

On Time-Splitting Pseudospectral Discretization for Nonlinear Klein-Gordon Equation in Nonrelativistic Limit Regime

Xuanchun Dong¹, Zhiguo Xu^{3,1} and Xiaofei Zhao^{2,*}

¹ Beijing Computational Science Research Center, Beijing 100084, P.R. China.

² Department of Mathematics, National University of Singapore, Singapore 119076, Singapore.

³ College of Mathematics, Jilin University, Changchun 130012, P.R. China.

Received 28 August 2013; Accepted (in revised version) 19 February 2014

Available online 23 May 2014

Abstract. In this work, we are concerned with a time-splitting Fourier pseudospectral (TSFP) discretization for the Klein-Gordon (KG) equation, involving a dimensionless parameter $\varepsilon \in (0, 1]$. In the nonrelativistic limit regime, the small ε produces high oscillations in exact solutions with wavelength of $\mathcal{O}(\varepsilon^2)$ in time. The key idea behind the TSFP is to apply a time-splitting integrator to an equivalent first-order system in time, with both the nonlinear and linear subproblems exactly integrable in time and, respectively, Fourier frequency spaces. The method is fully explicit and time reversible. Moreover, we establish rigorously the optimal error bounds of a second-order TSFP for fixed $\varepsilon = \mathcal{O}(1)$, thanks to an observation that the scheme coincides with a type of trigonometric integrator. As the second task, numerical studies are carried out, with special efforts made to applying the TSFP in the nonrelativistic limit regime, which are geared towards understanding its temporal resolution capacity and meshing strategy for $\mathcal{O}(\varepsilon^2)$ -oscillatory solutions when $0 < \varepsilon \ll 1$. It suggests that the method has uniform spectral accuracy in space, and an asymptotic $\mathcal{O}(\varepsilon^{-2}\Delta t^2)$ temporal discretization error bound (Δt refers to time step). On the other hand, the temporal error bounds for most trigonometric integrators, such as the well-established Gautschi-type integrator in [6], are $\mathcal{O}(\varepsilon^{-4}\Delta t^2)$. Thus, our method offers much better approximations than the Gautschi-type integrator in the highly oscillatory regime. These results, either rigorous or numerical, are valid for a splitting scheme applied to the classical relativistic NLS reformulation as well.

AMS subject classifications: 35L70, 65M12, 65M15, 65M70

Key words: Klein-Gordon equation, high oscillation, time-splitting, trigonometric integrator, error estimate, meshing strategy.

*Corresponding author. *Email addresses:* dong.xuanchun@gmail.com (X. Dong), zhxfnus@gmail.com (X. Zhao), xuzg11@csrc.ac.cn (Z. Xu)

1 Introduction

The relativistic Klein-Gordon (KG) equation in d -dimensions ($d = 1, 2, 3$) reads, under a proper non-dimensionalization [6, 26–28, 30, 31, 43],

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

with initial conditions:

$$u(\mathbf{x}, 0) = \phi_1(\mathbf{x}), \quad \partial_t u(\mathbf{x}, 0) = \frac{1}{\varepsilon^2} \phi_2(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \quad (1.1b)$$

The KG equation is also known as the relativistic version of the Schrödinger equation and used to describe the motion of a spinless particle; see, e.g. [13, 34] for its derivation. In this work, $u = u(\mathbf{x}, t)$ is considered to be a real-valued scalar field, the dimensionless parameter $\varepsilon > 0$ is inversely proportional to the speed of light, ϕ_1 and ϕ_2 are two given real-valued functions independent of ε . $f(\cdot)$ is a real-valued function describing the non-linear interaction, independent of ε and satisfying $f(0) = 0$. The KG equation (1.1) is time symmetry or time reversible and conserves the energy, provided that $u(\cdot, t) \in H^1(\mathbb{R}^d)$ and $\partial_t u(\cdot, t) \in L^2(\mathbb{R}^d)$,

$$\begin{aligned} E(t) &:= \int_{\mathbb{R}^d} \left[\varepsilon^2 (\partial_t u(\mathbf{x}, t))^2 + |\nabla u(\mathbf{x}, t)|^2 + \frac{1}{\varepsilon^2} (u(\mathbf{x}, t))^2 + F(u(\mathbf{x}, t)) \right] dx \\ &\equiv \int_{\mathbb{R}^d} \left[\frac{1}{\varepsilon^2} (\phi_2(\mathbf{x}))^2 + |\nabla \phi_1(\mathbf{x})|^2 + \frac{1}{\varepsilon^2} (\phi_1(\mathbf{x}))^2 + F(\phi_1(\mathbf{x})) \right] dx := E(0), \quad t \geq 0, \end{aligned} \quad (1.2)$$

with $F(u) = 2 \int_0^u f(\rho) d\rho$, $u \in \mathbb{R}$.

When $\varepsilon > 0$ in (1.1) is fixed, e.g. $\varepsilon = 1$, corresponding to the $\mathcal{O}(1)$ -speed of light regime, a surge of analysis and numerics results have been reported in literatures. For instance, the Cauchy problem was considered in [2, 10, 23, 24, 38]. In particular, global existence of solutions was established in [10] for $F(u) \geq 0$ (defocusing case); and possible blow-up was shown in [2] for $F(u) < 0$ (focusing case). For more results in this regime, we refer the readers to [29, 33, 36, 40] and references given therein. Along the numerical aspect, many numerical schemes have been proposed in literatures. The classical numerical methods are the standard finite difference time domain methods including energy conservative, semi-implicit and explicit finite difference discretizations [1, 15, 25, 32, 41] and some other approaches such as finite element or spectral discretization in space coupled with appropriate time integrator, like standard finite difference or Gautschi-type exponential integrator [6, 12, 14, 42]. For comparisons of different numerical methods, we refer the readers to [6, 22, 32].

Over the past decade, more attentions have been paid to the regime $0 < \varepsilon \ll 1$ in (1.1), which corresponds to the nonrelativistic limit or the speed of light goes to infinity. In this regime, the analysis and efficient simulation are mathematically rather complicated

issues; see, e.g. [6, 26–28, 30, 31, 43]. The analysis difficulty mainly lies in the unbounded energy $E(t)$ in (1.2) when $\varepsilon \rightarrow 0$. Recently, Machihara *et al.* [27] studied the nonrelativistic limit in the energy space, and Masmoudi *et al.* [28] analyzed the similar limit in a strong topology of the energy space. For more recent progresses made to understand this limit, we refer to [30, 31, 43]. Based on their results, the solution propagates waves with wavelength of $\mathcal{O}(\varepsilon^2)$ and $\mathcal{O}(1)$ in time and, respectively, in space when $0 < \varepsilon \ll 1$. The highly oscillatory nature in time provides severe numerical burdens, making the computation in the nonrelativistic limit regime is extremely challenging. Even for the stable numerical discretizations (or under stability restrictions on meshing strategies) the approximations may come out completely wrong unless the temporal oscillation is fully resolved numerically, i.e., using many time steps per wavelength of $\mathcal{O}(\varepsilon^2)$. Due to the non-oscillation fact in space, there would be no spatial meshing constraint, i.e. mesh size Δx is independent of ε .

Since the nonrelativistic limit behavior of KG equation is largely unknown, a temporal discretization that performs well (allowing largest possible time step for a given ε) is of great importance for the investigation and prediction of the limits. In the recent work [6], frequently used finite difference time discretization and a Gautschi-type exponential integrator were analyzed rigorously by paying particular attention on how their error bounds depend on small ε . These results show that for those finite difference integration, one needs the time step $\Delta t = \mathcal{O}(\varepsilon^3)$ in order to guarantee ‘correct’ approximations for ε small. On the other hand, the Gautschi-type exponential integrator was well demonstrated in literatures that it has favorable properties compared to standard time integrators for oscillatory second-order differential equations [17, 18, 20, 21]. When the Gautschi-type integrator is applied to the KG equation with small ε , it was proven in [6] that, in order to guarantee ‘correct’ approximations, one needs the meshing strategy constraint that $\Delta t = \mathcal{O}(\varepsilon^2)$ for nonlinear problem. Although the Gautschi-type integrator allows the time step one order of magnitude larger than the finite difference integration, it is an interesting problem to find a temporal discretization which can further loosen the meshing strategy restriction when $0 < \varepsilon \ll 1$.

In this work, we study a time-splitting integrator (or so-called split-step method), coupled with Fourier pseudospectral (TSFP) discretization in space, for the nonlinear KG equation (1.1). The starting point is to rewrite the second-order equation as an equivalent but simple form of first-order system in time. The key ideas of the method are: (i) split the evolution system in a proper way such that the nonlinear subproblem can be integrated exactly in time space; (ii) solve the linear subproblem in phase space by applying the Fourier pseudospectral approximation to the spatial derivative and integrating the equations (which is a first-order linear ODE system) about the Fourier coefficients exactly. We shall study simultaneously this TSFP method in various regimes, ranging from the smooth for $\varepsilon = \mathcal{O}(1)$ to the highly oscillatory for $0 < \varepsilon \ll 1$. Moreover, an alternative time-splitting scheme can also be constructed based on a classical first-order reformulation, which results in a relativistic (i.e. fractional) nonlinear Schrödinger (NLS) equation. Numerical analysis for the fractional NLS is a difficult issue; however, the results ob-

tained in this work, either rigorously or numerically, still hold for the splitting integrator applied to the fractional NLS, in the view of a fact that each splitting step in the proposed TSFP indeed solves an equivalent subproblem based on the classical reformulation, and vice versa.

The time-splitting schemes for evolution equations can even date back to 1970s [19]; however, few results are available so far when they are applied to the KG equation (1.1), even with $\varepsilon = \mathcal{O}(1)$. The first goal of this paper is to establish rigorous error estimates of the proposed TSFP for solving KG (1.1) with fixed $\varepsilon = \mathcal{O}(1)$. This is achieved thanks to the observation that the TSFP coincides with a type of trigonometric integrator Fourier pseudospectral discretization. On the other hand, although time-splitting schemes are widely used to compute the solutions to NLS (see, e.g. [4,44]) and especially successful for the semiclassical NLS whose solutions exhibit spatial-temporal oscillations (see [7, 8]), it does not give any clue to their performance for KG equation in highly oscillatory regime. Thus, the second purpose of this work is to investigate numerically the performance of the TSFP for solving KG (1.1) when $\varepsilon \rightarrow 0$, with special attentions paid to how does the convergence affected by ε . Our extensive numerical experiments show that the TSFP has uniform spectral accuracy in space, and possesses a similar temporal convergence regime as the Gautschi-type integrator for ε vanishing. However, within the convergence regime, the temporal discretization error bound for TSFP is suggested to be $\mathcal{O}(\varepsilon^{-2}\Delta t^2)$, whereas it has been shown both rigorously and numerically the temporal error for Gautschi-type integrator is $\mathcal{O}(\varepsilon^{-4}\Delta t^2)$ [6]. Therefore, the time-splitting pseudospectral discretization offers compelling better approximations over other schemes, especially in the nonrelativistic limit regime.

The rest paper is organized as follows. In Section 2, we discuss the derivation of the TSFP method, to which rigorous convergence results are given in Section 3. In Section 4, extensive numerical studies are carried out to investigate the accuracy and temporal resolution capacity of TSFP. Comparisons between TSFP and Gautschi-type schemes are also made. Finally, some concluding remarks are drawn in Section 5. Throughout this paper, we adopt the standard Sobolev spaces and the corresponding norms, and also denote $A \lesssim B$ to represent that there exists a generic constant $C > 0$, which is independent of Δt , Δx and ε if any, such that $|A| \leq CB$.

2 Numerical methods

In this section, we shall first present the time-splitting Fourier pseudospectral (TSFP) method, which is based on the application of Fourier pseudospectral approximation to spatial derivative followed by the time-splitting integrator to time discretization. Then, two alternative approaches to the TSFP will be discussed. Also, the Gautschi-type exponential integrator Fourier pseudospectral scheme will be recalled for comparison reasons.

For simplicity of notations, we only present the methods in one space dimension (1D), and generalization to higher dimensions are straightforward due to the tensor product

grids. In practice, we truncate the whole space problem onto an interval $\Omega = (a, b)$ with periodic boundary conditions. In 1D, the KG equation (1.1) with periodic boundary conditions collapses to

$$\varepsilon^2 \partial_{tt} u(x, t) - \partial_{xx} u(x, t) + \frac{1}{\varepsilon^2} u(x, t) + f(u(x, t)) = 0, \quad a < x < b, \quad t > 0, \quad (2.1a)$$

$$u(a, t) = u(b, t), \quad \partial_x u(a, t) = \partial_x u(b, t), \quad t \geq 0, \quad (2.1b)$$

$$u(x, 0) = \phi_1(x), \quad \partial_t u(x, 0) = \frac{1}{\varepsilon^2} \phi_2(x), \quad a \leq x \leq b. \quad (2.1c)$$

Such boundary conditions are inspired by the physical backgrounds as well as most studies in literatures; see, e.g. [6] and references given therein.

2.1 Time-splitting Fourier pseudospectral (TSFP) discretization

As a preparatory step, we begin by recalling the construction of a time-splitting (or split-step) integrator for a general evolution system in the form:

$$\partial_t \mathbf{y} = \Phi(\mathbf{y}) = \mathcal{A}\mathbf{y} + \mathcal{B}\mathbf{y}, \quad (2.2)$$

where the mapping $\Phi(\mathbf{y})$ is usually a nonlinear operator and the decoupling $\Phi(\mathbf{y}) = \mathcal{A}\mathbf{y} + \mathcal{B}\mathbf{y}$ (or called operator-splitting) can be quite arbitrary; in particular, \mathcal{A} and \mathcal{B} can be two non-commutative operators. With a given time step $\Delta t > 0$, let $t_n = n\Delta t$, $n = 0, 1, 2, \dots$, and \mathbf{y}^n be the approximation of $\mathbf{y}(t_n)$. A commonly used second-order time-splitting integrator for (2.2), $\mathbf{y}^{n+1} = [\Phi_2(\Delta t)](\mathbf{y}^n)$, can be constructed due to the Strang formula [39],

$$\mathbf{y}^{(1)} = \exp\left(\frac{1}{2}\Delta t \mathcal{A}\right) \mathbf{y}^n, \quad \mathbf{y}^{(2)} = \exp(\Delta t \mathcal{B}) \mathbf{y}^{(1)}, \quad \mathbf{y}^{n+1} = \exp\left(\frac{1}{2}\Delta t \mathcal{A}\right) \mathbf{y}^{(2)}, \quad (2.3)$$

which is explicit and symmetric, i.e., $\Phi_2(\Delta t)\Phi_2(-\Delta t) = 1$. A fourth-order symplectic time integrator for (2.2), $\mathbf{y}^{n+1} = [\Phi_4(\Delta t)](\mathbf{y}^n)$, is constructed as follows (cf. [9, 45]):

$$\Phi_4(\Delta t) = \Phi_2(\omega\Delta t)\Phi_2((1-2\omega)\Delta t)\Phi_2(\omega\Delta t), \quad (2.4)$$

where

$$\omega = \frac{1}{3} \left(2 + 2^{1/3} + 2^{-1/3} \right). \quad (2.5)$$

Clearly, the above fourth-order integrator is still explicit and time reversible. It is also possible to construct higher-order symplectic integrators (cf. [45]). In general, the operators \mathcal{A} and \mathcal{B} may be interchanged without affecting the accuracy order of the method.

Introducing $v(x,t) = \partial_t u(x,t)$, then (2.1) is equivalent to the following first-order-in-time system,

$$\partial_t u(x,t) - v(x,t) = 0, \quad a < x < b, \quad t > 0, \tag{2.6a}$$

$$\varepsilon^2 \partial_t v(x,t) - \partial_{xx} u(x,t) + \frac{1}{\varepsilon^2} u(x,t) + f(u(x,t)) = 0, \quad a < x < b, \quad t > 0, \tag{2.6b}$$

$$u(a,t) = u(b,t), \quad \partial_x u(a,t) = \partial_x u(b,t), \quad v(a,t) = v(b,t), \quad t \geq 0, \tag{2.6c}$$

$$u(x,0) = \phi_1(x), \quad v(x,0) = \frac{1}{\varepsilon^2} \phi_2(x), \quad a \leq x \leq b. \tag{2.6d}$$

We now rewrite the system (2.6a)-(2.6b) in the form of (2.2) with

$$\mathbf{y} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \mathcal{A} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ -\varepsilon^{-2} f(u) \end{pmatrix}, \quad \mathcal{B} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ \varepsilon^{-2} \partial_{xx} u - \varepsilon^{-4} u \end{pmatrix}. \tag{2.7}$$

Thus, the key to an efficient implementation of the time-splitting integrator $\Phi_2(\Delta t)$ or $\Phi_4(\Delta t)$ is to solve efficiently the following two subproblems:

$$\partial_t u(x,t) = 0, \quad a < x < b, \quad t > 0, \tag{2.8a}$$

$$\partial_t v(x,t) + \frac{1}{\varepsilon^2} f(u(x,t)) = 0, \quad a < x < b, \quad t > 0, \tag{2.8b}$$

and

$$\partial_t u(x,t) - v(x,t) = 0, \quad a < x < b, \quad t > 0, \tag{2.9a}$$

$$\partial_t v(x,t) - \frac{1}{\varepsilon^2} \partial_{xx} u(x,t) + \frac{1}{\varepsilon^4} u(x,t) = 0, \quad a < x < b, \quad t > 0, \tag{2.9b}$$

$$u(a,t) = u(b,t), \quad \partial_x u(a,t) = \partial_x u(b,t), \quad v(a,t) = v(b,t), \quad t \geq 0. \tag{2.9c}$$

The solutions to (2.8) are trivial by noting that (2.8a) leaves $u(x,t)$ invariant in t and therefore (2.8b) can be integrated exactly, i.e., for $t \geq t_s$ (t_s any given time),

$$u(x,t) = u(x,t_s), \quad v(x,t) = v(x,t_s) - \frac{1}{\varepsilon^2} (t - t_s) f(u(x,t_s)), \quad a < x < b, \quad t \geq t_s. \tag{2.10}$$

Now, the key remains to find an efficient and accurate method for (2.9). We shall solve (2.9) below in phase space by applying the Fourier spectral or pseudospectral approximation in space discretization; and in particular, the equations about the Fourier coefficients are linear ODEs which can be solved exactly.

Choose mesh size $\Delta x = (b - a) / M$ with M being an even positive integer, and denote grid points as $x_j := a + j\Delta x, j = 0, 1, \dots, M$. Let $X_M := \text{span}\{e^{i\mu_l(x-a)}, \mu_l = \frac{2\pi l}{b-a}, x \in \Omega, -M/2 \leq l \leq M/2 - 1\}$ and $Y_M := \{w = (w_0, w_1, \dots, w_M) \in \mathbb{R}^{M+1} : w_0 = w_M\}$. For a general periodic function $w(x)$ on $\bar{\Omega} = [a, b]$ and a vector $w \in Y_M$, let $P_M : L^2(\Omega) \rightarrow X_M$ be the standard

L^2 - projection operator onto X_M , $I_M: C(\Omega) \rightarrow X_M$ and $I_M: Y_M \rightarrow X_M$ be the trigonometric interpolation operator [37], i.e.

$$(P_M w)(x) = \sum_{l=-M/2}^{M/2-1} \hat{w}_l e^{i\mu_l(x-a)}, \quad (I_M w)(x) = \sum_{l=-M/2}^{M/2-1} \tilde{w}_l e^{i\mu_l(x-a)}, \quad x \in \Omega, \quad (2.11)$$

where

$$\hat{w}_l = \frac{1}{b-a} \int_a^b w(x) e^{-i\mu_l(x-a)} dx, \quad \tilde{w}_l = \frac{1}{M} \sum_{j=0}^{M-1} w_j e^{-i\mu_l(x_j-a)}, \quad (2.12)$$

with w_j interpreted as $w(x_j)$ for a function w . It is easy to check that P_M and I_M are identity operators on X_M .

The Fourier spectral method for (2.9) is to find $u_M(x, t) \in X_M$ and $v_M(x, t) \in X_M$ (cf. [37]), i.e.,

$$u_M(x, t) = \sum_{l=-M/2}^{M/2-1} \hat{u}_l(t) e^{i\mu_l(x-a)}, \quad v_M(x, t) = \sum_{l=-M/2}^{M/2-1} \hat{v}_l(t) e^{i\mu_l(x-a)}, \quad x \in \Omega, \quad t \geq 0, \quad (2.13)$$

such that

$$\partial_t u_M(x, t) - v_M(x, t) = 0, \quad x \in \Omega, \quad t \geq 0, \quad (2.14a)$$

$$\partial_t v_M(x, t) - \frac{1}{\epsilon^2} \partial_{xx} u_M(x, t) + \frac{1}{\epsilon^4} u_M(x, t) = 0, \quad x \in \Omega, \quad t \geq 0. \quad (2.14b)$$

Plugging (2.13) into (2.14), noticing the orthogonality of Fourier functions, we find

$$\frac{d}{dt} \hat{u}_l(t) - \hat{v}_l(t) = 0, \quad \frac{d}{dt} \hat{v}_l(t) + \beta_l^2 \hat{u}_l(t) = 0, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1, \quad t \geq 0.$$

where $\beta_l = \epsilon^{-2} \sqrt{\epsilon^2 \mu_l^2 + 1}$. The above system is a first-order linear ODE system, whose analytical solutions can be obtained directly, i.e. for $t \geq t_s$ (t_s any given time) and $l = -M/2, \dots, M/2 - 1$,

$$\hat{u}_l(t) = \cos(\beta_l(t - t_s)) \hat{u}_l(t_s) + \frac{\sin(\beta_l(t - t_s))}{\beta_l} \hat{v}_l(t_s), \quad (2.15a)$$

$$\hat{v}_l(t) = -\beta_l \sin(\beta_l(t - t_s)) \hat{u}_l(t_s) + \cos(\beta_l(t - t_s)) \hat{v}_l(t_s). \quad (2.15b)$$

The above procedure for solving (2.9) is not suitable in practice due to the difficulty in evaluating the integrals in (2.12). Thus, we shall approximate the integrals in (2.12) by a quadrature rule on the grids $\{x_j\}_{j=0}^M$, i.e., replacing the projections by the interpolations, which refers to the Fourier pseudospectral approximation [37].

For $j = 0, 1, \dots, M$ and $n = 0, 1, \dots$, let u_j^n and v_j^n be the approximations of $u(x_j, t_n)$ and $v(x_j, t_n)$, denote by u^n and v^n the solution vectors with components u_j^n and v_j^n , and choose

$u_j^0 = \phi_1(x_j)$ and $v_j^0 = \phi_2(x_j)/\varepsilon^2$. Then the second-order time-splitting Fourier pseudospectral (TSFP) discretization for the 1D KG equation (2.6) is given by

$$u_j^{n,+} = u_j^n, \quad v_j^{n,+} = v_j^n - \frac{\Delta t}{2\varepsilon^2} f(u_j^n), \tag{2.16a}$$

$$u_j^{n+1,-} = \mathcal{L}_u(\Delta t, u^{n,+}, v^{n,+})_j, \quad v_j^{n+1,-} = \mathcal{L}_v(\Delta t, u^{n,+}, v^{n,+})_j, \tag{2.16b}$$

$$u_j^{n+1} = u_j^{n+1,-}, \quad v_j^{n+1} = v_j^{n+1,-} - \frac{\Delta t}{2\varepsilon^2} f(u_j^{n+1,-}). \tag{2.16c}$$

Here, $\mathcal{L}_u(\tau, U, V)_j$ and $\mathcal{L}_v(\tau, U, V)_j$ ($j = 0, 1, \dots, M$) are computed from any $\tau \in \mathbb{R}$, $U = (U_0, U_1, \dots, U_M)^T$ and $V = (V_0, V_1, \dots, V_M)^T$:

$$\begin{aligned} \mathcal{L}_u(\tau, U, V)_j &= \sum_{l=-M/2}^{M/2-1} \left[\cos(\beta_l \tau) \tilde{U}_l + \frac{\sin(\beta_l \tau)}{\beta_l} \tilde{V}_l \right] e^{i\mu_l(x_j-a)}, \\ \mathcal{L}_v(\tau, U, V)_j &= \sum_{l=-M/2}^{M/2-1} \left[-\beta_l \sin(\beta_l \tau) \tilde{U}_l + \cos(\beta_l \tau) \tilde{V}_l \right] e^{i\mu_l(x_j-a)}, \\ \tilde{U}_l &= \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-i\mu_l(x_j-a)}, \quad \tilde{V}_l = \frac{1}{M} \sum_{j=0}^{M-1} V_j e^{-i\mu_l(x_j-a)}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \end{aligned}$$

A fourth-order TSFP discretization for (2.6) can be constructed according to (2.4) in a similar way. We omit the details here for brevity.

The time discretization error of the TSFP discretization is only the splitting error, which is second/fourth order in Δt . Moreover, TSFP is explicit, time symmetric and easy to extend to 2D and 3D. The memory cost is $\mathcal{O}(M)$ and computational load per time step is $\mathcal{O}(M \ln M)$ thanks to FFT.

Remark 2.1. Clearly, (2.16a) and (2.16c) imply that $u^{n+1,+} = u^{n+1} = u^{n+1,-}$, so the TSFP (2.16) can be implemented according to

$$u_j^{n+1} = \mathcal{L}_u(\Delta t, u^n, v^{n,+})_j, \quad v_j^{n+1,-} = \mathcal{L}_v(\Delta t, u^n, v^{n,+})_j, \tag{2.17a}$$

$$v_j^{n+1,+} = v_j^{n+1,-} - \frac{\Delta t}{\varepsilon^2} f(u_j^{n+1}). \tag{2.17b}$$

Thus, it is not necessary to output v^{n+1} unless it is of interests.

Remark 2.2. Note that for the special case $f(u) = 0$, i.e., the linear problem, the TSFP collapses to the following one-step formula:

$$u_j^{n+1} = \mathcal{L}_u((n+1)\Delta t, u^0, v^0)_j, \quad v_j^{n+1} = \mathcal{L}_v((n+1)\Delta t, u^0, v^0)_j, \quad n = 0, 1, \dots, \tag{2.18}$$

thereby introducing no time discretization error.

2.2 TSFP discretization based on relativistic NLS reformulation

Classical reformulation of the KG equation into a first-order-in-time system results in a coupled system of relativistic (i.e. fractional) NLS equations [27,28]. Also, if the unknown u in (1.1a) is real-valued, then the system collapses to a single relativistic NLS, which is the case considered in this work (see Remark 2.3 below for the discussion about complex-valued unknown). Here, we shall discuss a splitting method based on this fractional NLS, and show it is essentially equivalent to the TSFP discussed in the last subsection.

Define the positive operator

$$\langle \Delta \rangle_\varepsilon = \varepsilon^{-1} \sqrt{-\varepsilon^2 \Delta + 1},$$

then the KG equation (1.1a) can be written as

$$\varepsilon^2 \partial_{tt} u(\mathbf{x}, t) + \langle \Delta \rangle_\varepsilon^2 u(\mathbf{x}, t) + f(u(\mathbf{x}, t)) = 0. \quad (2.19)$$

Introducing

$$\psi(\mathbf{x}, t) = u(\mathbf{x}, t) - i\varepsilon \langle \Delta \rangle_\varepsilon^{-1} v(\mathbf{x}, t), \quad (2.20)$$

with $v(\mathbf{x}, t) = \partial_t u(\mathbf{x}, t)$. Substituting (2.20) into (1.1a), a simple calculation shows that (1.1a) is equivalent to the following relativistic NLS:

$$i\partial_t \psi + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon \psi + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon^{-1} f\left(\frac{1}{2}(\psi + \bar{\psi})\right) = 0. \quad (2.21)$$

Here and after, \bar{g} denotes the complex conjugate of g . Now, a TSFP can be readily constructed for solving (2.21), which consists of a sequence of solving the following two subproblems:

$$i\partial_t \psi + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon^{-1} f\left(\frac{1}{2}(\psi + \bar{\psi})\right) = 0, \quad (2.22)$$

and

$$i\partial_t \psi + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon \psi = 0. \quad (2.23)$$

First, the fractional differential operator $\langle \Delta \rangle_\varepsilon$ can be discretized by applying Fourier spectral or pseudospectral discretization in space, provided that the periodic boundary conditions are posed; see, e.g. [5] for a detailed discussion about a similar operator. Next, adding (2.22) to its complex conjugate, and taking the imaginary part, we can see it leaves $(\psi + \bar{\psi})$ unchanged in time, and therefore (2.22) is exactly integrable in time. For solving (2.23), one can integrate the linear ODEs about Fourier coefficients exactly in phase space, which is quite similar to the solver to (2.9) and we omit the details here for brevity.

To show the equivalence of the TSFP method for (2.6) and the one for (2.21) with periodic boundary conditions, we only need to show the subproblem (2.8) (subproblem (2.9)) coincides with (2.22) ((2.23)) in terms of finding u and v , since in both cases the subproblems are integrated exactly in time. This is easy to check due to a simple calculation.

Substituting (2.20) into (2.22) or (2.8), noting that both u and v are real-valued, it is ready to check that (2.22) essentially coincides with (2.8), and vice versa. The equivalence of (2.9) and (2.23) can be checked in a similar way. In view of such equivalence, In the sequel we shall focus on the convergence and numerical resolution for oscillatory solutions of the TSFP method (2.16).

Remark 2.3. If the KG equation (1.1a) is considered for complex-valued unknown $u(\mathbf{x}, t)$ with nonlinear term $f(u)$ satisfying gauge invariance, introducing

$$\psi_+(\mathbf{x}, t) = u(\mathbf{x}, t) - i\varepsilon \langle \Delta \rangle_\varepsilon^{-1} v(\mathbf{x}, t), \quad \psi_-(\mathbf{x}, t) = \bar{u}(\mathbf{x}, t) - i\varepsilon \langle \Delta \rangle_\varepsilon^{-1} \bar{v}(\mathbf{x}, t), \quad (2.24)$$

then (1.1a) is equivalent to a form of coupled relativistic NLS [27, 28]:

$$i\partial_t \psi_+ + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon \psi_+ + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon^{-1} f\left(\frac{1}{2}(\psi_+ + \bar{\psi}_-)\right) = 0, \quad (2.25a)$$

$$i\partial_t \psi_- + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon \psi_- + \frac{1}{\varepsilon} \langle \Delta \rangle_\varepsilon^{-1} f\left(\frac{1}{2}(\bar{\psi}_+ + \psi_-)\right) = 0. \quad (2.25b)$$

For this coupled relativistic NLS, a TSFP can be constructed in a similar manner as before, which again coincides with the TSFP applied to (2.6).

2.3 Alternative approach to TSFP from trigonometric integrator

As a fact pointed out in [21, Section XIII.1.3], for the first-order-in-time evolution equations, the split-step method is reduced to a trigonometric integrator proposed by P. Deuffhard [16]. Here, we discuss an alternative approach to derive the proposed TSFP (2.16) via using the Deuffhard-type trigonometric integrator with Fourier pseudospectral discretization in space, which in consequence gives rise to a simple way to analyze the convergence of the splitting method.

Similar as the solver to (2.9), we seek for $u_M(x, t), v_M(x, t) \in X_M$ defined in (2.13) as spatial approximations to solutions $u(x, t)$ and $v(x, t)$, respectively. Plugging $u_M(x, t)$ into (2.1) and applying the L^2 -projection, we get

$$\varepsilon^2 \partial_{tt} u_M - \partial_{xx} u_M + \frac{1}{\varepsilon^2} u_M + P_M f(u_M) = 0.$$

Noticing the orthogonality of Fourier bases, we get

$$\varepsilon^2 \frac{d^2}{dt^2} \hat{u}_l(t) + \mu_l^2 \hat{u}_l(t) + \frac{1}{\varepsilon^2} \hat{u}_l(t) + \widehat{f(u_M)}_l(t) = 0, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1, \quad t > 0. \quad (2.26)$$

By using the variation-of-constant formula and noting $v = \partial_t u$, we get for any $t \geq t_s \geq 0$,

$$\hat{u}_l(t) = \cos(\beta_l(t - t_s)) \hat{u}_l(t_s) + \frac{\sin(\beta_l(t - t_s))}{\beta_l} \hat{v}_l(t_s) - \int_{t_s}^t \frac{\sin(\beta_l(t - s))}{\varepsilon^2 \beta_l} \widehat{f(u_M)}_l(s) ds. \quad (2.27)$$

Taking derivative with respect to t on both sides of (2.27), we get

$$\begin{aligned} \widehat{v}_l(t) = & -\beta_l \sin(\beta_l(t-t_s)) \widehat{u}_l(t_s) + \cos(\beta_l(t-t_s)) \widehat{v}_l(t_s) \\ & - \int_{t_s}^t \frac{\cos(\beta_l(t-s))}{\varepsilon^2} \widehat{f(u_M)}_l(s) ds. \end{aligned} \tag{2.28}$$

Applying the standard trapezoidal rule to the two unknown integrations in (2.27) and (2.28), we get

$$\begin{aligned} \widehat{u}_l(t) \approx & \cos(\beta_l(t-t_s)) \widehat{u}_l(t_s) + \frac{\sin(\beta_l(t-t_s))}{\beta_l} \widehat{v}_l(t_s) - \frac{t-t_s}{2\varepsilon^2 \beta_l} \sin(\beta_l(t-t_s)) \widehat{f(u_M)}_l(t_s), \\ \widehat{v}_l(t) \approx & -\beta_l \sin(\beta_l(t-t_s)) \widehat{u}_l(t_s) + \cos(\beta_l(t-t_s)) \widehat{v}_l(t_s) \\ & - \frac{t-t_s}{2\varepsilon^2} \left[\cos(\beta_l(t-t_s)) \widehat{f(u_M)}_l(t_s) + \widehat{f(u_M)}_l(t) \right]. \end{aligned}$$

Replacing the above Fourier spectral approximations by pseudospectral discretization, we obtain the following Deuffhard-type integrator Fourier pseudospectral (DIFP) method. For $j=0,1,\dots,M$ and $n=0,1,\dots$, choosing $u_j^0 = \phi_1(x_j)$ and $v_j^0 = \phi_2(x_j)/\varepsilon^2$, then

$$u_j^{n+1} = \sum_{l=-M/2}^{M/2-1} \widehat{u}_l^{n+1} e^{i\mu_l(x_j-a)}, \quad v_j^{n+1} = \sum_{l=-M/2}^{M/2-1} \widehat{v}_l^{n+1} e^{i\mu_l(x_j-a)}, \quad n \geq 0, \tag{2.29a}$$

$$\widehat{u}_l^{n+1} = \cos(\beta_l \Delta t) \widehat{u}_l^n + \frac{\sin(\beta_l \Delta t)}{\beta_l} \widehat{v}_l^n - \frac{\Delta t}{2\varepsilon^2 \beta_l} \sin(\beta_l \Delta t) \widetilde{f}_l^n, \tag{2.29b}$$

$$\widehat{v}_l^{n+1} = -\beta_l \sin(\beta_l \Delta t) \widehat{u}_l^n + \cos(\beta_l \Delta t) \widehat{v}_l^n - \frac{\Delta t}{2\varepsilon^2} \left[\cos(\beta_l \Delta t) \widetilde{f}_l^n + \widetilde{f}_l^{n+1} \right], \tag{2.29c}$$

where,

$$\widetilde{u}_l^n = \frac{1}{M} \sum_{j=0}^{M-1} u_j^n e^{-i\mu_l(x_j-a)}, \quad \widetilde{v}_l^n = \frac{1}{M} \sum_{j=0}^{M-1} v_j^n e^{-i\mu_l(x_j-a)}, \quad \widetilde{f}_l^n = \frac{1}{M} \sum_{j=0}^{M-1} f(u_j^n) e^{-i\mu_l(x_j-a)}.$$

A simple calculation shows that

Proposition 2.1. The TSFP method (2.16) coincides with the DIFP method (2.29).

Proof. First, the initial choices of TSFP and DIFP methods are the same, i.e. $u_j^0 = \phi_1(x_j)$, $v_j^0 = \phi_2(x_j)/\varepsilon^2$. From the TSFP method (2.16), plugging (2.16a) into (2.16b) and noticing from (2.16c) that $u^{n+1} = u^{n+1,-}$, we get

$$\widehat{u}_l^{n+1} = \cos(\beta_l \Delta t) \widehat{u}_l^n + \frac{\sin(\beta_l \Delta t)}{\beta_l} \left(\widehat{v}_l^n - \frac{\Delta t}{2\varepsilon^2} \widetilde{f}_l^n \right), \tag{2.30}$$

which is indeed (2.29b) in the DIFP method, and

$$\widehat{v}_l^{n+1,-} = -\beta_l \sin(\beta_l \Delta t) \widehat{u}_l^n + \cos(\beta_l \Delta t) \left(\widehat{v}_l^n - \frac{\Delta t}{2\varepsilon^2} \widetilde{f}_l^n \right). \tag{2.31}$$

Plugging (2.31) into (2.16c), we are led to (2.29c) in DIFP method, which completes the proof. \square

2.4 Gautschi-type integrator Fourier pseudospectral method

For benchmark comparisons, we recall another numerical method, a Gautschi-type trigonometric integrator Fourier pseudospectral (GIFP) discretization for (2.6), which was proposed and analyzed rigorously in [6]. The detailed method is as follows:

$$u_j^{n+1} = \sum_{l=-M/2}^{M/2-1} \tilde{u}_l^{n+1} e^{i\mu_l(x_j-a)}, \quad v_j^{n+1} = \sum_{l=-M/2}^{M/2-1} \tilde{v}_l^{n+1} e^{i\mu_l(x_j-a)}, \quad (2.32)$$

for $j=0,1,\dots,M, n=0,1,\dots$, with,

$$\tilde{u}_l^1 = \left[\cos(\beta_l^0 \Delta t) + \frac{\alpha^0 (1 - \cos(\beta_l^0 \Delta t))}{(\varepsilon \beta_l^0)^2} \right] \tilde{u}_l^0 + \frac{\sin(\beta_l^0 \Delta t)}{\beta_l^0} \tilde{v}_l^0 + \frac{\cos(\beta_l^0 \Delta t) - 1}{(\varepsilon \beta_l^0)^2} \tilde{f}_l^0, \quad (2.33)$$

$$\tilde{v}_l^1 = -\beta_l \sin(\beta_l \Delta t) \tilde{v}_l^0 + \cos(\beta_l \Delta t) \tilde{v}_l^0 - \frac{\sin(\beta_l \Delta t)}{\varepsilon^2 \beta_l} \tilde{f}_l^0, \quad (2.34)$$

$$\tilde{u}_l^{n+1} = -\tilde{u}_l^{n-1} + 2 \left[\cos(\beta_l^n \Delta t) + \frac{\alpha^n (1 - \cos(\beta_l^n \Delta t))}{(\varepsilon \beta_l^n)^2} \right] \tilde{u}_l^n + \frac{2(\cos(\beta_l^n \Delta t) - 1)}{(\varepsilon \beta_l^n)^2} \tilde{f}_l^n, \quad (2.35)$$

$$\tilde{v}_l^{n+1} = \tilde{v}_l^{n-1} - 2\beta_l \sin(\beta_l \Delta t) \tilde{u}_l^n - 2 \frac{\sin(\beta_l \Delta t)}{\varepsilon^2 \beta_l} \tilde{f}_l^n, \quad n \geq 1. \quad (2.36)$$

Here,

$$\beta_l^n = \frac{1}{\varepsilon^2} \sqrt{1 + \varepsilon^2 (\mu_l^2 + \alpha^n)}, \quad \alpha^n = \max \left\{ \alpha^{n-1}, \max_{u_j^n \neq 0} f(u_j^n) / u_j^n \right\}, \quad \alpha^{-1} = 0, \quad n = 0, 1, \dots,$$

where α^n is introduced to ensure the unconditional stability.

As it is shown in [6], the GIFP method is spectrally accurate in space, brings in no time discretization error for linear problem and is second-order accurate in time for nonlinear problem. Moreover, in order to capture the $\mathcal{O}(\varepsilon^2)$ -oscillation in time when $0 < \varepsilon \ll 1$, the meshing strategy constraint for GIFP method is $\Delta t = \mathcal{O}(\varepsilon^2)$ and $\Delta x = \mathcal{O}(1)$ for nonlinear problem, and $\Delta t = \mathcal{O}(1)$ and $\Delta x = \mathcal{O}(1)$ for linear problem.

3 Convergence analysis of TSFP

In this section, we will establish rigorously the error bounds of the TSFP method (2.16) in the energy space $H^1 \times L^2$ for fixed $\varepsilon = \varepsilon_0 = \mathcal{O}(1)$ (the $\mathcal{O}(1)$ -speed of light regime). Without loss of generality and for simplicity of notations, we set $\varepsilon = 1$ in this section. The rigorous arguments are achieved thanks to its equivalent formulation, i.e. the DIFP method (2.29).

3.1 Main results on error bounds in energy space

Let T_* be the maximum existence time for the solutions to the KG (2.1). In order to establish the error estimates for the TSFP method, we make the following assumptions on the nonlinearity and the exact solutions: for $0 < T < T_*$,

$$f \in C^k(\mathbb{R}), \quad u \in C\left([0, T]; W^{1,\infty} \cap H_p^{m_0+1}\right) \cap C^1\left([0, T]; W^{1,4} \cap H_p^{m_0}\right) \cap C^2\left([0, T]; H^1\right), \quad (A)$$

for some integer $k, m_0 \geq 2$. Under assumption (A), we let

$$m = \min\{k, m_0\}, \quad K_1 = \|u\|_{L^\infty([0, T]; L^\infty(\Omega) \cap H^1(\Omega))}, \quad K_2 = \|\partial_t u\|_{L^\infty([0, T]; L^2(\Omega))}.$$

Denote the trigonometric interpolations of numerical solutions as $u_I^n(x) := I_M(u^n)(x)$, $v_I^n(x) := I_M(v^n)(x)$, and define the ‘error’ functions as

$$e^n(x) := u(x, t_n) - u_I^n(x), \quad \dot{e}^n(x) := \partial_t u(x, t_n) - v_I^n(x), \quad x \in \Omega, \quad n \geq 0,$$

then we have the following main error estimate result:

Theorem 3.1. *Let u^n and v^n be the numerical approximations obtained from the TSFP method (2.16). Under the assumption (A), there exist two constants $\tau_0, h_0 > 0$, independent of Δt (or n) and Δx , such that for any $0 < \Delta t < \tau_0$, $0 < \Delta x < h_0$,*

$$\|e^n\|_{H^1} + \|\dot{e}^n\|_{L^2} \lesssim \Delta t^2 + \Delta x^m, \quad n = 0, \dots, \frac{T}{\Delta t}, \quad (3.1a)$$

$$\|u_I^n\|_{H^1} \leq K_1 + 1, \quad \|v_I^n\|_{L^2} \leq K_2 + 1, \quad \|u^n\|_{l^\infty} \leq K_1 + 1. \quad (3.1b)$$

3.2 Proof of Theorem 3.1

Thanks to the DIFP formulism, the error estimates can be done in analogous lines as [6] by means of ‘energy method’. We first introduce the following notations.

Suppose u, v are exact solutions to KG equation (2.6). Denote the L^2 -projected solutions as[†]

$$\begin{aligned} u_M(x, t) &:= P_M(u(x, t)) = \sum_{l=-M/2}^{M/2-1} \hat{u}_l(t) e^{i\mu_l(x-a)}, \\ v_M(x, t) &:= P_M(v(x, t)) = \sum_{l=-M/2}^{M/2-1} \hat{v}_l(t) e^{i\mu_l(x-a)}, \end{aligned} \quad x \in \Omega, \quad t \geq 0, \quad (3.2)$$

and the projected error functions as

$$\begin{aligned} e_M^n(x) &:= P_M(e^n(x)) = \sum_{l=-M/2}^{M/2-1} \hat{e}_l^n e^{i\mu_l(x-a)}, \\ \dot{e}_M^n(x) &:= P_M(\dot{e}^n(x)) = \sum_{l=-M/2}^{M/2-1} \hat{\dot{e}}_l^n e^{i\mu_l(x-a)}, \end{aligned} \quad n = 0, \dots, \frac{T}{\Delta t}. \quad (3.3)$$

[†]We remark that the u_M defined in (3.2) is different with the one in spectral method formulation, since for the later case $u_M \approx P_M u$, and similar for v_M . Here we adopt the same notations for simplicity.

Then we should have

$$\widehat{e}_l^n = \widehat{u}_l(t_n) - \widetilde{u}_l^n, \quad \widehat{e}_l^n = \widehat{v}_l(t_n) - \widetilde{v}_l^n, \quad n = 0, \dots, \frac{T}{\Delta t}. \quad (3.4)$$

Based on (2.29b) and (2.29c), define the local truncation errors as

$$\widehat{\zeta}^n(x) := \sum_{l=-M/2}^{M/2-1} \widehat{\zeta}_l e^{i\mu_l(x-a)}, \quad \check{\zeta}^n(x) := \sum_{l=-M/2}^{M/2-1} \check{\zeta}_l e^{i\mu_l(x-a)}, \quad x \in \Omega, \quad n = 0, \dots, \frac{T}{\Delta t} - 1. \quad (3.5)$$

where

$$\widehat{\zeta}_l^n = \widehat{u}_l(t_{n+1}) - \cos(\beta_l \Delta t) \widehat{u}_l(t_n) - \frac{\sin(\beta_l \Delta t)}{\beta_l} \widehat{v}_l(t_n) + \frac{\Delta t}{2\beta_l} \sin(\beta_l \Delta t) (\widehat{f_M})_l(t_n), \quad (3.6a)$$

$$\begin{aligned} \check{\zeta}_l^n &= \widehat{v}_l(t_{n+1}) + \beta_l \sin(\beta_l \Delta t) \widehat{u}_l(t_n) - \cos(\beta_l \Delta t) \widehat{v}_l(t_n) \\ &\quad + \frac{\Delta t}{2} \left(\cos(\beta_l \Delta t) (\widehat{f_M})_l(t_n) + (\widehat{f_M})_l(t_{n+1}) \right), \end{aligned} \quad (3.6b)$$

and $\beta_l = \sqrt{\mu_l^2 + 1}$, with $(\widehat{f_M})_l(t)$ the Fourier coefficient of $f(u_M(x, t))$. Subtracting the local truncation errors (3.6) from the scheme (2.29b) and (2.29c), we are led to the error equations

$$\widehat{e}_l^{n+1} = \cos(\beta_l \Delta t) \widehat{e}_l^n + \frac{\sin(\beta_l \Delta t)}{\beta_l} \widehat{e}_l^n + \widehat{\zeta}_l^n - \widehat{\eta}_l^n, \quad (3.7a)$$

$$\check{e}_l^{n+1} = -\beta_l \sin(\beta_l \Delta t) \widehat{e}_l^n + \cos(\beta_l \Delta t) \check{e}_l^n + \check{\zeta}_l^n - \check{\eta}_l^n, \quad (3.7b)$$

where

$$\widehat{\eta}_l^n = \frac{\Delta t}{2\beta_l} \sin(\beta_l \Delta t) \left((\widehat{f_M})_l(t_n) - \widetilde{f}_l^n \right), \quad (3.8a)$$

$$\check{\eta}_l^n = \frac{\Delta t}{2} \left[\cos(\beta_l \Delta t) \left((\widehat{f_M})_l(t_n) - \widetilde{f}_l^n \right) + \left((\widehat{f_M})_l(t_{n+1}) - \widetilde{f}_l^{n+1} \right) \right], \quad (3.8b)$$

with the nonlinear error functions defined as

$$\eta^n(x) := \sum_{l=-M/2}^{M/2-1} \widehat{\eta}_l(t_n) e^{i\mu_l(x-a)}, \quad \check{\eta}^n(x) := \sum_{l=-M/2}^{M/2-1} \check{\eta}_l(t_n) e^{i\mu_l(x-a)}, \quad x \in \Omega.$$

Define the error energy as

$$\mathcal{E}(P, Q) := \|P\|_{H^1}^2 + \|Q\|_{L^2}^2, \quad (3.9)$$

for two arbitrary functions $P(x)$ and $Q(x)$.

In order to proceed with the proof of Theorem 3.1, we give the following lemmas. First, we have estimates for the local truncation errors, stated in the following lemma.

Lemma 3.1. *Based on assumption (A), we have estimates for the local truncation errors as*

$$\|\xi^n\|_{H^1} + \|\hat{\xi}^n\|_{L^2} \lesssim \Delta t^3 + \Delta t \cdot \Delta x^{m_0+1}, \quad n=0, \dots, \frac{T}{\Delta t} - 1. \tag{3.10}$$

Proof. Applying L^2 -projection on both sides of (2.1), due to the orthogonality and variation-of-constant formula, the Fourier coefficients $\hat{u}_l(t_n), \hat{v}_l(t_n)$ should satisfy

$$\hat{u}_l(t_{n+1}) = \cos(\beta_l \Delta t) \hat{u}_l(t_n) + \frac{\sin(\beta_l \Delta t)}{\beta_l} \hat{v}_l(t_n) - \int_{t_n}^{t_{n+1}} \frac{\sin(\beta_l(t_{n+1}-s))}{\beta_l} \hat{f}_l(s) ds, \tag{3.11a}$$

$$\hat{v}_l(t_{n+1}) = -\beta_l \sin(\beta_l \Delta t) \hat{u}_l(t_n) + \cos(\beta_l \Delta t) \hat{v}_l(t_n) - \int_{t_n}^{t_{n+1}} \cos(\beta_l(t_{n+1}-s)) \hat{f}_l(s) ds, \tag{3.11b}$$

for $n=0, \dots, T/\Delta t$, where $\hat{f}_l(t)$ denotes the Fourier coefficient of $f(u(x,t))$ for short, provided that no confusion occurs. Subtracting (3.11) from the local truncation errors (3.6), we get

$$\hat{\xi}_l^n = \frac{\Delta t}{2\beta_l} \sin(\beta_l \Delta t) (\widehat{f_M})_l(t_n) - \int_{t_n}^{t_{n+1}} \frac{\sin(\beta_l(t_{n+1}-s))}{\beta_l} \hat{f}_l(s) ds, \tag{3.12a}$$

$$\hat{\xi}_l^n = \frac{\Delta t}{2} \left[\cos(\beta_l \Delta t) (\widehat{f_M})_l(t_n) + (\widehat{f_M})_l(t_{n+1}) \right] - \int_{t_n}^{t_{n+1}} \cos(\beta_l(t_{n+1}-s)) \hat{f}_l(s) ds. \tag{3.12b}$$

For a general function $g(s) \in C^2$, we have the quadrature error for the standard trapezoidal rule written in the second-order Peano form [11],

$$\frac{\tau}{2} (g(0) + g(\tau)) - \int_0^\tau g(s) ds = \frac{\tau^3}{2} \int_0^1 \theta(1-\theta) g''(\theta\tau) d\theta.$$

Rewriting $\hat{\xi}_l^n$ in (3.12a) as

$$\begin{aligned} \hat{\xi}_l^n &= \frac{\Delta t}{2\beta_l} \sin(\beta_l \Delta t) \hat{f}_l(t_n) - \int_{t_n}^{t_{n+1}} \frac{\sin(\beta_l(t_{n+1}-s))}{\beta_l} \hat{f}_l(s) ds \\ &\quad + \frac{\Delta t}{2\beta_l} \sin(\beta_l \Delta t) \left[(\widehat{f_M})_l(t_n) - \hat{f}_l(t_n) \right], \end{aligned}$$

we then have

$$\begin{aligned} |\hat{\xi}_l^n| &\lesssim \Delta t^3 \int_0^1 \theta(1-\theta) \left[\beta_l \left| \hat{f}_l(t_n + \theta\Delta t) \right| + \left| \frac{d}{ds} \hat{f}_l(t_n + \theta\Delta t) \right| + \frac{1}{\beta_l} \left| \frac{d^2}{ds^2} \hat{f}_l(t_n + \theta\Delta t) \right| \right] d\theta \\ &\quad + \frac{\Delta t}{\beta_l} \left| (\widehat{f_M})_l(t_n) - \hat{f}_l(t_n) \right|. \end{aligned} \tag{3.13}$$

Taking square on both sides of (3.13) and using the Cauchy's inequality, we get

$$\begin{aligned} |\hat{\xi}_l^n|^2 &\lesssim \Delta t^6 \int_0^1 \left[\beta_l^2 \left| \hat{f}_l(t_n + \theta\Delta t) \right|^2 + \left| \frac{d}{ds} \hat{f}_l(t_n + \theta\Delta t) \right|^2 + \frac{1}{\beta_l^2} \left| \frac{d^2}{ds^2} \hat{f}_l(t_n + \theta\Delta t) \right|^2 \right] d\theta \\ &\quad + \frac{\Delta t^2}{\beta_l^2} \left| (\widehat{f_M})_l(t_n) - \hat{f}_l(t_n) \right|^2. \end{aligned} \tag{3.14}$$

Multiplying both sides of (3.14) by $\beta_l^2 = (1 + \mu_l^2)$ and then summing up for $l = -M/2, \dots, M/2 - 1$, thanks to the Parseval's identity, we get

$$\begin{aligned} \|\zeta^n\|_{H^1}^2 &\lesssim \Delta t^6 \cdot \sup_{0 < t < T} \left[\|f(u)\|_{H^2}^2 + \|\partial_t f(u)\|_{H^1}^2 + \|\partial_{tt} f(u)\|_{L^2}^2 \right] \\ &\quad + \Delta t^2 \|f(u_M(\cdot, t_n)) - f(u(\cdot, t_n))\|_{L^2}^2. \end{aligned}$$

Then based on assumption (A) and the standard projection error bounds [37], we have

$$\|\zeta^n\|_{H^1}^2 \lesssim \Delta t^6 + \Delta t^2 \cdot \Delta x^{2(m_0+1)}, \quad n = 0, \dots, \frac{T}{\Delta t} - 1. \tag{3.15}$$

Similarly, from (3.12b) we could get

$$\|\xi^n\|_{L^2}^2 \lesssim \Delta t^6 + \Delta t^2 \cdot \Delta x^{2(m_0+1)}, \quad n = 0, \dots, \frac{T}{\Delta t} - 1. \tag{3.16}$$

Combining (3.15) and (3.16) gives the assertion (3.10). □

For the nonlinear error terms, we have estimates stated as the following lemma.

Lemma 3.2. *Based on assumption (A), and assume (3.1b) holds for some $0 \leq n \leq T/\Delta t - 1$ (which will be given by induction later), then we have*

$$\|\eta^n\|_{H^1} + \|\dot{\eta}^n\|_{L^2} \lesssim \Delta t \left(\|e_M^n\|_{L^2} + \|e_M^{n+1}\|_{L^2} \right) + \Delta t \cdot \Delta x^m. \tag{3.17}$$

Proof. From (3.8), we have

$$|\widehat{\eta}_l^n| \leq \frac{\Delta t}{2\beta_l} \left| (\widehat{f_M})_l(t_n) - \widetilde{f}_l^n \right|, \quad |\dot{\widehat{\eta}}_l^n| \leq \frac{\Delta t}{2} \left[\left| (\widehat{f_M})_l(t_n) - \widetilde{f}_l^n \right| + \left| (\widehat{f_M})_l(t_{n+1}) - \widetilde{f}_l^{n+1} \right| \right].$$

For the first part, similarly as before, we can get

$$\|\eta^n\|_{H^1}^2 \leq \frac{\Delta t^2}{4} \|P_M f(u_M(\cdot, t_n)) - I_M f(u^n)\|_{L^2}^2.$$

Under assumption (A), we should have $f(u_M(x, t)) \in C([0, T]; H_p^m)$, then

$$\begin{aligned} \|\eta^n\|_{H^1}^2 &\leq \frac{\Delta t^2}{4} \|I_M f(u_M(\cdot, t_n)) - I_M f(u^n)\|_{L^2}^2 + \frac{\Delta t^2}{4} \|P_M f(u_M(\cdot, t_n)) - I_M f(u_M(\cdot, t_n))\|_{L^2}^2 \\ &\lesssim \frac{\Delta t^2}{4} \|I_M f(u_M(\cdot, t_n)) - I_M f(u^n)\|_{L^2}^2 + \frac{\Delta t^2}{4} \Delta x^{2m}. \end{aligned} \tag{3.18}$$

By Parseval's identity, together with the assumption (A) and (3.1b), we have

$$\begin{aligned} &\|I_M f(u_M(\cdot, t_n)) - I_M f(u^n)\|_{L^2} = \|f(u_M(\cdot, t_n)) - f(u^n)\|_{l^2} \\ &= \left\| \int_0^1 f'(s u_M(\cdot, t_n) + (1-s)u^n) ds \cdot (u_M(\cdot, t_n) - u^n) \right\|_{l^2} \\ &\lesssim \|u_M(\cdot, t_n) - u^n\|_{l^2} = \|u_M(\cdot, t_n) - I_M u^n\|_{L^2} = \|e_M^n\|_{L^2}, \end{aligned} \tag{3.19}$$

where $f'(\rho) = \frac{d}{d\rho}f(\rho)$ and $\|\cdot\|_{l^2}$ denotes the standard l^2 -norm. Plugging the above estimate back to (3.18), we get

$$\|\eta^n\|_{H^1} \lesssim \Delta t \|e_M^n\|_{L^2} + \Delta t \cdot \Delta x^m. \tag{3.20}$$

For the second part, similarly, we can have

$$\begin{aligned} \|\dot{\eta}^n\|_{L^2}^2 &\leq \frac{\Delta t^2}{4} \left[\|P_M f(u_M(\cdot, t_n)) - I_M f(u^n)\|_{L^2}^2 + \|P_M f(u_M(\cdot, t_{n+1})) - I_M f(u^{n+1})\|_{L^2}^2 \right] \\ &\lesssim \Delta t^2 \|e_M^n\|_{L^2}^2 + \Delta t^2 \cdot \Delta x^{2m} + \Delta t^2 \|I_M f(u_M(\cdot, t_{n+1})) - I_M f(u^{n+1})\|_{L^2}^2. \end{aligned} \tag{3.21}$$

To carry out a similar argument as (3.19), now we only need to show the maximum value of the numerical solution at t_{n+1} level, i.e. $\|u^{n+1}\|_{l^\infty}$, is bounded by some generic constant under assumption (3.1b). By the Sobolev's inequality,

$$\|u^{n+1}\|_{l^\infty} \leq \|u_I^{n+1}\|_{L^\infty} \lesssim \|u_I^{n+1}\|_{H^1} = \sqrt{(b-a) \sum_{l=-M/2}^{M/2-1} (1+\mu_l^2) |\tilde{u}_l^{n+1}|^2}.$$

From (2.29b), we get

$$|\tilde{u}_l^{n+1}| \leq |\tilde{u}_l^n| + \frac{1}{\beta_l} |\tilde{v}_l^n| + \frac{\Delta t}{2\beta_l} |\tilde{f}_l^n|.$$

Then with $\Delta t \leq 1$ similarly as before, we can get

$$\begin{aligned} \|u_I^{n+1}\|_{H^1} &\leq 2\|u_I^n\|_{H^1} + 2\|v_I^n\|_{L^2} + \|f(u^n)\|_{l^2} \leq 2\|u_I^n\|_{H^1} + 2\|v_I^n\|_{L^2} + \|f(u^n)\|_{l^\infty} \\ &\leq 2(K_1 + K_2 + 2) + \|f(\cdot)\|_{L^\infty(0, K_1+1)}. \end{aligned}$$

Thus, following the argument done as (3.19), we can get a further estimate for (3.21) as

$$\|\dot{\eta}^n\|_{L^2} \lesssim \Delta t \left(\|e_M^n\|_{L^2} + \|e_M^{n+1}\|_{L^2} \right) + \Delta t \cdot \Delta x^m. \tag{3.22}$$

Combing (3.20) and (3.22), we finish the proof. □

With the error energy functional notation (3.9), it is ready to show the following fact.

Lemma 3.3. For $n = 0, \dots, T/\Delta t - 1$, we have

$$\mathcal{E}(e_M^{n+1}, \dot{e}_M^{n+1}) - \mathcal{E}(e_M^n, \dot{e}_M^n) \leq \Delta t \mathcal{E}(e_M^n, \dot{e}_M^n) + \frac{2}{\Delta t} [\mathcal{E}(\tilde{\zeta}^n, \dot{\tilde{\zeta}}^n) + \mathcal{E}(\eta^n, \dot{\eta}^n)]. \tag{3.23}$$

Proof. Multiplying (3.7a) with its complex conjugate, and by Cauchy's inequality, we have

$$\left| \tilde{e}_l^{n+1} \right|^2 \leq (1 + \Delta t) \left| \cos(\beta_l \Delta t) \tilde{e}_l^n + \frac{\sin(\beta_l \Delta t)}{\beta_l} \tilde{e}_l^n \right|^2 + \frac{1}{\Delta t} \left| \tilde{\zeta}_l^n - \tilde{\eta}_l^n \right|^2. \tag{3.24}$$

Similarly for (3.7b), we have

$$\left| \widehat{e}_l^{n+1} \right|^2 \leq (1 + \Delta t) \left| -\beta_l \sin(\beta_l \Delta t) \widehat{e}_l^n + \cos(\beta_l \Delta t) \widehat{e}_l^n \right|^2 + \frac{1}{\Delta t} \left| \widehat{\zeta}_l^n - \widehat{\eta}_l^n \right|^2. \quad (3.25)$$

Multiplying (3.24) by $\beta_l^2 = 1 + \mu_l^2$ and then adding to (3.25), we get

$$\beta_l^2 \left| \widehat{e}_l^{n+1} \right|^2 + \left| \widehat{e}_l^{n+1} \right|^2 \leq (1 + \Delta t) \left(\beta_l^2 \left| \widehat{e}_l^n \right|^2 + \left| \widehat{e}_l^n \right|^2 \right) + \frac{1}{\Delta t} \left(\beta_l^2 \left| \widehat{\zeta}_l^n - \widehat{\eta}_l^n \right|^2 + \left| \widehat{\zeta}_l^n - \widehat{\eta}_l^n \right|^2 \right).$$

Summing the above inequalities up for $l = -M/2, \dots, M/2 - 1$, and noticing (3.9) we get

$$\mathcal{E}(e_M^{n+1}, e_M^{n+1}) \leq (1 + \Delta t) \mathcal{E}(e_M^n, e_M^n) + \frac{1}{\Delta t} \mathcal{E}(\zeta^n - \eta^n, \zeta^n - \eta^n),$$

and then by applying Cauchy's inequality again, we get assertion (3.23). □

Now, combining the Lemmas 3.1-3.3, we give the proof of Theorem 3.1 with the help of mathematical induction argument [6], or the so-called cut-off technique [3] for the boundedness of numerical solutions.

Proof of Theorem 3.1. For $n = 0$, from the scheme and assumption (A), we have

$$\|e^0\|_{H^1} + \|e^0\|_{L^2} = \|\phi_1 - I_M \phi_1\|_{H^1} + \|\phi_2 - I_M \phi_2\|_{L^2} \lesssim \Delta x^{m_0}.$$

Then by triangle inequality,

$$\|u_l^0\|_{H^1} \leq \|\phi_1\|_{H^1} + \|e^0\|_{H^1} \leq K_1 + 1, \quad \|v_l^0\|_{L^2} \leq \|\phi_2\|_{L^2} + \|e^0\|_{L^2} \leq K_2 + 1, \quad 0 < \Delta x \leq h_1,$$

for some $h_1 > 0$ independent of Δt and Δx , and obviously $\|u^0\|_{l^\infty} \leq K_1 + 1$. Thus, (3.1) is true for $n = 0$.

Assume (3.1) holds for $n \leq N \leq T/\Delta t - 1$. Now we need to show the results are still true for $n = N + 1$. First, by triangle inequality,

$$\begin{aligned} \|e^n\|_{H^1} + \|e^n\|_{L^2} &\leq \|e_M^n\|_{H^1} + \|e_M^n\|_{L^2} + \|u(\cdot, t_n) - u_M(\cdot, t_n)\|_{H^1} + \|v(\cdot, t_n) - v_M(\cdot, t_n)\|_{L^2} \\ &\lesssim \|e_M^n\|_{H^1} + \|e_M^n\|_{L^2} + \Delta x^{m_0}. \end{aligned} \quad (3.26)$$

Then from Lemma 3.3, we have for $n = 0, \dots, N$,

$$\mathcal{E}(e_M^{n+1}, e_M^{n+1}) - \mathcal{E}(e_M^n, e_M^n) \lesssim \Delta t \mathcal{E}(e_M^n, e_M^n) + \frac{1}{\Delta t} [\mathcal{E}(\zeta^n, \zeta^n) + \mathcal{E}(\eta^n, \eta^n)].$$

Since (3.1b) is assumed to be true for all $n \leq N$, we can plug the estimates in Lemma 3.1 and Lemma 3.2 into the above estimate and get

$$\begin{aligned} \mathcal{E}(e_M^{n+1}, e_M^{n+1}) - \mathcal{E}(e_M^n, e_M^n) &\lesssim \Delta t \left[\mathcal{E}(e_M^{n+1}, e_M^{n+1}) + \mathcal{E}(e_M^n, e_M^n) \right] + \Delta t^5 + \Delta t (\Delta x^{2m} + \Delta x^{2(m_0+1)}) \\ &\lesssim \Delta t \left[\mathcal{E}(e_M^{n+1}, e_M^{n+1}) + \mathcal{E}(e_M^n, e_M^n) \right] + \Delta t^5 + \Delta t \cdot \Delta x^{2m}. \end{aligned} \quad (3.27)$$

Summing (3.27) up for $n=0, \dots, N$, and then by the discrete Gronwall's inequality, we get

$$\mathcal{E}(e_M^{N+1}, \dot{e}_M^{N+1}) \lesssim \Delta t^4 + \Delta x^{2m}.$$

Thus, we have $\|e_M^{N+1}\|_{H^1} + \|\dot{e}_M^{N+1}\|_{L^2} \leq \Delta t^2 + \Delta x^m$, which together with (3.26) show (3.1b) is valid for $n=N+1$. Then by triangle inequality,

$$\begin{aligned} \|u_I^{N+1}\|_{H^1} &\leq \|u(\cdot, t_{N+1})\|_{H^1} + \|e^{N+1}\|_{H^1} \leq K_1 + 1, & 0 < \Delta t \leq \tau_1, \quad 0 < \Delta x \leq h_2, \\ \|v_I^{N+1}\|_{L^2} &\leq \|v(\cdot, t_{N+1})\|_{L^2} + \|\dot{e}^{N+1}\|_{L^2} \leq K_2 + 1, \end{aligned}$$

for some $\tau_1, h_2 > 0$ independent of Δt and Δx . Noting the Sobolev's inequality $\|e^N\|_{L^\infty} \lesssim \|e^N\|_{H^1}$, we also have

$$\|u^{N+1}\|_{L^\infty} \leq \|u_I^{N+1}\|_{L^\infty} \leq \|u(\cdot, t_{N+1})\|_{L^\infty} + \|e^{N+1}\|_{L^\infty} \leq K_1 + 1, \quad 0 < \Delta t \leq \tau_2, \quad 0 < \Delta x \leq h_3,$$

for some $\tau_2, h_3 > 0$ independent of Δt and Δx . Therefore, the proof is completed by choosing $\tau_0 = \min\{\tau_1, \tau_2\}$ and $h_0 = \min\{h_1, h_2, h_3\}$. □

Remark 3.1. We would like to remark that although the error estimate arguments are given for 1D, the results and proof for higher dimensions can be achieved in the same spirit. In higher dimensional space, the Sobolev's inequality reads $\|\rho\|_{L^\infty} \lesssim \|\rho\|_{H^2}$, then one only need to rise the energy space for error functions to $H^2 \times H^1$, under a stronger regularity assumption than (A).

Remark 3.2. In the nonrelativistic limit regime, i.e. (2.1) with $0 < \varepsilon \ll 1$, following the analogous procedure made in this section, one can establish an error bound of the TSFP method (2.16) as

$$\|e^n\|_{H^1} + \varepsilon \|\nabla e^n\|_{H^1} + \varepsilon^2 \|\dot{e}^n\|_{H^1} \lesssim \frac{\Delta t^2}{\varepsilon^4} + \Delta x^m, \tag{3.28}$$

under a stronger regularity assumption than (A) and an oscillation assumption

$$\|u\| + \varepsilon^2 \|\partial_t u\| + \varepsilon^4 \|\partial_{tt} u\| \lesssim 1,$$

for certain norms. We omit the detailed arguments here for brevity.

The error bound (3.28) is quite similar to the one obtained in [6] for GIFFP, of which the ε -dependence $\Delta t \lesssim \varepsilon^2$ has been numerically shown to be optimal for GIFFP. On the other hand, the error bound (3.28) also agrees with the expectation since the local truncation errors mainly come from trapezoidal quadrature, which is second-order accurate with a factor before Δt^2 of the same order as $\partial_{tt} u$. However, our extensive numerical results, presented in the forthcoming section, will show that the ε -dependence in the estimate (3.28) is indeed not optimal for TSFP when $0 < \varepsilon \ll 1$. In fact, it suggests that the error of TSFP would asymptotically behave like

$$\|e^n\|_{H^1} \lesssim \frac{\Delta t^2}{\varepsilon^2} + \Delta x^m. \tag{3.29}$$

Thus, rigorous arguments towards an optimal error estimate of TSFP for $0 < \varepsilon \ll 1$ are still absent, for which our work is still on-going.

4 Numerical studies

We now report numerical results by using the TSFP method introduced in Section 2. We will first test the TSFP method in the $\mathcal{O}(1)$ -speed of light regime to justify our theoretical error estimate results Theorem 3.1. Then we will apply the TSFP method to solve the KG equation (2.1) for $\varepsilon \in (0,1)$, with special attention paid to its numerical resolution capacity for the temporally $\mathcal{O}(\varepsilon^2)$ -oscillatory solutions when $0 < \varepsilon \ll 1$, i.e., what meshing strategy requirement should be satisfied in order to obtain ‘correct’ approximations or the desired accurate digits in the nonrelativistic limit regime (ε -scalability). Since we are mainly interested in its temporal resolution in this regime, 1D problem is used as test examples.

We consider the nonlinearity as $f(u) = \lambda u^{p+1}$ with $p \geq 0$ and $\lambda \in \mathbb{R}$ in (2.1), i.e., the pure power case which is the typical nonlinearity investigated in literatures [6, 10, 26–29, 33, 35, 38, 40]. We set $p=2$ and choose the initial conditions (2.1c) as

$$u(x,0) = \frac{3\sin(x)}{e^{0.5x^2} + e^{-0.5x^2}}, \quad v(x,0) = \frac{2e^{-x^2}}{\sqrt{\pi}\varepsilon^2}, \quad x \in \mathbb{R}.$$

We truncate the problem onto a finite domain $\Omega = [-16, 16]$, i.e. $b = -a = 16$, which is large enough such that the periodic boundary conditions (2.1b) do not introduce significant aliasing errors relative to the whole space problem.

4.1 Accuracy tests for $\varepsilon = \mathcal{O}(1)$

We take fixed $\varepsilon=1$ (i.e., the $\mathcal{O}(1)$ -speed of light regime). In this case, there is no analytical solution and we let $u(x,t)$ be the ‘exact’ solution which is obtained numerically by using TSFP method with very fine mesh size and small time step, e.g., $\Delta x = 1/32$ and $\Delta t = 10^{-5}$. We solve the problem on the interval $\Omega = [-16, 16]$ till time $t = 2$ in two cases: $\lambda = 1$ (defocusing case) and $\lambda = -1$ (focusing case). Here, we test the spatial and temporal discretization errors separately, and then test the conservation of the conserved energy of TSFP. To quantify the numerical results, we present the error:

$$e^{\Delta t, \Delta x}(t = t_n) = \|e^n\|_{H^1} = \|u(\cdot, t_n) - u_1^n\|_{H^1}. \quad (4.1)$$

First, we test the discretization error in space, and in order to do this we take a very fine time step $\Delta t = 10^{-5}$ such that the error from time discretization is negligible compared to the spatial discretization error. Table 1 lists the errors (4.1) at time $t = 1$ with different mesh sizes Δx and parameter λ . Second, we test the discretization error in time, and mesh size is chosen as $\Delta x = 1/16$ such that the error from space discretization is negligible. Table 2 shows the errors (4.1) at time $t = 1$ with different time steps Δt and parameter λ . Third, we test the conservation of the conserved energy $E(t)$ (1.2). Here we choose a small mesh size in space such that the energy $E(t=0)$ which is approximated spectrally from the initial data is very close to the exact conserved energy. Table 3 lists the discrete

Table 1: Spatial discretization errors of TSFP at time $t=1$ for different mesh sizes Δx under $\Delta t=10^{-5}$.

$e^{\Delta t, \Delta x}$	$\Delta x=1$	$\Delta x=1/2$	$\Delta x=1/4$	$\Delta x=1/8$
$\lambda=1$	$3.71E-2$	$1.70E-3$	$1.34E-6$	$2.22E-12$
$\lambda=-1$	$3.95E-2$	$1.70E-3$	$1.58E-6$	$2.53E-12$

Table 2: Temporal discretization errors of TSFP at time $t=1$ for different time steps Δt under $\Delta x=1/16$ with convergence rate.

$e^{\Delta t, \Delta x}$	$\Delta t=1/5$	$\Delta t=1/10$	$\Delta t=1/20$	$\Delta t=1/40$	$\Delta t=1/80$
$\lambda=1$	$1.50E-3$	$3.65E-4$	$9.06E-5$	$2.26E-5$	$5.64E-6$
rate	-	2.03	2.01	2.00	2.00
$\lambda=-1$	$2.40E-3$	$6.14E-4$	$1.54E-4$	$3.84E-5$	$9.61E-6$
rate	-	1.97	2.00	2.00	2.00

Table 3: Conserved energy analysis of TSFP: $\Delta t=10^{-3}$ and $\Delta x=1/8$.

$E(t)$	$t=0$	$t=0.5$	$t=1.0$	$t=1.5$	$t=2.0$
$\lambda=1$	10.0957456	10.0957438	10.0957437	10.0957450	10.0957441
$\lambda=-1$	7.6534166	7.6534174	7.6534178	7.6534176	7.6534175

energy at different time points with $\Delta t=10^{-3}$ and $\Delta x=1/8$. Fig. 1 shows the convergence of the energy error as Δt decreases. Here for the focusing case, i.e. $\lambda=-1$, the results are only shown till $T=2$ because of the finite time blow up of the solution.

From Tables 1-3, Fig. 1 and additional results not shown here brevity, we can draw the following observations:

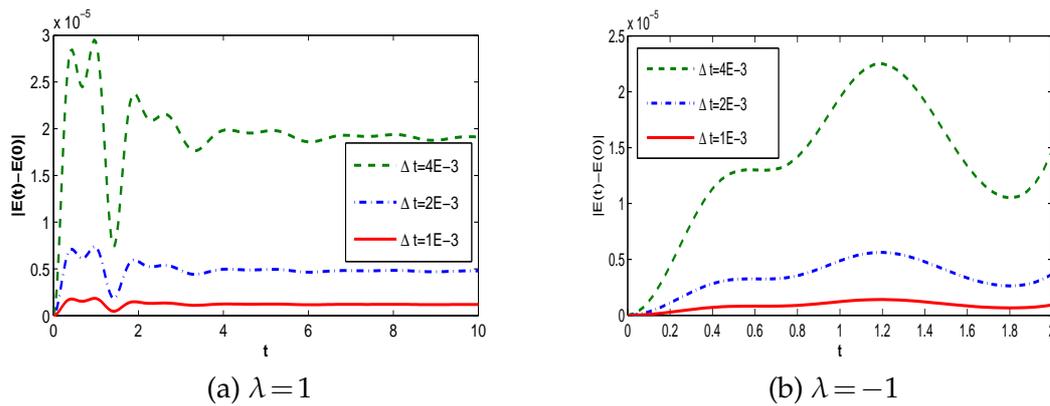


Figure 1: Energy error of TSFP in defocusing case ($\lambda=1$) and focusing case ($\lambda=-1$): $|E(t) - E(0)|$ for different Δt during the computing under $\Delta x=1/8$ and $\varepsilon=1$.

1. In the $\mathcal{O}(1)$ -speed of light regime, the TSFP (2.16) is of spectral-order accuracy in space, and is of second-order accuracy in time (cf. Tables 1 and 2), which verifies our error estimate (3.1a) and indicates the result is optimal.
2. TSFP conserves the energy very well. The energy obtained from the numerical solution is just a small fluctuation from the exact energy during the computation (cf. Table 3). As time step Δt decreases to zero, the energy error during the computing converges to zero (cf. Fig. 1).
3. Furthermore, the method is efficient, easy to implement, less memory requirement, and easy to extended to 2D and 3D problems.

4.2 Convergence and resolution studies for $0 < \varepsilon \ll 1$

We now consider $\varepsilon \in (0, 1)$ in (2.1), with special efforts made to the regime $0 < \varepsilon \ll 1$, i.e., the nonrelativistic limit regime. Here, we investigate the temporal and spatial errors of TSFP under different mesh sizes and time steps as $\varepsilon \rightarrow 0$. By doing so, we mainly want to study two questions. The first one is how do the convergence and accuracy of the numerical method be affected as ε decays. Then within the convergence regime, how do the error bounds depend on ε . Again, the ‘exact’ solution $u(x, t)$ is obtained by a similar way as before. Since the numerical results of TSFP are similar in defocusing and focusing cases, so we here only consider a defocusing case with $\lambda = 1$ as a numerical example.

The spatial error and temporal error here are computed in a similar way as before. For error analysis in space, either from our numerical experience or from the theoretical result (3.28) and estimates in [6], the spatial errors of TSFP and GIFP are almost the same due to the same spectral discretization used in space. Thus here we omit the results of GIFP for brevity and tabulate the spatial error of TSFP under different ε and mesh sizes Δx in Table 4. Table 5 shows the temporal error of TSFP, under different ε and time steps Δt , together with the results of GIFP for comparisons. To study the error bounds of the numerical methods inside the convergence regime, we plot the temporal discretization errors as a function of ε for some fixed Δt in log-scale. The results are shown in Fig. 2. Moreover, we

Table 4: Spatial error analysis of TSFP for different ε and Δx at time $t=1$ under $\Delta t=10^{-5}$.

TSFP	$\Delta x_0=1$	$\Delta x_0/2$	$\Delta x_0/4$	$\Delta x_0/8$
$\varepsilon_0=0.5$	7.99E-02	4.20E-03	3.01E-06	2.78E-12
$\varepsilon_0/2$	8.13E-02	5.40E-03	2.28E-06	2.84E-12
$\varepsilon_0/2^2$	2.77E-02	1.30E-03	1.07E-06	1.75E-12
$\varepsilon_0/2^3$	4.88E-02	4.60E-03	3.05E-06	1.67E-12
$\varepsilon_0/2^4$	8.24E-02	4.30E-03	2.74E-06	1.72E-12
$\varepsilon_0/2^6$	4.57E-02	5.00E-03	3.02E-06	1.89E-12

Table 5: Temporal error analysis of TSFP for different ε and Δt , and comparisons with GIFP at time $t=1$ under $\Delta x=1/16$ with convergence rate.

TSFP	$\Delta t_0=0.2$	$\Delta t_0/2^2$	$\Delta t_0/2^4$	$\Delta t_0/2^6$	$\Delta t_0/2^8$	$\Delta t_0/2^{10}$
$\varepsilon_0=0.5$	1.52E-02	5.66E-04	3.49E-05	2.18E-06	1.36E-07	8.37E-09
rate	–	2.37	2.01	2.00	2.00	2.01
$\varepsilon_0/2^1$	1.27E-01	3.80E-03	1.97E-04	1.22E-05	7.62E-07	4.69E-08
rate	–	2.53	2.13	2.00	2.00	2.01
$\varepsilon_0/2^2$	1.13E-01	1.13E-01	8.93E-04	4.85E-05	3.01E-06	1.85E-07
rate	–	0.00	3.49	2.10	2.01	2.01
$\varepsilon_0/2^3$	1.21E-01	4.57E-02	4.52E-02	2.33E-04	1.27E-05	7.76E-07
rate	–	0.70	0.01	3.80	2.10	2.02
$\varepsilon_0/2^4$	1.35E-01	8.90E-03	1.02E-02	1.04E-02	6.75E-05	3.69E-06
rate	–	1.97	-0.98	-0.01	3.62	2.10
$\varepsilon_0/2^6$	1.38E-01	1.42E-02	3.70E-03	4.90E-04	6.70E-04	6.83E-04
rate	–	1.64	0.97	1.46	-0.22	-0.01
$\varepsilon_0/2^8$	8.80E-02	1.99E-02	2.10E-02	7.54E-04	2.32E-04	1.62E-05
rate	–	1.07	-0.04	2.40	0.85	1.92
$\varepsilon_0/2^{10}$	1.14E-01	5.43E-02	1.50E-03	1.00E-03	1.10E-03	3.78E-05
rate	–	0.53	2.58	0.26	-0.09	2.43
GIFP	$\Delta t_0=0.2$	$\Delta t_0/2^2$	$\Delta t_0/2^4$	$\Delta t_0/2^6$	$\Delta t_0/2^8$	$\Delta t_0/2^{10}$
$\varepsilon_0=0.5$	1.96E-02	7.82E-04	4.87E-05	3.04E-06	1.90E-07	1.20E-08
rate	–	2.32	2.00	2.00	2.00	1.99
$\varepsilon_0/2$	4.37E-01	1.16E-02	6.86E-04	4.29E-05	2.68E-06	1.68E-07
rate	–	2.61	2.04	2.00	2.00	2.00
$\varepsilon_0/2^2$	1.19E-01	4.98E-01	1.17E-02	7.09E-04	4.43E-05	2.77E-06
rate	–	-1.03	2.71	2.02	2.00	2.00
$\varepsilon_0/2^3$	1.68E-01	1.64E-01	4.73E-01	1.19E-02	7.06E-04	4.41E-05
rate	–	0.02	-0.46	2.66	2.04	2.00
$\varepsilon_0/2^4$	1.76E-01	1.85E-01	1.96E-01	6.85E-01	1.12E-02	6.63E-04
rate	–	-0.04	-0.04	-0.90	2.96	2.04
$\varepsilon_0/2^6$	1.13E-01	2.04E-01	2.22E-01	2.17E-01	2.20E-01	8.77E-01
rate	–	-0.42	-0.06	0.02	-0.01	-0.99
$\varepsilon_0/2^8$	1.53E-01	1.94E-01	2.01E-01	4.00E-01	2.13E-01	5.72E-01
rate	–	-0.17	-0.03	-0.04	0.45	-0.71
$\varepsilon_0/2^{10}$	1.76E-01	1.99E-01	2.09E-01	2.12E-01	2.16E-01	2.14E-01
rate	–	-0.08	-0.04	-0.01	-0.01	0.01

study the performance of TSFP in temporal approximations in Table 6 under the meshing strategy $\Delta = \mathcal{O}(\varepsilon^2)$, which is the exact ε -scalability of GIFP [6].

From Tables 5-6, Fig. 2 and additional results not shown here brevity, we can draw

Table 6: ε -scalability analysis: temporal error at time $t=1$ with $\Delta x=1/16$ for different Δt and ε under meshing requirement $\Delta t=c \cdot \varepsilon^2$.

TSFP	$\varepsilon_0=0.5$	$\varepsilon_0/2$	$\varepsilon_0/4$	$\varepsilon_0/8$	$\varepsilon_0/16$
$c=0.8$	1.52E-02	3.80E-03	8.93E-04	2.33E-04	6.75E-05
$c=0.4$	2.40E-03	8.10E-04	1.99E-04	5.20E-05	1.52E-05
$c=0.2$	5.66E-04	1.97E-04	4.85E-05	1.27E-05	3.69E-06
GIFP	$\varepsilon_0=0.5$	$\varepsilon_0/2$	$\varepsilon_0/4$	$\varepsilon_0/8$	$\varepsilon_0/16$
$c=0.8$	1.96E-02	1.16E-02	1.17E-02	1.19E-02	1.12E-02
$c=0.4$	3.20E-03	2.70E-03	2.90E-03	2.90E-03	2.70E-03
$c=0.2$	7.82E-04	6.86E-04	7.09E-04	7.06E-04	6.63E-04

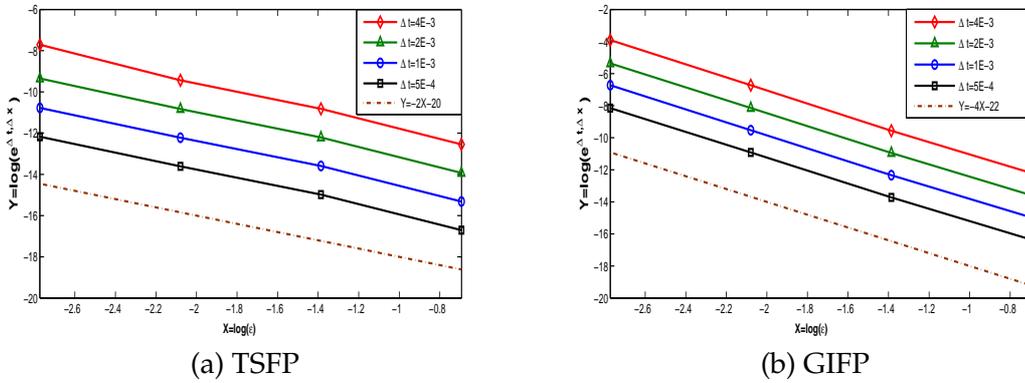


Figure 2: Dependence of the temporal discretization error on ε (in log-scale) for different Δt at $t=1$ under $\Delta x=1/8$: (a) for TSFP and (b) for GIFP.

the following observations:

1. TSFP has uniform spectral accuracy in space for all $\varepsilon \in (0,1]$ (cf. each column in Table 4). The spatial discretization error is totally independent of ε . Thus the spatial resolution of TSFP is

$$\Delta x = \mathcal{O}(1), \quad 0 < \varepsilon \ll 1,$$

i.e. the mesh size can be chosen independent of ε , which is the same as GIFP [6].

2. As ε vanishes, both TSFP and GIFP are second-order accurate in time when Δt is sufficiently small, i.e. within the convergence regime, e.g. $\Delta t \lesssim \varepsilon^2$, (cf. the upper diagonal part of Table 5), and both methods either have some convergence order reductions or lose the convergence outside the convergence regime (cf. the lower diagonal part of Table 5). Between the two numerical methods, TSFP always offers better temporal approximations than GIFP under the same time step, especially when ε becomes small (cf. Table 5).

3. The temporal discretization error bound of GIFP within the convergence regime behaves like $\mathcal{O}(\varepsilon^{-4}\Delta t^2)$ (cf. Fig. 2(b)) and the ε -scalability is $\Delta t = \mathcal{O}(\varepsilon^2)$ which are consistent with the results in [6]. Fig. 2(a) indicates that the temporal error bound of TSFP would asymptotically behave like $\mathcal{O}(\varepsilon^{-2}\Delta t^2)$ within the convergence regime, which on the other hand indicates that the estimate provided in (3.28) is not optimal in time. Table 6 illustrates a clearly second convergence in terms of ε for the temporal error of TSFP as $\varepsilon \rightarrow 0$ under the mesh strategy $\Delta t = \mathcal{O}(\varepsilon^2)$, while GIFP shows no convergence, which again indicate the temporal error bounds for the two methods and shows that TSFP will dominant in the highly oscillatory regime.

5 Conclusions

A time-splitting Fourier pseudospectral (TSFP) discretization was proposed and analyzed for solving the Klein-Gordon (KG) equation in a scaling involving a dimensionless parameter $\varepsilon \in (0,1]$, which is inversely proportional to the speed of light. Special attention was paid to the highly oscillatory regime, i.e., the nonrelativistic limit regime $0 < \varepsilon \ll 1$ or the speed of light goes to infinity. The TSFP under study was derived for a simple equivalent first-order-in-time form of the KG equation. It was shown that this method coincides with a splitting method applied to the classical relativistic NLS reformulation. We also showed that the TSFP is essentially equivalent to a trigonometric integrator pseudospectral method. Rigorous error estimates of the TSFP method were achieved for the regime $\varepsilon = \mathcal{O}(1)$. Extensive numerical studies were carried out, which demonstrated that the TSFP is quite effective in capturing the temporally $\mathcal{O}(\varepsilon^2)$ -oscillatory solutions in the nonrelativistic limit regime. When $0 < \varepsilon \ll 1$, the TSFP has uniform spectral accuracy in spatial discretization and offers compelling better temporal approximations over the well-established Gautschi-type pseudospectral method rigorously studied in [6]. In fact, our numerical results suggest that the temporal discretization error of TSFP within the convergence regime is $\mathcal{O}(\varepsilon^{-2}\Delta t^2)$, whereas that of Gautschi-type method is $\mathcal{O}(\varepsilon^{-4}\Delta t^2)$, where Δt denotes the time step. Rigorous arguments for optimal error bounds of the TSFP when $0 < \varepsilon \ll 1$ are of great interests and we propose to do it in a future work.

Acknowledgments

This work was supported by the Singapore A*STAR SERC PSF-Grant 1321202067. The authors would like to denote their thanks to Prof. Weizhu Bao for the stimulating discussions and helpful advice. Part of this work was done when the authors were visiting the Institute for Mathematical Science, National University of Singapore in 2011/12.

References

- [1] M.J. Ablowitz, M.D. Kruskal, and J.F. Ladik, Solitary wave collisions, *SIAM J. Appl. Math.*, 36 (1979), pp. 428–437.

- [2] D.D. Bařnov and E. Minchev, Nonexistence of global solutions of the initial-boundary value problem for the nonlinear Klein-Gordon equation, *J. Math. Phys.* 36 (1995), pp. 756–762.
- [3] W. Bao and Y. Cai, Optimal error estimates of finite difference methods for the Gross-Pitaevskii equation with angular momentum rotation, *Math. Comp.*, 82 (2013), pp. 99–128.
- [4] W. Bao and Y. Cai, Mathematical theory and numerical methods for Bose-Einstein condensation, *Kinet. Relat. Mod.*, 6 (2013), pp. 1–135.
- [5] W. Bao and X. Dong, Numerical methods for computing ground states and dynamics of nonlinear relativistic Hartree equation for boson stars, *J. Compt. Phys.*, 230 (2011), pp. 5449–5469.
- [6] W. Bao and X. Dong, Analysis and comparison of numerical methods for Klein-Gordon equation in nonrelativistic limit regime, *Numer. Math.*, 120 (2012), pp. 189–229.
- [7] W. Bao, S. Jin, and P.A. Markowich, Time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime, *J. Compt. Phys.*, 175 (2002), pp. 487–524.
- [8] W. Bao, S. Jin, and P.A. Markowich, Numerical studies of time-splitting spectral discretizations of nonlinear Schrödinger equations in the semiclassical regime, *SIAM J. Sci. Compt.*, 25 (2003), pp. 27–64.
- [9] W. Bao and J. Shen, A fourth-order time-splitting Laguerre-Hermite pseudo-spectral method for Bose-Einstein condensates, *SIAM J. Sci. Comput.*, 26 (2005), pp. 2010–2028.
- [10] P. Brenner and W. Von Wahl, Global classical solutions of non-Linear wave equations, *Math. Z.*, 176 (1981), pp. 87–121.
- [11] D. Germund and B. Åke, *Numerical Methods in Scientific Computing: Vol. 1*, SIAM, 2008.
- [12] W. Cao and B. Guo, Fourier collocation method for solving nonlinear Klein-Gordon equation, *J. Comput. Phys.* 108 (1993), pp. 296–305.
- [13] A.S. Davydov, *Quantum Mechanics*, 2nd Edition, Pergamon, Oxford, 1976.
- [14] E.Y. Deeba and S.A. Khuri, A decomposition method for solving the nonlinear Klein-Gordon equation, *J. Comput. Phys.*, 124 (1996), pp. 442–448.
- [15] D.B. Duncan, Symplectic finite difference approximations of the nonlinear Klein-Gordon equation, *SIAM J. Numer. Anal.* 34 (1997), pp. 1742–1760.
- [16] P. Deuffhard, A study of extrapolation methods based on multistep schemes without parasitic solutions, *ZAMP.*, 30 (1979), pp. 177–189.
- [17] V. Grimm, A note on the Gautschi-type method for oscillatory second-order differential equations, *Numer. Math.*, 102 (2005), pp. 61–66.
- [18] V. Grimm, On error bounds for the Gautschi-type exponential integrator applied to oscillatory second-order differential equations, *Numer. Math.*, 100 (2005), pp. 71–89.
- [19] R.H. Hardin and F.D. Tappert, Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations, *SIAM Rev.* 15 (1973), pp. 423.
- [20] M. Hochbruck and Ch. Lubich, A Gautschi-type method for oscillatory second-order differential equations, *Numer. Math.*, 83 (1999), pp. 402–426.
- [21] E. Hairer, Ch. Lubich and G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, 2002.
- [22] S. Jiménez and L. Vázquez, Analysis of four numerical schemes for a nonlinear Klein-Gordon equation, *Appl. Math. Comput.* 35 (1990), pp. 61–94.
- [23] S. Ibrahim, M. Majdoub and N. Masmoudi, Global solutions for a semilinear, two-dimensional Klein-Gordon equation with exponential-type nonlinearity, *Comm. Pure Appl. Math.* 59, (2006), pp. 1639–1658.
- [24] R. Kosecki, The unit condition and global existence for a class of nonlinear Klein-Gordon equations, *J. Diff. Eqn.* 100 (1992), pp. 257–268.

- [25] S. Li and L. Vu-Quoc, Finite difference calculus invariant structure of a class of algorithms for the nonlinear Klein-Gordon equation, *SIAM J. Numer. Anal.*, 32 (1995), pp. 1839–1875.
- [26] S. Machihara, The nonrelativistic limit of the nonlinear Klein-Gordon equation, *Funkcial. Ekvac.*, 44 (2), pp. 243–252.
- [27] S. Machihara, K. Nakanishi and T. Ozawa, Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations, *Math. Ann.*, 322 (2002), pp. 603–621.
- [28] N. Masmoudi and K. Nakanishi, From nonlinear Klein-Gordon equation to a system of coupled nonlinear Schrödinger equations, *Math. Ann.*, 324 (2002), pp. 359–389.
- [29] C. Morawetz and W. Strauss, Decay and scattering of solutions of a nonlinear relativistic wave equation, *Comm. Pure Appl. Math.*, 25 (1972), pp. 1–31.
- [30] B. Najman, The nonrelativistic limit of the nonlinear Klein-Gordon equation, *Nonlinear Anal.* 15 (1990), pp. 217–228.
- [31] B. Najman, The nonrelativistic limit of Klein-Gordon and Dirac equations, *Differential equations with applications in biology, physics, and engineering (Leibnitz, 1989)*, pp. 291–299, *Lecture Notes in Pure and Appl. Math.*, 133, Dekker, New York, 1991.
- [32] P.J. Pascual, S. Jiménez and L. Vázquez, Numerical simulations of a nonlinear Klein-Gordon model, *Applications. Computational physics (Granada, 1994)*, pp. 211–270, *Lecture Notes in Phys.*, 448, Springer, Berlin, 1995.
- [33] H. Pecher, Nonlinear Small Data Scattering for the Wave and Klein-Gordon Equation, *Math. Z.*, 185 (1984), pp. 261–270.
- [34] J.J. Sakurai, *Advanced Quantum Mechanics*, Addison Wesley, New York, 1967.
- [35] I.E. Segal, The Global Cauchy Problem for a Relativistic Scalar Field with Power Interaction, *Bull. Soc. Math. Fr.*, 91 (1963), pp. 129–135.
- [36] J. Shatah, Normal forms and quadratic nonlinear Klein-Gordon equations, *Commun. Pure Appl. Math.* 38 (1985), pp. 685–696.
- [37] J. Shen and T. Tang, *Spectral and High-Order Methods with Applications*, Science Press, Beijing, 2006.
- [38] J.C.H. Simon and E. Taflin, The cauchy problem for non-linear Klein-Gordon equations, *Commun. Math. Phys.*, 152 (1993), pp. 433–478.
- [39] G. Strang, On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.*, 5 (1968), pp. 506–517.
- [40] W. Strauss, Decay and Asymptotics for $\square u = f(u)$, *J. Funct. Anal.*, 2 (1968), pp. 409–457.
- [41] W. Strauss and L. Vázquez, Numerical solution of a nonlinear Klein-Gordon equation, *J. Comput. Phys.*, 28 (1978), pp. 271–278.
- [42] Y. Tourigny, Product approximation for nonlinear Klein-Gordon equations, *IMA J. Numer. Anal.*, 9 (1990), pp. 449–462.
- [43] M. Tsutsumi, Nonrelativistic approximation of nonlinear Klein-Gordon equations in two space dimensions, *Nonlinear Anal.* 8 (1984), pp. 637–643.
- [44] J.A.C. Weideman and B.M. Herbst, Split-step methods for the solution of the nonlinear Schrödinger equation, *SIAM J. Numer. Anal.*, 23 (1986), pp. 485–507.
- [45] H. Yoshida, Construction of higher order symplectic integrators, *Phys. Lett. A*, 150 (1990), pp. 262–268.