DOI 10.1007/s10958-019-04301-1
Journal of Mathematical Sciences, Vol. 239, No. 3, June, 2019

APPROXIMATING THE NONLINEAR SCHRÖDINGER EQUATION BY A TWO LEVEL LINEARLY IMPLICIT FINITE ELEMENT METHOD

M. Asadzadeh*

Chalmers University of Technology and Göteborg University
SE–412 96, Göteborg, Sweden
mohammad@chalmers.se

C. Standar

Chalmers University of Technology and Göteborg University
SE–412 96, Göteborg, Sweden
standarc@chalmers.se

UDC 517.9

We study a numerical scheme for an initial- and Dirichlet boundary-value problem for a nonlinear Schrödinger equation. For the proposed fully discrete scheme we show convergence both in the L_2– and H^1–norms. Bibliography: 16 *titles.*

1 Introduction

We consider an initial- and Dirichlet boundary–value problem for a nonlinear Schrödinger equation and approximate the solution by using a local (nonuniform) two level scheme in time [1, 2] combined with an optimal finite element strategy for the discretization in the spatial variable based on studies outlined, for example, in [3, 4].

Let $T > 0$ be a final time, and let $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, be an arbitrary convex and simply connected spatial domain. We consider the following initial- and Dirichlet boundary-value problem for a nonlinear Schrödinger equation: Find a function $u : [0, T] \times \overline{D} \to \mathbb{C}$ such that

$$u_t = i\Delta u + if(|u|^2)u + g(t, x) \quad \forall (t, x) \in (0, T] \times D, \tag{1.1}$$

$$u(t, \cdot)|_{\partial D} = 0 \quad \forall t \in (0, T], \tag{1.2}$$

$$u(0, x) = u_0 \quad \forall x \in D, \tag{1.3}$$

where $u_0 : \overline{D} \to \mathbb{C}$, $f \in C^3([0, \infty); \mathbb{R})$, and $g \in C^3([0, T) \times \mathbb{R}^d; \mathscr{C}_h)$. For this problem we consider a fully discrete optimal space-time numerical scheme based on a spatial discretization strategy as in [3] combined with two-level (half-step) Crank–Nicolson and Backward Euler temporal

* To whom the correspondence should be addressed.

Translated from *Problemy Matematicheskogo Analiza* **97**, 2019, pp. 3-14.

1072-3374/19/2393-0233 © 2019 Springer Science+Business Media, LLC

discretizations. (For locally Lipschitz f, $g : \mathbb{C} \to \mathbb{C}$ the well-posedness of the problem with a sufficiently smooth solution requires further smoothness and compatibility assumptions, which will be considered in the semi-discrete approximation below.)

The nonlinear Schrödinger equation is modeling several physical phenomena describing, for example, quantum effects, with a solution that describes molecular, atomic, subatomic as well as macroscopic systems. In particular, the cubic nonlinear Schrödinger equation (when $f(x) = \lambda x$ for a real number λ) is of vital interest in applications such as nonlinear optics, oceanography, and plasma physics. For a survey of significant mathematical results on the Schrödinger equation we refer to an early work [5]. Previous studies related to this work can be found, for example, in [6, 7]. The study in [6] concerns an initial value problem for a radially symmetric nonlinear Schrödinger equation in 2 and 3 spatial dimensions discretized by a standard Galerkin method combined with a Crank-Nicolson type time-stepping. While in [7], the authors consider an implicit Runge-Kutta temporal approximation combined with the Galerkin method for the spatial domain. In both studies, the spatial scheme is on the background and the focus of analysis is on the time discretization. We give a brief approach to a more standard spatial discretization. In this part, we relay on the investigations in [3], where also a second order accurate temporal discretization based on a Crank-Nicolson scheme is studied. There is a more abstract approach [8] that relies on the nonlinear stability theory developed in [4]. In [8], a pointwise error bound is established and H^1 optimal estimates are derived for both backward-Euler and Crank-Nicolson, temporal, schemes. Somewhat more elaborate studies employing the discontinuous Galerkin strategy for spatial discretization are given in [9] and [10], where the discontinuous Galerkin method for the coupled nonlinear Schrödinger equations is considered. More specifically, while [9] concerns the L_2-stability and implementations, the work [10] is devoted to multiscale variational approach for the space-time discretization of a coupled nonlinear Schrödinger equation and corresponding implementations. A related study, with a finite difference approach, is given in [11] for a linear Schrödinger-type equation.

In this paper, we extend the uniform time scheme studied in [12] to the case of a two-time-level, nonuniform, linearly implicit finite element scheme. The finite element approach for the spatial discretization is widely studied in full detail inheriting some crucial results from the nonlinear heat equation. Therefore, as mentioned above, the finite element discretization will appear as a background scheme with its crucial results presented in overview form. Hence, although fully discrete, most of the new contribution concerns temporal approximation. In this setting, assuming the regularity of the exact solution $u \in H^2((0,t]; H^r(D))$, we prove the convergence rate of order $\mathscr{O}(k^2 + h^r)$ in the $L_2(D)$-norm and a gradient estimate with accuracy $\mathscr{O}(k + h^{r-1})$ in the $H_0^1(D)$-norm. The L_2-estimate is optimal. As for the gradient estimate, comparing with the theoretical result [8], our gradient estimate is sharp as well. Whereas compared to [12], where the spatial gradient estimate is of order $\mathscr{O}(k^2 + h^{r-1})$, due to the fact that there is no time derivative involved, our result is suboptimal.

An outline of this paper is as follows. In Section 2, we introduce some notation and preliminaries necessary in the analysis. In Section 3, we introduce two related spatial discretization strategies, study the convergence of the simplest one, and derive the optimal semi-discrete error estimates. The results are of overview nature and are presented for the sake of completeness. Section 4 is devoted to the study of a two-level nonuniform grid time discretization of the (background) Galerkin finite element solution obtained in Section 3. In this section, we also include the consistency of the temporal scheme. The convergence analysis is singled out and presented in the concluding Section 5. Finally, in Section 6, we give a conclusion of the results of the paper. Throughout the paper, C denotes a generic constant that might be different at different appearances and is independent of all involved parameters and functions unless otherwise explicitly stated.

2 Notation and Preliminaries

We employ the $L_2(D)$-based complex inner product and the bilinear form

$$(u,v) = \int\limits_D u(x)\overline{v(x)}\,dx, \quad a(u,v) = \int\limits_D \nabla u(x)\overline{\nabla v(x)}\,dx$$

respectively. For a multiindex $\alpha = (\alpha_1, \alpha_2 \dots, \alpha_d)$, $\alpha_i \geqslant 0$, $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$, we recall the standard Sobolev space

$$H^s(D) := \{v : \mathscr{D}^\alpha v \in L_2(D); |\alpha| \leqslant s\},$$

of all complex-valued functions in $L_2(D)$ having all their distribution derivatives of order at most s in $L_2(D)$. The space $H^s(D)$ is associated with the norm

$$\|v\|_s = \|v\|_{H^s(D)} = \Big(\sum_{|\alpha|\leqslant s} \|\mathscr{D}^\alpha v\|^2_{L_2(D)}\Big)^{1/2}$$

and the seminorm

$$|v|_s = \Big(\sum_{|\alpha|=s} \|\mathscr{D}^\alpha v\|^2_{L_2(D)}\Big)^{1/2}.$$

Hence $L_2 = H^0$, and we denote $\|v\| := \|v\|_{L_2} = \|v\|_0$. Further, we define

$$H^1_0(D) := \{v \in H^1(D); v = 0 \quad \text{on} \quad \partial D\}.$$

We also use the space $C^m(D)$ of continuously differentiable functions on D consisting of all complex-valued functions v with all their partial derivatives $\mathscr{D}^\alpha v$ of order $|\alpha| \leqslant m$ being continuous in D. Further, $C^m(\overline{D})$ presents the space of functions $v \in C^m(D)$ for which $\mathscr{D}^\alpha v$ is bounded and uniformly continuous in D for $|\alpha| \leqslant m$. The norm in $C^m(\overline{D})$ is defined by

$$\|v\|_{C^m(\overline{D})} = \max_{|\alpha|\leqslant m} \sup_{x\in\overline{D}} |\mathscr{D}^\alpha v(x)|.$$

In this setting, it is easy to verify that the solution $u(t) := u(t, \cdot)$ satisfies the boundedness relation in the sense that

$$\|u(t)\| \leqslant \|u_0\| + \int\limits_0^t \|g(\tau, \cdot)\|\,d\tau, \quad 0 \leqslant t \leqslant T. \tag{2.1}$$

Frequently, an abstract and extended version of the Sobolev spaces to time dependent functions appear in our reference literature (cf., for example, [13]). Below, we include a brief notation.

For a Banach space X with the norm $\|\cdot\|_X$ the function space $L_p(0,T;X)$ consists of all strongly measurable functions $\boldsymbol{u}:[0,T]\to X$ such that

$$\|\boldsymbol{u}\|_{L_p(0,T;X)} := \begin{cases} \left(\int\limits_0^T \|\boldsymbol{u}(t)\|_X^p \, dt \right)^{1/p}, & 1 \leqslant p < \infty, \\ \operatorname*{ess\,sup}\limits_{0\leqslant t\leqslant T} \|\boldsymbol{u}(t)\|_X, & p = \infty. \end{cases}$$

Then the Sobolev space $W^{s,p}(0,T;X)$ is defined by the boundedness of the norms of its elements; namely,

$$\|\boldsymbol{u}\|_{W^{s,p}(0,T;X)} = \begin{cases} \left(\sum\limits_{m=0}^{s} \int\limits_0^T \left\| \frac{\partial^m \boldsymbol{u}}{\partial t^m} \right\|_X^p dt \right)^{1/p}, & 1 \leqslant p < \infty, \\ \max\limits_{0\leqslant m\leqslant s} \operatorname*{ess\,sup}\limits_{t\in[0,T]} \left\| \frac{\partial^m \boldsymbol{u}}{\partial t^m}(t,\cdot) \right\|_X, & p = \infty. \end{cases}$$

It is obvious that 0 and T can be replaced by any t_1 and t_2 with $0 \leqslant t_1 < t_2$ and hence $[0,T]$ is replaced by $[t_1,t_2]$.

3 Spatial Discretization Scheme

The spatial discretization for the Schrödinger equation is by now a standard procedure considered by several authors. At first glance, it can be viewed as an extension of the results for the heat equation. This however is not a straight-forward strategy, due to the complex terms in the nonlinear Schrödinger equation. Nevertheless, below, for the sake of completeness we introduce two equivalent spatial discretization strategies outlined in [8] and [3] respectively:

SI. Let N be a positive integer, and for $n=1,\dots,N$ let $\{S_h^n\}_{h\in(0,1)} \subset H_0^1(D)$ be a family of finite dimensional subspaces. Consider a partition of the time interval $[0,T]$ into not necessarily uniform subintervals $I_n := (t_{n-1},t_n)$. Let $k_n := |I_n| = t_n - t_{n-1}$ be the length of I_n. Denoting by $u_h^n \in S_h^n$ an approximation of $u(t_n)$, we construct a vector $u_h = (u_h^0, u_h^1, u_h^2, \dots, u_h^n) \in X_h$ by solving a discretized problem of the form

$$\boldsymbol{\Phi}_h(u_h) = 0,$$

where $\boldsymbol{\Phi}_h : X_h \to Y_h$ is a nonlinear mapping such that

$$\begin{cases} X_h = S_h^0 \times S_h^1 \times \cdots \times S_h^N, \\ Y_h = (S_h^0)^* \times (S_h^1)^* \times \cdots \times (S_h^N)^*. \end{cases}$$

where $(S_h^n)^*$ is the dual space of $(S_h^n, \|\cdot\|)$, $n=0,\dots,N$, equipped with the norm

$$\|\cdot\|_* = \sup_{\varphi\in S_h^n} \frac{|\langle\cdot,\varphi\rangle|}{\|\varphi\|}.$$

Note that we have chosen $N+1$ different finite element spaces with S_h^n corresponding to a discrete time level t_n, $n=0,\dots,N$. A detailed stability and convergence analysis in this is given in [4] (cf. also the analysis in [8]).

We focus on a simpler strategy, based on the approach [3] as outlined below:

SII. At each time level, let $\{S_h\}_{h\in(0,1)}$ be a family of finite dimensional subspaces of $\mathscr{H}_0 := H_0^1(\Omega)\cap C(\overline{D})$ satisfying the approximation property

$$\inf_{\chi\in S_h}\Big\{\|v-\chi\|+h\|v-\chi\|_1\Big\}\leqslant Ch^s\|v\|_s \quad \forall v\in H^s(D)\cap H_0^1(D),\quad 2\leqslant s\leqslant r, \tag{3.1}$$

for all $h\in(0,1)$, where s is an integer. Here, for a quasi-uniform family of partitions of D, $\{S_h\}$ is the set of all continuous functions with their real and imaginary parts being piecewise polynomials of degree $r-1$ on D, where $r\geqslant 2$. Then for $\varphi\in S_h$ we have the following inverse inequalities due to [14]:

$$\|\varphi\|_{L_\infty(D)}\leqslant C\begin{cases} h^{-\frac{d}{2}}\|\varphi\|, & d=1,2,3,\\ h^{1-\frac{d}{2}}(\log(1/h))^{1-1/d}\|\nabla\varphi\|, & d=2,3, \end{cases} \qquad \forall\varphi\in S_h. \tag{3.2}$$

Now, using (3.1), (3.2) and assuming the existence of certain operator bounds, one can deduce (cf. [3])

$$\lim_{h\to 0}\sup_{0\leqslant t\leqslant T}\inf_{\varphi\in S_h}\{|u(t)-\varphi|_\infty+h^{-\frac{d}{2}}\|u(t)-\varphi\|\}=0. \tag{3.3}$$

To proceed, for $h\in(0,1)$ we define the discrete Laplacian operator $\Delta_h: S_h\to S_h$ by

$$(\Delta_h\varphi,\chi)=-(\nabla\varphi,\nabla\chi)\quad\forall\varphi,\,\chi\in S_h,$$

and the elliptic projection operator $R_h: H^1(D)\to S_h$ by

$$(\nabla R_h v,\nabla\chi)=(\nabla v,\nabla\chi)\quad\forall v\in H^1(D),\quad\forall\chi\in S_h.$$

Then (cf., for example, [15, 16]) R_h satisfies the approximation properties (3.1); namely,

$$\|R_h v-v\|+h\|R_h v-v\|_1\leqslant C_R h^s\|v\|_s\quad\forall v\in H^s(D)\cap H_0^1(D),\quad 2\leqslant s\leqslant r, \tag{3.4}$$

and

$$\|R_h v(t)-v(t)\|_{L_\infty(D)}\leqslant C_R(\log(1/h)^{\overline{s}}h^s\|v\|_{W^{s,\infty}(D)}$$

for all $h\in(0,1)$, where $\overline{s}$ is equal to 1 if $s=2$ and $d\geqslant 2$ and 0 otherwise. Further,

$$\|\nabla R_h v\|\leqslant\|\nabla v\|\quad\forall v\in H^1(D),\quad\forall h\in(0,1).$$

Now, the time continuous variational formulation for the problem (1.1)–(1.3) reads as follows: Find $u\in H_0^1(D)$ such that

$$\begin{aligned}&(u_t,\chi)+i(\nabla u,\nabla\chi)-i(f(|u|^2)u,\chi)=(g,\chi)\quad\forall\chi\in H_0^1(D),\\ &u(0,x)=u_0(x).\end{aligned}$$

The corresponding finite element problem is formulated as follows: Find $u_h\in S_h$ such that

$$\begin{aligned}&(u_{h,t},\chi)+i(\nabla u_h,\nabla\chi)-i(f(|u_h|^2)u_h,\chi)=(g,\chi)\quad\forall\chi\in S_h,\\ &u_h(0,x)=u_{0,h}(x),\end{aligned} \tag{3.5}$$

where $u_{0,h}$ is an approximation of u_0 in S_h.

Spatial/semidiscrete error estimate. As we mentioned in the introduction, the finite element schemes for the spatial discretization of the Schrödinger equation (1.1) are fully considered in the literature. This section is a brief review of two equivalent spatial discretization strategies adequate in our fully discrete study.

As a crucial property of the finite element scheme (3.5), in analogy with (2.1), we can easily verify that the L_2-norm of the semidiscrete solution $u_h(t) := u_h(\cdot, t)$ is bounded in the following sense:

$$\|u_h(t)\| \leqslant \|u_{0,h}\| + \int_0^t \|g(\tau,\cdot)\|\, d\tau, \quad 0 \leqslant t \leqslant T. \tag{3.6}$$

Then the existence of a unique solution u_h for (3.5) would follow recalling that f is locally Lipschitz together with the relations (3.2) and (3.6).

The convergence estimate for this semidiscrete problem is derived in [3]. Below, for the sake of completeness we outline a concise approach to their proof.

Theorem 3.1. *Assume that $u \in H^s(D)$. Then the finite element solution $u_h(t)$ for (3.5) inherits the convergence rate for an appropriately chosen approximation $u_{0,h}$ for u_0 and yields the optimal convergence rate; namely,*

$$\|u_0 - u_{0,h}\| \leqslant Ch^s \Longrightarrow \max_{0\leqslant t\leqslant T} \|u(t) - u_h(t)\| \leqslant Ch^s\|u\|_s.$$

Proof. For $\varepsilon > 0$ we assume that $M_\varepsilon = \{z \in \mathbb{C} : \exists (x,t) \in \overline{D} \times [0,T]\ |z - u(x,t)| < \varepsilon\}$ and define a globally Lipschitz function $f_\varepsilon : \mathbb{C} \to \mathbb{C}$ such that $f_\varepsilon(z) = f(z)$ for $z \in M_\varepsilon$. Now, we introduce an auxiliary function $w_h : [0,t] \to S_h$ as the unique solution of the problem

$$\begin{aligned} &(w_{h,t}, \chi) + i(\nabla w_h, \nabla\chi) - i(f_\varepsilon(|w_h|^2)w_h, \chi) = (g,\chi) \quad \forall \chi \in S_h, \\ &w_h(0,x) = u_{0,h}(x), \end{aligned} \tag{3.7}$$

The proof is now based on first establishing the auxiliary estimate

$$\max_{0\leqslant t\leqslant T} \|u(t) - w_h(t)\| \leqslant C\Big(\|u_0 - u_{0,h}\| + h^s\|u\|_s\Big) \tag{3.8}$$

and then justifying the fact that u_h and w_h coincide for sufficiently small h. To show (3.8), we use the split $u - w_h = (u - R_h u) + (R_h u - w_h) := \rho + \theta$. Now, a combination of (1.1), (3.5), and (3.7) yields

$$(\theta_t, \chi) + i(\nabla\theta, \nabla\chi) = -(\rho_t, \chi) - i(f_\varepsilon(|w_h|^2)w_h - f_\varepsilon(|u|^2)u, \chi), \quad 0 \leqslant t \leqslant T,$$

where we used the fact that f_ε coincides with f in M_ε. Next, we take $\chi = \theta$ and consider the real part to get

$$\frac{1}{2}\frac{d}{dt}\|\theta(t)\|^2 \leqslant \Big(\|f_\varepsilon(|w_h|^2)w_h - f_\varepsilon(|u|^2)u\| + \|\rho_t\|\Big)\|\theta(t)\|.$$

Consequently

$$\begin{aligned} \frac{d}{dt}\|\theta(t)\| &\leqslant L\|w_h - u\|\|w_h + u\|_\infty\|w_h\| + \|w_h - u\|\|f_\varepsilon(|u|^2)\| + \|\rho_t\| \\ &\leqslant C(\|\theta(t)\| + \|\rho(t)\| + \|\rho_t(t)\| + \|u_0\| + \|u_{0,h}\|) < \end{aligned}$$

where we used the Lipschitz continuity of f with Lipschitz constant L. Here, we assumed that $\max(\|u\|_\infty, \|w_h\|_\infty) < C(t)$, which can be motivated by the stability estimates (2.1) and (3.6) (the latter for w_h replacing u_h). By the Grönwall lemma, the property of the elliptic operator R_h, and (3.4), we deduce the desired estimate (3.8). Combining the assumption of the theorem: $\|u_0 - u_{0,h}\| \leqslant Ch^s$ and (3.8), we get

$$\max_{0\leqslant t\leqslant T} \|u(t) - w_h(t)\| \leqslant Ch^s. \tag{3.9}$$

By (3.2), (3.4), and (3.9), for $0 \leqslant t \leqslant T$ and $\chi \in S_h$ we have

$$\begin{aligned}\|u(t) - w_h(t)\|_\infty &\leqslant \|u(t) - \chi\|_\infty + \|\chi - R_h u(t)\|_\infty + \|R_h u(t) - w_h(t)\|_\infty \\ &\leqslant \|u(t) - \chi\|_\infty + Ch^{-\frac{d}{2}}(\|u(t) - \chi\| + \|\rho(t)\| + \|\theta(t)\|) \\ &\leqslant \|u(t) - \chi\|_\infty + Ch^{-\frac{d}{2}}\|u(t) - \chi\| + Ch^{s-\frac{1}{2}}.\end{aligned}$$

Recalling (3.3), we deduce

$$\exists h_0 > 0 : \forall h \leqslant h_0, w_h(x,t) \in M_\varepsilon, \quad (x,t) \in \overline{D} \times [0,T].$$

For such h we have $u_h = w_h$, and the proof follows from (3.9). □

4 Time Discretization Scheme

Let $N \in \mathbb{N}$, and let $\{t_n\}_{n=0}^N$ be the nodes of a nonuniform partition of the time interval $[0,T]$, i.e. $t_n < t_{n+1}$ for $n = 0,1,\dots,N-1$, $t_0 = 0$ and $t_N = T$. Then we set $k_n := t_n - t_{n-1}$ for $n = 0,1,\dots,N$ and consider the following two time-step numerical scheme:

Step 1. We set

$$U_h^0 = u_{0,h},$$

where $u_{0,h} = R_h u_0$.

Step 2. For $n = 1,2,\dots,N$ we first find $U_h^{n-\frac{1}{2}} \in S_h$ such that

$$\begin{aligned}\left(\frac{U_h^{n-\frac{1}{2}} - U_h^{n-1}}{k_n/2}, \chi\right) =& i\left(\frac{\nabla U_h^{n-\frac{1}{2}} + \nabla U_h^{n-1}}{2}, \nabla\chi\right) \\ &+ i\left(f(|U_n^{n-1}|^2)\frac{U_h^{n-\frac{1}{2}} + U_h^{n-1}}{2}, \chi\right) + (g(t_{n-1},\cdot),\chi)\end{aligned}$$

for all $\chi \in S_h$ and then find $U_h^n \in S_h$ such that

$$\begin{aligned}\left(\frac{U_h^n - U_h^{n-1}}{k_n}, \chi\right) =& i\left(\frac{\nabla U_h^n + \nabla U_h^{n-1}}{2}, \nabla\chi\right) \\ &+ i\left(f(|U_h^{n-\frac{1}{2}}|^2)\frac{U_h^n + U_h^{n-1}}{2}, \chi\right) + \left(g(t_{n-\frac{1}{2}},\cdot),\chi\right)\end{aligned}$$

for all $\chi \in S_h$.

Consistency. Below, we show the consistency of the scheme defined at step 2. To this approach, for $n = 1, 2, \dots, N$, we define $r^{n-\frac{1}{2}}$ and r^n by

$$\begin{aligned}\frac{u^{n-\frac{1}{2}} - u^{n-1}}{k_n/2} =& i\Delta\Big(\frac{u^{n-\frac{1}{2}} + u^{n-1}}{2}\Big) + if(|u^{n-1}|^2)\frac{u^{n-\frac{1}{2}} + u^{n-1}}{2} \\ &+ g(t_{n-1}, \cdot) + r^{n-\frac{1}{2}}\end{aligned} \tag{4.1}$$

and

$$\begin{aligned}\frac{u^n - u^{n-1}}{k_n} =& i\Delta\Big(\frac{u^n + u^{n-1}}{2}\Big) + if(|u^{n-\frac{1}{2}}|^2)\frac{u^n + u^{n-1}}{2} \\ &+ g(t_{n-\frac{1}{2}}, \cdot) + r^n\end{aligned} \tag{4.2}$$

respectively, where $u^n = u(t_n, \cdot)$ for $n = 0, 1, \dots, N$. Then we have the following estimates for $r^{n-\frac{1}{2}}$ and r^n:

Proposition 4.1. *Assume that there is a constant C_1 such that*

$$\max\Big(\|\partial_t u\|_\infty,\ \|\partial_t^2 u\|_\infty,\ \|\Delta\partial_t u\|_\infty\Big) < C_1. \tag{4.3}$$

Then

$$\|r^{n-\frac{1}{2}}\| \leqslant Ck_n.$$

If, in addition to (4.3)*, there is a constant C_2 such that*

$$\max\Big(\|\partial_t^3 u\|_\infty,\ \|\Delta\partial_t^2 u\|_\infty\Big) < C_2, \tag{4.4}$$

then

$$\|r^n\| \leqslant Ck_n^2.$$

Proof. We start by proving the second assertion. Subtracting the Schrödinger equation at time $t = t_{n-\frac{1}{2}}$ from Equation (4.2), we have

$$\begin{aligned}r^n =& \frac{u^n - u^{n-1} - k_n u_t^{n-\frac{1}{2}}}{k_n} - i\Delta\Big(\frac{u^n + u^{n-1} - 2u^{n-\frac{1}{2}}}{2}\Big) \\ &- if(|u^{n-\frac{1}{2}}|^2)\Big(\frac{u^n + u^{n-1} - 2u^{n-\frac{1}{2}}}{2}\Big) := J_1 - J_2 - J_3.\end{aligned}$$

We estimate each $\|J_i\|$, $i = 1, 2, 3$, separately. For this purpose we expand u^n and u^{n-1} in the Taylor series about $t_{n-\frac{1}{2}}$ of degree 2 for J_1 and of degree 1 for J_2 and J_3. As for J_1, by cancellations in the Taylor expansions we end up with

$$J_1 = \frac{1}{6k_n}\int_{t_{n-\frac{1}{2}}}^{t_n} \partial_t^3 u(t, \cdot)(t_n - t)^2\, dt - \frac{1}{6k_n}\int_{t_{n-1}}^{t_{n-\frac{1}{2}}} \partial_t^3 u(t, \cdot)(t_{n-1} - t)^2\, dt.$$

Since $(t_n - t)^2 \leqslant k_n^2/4$ on $(t_{n-\frac{1}{2}}, t_n)$, likewise $(t_{n-1} - t)^2 \leqslant k_n^2/4$ on $(t_{n-1}, t_{n-\frac{1}{2}})$, we use (4.4) to get

$$\|J_1\| \leqslant \frac{k_n}{24} \int_{t_{n-1}}^{t_n} \|\partial_t^3 u(t,\,\cdot)\|\, dt \leqslant C k_n^2.$$

Similarly,

$$J_2 = \frac{i\Delta}{4}\Bigg(\int_{t_{n-\frac{1}{2}}}^{t_n} \partial_t^2 u(t,\,\cdot)(t_n - t)\, dt - \int_{t_{n-1}}^{t_{n-\frac{1}{2}}} \partial_t^2 u(t,\,\cdot)(t_{n-1} - t)\, dt \Bigg),$$

where, using $(t_n - t) \leqslant k_n/2$ on the interval $(t_{n-\frac{1}{2}}, t_n)$ and $(t_{n-1} - t) \leqslant k_n/2$ on $(t_{n-1}, t_{n-\frac{1}{2}})$ together with (4.4), we find

$$\|J_2\| \leqslant \frac{k_n}{8} \int_{t_{n-1}}^{t_n} \|\Delta \partial_t^2 u(t,\,\cdot)\|\, dt \leqslant C k_n^2.$$

For J_3 we have

$$J_3 = \frac{i}{4} f\Big(|u^{n-\frac{1}{2}}|^2\Big)\Bigg(\int_{t_{n-\frac{1}{2}}}^{t_n} \partial_t^2 u(t,\,\cdot)(t_n - t)\, dt - \int_{t_{n-1}}^{t_{n-\frac{1}{2}}} \partial_t^2 u(t,\,\cdot)(t_{n-1} - t)\, dt \Bigg).$$

Therefore, using (4.3) combined with the assumption on f, we have

$$\|J_3\| \leqslant C k_n \int_{t_{n-1}}^{t_n} \|\partial_t^2 u(t,\,\cdot)\|\, dt \leqslant C k_n^2.$$

Consequently, we have proved the second assertion

$$\|r^n\| \leqslant C k_n^2.$$

The proof of the first assertion follows a similar path. However for the sake of completeness we include it as well. This time, we subtract the Schrödinger equation at the time level $t = t_{n-1}$ from the equation (4.1), which yields

$$\begin{aligned} r^{n-1/2} =& \frac{u^{n-1/2} - u^{n-1} - \frac{k_n}{2} u_t^{n-1}}{k_n/2} - i\Delta\Big(\frac{u^{n-1/2} - u^{n-1}}{2}\Big) \\ & - i f(|u^{n-1}|^2)\Big(\frac{u^{n-1/2} - u^{n-1}}{2}\Big) = S_1 - S_2 - S_3. \end{aligned}$$

Below we estimate the norms $\|S_i\|$, $i = 1, 2, 3$, by using the Taylor expansion of $u^{n-1/2}$ about t_{n-1}, of order 1 for S_1 and order 0 for S_2 and S_3. For S_1 we have

$$S_1 = \frac{1}{k_n} \int_{t_{n-1}}^{t_{n-1/2}} \partial_t^2 u(t,\,\cdot)(t_{n-1/2} - t)\, dt.$$

Using (4.3), we deduce

$$\|S_1\| \leqslant \frac{1}{2} \int\limits_{t_{n-1}}^{t_{n-1/2}} \|\partial_t^2 u(t,\,\cdot)\|\,dt \leqslant Ck_n. \tag{4.5}$$

For S_2 we have

$$S_2 = \frac{i\Delta}{2} \int\limits_{t_{n-1}}^{t_{n-1/2}} \partial_t u(t,\,\cdot)\,dt,$$

which, by (4.3), can be estimated as

$$\|S_2\| \leqslant \frac{1}{2} \int\limits_{t_{n-1}}^{t_{n-1/2}} \|\Delta\partial_t u(t,\,\cdot)\|\,dt \leqslant Ck_n. \tag{4.6}$$

Finally, we have

$$S_3 = \frac{i}{2} f\Big(|u^{n-1}|^2\Big) \int\limits_{t_{n-1}}^{t_{n-1/2}} \partial_t u(t,\,\cdot)\,dt,$$

Once again using (4.3), we can estimate S_3 as

$$\|S_3\| \leqslant \frac{1}{2} \|f\Big(|u^{n-1}|^2\Big)\| \int\limits_{t_{n-1}}^{t_{n-1/2}} \|\partial_t u(t,\,\cdot)\|\,dt \leqslant Ck_n. \tag{4.7}$$

Summing up the estimates (4.5), (4.6), and (4.7), we obtain the first assertion of the theorem and the estimate for $r^{n-1/2}$. The proof is complete. □

5 Convergence Analysis

In this part, we rely on the result of [12].

Lemma 5.1. *Let $u_1, u_2 \in C(\overline{D})$, and let $g \in C^1([0,\infty); \mathbf{R})$. Then*

$$\|g\left(|u_1|^2\right) - g\left(|u_2|^2\right)\| \leqslant \sup_{x \in I(u_1,u_2)} |g'(x)|(\|u_1\|_\infty + \|u_2\|_\infty)\|u_1 - u_2\|$$

with $I(u_1, u_2) := [0, \max\{\|u_1\|_\infty^2, \|u_2\|_\infty^2\}]$.

Now, we are ready to formulate the main result of this paper.

Theorem 5.1. *Let $e^n := U_h^n - u^n$ be the error at the time level $t = t_n$. Assume that u satisfies the conditions (4.3) and (4.4). Then there is a constant C such that*

$$\|e^n\| \leqslant C(k^2 + h^r),$$

and

$$\|\nabla e^n\| \leqslant C(k + h^{r-1}),$$

with $k := \max_{1 \leqslant n \leqslant N} k_n$.

Proof. We start by proving the L_2 estimate. For this purpose, we rely on the classical approach and split the error e^n into the Ritz projection error at the time level n and the error between the fully approximate solution U_h^n and the Ritz projection:

$$e^n := U_h^n - u^n = (U_h^n - R_h u^n) + (R_h u^n - u^n) =: \theta^n + \rho^n.$$

We invoke the standard estimate for the Ritz projection error ρ^n from the theory and focus on the estimates for θ^n. Note first that θ^n satisfies the equation

$$\begin{aligned}\left(\frac{\theta^n - \theta^{n-1}}{k_n}, \chi\right) = &- i\left(\nabla\frac{\theta^n + \theta^{n-1}}{2}, \nabla\chi\right) - \left(\frac{\rho^n - \rho^{n-1}}{k_n}, \chi\right)\\ &+ i(\omega^n, \chi) - (r^n, \chi),\end{aligned} \tag{5.1}$$

with

$$\omega^n = f\left(|U_h^{n-1/2}|^2\right)\frac{U_h^n + U_h^{n-1}}{2} - f\left(|u^{n-1/2}|^2\right)\frac{u^n + u^{n-1}}{2}.$$

Choosing $\chi = \theta^n + \theta^{n-1}$ leads to

$$\begin{aligned}\|\theta^n\|^2 - \|\theta^{n-1}\|^2 = &- \frac{ik_n}{2}\|\nabla(\theta^n + \theta^{n-1})\|^2 - \left(\rho^n - \rho^{n-1}, \theta^n + \theta^{n-1}\right)\\ &+ ik_n(\omega^n, \theta^n + \theta^{n-1}) - k_n(r^n, \theta^n + \theta^{n-1}).\end{aligned}$$

It is obvious that the first term on the right-hand side above is purely imaginary. Taking the real part of the other terms on the right-hand side, we end up with

$$\begin{aligned}\|\theta^n\|^2 - \|\theta^{n-1}\|^2 = &- \mathrm{Re}\left(\rho^n - \rho^{n-1}, \theta^n + \theta^{n-1}\right)\\ &- k_n\Big[\mathrm{Im}\,(\omega^n,\, \theta^n + \theta^{n-1}) + \mathrm{Re}\,(r^n,\, \theta^n + \theta^{n-1})\Big].\end{aligned} \tag{5.2}$$

For the first term on the right-hand side we use the mean value theorem together with (3.4) to get the estimate

$$|\left(\rho^n - \rho^{n-1}, \theta^n + \theta^{n-1}\right)| \leqslant Ck_n h^r\left(\|\theta^n\| + \|\theta^{n-1}\|\right). \tag{5.3}$$

The estimate for r^n was derived in the consistency section. It remains to estimate ω^n. To do so, we split

$$\omega^n = \omega_1^n + \omega_2^n,$$

where

$$\begin{aligned}\omega_1^n &:= \left(f\left(|U_h^{n-1/2}|^2\right) - f\left(|u^{n-1/2}|^2\right)\right)\frac{U_h^n + U_h^{n-1}}{2},\\ \omega_2^n &:= f\left(|u^{n-1/2}|^2\right)\left(\frac{e^n + e^{n-1}}{2}\right).\end{aligned}$$

To estimate ω_1^n and ω_2^n, we assume that there exists $\delta > 0$ such that

$$\sup_{t\in[0,T]}\|u(t,\,\cdot)\|_\infty + \max_n\|U_h^n\|_\infty < \delta.$$

By Lemma 5.1, we have the estimates

$$\begin{aligned}\|\omega_1^n\| &\leqslant \delta \sup_{x\in[0,\delta^2]}|f'(x)|\cdot\left(\|U_h^{n-1/2}\|_\infty + \|u^{n-1/2}\|_\infty\right)\|U_h^{n-1/2} - u^{n-1/2}\|\\ &\leqslant C\delta^2\|e^{n-1/2}\|.\end{aligned} \tag{5.4}$$

We also have

$$\|\omega_2^n\| \leqslant \widetilde{C}\Big(\|e^n\| + \|e^{n-1}\|\Big), \tag{5.5}$$

where

$$\widetilde{C} = \frac{1}{2} \sup_{x\in[0,\delta^2]} |f(x)|.$$

Inserting (5.3), (5.4), (5.5) and the estimate for r^n into (5.2), we get

$$\begin{aligned}\|\theta^n\|^2 - \|\theta^{n-1}\|^2 \leqslant &\Big(Ck_n\delta^2\|e^{n-1/2}\| + \widetilde{C}k_n\left(\|e^n\| + \|e^{n-1}\|\right) + Ck_n^3 + Ck_nh^r\Big)\\ &\times\Big(\|\theta^n\| + \|\theta^{n-1}\|\Big).\end{aligned}$$

Consequently,

$$\begin{aligned}\|\theta^n\| - \|\theta^{n-1}\| &\leqslant Ck_n\delta^2\|e^{n-1/2}\| + \widetilde{C}k_n\left(\|e^n\| + \|e^{n-1}\|\right) + Ck_n(k_n^2 + h^r)\\ &\leqslant Ck_n\delta^2\|\theta^{n-1/2}\| + \widetilde{C}k_n\left(\|\theta^n\| + \|\theta^{n-1}\|\right) + Ck_n(k_n^2 + h^r).\end{aligned} \tag{5.6}$$

Hence

$$(1-\widetilde{C}k_n)\|\theta^n\| \leqslant Ck_n\delta^2\|\theta^{n-1/2}\| + (1+\widetilde{C}k_n)\|\theta^{n-1}\| + Ck_n(k_n^2 + h^r).$$

A similar argument for $\theta^{n-1/2}$, using the estimate for $r^{n-1/2}$, reads as follows:

$$\Big(1 - \frac{\widetilde{C}k_n}{2}\Big)\|\theta^{n-1/2}\| \leqslant \Big(1 + \frac{\widetilde{C}k_n}{2} + Ck_n\delta^2\Big)\|\theta^{n-1}\| + Ck_n(k_n + h^r). \tag{5.7}$$

To proceed we assume that $\widetilde{C}k < 1$. Then a combination of (5.6) and (5.7) gives that

$$\|\theta^n\| - \|\theta^{n-1}\| \leqslant \widetilde{C}k_n\|\theta^n\| + C_n\|\theta^{n-1}\| + Ck_n(k_n^2 + h^r),$$

where

$$C_n := \Big(\frac{Ck_n\delta^2(1 + \frac{\widetilde{C}k_n}{2} + Ck_n\delta^2)}{1 - \frac{\widetilde{C}k_n}{2}} + \widetilde{C}k_n\Big).$$

Relabeling n to j and summing over $j = 1, \ldots, n$, we thus get

$$\|\theta^n\| - \|\theta^0\| \leqslant \sum_{j=1}^{n}\Big(\|\theta^j\| - \|\theta^{j-1}\|\Big) \leqslant \sum_{j=1}^{n}\Big(\widetilde{C}k_j\|\theta^j\| + C_j\|\theta^{j-1}\| + Ck_j(k^2 + h^r)\Big).$$

Hence

$$(1-\widetilde{C}k)\|\theta^n\| \leqslant Ct_n(k^2 + h^r) + (1 + C_1)\|\theta^0\| + \sum_{j=1}^{n-1}[\widetilde{C}k_j + C_{j+1}]\|\theta^j\|.$$

Then

$$\|\theta^n\| \leqslant \frac{Ct_n(k^2+h^r)}{1-\widetilde{C}k} + \frac{1+C_1}{1-\widetilde{C}k}\|\theta^0\| + \sum_{j=1}^{n-1}\frac{\widetilde{C}k_j + C_{j+1}}{1-\widetilde{C}k}\|\theta^j\|,$$

so that from the discrete Grönwall inequality we get

$$\|\theta^n\| \leqslant \frac{Ct_n(k^2+h^r)}{1-\widetilde{C}k} \exp\Big[\frac{1+C_1}{1-\widetilde{C}k} + \sum_{j=1}^{n-1} \frac{\widetilde{C}k_j + C_{j+1}}{1-\widetilde{C}k}\Big]. \tag{5.8}$$

Thus, for the error e^n we have the L_2 estimate

$$\|e^n\| \leqslant \|\theta^n\| + \|\rho^n\| \leqslant Ct_n(k^2+h^r).$$

Combining (5.7) and (5.8) we also have

$$\|\theta^{n-1/2}\| \leqslant Ct_n(k^2+h^r),$$

and therefore we have the same L_2 estimate for the error $e^{n-1/2}$ in the intermediate time level:

$$\|e^{n-1/2}\| \leqslant Ct_n(k^2+h^r).$$

It remains to derive the L_2-estimate for the gradient of the error $\|\nabla e^n\|$. Even here, the estimate for $\|\nabla \rho^n\|$ is an approximation theory result, and we need to give an error bound for $\|\nabla\theta^n\|$. For this purpose we choose $\chi = \theta^n - \theta^{n-1}$ in (5.1), which yields

$$\begin{aligned}\frac{1}{k_n}\|\theta^n - \theta^{n-1}\|^2 = -\frac{i}{2}\left(\|\nabla\theta^n\|^2 - \|\nabla\theta^{n-1}\|^2\right) + 2i\,\mathrm{Im}\,(\nabla\theta^n, \nabla\theta^{n-1})\\ -\left(\frac{\rho^n-\rho^{n-1}}{k_n}, \theta^n-\theta^{n-1}\right) + i(\omega^n, \theta^n-\theta^{n-1}) - (r^n, \theta^n-\theta^{n-1}).\end{aligned}$$

Taking the imaginary part of the above relation, we get

$$\begin{aligned}\frac{1}{2}\left(\|\nabla\theta^n\|^2 - \|\nabla\theta^{n-1}\|^2\right) = -\,\mathrm{Im}\left(\frac{\rho^n-\rho^{n-1}}{k_n}, \theta^n-\theta^{n-1}\right)\\ + \mathrm{Re}\,(\omega^n, \theta^n-\theta^{n-1}) - \mathrm{Im}\,(r^n, \theta^n-\theta^{n-1}).\end{aligned}$$

From the Cauchy–Schwarz inequality and the triangle inequality it follows that

$$\|\nabla\theta^n\|^2 - \|\nabla\theta^{n-1}\|^2 \leqslant 2\left(\|\theta^n\| + \|\theta^{n-1}\|\right)\left(\left\|\frac{\rho^n-\rho^{n-1}}{k_n}\right\| + \|\omega^n\| + \|r^n\|\right).$$

Note that we already have estimated the terms in the second factor on the right-hand side. For the θ-terms we use the Poincaré inequality. Therefore,

$$\|\nabla\theta^n\|^2 - \|\nabla\theta^{n-1}\|^2 \leqslant C(k^2+h^r)\left(\|\nabla\theta^n\| + \|\nabla\theta^{n-1}\|\right).$$

Canceling common factors, we get

$$\|\nabla\theta^n\| \leqslant \|\nabla\theta^{n-1}\| + C(k^2+h^r).$$

Applying the above inequality iteratively, we find

$$\|\nabla\theta^n\| \leqslant C(k+h^{r-1}) \tag{5.9}$$

under the assumption that k is proportional to h. The desired estimate for $\|\nabla e^n\|$ follows from (5.9) and (3.4). □

Remark 5.1. Regarding the estimate of $\|\nabla e^n\|$, one would expect an order of $\mathscr{O}(k^2+h^{r-1})$ since we only have spatial derivatives (cf. [12]). However, in [8] (and here) an *optimal* error estimate for the H^1-norm of order $\mathscr{O}(k+h^{r-1})$ is derived. On the other hand, compared to the gradient estimate in [12], our H^1-norm estimate is suboptimal.

6 Concluding Remarks

In this paper, we considered a discretization of a nonlinear Schrödinger equation based on the two-level time stepping scheme with underlying finite element spatial discretization. The nonlinearity is of cubic type with crucial applications. In the spatial discretization, we follow the strategy of [3] based on the classical estimates in [16] and [15]. We also briefly outline a more abstract form of a time-space scheme by [8] and its convergence properties derived in [4]. The crucial steps in the spatial discretization include the split of the error via L_2, H^1 and elliptic projections and then proceed with the argument of dominating the error between approximation and projection with that of the projection error (an error between the exact solution and corresponding projection). Here, both solution and gradient estimates are derived.

Then these results are further used in half-steps in time following the results [12] and the references therein. In this part, we prove the consistency of the numerical scheme, derive stability estimates, and establish the convergence analysis. In the temporal discretization, we considered the $L_\infty(L_2)$ approximations. The L_2 results are optimal of accuracy $\mathscr{O}(k^2 + h^r)$ due to the maximal available regularity of the exact solution, where h and k are spatial and temporal mesh parameters respectively. For the gradient estimates we prove the convergence of order $\mathscr{O}(k + h^{r-1})$.

Acknowledgments

The work is partially supported by the Swedish Foundation of Strategic Research (SSF) and Swedish Research Council (VR).

References

1. C. Besse, "Relaxation scheme for the nonlinear Schrödinger equation and Davey-Stewartson systems" [in French], *C. R. Acad. Sci., Paris, Ser. I, Math.* **326**, No. 12, 1427–1432 (1998).

2. C. Besse, "A relaxation scheme for the nonlinear Schrödinger equation," *SIAM J. Numer. Anal.* **42**, No. 3, 934–952 (2004).

3. G. D. Akrivis, V. A. Dougalis, and O. A. Karakashian, "On fully discrete Galerkin methods of second order temporal accuracy for the nonlinear Schrödinger equation," *Numer. Math.* **59**, No. 1, 31–53 (1991).

4. J. C. Lòpez-Marcos and J. M. Sanz-Serna, "Stability and convergence in numerical analysis II: A definitions of stability of nonlinear problems," In: *Numerical Treatment of Differential Equations*, B.G. Teubner, Stuttgart etc. (1991).

5. W. Strauss, "Mathematical aspects of classical nonlinear field equations," *Lect. Notes Phys.* **98**, 123–149 (1979).

6. G. D. Akrivis, V. A. Dougalis, O. A. Karakashian, and W. R. McKinney, "Numerical approximation of blow-up of radially symmetric solutions of the nonlinear Schrödinger equation," *SIAM J. Sci. Comput.* **25**, No. 1, 186–212 (2003).

7. O. Karakashian, G. D. Akrivis, and V. A. Dougalis, "On optimal order error estimates for the nonlinear Schrödinger equation," *SIAM J. Numer. Anal.* **30**, No. 2, 377–400 (1993).

8. Y. Tourigny, "Optimal H^1 estimates for two-time-discrete Galerkin approximations of a nonlinear Schrödinger equation," *IMA J. Numer. Anal.* **11**, No. 4, 509–523 (1991).

9. Y. Xu and C. W. Shu, "Optimal error estimates of the semidiscrete local discontinuous Galerkin methods for high order wave equations," *SIAM J. Numer. Anal.* **50**, No. 1, 79–104 (2012).

10. M. Asadzadeh, D. Rostamy, and F. Zabihi, "Discontinuous Galerkin and multiscale variational schemes for a coupled damped nonlinear system of Schrödinger equations," *Numer. Methods Partial Differ. Equations* **29**, No. 6, 1912–1945 (2013).

11. G. D. Akrivis, V. A. Dougalis, and G. E. Zouraris, "Finite difference schemes for the "parabolic" equation in a variable depth environment with a rigid bottom boundary condition," *SIAM J. Numer. Anal.* **39**, No. 2, 539–565 (2001).

12. G. E. Zouraris, "On the convergence of the linear two-step finite element method for the nonlinear Schrödinger equation," *Math. Model. Numer. Anal.* **35**, No. 3, 389–405 (2001).

13. L. C. Evans, *Partial Differential Equations*, Am. Math. Soc., Providence, RI (2002).

14. E. Süli, "Convergence and nonlinear stability of the Lagrange-Galerkin method for the Navier-Stokes equations," *Numer. Math.* **53**, No. 3, 459–483 (1988).

15. V. Thomée, *Galerkin Finite-Element Methods for Parabolic Problems*, Springer, Berlin etc. (2006).

16. A. H. Schatz and L. B. Wahlbin, "On the quasi-optimality in L_∞ od the H^1-into finite element spaces," *Math. Comput.* **38**, 1–22 (1982).

Submitted on October 29, 2018