SIAM J. NUMER. ANAL.
Vol. 59, No. 3, pp. 1566–1591

HIGH-ORDER MASS- AND ENERGY-CONSERVING SAV-GAUSS COLLOCATION FINITE ELEMENT METHODS FOR THE NONLINEAR SCHRÖDINGER EQUATION*

XIAOBING FENG†, BUYANG LI‡, AND SHU MA§

Abstract. A family of arbitrarily high-order fully discrete space-time finite element methods are proposed for the nonlinear Schrödinger equation based on the scalar auxiliary variable formulation, which consists of a Gauss collocation temporal discretization and the finite element spatial discretization. The proposed methods are proved to be well-posed and conserving both mass and energy at the discrete level. An error bound of the form $O(h^p + \tau^{k+1})$ in the $L^\infty(0,T;H^1)$-norm is established, where h and τ denote the spatial and temporal mesh sizes, respectively, and (p,k) is the degree of the space-time finite elements. Numerical experiments are provided to validate the theoretical results on the convergence rates and conservation properties. The effectiveness of the proposed methods in preserving the shape of a soliton wave is also demonstrated by numerical results.

Key words. nonlinear Schrödinger equation, mass- and energy-conservation, high-order conserving schemes, SAV-Gauss collocation finite element method, error estimates

AMS subject classifications. 65N12, 65N15

DOI. 10.1137/20M1344998

1. Introduction.

This paper is concerned with the development and analysis of high-order fully discrete numerical methods for the following initial boundary value problem of the nonlinear Schrödinger (NLS) equation:

$$\mathrm{i}\partial_t u - \Delta u - f(|u|^2)u = 0 \quad \text{in } \Omega \times (0,T], \tag{1.1a}$$

$$u = 0 \quad \text{on } \partial\Omega \times (0,T], \tag{1.1b}$$

$$u = u_0 \quad \text{in } \Omega \times \{0\}, \tag{1.1c}$$

where $\Omega \subset \mathbb{R}^d$ is a polygonal or polyhedral domain with boundary $\partial\Omega$, and $u : \Omega \to \mathbb{C}$ is a complex-valued function, with $\mathrm{i} = \sqrt{-1}$, and $f : \mathbb{R}_+ \to \mathbb{R}$ is the derivative of some function $F : \mathbb{R}_+ \to \mathbb{R}$. The best known examples are

$$f(s) = \pm s^{\frac{q-1}{2}} \quad \text{and} \quad F(s) = \pm \frac{2}{q+1} s^{\frac{q+1}{2}} \quad \text{with} \quad q > 1, \tag{1.2}$$

where the "−" and "+" cases are often referred to as defocusing and focusing models, respectively. In the focusing case, the solution will blow up in $L^\infty(\Omega)$ within finite time when the initial energy is negative; see [7, 34]. The NLS equation (1.1) arises

*Received by the editors June 11, 2020; accepted for publication (in revised form) March 10, 2021; published electronically June 9, 2021.
https://doi.org/10.1137/20M1344998
Funding: The work of the first author was partially supported by the National Science Foundation grants DMS-1620168 and DMS-2012414. The work of the second author was partially supported by an internal grant of the Hong Kong Polytechnic University, project ZZKQ. The work of the third author was partially supported by the Hong Kong RGC grant 15300817.
†Department of Mathematics, The University of Tennessee, Knoxville, TN 37996 USA (xfeng@math.utk.edu).
‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (buyang.li@polyu.edu.hk).
§Corresponding author. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (maisie.ma@connect.polyu.hk).

from many applications in physics and engineering, and is one of the fundamental equations in mathematical physics [7, 34, 42, 26, 28].

It is well known that the solutions of (1.1) conserve mass and energy in the sense that for all $t \geq 0$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}|u|^2\mathrm{d}x = 0 \qquad \text{(mass conservation)}, \tag{1.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}\left(\frac{1}{2}|\nabla u|^2 - \frac{1}{2}F(|u|^2)\right)\mathrm{d}x = 0 \qquad \text{(energy conservation)}. \tag{1.4}$$

The development of numerical methods that can retain these conservation properties in numerical solutions is important for long-time numerical simulation and, therefore, has been one of the research focuses in numerical approximation to the NLS equation.

There exists a large amount of literature on numerical solutions and numerical analysis of the NLS equation; see [10, 32, 27, 21, 1, 2, 4, 5, 6, 35, 20, 24, 37, 14, 13]. To the best of our knowledge, all the existing mass- and energy-conserving methods have only second-order accuracy in time and are of the Crank–Nicolson type. No higher-order time-stepping schemes, which conserve both mass and energy, have been reported in the literature. Moreover, the existing error estimates for nonlinearly implicit schemes for the NLS equation generally require certain grid-ratio conditions. The standard grid-ratio conditions in the literature are $\tau = o(h^{\frac{d}{4}})$ for the cubic NLS equation and $\tau = o(h^{\frac{d}{2}})$ for general nonlinearity, where h and τ denote the spatial and temporal mesh sizes. Karakashian and Makridakis [21, 22] proposed some continuous and discontinuous space-time Galerkin finite element methods for the cubic NLS equation and proved optimal-order convergence under a weaker grid-ratio condition $\tau^{k-1}|\ln h| \to 0$ in two dimensions, where $k \geq 2$ is the degree of finite elements in time. For the defocusing cubic NLS equation (or the focusing cubic NLS equation with sufficiently small initial data), using the energy conservation of the numerical scheme, error estimates were established without grid-ratio condition in [16, 36]. For general nonlinearity (possibly focusing), Wang [35] established an error estimate for a linearized semi-implicit scheme without grid-ratio condition; Henning and Peterseim [19] established an error estimate for the nonlinearly implicit Crank–Nicolson finite element method without grid-ratio condition. Both [35] and [19] used an error splitting technique in which they proved boundedness of the numerical solutions by establishing an L^∞-norm error estimate between the fully discrete and the semidiscrete-in-time numerical solutions. The error splitting technique allows us to avoid grid-ratio conditions in using the inverse inequality.

The objective of this paper is to develop a family of arbitrarily higher-order mass- and energy-conserving fully discrete space-time finite element methods based on the scalar auxiliary variable (SAV) formulation of the NLS equation, and to establish the existence, uniqueness, and optimal-order convergence of numerical solutions without grid-ratio condition. Two key ideas are utilized in our construction of the method. First, the SAV reformulation of the NLS equation is used. This approach was introduced in [30, 29] as an enhanced version of the invariant energy quadratization (IEQ) approach [38, 39, 40, 41], for developing energy-decay methods for dissipative (gradient flow) systems. Here we adapt the SAV approach to the dispersive NLS equation, and the SAV reformulation is essential to enable our methods to maintain the energy conservation property at the discrete level. Second, the Gauss collocation method is used for time discretization in the SAV formulation of the NLS equation. The method can be viewed as an efficient implementation of the space-time finite element methods for the SAV formulation with Gauss quadrature in time. The Gauss collocation method was combined with IEQ and SAV to preserve energy decay in solving phase field equations in [3, 17, 18]. We adopt this method here to preserve mass conservation without affecting the energy conservation structure of the SAV formulation.

The SAV formulation introduces new difficulties to error analysis for the NLS equation due to the presence of $\partial_t u$ in the equation of r (see (2.2b)), which leads to a consistency error of suboptimal order in time and introduces new difficulty in obtaining the stability estimate. These difficulties are overcome by combining three techniques. First, inspired by the error analysis of Karakashian and Makridakis [22], our proof makes use of properties of the Legendre polynomials on each interval I_n, rewriting the Gauss collocation method into a space-time Galerkin finite element method, which makes it easier to choose suitable test functions in the error estimation. Second, we introduce a temporal Ritz projection and use a superapproximation result of the temporal local L^2 projection to eliminate the suboptimal temporal consistency error caused by $\partial_t u$ in the equation of r. Third, we estimate the time derivative of the error in $H^{-1}(\Omega)$ using a duality argument, which leads to an optimal-order H^1-norm error estimate. We prove the existence, uniqueness, and optimal-order convergence of numerical solutions based on Schaefer's fixed point theorem in an L^∞-neighborhood of the exact solution. This allows us to avoid grid-ratio conditions for the NLS equation with general nonlinearity.

The rest of this paper is organized as follows. In section 2, we present the SAV reformulation of the NLS equation and introduce our SAV space-time Gauss collocation finite element method. In section 3, we first present an integral reformulation of the proposed method and then establish its mass and energy conservation properties. We also derive a consistency error estimate for the method, which is vitally used to prove an error estimate in the subsequent section. In section 4, we first establish the well-posedness of the numerical method and then prove an error bound of the form $O(h^p + \tau^{k+1})$ in the energy norm, where τ and h denote the temporal and spatial mesh sizes, respectively, with (p, k) denoting the degree of polynomials in the space-time finite element method. Finally, in section 5, we present a few numerical tests to validate the theoretical results, and to demonstrate the effectiveness of the proposed method in preserving the shape of a soliton wave.

Throughout this paper, unless stated otherwise, C will be used to denote a generic positive constant which is independent of τ, h, n, and N, but may depend on T and the regularity of solution.

2. Formulation of the SAV-Gauss collocation finite element method.

In this section, we construct a Gauss collocation finite element method based on the SAV reformulation of the NLS equation.

2.1. Function spaces. Let $H^k(\Omega)$, $k \geq 0$, be the conventional complex-valued Sobolev space of functions on Ω, and denote

$$L^2(\Omega) = H^0(\Omega) \quad \text{and} \quad H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}.$$

We denote by $(\cdot,\cdot)$ and $\|\cdot\|$ the inner product and norm of the complex-valued Hilbert space $L^2(\Omega)$, respectively, defined by

$$(u, v) := \int_\Omega u\,\overline{v}\,\mathrm{d}x \quad \text{and} \quad \|u\| := \sqrt{(u,u)}.$$

For $m, s \geq 0$ and $1 \leq p \leq \infty$, the notation $W^{m,p}(0,T; H^s(\Omega))$ stands for the space-time Sobolev space of functions which are $W^{m,p}$ in time and H^s in space; see [11, Chapter 5.9]. We abbreviate the norms of $H^s(\Omega)$ and $W^{m,p}(0,T;H^s(\Omega))$ as $\|\cdot\|_{H^k}$ and $\|\cdot\|_{W^{m,p}(I_n;H^s)}$, respectively, omitting the dependence on Ω in the subscripts.

2.2. The SAV reformulation of (1.1). The SAV formulation of the NLS equation (cf. [29]) introduces an SAV

$$r = \sqrt{\int_\Omega \tfrac{1}{2}F(|u|^2)\mathrm{d}x + c_0} \quad \text{with} \quad g(u) = \frac{f(|u|^2)}{\sqrt{\int_\Omega \frac{1}{2}F(|u|^2)\mathrm{d}x + c_0}} \tag{2.1}$$

with a positive c_0 (which guarantees that the function r has a positive lower bound), and reformulate (1.1) as

$$\mathrm{i}\partial_t u - \Delta u - rg(u)u = 0 \qquad \text{in } \Omega \times (0,T], \tag{2.2a}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \mathrm{Re}\big(\tfrac{1}{2}g(u)u, \partial_t u\big) \qquad \text{in } \Omega \times (0,T], \tag{2.2b}$$

$$u = 0 \qquad \text{on } \partial\Omega \times (0,T], \tag{2.2c}$$

$$u = u_0, \quad r = r_0 \qquad \text{in } \Omega \times \{0\}, \tag{2.2d}$$

where $r_0 = \sqrt{\int_\Omega \frac{1}{2}F(|u_0|^2)\mathrm{d}x + c_0}$. The mass and energy conservation in the SAV formulation are

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |u|^2\mathrm{d}x = 0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\int_\Omega |\nabla u|^2\mathrm{d}x - r^2 + c_0\right) = 0. \tag{2.3}$$

2.3. Space-time finite element spaces. Let $\mathcal{T}_h$ be a shape-regular and quasi-uniform triangulation of Ω with mesh size $h \in (0,1)$ and $\{t_n\}_{n=0}^N$ be a uniform partition of $[0,T]$ with the time step size $\tau \in (0,1)$, where N is a positive integer and hence $\tau = \frac{T}{N}$. For an integer $p \geq 1$ we denote by $\mathbb{Q}^p$ the space of complex-valued polynomials of degree $\leq p$ in space, and we denote by S_h the complex-valued Lagrange finite element space subject to the triangulation of Ω, defined by

$$S_h = \big\{v \in C(\overline{\Omega}) : v|_K \in \mathbb{Q}^p \text{ for all } K \in \mathcal{T}_h, \ v = 0 \text{ on } \partial\Omega\big\},$$

where $C(\overline{\Omega})$ denotes the space of complex-valued uniformly continuous functions on Ω. Then S_h is a complex Hilbert space with the inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$.

For an integer $k \geq 1$, let $\mathbb{P}^k$ denote the space of real-valued polynomials of degree $\leq k$ in t. For a Banach space X, such as $X = L^2(\Omega)$ or $X = S_h$, we define the following tensor-product space:

$$\mathbb{P}^k \otimes X := \mathrm{span}\Big\{p(t)\phi(x) : p \in \mathbb{P}^k, \phi \in X\Big\} = \Big\{\textstyle\sum_{j=0}^k t^j \phi_j : \phi_j \in X\Big\}. \tag{2.4}$$

Moreover, let $P_h : L^2(\Omega) \to S_h$ denote the L^2 projection operator defined by

$$\big(w - P_h w, v_h\big) = 0 \quad \forall\, v_h \in S_h \ \forall\, w \in L^2(\Omega).$$

The following stability properties are well known (cf. [8]):

$$\|P_h w\| \leq \|w\| \qquad \forall\, w \in L^2(\Omega), \tag{2.5a}$$

$$\|P_h w\|_{H^1} \leq C\|w\|_{H^1} \qquad \forall\, w \in H_0^1(\Omega), \tag{2.5b}$$

where C depends only on the shape regularity and quasi-uniformity of the mesh.

We also introduce the global space-time finite element spaces

$$X_{\tau,h} = \{v_h \in C([0,T]; S_h) : v_h|_{I_n} \in \mathbb{P}^k \otimes S_h \text{ for } n = 1,\dots,N\}, \tag{2.6}$$

$$Y_{\tau,h} = \{q_h \in C([0,T]) : q_h|_{I_n} \in \mathbb{P}^k \text{ for } n = 1,\dots,N\}. \tag{2.7}$$

2.4. SAV-Gauss collocation finite element method. Let c_j and w_j, $j = 1, \dots, k$, be the nodes and weights of the k-point Gauss quadrature rule in the interval $[-1, 1]$ (see [31, Table 3.1]), and let $t_{nj} = t_{n-1} + (1 + c_j)\tau/2$, $j = 1, \dots, k$, denote the Gauss points in the interval $I_n = [t_{n-1}, t_n]$. We define the following Gauss collocation finite element method for (2.2).

Main Algorithm.

Step 1: Set $u_h^0 := I_h u_0$ and $r_h^0 := r_0$, where I_h is the Lagrange interpolation operator onto the finite element space. Determine $(u_h, r_h) \in X_{\tau,h} \times Y_{\tau,h}$ by the following two steps.

Step 2: For $n = 1, 2, \dots, N$, define $\{(u_h(t_{nj}), r_h(t_{nj}))\}_{j=1}^k \subset S_h \times \mathbb{R}$ by solving recursively (in n) the following nonlinear (algebraic) system:

$$\text{(2.8a)} \quad \mathrm{i}\big(\partial_t u_h(t_{nj}), v_h\big) + \big(\nabla u_h(t_{nj}), \nabla v_h\big) - \big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), v_h\big) = 0 \qquad \forall\, v_h \in S_h,$$

$$\text{(2.8b)} \quad \partial_t r_h(t_{nj}) = \frac{1}{2} \mathrm{Re}\big(g(u_h(t_{nj})) u_h(t_{nj}), \partial_t u_h(t_{nj})\big),$$

$$\text{(2.8c)} \quad u_h(t_{n-1}) = u_h^{n-1}, \quad \text{and} \quad r_h(t_{n-1}) = r_h^{n-1}.$$

Step 3: Set $u_h^n := u_h(t_n)$ and $r_h^n := r_h(t_n)$.

Remark 2.1. (a) We note that in (2.8a) and (2.8b), $\partial_t u_h(t_{nj}) = \partial_t u_h(t)|_{t=t_{nj}}$ and $\partial_t r_h(t_{nj}) = \partial_t r_h(t)|_{t=t_{nj}}$. Main Algorithm actually computes $\{(u_h(t_{nj}), r_h(t_{nj}))\}_{j=1}^k$ for each $n \geq 1$; however, since any kth-order polynomial on I_n is uniquely determined by its initial value at t_{n-1} and its values at the k Gauss points t_{nj}, $j = 1, \dots, k$, then the Gauss-point values generated by Main Algorithm uniquely determine the pair $(u_h, r_h) \in X_{\tau,h} \times Y_{\tau,h}$.

(b) Each of (2.8a) and (2.8b) consists of nonlinear algebraic equations; note that the test function v_h can be different for different j's, and one "initial condition" is prescribed for each of u_h and r_h. The number of equations imposed is the same as the degrees of freedom which equals the dimension of the space $\mathbb{P}^k \otimes S_h$ for each n.

(c) Main Algorithm can be obtained by applying the Gauss quadrature rule (in time) to a (continuous) space-time finite element method for (2.2); see section 3.1.

(d) In practical computation, we solve for the solution of the nonlinear scheme (2.8) by Newton's method: For given $\{(u_h^{\ell-1}(t_{nj}), r_h^{\ell-1}(t_{nj}))\}_{j=1}^k \subset S_h \times \mathbb{R}$, find

$$\big\{(u_h^\ell(t_{nj}), r_h^\ell(t_{nj}))\big\}_{j=1}^k \subset S_h \times \mathbb{R}$$

satisfying the linearized equations

$$\begin{aligned} \text{(2.9a)} \quad & \mathrm{i}\big(\partial_t u_h^\ell(t_{nj}), v_h\big) + \big(\nabla u_h^\ell(t_{nj}), \nabla v_h\big) \\ &= \big(r_h^\ell(t_{nj}) g(u_h^{\ell-1}(t_{nj})) u_h^{\ell-1}(t_{nj}), v_h\big) \\ &\quad + \big(r_h^{\ell-1}(t_{nj}) g_1(u_h^{\ell-1}(t_{nj}))(u_h^\ell(t_{nj}) - u_h^{\ell-1}(t_{nj})), v_h\big) \\ &\quad + \big(r_h^{\ell-1}(t_{nj}) g_2(u_h^{\ell-1}(t_{nj}))(\bar{u}_h^\ell(t_{nj}) - \bar{u}_h^{\ell-1}(t_{nj})), v_h\big) \qquad \forall\, v_h \in S_h, \end{aligned}$$

(2.9b)

$$\begin{aligned} \partial_t r_h^\ell(t_{nj}) &= \frac{1}{2} \mathrm{Re}\big(g(u_h^{\ell-1}(t_{nj})) u_h^{\ell-1}(t_{nj}), \partial_t u_h^\ell(t_{nj})\big) \\ &\quad + \frac{1}{2} \mathrm{Re}\big(g_1(u_h^{\ell-1}(t_{nj}))(u_h^\ell(t_{nj}) - u_h^{\ell-1}(t_{nj})), \partial_t u_h^{\ell-1}(t_{nj})\big), \\ &\quad + \frac{1}{2} \mathrm{Re}\big(g_2(u_h^{\ell-1}(t_{nj}))(\bar{u}_h^\ell(t_{nj}) - \bar{u}_h^{\ell-1}(t_{nj})), \partial_t u_h^{\ell-1}(t_{nj})\big), \end{aligned}$$

(2.9c)
$$u_h^\ell(t_{n-1}) = u_h^{n-1}, \quad \text{and} \quad r_h^\ell(t_{n-1}) = r_h^{n-1},$$

where

$$g_1(u) := \partial_u[g(u)u] \quad \text{and} \quad g_2(u) := \partial_{\bar u}[g(u)u],$$

and $\partial_{\bar u}$ denotes the differentiation with respect to $\bar u$ in the expression

$$g(u)u = \frac{f(u\bar u)u}{\sqrt{\int_\Omega \frac{1}{2}F(u\bar u)\mathrm{d}x + c_0}}.$$

The iteration in ℓ is set to stop when the desired tolerance error is achieved.

3. Conservation, stability, and consistency analysis.

3.1. A reformulation of scheme (2.8a)–(2.8b). In this subsection, we present several integral identities and inequalities, including a reformulation of Main Algorithm. These identities and inequalities will be used in the subsequent analysis of existence, uniqueness, and convergence of numerical solutions.

Consider the interval $I_n = [t_{n-1}, t_n]$; then we define $P_\tau^n : L^2(I_n; L^2(\Omega)) \to \mathbb{P}^{k-1} \otimes L^2(\Omega)$ to be the L^2 projection defined by

$$\int_{I_n} (u - P_\tau^n u, v)\,\mathrm{d}t = 0 \quad \forall\, v \in \mathbb{P}^{k-1} \otimes L^2(\Omega). \tag{3.1}$$

Thus $u - P_\tau^n u$ is orthogonal to all temporal polynomials of degree $\le k-1$, which means that if $u \in \mathbb{P}^k \otimes L^2(\Omega)$, then

$$u - P_\tau^n u = \phi_{n-1} L_k, \tag{3.2}$$

where $\phi_{n-1} \in L^2(\Omega)$ and

$$L_k(t) := \widehat{L}_k\left(\frac{2t - t_{n-1} - t_n}{\tau}\right) \tag{3.3}$$

is the shifted Legendre polynomial (orthogonal to polynomials of lower degree on I_n). The temporal L^2 projection operator P_τ^n has the following approximation property (cf. [9]):

$$\max_{t\in I_n} \|v - P_\tau^n v\|_X \le C\tau^m \max_{t\in I_n} \|\partial_t^m v\|_X, \quad 0 \le m \le k, \tag{3.4}$$

for all $v \in C^k([0,T]; X)$, where $X = \mathbb{R}$ or $X = H^s(\Omega)$ for some $s \in \mathbb{R}$.

Since the k-point Gauss quadrature holds exactly for polynomials of degree $2k-1$ (cf. [15, p. 222]), and the Gauss points t_{nj}, $j = 1, \dots, k$, are the roots of the Legendre polynomial $L_k(t)$ (cf. [23, p. 33]), it follows that the following two identities hold:

$$\int_{I_n} v(t)\mathrm{d}t = \frac{\tau}{2}\sum_{j=1}^k v(t_{nj})w_j \qquad \forall\, v \in \mathbb{P}^{2k-1} \otimes S_h, \tag{3.5}$$

$$v(t_{nj}) = P_\tau^n v(t_{nj}) \qquad \forall\, v \in \mathbb{P}^k \otimes S_h. \tag{3.6}$$

Setting $v_h = \frac{\tau}{2} v_h(t_{nj}) w_j$ in (2.8a) and summing up the results for $j = 1, \dots, k$, and using (3.5)–(3.6) in the first two terms yield the following integral identity:

$$\begin{aligned}(3.7)\qquad &\int_{I_n} \mathrm{i}\big(\partial_t u_h, v_h\big)\mathrm{d}t + \int_{I_n} (\nabla P_\tau^n u_h, \nabla v_h)\mathrm{d}t \\ &\quad - \frac{\tau}{2}\sum_{j=1}^{k} w_j (r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), v_h(t_{nj})) = 0 \qquad \forall\, v_h \in \mathbb{P}^k \otimes S_h.\end{aligned}$$

Similarly, multiplying (2.8b) by $\frac{\tau}{2} q_h(t_{nj}) w_j$ and summing up the results for $j = 1, \dots, k$, and using (3.5) in the first term, we have

$$(3.8)\qquad \int_{I_n} \partial_t r_h q_h \mathrm{d}t = \frac{\tau}{2}\sum_{j=1}^{k} \frac{w_j}{2} \mathrm{Re}\big(g(u_h(t_{nj})) u_h(t_{nj}), \partial_t u_h(t_{nj})\, q_h(t_{nj})\big) \quad \forall\, q_h \in \mathbb{P}^k.$$

(3.7)–(3.8) provides a reformulation of Main Algorithm. The above reformulation will be crucially used later to show mass and energy conservations, as well as existence, uniqueness, and convergence of numerical solutions.

From (3.2) we get

$$\begin{aligned}\|\phi_{n-1}\| &= \frac{1}{|L_k(t_{n-1})|}\|u_h(t_{n-1}) - P_\tau^n u_h(t_{n-1})\| \\ &\le C\|u_h(t_{n-1})\| + C\left(\frac{1}{\tau}\int_{I_n}\|P_\tau^n u_h(t)\|^2 \mathrm{d}t\right)^{\frac{1}{2}},\end{aligned}$$

where we have used the inverse inequality in time. Thus, by using (3.2) again, we obtain the following inequality:

$$(3.9)\qquad \int_{I_n}\|u_h\|^2\mathrm{d}t \le C\int_{I_n}\|P_\tau^n u_h\|^2\mathrm{d}t + C\tau\|u_h(t_{n-1})\|^2 \qquad \forall\, u_h \in \mathbb{P}^k \otimes S_h.$$

By using the two identities (3.5)–(3.6), one can also prove the following inequality:

$$(3.10)\qquad \frac{\tau}{2}\sum_{j=1}^{k} w_j\|v_h(t_{nj})\|^2 = \int_{I_n}\|P_\tau^n v_h(t)\|^2\mathrm{d}t \le \int_{I_n}\|v_h(t)\|^2\mathrm{d}t \quad \forall\, v_h \in \mathbb{P}^k \otimes S_h.$$

The inequalities (3.9)–(3.10) will be frequently used in the subsequent error analysis.

3.2. Mass and energy conservation properties. In this subsection, we prove the following conservation properties of the numerical solution, which comprise the first main theorem of this paper.

THEOREM 3.1. *Let $(u_h, r_h) \in X_{\tau,h} \times Y_{\tau,h}$ be a solution of Main Algorithm, then the following mass and energy conservations hold:*

$$\begin{aligned}\frac{1}{2}\|u_h(t_n)\|^2 &= \frac{1}{2}\|u_h(t_0)\|^2 && \textit{for } n \ge 1,\\ \frac{1}{2}\|\nabla u_h(t_n)\|^2 - |r_h(t_n)|^2 + c_0 &= \frac{1}{2}\|\nabla u_h(t_0)\|^2 - |r_h(t_0)|^2 + c_0 && \textit{for } n \ge 1.\end{aligned}$$

Proof. Setting $v_h = u_h \in \mathbb{P}^k \otimes S_h$ in (3.7) and taking the imaginary part yield

$$\operatorname{Im}\int_{I_n} \mathrm{i}\big(\partial_t u_h, u_h\big)\mathrm{d}t = -\operatorname{Im}\int_{I_n} (\nabla P_\tau^n u_h, \nabla u_h)\mathrm{d}t + \operatorname{Im}\Big[\frac{\tau}{2}\sum_{j=1}^{k} w_j (r_h(t_{nj}) g(u_h(t_{nj})), |u_h(t_{nj})|^2)\Big] = 0, \tag{3.11}$$

where we have used the definition of the projection operator P_τ^n, which implies

$$\operatorname{Im}\int_{I_n} (\nabla P_\tau^n u_h, \nabla u_h)\mathrm{d}t = \operatorname{Im}\int_{I_n} (\nabla P_\tau^n u_h, \nabla P_\tau^n u_h)\mathrm{d}t = 0.$$

Then the mass conservation follows from (3.11) and the identity

$$\operatorname{Im}\int_{I_n} \mathrm{i}\big(\partial_t u_h, u_h\big)\mathrm{d}t = \frac{1}{2}\|u_h(t_n)\|^2 - \frac{1}{2}\|u_h(t_{n-1})\|^2.$$

Alternatively, setting $v_h = \partial_t u_h$ and $q_h = 2r_h$ in (3.7) and (3.8), respectively, and taking the real parts yield

$$\operatorname{Re}\int_{I_n} (\nabla P_\tau^n u_h, \nabla \partial_t u_h)\mathrm{d}t = \frac{\tau}{2}\operatorname{Re}\sum_{j=1}^{k} w_j (r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), \partial_t u_h(t_{nj})), \tag{3.12}$$

$$|r_h(t_n)|^2 - |r_h(t_{n-1})|^2 = \frac{\tau}{2}\operatorname{Re}\sum_{j=1}^{k} w_j \big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), \partial_t u_h(t_{nj})\big). \tag{3.13}$$

Since

$$\operatorname{Re}\int_{I_n} (\nabla P_\tau^n u_h, \nabla \partial_t u_h)\mathrm{d}t = \operatorname{Re}\int_{I_n} (P_\tau^n \nabla u_h, \nabla \partial_t u_h)\mathrm{d}t = \operatorname{Re}\int_{I_n} (\nabla u_h, \nabla \partial_t u_h)\mathrm{d}t = \frac{1}{2}\|\nabla u_h(t_n)\|^2 - \frac{1}{2}\|\nabla u_h(t_{n-1})\|^2,$$

it follows that

$$\frac{1}{2}\|\nabla u_h(t_n)\|^2 - \frac{1}{2}\|\nabla u_h(t_{n-1})\|^2 = \frac{\tau}{2}\operatorname{Re}\sum_{j=1}^{k} w_j \big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), \partial_t u_h(t_{nj})\big). \tag{3.14}$$

Subtracting (3.13) from (3.14) yields

$$\frac{1}{2}\|\nabla u_h(t_n)\|^2 - |r_h(t_n)|^2 = \frac{1}{2}\|\nabla u_h(t_{n-1})\|^2 - |r_h(t_{n-1})|^2 \quad \text{for } n \geq 1. \tag{3.15}$$

Thus, the energy conservation holds. The proof is complete. ∎

3.3. An upper bound of mass at internal stages. In this subsection, we prove that the average mass of numerical solutions at internal stages has an upper bound unconditionally (independent of the regularity of solutions). This property furthermore strengthens the stability of numerical solutions when the exact solution is not smooth (for example, close to blowup).

THEOREM 3.2. *Let $(u_h, r_h) \in X_{\tau,h} \times Y_{\tau,h}$ be a solution of Main Algorithm, then the following inequalities hold:*

$$\max_{1\le n\le N} \frac{1}{\tau}\int_{I_n} \|P_\tau^n u_h\|^2 \mathrm{d}t \le \|u_h(0)\|^2, \tag{3.16a}$$

$$\max_{1\le n\le N}\max_{1\le j\le k} \|u_h(t_{nj})\| \le C\|u_h(0)\|, \tag{3.16b}$$

where C is a constant independent of τ, h, and the regularity of the solution.

Proof. By the definition of the temporal L^2 projection P_τ^n, we get

$$\begin{aligned}
&\int_{I_n} \|P_\tau^n u_h(t)\|^2 \mathrm{d}t \\
&= \mathrm{Re}\int_{I_n} (u_h(t), P_\tau^n u_h(t))\mathrm{d}t \\
&= \mathrm{Re}\,(u_h(t_{n-1}), P_\tau^n u_h(t_{n-1}))\tau + \mathrm{Re}\int_{I_n} (\partial_t u_h(t), (t_n - t)P_\tau^n u_h(t))\mathrm{d}t \\
&\quad + \mathrm{Re}\int_{I_n} (u_h(t), (t_n - t)\partial_t P_\tau^n u_h(t))\mathrm{d}t =: J_1 + J_2 + J_3,
\end{aligned} \tag{3.17}$$

where we have interchanged the order of integration in deriving the second to last equality. It can be shown (cf. [12]) that $J_2 = 0$ and

$$\begin{aligned}
J_1 &\le \frac{\tau}{2}\|u_h(t_{n-1})\|^2 + \frac{\tau}{2}\|P_\tau^n u_h(t_{n-1})\|^2, \\
J_3 &= -\frac{\tau}{2}\|P_\tau^n u_h(t_{n-1})\|^2 + \int_{I_n} \frac{1}{2}\|P_\tau^n u_h(t)\|^2 \mathrm{d}t.
\end{aligned}$$

Substituting the estimates of J_1, J_2, and J_3 into (3.17) gives (3.16a).

Substituting (3.16a) into (3.9) and using the mass conservation property again, we obtain $\int_{I_n} \|u_h\|^2 \mathrm{d}t \le C\tau\|u_h(0)\|^2$, and an application of the inverse inequality yield (3.16b). The proof is complete. ∎

3.4. Temporal and spatial Ritz projections. Let $I_\tau^n u$ and $I_\tau^n r$ be the temporal Lagrange interpolation polynomials of u and r, respectively, interpolated at the $k+1$ points t_{n-1} and t_{nj}, $j = 1, \dots, k$. It is well known that the following approximation property (cf. [9]) holds:

$$\max_{t\in I_n}\big(\|v - I_\tau^n v\|_X + \tau\|\partial_t(v - I_\tau^n v)\|_X\big) \le C\tau^{m+1}\max_{t\in I_n}\|\partial_t^{m+1} v\|_X \tag{3.18}$$

for all $v \in C^{m+1}([0,T];X)$, $0 \le m \le k$, and $X = \mathbb{R}$ or $X = H^s(\Omega)$ for some $s \in \mathbb{R}$. We also define a temporal Ritz projection operator $R_\tau^n : W^{1,\infty}(I_n; L^2(\Omega)) \to \mathbb{P}^k \otimes L^2(\Omega)$ as follows:

$$\int_{I_n} (\partial_t(u - R_\tau^n u), v)\mathrm{d}t = 0 \quad \forall\, v \in \mathbb{P}^{k-1} \otimes L^2(\Omega), \tag{3.19}$$

$$u(t_{n-1}) - R_\tau^n u(t_{n-1}) = 0. \tag{3.20}$$

By using this property and the shifted Legendre polynomials defined in (3.3), we can express the temporal Ritz projection as

$$R_\tau^n u(t) = u(t_{n-1}) + \sum_{j=0}^{k-1} \frac{\int_{I_n} L_j(s)\partial_s u(s)\mathrm{d}s}{\int_{I_n} |L_j(s)|^2 \mathrm{d}s}\int_{t_{n-1}}^t L_j(s)\mathrm{d}s, \tag{3.21}$$

which implies that if $X \subset L^2(\Omega)$ is a Banach space and $u \in W^{1,\infty}(I_n; X)$, then $R_\tau^n u$ is automatically in $\mathbb{P}^k \otimes X$. It can be shown that R_τ^n satisfies the following approximation property; see [12, Lemma 3.3].

LEMMA 3.3. *Let $X = \mathbb{R}$ or $H^s(\Omega)$ for some $s \geq 0$. For $u \in W^{m+1,\infty}(I_n; X)$, with $0 \leq m \leq k$, the following approximation property holds:*

$$\|u - R_\tau^n u\|_{L^\infty(I_n;X)} + \tau\|\partial_t(u - R_\tau^n u)\|_{L^\infty(I_n;X)} \leq C\tau^{m+1}\|u\|_{W^{m+1,\infty}(I_n;X)}.$$

In addition to the above optimal-order approximation result, we also have the following superconvergence result.

LEMMA 3.4 (a superapproximation property). *Let $X = \mathbb{R}$ or $H^s(\Omega)$ for some $s \geq 0$. If $w \in W^{k,\infty}(I_n; W^{s,\infty}(\Omega))$ and $v \in \mathbb{P}^{k-1} \otimes X$, then*

$$\|wv - P_\tau^n(wv)\|_{L^2(I_n;X)} \leq C\tau\|v\|_{L^2(I_n;X)}.$$

Proof. We only give a proof for the case $X = H^s(\Omega)$ because the other cases are similar. By applying (3.4) with $m = k$, we have

$$\begin{aligned}
\|wv - P_\tau^n(wv)\|_{L^2(I_n;H^s)} &\leq C\tau^{\frac{1}{2}}\|wv - P_\tau^n(wv)\|_{L^\infty(I_n;H^s)} \\
&\leq C\tau^{k+\frac{1}{2}}\|\partial_t^k(wv)\|_{L^\infty(I_n;H^s)} \\
&\leq C\sum_{m=0}^{k-1}\tau^{k+\frac{1}{2}}\|\partial_t^{k-m}w\partial_t^m v\|_{L^\infty(I_n;H^s)} \quad (\text{since } \partial_t^k v = 0) \\
&\leq C\sum_{m=0}^{k-1}\tau^{k+\frac{1}{2}}\|\partial_t^{k-m}w\|_{L^\infty(I_n;W^{s,\infty})}\|\partial_t^m v\|_{L^\infty(I_n;H^s)} \\
&\leq C\sum_{m=0}^{k-1}\tau^{k+\frac{1}{2}-m}\|v\|_{L^\infty(I_n;H^s)} \\
&\leq C\tau^{\frac{3}{2}}\|v\|_{L^\infty(I_n;H^s)} \\
&\leq C\tau\|v\|_{L^2(I_n;H^s)}.
\end{aligned}$$

Here we have used the inverse inequality in time twice. The proof is complete. ∎

Finally, we also recall the (spatial) Ritz projection operator $R_h : H_0^1(\Omega) \to S_h$ defined by

$$\big(\nabla(w - R_h w), \nabla v_h\big) = 0 \qquad \forall v_h \in S_h \ \ \forall\, w \in H_0^1(\Omega),$$

and the discrete Laplacian operator $\Delta_h : S_h \to S_h$ defined by

$$(\Delta_h \phi_h, \chi_h) := -(\nabla\phi_h, \nabla\chi_h) \quad \forall\, \phi_h, \chi_h \in S_h. \tag{3.22}$$

It is known [8] that there hold the following identities:

$$P_h\Delta v = \Delta_h R_h v \qquad \forall\, v \in H_0^1(\Omega), \tag{3.23a}$$

$$R_\tau^n R_h v = R_h R_\tau^n v \qquad \forall\, v \in W^{1,\infty}(I_n; H_0^1(\Omega)), \tag{3.23b}$$

$$R_\tau^n \Delta_h v_h = \Delta_h R_\tau^n v_h \qquad \forall\, v \in W^{1,\infty}(I_n; S_h). \tag{3.23c}$$

Moreover, there holds the following approximation property (cf. [8]):

$$\|v - R_h v\|_{H^1} \leq Ch^p\|v\|_{H^{p+1}} \quad \forall\, v \in H_0^1(\Omega) \cap H^{p+1}(\Omega). \tag{3.24}$$

3.5. Consistency of scheme (2.8a)–(2.8b). We define a pair of intermediate solutions (for comparison with the numerical solutions)

$$u_h^* = R_\tau^n R_h u \quad \text{and} \quad r_h^* = R_\tau^n r,$$

and the following consistency error functions:

$$d_u^n := \mathrm{i}\partial_t R_\tau^n (R_h u - u) + \Delta_h R_h (u - R_\tau^n u) + r g(u) u - I_\tau^n [r_h^* g(u_h^*) u_h^*], \tag{3.25}$$

$$d_r^n := \frac{1}{2} \mathrm{Re}\big[\big(g(u)u, \partial_t u\big) - I_\tau^n \big(g(u_h^*) u_h^*, \partial_t u_h^*\big)\big]. \tag{3.26}$$

It is easy to check that there hold

$$\begin{aligned} &\int_{I_n} \mathrm{i}\big(\partial_t u_h^*, v_h\big) \mathrm{d}t + \int_{I_n} \big(\nabla P_\tau^n u_h^*, \nabla v_h\big) \mathrm{d}t \\ &\quad - \frac{\tau}{2} \sum_{j=1}^k w_j \big(r_h^*(t_{nj}) g(u_h^*(t_{nj})) u_h^*(t_{nj}), v_h(t_{nj})\big) \\ &= \int_{I_n} (P_\tau^n d_u^n, v_h) \mathrm{d}t, \end{aligned} \tag{3.27}$$

$$\begin{aligned} &\int_{I_n} \partial_t r_h^* q_h \mathrm{d}t \\ &= \frac{\tau}{4} \sum_{j=1}^k w_j \mathrm{Re}\big(q_h(t_{nj}) g(u_h^*(t_{nj})) u_h^*(t_{nj}), \partial_t u_h^*(t_{nj})\big) + \int_{I_n} P_\tau^n d_r^n q_h \mathrm{d}t. \end{aligned} \tag{3.28}$$

THEOREM 3.5. *Suppose that the solution of* (1.1) *is sufficiently smooth, then* $d_u^n \in C(I_n; H_0^1(\varOmega))$ *and there hold*

$$\sup_{t\in I_n} \|d_u^n\|_{H^1} \le C(h^p + \tau^{k+1}) \quad \textit{and} \quad \sup_{t\in I_n} |P_\tau^n d_r^n| \le C(h^p + \tau^{k+1}). \tag{3.29}$$

Proof. Since the spatial Ritz projection R_h maps $H_0^1(\varOmega)$ into $S_h \subset H_0^1(\varOmega)$, and the temporal Ritz projection R_τ^n maps $W^{1,\infty}(I_n; H_0^1(\varOmega))$ into $\mathbb{P}^k \otimes H_0^1(\varOmega)$, it follows that every term in (3.25) is in $C(I_n; H_0^1(\varOmega))$. This implies $d_u^n \in C(I_n; H_0^1(\varOmega))$.

By using the triangle inequality, from (3.25) we get

$$\begin{aligned} \max_{t\in I_n} \|d_u^n\|_{H^1} &\le \max_{t\in I_n} \big(\|\partial_t R_\tau^n (R_h u - u)\|_{H^1} + \|\Delta_h R_h (u - R_\tau^n u)\|_{H^1}\big) \\ &\quad + \max_{t\in I_n} \big(\|r g(u) u - I_\tau^n [r g(u) u]\|_{H^1} + \|r g(u) u - r_h^* g(u_h^*) u_h^*\|_{H^1}\big) \\ &=: D_1^u + D_2^u + D_3^u + D_4^u. \end{aligned} \tag{3.30}$$

Choosing $m = 0$ in Lemma 3.3, we obtain the following stability result:

$$\|R_\tau^n u\|_{W^{1,\infty}(I_n; H^s)} \le C \|u\|_{W^{1,\infty}(I_n; H^s)}. \tag{3.31}$$

Using (3.31) and (3.24), we can estimate D_1^u as follows:

$$\begin{aligned} D_1^u = \max_{t\in I_n} \|\partial_t R_\tau^n (R_h u - u)\|_{H^1} &\le \|R_h u - u\|_{W^{1,\infty}(I_n; H^1)} \\ &\le C h^p \|R_h u - u\|_{W^{1,\infty}(I_n; H^{p+1})}. \end{aligned}$$

Similarly, using identity (3.23) and Lemma 3.3, we have

$$D_2^u = \max_{t\in I_n} \|\Delta_h R_h(u - R_\tau^n u)\|_{H^1} \le \max_{t\in I_n} \|u - R_\tau^n u\|_{H^3} \le C\tau^{k+1}\|u\|_{W^{k+1,\infty}(I_n;H^3)},$$

and

$$D_3^u = \max_{t\in I_n} \|rg(u)u - I_\tau^n[rg(u)u]\|_{H^1} \le C\tau^{k+1}.$$

By using the triangle inequality, we decompose D_4^u into two parts,

$$D_4^u \le \max_{t\in I_n} \big(\|rg(u)u - rg(R_hu)R_hu\|_{H^1} + \|rg(R_hu)R_hu - R_\tau^n rg(R_\tau^n R_hu)R_\tau^n R_hu\|_{H^1}\big) \le Ch^p + C\tau^{k+1}.$$

Then, substituting the estimates of D_j^u, $j = 1, 2, 3, 4$, into (3.30), we obtain the desired estimate for $\|d_u^n\|_{H^1}$.

To estimate $|P_\tau^n d_r^n|$, we rewrite (3.26) as

$$d_r^n = \frac{1}{2}\mathrm{Re}\Big[\big(g(u)u, \partial_t(u - u_h^*)\big) + \big(g(u)u - g(u_h^*)u_h^*, \partial_t u_h^*\big)\Big] + \frac{1}{2}\mathrm{Re}\Big[\big(g(u_h^*)u_h^*, \partial_t u_h^*\big) - I_\tau^n\big(g(u_h^*)u_h^*, \partial_t u_h^*\big)\Big]$$

and test this expression by $P_\tau^n v$ in the time interval I_n with $v \in \mathbb{P}^k$. This yields

$$\begin{aligned} \int_{I_n} P_\tau^n d_r^n v \,\mathrm{d}t &= \int_{I_n} d_r^n P_\tau^n v \,\mathrm{d}t \\ &\le \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, \partial_t(u - u_h^*)\big)P_\tau^n v\,\mathrm{d}t \\ &\quad + C\tau^{\frac12}\|(g(u)u - g(u_h^*)u_h^*, \partial_t u_h^*)\|_{L^\infty(I_n)}\|v\|_{L^2(I_n)} \\ &\quad + C\tau^{k+\frac32}\|\partial_t^{k+1}\big(g(u_h^*)u_h^*, \partial_t u_h^*\big)\|_{L^\infty(I_n)}\|v\|_{L^2(I_n)} \\ &\le \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, \partial_t(u - u_h^*)\big)P_\tau^n v\,\mathrm{d}t + C\tau^{\frac12}(h^p + \tau^{k+1})\|v\|_{L^2(I_n)}. \end{aligned} \tag{3.32}$$

The first term on the right-hand side of (3.32) can be estimated as follows:

$$\begin{aligned} \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, \partial_t(u - u_h^*)\big)P_\tau^n v\,\mathrm{d}t &= \int_{I_n} (g(u)u, \partial_t(u - R_\tau^n u))P_\tau^n v\,\mathrm{d}t \\ &\quad + \int_{I_n} (g(u)u, \partial_t R_\tau^n(u - R_hu))P_\tau^n v\,\mathrm{d}t \\ &=: D_1^r + D_2^r, \end{aligned} \tag{3.33}$$

$$\begin{aligned} D_1^r &= \int_{I_n} \big(g(u)uP_\tau^n v, \partial_t(u - R_\tau^n u)\big)\,\mathrm{d}t \\ &= \int_{I_n} \big(g(u)uP_\tau^n v - P_\tau^n(g(u)uP_\tau^n v), \partial_t(u - R_\tau^n u)\big)\,\mathrm{d}t \\ &\le C\tau^{\frac12}\|g(u)uP_\tau^n v - P_\tau^n(g(u)uP_\tau^n v)\|_{L^2(I_n;L^2)}\|\partial_t(u - R_\tau^n u)\|_{L^\infty(I_n;L^2)} \end{aligned}$$

$$\leq C\tau^{\frac{3}{2}}\|P_\tau^n v\|_{L^2(I_n;L^2)}\|\partial_t(u-R_\tau^n u)\|_{L^\infty(I_n;L^2)} \quad \text{(we have used Lemma 3.4)}$$
$$\leq C\tau^{k+\frac{3}{2}}\|v\|_{L^2(I_n)}\|\partial_t^{k+1}u\|_{L^\infty(I_n;L^2)} \quad \text{(we have used Lemma 3.3)},$$

$$D_2^r \leq C\tau^{\frac{1}{2}}\|g(u)u\|_{L^\infty(I_n;L^2)}\|u-R_h u\|_{W^{1,\infty}(I_n;L^2)}\|v\|_{L^2(I_n)}$$
$$\leq C\tau^{\frac{1}{2}}h^p\|v\|_{L^2(I_n)}\|u\|_{W^{1,\infty}(I_n;H^{p+1})}.$$

Substituting these estimates into (3.32), we obtain

$$\left|\int_{I_n} P_\tau^n d_r^n v\,\mathrm{d}t\right| \leq C\tau^{\frac{1}{2}}(h^p+\tau^{k+1})\|v\|_{L^2(I_n)}.$$

Since this inequality holds for arbitrary $v \in L^2(I_n)$, it follows that

$$\|P_\tau^n d_r^n\|_{L^2(I_n)} \leq C\tau^{\frac{1}{2}}(h^p+\tau^{k+1}).$$

Then, using the inverse inequality in time, we obtain the desired estimate for $|P_\tau^n d_r^n|$. ∎

4. Well-posedness and convergence analysis. We define the error functions $e_h^u = u_h - u_h^*$ and $e_h^r = r_h - r_h^*$, with the following abbreviations:

$$e_{nj}^u = e_h^u(t_{nj}), \quad e_{nj}^r = e_h^r(t_{nj}), \quad u_{nj} = u_h(t_{nj}), \quad r_{nj} = r_h(t_{nj}),$$
$$u_{nj}^* = u_h^*(t_{nj}), \quad r_{nj}^* = r_h^*(t_{nj}), \quad v_{nj} = v_h(t_{nj}), \quad q_{nj} = q_h(t_{nj}).$$

Subtracting (3.27)–(3.28) from (3.7)–(3.8), we obtain the following error equations:

$$\mathrm{i}\int_{I_n}\left(\partial_t e_h^u, v_h\right)\mathrm{d}t = -\int_{I_n}\left(\nabla P_\tau^n e_h^u, \nabla v_h\right)\mathrm{d}t + \frac{\tau}{2}\sum_{j=1}^k w_j\Big(e_{nj}^r g(u_{nj})u_{nj}, v_{nj}\Big)$$
$$+\frac{\tau}{2}\sum_{j=1}^k w_j\Big(r_{nj}^*\big[g(u_{nj})u_{nj}-g(u_{nj}^*)u_{nj}^*\big], v_{nj}\Big) - \int_{I_n}(P_\tau^n d_u^n, v_h)\mathrm{d}t, \tag{4.1a}$$

$$\int_{I_n}\partial_t e_h^r q_h \mathrm{d}t = \frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj}\big(g(u_{nj})u_{nj}-g(u_{nj}^*)u_{nj}^*\big), \partial_t u_h^*(t_{nj})\big)$$
$$+\frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj}g(u_{nj})u_{nj}, \partial_t e_h^u(t_{nj})\big) - \int_{I_n} P_\tau^n d_r^n q_h \mathrm{d}t, \tag{4.1b}$$

which hold for all test functions $v_h \in \mathbb{P}^k \otimes S_h$ and $q_h \in \mathbb{P}^k$.

Remark 4.1. If (4.1) has a solution $(e_h^u, e_h^r) \in X_{\tau,h} \times Y_{\tau,h}$, then $u_h = u_h^* + e_h^u$ and $r_h = r_h^* + e_h^r$ give a solution of the numerical scheme (2.8). In the following, we prove existence of (e_h^u, e_h^r) to (4.1) by using Schaefer's fixed point theorem, which is quoted below.

THEOREM 4.1 (Schaefer's fixed point theorem [11, Chapter 9.2, Theorem 4]). *Let B be a Banach space and $M : B \to B$ be a continuous and compact mapping. If the set*

$$\big\{\phi \in B :\ \exists\,\theta \in [0,1] \ \text{ such that } \ \phi = \theta M(\phi)\big\} \tag{4.2}$$

is bounded in B, then the mapping M has at least one fixed point.

We define

$$X^*_{\tau,h} = \Big\{ v_h \in X_{\tau,h} : \max_{1\le n\le N}\max_{1\le j\le k} \|v_h(t_{nj}) - u_h^*(t_{nj})\|_{L^\infty\cap H^1} \le \frac{1}{2}\Big\}, \tag{4.3}$$

$$Y^*_{\tau,h} = \Big\{ q_h \in Y_{\tau,h} : \max_{1\le n\le N}\max_{1\le j\le k} |q_h(t_{nj}) - r_h^*(t_{nj})| \le \frac{1}{2}\Big\}, \tag{4.4}$$

where the norm $\|\cdot\|_{L^\infty\cap H^1}$ is defined as

$$\|\phi_h\|_{L^\infty\cap H^1} := \max\big(\|\phi_h\|_{L^\infty}, \|\phi_h\|_{H^1}\big).$$

For any element $(\phi_h, \varphi_h) \in X_{\tau,h}\times Y_{\tau,h}$, we define two associated numbers

$$\rho[\phi_h] := \min\left(\frac{1}{\max\limits_{1\le n\le N}\max\limits_{1\le j\le k}\|\phi_h(t_{nj})\|_{L^\infty\cap H^1}}, 1\right), \tag{4.5a}$$

$$\rho[\varphi_h] := \min\left(\frac{1}{\max\limits_{1\le n\le N}\max\limits_{1\le j\le k}|\varphi_h(t_{nj})|}, 1\right), \tag{4.5b}$$

which are continuous with respect to (ϕ_h, φ_h) (because all norms are equivalent in the finite-dimensional space $X_{\tau,h}\times Y_{\tau,h}$). Furthermore, the two numbers defined above satisfy the following estimates:

$$\max_{1\le n\le N}\max_{1\le j\le k}\|\rho[\phi_h]\phi_h(t_{nj})\|_{L^\infty\cap H^1} \le 1, \tag{4.6}$$

$$\max_{1\le n\le N}\max_{1\le j\le k}|\rho[\varphi_h]\varphi_h(t_{nj})| \le 1. \tag{4.7}$$

Then we define

$$u^\phi := u_h^* + \rho[\phi_h]\phi_h \quad \text{and} \quad r^\varphi := r_h^* + \rho[\varphi_h]\varphi_h \tag{4.8}$$

with the following abbreviations:

$$u^\phi_{nj} = u^\phi_h(t_{nj}) \quad \text{and} \quad \varphi_{nj} = \varphi_h(t_{nj}),$$

and define $(e_h^u, e_h^r) \in X_{\tau,h}\times Y_{\tau,h}$ to be the solution of the following linear equations:

$$\begin{aligned}\mathrm{i}\int_{I_n}\big(\partial_t e_h^u, v_h\big)\,\mathrm{d}t &+ \int_{I_n}\big(\nabla P_\tau^n e_h^u, \nabla v_h\big)\,\mathrm{d}t \\ &= \frac{\tau}{2}\sum_{j=1}^k w_j\Big(\varphi_{nj} g(u^\phi_{nj})u^\phi_{nj}, v_{nj}\Big) \\ &\quad + \frac{\tau}{2}\sum_{j=1}^k w_j\Big(r^*_{nj}\big[g(u^\phi_{nj})u^\phi_{nj} - g(u^*_{nj})u^*_{nj}\big], v_{nj}\Big) - \int_{I_n}(P_\tau^n d^u, v_h)\mathrm{d}t\end{aligned} \tag{4.9}$$

and

$$\begin{aligned}\int_{I_n}\partial_t e_h^r q_h\,\mathrm{d}t &= \frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj}\big(g(u^\phi_{nj})u^\phi_{nj} - g(u^*_{nj})u^*_{nj}\big), \partial_t u_h^*(t_{nj})\big) \\ &\quad + \frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj} g(u^\phi_{nj})u^\phi_{nj}, \partial_t\phi_h(t_{nj})\big) - \int_{I_n} P_\tau^n d^r q_h \mathrm{d}t\end{aligned} \tag{4.10}$$

for all $v_h \in \mathbb{P}^k \otimes S_h$ and $q_h \in \mathbb{P}^k$, $n = 1, \ldots, N$. We denote by $M : X_{\tau,h} \times Y_{\tau,h} \to X_{\tau,h} \times Y_{\tau,h}$ the mapping from (ϕ_h, φ_h) to (e_h^u, e_h^r), and define the set

$$\mathfrak{B} = \big\{(\phi_h, \varphi_h) \in X_{\tau,h} \times Y_{\tau,h} : \ \exists\, \theta \in [0,1] \ \text{such that} \ (\phi_h, \varphi_h) = \theta M(\phi_h, \varphi_h)\big\}, \tag{4.11}$$

and the following norm on $X_{\tau,h} \times Y_{\tau,h}$: For any $(\phi_h, \varphi_h) \in X_{\tau,h} \times Y_{\tau,h}$

$$\|(\phi_h, \varphi_h)\|_{X_{\tau,h} \times Y_{\tau,h}} := \|\phi_h\|_{L^\infty(0,T;H^1)} + \|\varphi_h\|_{L^\infty(0,T)}. \tag{4.12}$$

It is straightforward to show the following result (see [12, proof of Lemma 4.2]).

LEMMA 4.2. *The mapping $M : X_{\tau,h} \times Y_{\tau,h} \to X_{\tau,h} \times Y_{\tau,h}$ is well-defined, continuous, and compact.*

Moreover, there holds the following key technical lemma.

LEMMA 4.3. *Let $1 \le d \le 3$ and assume that the solution of the NLS equation (1.1) is sufficiently smooth. Then there exist positive constants τ_0 and h_0 such that when $\tau \le \tau_0$ and $h \le h_0$, the following statement holds: If $(\phi_h, \varphi_h) \in \mathfrak{B}$ and $(e_h^u, e_h^r) = M(\phi_h, \varphi_h)$, then*

$$\begin{aligned} \|e_h^u\|_{L^\infty(0,T;H^1)} + \|e_h^r\|_{L^\infty(0,T)} \le\ & C\big(\|e_h^u(0)\|_{H^1} + |e_h^r(0)|\big) \\ & + C \max_{1\le n\le N} \max_{t\in I_n}\big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big), \end{aligned} \tag{4.13}$$

$$\max_{1\le n\le N} \max_{1\le j\le k} \|e_h^u(t_{nj})\|_{L^\infty \cap H^1} \le \frac{1}{2} \quad \textit{and} \quad \max_{1\le n\le N} \max_{1\le j\le k} |e_h^r(t_{nj})| \le \frac{1}{2}, \tag{4.14}$$

$$\rho[\phi_h] = 1, \quad \rho[\varphi_h] = 1. \tag{4.15}$$

Proof. Since the proof is very long and technical, below we only outline the main steps and ingredients of the proof and refer the interested reader to [12] for the details.

If $(\phi_h, \varphi_h) \in \mathfrak{B}$ and $(e_h^u, e_h^r) = M(\phi_h, \varphi_h)$, then

$$(\phi_h, \varphi_h) = \theta M(\phi_h, \varphi_h) = (\theta e_h^u, \theta e_h^r),$$

which implies $\phi_h = \theta e_h^u$ and $\varphi_h = \theta e_h^r$. In this case, (4.9)–(4.10) can be rewritten as

$$\begin{aligned} \mathrm{i}\int_{I_n} \big(\partial_t e_h^u, v_h\big)\,\mathrm{d}t = & -\int_{I_n} \big(\nabla P_\tau^n e_h^u, \nabla v_h\big)\,\mathrm{d}t + \frac{\theta\tau}{2}\sum_{j=1}^k w_j\Big(e_{nj}^r g(u_{nj}^\phi)u_{nj}^\phi, v_{nj}\Big) \\ & + \frac{\tau}{2}\sum_{j=1}^k w_j\Big(r_{nj}^*\big[g(u_{nj}^\phi)u_{nj}^\phi - g(u_{nj}^*)u_{nj}^*\big], v_{nj}\Big) \\ & - \int_{I_n} (P_\tau^n d_u^n, v_h)\mathrm{d}t, \end{aligned} \tag{4.16}$$

$$\begin{aligned} \int_{I_n} \partial_t e_h^r q_h \mathrm{d}t = & \frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj}\big(g(u_{nj}^\phi)u_{nj}^\phi - g(u_{nj}^*)u_{nj}^*\big), \partial_t u_h^*(t_{nj})\big) \\ & + \frac{\theta\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj} g(u_{nj}^\phi)u_{nj}^\phi, \partial_t e_h^u(t_{nj})\big) - \int_{I_n} P_\tau^n d_r^n\, q_h \mathrm{d}t, \end{aligned} \tag{4.17}$$

which hold for all $v_h \in \mathbb{P}^k \otimes S_h$ and $q_h \in \mathbb{P}^k$, $n = 1, \ldots, N$. It remains to derive estimates for e_h^u and e_h^r based on the above equations.

From (4.6)–(4.7) and definition (4.8) we get

$$\begin{aligned}(4.18)\quad &\max_{1\le n\le N}\max_{1\le j\le k}\|u^{\phi}(t_{nj})\|_{L^\infty\cap H^1}+\max_{1\le n\le N}\max_{1\le j\le k}|r^{\varphi}(t_{nj})|\\ &\le \max_{1\le n\le N}\max_{1\le j\le k}\|u_h^*(t_{nj})\|_{L^\infty\cap H^1}+\max_{1\le n\le N}\max_{1\le j\le k}|r_h^*(t_{nj})|\\ &\quad+\max_{1\le n\le N}\max_{1\le j\le k}\|\rho[\phi_h]\phi_h(t_{nj})\|_{L^\infty\cap H^1}+\max_{1\le n\le N}\max_{1\le j\le k}|\rho[\varphi_h]\varphi_h(t_{nj})|\\ &\le \|u_h^*\|_{L^\infty(0,T;L^\infty\cap H^1)}+\|r_h^*\|_{L^\infty(0,T)}+2.\end{aligned}$$

Thus $\|u^{\phi}(t_{nj})\|_{L^\infty\cap H^1}$ and $|r^{\varphi}(t_{nj})|$ are bounded uniformly with respect to τ and h.

The major part of the remaining proof is devoted to proving the following three inequalities:

(4.19)
$$\int_{I_n}\|e_h^u\|_{H^1}^2\mathrm{d}t\le C\tau\|e_h^u(t_{n-1})\|_{H^1}^2+C\tau^2\int_{I_n}|e_h^r|^2\mathrm{d}t+C\tau^3\max_{t\in I_n}\|d_u^n\|_{H^1}^2,$$

$$(4.20)\quad \int_{I_n}|e_h^r|^2\mathrm{d}t\le C\tau\big[\|e_h^u(t_{n-1})\|_{H^1}^2+|e_h^r(t_{n-1})|^2+\tau^2\max_{t\in I_n}\big(\|d_u^n\|_{H^1}^2+|P_\tau^n d_r^n|^2\big)\big],$$

(4.21)
$$\begin{aligned}&\|\nabla e_h^u(t_n)\|^2+|e_h^r(t_n)|^2-\|\nabla e_h^u(t_{n-1})\|^2-|e_h^r(t_{n-1})|^2+\int_{I_n}\|\partial_t e_h^u\|_{H^{-1}}^2\mathrm{d}t\\ &\le C\int_{I_n}\big(\|e_h^u\|_{H^1}^2+|e_h^r|^2\big)\mathrm{d}t+C\int_{I_n}(\|d_u^n\|_{H^1}^2+|P_\tau^n d_r^n|^2)\mathrm{d}t.\end{aligned}$$

In particular, (4.19) can be obtained by substituting $v_h=(-\Delta_h)P_\tau^n[P_\tau^n e_h^u(t)(t_n-t)]$ into (4.16) and considering the imaginary part; (4.20) can be obtained by substituting $q_h=P_\tau^n\big[P_\tau^n e_h^r(t)(t_n-t)\big]$ into (4.17); (4.21) is obtained by setting $v_h=\partial_t e_h^u$ into (4.16) and considering the real part, setting $q_h=2e_h^r$ into (4.17), and estimating $\int_{I_n}\|\partial_t e_h^u\|_{H^{-1}}^2\mathrm{d}t$ via a duality argument using (4.16). More details can be found in [12, proof of Lemma 4.3].

To complete the proof, substituting (4.19)–(4.20) into (4.21), we obtain

$$\begin{aligned}(4.22)\quad &\big(\|\nabla e_h^u(t_n)\|^2+|e_h^r(t_n)|^2\big)-\big(\|\nabla e_h^u(t_{n-1})\|^2+|e_h^r(t_{n-1})|^2\big)+\int_{I_n}\|\partial_t e_h^u\|_{H^{-1}}^2\mathrm{d}t\\ &\le C\tau\big(\|\nabla e_h^u(t_{n-1})\|^2+|e_h^r(t_{n-1})|^2\big)+C\int_{I_n}(\|d_u^n\|_{H^1}^2+|P_\tau^n d_r^n|^2)\mathrm{d}t.\end{aligned}$$

It follows from Gronwall's inequality that

$$\begin{aligned}(4.23)\quad &\max_{1\le n\le N}\big(\|\nabla e_h^u(t_n)\|^2+|e_h^r(t_n)|^2\big)+C\int_0^T\|\partial_t e_h^u\|_{H^{-1}}^2\,\mathrm{d}t\\ &\le C(\|\nabla e_h^u(0)\|^2+|e_h^r(0)|^2)+C\sum_{n=1}^N\int_{I_n}(\|d_u^n\|_{H^1}^2+|P_\tau^n d_r^n|^2)\mathrm{d}t.\end{aligned}$$

Then, substituting the above inequality into (4.19)–(4.20) and using the temporal inverse inequality, we obtain

$$\begin{aligned}(4.24)\quad \max_{t\in[0,T]}\big(\|e_h^u(t)\|_{H^1}^2+|e_h^r(t)|^2\big)&\le C(\|e_h^u(0)\|_{H^1}^2+|e_h^r(0)|^2)\\ &\quad+C\max_{1\le n\le N}\max_{t\in I_n}\big(\|d_u^n\|_{H^1}^2+|P_\tau^n d_r^n|^2\big).\end{aligned}$$

Hence, (4.13) holds.

When τ and h are sufficiently small, inequality (4.24) implies that

$$\max_{t\in[0,T]} \|e_h^u(t)\|_{H^1} \le \frac{1}{2} \qquad \text{and} \qquad \max_{t\in[0,T]} |e_h^r(t)| \le \frac{1}{2}. \tag{4.25}$$

On the one hand, by the inverse inequality, we have

$$\begin{aligned}\max_{t\in[0,T]} \|e_h^u(t)\|_{L^\infty} &\le C\ell_h \max_{t\in[0,T]} \|e_h^u(t)\|_{H^1} \\ &\le C\ell_h \Big[\|e_h^u(0)\|_{H^1} + |e_h^r(0)| + \max_{1\le n\le N} \max_{t\in I_n} \big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big],\end{aligned} \tag{4.26}$$

where

$$\ell_h = \begin{cases} 1 & \text{if } d = 1, \\ \ln(2+1/h) & \text{if } d = 2, \\ h^{-\frac{1}{2}} & \text{if } d = 3. \end{cases}$$

On the other hand, by choosing a test function v in (4.16) satisfying the properties $v(t_{nj}) = 1$ and $v(t_{ni}) = 0$ for $i \ne j$, and using property (3.6), we obtain

$$\begin{aligned}\|\Delta_h e_{nj}^u\| &= \Big\| \mathrm{i}\partial_t e_{nj}^u - \theta P_h\big[e_{nj}^r g(u_{nj}^\phi) u_{nj}^\phi\big] + P_h d_{nj}^u \\ &\qquad - P_h\big[r_{nj}^*(g(u_{nj}^\phi)u_{nj}^\phi - g(u_{nj}^*)u_{nj}^*)\big]\Big\| \\ &\le C\tau^{-1}\Big[\|e_h^u(0)\|_{H^1} + |e_h^r(0)| + \max_{1\le n\le N} \max_{t\in I_n} \big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big],\end{aligned} \tag{4.27}$$

where we have used (4.23)–(4.24) and an inverse inequality in time in estimating $\partial_t e_{nj}^u$. By the discrete Sobolev embedding inequality, for $1 \le d \le 3$ we have

$$\begin{aligned}\|e_{nj}^u\|_{L^\infty} &\le C\|e_{nj}^u\|_{H^1}^{\frac{1}{2}} \|\Delta_h e_{nj}^u\|^{\frac{1}{2}} \\ &\le C\tau^{-\frac{1}{2}}\Big[\|e_h^u(0)\|_{H^1} + |e_h^r(0)| + \max_{1\le n\le N} \max_{t\in I_n} \big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big],\end{aligned} \tag{4.28}$$

where we have used (4.24) and (4.27) in the last inequality. Then, combining (4.26) and (4.28) yields

$$\begin{aligned}\max_{1\le n\le N} \max_{1\le j\le k} \|e^u(t_{nj})\|_{L^\infty} &\le C\min(\ell_h, \tau^{-\frac{1}{2}})\Big[\|e_h^u(0)\|_{H^1} + |e_h^r(0)| \\ &\qquad + \max_{1\le n\le N} \max_{t\in I_n} \big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big] \\ &\le C\big(h^{p-\frac{1}{2}} + \tau^{k+\frac{1}{2}}\big),\end{aligned}$$

where we have used the consistency estimate from Theorem 3.5. When τ and h are sufficiently small, the inequality above implies

$$\max_{1\le n\le N} \max_{1\le j\le k} \|e^u(t_{nj})\|_{L^\infty} \le \frac{1}{2}\,. \tag{4.29}$$

This together with (4.25) gives (4.14).

Furthermore, since $\phi_h = \theta e_h^u$ and $\varphi_h = \theta e_h^r$, it follows that

$$\max_{1\le n\le N} \max_{1\le j\le k} \|\phi_h(t_{nj})\|_{L^\infty\cap H^1} \le \frac{1}{2} \quad \text{and} \quad \max_{1\le n\le N} \max_{1\le j\le k} |\varphi_h(t_{nj})| \le \frac{1}{2},$$

which imply $\rho[\phi_h] = \rho[\varphi_h] = 1$ in view of definitions in (4.5). This proves (4.15). $\square$

We now are ready to state and prove the existence, uniqueness, and convergence of numerical solutions, which comprise the second main theorem of this paper.

THEOREM 4.4. *Let* $1 \le d \le 3$ *and assume that the solution of the NLS equation* (1.1) *is sufficiently smooth. Then there exist positive constants* τ_0 *and* h_0 *such that when* $\tau \le \tau_0$ *and* $h \le h_0$, *the numerical method* (2.8) *has a unique solution* $(u_h, r_h) \in X^*_{\tau,h} \times Y^*_{\tau,h}$. *Moreover, this solution satisfies the following error estimate:*

$$\max_{t\in[0,T]}\Big(\|u_h(t) - u_h^*(t)\|_{H^1} + |r_h(t) - r_h^*(t)|\Big) \le C(h^p + \tau^{k+1}). \tag{4.30}$$

Proof. Step 1: *Existence.* By the definition of $\mathfrak{B}$, if $(\phi_h, \varphi_h) \in \mathfrak{B}$ and $(e_h^u, e_h^r) = M(\phi_h, \varphi_h)$, then $\phi_h = \theta e_h^u$ and $\varphi_h = \theta e_h^r$. Thus (4.13) implies

$$\|(\phi_h, \varphi_h)\|_{X_{\tau,h}\times Y_{\tau,h}} = \|\phi_h\|_{L^\infty(0,T;H^1)} + \|\varphi_h\|_{L^\infty(0,T)} \le C, \tag{4.31}$$

which together with Schaefer's fixed point theorem imply the existence of a fixed point (ϕ_h, φ_h) for the mapping M (corresponding to $\theta = 1$) with

$$(e_h^u, e_h^r) = (\phi_h, \varphi_h), \quad u^\phi = u_h^* + \phi_h, \quad \text{and} \quad r^\phi = r_h^* + \varphi_h,$$

satisfying (4.9)–(4.10), where we have used (4.15) in the expression (4.8). Consequently, (e_h^u, e_h^r) is a solution of (4.1) with $(u_h, r_h) = (u_h^\phi, r_h^\varphi) = (u_h^* + e_h^u, r_h^* + e_h^r)$. Hence, in view of the discussions in Remark 4.1, (u_h, r_h) is a solution of the numerical scheme (2.8), and (4.14) implies (u_h, r_h) is in the set $X^*_{\tau,h} \times Y^*_{\tau,h}$ defined in (4.3)–(4.4). This proves the existence of a numerical solution in $X^*_{\tau,h} \times Y^*_{\tau,h}$.

Step 2: *Uniqueness.* Suppose that (u_h, r_h) and $(\widetilde{u}_h, \widetilde{r}_h)$ in $X^*_{\tau,h} \times Y^*_{\tau,h}$ are two pairs of numerical solutions, and set $e_h^u = u_h - \widetilde{u}_h$ and $e_h^r = r_h - \widetilde{r}_h$ (abusing the notation). Subtracting the corresponding equations satisfied by (u_h, r_h) and $(\widetilde{u}_h, \widetilde{r}_h)$ shows that (e_h^u, e_h^r) satisfies (4.1) with $e_h^u(0) = e_h^r(0) = 0$ and $d_u^n = d_r^n = 0$. In the meantime, the definition in (4.3)–(4.4) implies

$$\|e_h^u(t_{nj})\|_{L^\infty\cap H^1} \le 1 \quad \text{and} \quad |e_h^r(t_{nj})| \le 1. \tag{4.32}$$

Accordingly, (e_h^u, e_h^r) is a fixed point of the mapping M (corresponding to $\theta = 1$ in $\mathfrak{B}$) in the case $e_h^u(0) = e_h^r(0) = 0$ and $d_u^n = d_r^n = 0$. Hence, an application of (4.13) yields

$$\begin{aligned}\|e_h^u\|_{L^\infty(0,T;H^1)} + \|e_h^r\|_{L^\infty(0,T)} \le C\Big[&\|e_h^u(0)\|_{H^1} + |e_h^r(0)| \\ &+ \max_{1\le n\le N}\max_{t\in I_n}\big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big] = 0.\end{aligned}$$

Thus, $(u_h, r_h) = (\widetilde{u}_h, \widetilde{r}_h)$ and the uniqueness of the numerical solution is proved.

Step 3: *Error estimate.* Since the error functions $e_h^u = u_h - u_h^*$ and $e_h^r = r_h - r_h^*$ satisfy (4.1) and (4.32), it follows that (e_h^u, e_h^r) is a fixed point of the mapping M (corresponding to $\theta = 1$ in $\mathfrak{B}$). Hence, an application of (4.13) yields

$$\begin{aligned}&\|e_h^u\|_{L^\infty(0,T;H^1)} + \|e_h^r\|_{L^\infty(0,T)} \\ &\qquad \le C\Big[\|e_h^u(0)\|_{H^1} + |e_h^r(0)| + \max_{1\le n\le N}\max_{t\in I_n}\big(\|d_u^n\|_{H^1} + |P_\tau^n d_r^n|\big)\Big].\end{aligned}$$

Substituting the consistency error estimates from Theorem 3.5 into the above inequality yields the desired estimate (4.30). The proof is complete. ∎

Remark 4.2. For the periodic and Neumann boundary conditions, the mass and energy conservations in Theorem 3.1 and the error estimate in Theorem 4.4 can be proved similarly.

Table 1
Time discretization errors of the proposed method, with $h = \frac{2L}{5000}$ *and* $T = 1$.

k	τ	$p = 3$	
		$\|u(x,t) - u_h(x,t)\|_{L^\infty(0,T;H^1)}$	Order
	1/60	3.7964E–05	–
	1/70	2.3429E–05	3.1312
2	1/80	1.5460E–05	3.1132
	1/90	1.0733E–05	3.0985
	1/100	7.7542E–06	3.0853
	1/20	3.4019E–05	–
	1/25	1.3821E–05	4.0364
3	1/30	6.6322E–06	4.0275
	1/35	3.5689E–06	4.0200
	1/40	2.0886E–06	4.0123
	1/8	1.2291E–04	–
	1/12	1.5120E–05	5.1681
4	1/14	6.8492E–06	5.1369
	1/16	3.4634E–06	5.1067
	1/20	1.1555E–06	4.9192

5. Numerical experiments. In this section, we present some one-dimensional numerical tests to validate the theoretical results proved in Theorems 3.1 and 4.4 about the mass and energy conservations, and the convergence rates of the proposed method. All the computations are performed using the software package FEniCS (https://fenicsproject.org).

We consider the cubic NLS equation

$$\begin{aligned} \mathrm{i}\partial_t u - \partial_{xx} u - 2|u|^2 u &= 0 \qquad \text{in } (-L, L) \times (0, T], \\ u|_{t=0} &= u_0 \qquad \text{in } (-L, L) \quad \text{with } L = 20, \end{aligned} \tag{5.1}$$

subject to the periodic boundary condition. We choose $u_0 = \mathrm{sech}(x)\exp(2\mathrm{i}x)$ so that the exact solution is given by

$$u(x,t) = \mathrm{sech}(x + 4t)\exp(\mathrm{i}(2x + 3t)). \tag{5.2}$$

This example contains a soliton wave and is often used as a benchmark for meansuring the effectiveness of numerical methods for the NLS equation; see [33, 37, 25].

5.1. Convergence rates. We solve problem (5.1) by the proposed method (2.8) and compare the numerical solutions with the exact solution (5.2). Newton's method is used to solve the nonlinear system. The iteration is set to stop when the error is below 10^{-10}.

The time discretization errors are presented in Table 1, where we have used finite elements of degree 3 with a sufficiently spatial mesh $h = 2L/5000$ so that the error from spatial discretization is negligibly small in observing the temporal convergence rates. From Table 1 we see that the error of time discretization is $O(\tau^{k+1})$, which is consistent with the result proved in Theorem 4.4.

The spatial discretization errors are presented in Table 2, where we have chosen $k = 3$ with a sufficiently small time stepsize $\tau = 1/1000$ so that the time discretization error is negligibly small compared to the spatial error. From Table 2 we see that the spatial discretization errors are $O(h^p)$ in the H^1 norm. This is also consistent with the result proved in Theorem 4.4.

TABLE 2
Spatial discretization errors of the proposed method with $\tau = \frac{1}{1000}$ and $T = 1$.

p	M	$k=3$	
		$\|u(x,t) - u_h(x,t)\|_{L^\infty(0,T;H^1)}$	Order
	1400	5.8670E–02	–
	1600	5.1134E–02	1.0295
1	1800	4.5330E–02	1.0229
	2000	4.0719E–02	1.0183
	2200	3.6964E–02	1.0149
	240	1.9306E–02	–
	260	1.6438E–02	2.0094
2	280	1.4167E–02	2.0062
	300	1.2338E–02	2.0041
	320	1.0842E–02	2.0027
	90	1.6147E–02	–
	100	1.1661E–02	3.0894
3	110	8.7112E–03	3.0599
	120	6.6844E–03	3.0436
	130	5.2435E–03	3.0334

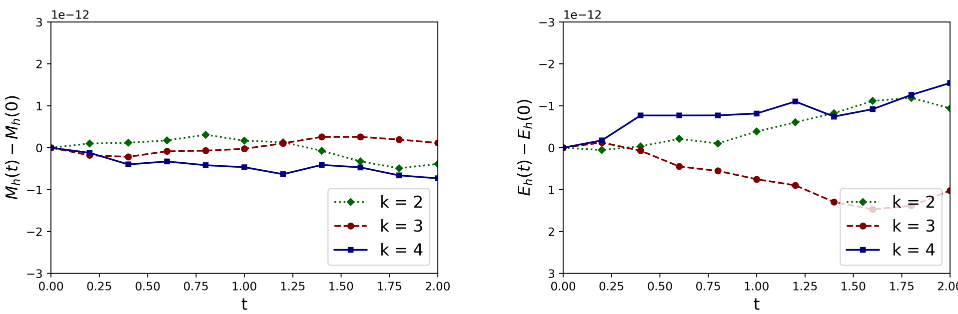

FIG. 1. *Evolution of mass $M_h(t) - M_h(0)$ and SAV energy $E_h(t) - E_h(0)$ with $p = 3$ and $\tau = h = 0.2$.*

5.2. Mass and energy conservations. We denote the mass and SAV energy of a numerical solution by

$$M_h(t) = \int_\Omega |u_h(t)|^2 \mathrm{d}x \quad \text{and} \quad E_h(t) = \frac{1}{2}\int_\Omega |\nabla u_h(t)|^2 \mathrm{d}x - r_h(t)^2, \tag{5.3}$$

respectively. The evolution of mass and SAV energy of the numerical solutions is presented in Figure 1 with $\tau = 0.2$ and $h = 0.2$. It is shown that

$$\text{mass} = 2 + O(10^{-12}) \quad \text{and} \quad \text{SAV energy} = -7.33358048516 + O(10^{-12}),$$

which are much smaller than the error of the numerical solutions, as shown in Figure 2. This shows the effectiveness of the proposed method in preserving mass and energy (independent of the error of numerical solutions). The number of iterations at each time level is presented in Figure 3 to show the effectiveness of Newton's method.

5.3. Comparison of different methods in preserving the shape of a soliton. The graph of $|u(x,t)|$ is a soliton propagating towards the left. Its shape remains

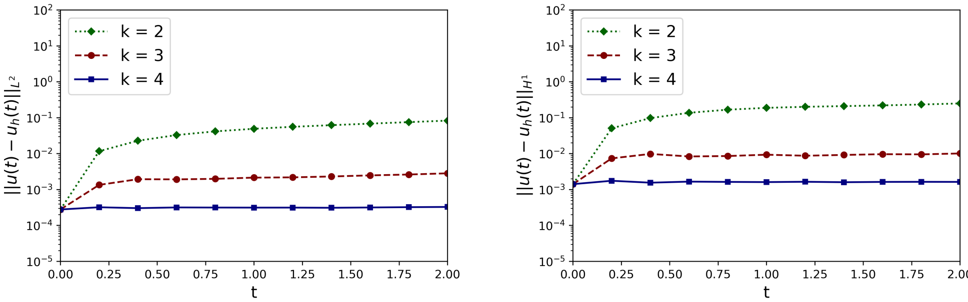

FIG. 2. *Evolution of the error of the numerical solution with* $p = 3$ *and* $\tau = h = 0.2$.

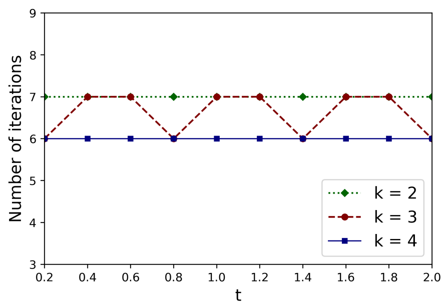

FIG. 3. *Number of iterations at each time level with* $p = 3$ *and* $\tau = h = 0.2$.

unchanged for all $t \geq 0$. The graphs of numerical solutions given by several different numerical methods using the same mesh sizes are presented in Figures 4 and 5. All the methods preserve mass and energy conservations. The numerical results show the effectiveness of the proposed method in preserving the shape of the soliton.

5.4. Capability of solving focusing nonlinearity. We consider the cubic NLS equation

$$\begin{aligned} \mathrm{i}\partial_t u - \partial_{xx} u - \partial_{yy} u + 2|u|^2 u &= 0 && \text{in } \Omega \times (0, T], \\ u|_{t=0} &= u_0 && \text{in } \Omega \end{aligned} \tag{5.4}$$

in two-dimensional space $\Omega = [0, 1] \times [0, 1]$ subject to the periodic boundary condition. We choose $u_0 = \exp(2\pi\mathrm{i}(x + y))$ so that the exact solution is given by

$$u(x, t) = \exp(\mathrm{i}(2\pi x + 2\pi y + (2 + 8\pi^2)t)), \tag{5.5}$$

which admits a progressive plane wave solution; see [37].

We solve problem (5.4) by the proposed method (2.8) and compare the numerical solutions with the exact solution (5.5). Newton's method is used to solve the nonlinear system. The iteration is stopped when the error is below 10^{-10}.

The time discretization errors are presented in Table 3, where we have used finite elements of degree 3 with a sufficiently spatial mesh $h = 1/80$ so that the error from spatial discretization is negligibly small in observing the temporal convergence rates. From Table 3 we see that the error of the time discretization is $O(\tau^{k+1})$, which is consistent with the result proved in Theorem 4.4.

The spatial discretization errors are presented in Table 4, where we have chosen $k = 3$ with a sufficiently small time stepsize $\tau = 1/1000$ so that the time discretization

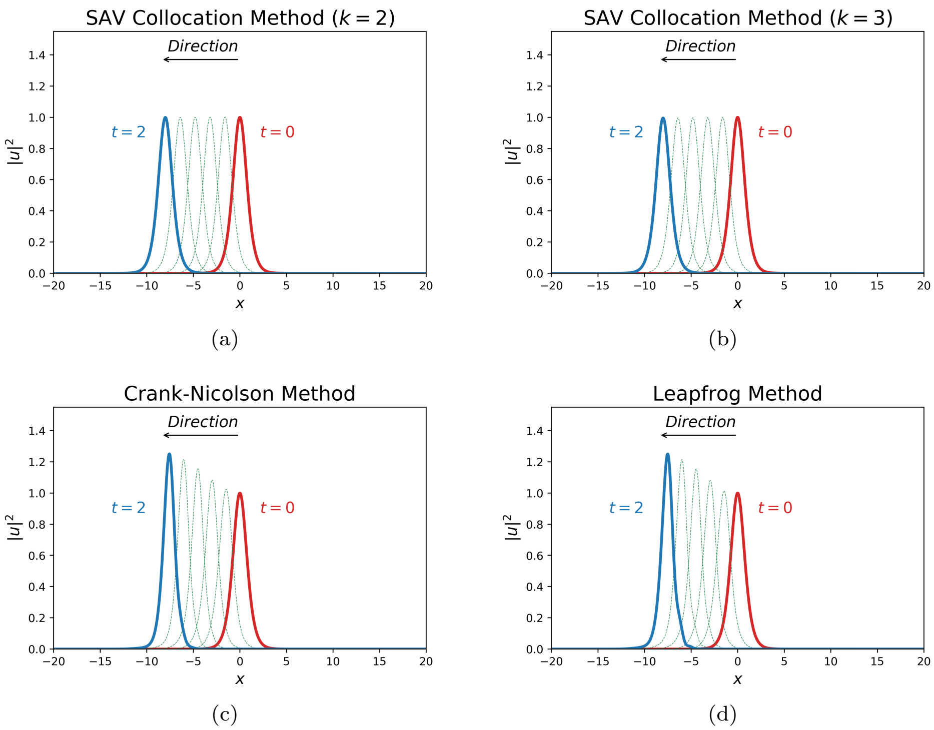

FIG. 4. *Soliton propagation when* $t \in [0, 2]$*: Numerical solutions with* $p = 1$*,* $M = 1200$*, and* $\Delta t = 0.1$.

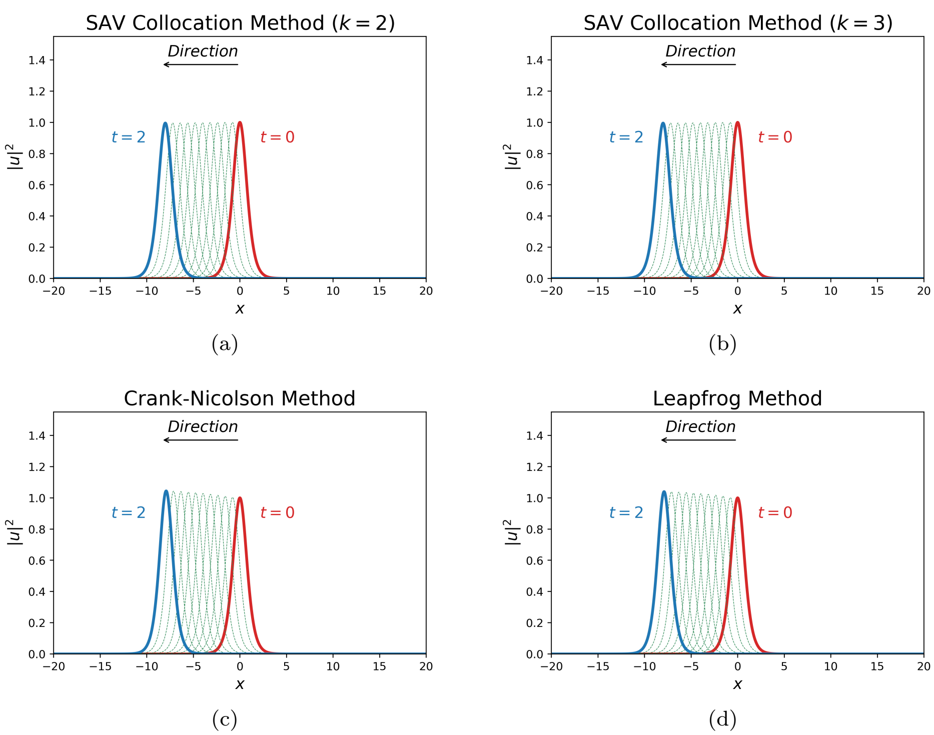

FIG. 5. *Soliton propagation when* $t \in [0, 2]$*: Numerical solutions with* $p = 1$*,* $M = 1200$*, and* $\Delta t = 0.05$.

TABLE 3
Time discretization errors of the proposed method with $h = \frac{1}{80}$ and $T = 0.1$.

k	τ	$p = 3$ $\|u(x,t) - u_h(x,t)\|_{L^\infty(0,T;H^1)}$	Order
2	1/460	5.0023E–04	–
	1/480	4.3780E–04	3.1321
	1/500	3.8572E–04	3.1027
	1/520	3.4198E–04	3.0686
	1/540	3.0504E–04	3.0290
3	1/60	1.6206E–02	–
	1/80	4.9792E–03	4.1022
	1/100	2.0173E–03	4.0490
	1/120	9.6960E–04	4.0183
	1/140	5.2530E–04	3.9761
4	1/30	3.6941E–02	–
	1/40	8.0993E–03	5.2750
	1/50	2.5534E–03	5.1731
	1/60	1.0078E–03	5.0989
	1/70	4.6554E–04	5.0104

TABLE 4
Spatial discretization errors of the proposed method with $\tau = \frac{1}{1000}$ and $T = 0.1$.

p	h	$k = 3$ $\|u(x,t) - u_h(x,t)\|_{L^\infty(0,T;H^1)}$	Order
1	1/70	5.6297E–01	–
	1/80	4.8304E–01	1.1466
	1/90	4.2346E–01	1.1178
	1/100	3.7726E–01	1.0964
	1/110	3.4035E–01	1.0803
2	1/10	4.9467E–01	–
	1/15	2.0992E–01	2.1141
	1/20	1.1748E–01	2.0178
	1/25	7.5177E–02	2.0005
	1/30	5.2233E–02	1.9972
3	1/12	2.1955E–02	–
	1/14	1.3738E–02	3.0412
	1/16	9.1747E–03	3.0236
	1/18	6.4327E–03	3.0144
	1/20	4.6849E–03	3.0092

error is negligibly small compared to the spatial error. From Table 4 we see that the spatial discretization errors are $O(h^p)$ in the H^1 norm. This is also consistent with the result proved in Theorem 4.4.

The evolution of mass and SAV energy of the numerical solutions are presented in Figure 6 with $\tau = 0.01$ and $h = 0.1$. It is shown that

$$\text{mass} = 1.000397142598 + O(10^{-12}) \quad \text{and} \quad \text{SAV energy} = 80.45628698537 + O(10^{-11}),$$

which are much smaller than the error of the numerical solutions, as shown in Figure 7. This shows the effectiveness of the proposed method in preserving mass and energy

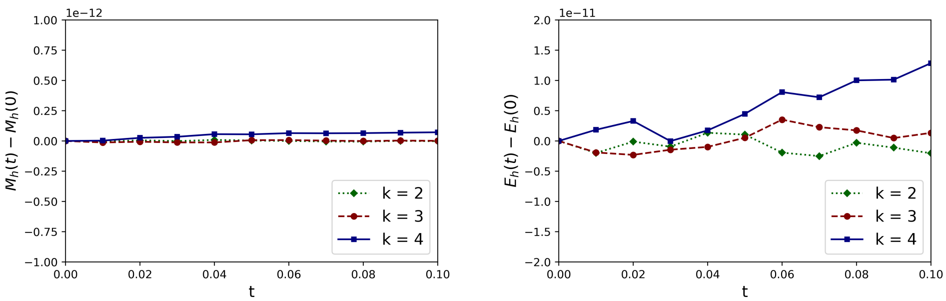

Fig. 6. *Evolution of mass* $M_h(t) - M_h(0)$ *and SAV energy* $E_h(t) - E_h(0)$ *with* $p = 3$, $\tau = 0.01$, *and* $h = 0.1$.

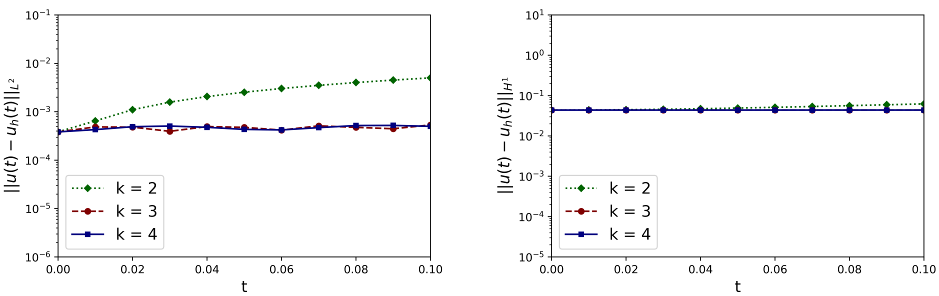

Fig. 7. *Evolution of error of the numerical solution, with* $p = 3$, $\tau = 0.01$, *and* $h = 0.1$.

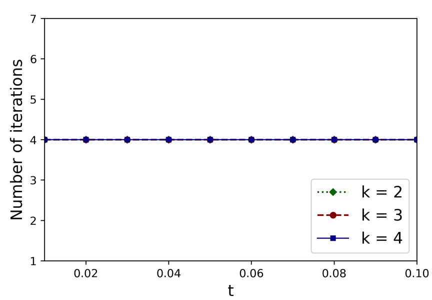

Fig. 8. *Number of iterations at each time level with* $p = 3$, $\tau = 0.01$, *and* $h = 0.1$.

(independent of the error of numerical solutions). The number of iterations at each time level is presented in Figure 8 to show the effectiveness of Newton's method.

REFERENCES

[1] G. D. Akrivis, V. A. Dougalis, and O. A. Karakashian, *On fully discrete Galerkin methods of second-order temporal accuracy for the NLS equation*, Numer. Math., 59 (1991), pp. 31–53.

[2] G. Akrivis, *Finite difference discretization of the cubic Schrödinger equation*, IMA J. Numer. Anal. 13 (1993), pp. 115–124.

[3] G. Akrivis, B. Li, and D. Li, *Energy-decaying extrapolated RK–SAV methods for the Allen–Cahn and Cahn–Hilliard equations*, SIAM J. Sci. Comput., 41 (2019), pp. A3703–A3727.

[4] X. Antoine, W. Bao, and C. Besse, *Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations*, Comput. Phys. Commun., 184 (2013), pp. 2621–2633.

[5] 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.
[6] W. Bao, Q. Tang, and Z. Xu, *Numerical methods and comparison for computing dark and bright solitons in the NLS equation*, J. Comput. Phys., 235 (2013), pp. 423–445.
[7] J. Bourgain, *Global Solutions of Nonlinear Schrödinger Equations*, Amer. Math. Soc. Colloq. Pub. 46, American Mathematical Society, Providence, RI, 1999.
[8] S. C. Brenner and L. R. Scott, *The Mathematical Theory of FEMs*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008.
[9] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*, Springer, Berlin, 2007.
[10] M. Delfour, M. Fortin, and G. Payre, *Finite-difference solutions of a non-linear Schrödinger equation*, J. Comput. Phys., 44 (1981), pp. 277–288.
[11] L. C. Evans, *Partial Differential Equations*, 2nd ed., Grad. Stud. Math. 19, AMS, Providence, RI, 2010.
[12] X. Feng, B. Li, and S. Ma, *High-Order Mass- and Energy-Conserving SAV–Gauss Collocation Finite Element Methods for the Nonlinear Schrödinger Equation*, preprint, arXiv:2006.05073.
[13] X. Feng, H. Liu, and S. Ma, *Mass- and energy-conserved numerical schemes for nonlinear Schrödinger equations*, Commun. Comput. Phys., 26 (2019), pp. 1365–1396.
[14] Z. Gao and S. Xie, *Fourth-order alternating direction implicit compact finite difference schemes for two-dimensional Schrödinger equations*, Appl. Numer. Math., 61 (2011), pp. 593–614.
[15] G. H. Golub and J. H. Welsch, *Calculation of Gauss quadrature rules*, Math. Comput., 23 (1969), pp. 221–230.
[16] Y. Gong, Q. Wang, and Y. Wang, *A conservative Fourier pseudo-spectral method for the nonlinear Schrödinger equation*, J. Comput. Phys., 328 (2017), pp. 354–370.
[17] Y. Gong and J. Zhao, *Energy-stable Runge–Kutta schemes for gradient flow models using the energy quadratization approach*, Appl. Math. Lett., 94 (2019), pp. 224–231.
[18] Y. Gong, J. Zhao, and Q. Wang, *Arbitrarily high-order unconditionally energy stable schemes for thermodynamically consistent gradient flow models*, SIAM J. Sci. Comput. 42 (2020), pp. B135–156.
[19] P. Henning and D. Peterseim, *Crank–Nicolson Galerkin approximations to nonlinear Schrödinger equations with rough potentials*, Math. Models Methods Appl. Sci., 27 (2017), pp. 2147–2184.
[20] J. Hong, Y. Liu, H. Munthe-Kaas, and A. Zanna, *Globally conservative properties and error estimation of a multi-symplectic scheme for Schrödinger equations with variable coefficients*, Appl. Numer. Math., 56 (2006), pp. 814–843.
[21] O. Karakashian and C. Makridakis, *A space-time finite element method for the nonlinear Schrödinger equation: The discontinuous Galerkin method*, Math. Comp., 67 (1998), pp. 479–499.
[22] O. Karakashian and C. Makridakis, *A space-time finite element method for the nonlinear Schrödinger equation: The continuous Galerkin method*, SIAM J. Numer. Anal., 36 (1999), pp. 1779–1807.
[23] D. A. Kopriva, *Implementing Spectral Methods for Partial Differential Equations: Algorithms for Scientists and Engineers*, Springer, Dordrecht, Netherlands, 2009.
[24] H. Liu, Y. Huang, W. Lu, and N. Yi, *On accuracy of the mass-preserving DG method to multi-dimensional Schrödinger equations*, IMA J. Numer. Anal., 39 (2019), pp. 760–791
[25] W. Lu, and Y. Huang, and H. Liu, *Mass preserving discontinuous Galerkin methods for Schrödinger equations*, J. Comput. Phys., 282 (2015), pp. 210–226.
[26] D. E. Pelinovsky, V. V. Afanasjev, and Y. S. Kivshar, *Nonlinear theory of oscillating, decaying, and collapsing solitons in the generalized nonlinear Schrödinger equation*, Phys. Rev. E(3), 53 (1996), pp. 1940–1953.
[27] J. M. Sanz-Serna, *Methods for the numerical solution of the nonlinear Schrödinger equation*, Math. Comp., 43 (1984), pp. 21–27
[28] H. W. Schürmann, *Traveling-wave solutions of the cubic-quintic nonlinear Schrödinger equation*, Phys. Rev. E(3), 54 (1996), pp. 4312–4320.
[29] J. Shen, J. Xu, and J. Yang, *A new class of efficient and robust energy stable schemes for gradient flows*, SIAM Rev., 61 (2019), pp. 474–506.
[30] J. Shen, J. Xu, and J. Yang, *The scalar auxiliary variable (SAV) approach for gradient flows*, J. Comput. Phys., 353 (2018), pp. 407–416.
[31] J. Shen, T. Tang, and L. L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer Ser. Comput. Math. 41, Springer, Berlin, 2011.

[32] W. A. Strauss and L. Vazquez, *Numerical solution of a nonlinear Klein–Gordon equation*, J. Comput. Phys., 28 (1978), pp. 271–278.

[33] N. Taghizadeh, M. Mirzazadeh, and F. Farahrooz, *Exact solutions of the nonlinear Schrödinger equation by the first integral method*, J. Math. Anal. Appl., 374 (2011), pp. 549–553.

[34] T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*, American Mathematical Society, Providence, RI, 2006.

[35] J. Wang, *A new error analysis of Crank–Nicolson Galerkin FEMs for a generalized nonlinear Schrödinger equation*, J. Sci. Comput., 60 (2014), pp. 390–407.

[36] T. Wang, B. Guo and Q. Xu, *Fourth-order compact and energy conservative difference schemes for the nonlinear Schrödinger equation in two dimensions*, J. Comput. Phys., 243 (2013), pp. 382–399.

[37] Y. Xu and C. W. Shu, *Local discontinuous Galerkin methods for nonlinear Schrödinger equations*, J. Comput. Phys., 205 (2005), pp. 72–97.

[38] X. Yang, *Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends*, J. Comput. Phys. 327 (2016), pp. 294–316.

[39] X. Yang and L. Ju, *Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model*, Comput. Methods Appl. Mech. Engrg., 315 (2017), pp. 691–712.

[40] X. Yang and L. Ju, *Linear and unconditionally energy stable schemes for the binary fluid-surfactant phase field model*, Comput. Methods Appl. Mech. Engrg., 318 (2017), pp. 1005–1029.

[41] X. Yang, J. Zhao, Q. Wang, and J. Shen, *Numerical approximations for a three component Cahn–Hilliard phase-field model based on the invariant energy quadratization method*, Math. Models Methods Appl. Sci., 27 (2017), pp. 1993–2030.

[42] N. J. Zabusky and M. D. Kruskal, *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett., 15 (1965), pp. 240–243.