \frac{4}{m(m+2)}} \|\psi_0\|_{M_{2,1}^{s+3+1/m}} + c^{1 + \frac{3d}{2} + \frac{2}{m+2} + \frac{8}{m(m+2)}} \sum_{m+1 \leq R < M} c^{\frac{4}{R^2-1}} R^d \|\psi\|_D^R. \tag{237}$$

and for $c \geq 1$ sufficiently large we can conclude by a standard contraction mapping argument (see, for example, the proof of Theorem 1.1 in [16]), by choosing

$$\delta > \delta_0(d, m, 2) := \max \left(\frac{d}{2} + \frac{4}{m(m+2)}, \frac{1}{m} + \frac{3d}{2m} + \frac{2}{m(m+2)} + \frac{8}{m^2(m+2)} + \frac{4}{m^3} \right). \tag{238}$$

□

Remark 23 By arguing in the same way for the general case $r > 2$ we end up with the condition

$$\delta > \delta_0(d, m, r) := \max \left(d \left(1 - \frac{1}{r} \right) + \frac{4r}{m^2(m+2(r-1))}, \frac{r-1}{m} + \frac{3d}{2m} + \frac{2rm+8(r-1)^2}{m^2(m+2(r-1))} + \frac{4(r-1)^2}{m^3} \right). \tag{239}$$

Remark 24 The quantity $\delta_0(d, l, r)$ defined in Corollary 1 is actually the right-hand side of (240) with m replaced by $2(l-1)$.

In order to prove the second part of Proposition 7 we will exploit another contraction mapping argument, like in the proof of Theorem 1 in [28] (which in turn is based on the proof of Theorem 4.1 of [30]). In the following, we denote by a $(Q_\alpha)_{\alpha \in \mathbb{Z}^d}$ a fixed family of nonoverlapping cubes of size R such that $\mathbb{R}^d = \bigcup_\alpha Q_\alpha$.

Lemma 16 *Let $d \geq 2$ and $r \geq 2$, then the following estimates hold.*

– (Local smoothing, homogeneous case)

$$\sup_{\alpha \in \mathbb{Z}^d} \left(\int_{Q_\alpha} \int_{\mathbb{R}} |D_x^{r-1/2} \mathcal{U}_r(t) \psi_0(x)|^2 dt dx \right)^{1/2} \lesssim c^{r-1} R^{1/2} \|\psi_0\|_{L^2}, \tag{241}$$

$$\left\| D_x^r \int_I \mathcal{U}_r(t-\tau) \psi(\tau, \cdot) d\tau \right\|_2 \lesssim c^{r-1} R^{1/2} \sum_{\alpha \in \mathbb{Z}^d} \left(\int_{Q_\alpha} \int_I |\psi(t, x)|^2 dt dx \right)^{1/2}; \tag{242}$$

– (Local smoothing, inhomogeneous case) *the solution of the inhomogeneous Cauchy problem*

$$-i \psi_t = A_{c,r} \psi + f(t, x), \quad t \in I, x \in \mathbb{R}^d,$$

such that $\psi_0 \equiv 0$ satisfies

$$\sup_{\alpha \in \mathbb{Z}^d} \|D_x^{2(r-1)} \psi\|_{L_x^2(Q_\alpha); L_t^2(I)} \lesssim c^{2(r-1)} RT^{1/(4d)} \sum_{\alpha \in \mathbb{Z}^d} \|f\|_{L_x^2(Q_\alpha); L_t^2(I)} \tag{243}$$

– (Maximal function estimate) *For any $s > d + \frac{1}{2}$ we have*

$$\left(\int_{\mathbb{R}^d} \sup_{|t| \lesssim c^{2(r-1)}} |\mathcal{U}_r(t) \psi_0(x)|^2 dx \right)^{1/2} \lesssim c^{d(1-\frac{1}{r})} \|\psi_0\|_{H^s}. \tag{244}$$

Proof (sketch) The proof in the case $r = 2$ can be obtained simply by rescaling Lemma 3, Lemma 4, Lemma 5 and Lemma 6 of [28]. The proof in the case $r > 2$ can be obtained by considering the operator $\mathcal{U}_r(t)$ and $\mathcal{A}_r(t)$ instead of $\mathcal{U}_2(t)$ and $\mathcal{A}_2(t)$. □

Proof (Proposition 7, part (ii), case $r = 2$) We will prove the result only for $s = s_0$, since the general case follows from commutator estimates. For simplicity, we only deal with the case

$$P((\partial_x^\alpha \psi)_{|\alpha| \leq 2}, (\partial_x^\alpha \bar{\psi})_{|\alpha| \leq 2}) = \partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi.$$

More precisely, we fix a positive constant $\nu < 1/3$, and we define the space Z_I^δ of all function $\phi : I \times \mathbb{R}^d \rightarrow \mathbb{C}$ such that the following three conditions hold

$$\|\phi\|_{L^\infty(I)H^{s_0}} \leq c^{-\delta}, \tag{245}$$

$$\sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\partial_x^\beta \phi(t, x)|^2 dx dt \right)^{1/2} \leq T^\nu, \tag{246}$$

$$\left(\sup_{t \in I} \sup_{x \in Q_\alpha} |D_x^2 \phi(t, x)|^2 \right)^{1/2} \leq c^{-\delta}. \tag{247}$$

We want to show that the map

$$\begin{aligned} \mathcal{F} : Z_I^\delta &\rightarrow Z_I^\delta, \\ \psi(t) &\mapsto \mathcal{U}_2(t)\psi_0 + i\mathcal{A}_2 P((\partial_x^\alpha \psi)_{|\alpha| \leq 2}, (\partial_x^\alpha \bar{\psi})_{|\alpha| \leq 2}) \end{aligned}$$

is a contraction mapping.

We can observe that for any $\beta \in \mathbb{Z}^d$ with $|\beta| = s_0 - \frac{3}{2}$

$$\begin{aligned} \partial_x^\beta (\partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi) &= \partial_x^\beta \partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi + \partial_{x_j}^2 \psi \partial_x^\beta \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi + \partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_x^\beta \partial_{x_m}^2 \psi \\ &\quad + R\left((\partial_x^\gamma \psi)_{2 \leq |\gamma| \leq s_0 - 1/2}\right). \end{aligned}$$

Now, for any $\psi \in Z_I^\delta$ we have

$$\begin{aligned} &\sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\partial_x^\beta \psi(t, x)|^2 dx dt \right)^{1/2} \\ &\lesssim \sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\mathcal{U}_2(t) \partial_x^\beta \psi_0(x)|^2 dx dt \right)^{1/2} \\ &\quad + \sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} \left| \int_0^t \mathcal{U}_2(t-\tau) \partial_x^\beta (\partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi) d\tau \right|^2 dx dt \right)^{1/2} \\ &\stackrel{(241), (242)}{\lesssim} cT^{1/3} \|\psi_0\|_{H^{s_0}} + c^2 T^{1/(4d)} \sum_{|\beta|=s_0-3/2} \sum_{j, k, m=1}^d \sum_{\alpha \in \mathbb{Z}^d} \|\partial_x^\beta \partial_{x_j}^2 \psi \partial_{x_k}^2 \psi \partial_{x_m}^2 \psi\|_{L_x^2(Q_\alpha; L_t^2(I))} \\ &\quad + c^2 \int_0^T \|D_x^{1/2} R((\partial_x^\gamma \psi)_{2 \leq |\gamma| \leq s_0 - 1/2})\|_{L^2} dt \\ &\lesssim cT^{1/3} \|\psi_0\|_{H^{s_0}} + c^2 T^{1/(4d)} \\ &\quad \sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\partial_x^\beta \psi|^2 dx dt \right)^{1/2} \left(\sum_{\alpha \in \mathbb{Z}^d} \sup_{t \in I} \sup_{x \in Q_\alpha} |D_x^2 \psi|^2 \right) \\ &\quad + c^2 T \sup_{t \in I} \|\psi\|_{H^{s_0}}^3 \\ &\lesssim c^{1-\delta} T^{1/3} + c^2 T^{1/(4d)} T^\nu c^{-2\delta} + c^2 T c^{-3\delta} \leq T^\nu, \end{aligned} \tag{248}$$

where in the last inequality we have chosen $\delta \gg 1$ such that

$$c^{1-\delta}T^{-\nu+1/3} + c^{2(1-\delta)}T^{1/(4d)} + c^{2-3\delta}T^{1-\nu} \lesssim 1, \quad T = \mathcal{O}(c^{2(r-1)}). \tag{249}$$

Next, we have that for any $\psi \in Z_I^\delta$

$$\begin{aligned} \|\psi\|_{L^\infty(I)H^{s_0}} &\leq \|\psi_0\|_{H^{s_0}} + \sup_{t \in I} \int_0^t \|\mathcal{W}_2(t-\tau)d_{x_j}^2\psi(\tau)\partial_{x_k}^2\psi(\tau)\partial_{x_m}^2\psi(\tau)\|_{L^2}d\tau \\ &\quad + \sup_{t \in I} \left\| D_x^{3/2} \int_0^t \mathcal{W}_2(t-\tau)D_x^{s_0-3/2}d_{x_j}^2\psi(\tau)\partial_{x_k}^2\psi(\tau)\partial_{x_m}^2\psi(\tau)d\tau \right\|_{L^2} \\ &\stackrel{(242)}{\lesssim} \|\psi_0\|_{H^{s_0}} + T \sup_{t \in I} \|\partial_{x_j}^2\psi(t)\partial_{x_k}^2\psi(t)\partial_{x_m}^2\psi(t)\|_{L^2} \\ &\quad + c^{r-1} \sum_{\alpha \in \mathbb{Z}^d} \left(\int_{Q_\alpha} \int_I |D_x^{s_0-3/2}(\partial_{x_j}^2\psi(t)\partial_{x_k}^2\psi(t)\partial_{x_m}^2\psi(t))|^2 dt dx \right)^{1/2} \\ &\lesssim \|\psi_0\|_{H^{s_0}} + T \sup_{t \in I} \|\psi\|_{H^{\frac{d}{3}+2}}^3 \\ &\quad + c \sum_{j,k,m=1}^d \sum_{\alpha \in \mathbb{Z}^d} \left(\int_{Q_\alpha} \int_I |D_x^{s_0-3/2}\partial_{x_j}^2\psi(t)\partial_{x_k}^2\psi(t)\partial_{x_m}^2\psi(t)|^2 dt dx \right)^{1/2} \\ &\quad + c \sum_{\alpha \in \mathbb{Z}^d} \left(\int_{Q_\alpha} \int_I |R(D_x^\gamma\psi)_{2 \leq |\gamma| \leq s_0-1/2}|^2 dt dx \right)^{1/2} \\ &\lesssim \|\psi_0\|_{H^{s_0}} + T \|\psi\|_{L^\infty(I)H^{\frac{d}{3}+2}}^3 \\ &\quad + c \sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\partial_x^\beta\psi|^2 dx dt \right)^{1/2} \sum_{\alpha \in \mathbb{Z}^d} \sup_{t \in I} \sup_{x \in Q_\alpha} |D_x^2\psi|^2 \\ &\quad + cT^{1/2} \|\psi\|_{L^\infty(I)H^{s_0}}^3 \\ &\lesssim \|\psi_0\|_{H^{s_0}} + (T + cT^{1/2})c^{-3\delta} + cT^\nu c^{-2\delta} \\ &\lesssim c^{-\delta}, \end{aligned} \tag{250}$$

where in the last inequality we have chosen $\delta \gg 1$ such that

$$(T + cT^{1/2})c^{-3\delta} + T^\nu c^{1-2\delta} \lesssim \frac{1}{2}, \quad T = \mathcal{O}(c^{2(r-1)}). \tag{251}$$

Then, we have that for any $\psi \in Z_I^\delta$

$$\begin{aligned} &\left(\sum_{\alpha \in \mathbb{Z}^d} \sup_{t \in I} \sup_{x \in Q_\alpha} |D_x^2\psi(t,x)|^2 \right)^{1/2} \\ &\stackrel{T \lesssim c^{2(r-1)}, (244)}{\lesssim} \|\psi_0\|_{H^{s_0}} + c^{d(1-\frac{1}{r})}T \|\psi\|_{L^\infty(I)H^{s_0}}^3 \\ &\lesssim \|\psi_0\|_{H^{s_0}} + c^{d(1-\frac{1}{r})}T c^{-3\delta} \\ &\lesssim c^{-\delta}, \end{aligned} \tag{252}$$

where in the last inequality we have chosen $\delta \gg 1$ such that

$$c^{d(1-\frac{1}{r})-2\delta} T \lesssim 1, \quad T = \mathcal{O}(c^{2(r-1)}). \tag{253}$$

Finally, if for any $\phi \in Z_1^\delta$ we set $\Lambda_T(\phi)$ as the maximum between the three following quantities,

$$\begin{aligned} & \sum_{\alpha \in \mathbb{Z}^d} \sup_{t \in I} \sup_{x \in Q_\alpha} |D_x^2 \psi(t, x)|^2, \\ & \|\phi\|_{L^\infty(I)H^{s_0}}, \\ & c^{-\delta} T^{-\nu} \sum_{|\beta|=s_0+1/2} \sup_{\alpha \in \mathbb{Z}^d} \left(\int_I \int_{Q_\alpha} |\partial_x^\beta \phi(t, x)|^2 dx dt \right)^{1/2}, \end{aligned}$$

we can observe that for any $\phi_1, \phi_2 \in Z_1^\delta$

$$\Lambda_T(\mathcal{F}(\phi_1) - \mathcal{F}(\phi_2)) \leq K T^\nu c^{-2\delta} \Lambda_T(\phi_1 - \phi_2),$$

where K is a positive constant which does not depend on c . Hence if we choose $\delta \gg 1$ such that (251), (249), (253) and

$$K T^\nu c^{-2\delta} \leq \frac{1}{2} \tag{254}$$

hold true, we can conclude. □

9 Long-time approximation

Now we study the evolution of the error between the approximate solution ψ_a , namely the solution of (99), and the original solution ψ of (3) for long (that means, c -dependent) time intervals. First we prove a result for the linear case, then we consider the approximation of small radiation solutions in the nonlinear case, and finally we make some remarks about the approximation of standing waves and soliton solutions.

9.1 Linear case

Fix $r \geq 1$, and take $\psi_0 \in H^{k+k_0}$, where $k_0 > 0$ and $k \gg 1$ are the ones in Theorem 7.

Now, we want to estimate the space–time norm of the error $\delta = \psi - \psi_a$. In the linear case we can observe that δ satisfies

$$\dot{\delta} = ic \langle \nabla \rangle_c \delta + \frac{1}{c^{2r}} X_{\mathcal{F}^{(r)} \circ \mathcal{R}^{(r)}}(\psi_a(t), \bar{\psi}_a(t)). \tag{255}$$

Proof (Theorem 3) By applying the Strichartz estimate (14) (choose $p = +\infty, q = 2, r = +\infty, s = 2$), together with estimate (31) for the vector field of the remainder $\mathcal{R}^{(r)}$, estimate (32) for the canonical transformation $\mathcal{F}^{(r)}$, and estimate (108) (choose $p = q = 2$), we can deduce Theorem 3. □

9.2 The nonlinear case: radiation solutions

Now, assume that we want to recover the approach of Sect. 9.1 to approximate radiation solutions of the NLKG equation for long (c -dependent) timescales.

We pursue such a program by a perturbative argument, considering a small radiation solution $\psi_r = \eta_{rad,r}$ of the normalized system (98) that exists up to times of order $\mathcal{O}(c^{2(r-1)})$, $r > 1$.

As an application of Proposition 1, we consider the following case. Fix $r > 1$, let $\sigma > 0$ and let $\psi_r = \eta_{rad}$ be a radiation solution of (98), namely such that

$$\eta_{rad,0} := \eta_{rad}(0) \in H^{k+k_0+\sigma+d/2}(\mathbb{R}^d), \tag{256}$$

where $k_0 > 0$ and $k \gg 1$ are the ones in Theorem 7.

Let $\delta(t)$ be a solution of (103); then, by Duhamel formula

$$\delta(t) := \mathcal{U}(t, 0)\delta_0 = e^{itc\langle \nabla \rangle_c} \delta_0 + \int_0^t e^{i(t-s)c\langle \nabla \rangle_c} dP(\psi_a(s)) \mathcal{U}(s, 0)\delta_0 ds. \tag{257}$$

Now fix $T \lesssim c^{2(r-1)}$; we want to estimate the local-in-time norm in the space $L^\infty([0, T])H^k(\mathbb{R}^d)$ of the error $\delta(t)$.

By (13) we can estimate the first term. We can estimate the second term by (14): Hence for any (p, q) Schrödinger-admissible exponents

$$\begin{aligned} & \left\| \int_0^t e^{i(t-s)c\langle \nabla \rangle_c} dP(\psi_a(s))\delta(s) ds \right\|_{L_t^\infty([0, T])H_x^k} \\ & \lesssim c^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} \|\langle \nabla \rangle_c^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} dP(\psi_a(t))\delta(t)\|_{L_t^{p'}([0, T])W_x^{k, q'}} \\ & \lesssim c^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} \|\langle \nabla \rangle_c^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} dP(\eta_{rad}(c^2t))\delta(t)\|_{L_t^{p'}([0, T])W_x^{k, q'}} \\ & \quad + c^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}} \|\langle \nabla \rangle_c^{\frac{1}{p} - \frac{1}{q} + \frac{1}{2}} [dP(\psi_a(t)) - dP(\eta_{rad}(c^2t))]\delta(t)\|_{L_t^{p'}([0, T])W_x^{k, q'}} \\ & =: I_p + II_p, \end{aligned}$$

but recalling (100) one has that

$$I_p \lesssim \frac{|\lambda|}{2^{l-1/2}(2l)(2l-1)} c^{\frac{1}{q} - \frac{1}{p} + \frac{1}{2}} \left\| \langle \nabla \rangle_c^{\frac{1}{p} - \frac{1}{q} - \frac{1}{2}} \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad} + \bar{\eta}_{rad}) \right]^{2(l-1)} \delta(t) \right\|_{L_t^{p'}([0, T])W_x^{k, q'}}.$$

Now fix a real number $\rho \in]0, 1[$, and choose

$$p = 2 + \rho, \tag{258}$$

$$q = \frac{2dp}{dp - 4} = \frac{4d + 2d\rho}{2d + d\rho - 4} = 2 + \frac{8}{2d - 4 + d\rho}, \tag{259}$$

we get (since $\|(c/\langle \nabla \rangle_c)^{\frac{1}{q} - \frac{1}{p} - \frac{1}{2}}\|_{L^{q'} \rightarrow L^q} \leq 1$)

$$I_2 \leq \frac{|\lambda|}{2^{l-1/2}(2l)(2l-1)} \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)) \right]^{2(l-1)} \delta(t) \right\|_{L_t^{2 - \frac{\rho}{1+\rho}}([0, T])W_x^{k, q'}}.$$

Now, since by Hölder inequality

$$\begin{aligned} & \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)) \right]^{2(l-1)} \delta(t) \right\|_{L_t^{2-\frac{\rho}{1+\rho}}([0,T])W_x^{k,q'}} \\ & \leq \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)) \right]^{2(l-1)} \right\|_{L_t^{2-\frac{\rho}{1+\rho}}([0,T])W_x^{k,d(1+\rho/2)}} \|\delta(t)\|_{L_t^\infty([0,T])H_x^k}, \end{aligned}$$

and by Sobolev product theorem (recall that $l \geq 2$, and that $k \gg 1$) we can deduce that

$$\begin{aligned} & \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)) \right]^{2(l-1)} \right\|_{L_t^{2-\frac{\rho}{1+\rho}}([0,T])W_x^{k,d(1+\rho/2)}} \\ & \leq \left[\int_0^T \left\| \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} (\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)) \right] \right\|_{W_x^{k,d(1+\rho/2)}}^{2(l-1)(2-\frac{\rho}{1+\rho})} dt \right]^{\frac{1}{(2-\frac{\rho}{1+\rho})}} \\ & \leq \|\eta_{rad}(c^2t) + \bar{\eta}_{rad}(c^2t)\|_{L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])W_x^{k,d(1+\rho/2)}}^{2(l-1)}, \end{aligned}$$

but since by Proposition 6 we have that for any $\sigma > 0$

$$\begin{aligned} L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])W_x^{k,d(1+\rho/2)} & \supseteq L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])M_{d(1+\rho/2),1,x}^k \\ & \supseteq L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])M_{2,1,x}^k \\ & \supseteq L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])H_x^{k+\sigma+d/2} \\ & \supseteq L_t^\infty([0,T])H_x^{k+\sigma+d/2}, \end{aligned}$$

we have that

$$\|\eta_{rad}\|_{L_t^{2(l-1)(2-\frac{\rho}{1+\rho})}([0,T])W_x^{k,d(1+\rho/2)}}^{2(l-1)} \lesssim T^{\frac{1+\rho}{2+\rho}} \|\eta_{rad}\|_{L_t^\infty([0,T])H_x^{k+\sigma+d/2}}^{2(l-1)}, \tag{260}$$

but by Corollary 1 the right-hand side of (260) is finite and does not depend on $c \geq 1$ for

$$\|\eta_{rad,0}\|_{H_x^{k+k_0+\sigma+d/2}} \lesssim c^{-\alpha}, \tag{261}$$

$$\alpha > \max \left(\delta_0(d, l, r), \delta_1(d, l, r), \frac{r-1}{l-1} \right) := \alpha^*(d, l, r). \tag{262}$$

where $c \geq c_0$ is sufficiently large, and where $\delta_0(d, l, r)$ and $\delta_1(d, l, r)$ are defined in Corollary 1.

Furthermore, via (32) one can show that there exists $c_{r,k} > 0$ sufficiently large such that for $c \geq c_{r,k}$ the term II_2 can be bounded by $\frac{1}{c^2} I_2$.

This means that we can estimate the $L^\infty([0, T])H^k$ norm of the error only for a small (with respect to c) radiation solution, which is the statement of Proposition 16.

To summarize, we get the following result.

Proposition 16 Consider (58) on \mathbb{R}^d , $d \geq 2$. Let $r > 1$, and fix $k_1 \gg 1$. Assume that $l \geq 2$ and $r < \frac{d}{2}(l-1)$. Then $\exists k_0 = k_0(r) > 0$ such that for any $k \geq k_1$ and for any $\sigma > 0$ the following

holds: Consider the solution η_{rad} of (98) with the initial datum $\eta_{rad,0} \in H^{k+k_0+\sigma+d/2}(\mathbb{R}^d)$, and call δ the difference between the solution of the approximate equation (99) and the original solution of the Hamilton equation for (58). Assume that $\delta_0 := \delta(0)$ satisfies

$$\|\delta_0\|_{H_x^k} \lesssim \frac{1}{c^2}.$$

Then there exist $\alpha^* := \alpha^*(d, l, r) > 0$ and there exists $c^* := c^*(r, k) > 1$, such that for any $\alpha > \alpha^*$ and for any $c > c^*$, if $\eta_{rad,0}$ satisfies

$$\|\eta_{rad,0}\|_{H^{k+k_0+\sigma+d/2}} \lesssim c^{-\alpha},$$

then

$$\sup_{t \in [0, T]} \|\delta(t)\|_{H_x^k} \lesssim \frac{1}{c^2}, \quad T \lesssim c^{2(r-1)}.$$

By exploiting (32) and Proposition 16, we obtain Theorem 4.

9.3 The nonlinear case: standing waves solutions

Now we consider the approximation of another important type of solutions, the so-called standing waves solutions.

The issue of (in)stability of standing waves and solitons has a long history: For the NLS equation and the NLKG the orbital stability of standing waves has been discussed first in [53]; for the NLS, the orbital stability of one soliton solutions has been treated in [26], while the asymptotic stability has been discussed in [20] for one soliton solutions, and in [48] and [47] for N-solitons. For the higher-order Schrödinger equation we mention [37], which deals with orbital stability of standing waves for fourth-order NLS-type equations. For the NLKG equation, the instability of solitons and standing waves has been studied in [29,45,54].

As for the case of radiation solution, we should fix $r \geq 1$, and consider a standing wave solution ψ_r of (98), namely of the form

$$\psi_r(t, x) = e^{it\omega} \eta_\omega(x), \tag{263}$$

where $\omega \in \mathbb{R}$, and $\eta_\omega \in \mathcal{S}(\mathbb{R}^d)$ solves

$$-\omega \eta_\omega = X_{H_{simp}}(\eta_\omega).$$

Remark 25 Of course the existence of a standing wave for the simplified equation (98) is a far from trivial question. For $r = 1$ [26] deals with the NLS equation; for $r \geq 2$, [37] deals with the one-dimensional fourth-order NLS-type equation.

We also point out that in the case of a standing wave solution, if $\delta(t)$ satisfies (103), then by Duhamel formula

$$\dot{\delta} = ic \langle \nabla \rangle_c \delta(t) + dP(\psi_a(t), \bar{\psi}_a(t))\delta(t).$$

Since

$$P(e^{it\omega} \eta_\omega, e^{-it\omega} \bar{\eta}_\omega) = 2^{l-1/2} \left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \text{Re}(e^{it\omega} \eta_\omega) \right]^{2l-1},$$

we have that

$$dP(\eta_\omega, \bar{\eta}_\omega)e^{it\omega}h = 2^{l-1/2} \left(\frac{c}{\langle \nabla \rangle_c} \right) \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \cos(\omega t) \eta_\omega \right]^{2(l-1)} (e^{it\omega}h + e^{-it\omega}\bar{h}),$$

and by setting $\delta = e^{-it\omega}h$, one gets

$$-i\dot{h} = (c\langle \nabla \rangle_c + \omega)h + 2^{l-1/2} \cos^{2(l-1)}(\omega t) \left(\frac{c}{\langle \nabla \rangle_c} \right) \left[\left(\frac{c}{\langle \nabla \rangle_c} \right)^{1/2} \eta_\omega \right]^{2(l-1)} (h + e^{-2it\omega}\bar{h}) \tag{264}$$

$$+ [dP(\psi_a(s), \bar{\psi}_a(s)) - dP(\eta_\omega, \bar{\eta}_\omega)]h. \tag{265}$$

Equation (264) is a Salpeter spinless equation with a periodic time-dependent potential; therefore, in order to get some information about the error, one would need the corresponding Strichartz estimates for Eq. (264). Unfortunately, in the literature of dispersive estimates there are only few results for PDEs with time-dependent potentials, and the majority of them is of perturbative nature; for the Schrödinger equation, we mention [22,25], in which Strichartz estimates are proved in a nonperturbative framework.

Remark 26 By using Proposition 3 one can show that the NLKG can be approximated by the simplified equation (3) locally uniformly in time, up to an error of order $\mathcal{O}(c^{-2r})$.

Remark 27 One could ask whether one could get a similar result for more general (in particular, moving) soliton solution of (98). Apart from the issue of existence and stability for such solutions, one can check that, provided that a moving soliton solution for (98) exists, then the error $\delta(t)$ must solve a (264)-type equation, namely a spinless Salpeter equation with a time-dependent moving potential. Unfortunately, since Eq. (264), unlike KG, is not manifestly covariant, one cannot apparently reduce to an analogue equation, and once again one cannot justify the approximation over the $\mathcal{O}(1)$ timescale.

Acknowledgements The author would like to thank Professor Dario Bambusi, for introducing him to the problem and for many valuable discussions and suggestions.

A Proof of Lemma 4

In order to normalize system (33), we used an adaptation of Theorem 4.4 in [2]. The result is based on the method of Lie transform that we will recall in the following.

Let $k \geq k_1$ and $p \in (1, +\infty)$ be fixed.

Given an auxiliary function χ analytic on $W^{k,p}$, we consider the auxiliary differential equation

$$\dot{\psi} = i\nabla_{\bar{\psi}}\chi(\psi, \bar{\psi}) =: X_\chi(\psi, \bar{\psi}) \tag{266}$$

and denote by Φ_χ^t its time- t flow. A simple application of Cauchy inequality gives

Lemma 17 *Let χ and its symplectic gradient be analytic in $B_{k,p}(R)$. Fix $\delta < R$, and assume that*

$$\sup_{B_{k,p}(R-\delta)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \leq \delta.$$

Then, if we consider the time- t flow Φ_χ^t of X_χ we have that for $|t| \leq 1$

$$\sup_{B_{k,p}(R-\delta)} \|\Phi_\chi^t(\psi, \bar{\psi}) - (\psi, \bar{\psi})\|_{k,p} \leq \sup_{B_{k,p}(R-\delta)} \|X_\chi(\psi, \bar{\psi})\|_{k,p}.$$

Definition 1 The map $\Phi := \Phi_\chi^1$ will be called the *Lie transform* generated by χ .

Remark 28 Given G analytic on $W^{k,p}$, consider the differential equation

$$\dot{\psi} = X_G(\psi, \bar{\psi}), \tag{267}$$

where by X_G we denote the vector field of G . Now define

$$\Phi^*G(\phi, \bar{\phi}) := G \circ \Phi(\psi, \bar{\psi}).$$

In the new variables $(\phi, \bar{\phi})$ defined by $(\psi, \bar{\psi}) = \Phi(\phi, \bar{\phi})$ Eq.(267) is equivalent to

$$\dot{\phi} = X_{\Phi^*G}(\phi, \bar{\phi}). \tag{268}$$

Using the relation

$$\frac{d}{dt}(\Phi_\chi^t)^*G = (\Phi_\chi^t)^*\{\chi, G\},$$

we formally get

$$\Phi^*G = \sum_{l=0}^{\infty} G_l, \tag{269}$$

$$G_0 := G, \tag{270}$$

$$G_l := \frac{1}{l} \{\chi, G_{l-1}\}, \quad l \geq 1. \tag{271}$$

In order to estimate the terms appearing in (269) we exploit the following results

Lemma 18 Let $R > 0$, and assume that χ, G are analytic on $B_{k,p}(R)$.

Then, for any $d \in (0, R)$ we have that $\{\chi, G\}$ is analytic on $B_{k,p}(R - d)$, and

$$\sup_{B_{k,p}(R-d)} \|X_{\{\chi,G\}}(\psi, \bar{\psi})\|_{k,p} \lesssim \frac{2}{d}. \tag{272}$$

Lemma 19 Let $R > 0$, and assume that χ, G are analytic on $B_{k,p}(R)$. Let $l \geq 1$, and consider G_l as defined in (269); for any $d \in (0, R)$, we have that G_l is analytic on $B_{k,p}(R - d)$, and

$$\sup_{B_{k,p}(R-d)} \|X_{G_l}(\psi, \bar{\psi})\|_{k,p} \lesssim \left(\frac{2e}{d}\right)^l. \tag{273}$$

Proof Fix l , and denote $\delta := d/l$. We look for a sequence $C_m^{(l)}$ such that

$$\sup_{B_{k,p}(R-m\delta)} \|X_{G_m}(\psi, \bar{\psi})\|_{k,p} \lesssim C_m^{(l)}, \quad \forall m \leq l.$$

By (272) we can define the sequence

$$\begin{aligned}
 C_0^{(l)} &:= \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p}, \\
 C_m^{(l)} &= \frac{2}{\delta m} C_{m-1}^{(l)} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \\
 &= \frac{2l}{dm} C_{m-1}^{(l)} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p}.
 \end{aligned}$$

One has

$$C_l^{(l)} = \frac{1}{l!} \left(\frac{2l}{d} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \right)^l \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p},$$

and by using the inequality $l! < l!e^l$ we can conclude. □

Remark 29 Let $k \geq k_1, p \in (1, +\infty)$, and assume that χ, F are analytic on $B_{k,p}(R)$. Fix $d \in (0, R)$, and assume also that

$$\sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \leq d/3,$$

Then for $|t| \leq 1$

$$\sup_{B_{k,p}(R-d)} \|X_{(\Phi_\chi^t)^* F - F}(\psi, \bar{\psi})\|_{k,p} = \sup_{B_{k,p}(R-d)} \|X_{F \circ \Phi_\chi^t - F}(\psi, \bar{\psi})\|_{k,p} \tag{274}$$

$$\stackrel{(272)}{\leq} \frac{5}{d} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \sup_{B_{k,p}(R)} \|X_F(\psi, \bar{\psi})\|_{k,p}. \tag{275}$$

Lemma 20 Let $k \geq k_1, p \in (1, +\infty)$, and assume that G is analytic on $B_{k,p}(R)$, and that h_0 satisfies PER. Then there exists χ analytic on $B_{k,p}(R)$ and Z analytic on $B_{k,p}(R)$ with Z in normal form, namely $\{h_0, Z\} = 0$, such that

$$\{h_0, \chi\} + G = Z. \tag{276}$$

Furthermore, we have the following estimates on the vector fields

$$\sup_{B_{k,p}(R)} \|X_Z(\psi, \bar{\psi})\|_{k,p} \leq \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p}, \tag{277}$$

$$\sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \lesssim \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p}. \tag{278}$$

Proof One can check that the solution of (276) is

$$\chi(\psi, \bar{\psi}) = \frac{1}{T} \int_0^T t [G(\Phi^t(\psi, \bar{\psi})) - Z(\Phi^t(\psi, \bar{\psi}))] dt,$$

with $T = 2\pi$. Indeed,

$$\begin{aligned} \{h_0, \chi\}(\psi, \bar{\psi}) &= \frac{d}{ds} \Big|_{s=0} \chi(\Phi^s(\psi, \bar{\psi})) \\ &= \frac{1}{2\pi} \int_0^{2\pi} t \frac{d}{ds} \Big|_{s=0} [G(\Phi^{t+s}(\psi, \bar{\psi})) - Z(\Phi^{t+s}(\psi, \bar{\psi}))] dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} t \frac{d}{dt} [G(\Phi^t(\psi, \bar{\psi})) - Z(\Phi^t(\psi, \bar{\psi}))] dt \\ &= \frac{1}{2\pi} [tG(\Phi^t(\psi, \bar{\psi})) - tZ(\Phi^t(\psi, \bar{\psi}))]_{t=0}^{2\pi} \\ &\quad - \frac{1}{2\pi} \int_0^{2\pi} [G(\Phi^t(\psi, \bar{\psi})) - Z(\Phi^t(\psi, \bar{\psi}))] dt \\ &= G(\psi, \bar{\psi}) - Z(\psi, \bar{\psi}). \end{aligned}$$

Finally, (277) follows from the fact that

$$X_\chi(\psi, \bar{\psi}) = \frac{1}{T} \int_0^T t \Phi^{-t} \circ X_{G-Z}(\Phi^t(\psi, \bar{\psi})) dt$$

by applying property (25). □

Lemma 21 *Let $k \geq k_1$, $p \in (1, +\infty)$, and assume that G is analytic on $B_{k,p}(R)$, and that h_0 satisfies PER. Let χ be analytic on $B_{k,p}(R)$, and assume that it solves (276). For any $l \geq 1$ denote by $h_{0,l}$ the functions defined recursively as in (269) from h_0 . Then for any $d \in (0, R)$ one has that $h_{0,l}$ is analytic on $B_{k,p}(R - d)$, and*

$$\sup_{B_{k,p}(R-d)} \|X_{h_{0,l}}(\psi, \bar{\psi})\|_{k,p} \leq 2 \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p} \left(\frac{5}{d} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \right)^l. \tag{279}$$

Proof By using (276) one gets that $h_{0,1} = Z - G$ is analytic on $B_{k,p}(R)$. Then by exploiting (275) one gets the result. □

Lemma 22 *Let $k_1 \gg 1$, $p \in (1, +\infty)$, $R > 0$, $m \geq 0$, and consider the Hamiltonian*

$$H^{(m)}(\psi, \bar{\psi}) = h_0(\psi, \bar{\psi}) + \varepsilon \hat{h}(\psi, \bar{\psi}) + \varepsilon Z^{(m)}(\psi, \bar{\psi}) + \varepsilon^{m+1} F^{(m)}(\psi, \bar{\psi}). \tag{280}$$

Assume that h_0 satisfies PER and INV, that \hat{h} satisfies NF, and that

$$\begin{aligned} \sup_{B_{k,p}(R)} \|X_{\hat{h}}(\psi, \bar{\psi})\|_{k,p} &\leq F_0, \\ \sup_{B_{k,p}(R)} \|X_{F^{(0)}}(\psi, \bar{\psi})\|_{k,p} &\leq F. \end{aligned}$$

Fix $\delta < R/(m + 1)$, and assume also that $Z^{(m)}$ are analytic on $B_{k,p}(R - m\delta)$, and that

$$\begin{aligned} \sup_{B_{k,p}(R-m\delta)} \|X_{Z^{(0)}}(\psi, \bar{\psi})\|_{k,p} &= 0, \\ \sup_{B_{k,p}(R-m\delta)} \|X_{Z^{(m)}}(\psi, \bar{\psi})\|_{k,p} &\leq F \sum_{i=0}^{m-1} \varepsilon^i K_s^i, \quad m \geq 1, \\ \sup_{B_{k,p}(R-m\delta)} \|X_{F^{(m)}}(\psi, \bar{\psi})\|_{k,p} &\leq F K_s^m, \quad m \geq 1, \end{aligned} \tag{281}$$

with $K_s := \frac{2\pi}{\delta}(18F + 5F_0)$.

Then, if $\varepsilon K_s < 1/2$ there exists a canonical transformation $\mathcal{T}_\varepsilon^{(m)}$ analytic on $B_{k,p}(R - (m + 1)\delta)$ such that

$$\sup_{B_{k,p}(R-m\delta)} \|\mathcal{T}_\varepsilon^{(m)}(\psi, \bar{\psi}) - (\psi, \bar{\psi})\|_{k,p} \leq 2\pi\varepsilon^{m+1}F, \tag{282}$$

$H^{(m+1)} := H^{(m)} \circ \mathcal{T}^{(m)}$ has the form (280) and satisfies (281) with m replaced by $m + 1$.

Proof The key point of the lemma is to look for $\mathcal{T}_\varepsilon^{(m)}$ as the time-one map of the Hamiltonian vector field of an analytic function $\varepsilon^{m+1}\chi_m$. Hence consider the differential equation

$$(\dot{\psi}, \dot{\bar{\psi}}) = X_{\varepsilon^{m+1}\chi_m}(\psi, \bar{\psi}); \tag{283}$$

by standard theory we have that, if $\|X_{\varepsilon^{m+1}\chi_m}\|_{B_{k,p}(R-m\delta)}$ is sufficiently small and $(\psi_0, \bar{\psi}_0) \in B_{k,p}(R - (m + 1)\delta)$, then the solution of (283) exists for $|t| \leq 1$. Therefore, we can define $\mathcal{T}_{m,\varepsilon}^t : B_{k,p}(R - (m + 1)\delta) \rightarrow B_{k,p}(R - m\delta)$, and in particular the corresponding time-one map $\mathcal{T}_\varepsilon^{(m)} := \mathcal{T}_{m,\varepsilon}^1$, which is an analytic canonical transformation, ε^{m+1} -close to the identity. We have

$$\begin{aligned} (\mathcal{T}_\varepsilon^{(m+1)})^*(h_0 + \varepsilon\hat{h} + \varepsilon Z^{(m)} + \varepsilon^{m+1}F^{(m)}) &= h_0 + \varepsilon\hat{h} + \varepsilon Z^{(m)} \\ &+ \varepsilon^{m+1}[\{\chi_m, h_0\} + F^{(m)}] \\ &+ (h_0 \circ \mathcal{T}^{(m+1)} - h_0 - \varepsilon^{m+1}\{\chi_m, h_0\}) + \varepsilon(\hat{h} \circ \mathcal{T}^{(m+1)} - \hat{h}) \end{aligned} \tag{284}$$

$$+ \varepsilon(Z^{(m)} \circ \mathcal{T}^{(m+1)} - Z^{(m)}) + \varepsilon^{m+1}(F^{(m)} \circ \mathcal{T}^{(m+1)} - F^{(m)}). \tag{285}$$

It is easy to see that the first three terms are already normalized, that the term in the second line is the nonnormalized part of order $m+1$ that will vanish through the choice of a suitable χ_m , and that the last lines contains all the terms of order higher than $m+1$.

Now we want to determine χ_m in order to solve the so-called homological equation

$$\{\chi_m, h_0\} + F^{(m)} = Z_{m+1},$$

with Z_{m+1} in normal form. The existence of χ_m and Z_{m+1} is ensured by Lemma 20, and by applying (277) and the inductive hypothesis we get

$$\sup_{B_{k,p}(R-m\delta)} \|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p} \leq 2\pi F, \tag{286}$$

$$\sup_{B_{k,p}(R-m\delta)} \|X_{Z_{m+1}}(\psi, \bar{\psi})\|_{k,p} \leq 2\pi F. \tag{287}$$

Now define $Z^{(m+1)} := Z^{(m)} + \varepsilon^m Z_{m+1}$, and notice that by Lemma 17 we can deduce the estimate of $X_{Z^{(m+1)}}$ on $B_{k,p}(R - (m + 1)\delta)$ and (282) at level $m + 1$. Next, set $\varepsilon^{m+2}F^{(m+1)} := (286) + (287)$. Then we can use (275) and (279), in order to get

$$\begin{aligned} & \sup_{B_{k,p}(R-(m+1)\delta)} \|X_{\varepsilon^{m+2}F(m+1)}(\psi, \bar{\psi})\|_{k,p} \tag{288} \\ & \leq \left(\frac{10}{\delta} \varepsilon^m K_s^m \varepsilon F + \frac{5}{\delta} \varepsilon F_0 + \frac{5}{\delta} \varepsilon F \sum_{i=0}^{m-1} \varepsilon^i K_s^i + \frac{5}{\delta} \varepsilon F \varepsilon^m K_s^m \right) \varepsilon^{m+1} \sup_{B_{k,p}(R-m\delta)} \|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p} \\ & = \varepsilon^{m+2} \left(\frac{10}{\delta} \varepsilon^m K_s^m F + \frac{5}{\delta} F_0 + \frac{5}{\delta} F \sum_{i=0}^{m-1} \varepsilon^i K_s^i + \frac{5}{\delta} F \varepsilon^m K_s^m \right) \sup_{B_{k,p}(R-m\delta)} \|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p}. \tag{289} \end{aligned}$$

If $m = 0$, then the third term is not present, and (289) reads

$$\sup_{B_{k,p}(R-\delta)} \|X_{\varepsilon^2 F(1)}(\psi, \bar{\psi})\|_{k,p} \leq \varepsilon^2 \left(\frac{15}{\delta} F + \frac{5}{\delta} F \right) 2\pi F < \varepsilon^2 K_s F.$$

If $m \geq 1$, we exploit the smallness condition $\varepsilon K_s < 1/2$, and (289) reads

$$\sup_{B_{k,p}(R-(m+1)\delta)} \|X_{\varepsilon^{m+2}F(m+1)}(\psi, \bar{\psi})\|_{k,p} < \left(\frac{18}{\delta} \varepsilon F + \frac{5}{\delta} \varepsilon F_0 \right) 2\pi \varepsilon F \varepsilon^m K_s^m = \varepsilon^{m+2} F K_s^{m+1}.$$

□

Now fix $R > 0$.

Proof (Lemma 4) The Hamiltonian (33) satisfies the assumptions of Lemma 22 with $m = 0$, $F_{N,r}$ in place of $F^{(0)}$ and $h_{N,r}$ in place of \hat{h} , $F = K_{k,p}^{(F,r)} r 2^{2Nr}$, $F_0 = K_{k,p}^{(h,r)} r 2^{2Nr}$. (For simplicity we will continue to denote by F and F_0 the last two quantities.) So we apply Lemma 22 with $\delta = R/4$, provided that

$$\frac{8\pi}{R} (18F + 5F_0) \varepsilon < \frac{1}{2},$$

which is true due to (44). Hence there exists an analytic canonical transformation $\mathcal{T}_{\varepsilon,N}^{(1)} : B_{k,p}(3R/4) \rightarrow B_{k,p}(R)$ with

$$\sup_{B_{k,p}(3R/4)} \|\mathcal{T}_{\varepsilon,N}^{(1)}(\psi, \bar{\psi}) - (\psi, \bar{\psi})\|_{k,p} \leq 2\pi F \varepsilon,$$

such that

$$H_{N,r} \circ \mathcal{T}_{\varepsilon,N}^{(1)} = h_0 + \varepsilon h_{N,r} + \varepsilon Z_N^{(1)} + \varepsilon^2 \mathcal{R}_N^{(1)}, \tag{290}$$

$$Z_N^{(1)} := \langle F_{N,r} \rangle, \tag{291}$$

$$\begin{aligned} \varepsilon^2 \mathcal{R}_N^{(1)} & := \varepsilon^2 F^{(1)} \\ & = \left(h_0 \circ \mathcal{T}_{\varepsilon,N}^{(1)} - h_0 - \varepsilon \{ \chi_1, h_0 \} \right) + \varepsilon \left(\hat{h}_{N,r} \circ \mathcal{T}_{\varepsilon,N}^{(1)} - \hat{h}_{N,r} \right) + \varepsilon \left(Z_N^{(1)} \circ \mathcal{T}_{\varepsilon,N}^{(1)} - Z_N^{(1)} \right) \\ & \quad + \varepsilon^2 \left(F_{N,r} \circ \mathcal{T}_{\varepsilon,N}^{(1)} - F_{N,r} \right), \end{aligned} \tag{292}$$

$$\sup_{B_{k,p}(3R/4)} \|X_{h_{N,r} + Z_N^{(1)}}(\psi, \bar{\psi})\|_{k,p} \leq F_0 + F =: \tilde{\sigma} F_0, \tag{293}$$

$$\sup_{B_{k,p}(3R/4)} \|X_{\mathcal{R}_N^{(1)}}(\psi, \bar{\psi})\|_{k,p} \leq \frac{8\pi}{R} (18F + 5F_0) F =: \tilde{\sigma} F. \tag{294}$$

Again (290) satisfies the assumptions of Lemma 22 with $m = 0$, and $h_{N,r} + Z_N^{(1)}$ and $\mathcal{R}_N^{(1)}$ in place of $F^{(0)}$ and \hat{h} .

Now fix $\delta := \delta(R) = \frac{R}{4r}$, and apply r times Lemma 22; we get an Hamiltonian of the form (45), such that

$$\sup_{B_{k,p}(R/2)} \|X_{Z_N^{(r)}}(\psi, \tilde{\psi})\|_{k,p} \leq 2\tilde{\sigma} F, \tag{295}$$

$$\sup_{B_{k,p}(R/2)} \|X_{\mathcal{R}_N^{(r)}}(\psi, \tilde{\psi})\|_{k,p} \leq \tilde{\sigma} F. \tag{296}$$

□

B Interpolation theory for relativistic Sobolev spaces

In this section we show an analogue of Theorem 6.4.5 (7) in [11] for the relativistic Sobolev spaces $\mathcal{W}_c^{k,p}$, $k \in \mathbb{R}$, $1 < p < +\infty$. We recall that

$$\mathcal{W}_c^{k,p}(\mathbb{R}^3) := \left\{ u \in L^p : \|u\|_{\mathcal{W}_c^{k,p}} := \|c^{-k} \langle \nabla \rangle_c^k u\|_{L^p} < +\infty \right\}, \quad k \in \mathbb{R}, \quad 1 < p < +\infty.$$

In order to state the main result of this section, we exploit notations and well-known results coming from complex interpolation theory. (See [11] for a detailed introduction to this topic.)

In order to study the relativistic Sobolev spaces, we have to recall the notion of Fourier multipliers.

Definition 2 Let $1 < p < +\infty$, and $\rho \in \mathcal{S}'$. We call ρ a *Fourier multiplier* on $L^p(\mathbb{R}^d)$ if the convolution $(\mathcal{F}^{-1}\rho) * f \in L^p(\mathbb{R}^d)$ for all $f \in L^p(\mathbb{R}^d)$, and if

$$\sup_{\|f\|_{L^p}=1} \|(\mathcal{F}^{-1}\rho) * f\|_{L^p} < +\infty. \tag{297}$$

The linear space of all such ρ is denoted by M_p and is endowed with the above norm (297).

One can check that for any $p \in (1, +\infty)$ one has $M_p = M_{p'}$ (where $1/p + 1/p' = 1$), and that by Parseval’s formula $M_2 = L^\infty$. Furthermore, by Riesz–Thorin theorem one gets that for any $\rho \in M_{p_0} \cap M_{p_1}$ and for any $\theta \in (0, 1)$

$$\|\rho\|_{M_p} \leq \|\rho\|_{M_{p_0}}^{1-\theta} \|\rho\|_{M_{p_1}}^\theta, \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}. \tag{298}$$

In particular, one can deduce that $\|\cdot\|_{M_p}$ decreases with $p \in (1, 2]$, and that $M_p \subset M_q$ for any $1 < p < q \leq 2$.

More generally, if H_0 and H_1 are Hilbert spaces, one can introduce a similar definition of Fourier multiplier. We use the notation $\mathcal{S}'(H_0, H_1)$ in order to denote the space of all linear continuous maps from $\mathcal{S}(\mathbb{R}^d, H_0)$ to H_1 .

Definition 3 Let $1 < p < +\infty$, let H_0 and H_1 be two Hilbert spaces, and consider $\rho \in \mathcal{S}'(H_0, H_1)$. We call ρ a *Fourier multiplier* if the convolution $(\mathcal{F}^{-1}\rho) * f \in L^p(H_1)$ for all $f \in L^p(H_0)$, and if

$$\sup_{\|f\|_{L^p(H_0)}=1} \|(\mathcal{F}^{-1}\rho) * f\|_{L^p(H_1)} < +\infty. \tag{299}$$

The linear space of all such ρ is denoted by $M_p(H_0, H_1)$, and is endowed with the above norm (299).

Next we recall *Mihlin multiplier theorem* (Theorem 6.1.6 in [11]).

Theorem 9 *Let H_0 and H_1 be Hilbert spaces, and assume that $\rho : \mathbb{R}^d \rightarrow L(H_0, H_1)$ be such that*

$$|\xi|^\alpha \|D^\alpha \rho(\xi)\|_{L(H_0, H_1)} \leq K, \quad \forall \xi \in \mathbb{R}^d, |\alpha| \leq L$$

for some integer $L > d/2$. Then $\rho \in M_p(H_0, H_1)$ for any $1 < p < +\infty$, and

$$\|\rho\|_{M_p} \leq C_p K, \quad 1 < p < +\infty.$$

Now, recall the Littlewood–Paley functions $(\phi_j)_{j \geq 0}$ defined in (24), and introduce the maps $\mathcal{J} : \mathcal{S}' \rightarrow \mathcal{S}'$ and $\mathcal{P} : \mathcal{S}' \rightarrow \mathcal{S}'$ via formulas

$$(\mathcal{J} f)_j := \phi_j * f, \quad j \geq 0, \tag{300}$$

$$\mathcal{P} g := \sum_{j \geq 0} \tilde{\sigma} \phi_j * g_j, \quad j \geq 0, \tag{301}$$

where $g = (g_j)_{j \geq 0}$ with $g_j \in \mathcal{S}'$ for all j , and

$$\begin{aligned} \tilde{\sigma} \phi_0 &:= \phi_0 + \phi_1, \\ \tilde{\sigma} \phi_j &:= \phi_{j-1} + \phi_j + \phi_{j+1}, \quad j \geq 1. \end{aligned}$$

One can check that $\mathcal{P} \circ \mathcal{J} f = f \forall f \in \mathcal{S}'$, since $\tilde{\sigma} \phi_j * \phi_j = \phi_j$ for all j . We then introduce for $c \geq 1$ and $k \geq 0$ the space

$$l_c^{2,k} := \left\{ (z_j)_{j \in \mathbb{Z}} : c^{-k} \sum_{j \in \mathbb{Z}} (c^2 + |j|^2)^k |z_j|^2 < +\infty \right\}.$$

Theorem 10 *Let $c \geq 1, k \geq 0, 1 < p < +\infty$. Then $(\nabla)_c^k L^p$ is a retract of $L^p(l_c^{2,k})$, namely that the operators*

$$\begin{aligned} \mathcal{J} &: \mathcal{W}_c^{k,p} \rightarrow L^p(l_c^{2,k}) \\ \mathcal{P} &: L^p(l_c^{2,k}) \rightarrow \mathcal{W}_c^{k,p} \end{aligned}$$

satisfy $\mathcal{P} \circ \mathcal{J} = id$ on $\mathcal{W}_c^{k,p}$.

Proof First we show that $\mathcal{J} : \mathcal{W}_c^{k,p} \rightarrow L^p(l_c^{2,k})$ is bounded.

Since $\mathcal{J} f = (\mathcal{F}^{-1} \chi_c) * \mathcal{J}_c^k f$, where

$$\begin{aligned} (\chi_c(\xi))_j &:= (c^2 + |\xi|^2)^{-k/2} \hat{\phi}_j(\xi), \quad j \geq 0 \\ \mathcal{J}_c^k f &:= \mathcal{F}^{-1}((c^2 + |\xi|^2)^{k/2} \hat{f}), \end{aligned}$$

we have that for any $\alpha \in \mathbb{N}^d$

$$|\xi|^\alpha \|D^\alpha \chi_c(\xi)\|_{L(\mathbb{C}, l_c^{2,k})} \leq |\xi|^\alpha \sum_{j \geq 0} (2^{jk} c^k |D^\alpha (\chi_c(\xi))_j|) \leq K_\alpha$$

because the sum consists of at most two nonzero terms for each ξ . Thus $\mathcal{J} \in M_p(\mathcal{W}_c^{k,p}, L^p(l_c^{2,k}))$ by Mihlin multiplier theorem.

On the other hand, consider $\mathcal{P} : L^p(l_c^{2,k}) \rightarrow \mathcal{W}_c^{k,p}$.

Since $\mathcal{J}_c^k \circ \mathcal{P}g = (\mathcal{F}^{-1}\delta_c) * g^{(k)}$, where

$$\begin{aligned} g &= (g_j)_{j \geq 0}, \\ g^{(k)} &:= (2^{jk}g_j)_{j \geq 0}, \\ \delta_c(\xi)g &:= \sum_{j \geq 0} 2^{-jk}(c^2 + |\xi|^2)^{k/2} \tilde{\sigma} \phi_j(\xi)g_j, \end{aligned}$$

we have that for any $\alpha \in \mathbb{N}^d$

$$|\xi|^\alpha \|D^\alpha \delta_c(\xi)\|_{L(I_c^{2,k}, \mathbb{C})} \leq |\xi|^\alpha \left[\sum_{j \geq 0} (2^{-jk}c^{-k} |D^\alpha (c^2 + |\xi|^2)^{k/2} \tilde{\sigma} \phi_j(\xi)|)^2 \right]^{1/2} \leq K_\alpha,$$

because the sum consists of at most four nonzero terms for each ξ . Thus $\mathcal{P} \in M_p(L^p(I_c^{2,k}), \mathcal{W}_c^{k,p})$ by Mihlin multiplier theorem, and we can conclude. \square

Corollary 2 *Let $\theta \in (0, 1)$, and assume that $k_0, k_1 \geq 0$ ($k_0 \neq k_1$) and $p_0, p_1 \in (1, +\infty)$ satisfy*

$$\begin{aligned} k &= (1 - \theta)k_0 + \theta k_1, \\ \frac{1}{p} &= \frac{1 - \theta}{p_0} + \frac{\theta}{p_1}. \end{aligned}$$

Then $(\mathcal{W}_c^{k_0,p}, \mathcal{W}_c^{k_1,p})_\theta = \mathcal{W}_c^{k,p}$.

The previous corollary, combined with the classical *3 lines theorem* (Lemma 1.1.2 in [11]), immediately leads us to the following Proposition.

Proposition 17 *Let $k_0 \neq k_1$, $1 < p < +\infty$, and assume that $T : \mathcal{W}_c^{k_0,p} \rightarrow \mathcal{W}_c^{k_0,p}$ has norm M_0 , and that $T : \mathcal{W}_c^{k_1,p} \rightarrow \mathcal{W}_c^{k_1,p}$ has norm M_1 . Then*

$$T : \mathcal{W}_c^{k,p} \rightarrow \mathcal{W}_c^{k,p}, \quad k = (1 - \theta)k_0 + \theta k_1,$$

with norm $M \leq M_0^{1-\theta} M_1^\theta$.

Now we conclude with the proof of Theorem 5.

Proof (Theorem 5) Estimates (21) clearly follow from Proposition 1 if we can prove that for any α and for any $q \in [2, 6]$

$$\|\langle \nabla \rangle_c^\alpha \mathcal{W}_\pm \langle \nabla \rangle_c^{-\alpha} \|_{L^q \rightarrow L^q} \lesssim 1, \tag{302}$$

$$\|\langle \nabla \rangle_c^\alpha \mathcal{Z}_\pm \langle \nabla \rangle_c^{-\alpha} \|_{L^q \rightarrow L^q} \lesssim 1. \tag{303}$$

Indeed in this case one would have

$$\|\langle \nabla \rangle_c^{1/q-1/p} e^{it\mathcal{H}_c(x)} P_c(-\Delta + V)\psi_0\|_{L_t^p L_x^q} = \|\langle \nabla \rangle_c^{1/q-1/p} \mathcal{W}_\pm e^{it\langle \nabla \rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_t^p L_x^q},$$

but

$$\|\langle \nabla \rangle_c^{1/q-1/p} \mathcal{W}_\pm e^{it\langle \nabla \rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_x^q} \lesssim \|\langle \nabla \rangle_c^{1/q-1/p} e^{it\langle \nabla \rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_x^q},$$

hence

$$\begin{aligned} \|\langle \nabla \rangle_c^{1/q-1/p} e^{it\mathcal{H}(x)} P_c(-\Delta + V)\psi_0\|_{L_t^p L_x^q} &\lesssim c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \|\langle \nabla \rangle_c^{1/2} \mathcal{Z}_{\pm}\psi_0\|_{L^2} \\ &\lesssim c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \|\langle \nabla \rangle_c^{1/2} \psi_0\|_{L^2}. \end{aligned}$$

To prove (303) we first show that it holds for $\alpha = 2k, k \in \mathbb{N}$. We argue by induction. The case $k = 0$ is true by Theorem 6. Now, suppose that (303) holds for $\alpha = 2(k - 1)$, then

$$\begin{aligned} &\|(c^2 - \Delta)^k \mathcal{Z}_{\pm}(c^2 - \Delta)^{-k}\|_{L^q \rightarrow L^q} \\ &= \|(c^2 - \Delta)(c^2 - \Delta)^{k-1} \mathcal{Z}_{\pm}(c^2 - \Delta)^{-(k-1)}(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} \\ &\leq c^2 \|(c^2 - \Delta)^{k-1} \mathcal{Z}_{\pm}(c^2 - \Delta)^{-(k-1)}(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} \\ &\quad + \|\Delta(c^2 - \Delta)^{k-1} \mathcal{Z}_{\pm}(c^2 - \Delta)^{-(k-1)}(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} \\ &\leq c^2 \|(c^2 - \Delta)^{k-1} \mathcal{Z}_{\pm}(c^2 - \Delta)^{-(k-1)}(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} \\ &\quad + \|\Delta(c^2 - \Delta)^{-1}(c^2 - \Delta)^{k-1} \mathcal{Z}_{\pm}(c^2 - \Delta)^{-(k-1)}\|_{L^q \rightarrow L^q} \\ &\quad + \|\Delta(c^2 - \Delta)^{k-1} [\mathcal{Z}_{\pm}, (c^2 - \Delta)^{-1}](c^2 - \Delta)^{-(k-1)}\|_{L^q \rightarrow L^q} \\ &\lesssim c^2 \|(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} + \|\Delta(c^2 - \Delta)^{-1}\|_{L^q \rightarrow L^q} \lesssim 1, \end{aligned}$$

since

$$\|[\mathcal{Z}_{\pm}, (c^2 - \Delta)^{-1}]\|_{L^2 \rightarrow L^2} \lesssim \frac{|\xi|}{(c^2 + |\xi|^2)^2} \leq (c^2 + |\xi|^2)^{-3/2}.$$

Similarly we can show (303) for $\alpha = -2k, k \in \mathbb{N}$. By Proposition 17 one can extend the result to any $\alpha \in \mathbb{R}$ via interpolation theory. □

References

1. Alazard, T., Carles, R.: Semi-classical limit of Schrödinger–Poisson equations in space dimension $n \geq 3$. *J. Differ. Equ.* **233**(1), 241–275 (2007)
2. Bambusi, D.: Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Mathematische Zeitschrift* **230**(2), 345–387 (1999)
3. Bambusi, D.: Galerkin averaging method and Poincaré normal form for some quasilinear PDEs. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* **4**(4), 669–702 (2005)
4. Bambusi, D., Carati, A., Ponno, A.: The nonlinear Schrödinger equation as a resonant normal form. *Discrete Contin. Dyn. Syst. Ser. B* **2**(1), 109–128 (2002)
5. Bambusi, D., Cuccagna, S.: On dispersion of small energy solutions of the nonlinear Klein Gordon equation with a potential. *Am. J. Math.* **133**(5), 1421–1468 (2011)
6. Bambusi, D., Ponno, A.: On metastability in FPU. *Commun. Math. Phys.* **264**(2), 539–561 (2006)
7. Bao, W., Dong, X.: Analysis and comparison of numerical methods for the Klein–Gordon equation in the nonrelativistic limit regime. *Numerische Mathematik* **120**(2), 189–229 (2012)
8. Bao, W., Zhao, X.: A uniformly accurate (UA) multiscale time integrator Fourier pseudospectral method for the Klein–Gordon–Schrödinger equations in the nonrelativistic limit regime. *Numerische Mathematik* **135**, 833–873 (2016)
9. Baumstark, S., Faou, E., Schratz, K.: Uniformly accurate exponential-type integrators for Klein–Gordon equations with asymptotic convergence to the classical NLS splitting. *Math. Comput.* **87**, 1227–1254 (2017)
10. Bechouche, P., Mauser, N.J., Selberg, S.: Nonrelativistic limit of Klein–Gordon–Maxwell to Schrödinger–Poisson. *Am. J. Math.* **126**(1), 31–64 (2004)
11. Bergh, J., Lofstrom, J.: *Interpolation Spaces: An Introduction*. Springer, Berlin (1976)
12. Bouclet, J.M.: Littlewood–Paley decompositions on manifolds with ends. *Bull. Soc. Math. Fr.* **138**(1), 1–37 (2010)

13. Burq, N., Gérard, P., Tzvetkov, N.: Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds. *Am. J. Math.* **126**(3), 569–605 (2004)
14. Carles, R., Lucha, W., Moulay, E.: Higher-order Schrödinger and Hartree–Fock equations. *J. Math. Phys.* **56**(123), 122,301 (2015)
15. Carles, R., Moulay, E.: Higher order Schrödinger equations. *J. Phys. A Math. Theor.* **45**(39), 395,304 (2012)
16. Cazenave, T., Weissler, F.B.: The Cauchy problem for the critical nonlinear Schrödinger equation in H^s . *Nonlinear Anal. Theory Methods Appl.* **14**(10), 807–836 (1990)
17. Cho, Y., Ozawa, T.: On the semirelativistic Hartree-type equation. *SIAM J. Math. Anal.* **38**(4), 1060–1074 (2006)
18. Choi, W., Seok, J.: Nonrelativistic limit of standing waves for pseudo-relativistic nonlinear Schrödinger equations. *J. Math. Phys.* **57**(2), 021,510 (2016)
19. Cordero, E., Zucco, D.: Strichartz estimates for the vibrating plate equation. *J. Evol. Equ.* **11**(4), 827–845 (2011)
20. Cuccagna, S.: Stabilization of solutions to nonlinear Schrödinger equations. *Commun. Pure Appl. Math.* **54**(9), 1110–1145 (2001)
21. D’Ancona, P., Fanelli, L.: Strichartz and smoothing estimates for dispersive equations with magnetic potentials. *Commun. Partial Differ. Equ.* **33**(6), 1082–1112 (2008)
22. D’Ancona, P., Pierfelice, V., Visciglia, N.: Some remarks on the Schrödinger equation with a potential in $L^r_t L^s_x$. *Mathematische Annalen* **333**(2), 271–290 (2005)
23. Faou, E., Schratz, K.: Asymptotic preserving schemes for the Klein–Gordon equation in the non-relativistic limit regime. *Numerische Mathematik* **126**(3), 441–469 (2014)
24. Fujiwara, M.: Über die obere Schranke des absoluten Betrages der Wurzeln einer algebraischen Gleichung. *Tohoku Math. J. First Ser.* **10**, 167–171 (1916)
25. Goldberg, M.: Strichartz estimates for the Schrödinger equation with time-periodic $L^{n/2}$ potentials. *J. Funct. Anal.* **256**(3), 718–746 (2009)
26. Grillakis, M., Shatah, J., Strauss, W.: Stability theory of solitary waves in the presence of symmetry. I. *J. Funct. Anal.* **74**(1), 160–197 (1987)
27. Han-Kwan, D., Nguyen, T.T., Rousset, F.: Long time estimates for the Vlasov–Maxwell system in the non-relativistic limit. *Commun. Math. Phys.* **363**(2), 389–434 (2018)
28. Hao, C., Hsiao, L., Wang, B.: Well-posedness of Cauchy problem for the fourth order nonlinear Schrödinger equations in multi-dimensional spaces. *J. Math. Anal. Appl.* **328**(1), 58–83 (2007)
29. Imaikin, V., Komech, A., Vainberg, B.: On scattering of solitons for the Klein–Gordon equation coupled to a particle. *Commun. Math. Phys.* **268**(2), 321–367 (2006)
30. Kenig, C.E., Ponce, G., Vega, L.: Small solutions to nonlinear Schrödinger equations. *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis* **10**(3), 255–288 (1993)
31. Kim, J., Arnold, A., Yao, X.: Global estimates of fundamental solutions for higher-order Schrödinger equations. *Monatshefte für Mathematik* **168**(2), 253–266 (2012)
32. Kobayashi, M., Sugimoto, M.: The inclusion relation between Sobolev and modulation spaces. *J. Funct. Anal.* **260**(11), 3189–3208 (2011)
33. Lämmerzahl, C.: The pseudodifferential operator square root of the Klein–Gordon equation. *J. Math. Phys.* **34**(9), 3918–3932 (1993)
34. Lu, Y., Zhang, Z.: Partially strong transparency conditions and a singular localization method in geometric optics. *Arch. Ration. Mech. Anal.* **222**(1), 245–283 (2016)
35. Machihara, S.: The nonrelativistic limit of the nonlinear Klein–Gordon equation. *Funkcialaj Ekvacioj-Serio Internacia* **44**(2), 243–252 (2001)
36. Machihara, S., Nakanishi, K., Ozawa, T.: Nonrelativistic limit in the energy space for nonlinear Klein–Gordon equations. *Mathematische Annalen* **322**(3), 603–621 (2002)
37. Maeda, M., Segata, J.: Existence and stability of standing waves of fourth order nonlinear Schrödinger type equation related to vortex filament. *Funkcialaj Ekvacioj* **54**(1), 1–14 (2011)
38. Masmoudi, N., Nakanishi, K.: From nonlinear Klein–Gordon equation to a system of coupled nonlinear Schrödinger equations. *Mathematische Annalen* **324**(2), 359–389 (2002)
39. Masmoudi, N., Nakanishi, K.: Nonrelativistic limit from Maxwell–Klein–Gordon and Maxwell–Dirac to Poisson–Schrödinger. *Int. Math. Res. Not.* **2003**(13), 697–734 (2003)
40. Masmoudi, N., Nakanishi, K.: Energy convergence for singular limits of Zakharov type systems. *Inventiones mathematicae* **172**(3), 535–583 (2008)
41. Masmoudi, N., Nakanishi, K.: From the Klein–Gordon–Zakharov system to a singular nonlinear Schrödinger system. *Annales de l’Institut Henri Poincaré - Analyse Non Linéaire* **27**(4), 1073–1096 (2010)

42. Najman, B.: The nonrelativistic limit of the nonlinear Klein–Gordon equation. *Nonlinear Anal. Theory Methods Appl.* **15**(3), 217–228 (1990)
43. Nakanishi, K.: Nonrelativistic limit of scattering theory for nonlinear Klein–Gordon equations. *J. Differ. Equ.* **180**(2), 453–470 (2002)
44. Nakanishi, K., et al.: Transfer of global wellposedness from nonlinear Klein–Gordon equation to nonlinear Schrödinger equation. *Hokkaido Math. J.* **37**(4), 749–771 (2008)
45. Ohta, M., Todorova, G.: Strong instability of standing waves for the nonlinear Klein–Gordon equation and the Klein–Gordon–Zakharov system. *SIAM J. Math. Anal.* **38**(6), 1912–1931 (2007)
46. Robert, D., Tamura, H.: Semi-classical estimates for resolvents and asymptotics for total scattering cross-sections. *Annales de l’IHP Physique théorique* **46**(4), 415–442 (1987)
47. Rodnianski, I., Schlag, W., Soffer, A.: Asymptotic stability of N-soliton states of NLS (2003). arXiv preprint [arxiv:math/0309114](https://arxiv.org/abs/math/0309114)
48. Rodnianski, I., Schlag, W., Soffer, A.: Dispersive analysis of charge transfer models. *Commun. Pure Appl. Math.* **58**(2), 149–216 (2005)
49. Ruzhansky, M., Sugimoto, M., Wang, B.: Modulation spaces and nonlinear evolution equations. *Prog. Math.* **301**, 267–283 (2012)
50. Ruzhansky, M., Wang, B., Zhang, H.: Global well-posedness and scattering for the fourth order nonlinear Schrödinger equations with small data in modulation and Sobolev spaces. *Journal de Mathématiques Pures et Appliquées* **105**(1), 31–65 (2016)
51. Schneider, G.: Bounds for the nonlinear Schrödinger approximation of the Fermi–Pasta–Ulam system. *Appl. Anal.* **89**(9), 1523–1539 (2010)
52. Segata, J., et al.: Well-posedness for the fourth-order nonlinear Schrödinger-type equation related to the vortex filament. *Differ. Integral Equ.* **16**(7), 841–864 (2003)
53. Shatah, J., Strauss, W.: Instability of nonlinear bound states. *Commun. Math. Phys.* **100**(2), 173–190 (1985)
54. Soffer, A., Weinstein, M.I.: Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations. *Inventiones mathematicae* **136**(1), 9–74 (1999)
55. Sucher, J.: Relativistic invariance and the square-root Klein–Gordon equation. *J. Math. Phys.* **4**(1), 17–23 (1963)
56. Taylor, M.: *Partial Differential Equations III. Nonlinear Equations*. Applied Mathematical Sciences, vol. 117. Springer, New York (2011)
57. Tsutsumi, M.: Nonrelativistic approximation of nonlinear Klein–Gordon equations in two space dimensions. *Nonlinear Anal. Theory Methods Appl.* **8**(6), 637–643 (1984)
58. Wang, B., Han, L., Huang, C.: Global well-posedness and scattering for the derivative nonlinear Schrödinger equation with small rough data. *Annales de l’Institut Henri Poincaré - Analyse Non Linéaire* **26**(6), 2253–2281 (2009)
59. Wang, B., Hudzik, H.: The global Cauchy problem for the NLS and NLKG with small rough data. *J. Differ. Equ.* **232**(1), 36–73 (2007)
60. Yajima, K.: The $W^{k,p}$ -continuity of wave operators for Schrödinger operators. *J. Math. Soc. Jpn.* **47**(3), 551–581 (1995)