Higher order asymptotic analysis of the Klein-Gordon equation in the non-relativistic limit regime

Yong Lu* Zhifei Zhang†

Abstract

In this paper, we study the asymptotic behavior of nonlinear Klein-Gordon equations in the non-relativistic limit regime. By employing the techniques in geometric optics, we show that the Klein-Gordon equation can be approximated by nonlinear Schrödinger equations. In particular, we show error estimates which are of the same order as the initial error. Our result gives a mathematical verification for some numerical results obtained in [1, 2], and offers a rigorous justification for a technical assumption in the numerical studies [1].

Contents

1 Introduction

1.1 Setting

The Klein-Gordon equation is a relativistic version of the Schrödinger equation and is used to describe the motion of a spinless particle. The non-dimensional Klein-Gordon equation reads as

*Chern Institute of Mathematics & LPMC, Nankai University, Tianjin 300071, China, lvyong@amss.ac.cn.
†School of Mathematical Science, Peking University, Beijing 100871, China, zfzhang@math.pku.edu.cn.

follows

$$\varepsilon^2\partial_{tt}u - \Delta u + \frac{1}{\varepsilon^2}u + f(u) = 0, \quad t \geq 0, \quad x \in \mathbb{R}^d. \tag{1.1}$$

Here $u = u(t,x)$ is a real-valued (or complex-valued) field, and $f(u)$ is a real-valued function (or $f(u) = g(|u|^2)u$ if u is complex-valued). The non-dimensional parameter ε is proportional to the inverse of the speed of light.

For fixed ε, the well-posedness of the Klein-Gordon equation is well studied [5, 6]. In this paper, our concern is the asymptotic behavior of the solution in the non-relativistic limit ($\varepsilon \to 0$) with real initial data of the form

$$u(0) = u_{0,\varepsilon}, \quad (\partial_t u)(0) = \frac{1}{\varepsilon^2}u_{1,\varepsilon}. \tag{1.2}$$

1.2 Background and motivation

The non-relativistic limit of (1.1)-(1.2) has gained a lot interest both in analysis and in numerical computations, see [17, 21, 18, 12, 13, 15, 19, 1, 2] and references therein. In particular, for the complex valued unknown u and the typical polynomial nonlinearity $f(u) = \lambda|u|^q u$ with $0 \leq q < \frac{4}{d-2}$, Masmoudi and Nakanishi [15] showed that a wide class of solutions u to (1.1)-(1.2) can be described by using a system of coupled nonlinear Schrödinger equations. More precisely, for H^1 initial data of the form

$$u_{0,\varepsilon} = \varphi_0 + \varepsilon\varphi_\varepsilon, \quad u_{1,\varepsilon} = \psi_0 + \varepsilon\psi_\varepsilon, \tag{1.3}$$

it was shown in [15] that

$$u(t,x) = e^{it/\varepsilon^2}v_+ + e^{-it/\varepsilon^2}\overline{v}_- + R(t,x),$$

where $v = (v_+, v_-)$ satisfies

$$2iv_t - \Delta v + \widetilde{f}(v) = 0, \quad v(0) := (\varphi_0 - i\psi_0, \bar{\varphi}_0 + i\bar{\psi}_0), \tag{1.4}$$

with $\widetilde{f}(v) = \big(\widetilde{f}_+(v), \widetilde{f}_-(v)\big)$ defined by

$$\widetilde{f}_\pm(v) := \frac{1}{2\pi}\int_0^{2\pi} f(v_\pm + e^{i\theta}\overline{v}_\mp)d\theta.$$

The error term $R(t,x)$ satisfies the following estimate

$$\|R\|_{L^\infty(0,T;L^2)} = o(\varepsilon^{1/2}), \quad \text{for any } T < T^*, \tag{1.5}$$

where T^* is the maximal existence time of the coupled nonlinear Schrödinger equations (1.4).

Furthermore, if $f \in C^2$, for H^3 initial data of the form

$$u_{0,\varepsilon} = \varphi_0 + \varepsilon\varphi_1 + \varepsilon^2\varphi_{2,\varepsilon}, \quad u_{1,\varepsilon} = \psi_0 + \varepsilon\psi_1 + \varepsilon^2\psi_{2,\varepsilon}, \tag{1.6}$$

it was shown in [15] the following second order approximation result

$$\Big\|u - \big(e^{it/\varepsilon^2}(v_+ + \varepsilon w_+) + e^{-it/\varepsilon^2}(\overline{v_-} + \varepsilon\overline{w_-}\big)\Big\|_{L^\infty(0,T;H^1(\mathbb{R}^d))} = o(\varepsilon), \tag{1.7}$$

where $w = (w_+, w_-)$ is the solution to the following Cauchy problem of a linear Schröginder equation

$$2iw_t - \Delta w + D\widetilde{f}(v).w = 0, \quad w(0) := (\varphi_1 - i\psi_1, \bar{\varphi}_1 + i\bar{\psi}_1),$$

where we use the notation in [15]:

$$D\widetilde{f}(v).w := \partial_{z_1}\tilde{f}(v)w_+ + \partial_{z_2}\tilde{f}(v)w_- + \partial_{\bar{z}_1}\tilde{f}(v)\bar{w}_+ + \partial_{\bar{z}_2}\tilde{f}(v)\bar{w}_-.$$

In the previous studies in non-relativistic limit problems, the error estimates obtained are not optimal, in the sense that the error estimates over $O(1)$ time intervals are much amplified compared to the initial error estimates. For example, in [15], the $o(\varepsilon^{1/2})$ error estimate in (1.5) is obtained under $O(\varepsilon)$ initial error, and the $o(\varepsilon)$ error estimate (1.7) is obtained under $O(\varepsilon^2)$ initial error (see Theorem 1.3 and Theorem 1.4 in [15]).

The main goal of this paper is to obtain *optimal* error estimates: over $O(1)$ time interval, the error estimate is of the same order as the initial error estimate. As a result, we give better convergence rates for the error estimates compared to the result in [15]: $O(\varepsilon)$ in (1.5) and $O(\varepsilon^2)$ in (1.7), while the price to pay is to assume higher regularity for initial data. This is in particular highly important for numerical calculations. Our results also give uniform estimates for $\|\varepsilon^2\partial_t u\|_{H^\mu}$ for any $\mu > d/2$, which justifies the technical Assumption (A) and Theorem 4.1 in [1], which is the key to design a uniformly convergent numerical scheme (see Remark 2.6).

In [15], the authors study the problem mainly in energy spaces and the convergence results are obtained by using typical techniques in the study of dispersive equations, such as Strichartz estimates and Bourgain spaces. In this paper, we adopt a different point of view by treating the non-relativistic limit problem as the stability problem in the framework of *geometric optics* which stands for the study of highly oscillating solutions to hyperbolic systems. The main concern in mathematical geometric optics is the stability analysis of a family of approximate solutions named *WKB solutions*. We refer to [7] as well as [8, 10, 20, 9] for more details of basic concepts in geometric optics.

1.3 Statement of the results

We consider general nonlinearities $f(u)$ including nonlinearities of the form

$$f(u) = \lambda u^{q+1}, \quad q \geq 0,\ q \in \mathbb{Z}; \qquad f(u) = \lambda|u|^q u, \quad q \geq 0. \tag{1.8}$$

We point out that compared to [15], here we are working in more regular Sobolev spaces. Thus we need more regularity for f. But we do not need to control the growth of $f(u)$ with respect to u, thanks to the L^∞ norm of u by Sobolev embedding. This implies that we can handle the nonlinearities in (1.8) with q arbitrarily large, while in [15] it has to be assumed that $q < 4/(d-2)$.

We also point out that we obtain better convergence rates for the error estimates than the ones in [15]. For initial data in H^s, $s > d/2 + 4$ and nonlinearities $f \in C^m$, $m > s$, we improve the error in (1.5) from $o(\varepsilon^{1/2})$ to $O(\varepsilon)$. If f enjoys more regularity in C^m, $m > s+1$, we can improve the error in (1.7) from $o(\varepsilon)$ to $O(\varepsilon^2)$. The error estimates are obtained in the Sobolev space H^{s-4}. To prove such results, we employ the techniques in geometric optics.

The first result concerns the first order approximation. Our basic Schrödinger equation is

$$2iv_t - \Delta v + \tilde{f}(v) = 0, \quad v(0) = \frac{\varphi_0 - i\psi_0}{2}, \tag{1.9}$$

where

$$\tilde{f}(v) := \frac{1}{2\pi}\int_0^{2\pi} e^{-i\theta} f(e^{-i\theta}\bar{v} + e^{i\theta}v)d\theta. \tag{1.10}$$

By the classical theory for the local well-posedness of nonlinear Schrödinger equations (see for instance Chapter 8 of [16]), if $f \in C^m,\ m > s > d/2+4$, the Cauchy problem (1.9) admits a unique solution $v \in C\left([0, T_0^*); H^s\right)$ with $T_0^* > 0$ the maximal existence time.

Our first theorem states:

Theorem 1.1. *Suppose the real initial datum* (1.3) *satisfies the regularity assumption:*

$$\begin{aligned}&(\varphi_0, \psi_0) \in (H^s)^2 \quad \textit{independent of } \varepsilon,\\ &\left\{(\varphi_\varepsilon, \psi_\varepsilon, \varepsilon\nabla\varphi_\varepsilon)\right\}_{0<\varepsilon<0} \textit{ uniformly bounded in } (H^{s-4})^{d+2}\end{aligned} \tag{1.11}$$

with $s > d/2 + 4$ and the nonlinearity $f(\cdot) \in C^m,\ m > s,\ f(0) = 0$. Then the Cauchy problem (1.1)-(1.2) *admits a unique solution $u \in C\left([0, T_\varepsilon^*); H^{s-4}\right)$, where the maximal existence time $T_\varepsilon^* > 0$ satisfies*

$$\liminf_{\varepsilon\to 0} T_\varepsilon^* \geq T_0^*, \tag{1.12}$$

and for any $T < \min\{T_\varepsilon^*, T_0^*\}$, there exists a constant $C(T)$ independent of ε such that

$$\left\| u - \left(e^{it/\varepsilon^2} v + e^{-it/\varepsilon^2}\bar{v}\right)\right\|_{L^\infty(0,T;H^{s-4})} \leq C(T)\,\varepsilon, \tag{1.13}$$

where v is the solution to (1.9).

Compared to the result in [15, Theorem 1.3], here we obtain an better error estimate of order $O(\varepsilon)$ instead of $o(\varepsilon^{1/2})$.

The second result concerns the second order approximation which is an extension of the convergence result (1.7) obtained in [15, Theorem 1.4].

Theorem 1.2. *Under the assumptions in Theorem 1.1, if in addition $f \in C^m,\ m > s+1$, and the initial datum is of the form* (1.6) *satisfying*

$$\begin{aligned}&(\varphi_1, \psi_1) \in (H^s)^2 \quad \textit{independent of } \varepsilon,\\ &\left\{(\varphi_{2,\varepsilon}, \psi_{2,\varepsilon}, \varepsilon\nabla\varphi_{2,\varepsilon})\right\}_{0<\varepsilon<1} \textit{ uniformly bounded in } (H^{s-4})^{d+2},\end{aligned} \tag{1.14}$$

then for any $T < \min\{T_\varepsilon^*, T_0^*\}$, there exists a constant $C(T)$ independent of ε such that

$$\left\| u - \left(e^{it/\varepsilon^2}(v + \varepsilon w) + e^{-it/\varepsilon^2}(\bar{v} + \varepsilon\bar{w})\right)\right\|_{L^\infty(0,T;H^{s-4})} \leq C(T)\,\varepsilon^2, \tag{1.15}$$

where v is the solution to (1.9) *and w is the solution to the Cauchy problem*

$$2iw_t - \Delta w = \tilde{\tilde{f}}(w), \quad w(0) = \frac{\varphi_1 - i\psi_1}{2} \tag{1.16}$$

with

$$\tilde{\tilde{f}}(w) := \frac{1}{2\pi}\int_0^{2\pi} e^{-i\theta} f'(e^{-i\theta}\bar{v} + e^{i\theta}v)(e^{-i\theta}\bar{w} + e^{i\theta}w)d\theta. \tag{1.17}$$

We give several remarks on our results stated in the above two theorems.

Remarks 1.3.

- *The error estimates in* (1.13) *and* (1.15) *are* optimal *in the sense that they are of the same order as the initial error estimates. This means the initial error is not much amplified by the dynamics of the system over time interval $[0, T]$. This is called sometimes the linear stability phenomena of the approximate solution.*

- *Under the assumptions in Theorem 1.2, the Cauchy problem* (1.16) *admits a unique solution* $w \in C([0, T_0^*); H^s)$. *The maximal existence time is the same as that of* $v \in C([0, T_0^*); H^s)$ *because* $\tilde{\tilde{f}}(w)$ *in* (1.17) *is linear in* w.

- *For* $f(u) = \lambda u^{q+1}$, $q \geq 0$, $q \in \mathbb{Z}$, *we have* $f \in C^\infty$. *Thus our results apply to such nonlinearities. For general* $f(u) = \lambda |u|^q u$, *to make sure* $f \in C^m$ *for some* $m > s > d/2 + 4$, *we need to assume* $q > d/2 + 4$.

- *For the typical cubic nonlinearity* $f(u) = \lambda u^3$, *direct calculation implies that the nonlinearity* (1.10) *of the approximate Schrödinger equation is also cubic* $\tilde{f}(v) = 3\lambda v^3$.

- *By* (1.12), *there exists* $\varepsilon_0 > 0$ *such that for any* $0 < \varepsilon < \varepsilon_0$ *there holds* $T_\varepsilon^* \geq T_0^*$. *Then for* $\varepsilon < \varepsilon_0$, *the error estimates* (1.13) *and* (1.15) *hold for any* $T < T_0^*$.

- *Theorem 1.1 and Theorem 1.2 hold true if initial data* $u_{0,\varepsilon}$ *and* $u_{1,\varepsilon}$ *in* (1.1) *are independent of* ε, *i.e.* $\varphi_\varepsilon = \psi_\varepsilon = 0$ *in* (1.3).

- *The results in Theorem 1.1 and Theorem 1.2 can be generalized to the Klein-Gordon equation with complex-valued unknown* $u \in \mathbb{C}$. *The proof is rather similar.*

In the sequel, if there is no specification, C denotes a constant independent of ε. Precisely, associate with the proof of Theorem 1.1, we have $C = C(s, d, D_0)$ with

$$D_0 := \|(\varphi_0, \psi_0)\|_{H^s} + \sup_{0<\varepsilon<1} \|(\varphi_\varepsilon, \psi_\varepsilon, \varepsilon \nabla \varphi_\varepsilon)\|_{H^{s-4}}.$$

Associate with the proof of Theorem 1.2, we have $C = (s, d, D_1)$ with

$$D_1 := \|(\varphi_0, \psi_0)\|_{H^s} + \|(\varphi_1, \psi_1)\|_{H^s} + \sup_{0<\varepsilon<1} \|(\varphi_{2,\varepsilon}, \psi_{2,\varepsilon}, \varepsilon \nabla \varphi_{2,\varepsilon})\|_{H^{s-4}}.$$

However, the value of C may be different from line to line.

This paper is organized as follows: In Section 2, we rewrite the Klein-Gordon equation (1.1) as a symmetric hyperbolic system and reformulate this non-relativistic limit problem as the stability problem of WKB approximate solutions in geometric optics. In Section 3, we employ the standard WKB expansion to construct an approximate solution for which the leading terms solve nonlinear Schrödinger equations. Section 4 and Section 5 are devoted to the proof of Theorem 2.3 and Theorem 2.4, respectively. Theorem 1.1 and Theorem 1.2 are direct corollaries of 2.3 and Theorem 2.4, respectively.

2 Reformulation

In this section, we reformulate this non-relativistic limit problem as the stability problem of WKB approximate solutions in geometric optics.

2.1 The equivalent symmetric hyperbolic system

We rewrite the Klein-Gordon equation into a symmetric hyperbolic system by introducing

$$U := (w, v, u) := \left(\varepsilon \nabla^T u, \varepsilon^2 \partial_t u, u\right)^T := \left(\varepsilon(\partial_{x_1} u, \cdots, \partial_{x_d} u), \varepsilon^2 \partial_t u, u\right)^T.$$

Then the equation (1.1) is equivalent to

$$\partial_t U - \frac{1}{\varepsilon}A(\partial_x)U + \frac{1}{\varepsilon^2}A_0 U = F(U), \tag{2.1}$$

where

$$A(\partial_x) := \begin{pmatrix} 0_{d\times d} & \nabla & 0 \\ \nabla^T & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_0 := \begin{pmatrix} 0_{d\times d} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad F(U) = -\begin{pmatrix} 0 \\ f(u) \\ 0 \end{pmatrix}. \tag{2.2}$$

Here the notation $\nabla := (\partial_{x_1}, \cdots, \partial_{x_d})^T$, and $0_{d\times d}$ denotes zero matrix of order $d \times d$. In what follows, we will use 0_d to denote the zero column vector of dimension d.

Under the assumptions on the initial data in (1.2), (1.3) and (1.11), we have

$$U(0) = \left(\varepsilon \nabla^T(\varphi_0 + \varepsilon\varphi_\varepsilon), \psi_0 + \varepsilon\psi_\varepsilon, \varphi_0 + \varepsilon\varphi_\varepsilon\right)^T \tag{2.3}$$

which is uniformly bounded in Sobolev space H^{s-4} with respect to ε.

The differential operator on the left-hand side of (2.1) is symmetric hyperbolic with constant coefficients. In spite of the large prefactors $1/\varepsilon$ and $1/\varepsilon^2$ in front of $A(\partial_x)$ and A_0, the H^{s-4} estimate is uniform and independent of ε because $A(\partial_x)$ and A_0 are both anti-adjoint operators. The well-posedness of Cauchy problem (2.1)-(2.3) in $C\left([0, T_\varepsilon^*); H^{s-4}\right)$ is classical (see for instance Chapter 2 of Majda [14] or Chapter 7 of Métivier [16]).

Since the nonlinearity $F(U)$ only depends on $f(u)$, the classical existence time satisfies

$$T_\varepsilon^* \propto \frac{1}{C(f, \|u\|_{L^\infty})},$$

and there is a criterion for the lifespan

$$T_\varepsilon^* < \infty \Longrightarrow \lim_{t\to T_\varepsilon^*} \|u\|_{L^\infty} = \infty. \tag{2.4}$$

2.2 WKB expansion and approximate solution

We look for an approximate solution to (2.1) by using WKB expansion which is a typical technique in geometric optics. The main idea is as follows.

We make a formal power series expansion in ε for the solution and each term in the series is a trigonometric polynomial in $\theta := t/\varepsilon^2$:

$$U_a = \sum_{n=0}^{K_a+1} \varepsilon^n U_n, \qquad U_n = \sum_{p\in\mathcal{H}_n} e^{ip\theta} U_{n,p}, \qquad K_a \in \mathbb{Z}_+,\ \mathcal{H}_n \subset \mathbb{Z}. \tag{2.5}$$

The amplitudes $U_{n,p}(t, x)$ are not highly-oscillating (independent of θ) and satisfies $U_{n,-p} = \overline{U}_{n,p}$ due to the reality of U_a. Here $\mathcal{H}_n$ is the n-th order harmonics set and will be determined in the construction of U_a. The *zero-order* or *fundamental* harmonics set $\mathcal{H}_0$ is defined as $\mathcal{H}_0 := \{p \in \mathbb{Z} :\ \det(ip + A_0) = 0\}$. For the nonhomogeneous case with $A_0 \neq 0$, the set $\mathcal{H}_0$ is always finite. Indeed, with A_0 given in (2.2), we have

$$\mathcal{H}_0 = \{-1, 0, 1\}.$$

Higher order harmonics are generated by the fundamental harmonics and the nonlinearity of the system. In general there holds the inclusion $\mathcal{H}_n \subset \mathcal{H}_{n+1}$.

We plug (2.5) into (2.1) and deduce the system of order $O(\varepsilon^n)$:

$$\Phi_{n,p} := \partial_t U_{n,p} - A(\partial_x)U_{n+1,p} + \big(ip + A_0\big)U_{n+2,p} - F(U_a)_{n,p} = 0, \tag{2.6}$$

for any $n \in \mathbb{Z}$, $n \geq -2$ and $p \in \mathbb{Z}$. In (2.6), we imposed $U_n = 0$ for any $n \leq -1$. Then to solve (2.1), it is sufficient to solve $\Phi_{n,p} = 0$ for all $(n,p) \in \mathbb{Z}^2$. This is in general not possible because there are infinity of n. However, we can solve (2.1) approximately by solving $\Phi_{n,p} = 0$ up to some nonnegative order $-2 \leq n \leq K_a - 1$ with $K_a \geq 1$, then U_a solves (2.1) with a remainder of order $O(\varepsilon^{K_a})$ which is small and goes to zero in the limit $\varepsilon \to 0$. More precisely, we look for an approximate solution U_a of the form (2.5) satisfying

$$\begin{cases} \partial_t U_a - \dfrac{1}{\varepsilon}A(\partial_x)U_a + \dfrac{1}{\varepsilon^2}A_0 U_a = F(U_a) - \varepsilon^{K_a}R^\varepsilon, \\ U_a(0,x) = U(0,x) - \varepsilon^K \Psi^\varepsilon(x), \end{cases}$$

where $|R^\varepsilon|_{L^\infty} + |\Psi^\varepsilon|_{L^\infty}$ is bounded uniformly in ε. Parameters K_a and K describe the level of precision of the approximate solution U_a.

We state the results in constructing the approximate solution. The first one concerns a first order approximation:

Proposition 2.1. *Under the assumptions in Theorem 1.1, there exits an approximate solution $U_a^{(1)} \in C([0,T_0^*); H^{s-2})$ for some $T_0^* > 0$ solving*

$$\begin{cases} \partial_t U_a^{(1)} - \dfrac{1}{\varepsilon}A(\partial_x)U_a^{(1)} + \dfrac{1}{\varepsilon^2}A_0 U_a^{(1)} = F(U_a^{(1)}) - \varepsilon R_\varepsilon^{(1)}, \\ U_a^{(1)}(0) = U(0) - \varepsilon \Psi_\varepsilon^{(1)}, \end{cases} \tag{2.7}$$

where for any $T < T_0^*$, there holds the estimate

$$\sup_{0<\varepsilon<1} \Big(\|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^{s-4})} + \|\Psi_\varepsilon^{(1)}\|_{H^{s-4}} \Big) < +\infty. \tag{2.8}$$

Moreover, $U_a^{(1)}$ is of the form (2.5) *with $U_n \in C([0,T_0^*); H^{s-2})$, $0 \leq n \leq K_a + 1 = 2$; in particular, the leading term U_0 is of the form*

$$U_0 := e^{it/\varepsilon^2} v e_+ + e^{-it/\varepsilon^2} \bar{v} e_-$$

with v the solution to (1.9) *and*

$$e_\pm := (0_d^T, \pm i, 1)^T. \tag{2.9}$$

The second one concerns a second order approximation:

Proposition 2.2. *Under the assumptions in Theorem 1.2, there exits $U_a^{(2)} \in C([0,T_0^*); H^{s-3})$ solving*

$$\begin{cases} \partial_t U_a^{(2)} - \dfrac{1}{\varepsilon}A(\partial_x)U_a^{(2)} + \dfrac{1}{\varepsilon^2}A_0 U_a^{(2)} = F(U_a^{(2)}) - \varepsilon^2 R_\varepsilon^{(2)}, \\ U_a^{(2)}(0) = U(0) - \varepsilon^2 \Psi_\varepsilon^{(2)}, \end{cases}$$

where for any $T < T_0^*$, there holds the estimates

$$\sup_{0<\varepsilon<1} \Big(\|R_\varepsilon^{(2)}\|_{L^\infty(0,T;H^{s-4})} + \|\Psi_\varepsilon^{(2)}\|_{H^{s-4}} \Big) < +\infty. \tag{2.10}$$

Moreover, $U_a^{(2)}$ is of the form (2.5) *with $U_n \in C([0,T_0^*); H^{s-3})$, $0 \leq n \leq K_a + 1 = 3$; in particular, the leading term U_0 is the same as in Proposition 2.1 and the leading error term U_1 is*

$$U_1 := e^{it/\varepsilon^2} \big(w e_+ + (\nabla^T v, 0, 0)^T \big) + e^{-it/\varepsilon^2} \big(\bar{w} e_- + (\nabla^T \bar{v}, 0, 0)^T \big)$$

with v the solution to (1.9) *and w the solution to* (1.16).

2.3 Stability of approximate solutions

To prove Theorem 1.1 and Theorem 1.2, we turn to prove the following two theorems which state the stability of the approximate solutions in Proposition 2.1 and Proposition 2.2.

Theorem 2.3. *Under the assumptions in Theorem 1.1, the Cauchy problem* (2.1)-(2.3) *admits a unique solution* $U \in C\left([0, T_\varepsilon^*); H^{s-4}\right)$ *where the maximal existence time satisfies*

$$\liminf_{\varepsilon\to 0} T_\varepsilon^* \geq T_0^*, \tag{2.11}$$

where T_0^* *is the maximal existence time of the solution to* (1.9). *Moreover, for any* $T < \min\{T_\varepsilon^*, T_0^*\}$, *there exists a constant* $C(T)$ *independent of* ε *such that*

$$\left\| U - U_a^{(1)} \right\|_{L^\infty(0,T;H^{s-4})} \leq C(T)\, \varepsilon, \tag{2.12}$$

where $U_a^{(1)}$ *is given in Proposition 2.1.*

Theorem 2.4. *Under the assumption in Theorem 1.2, for any* $T < \min\{T_\varepsilon^*, T_0^*\}$, *there exists a constant* $C(T)$ *independent of* ε *such that the solution* $U \in C\left([0, T_\varepsilon^*); H^{s-4}\right)$ *to* (2.1)-(2.3) *satisfies*

$$\left\| U - U_a^{(2)} \right\|_{L^\infty(0,T;H^{s-4})} \leq C(T)\, \varepsilon^2, \tag{2.13}$$

where $U_a^{(2)}$ *is given in Proposition 2.2.*

We gives several remarks on our results, concerning the applications and possible further results.

Remark 2.5. *It is direct to observe that Theorem 1.1 is a corollary of Theorem 2.3 and Proposition 2.1, and Theorem 1.2 is a corollary of Theorem 2.4 and Proposition 2.2. Hence, it suffices to prove Proposition 2.1, Proposition 2.2, Theorem 2.3 and Theorem 2.4.*

Remark 2.6. *We obtain uniform estimates for the solution* U *in* $L^\infty([0, T]; H^{s-4})$ *for any* $T < T_\varepsilon^*$ *in Theorem 2.3 and Theorem 2.4. As a result, we obtain the corresponding uniform estimates for* $\varepsilon^2 \partial_t u$ *where* u *is the unique solution to* (1.1)-(1.2), *since* $\varepsilon^2 \partial_t u$ *is a component of* U *which is defined in* (2.1). *This gives a rigorous verification for the technical Assumption (A) in [1].*

Remark 2.7. *If the approximate solutions is well-posed up to longer time, such as* $O(|\log \varepsilon|)$ *or even* $O(1/\varepsilon)$, *by employing the techniques in [4], one may be able to obtain convergence results up to logarithm time as* $O(|\log \varepsilon|)$. *However, to achieve such logarithm time, by using the method in [4], there is a price on the convergence rates: one may only derive* $O(\sqrt{\varepsilon})$ *error estimate in* (2.12) *and* $O(\varepsilon)$ *error estimate in* (2.13). *Such long time behavior is systematically considered in another paper [11] for quadratic nonlinearities and long time of order* $O(1/\varepsilon)$ *stability is obtained.*

Remark 2.8. *As suggested by the referee, we find out that reformulated problem* (2.1)–(2.3) *falls into the general framework studied by Lannes in [3] after a rescaling in time; in particular, the problem in this paper corresponds to the particular case with zero initial wave number (without initial oscillation) and the zero group velocity. This setting allows us to consider the following particular ansatz:*

$$u^\varepsilon(t, x) = \varepsilon^p u(\varepsilon t, x, \theta), \quad t \in \mathbb{R},\ x \in \mathbb{R}^d,\ \theta := -t/\varepsilon,$$

where p *is taken in the same manner as in [3], such that one can gain* ε *from the nonlinearity, while an often used ansatz in general setting is: (see equation (2) in [3]):*

$$u^\varepsilon(t, x) = \varepsilon^p u(\varepsilon t, t, x, \theta), \quad t \in \mathbb{R},\ x \in \mathbb{R}^d,\ \theta := (k \cdot x - \omega t)/\varepsilon,$$

where k is the initial wave number and ω is the temporal frequency associated with the initial wave number.

The results obtained in [3] show the convergence of the approximate solutions to the exact solutions; however, no further convergence rate is provided. The main issue is the possible secular growth caused by the general structure of the systems (the differential operators and the nonlinearities). Here in this paper, we do not include initial oscillation (zero initial wave number). Together with the special structure of the Klein–Gordon equations, we do not see any secular growth in the stability analysis and we obtain optimal convergence rates.

3 Construction of WKB solutions

We now carry out the idea in Section 2.2 to construct approximate solutions satisfying the properties stated in Proposition 2.1 and Proposition 2.2.

3.1 WKB cascade

We start from considering $\Phi_{-2,p} = 0$ corresponding to the equations in the terms of order $O(\varepsilon^{-2})$. We reproduce such equations as follows

$$(ip + A_0)U_{0,p} = 0, \quad \text{for all } p. \tag{3.1}$$

It is easy to find that $(ip + A_0)$ are invertible except $p \in \mathcal{H}_0 = \{-1, 0, 1\}$. We then deduce from (3.1) that

$$U_{0,p} = 0, \quad \text{for all } p \text{ such that } |p| \geq 2. \tag{3.2}$$

This is in fact how we determine $\mathcal{H}_0$: for any $p \notin \mathcal{H}_0$, necessarily $U_{0,p} = 0$.

As in [15], we do not need to include the mean mode $U_{0,0}$ in the approximation. Hence, for simplicity, we take

$$U_{0,0} = 0. \tag{3.3}$$

For $p = 1$, (3.1) is equivalent to the so called polarization condition $U_{0,p} \in \ker(ip + A_0)$. This implies

$$U_{0,1} = g_0 e_+, \tag{3.4}$$

for some scalar function g_0.

For $p = -1$, reality implies

$$U_{0,-1} = \overline{U}_{0,1} = \bar{g}_0 e_-.$$

Here $e_\pm$ are defined as in (2.9) satisfying

$$\ker(i + A_0) = \text{span}\,\{e_+\}, \quad \ker(-i + A_0) = \text{span}\,\{e_-\}.$$

We continue to consider the equations in the terms of order $O(\varepsilon^{-1})$ which are $\Phi_{-1,p} = 0$:

$$-A(\partial_x)U_{0,p} + (ip + A_0)U_{1,p} = 0, \quad \text{for all } p. \tag{3.5}$$

When $p = 0$, by the choice of the leading mean mode in (3.3), equation (3.5) becomes

$$A_0 U_{1,0} = 0$$

which is equivalent to

$$U_{1,0} = (h_1^T, 0, 0)^T \quad \text{for some vector valued function } h_1 \in \mathbb{R}^d. \tag{3.6}$$

When $p = 1$, by (3.4), equation (3.5) is equivalent to

$$U_{1,1} = g_1 e_+ + (\nabla^T g_0, 0, 0)^T \quad \text{for some scalar function } g_1.$$

When $|p| \geq 2$, the invertibility of $(ip + A_0)$ and (3.2) imply

$$U_{1,p} = 0, \quad \text{for all } p \text{ such that } |p| \geq 2.$$

The equations which comprise all terms of order $O(\varepsilon^0)$ are

$$\partial_t U_{0,p} - A(\partial_x) U_{1,p} + (ip + A_0) U_{2,p} = F(U_0)_p, \quad \text{for all } p. \tag{3.7}$$

Here $F(U_0)_p$ is the p-th coefficient of the Fourier series of $F(U_0)$ in θ. Precisely,

$$F(U_0)_p = (0_d^T, -\tilde{f}_p, 0)^T, \ \tilde{f}_p := f(u_0)_p = \frac{1}{2\pi} \int_0^{2\pi} e^{-ip\theta} \check{f}(\theta) d\theta, \tag{3.8}$$

where

$$\check{f}(\theta) := f(u_0)(\theta) := f(e^{-i\theta} \bar{g}_0 + e^{i\theta} g_0). \tag{3.9}$$

Here we used the notation for the corresponding components:

$$U_n =: \begin{pmatrix} w_n \\ v_n \\ u_n \end{pmatrix}, \quad \text{for any } n \in \mathbb{Z}, \ n \geq 0. \tag{3.10}$$

Lemma 3.1. *For $\tilde{f}_p$ defined by* (3.8) *and* (3.9)*, we have the estimates for any $p \in \mathbb{Z}$:*

$$\begin{aligned} &\|\tilde{f}_p(t)\|_{H^\sigma} \leq C(f, \|g_0(t)\|_{L^\infty}) \|g_0(t)\|_{H^\sigma}, \quad \textit{for any } 0 \leq \sigma < m, \\ &\|\tilde{f}_p(t)\|_{H^\sigma} \leq \frac{C(f, \|g_0(t)\|_{L^\infty})}{1 + |p|} \|g_0(t)\|_{H^\sigma}, \quad \textit{for any } 0 \leq \sigma < m - 1, \end{aligned} \tag{3.11}$$

where the dependency of the constant C is as follows

$$C(f, \|g_0(t)\|_{L^\infty}) = C\left(\sum_{|\alpha| \leq m} \|f^{(\alpha)}(g_0(t))\|_{L^\infty}, \|g_0(t)\|_{L^\infty} \right).$$

Proof of Lemma 3.1. Since $f \in C^m, \ m > s > d/2 + 4$, then we have

$$\|f(u)\|_{H^\sigma} \leq C\left(\sum_{|\alpha| \leq m} \|f^{(\alpha)}(u)\|_{L^\infty}, \|u\|_{L^\infty} \right) \|u\|_{H^\sigma}, \quad \text{for any } 0 \leq \sigma < m. \tag{3.12}$$

For a proof of this fact (3.12), we refer to Theorem 5.2.6 in [16]. Then it is direct to deduce $(3.11)_1$. We can obtain $(3.11)_2$ by observing for any $p \neq 0$:

$$\tilde{f}_p := \frac{1}{2\pi} \int_0^{2\pi} (ip)^{-1} e^{-ip\theta} f'(e^{-i\theta} \bar{g}_0 + e^{i\theta} g_0)(-ie^{-i\theta} \bar{g}_0 + ie^{i\theta} g_0) d\theta.$$

□

When $p = 0$, equation (3.7) is equivalent to

$$U_{2,0} = (h_2^T, 0, \mathrm{div} h_1 - \tilde{f}_0)^T \quad \text{for some vector valued function } h_2 \in \mathbb{R}^d. \tag{3.13}$$

When $p = 1$, equation (3.7) is equivalent to

$$\begin{cases} 2i\partial_t g_0 - \Delta g_0 + \tilde{f}_1 = 0, \\ U_{2,1} = g_2 e_+ + (\nabla^T g_1, \partial_t g_0, 0)^T, \quad \text{for some scalar function } g_2. \end{cases} \tag{3.14}$$

Here the equation $(3.14)_1$ is exactly the nonlinear Schrödinger equation (1.9). The initial datum of g_0 is determined in such a way that $U_0(0) = (0_d^T, \psi_0, \varphi_0)^T$ which is the leading term of initial data $U(0)$ (see (2.3)). This imposes

$$g_0(0) = \frac{\varphi_0 - i\psi_0}{2}. \tag{3.15}$$

By the regularity assumption (1.11), initial data $g_0(0) \in H^s$, $s > d/2 + 4$, then by Lemma 3.1 and the classical theory for the local well-posedness of Schröginder equations (see for instance Chapter 8 of [16]), the Cauchy problem $(3.14)_1$-(3.15) admits a unique solution $g_0 \in C([0, T_0^*); H^s) \cap C^1([0, T_0^*); H^{s-2})$ where T_0^* is the maximal existence time. Moreover, there holds the estimate for any $T < T_0^*$:

$$\|\partial_t g_0\|_{L^\infty(0,T;H^{s-2})} \leq C\|g_0\|_{L^\infty(0,T;H^s)} \leq C\|(\phi_0, \psi_0)\|_{H^s}. \tag{3.16}$$

When $|p| \geq 2$, equation (3.7) is equivalent to

$$(ip + A_0)U_{2,p} = \tilde{f}_p \Longleftrightarrow U_{2,p} = (ip + A_0)^{-1}\tilde{f}_p. \tag{3.17}$$

3.2 First order approximation - Proof of Proposition 2.1

In this subsection, we are working under the assumptions in Theorem 1.1 and we will finish the proof of Proposition 2.1.

We stop the WKB expansion by taking

$$g_1 = g_2 = h_1 = h_2 = 0$$

in (3.6), (3.13) and $(3.14)_2$, we construct an approximate solution

$$U_a^{(1)} := U_0 + \varepsilon U_1 + \varepsilon^2 U_2,$$

where

$$\begin{aligned} &U_0 := e^{it/\varepsilon^2} g_0 e_+ + c.c., \qquad U_1 := e^{it/\varepsilon^2} \begin{pmatrix} \nabla g_0 \\ 0 \\ 0 \end{pmatrix} + c.c., \\ &U_2 := \begin{pmatrix} 0_d \\ 0 \\ -\tilde{f}_0 \end{pmatrix} + \left(e^{it/\varepsilon^2} \begin{pmatrix} 0_d \\ \partial_t g_0 \\ 0 \end{pmatrix} + c.c. \right) + \sum_{|p|\geq 2} e^{ipt/\varepsilon^2} U_{2,p}. \end{aligned} \tag{3.18}$$

Here $c.c.$ means complex conjugate and $z + c.c. = 2\Re z$ is two times of the real part of z.

Now we verify such $U_a^{(1)}$ fulfills the properties stated in Proposition 2.1.

By $(3.11)_2$ in Lemma 3.1 and (3.17), we have

$$\|U_{2,p}(t)\|_{H^{s-1}} \le \frac{C(f, \|g_0(t)\|_{L^\infty})}{(1+|p|)^2} \|g_0(t)\|_{H^{s-1}}, \quad |p| \ge 2.$$

By (3.16) and (3.18), we obtain for any $t < T_0^*$:

$$\|U_2(t)\|_{H^{s-2}} \le \sum_{p\in\mathbb{Z}} \|U_{2,p}(t)\|_{H^{s-2}} \le C\|g_0(t)\|_{H^s}. \tag{3.19}$$

Together with (3.16) and (3.18), we conclude $U_a^{(1)} \in C([0, T_0^*); H^{s-2})$ solving (2.7) with the remainders

$$R_\varepsilon^{(1)} := \frac{F(U_a^{(1)}) - F(U_0)}{\varepsilon} + \left(e^{it/\varepsilon^2} \begin{pmatrix} \nabla\partial_t g_0 \\ 0 \\ 0 \end{pmatrix} + c.c. \right) + A(\partial_x)U_2 - \varepsilon \sum_{p\in\mathbb{Z}} e^{ipt/\varepsilon^2} \partial_t U_{2,p},$$

$$\Psi_\varepsilon^{(1)} := \left(\varepsilon\nabla^T\varphi_\varepsilon, \psi_\varepsilon, \varphi_\varepsilon\right)^T - \varepsilon U_2(0).$$

By (1.3), (1.11), (2.3) and (3.19), we have the uniform estimate (2.8) for $\Psi_\varepsilon^{(1)}$.

It is left to show the uniform estimate (2.8) for $R_\varepsilon^{(1)}$. Direct calculation gives

$$\frac{F(U_a^{(1)}) - F(U_0)}{\varepsilon} = \int_0^1 F'(U_0 + \varepsilon\tau(U_1 + \varepsilon U_2)) \cdot (U_1 + \varepsilon U_2) d\tau.$$

Since $f \in C^m,\ m > s$, so des F. Then

$$\frac{1}{\varepsilon}\|F(U_a^{(1)}) - F(U_0)(t)\|_{H^\sigma} \le C\left(f, \sum_j \|U_j\|_{L^\infty}\right) \sum_j \|U_j\|_{H^\sigma}$$

for any $0 \le \sigma < m - 1$.

Again by (3.18) and (3.16), we have for any $t < T_0^*$:

$$\frac{1}{\varepsilon}\|F(U_a^{(1)}) - F(U_0)(t)\|_{H^{s-4}} \le C.$$

The proof of the uniform estimates for other terms in $R_\varepsilon^{(1)}$ is similar and rather direct by using the estimate (3.12). We omit the details.

This approximate solution $U_a^{(1)}$ will be used to prove Theorem 2.3 in Section 4.

3.3 Second order approximation - Proof of Proposition 2.2

First of all, we point out that in this subsection, we are working under the assumptions in Theorem 1.2 and we will finish the proof of Proposition 2.1.

We continue the WKB expansion process from the end of Section 3.1 where we achieved the equations of order $O(\varepsilon^0)$:

The equations in the terms of order $O(\varepsilon)$ are

$$\partial_t U_{1,p} - A(\partial_x)U_{2,p} + (ip + A_0)U_{3,p} = -(0_d^T, \tilde{\tilde{f}}_p, 0)^T, \quad \text{for all } p, \tag{3.20}$$

where

$$\tilde{\tilde{f}}_p := (f'(u_0)u_1)_p = \frac{1}{2\pi}\int_0^{2\pi} e^{-ip\theta} f'(e^{-i\theta}\bar{g}_0 + e^{i\theta}g_0)(e^{-i\theta}\bar{g}_1 + e^{i\theta}g_1)d\theta. \tag{3.21}$$

Here we used the notations in (3.10).

A similar proof as that of Lemma 3.1, we can obtain:

Lemma 3.2. *There holds the estimates for any $p \in \mathbb{Z}$:*

$$\begin{aligned}
&\|\tilde{\tilde{f}}_p(t)\|_{H^\sigma} \le C(f, \|(g_0,g_1)(t)\|_{L^\infty})\|(g_0,g_1)(t)\|_{H^\sigma}, \quad \textit{for any } 0 \le \sigma < m-1,\\
&\|\tilde{\tilde{f}}_p(t)\|_{H^\sigma} \le \frac{C(f, \|(g_0,g_1)(t)\|_{L^\infty}}{1+|p|}\|(g_0,g_1)(t)\|_{H^\sigma}, \quad \textit{for any } 0 \le \sigma < m-2,
\end{aligned}$$

where the dependency of the constant C is as follows

$$C\left(f, \|(g_0,g_1)(t)\|_{L^\infty}\right) = C\left(\sum_{|\alpha|\le m} \|f^{(\alpha)}(g_0(t))\|_{L^\infty}, \|(g_0,g_1)(t)\|_{L^\infty}\right).$$

When $p = 0$, equation (3.20) becomes

$$\partial_t U_{1,0} - A(\partial_x)U_{2,0} + A_0 U_{3,0} = -(0_d^T, \tilde{\tilde{f}}_0, 0)^T$$

which is equivalent to

$$\partial_t h_1 = 0, \quad U_{3,0} = (h_3^T, 0, \mathrm{div}h_2 - \tilde{\tilde{f}}_0)^T \tag{3.22}$$

for some vector valued function $h_3 \in \mathbb{R}^d$.

When $p = 1$, equation (3.20) becomes

$$\partial_t U_{1,1} - A(\partial_x)U_{2,1} + (i + A_0)U_{3,1} = -(0_d^T, \tilde{\tilde{f}}_1, 0)^T,$$

which is equivalent to

$$\begin{cases} 2i\partial_t g_1 - \Delta g_1 + \tilde{\tilde{f}}_1 = 0, \\ U_{3,1} = g_3 e_+ + (\nabla^T g_2, \partial_t g_1, 0)^T, \quad \text{for some scalar function } g_3. \end{cases} \tag{3.23}$$

We find that g_1 satisfies a Schrödinger equation where the source term $\tilde{\tilde{f}}_1$ is actually linear in g_1 (see (3.21)). The initial data $g_1(0)$ is determined such that $U_1(0) = (\nabla^T\varphi_0, \psi_1, \varphi_1)^T$ which is the first order ($O(\varepsilon)$) perturbation of $U(0)$ (see (2.3) and (1.6)). This imposes

$$g_1(0) + \bar{g}_1(0) = \varphi_1, \quad ig_1(0) - i\bar{g}_1(0) = \psi_1,$$

which is equivalent to

$$g_1 = \frac{\varphi_1 - i\psi_1}{2} \in H^s. \tag{3.24}$$

This is exactly the initial condition in (1.16), and we used the regularity assumption in (1.14).

Since $m > s+1,\ s > d/2 + 4$, by Lemma 3.2 and the classical theory, the Cauchy problem $(3.23)_1$-(3.24) admits a unique solution in Sobolev space $C\left([0, T_1^*), H^s\right)$. Here we have the maximal existence time $T_1^* = T_0^*$ where T_0^* is the maximal existence time for the solution $g_0 \in C([0,T_0^*), H^s)$ to $(3.14)_1$, because $\tilde{\tilde{f}}_1$ is linear in g_1. Moreover, there holds for any $T < T_0^*$:

$$\|\partial_t g_1\|_{L^\infty(0,T;H^{s-2})} \le C\|g_1\|_{L^\infty(0,T;H^s)} \le C\|(\phi_1,\psi_1)\|_{H^s}.$$

When $|p| \ge 2$, equation (3.20) becomes

$$-A(\partial_x)U_{2,p} + (ip + A_0)U_{3,p} = -(0_d^T, \tilde{\tilde{f}}_p, 0)^T,$$

which is equivalent to

$$U_{3,p} = (ip + A_0)^{-1}\left(A(\partial_x)U_{2,p} - (0_d^T, \tilde{\tilde{f}}_p, 0)^T\right).$$

We stop the WKB expansion and take

$$g_2 = g_3 = h_1 = h_2 = h_3 = 0$$

in (3.6), (3.13), (3.22) and $(3.23)_2$. Then we construct another approximate solution

$$U_a^{(2)} := U_0 + \varepsilon U_1 + \varepsilon^2 U_2 + \varepsilon^3 U_3,$$

where

$$U_0 := e^{it/\varepsilon^2} g_0 e_+ + c.c., \qquad U_1 := e^{it/\varepsilon^2}\begin{pmatrix}\nabla g_0\\ i g_1\\ g_1\end{pmatrix} + c.c.,$$

$$U_2 := \begin{pmatrix}0_d\\ 0\\ -\tilde{f}_0\end{pmatrix} + \left(e^{it/\varepsilon^2}\begin{pmatrix}\nabla g_1\\ \partial_t g_0\\ 0\end{pmatrix} + c.c.\right) + \sum_{|p|\ge 2} e^{ipt/\varepsilon^2} U_{2,p},$$

$$U_3 := \begin{pmatrix}0_d\\ 0\\ -\tilde{\tilde{f}}_0\end{pmatrix} + \left(e^{it/\varepsilon^2}\begin{pmatrix}0_d\\ \partial_t g_1\\ 0\end{pmatrix} + c.c.\right) + \sum_{|p|\ge 2} e^{ipt/\varepsilon^2} U_{3,p}.$$

Similar as the verification of $U_a^{(1)}$ at the end of Section 3.2 for the proof of Proposition 2.1, we can show that $U_a^{(2)}$ constructed above fulfills the properties given in Proposition 2.2. We omit the details here.

The approximate solution $U_a^{(2)}$ will be used to prove Theorem 2.4 in Section 5.

4 Proof of Theorem 2.3

Associate with the approximate solution $U_a^{(1)}$ in Proposition 2.1, we define the error

$$\dot{U} := \frac{U - U_a^{(1)}}{\varepsilon}, \tag{4.1}$$

where $U \in C([0,T_\varepsilon^*); H^{s-4})$ is the exact solution to (2.1)-(2.3). Then at least over the time interval $[0, \min\{T_\varepsilon^*, T_0^*\})$, $\dot{U}$ solves

$$\begin{cases} \partial_t \dot{U} - \dfrac{1}{\varepsilon}A(\partial_x)\dot{U} + \dfrac{1}{\varepsilon^2}A_0\dot{U} = \dfrac{1}{\varepsilon}\left(F(U_a^{(1)} + \varepsilon\dot{U}) - F(U_a^{(1)})\right) + R_\varepsilon^{(1)},\\ \dot{U}(0) = \Psi_\varepsilon^{(1)}, \end{cases} \tag{4.2}$$

where $R_\varepsilon^{(1)}$ and $\Psi_\varepsilon^{(1)}$ satisfy the uniform estimate in (2.8).

Concerning the well-posedness of Cauchy problem (4.2), we have the following proposition.

Proposition 4.1. *Under the assumptions in Theorem 1.1, the Cauchy problem* (4.2) *admits a unique solution* $\dot{U} \in C([0, \widetilde{T}_\varepsilon^*); H^{s-4})$ *where* $\widetilde{T}_\varepsilon^*$ *is the maximal existence time. Moreover, there holds*

$$\liminf_{\varepsilon \to 0} \widetilde{T}_\varepsilon^* \geq T_0^*, \tag{4.3}$$

and for any $T < \min\{T_0^*, \widetilde{T}_\varepsilon^*\}$*:*

$$\sup_{0<\varepsilon<1} \|\dot{U}\|_{L^\infty(0,T;H^{s-2})} \leq C(T). \tag{4.4}$$

Proof of Proposition 4.1. We calculate

$$\frac{1}{\varepsilon}\left(F(U_a^{(1)} + \varepsilon \dot{U}) - F(U_a^{(1)})\right) = F'(U_a)\dot{U} + \varepsilon \int_0^1 F''(U_a + \varepsilon \tau \dot{U})\dot{U}^2 \frac{(1-\tau)^2}{2} d\tau.$$

Since $U_a^{(1)} \in C([0, T_0^*); H^{s-2})$ and $F \in C^m,\ m > s$ with $s > d/2 + 4$, then for any $t < T_0^*$, there holds

$$\begin{aligned}
&\left\| \frac{1}{\varepsilon}\left(F(U_a^{(1)} + \varepsilon \dot{U}) - F(U_a^{(1)})\right)(t) \right\|_{H^{s-4}} \leq \\
&\left(C(F, \|\dot{U}_a(t)\|_{H^{s-4}}) + \varepsilon\, C(F, \|\dot{U}_a(t)\|_{H^{s-4}}, \|\dot{U}(t)\|_{H^{s-4}})\right) \|\dot{U}(t)\|_{H^{s-4}}.
\end{aligned} \tag{4.5}$$

The system in $\dot{U}$ is semi-linear symmetric hyperbolic and the initial datum is uniformly bounded in H^{s-4}. By (4.5), the local-in-time well-posedness of Cauchy problem (4.2) in Sobolev space H^{s-4} is classical (see for instance Chapter 7 of [16]). Moreover, if we denote $\widetilde{T}_\varepsilon^*$ to be the maximal existence time, the classical solution is in $C([0, \widetilde{T}_\varepsilon^*); H^{s-4})$ and there holds the estimate

$$\sup_{0<\varepsilon<1} \|\dot{U}\|_{L^\infty(0,T;H^{s-2})} \leq C(T), \quad \text{for any } T < \min\{\widetilde{T}_\varepsilon^*, T_0^*\}$$

and the criterion of the life-span

$$\widetilde{T}_\varepsilon^* < \infty \Longrightarrow \lim_{t \to \widetilde{T}_\varepsilon^*} \|\dot{U}\|_{L^\infty} = \infty. \tag{4.6}$$

It is left to prove (4.3) to finish the proof. Let $T < T_0^*$ be a arbitrary number. It is sufficient to show there exists $\varepsilon_0 > 0$ such that $\widetilde{T}_\varepsilon^* > T$ for any $0 < \varepsilon < \varepsilon_0$. By classical energy estimates in Sobolev spaces for semi-linear symmetric hyperbolic system, we have for any $t < \min\{T, \widetilde{T}_\varepsilon^*\}$:

$$\frac{d}{dt}\|\dot{U}(t)\|_{H^{s-4}} \leq C(T)\left(1 + \varepsilon\, C(\|\dot{U}(t)\|_{H^{s-4}})\right)\|\dot{U}(t)\|_{H^{s-4}} + \|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^s)}.$$

Here $C(\|\dot{U}(t)\|_{H^{s-4}})$ is continuous and increasing in $\|\dot{U}(t)\|_{H^{s-4}}$. Then Gronwall's inequality implies

$$\begin{aligned}
\|\dot{U}(t)\|_{H^{s-4}} \leq & \exp\left(\int_0^t \left(C(T)(1 + \varepsilon\, C(\|\dot{U}(\tau)\|_{H^{s-4}}))\right) d\tau\right) \|\dot{U}(0)\|_{H^{s-4}} \\
& + T\|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^{s-4})}.
\end{aligned} \tag{4.7}$$

Let

$$M(T) := \exp\left(2C(T)T\right)\|\dot{U}(0)\|_{H^{s-4}} + T\|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^{s-4})}.$$

We then define

$$\mathbf{T} := \sup\left\{t : \|\dot{U}\|_{L^\infty(0,t;H^{s-4})} \le M(T)\right\}.$$

If $\mathbf{T} \le \min\{T, \widetilde{T}_\varepsilon^*\}$, then for any $t < \mathbf{T}$, the inequality (4.7) implies

$$\|\dot{U}(t)\|_{H^{s-4}} \le \exp\left\{TC(T)[1 + \varepsilon\, C(M(T))]\right\} \|\dot{U}(0)\|_{H^{s-4}} + T\|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^{s-4})}.$$

Let

$$\varepsilon_0 := \{2C(M(T))\}^{-1}.$$

Then for any $0 < \varepsilon < \varepsilon_0$, there holds

$$\|\dot{U}(t)\|_{H^{s-4}} \le \exp\left(\frac{3}{2}C(T)T\right) \|\dot{U}(0)\|_{H^{s-4}} + T\|R_\varepsilon^{(1)}\|_{L^\infty(0,T;H^{s-4})}.$$

The classical continuation argument implies that

$$\mathbf{T} > \min\{T, \widetilde{T}_\varepsilon^*\}, \quad \text{for any } 0 < \varepsilon < \varepsilon_0. \tag{4.8}$$

By (4.6), we have $\widetilde{T}_\varepsilon^* \ge \mathbf{T}$. Together with (4.8), we deduce $\widetilde{T}_\varepsilon^* > T$. Since $T < T_0^*$ is a arbitrary number, we obtain (4.3) and complete the proof. □

Now we are ready to prove Theorem 2.3. Given $U_a^{(1)}$ as in Proposition 2.1 and $\dot{U}$ the solution of (4.2), we can reconstruct U through (4.1):

$$U = U_a^{(1)} + \varepsilon\dot{U}$$

which solves (2.1)-(2.3). This implies that the maximal existence time T_ε^* of the solution $U \in C([0, T_\varepsilon^*); H^{s-4})$ satisfies

$$T_\varepsilon^* \ge \min\{T_0^*, \widetilde{T}_\varepsilon^*\}.$$

By (4.3) in Proposition 4.1, we obtain (2.11) in Theorem 2.3.

Finally, by (3.16), (3.18) and (4.4), we deduce (2.12) and we complete the proof of Theorem 2.3.

5 Proof of Theorem 2.4

Associate with the approximate solution $U_a^{(2)}$ in Proposition 2.2, we define the error

$$\dot{V} := \frac{U - U_a^{(2)}}{\varepsilon^2}. \tag{5.1}$$

Then the equation and initial datum for $\dot{V}$ are

$$\begin{cases} \partial_t \dot{V} - \dfrac{1}{\varepsilon} A(\partial_x)\dot{V} + \dfrac{1}{\varepsilon^2} A_0 \dot{V} = \dfrac{1}{\varepsilon^2}\left(F(U_a^{(1)} + \varepsilon^2 \dot{V}) - F(U_a^{(1)})\right) + R_\varepsilon^{(2)}, \\ \dot{V}(0) = \Psi_\varepsilon^{(2)}, \end{cases} \tag{5.2}$$

where $R_\varepsilon^{(2)}$ and $\Psi_\varepsilon^{(2)}$ satisfy the uniform estimate in (2.10).

Similar to Proposition 4.1, we have

Proposition 5.1. *Under the assumptions in Theorem 2.4, the Cauchy problem* (5.2) *admits a unique solution* $\dot{V} \in C([0, \hat{T}^*_\varepsilon); H^{s-4})$ *where* $\hat{T}^*_\varepsilon$ *is the maximal existence time. Moreover, there holds*

$$\liminf_{\varepsilon\to 0} \hat{T}^*_\varepsilon \geq T^*_0 \tag{5.3}$$

and for any $T < \min\{T^*_0, \hat{T}^*_\varepsilon\}$*:*

$$\sup_{0<\varepsilon<1} \|\dot{V}\|_{L^\infty(0,T;H^{s-4})} \leq C(T). \tag{5.4}$$

The proof is the same as the proof of Proposition 4.1. Theorem 2.4 follows from Proposition 5.1 through a similar argument as in the end of Section 4.

Acknowledgments. We would like to thank the anonymous referee for the suggestions and particularly for pointing out reference [3]. The first author acknowledges the support of the project LL1202 in the programme ERC-CZ funded by the Ministry of Education, Youth and Sports of the Czech Republic. The second author thanks Professor Weizhu Bao for helpful discussions. Z. Zhang was partially supported by NSF of China under Grant 11371039 and 11425103.

References

[1] W. Bao, Y. Cai, X. Zhao, *A uniformly accurate multiscale time integrator pseudospectral method for the Klein-Gordon equation in the nonrelativistic limit regime*, SIAM J. Numer. Anal., 52 (2014), 2488-2511.

[2] W. Bao, X. Dong, *Analysis and comparison of numerical methods for the Klein-Gordon equation in the nonrelativistic limit regime*, Numer. Math., 120 (2012), 189-229.

[3] D. Lannes, *Dispersive effects for nonlinear geometrical optics with rectification*, Asymptot. Anal., 18 (1998), no. 1-2, 111-146.

[4] T. Colin, D. Lannes, *Justification of and long-wave correction to Davey-Stewartson systems from quadratic hyperbolic systems*, Discrete Contin. Dyn. Syst., 11 (2004), 83-100.

[5] J. Ginibre, G. Velo, *The global Cauchy problem for the nonlinear Klein-Gordon equation*, Math. Z., 189 (1985), 487-505.

[6] J. Ginibre, G. Velo, *The global Cauchy problem for the nonlinear Klein-Gordon equation II*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 6 (1989), 15-35.

[7] J.-L. Joly, G. Métivier, J. Rauch, *Transparent nonlinear geometric optics and Maxwell-Bloch equations*, J. Difference Equations, 166 (2000), 175-250.

[8] Y. Lu, *High-frequency limit of the Maxwell-Landau-Lifshitz system in the diffractive optics regime*, Asymptotic Anal., 82 (2013), 109-137.

[9] Y. Lu, *Higher-order resonances and instability of high-frequency WKB solutions*, J. Differential Equations, 260(3) (2016), 2296-2353.

[10] Y. Lu, B. Texier, *A stability criterion for high-frequency oscillations,* Mém. Soc. Math. Fr. 142 (2015), vi+130pp.

[11] Y. Lu, Z. Zhang, *Partially strong transparency conditions and a singular localization method in geometric optics*, Arch. Ration. Mech. Anal., 222(1) (2016), 245-283.

[12] S. Machihara, *The nonrelativistic limit of the nonlinear Klein-Gordon equation*, Funkcial. Ekvac., 44 (2001), 243-252.

[13] S. Machihara, K. Nakanishi, T. Ozawa, *Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations*, Math. Ann., 322 (2002), 603-621.

[14] A. Majda, *Compressible fluid flows and systems of conservation laws in several space variables*, Applied Math Sciences, Springer, vol. 53, 1984.

[15] N. Masmoudi, K. Nakanishi, *From nonlinear Klein-Gordon equation to a system of coupled nonliner Schrödinger equations*, Math. Ann., 324 (2002), 359-389.

[16] G. Métivier *Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems*, Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series, 5. Edizioni della Normale, Pisa, 2008. xii.

[17] C. Morawetz, W. Strauss, *Decay and scattering of solutions of a nonlinear relativistic wave equation*, Comm. Pure Appl. Math., 25 (1972), 1-31.

[18] B. Najman, *The nonrelativistic limit of the nonlinear Klein-Gordon equation*, Nonlinear Anal., 15 (1990), 217-228.

[19] W. Strauss, L. Vázquez, *Numerical solution of a nonlinear Klein-Gordon equation*, J. Comput. Phys., 28 (1978), 271-278.

[20] B. Texier, *Derivation of the Zakharov equations*, Arch. Ration. Mech. Anal., 184 (2007), 121-183.

[21] M. Tsutsumi, *Nonrelativistic approximation of nonlinear Klein-Gordon equations in two space dimensions*, Nonlinear Anal., 8 (1984), 637-643.