Advances in Computational Mathematics (2023) 49:77
https://doi.org/10.1007/s10444-023-10077-5

Mass- and energy-conserving Gauss collocation methods for the nonlinear Schrödinger equation with a wave operator

Shu Ma[1] · **Jilu Wang**[2] · **Mingyan Zhang**[3] · **Zhimin Zhang**[4]

Received: 3 December 2022 / Accepted: 29 September 2023 / Published online: 23 October 2023

Abstract

A fully discrete finite element method with a Gauss collocation in time is proposed for solving the nonlinear Schrödinger equation with a wave operator in the d-dimensional torus, $d \in \{1, 2, 3\}$. Based on Gauss collocation method in time and the scalar auxiliary variable technique, the proposed method preserves both mass and energy conservations at the discrete level. Existence and uniqueness of the numerical solutions to the nonlinear algebraic system, as well as convergence to the exact solution with order $O(h^p + \tau^{k+1})$ in the $L^\infty(0, T; H^1)$ norm, are proved by using Schaefer's fixed point theorem without requiring any grid-ratio conditions, where (p, k) is the degree of the space-time finite elements. The Newton iterative method is applied for solving the nonlinear algebraic system. The numerical results show that the proposed method preserves discrete mass and energy conservations up to machine precision, and requires only a few Newton iterations to achieve the desired accuracy, with optimal-order convergence in the $L^\infty(0, T; H^1)$ norm.

Keywords Nonlinear Schrödinger equation with wave operator · Mass and energy conservation · High order · Scalar auxiliary variable · Gauss collocation · Error estimate

Mathematics Subject Classification (2010) 65N12 · 65M70 · 65M15 · 35Q55

Communicated by: Ilaria Perugia

All authors contributed equally to this work

✉ Mingyan Zhang
myzhang@csrc.ac.cn

Extended author information available on the last page of the article

1 Introduction

We consider the nonlinear Schrödinger equation with a wave operator (NLSW) on the d-dimensional torus $\Omega = [0, L]^d$, $d = 1, 2, 3$, up to a given time $T > 0$, i.e.,

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

$$u = u_0, \quad \partial_t u = u_1 \qquad \text{on} \Omega \times \{0\}, \tag{1.1b}$$

subject to periodic boundary condition, where $u : \Omega \to \mathbb{C}$ is a complex-valued function, $\mathrm{i} = \sqrt{-1}$ denotes the imaginary unit on the complex plane, and $f : \overline{\mathbb{R}}_+ \to \mathbb{R}$ is the derivative of some smooth potential function $F : \overline{\mathbb{R}}_+ \to \mathbb{R}$. The best known examples for the nonlinear function in the NLSW equation 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 "−" and "+" correspond to the defocusing and focusing models of the NLSW equation, respectively. As one of the most important NLS-type equations, the NLSW equation arises from many physical applications, including the Langmiur wave envelop approximation in plasma [1], the modulated planar pulse approximation of the sine-Gordon equation for light bullets [2], the non-relativistic limit of the Klein-Gordon equation [3], water waves, bimolecular dynamics, protein chemistry and so on. Accordingly, the development of efficient numerical methods for the NLSW equation is of great practical importance.

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

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \left(\frac{1}{2}|u|^2 + \mathrm{Im}(u_t \overline{u}) \right) \mathrm{d}x = 0, \quad \text{(mass conservation)} \tag{1.3}$$

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

where $\overline{v}$ and $\mathrm{Im}(v)$ denote the conjugate and imaginary part of v, respectively. Since numerical blow-up may be encountered in nonconservative schemes for the NLS-type equation [4], the development of mass- and energy-conserving numerical methods is important for long-time numerical simulation, and therefore has generated much attention in the numerical approximation to the NLS-type equations.

The proposition of conservative numerical methods and numerical analysis of NLS-type equations have been widely studied in the past years. Delfour et al. [5] proposed the first second-order modified Crank-Nicolson scheme for the NLS equation to preserve both mass and energy conservations. Since then, mass- and energy-conserving numerical methods for NLS equation have been investigated extensively, see [4, 6–17]. For the NLSW equation, Bao and Cai [18] established uniform error estimates of the Crank–Nicolson finite difference method for the NLS equation perturbed by the wave operator with a perturbation strength, and Guo and Xu [19] utilized a Crank–Nicolson local discontinuous Galerkin method to imporve the spatial accuracy. Both schemes preserve energy conservation at the discrete level. Besides the Crank–Nicolson scheme, several fully discrete finite difference methods and compact finite difference methods were developed in [20–22] and in [23–25], respectively. In particular, all of these schemes have second-order accuracy in time and only preserve energy conservation at the discrete level. To avoid solving nonlinear systems, a second-order energy-conserving linearized leapfrog finite element method was proposed by Cai et al. in [26], and a linearly second-order energy-conserving scheme was proposed by Cheng et al. in [27, 28] based on the scalar auxiliary variable (SAV) approach, which was introduced in [29, 30] as an enhanced version of the invariant energy quadratization (IEQ) approach [31–34], to develop energy-decay methods for dissipative (gradient flow) systems. Up to now, only very few high-order numerical methods have been proposeed for the NLSW equation. Brugnano et al. [35] presented a high-order energy-conserving Runge-Kutta scheme by using the Hamiltonian boundary value methods. Li et al. [36] proposed a class of high-order symplectic Runge-Kutta schemes based on the SAV approach. However, all these schemes only preserve energy conservation at the discrete level. In addition, the high-order convergence results in [35, 36] were only confirmed in the numerical experiments, and no rigorous analysis for convergence of numerical methods has been established so far.

To the best of our knowledge, all the existing mass- and energy-conserving methods for solving the NLSW equation have only second-order accuracy in time and is of the Crank-Nicolson type. No higher-order time-stepping schemes, which conserve both mass and energy, have been reported in the literature. Most recently, the SAV approach was combined with Gauss collocation method in time to preserve mass and energy conservations when solving the NLS equations in [37]. This new scheme has motivated us to propose a high-order mass- and energy-conserving method for the NLSW equation.

The objective of this paper is to develop high-order conserving methods based on the SAV formulation of the NLSW equation and the space-time FEM, and in the meantime establish a complete theory of existence, uniqueness and convergence of numerical solutions without grid-ratio conditions. The proposed method can be made arbitrarily high-order in both time and space discretizations, and retains the mass and energy conservation properties at the discrete level. Three main difficulties are overcome in the proposeion and analysis of our numerical method:

- The SAV reformulation of the NLSW equation is used to preserve the energy conservation at the discrete level. Due to the presence of $\partial_{tt}u$ in the equation (1.1), we can not directly obtain the error estimates as in [37]. This difficulty is overcome in the current paper by introducing an auxiliary variable $v = \partial_t u$ to establish an equivalent system of (1.1). Then the adoption of the SAV reformulation to the equivalent system avoids the presence of $\partial_t u$ in the equation of r (see equation (3.3)).
- The introduction of the new variable v makes the $W^{1,\infty}$-stability of the temporal Ritz projection in [37] insufficient. This is compensated in this paper by presenting the L^∞-stability estimate of the temporal Ritz projection; see (2.22).
- In the analysis for optimal-order convergence, it is difficult to give an estimate of the error function for u. By exploiting the relationship of the error equation

for u and v, the error estimates are proved together. Existence and uniqueness are obtained based on Schaefer's fixed point theorem in an L^∞-neighborhood of the exact solution. This avoids grid-ratio conditions for the equation (1.1) with general nonlinearity.

The rest of this article is organized as follows. In Section 2, we provide some preliminary results about the space-time finite element space. In Section 3, based on the SAV reformulation of the NLSW equation, we propose the SAV space-time Gauss collocation finite element method with its integral reformulation. The mass and energy conservation properties, and a consistency error estimate for the proposed method are also presented. In Section 4, we prove the existence, uniqueness and an optimal-order H^1-norm error estimate of numerical solutions by using Schaefer's fixed point theorem, and propose a key technical lemma, which is proved in Section 5. Numerical experiments are provided to support theoretical results in Section 6.

2 Notation and preliminary results

2.1 Function space

Let $W^{m,p}(\Omega)$ be the standard complex-valued Sobolev space of functions on Ω with $m \geq 0$ and $1 \leq p \leq \infty$, and denote

$$H^m(\Omega) = W^{m,2}(\Omega) \quad \text{and} \quad L^2(\Omega) = H^0(\Omega).$$

For any two complex functions $u,\ v \in L^2(\Omega)$, we define the $L^2(\Omega)$ inner product and norm as

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

For $m, s \geq 0$ and $1 \leq p \leq \infty$, We adopt notation $W^{m,p}(0, T; H^s(\Omega))$ to denote the space-time Sobelev space; see [38]. We abbreviate the norms of $H^s(\Omega)$ and $W^{m,p}(0, T; H^s(\Omega))$ as $\|\cdot\|_{H^s}$ and $\|\cdot\|_{W^{m,p}(I_n;H^s)}$, respectively, omitting the dependence on Ω in the subscripts.

Throughout this paper, we denote by C a generic positive constant which may vary at different occurrences and be depend on T and the regularity of solution, but it is always be independent of τ, h, n and N.

2.2 Spatial L^2 and Ritz projections

Let $\mathcal{T}_h = \{K_j : j = 1, 2, \dots, J\}$ be a shape-regular and quasi-uniform division of Ω into simplices K_j, $j = 1, 2, \dots, J$, with mesh size $h = \max_j\{\mathrm{diam}K_j\}$ less than 1. For any integer $p \geq 1$, we denote by $\mathbb{Q}^p$ the space of complex-valued polynomials on Ω with degree smaller than or equal to p. We define the complex-valued Lagrange finite element space $\mathcal{S}_h$ by

$$\mathcal{S}_h = \{v \in H^1(\Omega) : v|_K \in \mathbb{Q}^p \text{ for all } K \in \mathcal{T}_h\}. \tag{2.1}$$

The L^2 projection operator $P_h : L^2(\Omega) \to \mathcal{S}_h$ is defined as follow:

$$(w - P_h w, v_h) = 0 \quad \forall\, v_h \in \mathcal{S}_h \quad \forall\, w \in L^2(\Omega). \tag{2.2}$$

The following stability properties are well-known (cf. [39]):

$$\|P_h w\| \le \|w\| \qquad \forall\, w \in L^2(\Omega), \tag{2.3a}$$

$$\|P_h w\|_{H^1} \le C\|w\|_{H^1} \qquad \forall\, w \in H^1(\Omega), \tag{2.3b}$$

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

The Ritz projection operator $R_h : H^1(\Omega) \to \mathcal{S}_h$ is defined by

$$\begin{aligned} &(\nabla(w - R_h w), \nabla v_h) = 0 \qquad \forall\, v_h \in \mathcal{S}_h \quad w \in H^1(\Omega), \\ &\int_\Omega (w - R_h w)\mathrm{d}x = 0 \end{aligned}$$

where the additional condition $\int_\Omega (w - R_h w)\mathrm{d}x = 0$ guarantees the uniqueness of $R_h w$ (otherwise $R_h w$ is only defined up to a constant). The discrete Laplacian operator $\Delta_h : \mathcal{S}_h \to \mathcal{S}_h$ is defined as

$$(\Delta_h \phi_h, \chi_h) = -(\nabla \phi_h, \nabla \chi_h) \quad \forall\, \phi_h,\ \chi_h \in \mathcal{S}_h. \tag{2.4}$$

It is easy to verify that the L^2 projection, Ritz projection and discrete Laplacian satisfy the following relations:

$$P_h \Delta v = \Delta_h R_h v \qquad \forall\, v \in H^1(\Omega), \tag{2.5a}$$

$$R_\tau^n R_h v = R_h R_\tau^n v \qquad \forall\, v \in W^{1,\infty}(I_n; H^1(\Omega)), \tag{2.5b}$$

$$R_\tau^n \Delta_h v_h = \Delta_h R_\tau^n v_h \qquad \forall\, v \in W^{1,\infty}(I_n; \mathcal{S}_h), \tag{2.5c}$$

and the following approximation property:

$$\|v - R_h v\|_{H^1} \le C h^p \|v\|_{H^{p+1}} \qquad \forall\, v \in H^{p+1}(\Omega). \tag{2.6}$$

The first relation of (2.5) is a simple consequence of the definitions of P_h, R_h and Δ_h. The second and third relations of (2.5) simply mean that the temporal and spatial discrete operators commute with each other. This is also a simple consequence of the definition of these operators. We refer to [39] for the rigorous proof of (2.6).

2.3 Space-time finite element space

Let $0 = t_0 < t_1 < \cdots < t_N = T$ be a uniform partition of time interval $[0, T]$ with stepsize $\tau := \frac{T}{N}$ less than 1. For any integer $k \geq 1$, we denote by $\mathbb{P}^k$ the space of real-valued polynomials in time with degree smaller than or equal to k. For a Banach space X, such as $X = L^2(\Omega)$ or $X = \mathcal{S}_h$, we define the tensor-product space:

$$\mathbb{P}^k \otimes X := \mathrm{span}\Big\{p(t)\phi(x) : p \in \mathbb{P}^k,\ \phi \in X\Big\} = \Big\{\textstyle\sum_{j=0}^k t^j \phi_j : \phi_j \in X\Big\}, \tag{2.7}$$

and the following global space-time finite element spaces :

$$X_{\tau,h} = \{v_h \in C([0, T]; \mathcal{S}_h) : v_h|_{I_n} \in \mathbb{P}^k \otimes \mathcal{S}_h \quad \text{for } n = 1, ..., N\}, \tag{2.8}$$

$$Y_{\tau,h} = \{q_h \in C([0, T]) : q_h|_{I_n} \in \mathbb{P}^k \quad \text{for } n = 1, ..., N\}. \tag{2.9}$$

2.4 Temporal L^2 projection

We denote by c_j and w_j, $j = 1, ...k$, the nodes and weights of the k-point Gauss quadrature rule on the interval $[-1, 1]$ (cf. [40, Table 3.1]), and denote by $t_{nj} = t_{n-1} + (1 + c_j)\tau/2$, $j = 1, ..., k$, the Gauss points on the subinterval $I_n = [t_{n-1}, t_n]$.

The temporal L^2 projection on the subinterval I_n, i.e., $P_\tau^n : L^2(I_n; L^2(\Omega)) \to \mathbb{P}^{k-1} \otimes L^2(\Omega)$, is defined by

$$\int_{I_n} (u - P_\tau^n u, v)\,\mathrm{d}t = 0 \quad \forall\, v \in \mathbb{P}^{k-1} \otimes L^2(\Omega). \tag{2.10}$$

Based on this definition, $u - P_\tau^n u$ is orthogonal to all the temporal polynomials of degree $\leq k - 1$. In particular, if $u \in \mathbb{P}^k \otimes L^2(\Omega)$ then

$$u - P_\tau^n u = \phi_{n-1} L_k, \tag{2.11}$$

where $\phi_{n-1} \in L^2(\Omega)$ and $L_k(t) := \widehat{L}_k((2t - t_{n-1} - t_n)/\tau)$ is the shifted Legendre polynomial of degree k which is orthogonal to all the polynomials of lower degree on I_n. It is well known that the following approximation property (cf. [41]):

$$\max_{t\in I_n} \|v - P_\tau^n v\|_X \leq C\tau^m \max_{t\in I_n} \|\partial_t^m v\|_X, \quad 0 \leq m \leq k, \tag{2.12}$$

for all $v \in C^k([0, T]; X)$, and the following superconvergence result (cf. [37, Lemma 3.4]):

$$\|wv - P_\tau^n(wv)\|_{L^2(I_n;X)} \leq C\tau\|v\|_{L^2(I_n;X)}, \tag{2.13}$$

for all $w \in W^{k,\infty}(I_n, W^{s,\infty}(\Omega))$ and $v \in \mathbb{P}^{k-1} \otimes X$, where $X = \mathbb{R}$ or $X = H^s(\Omega)$ for some $s \in \mathbb{R}$, and the constant C is dependent on w.

Since the k-point Gauss quadrature holds exactly for polynomials of degree $2k-1$ (cf. [42, p. 222]), it follows from (2.11) that the following two identities hold:

$$\int_{I_n} v(t)\,\mathrm{d}t = \frac{\tau}{2}\sum_{j=1}^{k} v(t_{nj})w_j \qquad \forall\, v \in \mathbb{P}^{2k-1} \otimes S_h, \tag{2.14}$$

$$v(t_{nj}) = P_\tau^n v(t_{nj}) \qquad \forall\, v \in \mathbb{P}^{k} \otimes S_h, \tag{2.15}$$

where the Gauss points t_{nj}, $j = 1, ..., k$ are the roots of the Legendre polynomial $L_k(t)$ (cf. [43, p. 33]). By using (2.11) and the two identities (2.14)–(2.15), the following two inequalities were proved in [37]:

$$\int_{I_n} \|u_h\|^2\,\mathrm{d}t \le C\int_{I_n} \|P_\tau^n u_h\|^2\,\mathrm{d}t + C\tau\|u_h(t_{n-1})\|^2 \qquad \forall\, u_h \in \mathbb{P}^k \otimes \mathcal{S}_h, \tag{2.16}$$

$$\frac{\tau}{2}\sum_{j=1}^{k} w_j\|v_h(t_{nj})\|^2 = \int_{I_n} \|P_\tau^n v_h(t)\|^2\,\mathrm{d}t \le \int_{I_n} \|v_h(t)\|^2\,\mathrm{d}t \qquad \forall\, v_h \in \mathbb{P}^k \otimes \mathcal{S}_h. \tag{2.17}$$

The integral identities (2.14)–(2.15) and inequalities (2.16)–(2.17) in this subsection will be used in the subsequent analysis of existence, uniqueness and convergence of numerical solutions.

2.5 Temporal Ritz projection

We denote by I_τ^n the temporal Lagrange interpolation operator at points t_{n-1} and t_{nj}, $j = 1, \ldots, k$. For the one-dimensional temporal Lagrange interpolation operator I_τ^n, the following approximation property holds (cf. [41]):

$$\max_{t\in I_n}\left(\|v - I_\tau^n v\|_X + \tau\|\partial_t(v - I_\tau^n v)\|_X\right) \le C\tau^{m+1}\max_{t\in I_n}\|\partial_t^{m+1}v\|_X, \tag{2.18}$$

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 any $s \in \mathbb{R}$.

We denote by $R_\tau^n : W^{1,\infty}(I_n; L^2(\Omega)) \to \mathbb{P}^k \otimes L^2(\Omega)$ the temporal Ritz projection defined by

$$\int_{I_n} (\partial_t(u - R_\tau^n u), v)\,\mathrm{d}t = 0 \quad \forall\, v \in \mathbb{P}^{k-1} \otimes L^2(\Omega), \tag{2.19}$$

$$u(t_{n-1}) - R_\tau^n u(t_{n-1}) = 0. \tag{2.20}$$

It has the following approximation property (cf. [37, Lemma 3.3]):

$$\|u - R_\tau^n u\|_{L^\infty(I_n,X)} + \tau \|\partial_t (u - R_\tau^n u)\|_{L^\infty(I_n,X)} \le C\tau^{m+1} \|u\|_{W^{m+1,\infty}(I_n;X)}, \quad (2.21)$$

for all $u \in W^{m+1,\infty}(I_n; X)$ with $0 \le m \le k$.

In addition to the approximation property in (2.21), we shall also use the following pointwise stability result.

Lemma 1 *The temporal Ritz projection defined in* (2.19) *satisfies the following pointwise stability estimate:*

$$\|R_\tau^n u\|_{L^\infty(I_n;X)} \le C\|u\|_{L^\infty(I_n;X)} \quad \forall u \in L^\infty(I_n; X), \quad (2.22)$$

where $X = \mathbb{R}$ *or* $X = H^s(\Omega)$ *for any* $s \in \mathbb{R}$.

Proof We consider the shifted Legendre polynomial $L_k(t) := \hat{L}_k\left(\frac{2t-t_{n-1}-t_n}{\tau}\right)$ on the interval $I_n = [t_{n-1}, t_n]$. Let $X = \mathbb{R}$ or $X = H^s(\Omega)$ for some $s \in \mathbb{R}$. Then, by using integration by parts and the properties of the Legendre polynomial $\hat{L}_k(t)$, $t \in [-1, 1]$ (see [40, (3.176b)]), we have

$$\begin{aligned}
&\left\| \int_{I_n} L_j(s)\partial_s u(s)\, \mathrm{d}s \right\|_X \\
&= \left\| L_j(t_n)u(t_n) - L_j(t_{n-1})u(t_{n-1}) - \int_{I_n} u(s)L_j^{'}(s)\, \mathrm{d}s \right\|_X \\
&= \left\| u(t_n) + (-1)^{j+1}u(t_{n-1}) - \frac{2}{\tau} \sum_{m=0, m+j \text{ odd}}^{j-1} (2m+1) \int_{I_n} u(s)L_m(s)\, \mathrm{d}s \right\|_X \\
&\le \|u(t_n)\|_X + \|u(t_{n-1})\|_X + \sum_{m=0, m+j \text{ odd}}^{j-1} 2(2m+1)\|u(t)\|_{L^\infty(I_n;X)} \\
&\le C\|u(t)\|_{L^\infty(I_n;X)}.
\end{aligned}$$

This, together with the expression (cf. [37, (3.21)])

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

yields the following result:

$$
\begin{aligned}
\|R_\tau^n u(t)\|_{L^\infty(I_n;X)} &\le \|u(t_{n-1})\|_X + \sum_{j=0}^{k-1} \frac{\int_{I_n} |L_j(s)|\,\mathrm{d}s}{\int_{I_n} |L_j(s)|^2\,\mathrm{d}s} \left\| \int_{I_n} L_j(s)\partial_s u(s)\,\mathrm{d}s \right\|_X \\
&\le \|u(t_{n-1})\|_X + \sum_{j=0}^{k-1} C\|u(t)\|_{L^\infty(I_n;X)} \\
&\le C\|u(t)\|_{L^\infty(I_n;X)},
\end{aligned}
$$

where we have used $\int_{I_n} |L_j(t)|\,\mathrm{d}t \le C\tau$ and $\int_{I_n} |L_j(t)|^2\,\mathrm{d}t = \frac{\tau}{2j+1}$; see [40, (3.174a)]. □

3 SAV-Gauss collocation finite element method

In this section, we propose a Gauss collocation finite element method based on the SAV reformulation of the NLSW equation, and prove some desired properties of the numerical method, including the mass and energy conservations.

3.1 The SAV reformulation of (1.1)

By introducing an auxiliary variable $v = \partial_t u$, the NLSW equation (1.1) can be rewritten as

$$
\begin{aligned}
\partial_t u &= v && \text{in } \Omega \times (0,T], && (3.1a)\\
\partial_t v + \mathrm{i}v - \Delta u - f(|u|^2)u &= 0 && \text{in } \Omega \times (0,T], && (3.1b)\\
u = u_0, \quad v &= u_1 && \text{in } \Omega \times \{0\}. && (3.1c)
\end{aligned}
$$

The SAV formulation (cf. [30]) of the equivalent problem (3.1) introduces a scalar auxiliary variable

$$
r = \sqrt{\int_\Omega \frac{1}{2} F(|u|^2)\,\mathrm{d}x + c_0} \quad \text{with} \quad g(u) = \frac{f(|u|^2)}{\sqrt{\int_\Omega \frac{1}{2} F(|u|^2)\,\mathrm{d}x + c_0}} \tag{3.2}
$$

with a positive c_0 (which guarantees that the function r has a positive lower bound), and reformulate (3.1) as

$$
\begin{aligned}
\partial_t u &= v && \text{in } \Omega \times (0,T], && (3.3a)\\
\partial_t v + \mathrm{i}v - \Delta u - rg(u)u &= 0 && \text{in } \Omega \times (0,T], && (3.3b)\\
\frac{\mathrm{d}r}{\mathrm{d}t} &= \frac{1}{2}\mathrm{Re}\big(g(u)u, v\big) && \text{in } \Omega \times (0,T], && (3.3c)\\
u = u_0, \quad v = u_1, \quad r &= r_0 && \text{in } \Omega \times \{0\}, && (3.3d)
\end{aligned}
$$

where $r_0 = \sqrt{\int_\Omega \frac{1}{2} F(|u_0|^2)\,\mathrm{d}x + c_0}$.

The mass and energy conservations in the SAV formulation (3.3) are

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \Big(\frac{1}{2}|u|^2 + \mathrm{Im}(v\overline{u})\Big)\,\mathrm{d}x = 0, \quad \text{(mass conservation)} \tag{3.4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\int_\Omega \frac{1}{2}\Big(|v|^2 + |\nabla u|^2\Big)\,\mathrm{d}x - r^2 + c_0\bigg) = 0. \quad \text{(energy conservation)} \tag{3.5}$$

Remark 1 For the NSLW equation (1.1), the following energy conservation holds

$$\frac{1}{2}\int_\Omega |u_t|^2 + |\nabla u|^2 - F(|u|^2)dx = \frac{1}{2}\int_\Omega |u_1|^2 + |\nabla u_0|^2 - F(|u_0|^2)dx, \quad \forall t > 0,$$

where $u_1 = \partial_t u|_{t=0}$ and $u_0 = u|_{t=0}$. To guarantee a positive lower bound for the function r, we determine the value of c_0 in the following manner in practical computation:

$$c_0 = \frac{1}{2}\int_\Omega |u_1|^2 + |\nabla u_0|^2 - F(|u_0|^2)dx.$$

It is noted that c_0 may be multiplied by a constant greater than 1 to adjust numerical behavior. The selection of c_0 primarily impacts the energy at each time step, but it does not affect the conservation properties and convergence orders of the numerical method.

3.2 The Gauss collocation method

Consider the interval $I_n = [t_{n-1}, t_n]$ with its Gauss points t_{nj} and weights w_j, $j = 1, \ldots, k$. We define the following Gauss collocation finite element method for the SAV formulation (3.3).

Set $(u_h^0, v_h^0, r_h^0) := (I_h u_0, I_h u_1, r_0)$.

For given $(u_h^{n-1}, v_h^{n-1}, r_h^{n-1}) \in \mathcal{S}_h \times \mathcal{S}_h \times \mathbb{R}$, we define $u_h|_{I_n} \in \mathbb{P}^k \otimes \mathcal{S}_h$, $v_h|_{I_n} \in \mathbb{P}^k \otimes \mathcal{S}_h$ and $r_h|_{I_n} \in \mathbb{P}^k$ by solving the following nonlinear algebraic system:

$$\big(\partial_t u_h(t_{nj}), \varphi_h\big) = \big(v_h(t_{nj}), \varphi_h\big) \quad \forall \varphi_h \in \mathcal{S}_h, \tag{3.6a}$$

$$\begin{aligned}\big(\partial_t v_h(t_{nj}), \psi_h\big) + \mathrm{i}\big(v_h(t_{nj}), \psi_h\big) + \big(\nabla u_h(t_{nj}), \nabla \psi_h\big)& \\ - \big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), \psi_h\big) = 0 \quad \forall \psi_h \in \mathcal{S}_h,&\end{aligned} \tag{3.6b}$$

$$\partial_t r_h(t_{nj}) = \frac{1}{2}\mathrm{Re}\big(g(u_h(t_{nj})) u_h(t_{nj}), v_h(t_{nj})\big), \tag{3.6c}$$

$$u_h(t_{n-1}) = u_h^{n-1}, \quad v_h(t_{n-1}) = v_h^{n-1} \;\text{ and }\; r_h(t_{n-1}) = r_h^{n-1}, \tag{3.6d}$$

which is obtained by collocating (3.3) at the nodes $t = t_{nj}$, $j = 1, \ldots, k$, and then imposing a continuity condition at $t = t_{n-1}$.

Set $u_h^n := u_h(t_n)$, $v_h^n := v_h(t_n)$ and $r_h^n := r_h(t_n)$.

Remark 2 (a) For each $n \geq 1$, algorithm (3.6) computes the numerical solution at the nodes t_{nj}, $j = 1, \ldots, k$ on the subinterval $I_n = [t_{n-1}, t_n]$. Since any polynomial of degree k on the subinterval I_n is uniquely determined by its initial value at t_{n-1} and its values at the nodes t_{nj}, $j = 1, \ldots, k$, it follows that the values of the numerical solution at the nodes uniquely determine a globally continuous numerical solution $(u_h, v_h, r_h) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$.

(b) The test functions φ_h and ψ_h can be different at different nodes t_{nj}. Hence, together with the one initial condition at $t = t_{n-1}$, the number of equations in the algebraic system is the same as the degree of freedoms of the underlying space $\mathbb{P}^k \otimes S_h$ from which we seek the numerical solution on the subinterval I_n.

(c) The following Newton iterative method can be applied to solve the nonlinear to system in the practical computation: For given $\left\{\left(u_h^{l-1}(t_{nj}), v_h^{l-1}(t_{nj}), r_h^{l-1}(t_{nj})\right)\right\}_{j=1}^k \subset \mathcal{S}_h \times \mathcal{S}_h \times \mathbb{R}$, find

$$\left\{\left(u_h^l(t_{nj}), v_h^l(t_{nj}), r_h^l(t_{nj})\right)\right\}_{j=1}^k \subset \mathcal{S}_h \times \mathcal{S}_h \times \mathbb{R}$$

satisfying the linearized equations

$$\left(\partial_t u_h^l(t_{nj}), \varphi_h\right) = \left(v_h^l(t_{nj}), \varphi_h\right) \quad \forall \varphi_h \in \mathcal{S}_h, \tag{3.7a}$$

$$\begin{aligned}\left(\partial_t v_h^l(t_{nj}), \psi_h\right) &+ \mathrm{i}\left(v_h^l(t_{nj}), \psi_h\right) + \left(\nabla u_h^l(t_{nj}), \nabla \psi_h\right) \\ &= \left(r_h^l(t_{nj}) g(u_h^{l-1}(t_{nj})) u_h^{l-1}(t_{nj}), \psi_h\right) \\ &\quad + \left(r_h^{l-1}(t_{nj}) g_1(u_h^{l-1}(t_{nj}))(u_h^l(t_{nj}) - u_h^{l-1}(t_{nj})), \psi_h\right) \\ &\quad + \left(r_h^{l-1}(t_{nj}) g_2(u_h^{l-1}(t_{nj}))(\overline{u}_h^l(t_{nj}) - \overline{u}_h^{l-1}(t_{nj})), \psi_h\right) \quad \forall \psi_h \in \mathcal{S}_h,\end{aligned} \tag{3.7b}$$

$$\begin{aligned}\partial_t r_h^l(t_{nj}) &= \frac{1}{2}\mathrm{Re}\left(g(u_h^{l-1}(t_{nj})) u_h^{l-1}(t_{nj}), v_h^l(t_{nj})\right) \\ &\quad + \frac{1}{2}\mathrm{Re}\left(g_1(u_h^{l-1}(t_{nj}))(u_h^l(t_{nj}) - u_h^{l-1}(t_{nj})), v_h^{l-1}(t_{nj})\right) \\ &\quad + \frac{1}{2}\mathrm{Re}\left(g_2(u_h^{l-1}(t_{nj}))(\overline{u}_h^l(t_{nj}) - \overline{u}_h^{l-1}(t_{nj})), v_h^{l-1}(t_{nj})\right),\end{aligned} \tag{3.7c}$$

$$u_h^l(t_{n-1}) = u_h^{n-1}, \quad v_h^l(t_{n-1}) = v_h^{n-1} \text{ and } r_h^l(t_{n-1}) = r_h^{n-1}, \tag{3.7d}$$

where

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

and denote by $\partial_{\overline{u}}$ be the differentiation with respect to $\overline{u}$ in the expression of

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

A tolerance error will be set to control the number of iterations.

3.3 An integral reformulation of the algorithm

By setting $\varphi_h = \frac{\tau}{2}\varphi_h(t_{nj})w_j$ in (3.6a) and $\psi_h = \frac{\tau}{2}\psi_h(t_{nj})w_j$ in (3.6b), respectively, summing up the results for $j = 1, ..., k$, and using (2.14)–(2.15), we obtain the following integral identities:

$$\int_{I_n} (\partial_t u_h, \varphi_h)\,\mathrm{d}t = \int_{I_n} (P_\tau^n v_h, \varphi_h)\,\mathrm{d}t \quad \forall \varphi_h \in \mathbb{P}^k \otimes \mathcal{S}_h, \tag{3.8}$$

$$\begin{aligned} &\int_{I_n} (\partial_t v_h, \psi_h)\,\mathrm{d}t + \mathrm{i}\int_{I_n} (P_\tau^n v_h, \psi_h)\,\mathrm{d}t + \int_{I_n} (\nabla P_\tau^n u_h, \nabla\psi_h)\,\mathrm{d}t \\ &\quad -\frac{\tau}{2}\sum_{j=1}^{k} w_j\big(r_h(t_{nj})g(u_h(t_{nj}))u_h(t_{nj}), \psi_h(t_{nj})\big)=0 \ \ \forall \psi_h \in \mathbb{P}^k \otimes \mathcal{S}_h. \end{aligned} \tag{3.9}$$

Similarly, multiplying (3.6c) by $\frac{\tau}{2}q_h(t_{nj})w_j$, summing up the results for $j = 1, 2, ...,$ k, and using (2.14) gives

$$\int_{I_n} \partial_t r_h q_h\,\mathrm{d}t = \frac{\tau}{2}\sum_{j=1}^{k}\frac{w_j}{2}\mathrm{Re}\big(g(u_h(t_{nj}))u_h(t_{nj}), v_h(t_{nj})q_h(t_{nj})\big) \quad \forall q_h \in \mathbb{P}^k. \tag{3.10}$$

Remark 3 (a) Note that (3.8)–(3.10) provide an integral reformulation of algorithm (3.6). By using the properties of the Legendre polynomials on each interval I_n, this integral reformulation rewrites the Gauss collocation method (3.6) into a space-time Galerkin FEM, which makes it easier to choose suitable test functions. Thus, (3.8)–(3.10) will be crucially used later to show mass and energy conservations, as well as existence, uniqueness, and convergence of numerical solutions.

(b) Algorithm (3.6) can be obtained by applying the Gauss quadrature rule (in time) to the (continuous) space-time finite element method for (3.3).

3.4 Mass and energy conservations of the algorithm

The main result of this subsection is the following theorem.

Theorem 2 *Let* $(u_h, v_h, r_h) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ *be a solution of algorithm* (3.6). *Then the following mass and energy conservations hold for* $n \geq 1$*:*

$$\frac{1}{2}\|u_h(t_n)\|^2 + \mathrm{Im}\big(v_h(t_n), u_h(t_n)\big) = \frac{1}{2}\|u_h^0\|^2 + \mathrm{Im}\,(v_h^0, u_h^0), \tag{3.11}$$

$$\frac{1}{2}\|v_h(t_n)\|^2 + \frac{1}{2}\|\nabla u_h(t_n)\|^2 - |r_h(t_n)|^2 + c_0 = \frac{1}{2}\|v_h^0\|^2 + \frac{1}{2}\|\nabla u_h^0\|^2 - |r_h^0|^2 + c_0. \tag{3.12}$$

Proof By setting $\varphi_h = u_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ in (3.8) and considering the real part of the result, we obtain

$$\mathrm{Re} \int_{I_n} (\partial_t u_h, u_h)\, \mathrm{d}t = \mathrm{Re} \int_{I_n} (P_\tau^n v_h, u_h)\, \mathrm{d}t. \tag{3.13}$$

Then, by choosing $\psi_h = u_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ in (3.9) and considering the imaginary part of the result, we obtain

$$\begin{aligned}
\mathrm{Im} \int_{I_n} (\partial_t v_h, u_h)\, \mathrm{d}t &= -\mathrm{Im} \int_{I_n} \mathrm{i}(P_\tau^n v_h, u_h)\, \mathrm{d}t - \mathrm{Im} \int_{I_n} (\nabla P_\tau^n u_h, \nabla u_h)\, \mathrm{d}t \\
&\quad + \mathrm{Im} \frac{\tau}{2} \sum_{j=1}^{k} w_j \big(r_h(t_{nj}) g(u_h(t_{nj})), |u_h(t_{nj})|^2\big) \\
&= -\mathrm{Re} \int_{I_n} (P_\tau^n v_h, u_h)\, \mathrm{d}t \\
&= \frac{1}{2} \|u_h(t_{n-1})\|^2 - \frac{1}{2} \|u_h(t_n)\|^2,
\end{aligned} \tag{3.14}$$

where we have used (3.13) in the last equality, and the following relation in the second to last equality:

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

By choosing $\varphi_h = v_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ in (3.8) and considering the imaginary part of the result, we have

$$\mathrm{Im} \int_{I_n} (\partial_t u_h, v_h)\, \mathrm{d}t = \mathrm{Im} \int_{I_n} (P_\tau^n v_h, v_h)\, \mathrm{d}t = \mathrm{Im} \int_{I_n} (P_\tau^n v_h, P_\tau^n v_h)\, \mathrm{d}t = 0, \tag{3.15}$$

which implies that (via integration by parts in time)

$$\begin{aligned}
-\mathrm{Im}\int_{I_n} (\partial_t v_h, u_h)\, \mathrm{d}t &= \mathrm{Im}\int_{I_n} (v_h, \partial_t u_h)\, \mathrm{d}t - \mathrm{Im}[(v_h(t_n), u_h(t_n)) - v_h(t_{n-1}), u_h(t_{n-1}))] \\
&= -\mathrm{Im}(v_h(t_n), u_h(t_n)) + \mathrm{Im}(v_h(t_{n-1}), u_h(t_{n-1})).
\end{aligned} \tag{3.16}$$

This identity can be combined with (3.14) to yield the mass conservation of the numerical solution, i.e.,

$$\frac{1}{2} \|u_h(t_n)\|^2 + \mathrm{Im}\big(v_h(t_n), u_h(t_n)\big) = \frac{1}{2} \|u_h(t_{n-1})\|^2 + \mathrm{Im}\big(v_h(t_{n-1}), u_h(t_{n-1})\big).$$

Alternatively, by setting $\psi_h = v_h$ and $q_h = 2r_h$ in (3.9) and (3.10), respectively, and considering the real parts, we have

$$\mathrm{Re}\int_{I_n} (\partial_t v_h, v_h)\,\mathrm{d}t + \mathrm{Re}\int_{I_n} (\nabla P_\tau^n u_h, \nabla v_h)\,\mathrm{d}t \tag{3.17}$$
$$= \frac{\tau}{2}\sum_{j=1}^{k} w_j \mathrm{Re}\big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), v_h(t_{nj})\big),$$
$$|r_h(t_n)|^2 - |r_h(t_{n-1})|^2 = \frac{\tau}{2}\sum_{j=1}^{k} w_j \mathrm{Re}\big(r_h(t_{nj}) g(u_h(t_{nj})) u_h(t_{nj}), v_h(t_{nj})\big). \tag{3.18}$$

Setting $\varphi_h = \Delta_h u_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ in (3.8) and taking the real part yield

$$\mathrm{Re}\int_{I_n} (\partial_t \nabla u_h, \nabla u_h)\,\mathrm{d}t = \mathrm{Re}\int_{I_n} \left(\nabla P_\tau^n v_h, \nabla u_h\right)\,\mathrm{d}t, \tag{3.19}$$

which gives

$$\mathrm{Re}\int_{I_n} (\nabla P_\tau^n u_h, \nabla v_h)\,\mathrm{d}t = \mathrm{Re}\int_{I_n} (P_\tau^n \nabla v_h, \nabla u_h)\,\mathrm{d}t = \frac{1}{2}\|\nabla u_h(t_n)\|^2 - \frac{1}{2}\|\nabla u_h(t_{n-1})\|^2. \tag{3.20}$$

Then, subtracting (3.18) from (3.17) and using (3.20), we obtain the energy conservation, i.e.,

$$\frac{1}{2}\|v_h(t_n)\|^2 + \frac{1}{2}\|\nabla u_h(t_n)\|^2 - |r_h(t_n)|^2$$
$$= \frac{1}{2}\|v_h(t_{n-1})\|^2 + \frac{1}{2}\|\nabla u_h(t_{n-1})\|^2 - |r_h(t_{n-1})|^2. \tag{3.21}$$

This completes the proof of Theorem 2. □

3.5 Consistency of the algorithm

For the comparison with the numerical solutions, we define the following intermediate solutions in terms of the temporal and spatial Ritz projections:

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

The consistency errors are defined as:

$$d_u^n := \partial_t R_\tau^n (R_h u - u) + (v - v_h^*), \tag{3.22a}$$

$$d_v^n := \partial_t R_\tau^n (R_h v - v) + \mathrm{i}(v_h^* - v) + \Delta_h R_h (u - R_\tau^n u) + r g(u) u - I_\tau^n [r_h^* g(u_h^*) u_h^*], \tag{3.22b}$$

$$d_r^n := \frac{1}{2}\mathrm{Re}\big[\big(g(u)u, v\big) - I_\tau^n\big(g(u_h^*)u_h^*, v_h^*\big)\big]. \tag{3.22c}$$

It is easy to check that there hold:

$$\int_{I_n} (\partial_t u_h^*, \varphi_h)\,\mathrm{d}t = \int_{I_n} (P_\tau^n v_h^*, \varphi_h)\,\mathrm{d}t + \int_{I_n} (P_\tau^n d_u^n, \varphi_h)\,\mathrm{d}t, \tag{3.23a}$$

$$\begin{aligned} &\int_{I_n} (\partial_t v_h^*, \psi_h)\,\mathrm{d}t \; + \mathrm{i}\int_{I_n} (P_\tau^n v_h^*, \psi_h)\,\mathrm{d}t + \int_{I_n} (\nabla P_\tau^n u_h^*, \nabla \psi_h)\,\mathrm{d}t \\ &\quad = \frac{\tau}{2}\sum_{j=1}^{k} w_j \big(r_h^*(t_{nj}) g(u_h^*(t_{nj})) u_h^*(t_{nj}), \psi_h(t_{nj})\big) + \int_{I_n} (P_\tau^n d_v^n, \psi_h)\,\mathrm{d}t, \end{aligned} \tag{3.23b}$$

$$\int_{I_n} \partial_t r_h^* q_h\,\mathrm{d}t = \frac{\tau}{4}\sum_{j=1}^{k} w_j \mathrm{Re}\big(q_h(t_{nj}) g(u_h^*(t_{nj})) u_h^*(t_{nj}), v_h^*(t_{nj})\big) + \int_{I_n} P_\tau^n d_r^n q_h\,\mathrm{d}t. \tag{3.23c}$$

The proof of the following theorem can be found in Appendix.

Theorem 3 *Suppose that the solution of (1.1) is sufficiently smooth. Then $d_u^n, d_v^n \in C(I_n; H^1(\Omega))$ and the following estimate is satisfied:*

$$\sup_{t\in I_n} \|d_u^n\|_{H^1} + \sup_{t\in I_n} \|d_v^n\|_{L^2} + \sup_{t\in I_n} |P_\tau^n d_r^n| \le C(h^p + \tau^{k+1}). \tag{3.24}$$

4 Well-posedness and convergence analysis

We denote by $e_h^u = u_h - u_h^*$, $e_h^v = v_h - v_h^*$ and $e_h^r = r_h - r_h^*$ the error functions of the numerical approximation. For the simplicity of notation, we use the following abbreviations:

$$\begin{aligned} &e_{nj}^u = e_h^u(t_{nj}), && e_{nj}^v = e_h^v(t_{nj}) && \text{and} \quad e_{nj}^r = e_h^r(t_{nj}), \\ &u_{nj} = u_h(t_{nj}), && v_{nj} = v_h(t_{nj}) && \text{and} \quad r_{nj} = r_h(t_{nj}), \\ &u_{nj}^* = u_h^*(t_{nj}), && v_{nj}^* = v_h^*(t_{nj}) && \text{and} \quad r_{nj}^* = r_h^*(t_{nj}), \\ &\varphi_{nj} = \varphi_h(t_{nj}), && \psi_{nj} = \psi_h(t_{nj}) && \text{and} \quad q_{nj} = q_h(t_{nj}). \end{aligned}$$

Subtracting (3.23a)–(3.23c) from (3.8)–(3.10), we have the following error equations:

$$\int_{I_n} (\partial_t e_h^u, \varphi_h)\,\mathrm{d}t = \int_{I_n} (P_\tau^n e_h^v, \varphi_h)\,\mathrm{d}t - \int_{I_n} (P_\tau^n d_u^n, \varphi_h)\,\mathrm{d}t, \tag{4.1a}$$

$$\int_{I_n} (\partial_t e_h^v, \psi_h)\,\mathrm{d}t = -\mathrm{i}\int_{I_n} (P_\tau^n e_h^v, \psi_h)\,\mathrm{d}t - \int_{I_n} (\nabla P_\tau^n e_h^u, \nabla \psi_h)\,\mathrm{d}t \tag{4.1b}$$

$$+\frac{\tau}{2}\sum_{j=1}^{k} w_j \big(e_{nj}^r g(u_{nj})u_{nj}, \psi_{nj}\big)$$

$$+\frac{\tau}{2}\sum_{j=1}^{k} w_j \Big(r_{nj}^*\big[g(u_{nj})u_{nj} - g(u_{nj}^*)u_{nj}^*\big], \psi_{nj}\Big) - \int_{I_n} (P_\tau^n d_v^n, \psi_h)\,\mathrm{d}t, \tag{4.1c}$$

$$\int_{I_n} \partial_t e_h^r q_h\,\mathrm{d}t = \frac{\tau}{4}\sum_{j=1}^{k} w_j \mathrm{Re}\Big(q_{nj}\big[g(u_{nj})u_{nj} - g(u_{nj}^*)u_{nj}^*\big], v_{nj}^*\Big)$$

$$+\frac{\tau}{4}\sum_{j=1}^{k} w_j \mathrm{Re}\big(q_{nj} g(u_{nj})u_{nj}, e_{nj}^v\big) - \int_{I_n} P_\tau^n d_r^n q_h\,\mathrm{d}t, \tag{4.1d}$$

which hold for all test functions $\varphi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$, $\psi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ and $q_h \in \mathbb{P}^k$.

Remark 4 For the given (u_h^*, v_h^*, r_h^*), if (4.1a)–(4.1d) has a solution $(e_h^u, e_h^v, e_h^r) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ then $u_h = u_h^* + e_h^u$, $v_h = v_h^* + e_h^v$ and $r_h = r_h^* + e_h^r$ form a solution of the algorithm (3.6).

We denote by $X_{\tau,h}^* \times \widetilde{X}_{\tau,h}^* \times Y_{\tau,h}^*$ the subset of $X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ defined by

$$X_{\tau,h}^* = \Big\{\varphi_h \in X_{\tau,h} : \max_{1\le n\le N}\max_{1\le j\le k} \|\varphi_h(t_{nj}) - u_h^*(t_{nj})\|_{L^\infty\cap H^1} \le \frac{1}{3}\Big\}, \tag{4.2a}$$

$$\widetilde{X}_{\tau,h}^* = \Big\{\psi_h \in X_{\tau,h} : \max_{1\le n\le N}\max_{1\le j\le k} \|\psi_h(t_{nj}) - v_h^*(t_{nj})\|_{L^2} \le \frac{1}{3}\Big\}, \tag{4.2b}$$

$$Y_{\tau,h}^* = \Big\{q_h \in Y_{\tau,h} : \max_{1\le n\le N}\max_{1\le j\le k} |q_h(t_{nj}) - r_h^*(t_{nj})| \le \frac{1}{3}\Big\}. \tag{4.2c}$$

The main result of this section is the following theorem on the existence, uniqueness and convergence of the numerical solutions.

Theorem 4 *(Well-posedness and convergence) Let $d \in \{1,2,3\}$ and assume that the solution of the NLSW 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 nonlinear system in (3.6) has a unique solution $(u_h, v_h, r_h) \in X_{\tau,h}^* \times \widetilde{X}_{\tau,h}^* \times Y_{\tau,h}^*$. Moreover, this solution satisfies the following error estimate:*

$$\max_{t\in[0,T]} \Big(\|u_h(t) - u_h^*(t)\|_{H^1} + \|v_h(t) - v_h^*(t)\|_{L^2} + |r_h - r_h^*(t)|\Big) \le C(h^p + \tau^{k+1}), \tag{4.3}$$

Remark 5 (a) For the Dirichlet and Neumann boundary conditions, the mass and energy conservations in Theorem 2 and the error estimate in Theorem 4 can be proved similarly.

(b) To avoid grid-ratio conditions for (1.1) with general nonlinearity, the well-posedness and convergence of numerical solutions in Theorem 4 are proved based on Schaefer's fixed point theorem in an L^∞-neighborhood of the exact solution, which is quoted below.

Lemma 5 *(Schaefer's fixed point theorem [38, 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*

$$\left\{\phi \in B : \phi = \theta M(\phi) \ \text{for some } 0 \le \theta \le 1\right\} \tag{4.4}$$

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

For any $(\alpha_h, \gamma_h, \beta_h) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$, we define a norm on $X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ by

$$\|(\alpha_h, \gamma_h, \beta_h)\|_{X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}} := \|\alpha_h\|_{L^\infty(0,T;H^1)} + \|\gamma_h\|_{L^\infty(0,T;L^2)} + \|\beta_h\|_{L^\infty(0,T)}, \tag{4.5}$$

and three associated numbers

$$\rho[\alpha_h] := \min\left(\frac{1}{\max\limits_{1\le n\le N}\max\limits_{1\le j\le k}\|\alpha_h(t_{nj})\|_{L^\infty\cap H^1}}, 1\right), \tag{4.6a}$$

$$\rho[\gamma_h] := \min\left(\frac{1}{\max\limits_{1\le n\le N}\max\limits_{1\le j\le k}\|\gamma_h(t_{nj})\|_{L^2}}, 1\right), \tag{4.6b}$$

$$\rho[\beta_h] := \min\left(\frac{1}{\max\limits_{1\le n\le N}\max\limits_{1\le j\le k}|\beta_h(t_{nj})|}, 1\right), \tag{4.6c}$$

which are continuous with respect to $(\alpha_h, \gamma_h, \beta_h)$ and satisfy

$$\max_{1\le n\le N}\max_{1\le j\le k}\|\rho[\alpha_h]\alpha_h(t_{nj})\|_{L^\infty\cap H^1} \le 1, \tag{4.7a}$$

$$\max_{1\le n\le N}\max_{1\le j\le k}\|\rho[\gamma_h]\gamma_h(t_{nj})\|_{L^2} \le 1, \tag{4.7b}$$

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

Then we define

$$u_h^\alpha := u_h^* + \rho[\alpha_h]\alpha_h, \quad v_h^\gamma := v_h^* + \rho[\gamma_h]\gamma_h \ \text{ and } \ r_h^\beta := r_h^* + \rho[\beta_h]\beta_h, \tag{4.8}$$

with the following abbreviations:

$$u_{nj}^\alpha = u_h^\alpha(t_{nj}), \qquad v_{nj}^\gamma = v_h^\gamma(t_{nj}) \ \text{ and } \ r_{nj}^\beta = r_h^\beta(t_{nj}),$$

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

$$\int_{I_n} (\partial_t e_h^u, \varphi_h)\,\mathrm{d}t = \int_{I_n} (P_\tau^n \gamma_h, \varphi_h)\,\mathrm{d}t - \int_{I_n} (P_\tau^n d_u^n, \varphi_h)\,\mathrm{d}t, \tag{4.9a}$$

$$\begin{aligned}\int_{I_n} (\partial_t e_h^v, \psi_h)\,\mathrm{d}t + \mathrm{i}\int_{I_n} (P_\tau^n e_h^v, \psi_h)\,\mathrm{d}t &= -\int_{I_n} (\nabla P_\tau^n \alpha_h, \nabla \psi_h)\,\mathrm{d}t - \int_{I_n} (P_\tau^n d_v^n, \psi_h)\,\mathrm{d}t \\ &\quad + \frac{\tau}{2}\sum_{j=1}^{k} w_j \Big(r_{nj}^* \big[g(u_{nj}^\alpha) u_{nj}^\alpha - g(u_{nj}^*) u_{nj}^* \big], \psi_{nj} \Big) \\ &\quad + \frac{\tau}{2}\sum_{j=1}^{k} w_j \big(\beta_{nj} g(u_{nj}^\alpha) u_{nj}^\alpha, \psi_{nj} \big),\end{aligned} \tag{4.9b}$$

$$\begin{aligned}\int_{I_n} \partial_t e_h^r q_h\,\mathrm{d}t &= \frac{\tau}{4}\sum_{j=1}^{k} w_j \mathrm{Re}\Big(q_{nj} \big[g(u_{nj}^\alpha) u_{nj}^\alpha - g(u_{nj}^*) u_{nj}^* \big], v_{nj}^* \Big) \\ &\quad + \frac{\tau}{4}\sum_{j=1}^{k} w_j \mathrm{Re}\big(q_{nj} g(u_{nj}^\alpha) u_{nj}^\alpha, \gamma_{nj} \big) \\ &\quad - \int_{I_n} P_\tau^n d_r^n q_h\,\mathrm{d}t,\end{aligned} \tag{4.9c}$$

which hold for all test functions $\varphi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$, $\psi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ and $q_h \in \mathbb{P}^k$, $n = 1, \ldots, N$. In view of (4.7), $(u_h^\alpha, v_h^\gamma, r_h^\beta) \in (H_0^1 \cap L^\infty) \times L^2 \times \mathbb{R}$ is well defined. Therefore, (4.9a)–(4.9c) is a linear system of equations for $(e_h^u, e_h^v, e_h^r) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$. If $\rho[\alpha_h] = 1$, $\rho[\gamma_h] = 1$ and $\rho[\beta_h] = 1$ then (4.9a)–(4.9c) reduce to (4.1a)–(4.1d).

We denote by $M : X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h} \to X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ the mapping from $(\alpha_h, \gamma_h, \beta_h) \to (e_h^u, e_h^v, e_h^r)$, and define the set

$$\mathcal{B} = \big\{ (\alpha_h, \gamma_h, \beta_h) \in X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h} : \exists \theta \in [0,1] \text{ s.t. } (\alpha_h, \gamma_h, \beta_h) = \theta M(\alpha_h, \gamma_h, \beta_h) \big\}. \tag{4.10}$$

By using similar techniques in the proof of [37, Lemma 4.2], it is straightforward to show the following Lemma (for which the proof is omitted here).

Lemma 6 *The mapping $M : X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h} \to X_{\tau,h} \times X_{\tau,h} \times Y_{\tau,h}$ is well defined, continuous and compact.*

The key to the proof of Theorem 4 is the following technical lemma.

Lemma 7 *Let $d \in \{1, 2, 3\}$ and assume that the solution of NLSW equation (1.1) is sufficiently smooth. Let $(\alpha_h, \gamma_h, \beta_h) \in \mathcal{B}$ and $(e_h^u, e_h^v, e_h^r) = M(\alpha_h, \gamma_h, \beta_h)$, then there exist positive constants τ_0 and h_0 such that when $\tau \le \tau_0$ and $h \le h_0$, the following estimates hold:*

$$\|e_h^u\|_{L^\infty(0,T;H^1)} + \|e_h^v\|_{L^\infty(0,T;L^2)} + \|e_h^r\|_{L^\infty(0,T)} \tag{4.11}$$

$$\le C\Big[\|e_h^u(0)\|_{H^1}+\|e_h^v(0)\|_{L^2}+|e_h^r(0)|+\max_{1\le n\le N}\max_{t\in I_n}\left(\|d_u^n\|_{H^1}+\|d_v^n\|_{L^2}+|P_\tau^n d_r^n|\right)\Big],$$

$$\max_{1\le n\le N}\max_{1\le j\le k}\|e_h^u(t_{nj})\|_{L^\infty\cap H^1}\le\frac{1}{3},\quad \max_{1\le n\le N}\max_{1\le j\le k}\|e_h^v(t_{nj})\|\le\frac{1}{3},$$

$$\max_{1\le n\le N}\max_{1\le j\le k}|e_h^r(t_{nj})|\le\frac{1}{3}, \tag{4.12}$$

$$\rho[\alpha_h]=1,\quad \rho[\gamma_h]=1\ \ \textit{and}\ \ \rho[\beta_h]=1. \tag{4.13}$$

In this section, we prove Theorem 4 by using Lemma 7. The proof of Lemma 7 is presented in the next section.

Proof of Theorem 4

Existence. By the definition of $\mathcal{B}$, if $(\alpha_h,\gamma_h,\beta_h)\in\mathcal{B}$ and $(e_h^u,e_h^v,e_h^r)=M(\alpha_h,\gamma_h,\beta_h)$, then

$$(\alpha_h,\gamma_h,\beta_h)=\theta M(\alpha_h,\gamma_h,\beta_h)=\theta(e_h^u,e_h^v,e_h^r),$$

thus, $\alpha_h=\theta e_h^u$, $\gamma_h=\theta e_h^v$ and $\beta_h=\theta e_h^r$ for some $0\le\theta\le 1$. Note that the estimate (4.11), we have

$$\|(\alpha_h,\gamma_h,\beta_h)\|_{X_{\tau,h}\times X_{\tau,h}\times Y_{\tau,h}}=\|\alpha_h\|_{L^\infty(0,T;H^1)}+\|\gamma_h\|_{L^\infty(0,T;L^2)}+\|\beta_h\|_{L^\infty(0,T)}\le C. \tag{4.14}$$

By using Schaefer's fixed point theorem, there exists a fixed point $(\alpha_h,\gamma_h,\beta_h)$ for the mapping M (corresponding to $\theta=1$ in $\mathcal{B}$) such that $(e_h^u,e_h^v,e_h^r)=(\alpha_h,\gamma_h,\beta_h)$ satisfying (4.9a)–(4.9c) with

$$u_h^\alpha=u_h^*+\alpha_h,\quad v_h^\gamma=v_h^*+\gamma_h\ \text{ and }\ r_h^\beta=r_h^*+\beta_h,$$

where we have used (4.13) in the expression (4.8). Then (4.13) guarantees that (e_h^u,e_h^v,e_h^r) is a solution of (4.1a)–(4.1d) with

$$(u_h,v_h,r_h)=(u_h^\alpha,v_h^\gamma,r_h^\beta)=(u_h^*+e_h^u,v_h^*+e_h^v,r_h^*+e_h^r).$$

Accordingly, (u_h,v_h,r_h) is a solution of the numerical scheme (3.6) as discussed in Remark 4. In addition, (4.12) implies $(u_h,v_h,r_h)\in X_{\tau,h}^*\times\widetilde{X}_{\tau,h}^*\times Y_{\tau,h}^*$ defined in (4.2), which shows the existence of a numerical solution in $X_{\tau,h}^*\times\widetilde{X}_{\tau,h}^*\times Y_{\tau,h}^*$.

Uniqueness. Suppose that (u_h,v_h,r_h) and $(\tilde u_h,\tilde v_h,\tilde r_h)$ in $X_{\tau,h}^*\times\widetilde{X}_{\tau,h}^*\times Y_{\tau,h}^*$ are two pairs of numerical solutions, i.e. both (u_h,v_h,r_h) and $(\tilde u_h,\tilde v_h,\tilde r_h)$ satisfy (3.8)–(3.10). Set $\tilde e_h^u=u_h-\tilde u_h$, $\tilde e_h^v=v_h-\tilde v_h$ and $\tilde e_h^r=r_h-\tilde r_h$. Thus, we have $(\tilde e_h^u,\tilde e_h^v,\tilde e_h^r)$ satisfies (4.1a)–(4.1d) with $u_{nj}^*=\tilde u_{nj}$, $r_{nj}^*=\tilde r_{nj}$ and $d_u^n=d_v^n=d_r^n=0$. In addition, the definition in (4.2) gives

$$\|\tilde e_h^u(t_{nj})\|_{L^\infty\cap H^1}\le 1,\quad \|\tilde e_h^v(t_{nj})\|_{L^2}\le 1\text{ and }|\tilde e_h^r(t_{nj})|\le 1,\qquad 1\le j\le k,\ 1\le n\le N.$$

Accordingly, $(\tilde e_h^u,\tilde e_h^v,\tilde e_h^r)$ is a fixed point of the mapping M (corresponding to $\theta=1$

in $\mathcal{B}$) in the case $\tilde{e}_h^u(0) = \tilde{e}_h^v(0) = \tilde{e}_h^r(0) = 0$ and $d_u^n = d_v^n = d_r^n = 0$. Hence, by using (4.11), we obtain

$$
\begin{aligned}
&\|\tilde{e}_h^u\|_{L^\infty(0,T;H^1)} + \|\tilde{e}_h^v\|_{L^\infty(0,T;L^2)} + \|\tilde{e}_h^r\|_{L^\infty(0,T)} \\
&\le C\Big(\|\tilde{e}_h^u(0)\|_{H^1} + \|\tilde{e}_h^v(0)\|_{L^2} + |\tilde{e}_h^r(0)| + \max_{1\le n\le N}\max_{t\in I_n}\big(\|d_u^n\|_{H^1} + \|d_v^n\|_{L^2} + |P_\tau^n d_r^n|\big)\Big) = 0.
\end{aligned} \tag{4.15}
$$

Error estimates. Since the error function $(e_h^u, e_h^v, e_h^r) = (u_h, v_h, r_h) - (u_h^*, v_h^*, r_h^*)$ satisfies (4.1a)–(4.1d) and (4.12), it follows that (e_h^u, e_h^v, e_h^r) is a fixed point of the mapping M (corresponding to $\theta = 1$ in $\mathcal{B}$). Thus, (4.11) yields

$$
\begin{aligned}
&\|e_h^u\|_{L^\infty(0,T;H^1)} + \|e_h^v\|_{L^\infty(0,T;L^2)} + |e_h^r|_{L^\infty(0,T)} \\
&\le C\Big(\|e_h^u(0)\|_{H^1} + \|e_h^v(0)\|_{L^2} + |e_h^r(0)| + \max_{1\le n\le N}\max_{t\in I_n}\big(\|d_u^n\|_{H^1} + \|d_v^n\|_{L^2} + |P_\tau^n d_r^n|\big)\Big),
\end{aligned} \tag{4.16}
$$

which together with the consistency error estimates in Theorem 3 yields the desired error estimate (4.3). The proof is complete.

5 Proof of Lemma 7

This section is devoted to the proof of Lemma 7, which is used to prove Theorem 4 in the last section. Note that $(\alpha_h, \gamma_h, \beta_h) \in \mathcal{B}$ and $(e_h^u, e_h^v, e_h^r) = M(\alpha_h, \gamma_h, \beta_h)$, it follows that

$$
(\alpha_h, \gamma_h, \beta_h) = \theta M(\alpha_h, \gamma_h, \beta_h) = (\theta e_h^u, \theta e_h^v, \theta e_h^r),
$$

which implies $\alpha_h = \theta e_h^u$, $\gamma_h = \theta e_h^v$ and $\beta_h = \theta e_h^r$. Then, (4.9a)–(4.9c) can be rewritten as

$$
\int_{I_n} (\partial_t e_h^u, \varphi_h)\,\mathrm{d}t = \theta \int_{I_n} (P_\tau^n e_h^v, \varphi_h)\,\mathrm{d}t - \int_{I_n} (P_\tau^n d_u^n, \varphi_h)\,\mathrm{d}t, \tag{5.1a}
$$

$$
\begin{aligned}
&\int_{I_n} (\partial_t e_h^v, \psi_h)\,\mathrm{d}t + \mathrm{i}\int_{I_n} (P_\tau^n e_h^v, \psi_h)\,\mathrm{d}t + \theta\int_{I_n} (\nabla P_\tau^n e_h^u, \nabla\psi_h)\,\mathrm{d}t \\
&\quad = \frac{\tau}{2}\sum_{j=1}^k w_j\Big(r_{nj}^*\big[g(u_{nj}^\alpha)u_{nj}^\alpha - g(u_{nj}^*)u_{nj}^*\big], \psi_{nj}\Big) \\
&\qquad + \frac{\theta\tau}{2}\sum_{j=1}^k w_j\big(e_{nj}^r g(u_{nj}^\alpha)u_{nj}^\alpha, \psi_{nj}\big) - \int_{I_n} (P_\tau^n d_v^n, \psi_h)\,\mathrm{d}t,
\end{aligned} \tag{5.1b}
$$

$$
\begin{aligned}
\int_{I_n} \partial_t e_h^r q_h\,\mathrm{d}t &= \frac{\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\Big(q_{nj}\big[g(u_{nj}^\alpha)u_{nj}^\alpha - g(u_{nj}^*)u_{nj}^*\big], v_{nj}^*\Big) \\
&\quad + \frac{\theta\tau}{4}\sum_{j=1}^k w_j \mathrm{Re}\big(q_{nj} g(u_{nj}^\alpha)u_{nj}^\alpha, e_{nj}^v\big) - \int_{I_n} P_\tau^n d_r^n q_h\,\mathrm{d}t,
\end{aligned} \tag{5.1c}
$$

which hold for all $\varphi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$, $\psi_h \in \mathbb{P}^k \otimes \mathcal{S}_h$ and $q_h \in \mathbb{P}^k$, $n = 1, \ldots, N$.

Inequality (4.7) and definition (4.8) imply that

$$\begin{aligned}
&\max_{1\le n\le N}\max_{1\le j\le k}\left(\|u^{\alpha}_{nj}\|_{H^1}+\|v^{\gamma}_{nj}\|_{L^2}+|r^{\beta}_{nj}|\right)\\
&\le \max_{1\le n\le N}\max_{1\le j\le k}\left(\|u^{*}_{nj}\|_{H^1}+\|v^{*}_{nj}\|_{L^2}+|r^{*}_{nj}|\right)\\
&\quad+\max_{1\le n\le N}\max_{1\le j\le k}\left(\|\rho[\alpha_h]\alpha_h(t_{nj})\|_{H^1}+\|\rho[\gamma_h]\gamma_h(t_{nj})\|_{L^2}+|\rho[\beta_h]\beta_h(t_{nj})|\right)\\
&\le \|u^*_h\|_{L^\infty(0,T;H^1)}+\|v^*_h\|_{L^\infty(0,T;L^2)}+\|r^*_h\|_{L^\infty(0,T)}+3.
\end{aligned} \tag{5.2}$$

Thus $\|u^{\alpha}_{nj}\|_{H^1}$, $\|v^{\gamma}_{nj}\|$ and $|r^{\beta}_{nj}|$ are bounded uniformly with respect to τ and h.

The rest of the proof is divided into four parts.

Part 1: Estimation of $\|e^u_h\|_{L^2(I_n;H^1)}$ *and* $\|e^v_h\|_{L^2(I_n;L^2)}$. Note that

$$\begin{aligned}
\int_{I_n}\|\nabla P^n_\tau e^u_h\|^2\,\mathrm{d}t &= \int_{I_n}\left(\nabla e^u_h(t),\nabla P^n_\tau e^u_h(t)\right)\mathrm{d}t\\
&=\int_{I_n}\Big[\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)+\int_{t_{n-1}}^{t}\partial_s\left(\nabla e^u_h(s),\nabla P^n_\tau e^u_h(s)\right)\mathrm{d}s\Big]\,\mathrm{d}t\\
&=\tau\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)+\int_{I_n}\partial_s\left(\nabla e^u_h(s),\nabla P^n_\tau e^u_h(s)\right)(t_n-s)\,\mathrm{d}s\\
&=\tau\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)+\mathrm{Re}\int_{I_n}\Big(\partial_t\nabla e^u_h(t),\nabla P^n_\tau e^u_h(t)(t_n-t)\Big)\,\mathrm{d}t\\
&\quad+\mathrm{Re}\int_{I_n}\Big(\nabla e^u_h(t),(t_n-t)\partial_t\nabla P^n_\tau e^u_h(t)\Big)\,\mathrm{d}t\\
&=\tau\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)+\mathrm{Re}\int_{I_n}\Big(\partial_t\nabla e^u_h(t),P^n_\tau\big[\nabla P^n_\tau e^u_h(t)(t_n-t)\big]\Big)\,\mathrm{d}t\\
&\quad+\mathrm{Re}\int_{I_n}\Big(\nabla P^n_\tau e^u_h(t),(t_n-t)\partial_t\nabla P^n_\tau e^u_h(t)\Big)\,\mathrm{d}t\\
&=:\tau\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)+J^u_1+J^u_2.
\end{aligned} \tag{5.3}$$

Using integration by parts, we have

$$J^u_2=\int_{I_n}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla P^n_\tau e^u_h(t)\|^2(t_n-t)\,\mathrm{d}t=-\frac{1}{2}\|\nabla P^n_\tau e^u_h(t_{n-1})\|^2\tau+\int_{I_n}\frac{1}{2}\|\nabla P^n_\tau e^u_h\|^2\,\mathrm{d}t,$$

which together with (5.3) yields

$$\frac{1}{2}\int_{I_n}\|\nabla P^n_\tau e^u_h\|^2\,\mathrm{d}t\le\tau|\left(\nabla e^u_h(t_{n-1}),\nabla P^n_\tau e^u_h(t_{n-1})\right)|+J^u_1. \tag{5.4}$$

Similarly, we can estimate $\int_{I_n}\|P_\tau^n e_h^u\|^2\,\mathrm{d}t$ to get

$$\frac{1}{2}\int_{I_n}\|P_\tau^n e_h^u\|^2\,\mathrm{d}t \le \tau\big|\big(e_h^u(t_{n-1}),\,P_\tau^n e_h^u(t_{n-1})\big)\big| + J_3^u, \tag{5.5}$$

where $J_3^u = \mathrm{Re}\int_{I_n}\big(\partial_t e_h^u(t),\,P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]\big)\,\mathrm{d}t$. Thus, from (5.4)–(5.5) we have

$$\begin{aligned}
&\frac{1}{2}\int_{I_n}\|P_\tau^n e_h^u\|_{H^1}^2\,\mathrm{d}t \\
&\quad\le \tau\big|\big(e_h^u(t_{n-1}),\,P_\tau^n e_h^u(t_{n-1})\big)\big| + \tau\big|\big(\nabla e_h^u(t_{n-1}),\,\nabla P_\tau^n e_h^u(t_{n-1})\big)\big| + J_1^u + J_3^u.
\end{aligned} \tag{5.6}$$

By choosing $\varphi_h = (-\Delta_h)P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]$ in (5.1a) and taking the real part, we have

$$\begin{aligned}
J_1^u &= \mathrm{Re}\int_{I_n}\Big(\partial_t e_h^u(t),\,(-\Delta_h)P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t \\
&= \theta\mathrm{Re}\int_{I_n}\Big(\nabla P_\tau^n e_h^v(t),\,P_\tau^n\big[\nabla P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t \\
&\quad - \mathrm{Re}\int_{I_n}\Big(\nabla P_\tau^n d_u^n,\,P_\tau^n\big[\nabla P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t =: J_{11}^u + J_{12}^u.
\end{aligned} \tag{5.7}$$

To estimate J_{11}^u, we set $\psi_h = P_\tau^n\big[P_\tau^n e_h^v(t)(t_n-t)\big]$ in (5.1b) and take the real part to get

$$\begin{aligned}
J_{11}^u &= \theta\mathrm{Re}\int_{I_n}\Big(\nabla P_\tau^n e_h^u(t),\,P_\tau^n\big[\nabla P_\tau^n e_h^v(t)(t_n-t)\big]\Big)\,\mathrm{d}t \\
&= -\mathrm{Re}\int_{I_n}\Big(\partial_t e_h^v(t),\,P_\tau^n\big[P_\tau^n e_h^v(t)(t_n-t)\big]\Big)\,\mathrm{d}t \\
&\quad - \mathrm{Re}\int_{I_n}\mathrm{i}\Big(P_\tau^n e_h^v(t),\,P_\tau^n\big[P_\tau^n e_h^v(t)(t_n-t)\big]\Big)\,\mathrm{d}t \\
&\quad + \frac{\theta\tau}{2}\sum_{j=1}^k w_j\mathrm{Re}\Big(e_{nj}^r g(u_{nj}^\alpha)u_{nj}^\alpha,\,P_\tau^n\big[P_\tau^n e_{nj}^v(t_n-t_{nj})\big]\Big) \\
&\quad + \frac{\tau}{2}\sum_{j=1}^k w_j\mathrm{Re}\Big(r_{nj}^*\big[g(u_{nj}^\alpha)u_{nj}^\alpha - g(u_{nj}^*)u_{nj}^*\big],\,P_\tau^n\big[P_\tau^n e_{nj}^v(t_n-t_{nj})\big]\Big) \\
&\quad - \mathrm{Re}\int_{I_n}\Big(P_\tau^n d_v^n,\,P_\tau^n\big[P_\tau^n e_h^v(t)(t_n-t)\big]\Big)\,\mathrm{d}t =: \sum_{m=1}^5 J_{11m}^u.
\end{aligned} \tag{5.8}$$

Using integration by parts, we have

$$J^u_{111} = \tau\big(e^v_h(t_{n-1}), P^n_\tau e^v_h(t_{n-1})\big) + \mathrm{Re}\int_{I_n}\Big(P^n_\tau e^v_h(t), \partial_t\big(P^n_\tau e^v_h(t)(t_n - t)\big)\Big)\,\mathrm{d}t \quad (5.9)$$
$$\le \tau\big|\big(e^v_h(t_{n-1}), P^n_\tau e^v_h(t_{n-1})\big)\big| - \frac{1}{2}\int_{I_n}\|P^n_\tau e^v_h\|^2\,\mathrm{d}t - \frac{1}{2}\tau\|P^n_\tau e^v_h(t_{n-1})\|^2.$$

Note that

$$J^u_{112} = \mathrm{Im}\int_{I_n}\Big(P^n_\tau e^v_h(t), P^n_\tau\big[P^n_\tau e^v_h(t)(t_n - t)\big]\Big)\,\mathrm{d}t = \mathrm{Im}\int_{I_n}\|P^n_\tau e^v_h(t)\|^2(t_n - t)\,\mathrm{d}t = 0. \quad (5.10)$$

Substituting (5.7)–(5.10) into (5.6) and using temporal inverse inequality, we obtain

$$\frac{1}{2}\int_{I_n}\|P^n_\tau e^u_h\|^2_{H^1}\,\mathrm{d}t + \frac{1}{2}\int_{I_n}\|P^n_\tau e^v_h\|^2\,\mathrm{d}t \quad (5.11)$$
$$\le \tau\big|\big(\nabla e^u_h(t_{n-1}), \nabla P^n_\tau e^u_h(t_{n-1})\big)\big| + \tau\big|\big(e^u_h(t_{n-1}), P^n_\tau e^u_h(t_{n-1})\big)\big|$$
$$+ \tau\big|\big(e^v_h(t_{n-1}), P^n_\tau e^v_h(t_{n-1})\big)\big| + \sum_{m=3}^{5} J^u_{11m} + J^u_{12} + J^u_3$$
$$\le C\tau\|e^u_h(t_{n-1})\|_{H^1}\Big(\frac{1}{\tau}\int_{I_n}\|P^n_\tau e^u_h\|^2_{H^1}\,\mathrm{d}t\Big)^{\frac{1}{2}} + C\tau\|e^v_h(t_{n-1})\|\Big(\frac{1}{\tau}\int_{I_n}\|P^n_\tau e^v_h\|^2\,\mathrm{d}t\Big)^{\frac{1}{2}}$$
$$+ \sum_{m=3}^{5} J^u_{11m} + J^u_{12} + J^u_3,$$

which implies

$$\int_{I_n}\|P^n_\tau e^u_h\|^2_{H^1}\,\mathrm{d}t + \int_{I_n}\|P^n_\tau e^v_h\|^2\,\mathrm{d}t \quad (5.12)$$
$$\le C\tau\|e^u_h(t_{n-1})\|^2_{H^1} + C\tau\|e^v_h(t_{n-1})\|^2 + 4\sum_{m=3}^{5} J^u_{11m} + 4J^u_{12} + 4J^u_3.$$

Now we bound the remaining terms on the right-hand side of (5.12) separately. By using identity (2.14) and inequality (2.17), we obtain

$$J^u_{113} \le \frac{\tau}{2}\sum_{j=1}^{k} w_j|e^r_{nj}|\big\|g(u^\alpha_{nj})u^\alpha_{nj}\big\|\,\big\|P^n_\tau\big[P^n_\tau e^v_{nj}(t_n - t_{nj})\big]\big\|$$
$$\le C\tau\sum_{j=1}^{k} w_j|e^r_{nj}|\big\|P^n_\tau\big[P^n_\tau e^v_{nj}(t_n - t_{nj})\big]\big\| \quad \text{(here (5.2) is used)}$$

$$
\begin{aligned}
&\le \frac{1}{16\tau^2}\int_{I_n} \big\| P_\tau^n\big[P_\tau^n e_h^v(t)(t_n-t)\big]\big\|^2\,\mathrm{d}t + C\tau^2\int_{I_n} |e_h^r|^2\,\mathrm{d}t\\
&\le \frac{1}{16}\int_{I_n} \|P_\tau^n e_h^v\|^2\,\mathrm{d}t + C\tau^2\int_{I_n} |e_h^r|^2\,\mathrm{d}t,\\
J_{114}^u &\le \frac{\tau}{2}\sum_{j=1}^k w_j|r_{nj}^*|\big\| P_\tau^n\big[g(P_\tau^n u_{nj}^\alpha)P_\tau^n u_{nj}^\alpha - g(P_\tau^n u_{nj}^*)P_\tau^n u_{nj}^*\big]\big\|\\
&\quad \times\big\| P_\tau^n\big[P_\tau^n e_{nj}^v(t_n-t_{nj})\big]\big\|\\
&\le C\tau\sum_{j=1}^k w_j\big\|P_\tau^n e_{nj}^u\big\|\big\| P_\tau^n\big[P_\tau^n e_{nj}^v(t_n-t_{nj})\big]\big\| \quad \text{(here (5.2) is used)}\\
&\le C\tau\int_{I_n} \|P_\tau^n e_h^u\|^2\,\mathrm{d}t + C\tau\int_{I_n}\|P_\tau^n e_h^v\|^2\,\mathrm{d}t,\\
J_{115}^u &\le C\tau^2\int_{I_n}\|P_\tau^n d_v^n\|^2\,\mathrm{d}t + \frac{1}{16}\int_{I_n}\|P_\tau^n e_h^v\|^2\,\mathrm{d}t \le C\tau^3\max_{t\in I_n}\|d_v^n\|^2\\
&\quad +\frac{1}{16}\int_{I_n}\|P_\tau^n e_h^v\|^2\,\mathrm{d}t,
\end{aligned}
$$

and by using (2.3b), we have

$$
\begin{aligned}
J_{12}^u &= \mathrm{Re}\int_{I_n}\Big(\nabla P_h P_\tau^n d_u^n,\, P_\tau^n\big[\nabla P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t\\
&\le \frac{1}{16}\int_{I_n}\|\nabla P_\tau^n e_h^u\|^2\,\mathrm{d}t + C\tau^3\max_{t\in I_n}\|d_u^n\|_{H^1}^2.
\end{aligned}
$$

Setting $\varphi_h = P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]$ in (5.1a) and taking the real part, we obtain

$$
\begin{aligned}
J_3^u &= \theta\mathrm{Re}\int_{I_n}\Big(P_\tau^n e_h^v(t),\, P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t\\
&\quad -\mathrm{Re}\int_{I_n}\Big(P_\tau^n d_u^n,\, P_\tau^n\big[P_\tau^n e_h^u(t)(t_n-t)\big]\Big)\,\mathrm{d}t\\
&\le C\tau^2\int_{I_n}\|P_\tau^n e_h^v\|^2\,\mathrm{d}t + \frac{1}{16}\int_{I_n}\|P_\tau^n e_h^u\|^2\,\mathrm{d}t + C\tau^3\max_{t\in I_n}\|d_u^n\|_{H^1}^2. \qquad (5.13)
\end{aligned}
$$

Substituting the estimates of J_{11m}^u, $m=3,4,5$, J_{12}^u and J_3^u into (5.12), for sufficiently small step size τ, we have

$$
\begin{aligned}
&\int_{I_n}\|P_\tau^n e_h^u\|_{H^1}^2\,\mathrm{d}t + \int_{I_n}\|P_\tau^n e_h^v\|^2\,\mathrm{d}t \qquad (5.14)\\
&\le C\tau\Big(\|e_h^u(t_{n-1})\|_{H^1}^2+\|e_h^v(t_{n-1})\|^2\Big)+C\tau^2\int_{I_n}|e_h^r|^2\,\mathrm{d}t+C\tau^3\max_{t\in I_n}\big(\|d_u^n\|_{H^1}^2+\|d_v^n\|^2\big).
\end{aligned}
$$

By using inequality (2.16), we can remove the operator P_τ^n in the above inequality, i.e.

$$\int_{I_n} \|e_h^u\|_{H^1}^2 \,\mathrm{d}t + \int_{I_n} \|e_h^v\|^2 \,\mathrm{d}t \tag{5.15}$$
$$\le C\tau\big(\|e_h^u(t_{n-1})\|_{H^1}^2 + \|e_h^v(t_{n-1})\|^2\big) + C\tau^2 \int_{I_n} |e_h^r|^2 \,\mathrm{d}t + C\tau^3 \max_{t\in I_n} \big(\|d_u^n\|_{H^1}^2 + \|d_v^n\|^2\big).$$

Part 2: Estimation of $\|e_h^r\|_{L^2(I_n)}$. To estimate the third term on the right-hand side of (5.15), we proceed similarly as (5.3), that is

$$\begin{aligned}\int_{I_n} |P_\tau^n e_h^r|^2 \,\mathrm{d}t &= \tau e_h^r(t_{n-1}) P_\tau^n e_h^r(t_{n-1}) + \int_{I_n} \partial_t e_h^r(t) P_\tau^n \big[P_\tau^n e_h^r(t)(t_n - t)\big] \,\mathrm{d}t \\ &\quad + \int_{I_n} P_\tau^n e_h^r(t)(t_n - t) \partial_t P_\tau^n e_h^r(t) \,\mathrm{d}t \\ &=: \tau e_h^r(t_{n-1}) P_\tau^n e_h^r(t_{n-1}) + J_1^r + J_2^r.\end{aligned} \tag{5.16}$$

By using integration by parts and temporal inverse inequality, we have

$$\int_{I_n} |P_\tau^n e_h^r|^2 \,\mathrm{d}t \le C\tau |e_h^r(t_{n-1})|^2 + 4J_1^r. \tag{5.17}$$

In order to estimate J_1^r, we choose $q_h = P_\tau^n \left[P_\tau^n e_h^r(t)(t_n - t)\right]$ in (5.1c) to derive

$$\begin{aligned} J_1^r &= \frac{\tau}{4} \sum_{j=1}^k w_j \mathrm{Re}\Big(P_\tau^n [P_\tau^n e_{nj}^r (t_n - t_{nj})] \big[g(u_{nj}^\alpha) u_{nj}^\alpha - g(u_{nj}^*) u_{nj}^*\big], v_{nj}^*\Big) \\ &\quad + \frac{\theta\tau}{4} \sum_{j=1}^k w_j \mathrm{Re}\Big(P_\tau^n [P_\tau^n e_{nj}^r (t_n - t_{nj})] g(u_{nj}^\alpha) u_{nj}^\alpha, e_{nj}^v\Big) \\ &\quad - \int_{I_n} P_\tau^n d_r^n P_\tau^n [P_\tau^n e_h^r(t)(t_n - t)] \,\mathrm{d}t =: \sum_{m=1}^3 J_{1m}^r. \end{aligned} \tag{5.18}$$

By using (2.17), we have

$$\begin{aligned} J_{11}^r &\le \frac{\tau}{4} \sum_{j=1}^k w_j \big|P_\tau^n \big[P_\tau^n e_{nj}^r (t_n - t_{nj})\big]\big| \big\|g(u_{nj}^\alpha) u_{nj}^\alpha - g(u_{nj}^*) u_{nj}^*\big\| \, \|v_{nj}^*\| \\ &\le C\tau \sum_{j=1}^k w_j \big|P_\tau^n \big[P_\tau^n e_{nj}^r (t_n - t_{nj})\big]\big| \, \|e_{nj}^u\| \end{aligned}$$

$$
\begin{aligned}
&\le \frac{1}{16\tau^2}\frac{\tau}{2}\sum_{j=1}^{k} w_j |P_\tau^n[P_\tau^n e_{nj}^r (t_n - t_{nj})]|^2 + C\tau^2 \frac{\tau}{2}\sum_{j=1}^{k} w_j \|e_{nj}^u\|^2 \\
&\le \frac{1}{16}\int_{I_n} |P_\tau^n e_h^r|^2 \,\mathrm{d}t + C\tau^2 \int_{I_n} \|e_h^u\|^2 \,\mathrm{d}t,
\end{aligned}
$$

$$
\begin{aligned}
J_{12}^r &\le \frac{\tau}{4}\sum_{j=1}^{k} w_j |P_\tau^n[P_\tau^n e_{nj}^r (t_n - t_{nj})]| \left\| g(u_{nj}^\alpha) u_{nj}^\alpha \right\| \|e_{nj}^v\| \\
&\le \frac{1}{16\tau^2}\frac{\tau}{2}\sum_{j=1}^{k} w_j |P_\tau^n[P_\tau^n e_{nj}^r (t_n - t_{nj})]|^2 + C\tau^2 \frac{\tau}{2}\sum_{j=1}^{k} w_j \|e_{nj}^v\|^2 \\
&\le \frac{1}{16}\int_{I_n} |P_\tau^n e_h^r|^2 \,\mathrm{d}t + C\tau^2 \int_{I_n} \|e_h^v\|^2 \,\mathrm{d}t,
\end{aligned}
$$

and

$$
\begin{aligned}
J_{13}^r &\le C\tau^2 \int_{I_n} |P_\tau^n d_r^n|^2 \,\mathrm{d}t + \frac{1}{16\tau^2}\int_{I_n} |P_\tau^n[P_\tau^n e_h^r(t)(t_n - t)]|^2 \,\mathrm{d}t \\
&\le C\tau^3 \max_{t\in I_n} |P_\tau^n d_r^n|^2 + \frac{1}{16}\int_{I_n} |P_\tau^n e_h^r|^2 \,\mathrm{d}t.
\end{aligned}
$$

Substituting the estimates of $J_{1m}^r, m = 1, 2, 3$ into (5.17)–(5.18), and using (2.16), we get

$$
\int_{I_n} |e_h^r|^2 \,\mathrm{d}t \le C\tau |e_h^r(t_{n-1})|^2 + C\tau^3 \max_{t\in I_n} |P_\tau^n d_r^n|^2 + C\tau^2 \int_{I_n} \left(\|e_h^u\|^2 + \|e_h^v\|^2\right) \mathrm{d}t. \tag{5.19}
$$

Then, combining (5.19) and (5.15), for sufficient small time step τ, we obtain

$$
\begin{aligned}
&\int_{I_n} \|e_h^u\|_{H^1}^2 \,\mathrm{d}t + \int_{I_n} \|e_h^v\|^2 \,\mathrm{d}t + \int_{I_n} |e_h^r|^2 \,\mathrm{d}t \\
&\le C\tau \left(\|e_h^u(t_{n-1})\|_{H^1}^2 + \|e_h^v(t_{n-1})\|^2 + |e_h^r(t_{n-1})|^2\right) \\
&\quad + C\tau^3 \max_{t\in I_n} \left(\|d_u^n\|_{H^1}^2 + \|d_v^n\|^2 + |P_\tau^n d_r^n|^2\right).
\end{aligned} \tag{5.20}
$$

Part 3: Estimation of $\|e_h^u(t_n)\|_{H^1}$, $\|e_h^v(t_n)\|$ *and* $|e_h^r(t_n)|$. By setting $\varphi_h = e_h^u$ and $-\Delta_h e_h^u$ in (5.1a), respectively, taking the real part and summing up the resulting identities, we get

$$
\begin{aligned}
&\frac{1}{2}\|e_h^u(t_n)\|_{H^1}^2 - \frac{1}{2}\|e_h^u(t_{n-1})\|_{H^1}^2 \\
&= \theta \mathrm{Re} \int_{I_n} \left(\nabla P_\tau^n e_h^u(t), \nabla P_\tau^n e_h^v(t)\right) \mathrm{d}t - \mathrm{Re}\int_{I_n} \left(\nabla P_\tau^n d_u^n, \nabla e_h^u(t)\right) \mathrm{d}t \\
&\quad + \theta \mathrm{Re} \int_{I_n} \left(P_\tau^n e_h^u(t), P_\tau^n e_h^v(t)\right) \mathrm{d}t - \mathrm{Re}\int_{I_n} \left(P_\tau^n d_u^n, e_h^u(t)\right) \mathrm{d}t.
\end{aligned} \tag{5.21}
$$

Setting $\psi_h = e_h^v$ in (5.1b) and taking the real part, we have

$$\begin{aligned}
&\frac{1}{2}\|e_h^v(t_n)\|^2 - \frac{1}{2}\|e_h^v(t_{n-1})\|^2 + \theta\mathrm{Re}\int_{I_n}\big(\nabla P_\tau^n e_h^u(t), \nabla P_\tau^n e_h^v(t)\big)\,\mathrm{d}t \\
&\quad = \frac{\tau}{2}\sum_{j=1}^k w_j \mathrm{Re}\Big(r_{nj}^*\big[g(u_{nj}^\alpha)u_{nj}^\alpha - g(u_{nj}^*)u_{nj}^*\big], e_{nj}^v\Big) \\
&\qquad + \frac{\theta\tau}{2}\sum_{j=1}^k w_j \mathrm{Re}\big(e_{nj}^r g(u_{nj}^\alpha)u_{nj}^\alpha, e_{nj}^v\big) - \mathrm{Re}\int_{I_n}\big(P_\tau^n d_v^n, P_\tau^n e_h^v(t)\big)\,\mathrm{d}t.
\end{aligned} \tag{5.22}$$

Summing up the above two identities gives

$$\begin{aligned}
&\frac{1}{2}\|e_h^u(t_n)\|_{H^1}^2 + \frac{1}{2}\|e_h^v(t_n)\|^2 - \frac{1}{2}\|e_h^u(t_{n-1})\|_{H^1}^2 - \frac{1}{2}\|e_h^v(t_{n-1})\|^2 \\
&\quad \le C\tau\sum_{j=1}^k w_j\big(|e_{nj}^r| + \|e_{nj}^u\|\big)\|e_{nj}^v\| + C\int_{I_n}\|d_u^n\|_{H^1}\|e_h^u\|_{H^1}\,\mathrm{d}t \\
&\qquad + C\int_{I_n}\|d_v^n\|\|e_h^v\|\,\mathrm{d}t + C\int_{I_n}\|e_h^u\|\|e_h^v\|\,\mathrm{d}t \\
&\quad \le C\int_{I_n}\big(\|e_h^u\|_{H^1}^2 + \|e_h^v\|^2 + |e_h^r|^2 + \|d_u^n\|_{H^1}^2 + \|d_v^n\|^2\big)\,\mathrm{d}t.
\end{aligned} \tag{5.23}$$

Similarly, setting $q_h = 2e_h^r$ in (5.1c) yields

$$\begin{aligned}
&|e_h^r(t_n)|^2 - |e_h^r(t_{n-1})|^2 \\
&\le C\tau\sum_{j=1}^k w_j|e_{nj}^r|\|e_{nj}^u\| + C\tau\sum_{j=1}^k w_j|e_{nj}^r|\|e_{nj}^v\| + C\int_{I_n}|P_\tau^n d_r^n||e_h^r|\,\mathrm{d}t \\
&\le C\int_{I_n}\big(\|e_h^u\|_{H^1}^2 + \|e_h^v\|^2 + |e_h^r|^2 + |P_\tau^n d_r^n|^2\big)\,\mathrm{d}t.
\end{aligned} \tag{5.24}$$

Part 4: Completion of the proof. Summing up (5.23) and (5.24), and substituting (5.20) into the resulting inequality, we obtain

$$\begin{aligned}
&\big(\|e_h^u(t_n)\|_{H^1}^2 + \|e_h^v(t_n)\|^2 + |e_h^r(t_n)|^2\big) - \big(\|e_h^u(t_{n-1})\|_{H^1}^2 + \|e_h^v(t_{n-1})\|^2 + |e_h^r(t_{n-1})|^2\big) \\
&\le C\tau\big(\|e_h^u(t_{n-1})\|_{H^1}^2 + \|e_h^v(t_{n-1})\|^2 + |e_h^r(t_{n-1})|^2\big) \\
&\quad + C\int_{I_n}\big(\|d_u^n\|_{H^1}^2 + \|d_v^n\|^2 + |P_\tau^n d_r^n|^2\big)\,\mathrm{d}t.
\end{aligned} \tag{5.25}$$

It follows from Gronwall's inequality that

$$\begin{aligned}&\max_{1\le n\le N}\left(\|e_h^u(t_n)\|_{H^1}^2+\|e_h^v(t_n)\|^2+|e_h^r(t_n)|^2\right)\\&\le C\left(\|e_h^u(0)\|_{H^1}^2+\|e_h^v(0)\|^2+|e_h^r(0)|^2\right)+C\sum_{n=1}^N\int_{I_n}\left(\|d_u^n\|_{H^1}^2+\|d_v^n\|^2+|P_\tau^n d_r^n|^2\right)\mathrm{d}t.\end{aligned}\tag{5.26}$$

Then, substituting this inequality into (5.20) and using temporal inverse inequality, we get

$$\begin{aligned}&\max_{t\in[0,T]}\left(\|e_h^u(t)\|_{H^1}^2+\|e_h^v(t)\|^2+|e_h^r(t)|^2\right)\\&\le C\left[\|e_h^u(0)\|_{H^1}^2+\|e_h^v(0)\|^2+|e_h^r(0)|^2+\max_{1\le n\le N}\max_{t\in I_n}\left(\|d_u^n\|_{H^1}^2+\|d_v^n\|^2+|P_\tau^n d_r^n|^2\right)\right],\end{aligned}\tag{5.27}$$

Hence, (4.11) holds.

Similarly as the proof in [37], for finite elements of degree $p\ge 1$ and sufficiently small stepsize τ and mesh size h, inequality (5.27) and error equation (5.1b) imply that

$$\max_{t\in[0,T]}\|e_h^u(t)\|_{L^\infty\cap H^1}\le\frac{1}{3},\quad \max_{t\in[0,T]}\|e_h^v(t)\|_{L^2}\le\frac{1}{3}\ \text{ and }\ \max_{t\in[0,T]}|e_h^r(t)|\le\frac{1}{3},\tag{5.28}$$

which gives (4.12).

Furthermore, it follows from $\alpha_h=\theta e_h^u$, $\gamma_h=\theta e_h^v$ and $\beta_h=\theta e_h^r$ with $0\le\theta\le 1$ that

$$\max_{1\le n\le N}\max_{1\le j\le k}\|\alpha_{nj}\|_{L^\infty\cap H^1}\le\frac{1}{3},\ \max_{1\le n\le N}\max_{1\le j\le k}\|\gamma_{nj}\|_{L^2}\le\frac{1}{3}\text{ and }\max_{1\le n\le N}\max_{1\le j\le k}|\beta_{nj}|\le\frac{1}{3},$$

which imply $\rho[\alpha_h]=\rho[\gamma_h]=\rho[\beta_h]=1$ in view of (4.6). This prove (4.13). The proof of Lemma 7 is complete.

6 Numerical experiments

In this section, we present numerical tests to support the theoretical results proved in Theorems 2 and 4 about the mass and energy conservations, and the convergence of the proposed method. All the computations are performed using the software package FEniCS (https://fenicsproject.org).

6.1 1D example

We consider the NLSW equation on the one-dimensional torus $\Omega = [0, 1]$ (which identifies the two end points of the unit interval)

$$\begin{aligned} &\partial_{tt}u + \mathrm{i}\partial_t u - \partial_{xx}u - 2|u|^2u = 0 && \text{on } [0,1]\times(0,T], \\ &u|_{t=0} = \sin(2\pi x), \quad \partial_t u|_{t=0} = \mathrm{i}\sin(2\pi x) && \text{on } [0,1], \end{aligned} \tag{6.1}$$

subject to periodic boundary condition. Problem (6.1) is solved by the proposed Gauss collocation method in (3.6). Newton's iteration is used to solve the nonlinear algebraic system with a tolerance error of 10^{-8}. Here, the mass and SAV energy of the numerical solution are defined as

$$M_h(t) = \int_\Omega \Big(\frac{1}{2}|u_h(t)|^2 + \mathrm{Im}[v_h(t)\,\overline{u}_h(t)]\Big)\mathrm{d}x \qquad t \geq 0, \tag{6.2}$$

$$E_h(t) = \int_\Omega \frac{1}{2}\Big(|v_h(t)|^2 + |\nabla u_h(t)|^2\Big)\mathrm{d}x - r_h(t)^2 + c_0 \qquad t \geq 0, \tag{6.3}$$

respectively. The evolution of mass and SAV energy of numerical solutions is presented in Fig. 1 with $\tau = 1/10$ and $h = 1/200$. We can observe clearly that the mass and SAV energy is conserved as the time increases, which is consistent with theoretical analysis.

The error $\|u_{h,\mathrm{ref}} - u_h\|_{\ell^\infty(0,T;H^1)}$ is computed by comparing the numerical solution with the reference solution given by sufficiently small stepsize and mesh size. The temporal discretization errors are presented in Table 1, where we have used finite elements of degree 3, and the reference solution $u_{h,\mathrm{ref}}$ is chosen to be the numerical solution with $\tau = 1/420$ and $h = 1/1024$. The spatial discretization errors are presented in Table 2, where we have chosen $k = 2$ and $u_{h,\mathrm{ref}}$ is computed with $\tau = 1/210$ and $h = 1/1024$. From Table 1 and 2, we see that the temporal and spatial convergence rates are $O(\tau^{k+1})$ and $O(h^p)$, respectively. This is consistent with the theoretical result proved in Theorem 4.

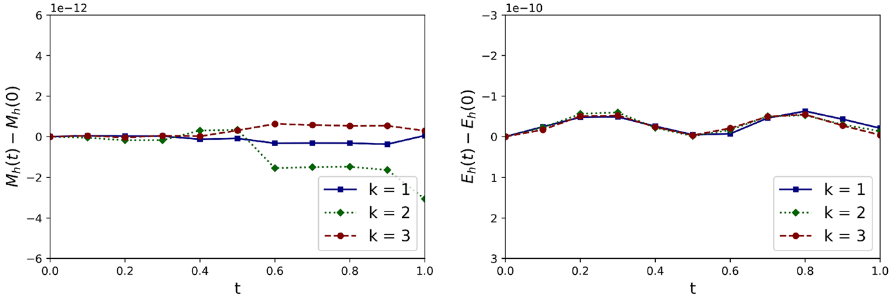

Fig. 1 Evolution of $M_h(t) - M_h(0)$ and $E_h(t) - E_h(0)$ in 1D, with $p = 3$, $h = 1/200$, and $\tau = 1/10$

Table 1 Time discretization errors of the proposed method, with $p = 3$, $h = 1/1024$ and $T = 2$

$k = 1$			$k = 2$			$k = 3$		
N	Error	Rate	N	Error	Rate	N	Error	Rate
80	9.213E–02	–	200	1.070E–05	–	200	6.253E–08	–
100	5.900E–02	2.00	240	6.044E–06	3.13	240	3.002E–08	4.02
120	4.082E–02	2.02	280	3.741E–06	3.11	280	1.603E–08	4.07
140	2.987E–02	2.02	320	2.498E–06	3.03	320	9.493E–09	3.92

6.2 2D example

Next, we consider the following nonlinear Schrödinger equation

$$\begin{aligned} &\partial_{tt}u + \mathrm{i}\partial_t u - \partial_{xx}u - \partial_{yy}u - 2|u|^2u = b && \text{on } \Omega \times (0, T], \\ &u|_{t=0} = \sin(\pi x)\sin(\pi y), \quad \partial_t u|_{t=0} = \mathrm{i}2\sin(\pi x)\sin(\pi y) && \text{on } \Omega, \end{aligned} \tag{6.4}$$

on the two-dimensional space $\Omega = [0, 1] \times [0, 1]$, subject to periodic boundary condition, and

$$b(x, y, t) = (2\pi^2 - 6)\sin(\pi x)\sin(\pi y)\exp(\mathrm{i}2t) - 2\sin^3(\pi x)\sin^3(\pi y)\exp(\mathrm{i}2t). \tag{6.5}$$

In this example, the exact solution of problem (6.4) is known and is given by

$$u(x, y, t) = \exp(\mathrm{i}2t)\sin(\pi x)\sin(\pi y). \tag{6.6}$$

Thus, we compare the numerical solutions with the exact solution (6.6) and compute the error $\|u - u_h\|_{\ell^\infty(0,T;H^1)}$. Here, we solve problem (6.4) by the proposed method for $k = 1, 2, 3$. Newton's iteration is used to solve the nonlinear algebraic system with a tolerance error of 10^{-8}.

The time discretization errors are presented in Table 3, where we have used finite elements of degree 3 with a sufficiently spatial mesh $h = 1/256$, so that the error of spatial discretization is negligibly small in observing the temporal convergence rates.

Table 2 Spatial discretization errors of the proposed method, with $k = 2$, $\tau = 1/210$ and $T = 2$

$p = 1$			$p = 2$			$p = 3$		
h	Error	Rate	h	Error	Rate	h	Error	Rate
1/16	5.154E–01	–	1/16	2.532E–02	–	1/16	8.618E–04	–
1/32	2.516E–01	1.03	1/32	6.323E–03	2.00	1/32	1.080E–04	3.00
1/64	1.246E–01	1.01	1/64	1.581E–03	2.00	1/64	1.349E–05	3.00
1/128	6.184E–02	1.01	1/128	3.953E–04	2.00	1/128	1.684E–06	3.00

Table 3 Time discretization errors of the proposed method, with $p = 3$, $h = 1/256$ and $T = 1$

$k = 1$			$k = 2$			$k = 3$		
N	Error	Rate	N	Error	Rate	N	Error	Rate
15	1.046E–02	–	15	4.606E–05	–	6	1.451E–05	–
20	5.901E–03	1.99	20	1.897E–05	3.08	9	2.849E–06	4.02
25	3.780E–03	2.00	25	9.615E–06	3.05	12	9.008E–07	4.00
30	2.627E–03	2.00	30	5.516E–06	3.04	15	3.700E–07	3.99

From Table 3, we see that the temporal convergence rate is $O(\tau^{k+1})$, which is consistent with our theoretical result.

The spatial discretization errors are presented in Table 4, where we have chosen $k = 2$ with a sufficiently small time stepsize $\tau = 1/128$, so that the time discretization error is negligibly small compared to the spatial error. From Table 4, we see that the spatial convergence rate is $O(h^p)$. This is also consistent with our theoretical result.

We further consider the case of $b(x, y, t) = 0$ and present the evolution of mass and SAV energy of the numerical solutions in Fig. 2 with $\tau = 0.1$ and $h = 0.1$. Again, it is shown that the mass and SAV energy is conserved as the time increases.

6.3 Comparison with other existing numerical method

In this example, we consider the following nonlinear Schrödinger equation

$$\begin{aligned} \partial_{tt}u + \mathrm{i}\partial_t u - \Delta u + |u|^2 u &= b && \text{on } \Omega \times (0, T], \\ u &= 0 && \text{on } \partial\Omega \times (0, T], \\ u|_{t=0} = \sin(\pi x)\sin(\pi y), \quad \partial_t u|_{t=0} &= -\mathrm{i}\sqrt{2}\pi \sin(\pi x)\sin(\pi y) && \text{on } \Omega, \end{aligned} \tag{6.7}$$

where $\Omega = [0, 1] \times [0, 1]$ and

$$b(x, y, t) = (\sqrt{2}\pi + \sin^2(\pi x)\sin^2(\pi y))\sin(\pi x)\sin(\pi y)\exp(-\mathrm{i}\sqrt{2}\pi t). \tag{6.8}$$

The exact solution of problem (6.7) is known and is given by

$$u(x, y, t) = \exp(-\mathrm{i}\sqrt{2}\pi t)\sin(\pi x)\sin(\pi y). \tag{6.9}$$

Table 4 Spatial discretization errors of the proposed method, with $k = 2$, $\tau = 1/128$ and $T = 1$

$p = 1$			$p = 2$			$p = 3$		
h	Error	Rate	h	Error	Rate	h	Error	Rate
1/10	3.526E–01	–	1/10	2.153E–02	–	1/10	9.876E–04	–
1/20	1.750E–01	1.01	1/20	5.399E–03	2.00	1/20	1.238E–04	3.00
1/30	1.165E–01	1.00	1/30	2.401E–03	2.00	1/30	3.669E–05	3.00
1/40	8.730E–02	1.00	1/40	1.351E–03	2.00	1/40	1.548E–06	3.00

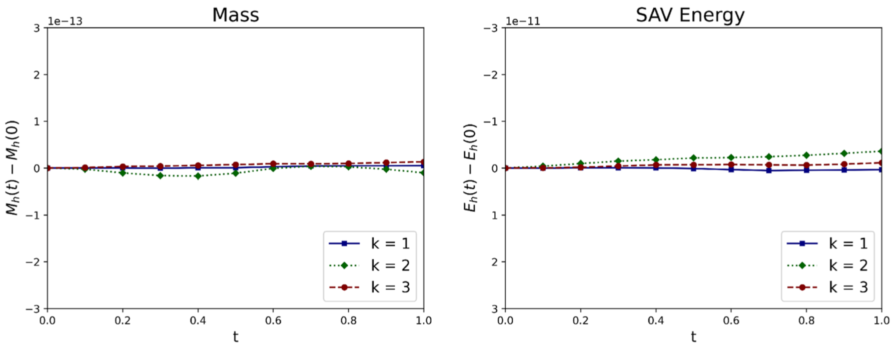

Fig. 2 Evolution of $M_h(t) - M_h(0)$ and $E_h(t) - E_h(0)$ in 2D, with $p = 3$, $h = 1/10$, and $\tau = 1/10$

Here, we compute the error $\|u - u_h\|_{\ell^\infty(0,T;L^2)}$ and compare our numerical scheme (3.6) (with three Gauss points) with the following linearized Leap-frog finite element method proposed in [26]:

$$(\delta_\tau^2 U_h^n, v_h) + (\nabla U_h^{\bar{n}}, \nabla v_h) + \mathrm{i}(D_\tau U_h^n, v_h) + (|U_h^n|^2 U_h^{\bar{n}}, v_h) = (b^n, v_h), \qquad (6.10)$$

with

$$\delta_\tau^2 U_h^n = \frac{U_h^{n+1} - 2U_h^n + U_h^{n-1}}{\tau^2}, \quad D_\tau U_h^n = \frac{U_h^{n+1} - U_h^{n-1}}{2\tau}, \quad U_h^{\bar{n}} = \frac{U_h^{n+1} + U_h^{n-1}}{2}.$$

We compare the errors of our SAV collocation method (3.6) with the errors of scheme (6.10) reported in [26]. Based on the results presented in Table 5, we can see that our method achieves spatial discretization errors comparable to those of scheme (6.10) and maintains the same spatial convergence rate, even when using a larger time step size. The time discretization errors are presented in Table 6, which show that the SAV collocation method (3.6) with three Gauss points exhibits significantly smaller errors compared to those of scheme (6.10), even applied on coarser grids. In this case, the absence of a convergence order for the SAV collocation method (3.6) can be attributed to the dominance of spatial errors.

Table 5 Spatial discretization errors of the proposed method and linearized method

	SAV scheme (3.6) with $\tau = 2^{-6}$		Scheme (6.10) with $\tau = 2^{-14}$	
h	$P1$ element	$P2$ element	$P1$ element	$P2$ element
1/16	1.019E-02	7.056E-05	1.136E-02	6.159E-05
1/32	2.555E-03	8.661E-06	2.843E-03	7.511E-06
1/64	6.392E-04	1.077E-06	7.126E-04	9.321E-07
rate	2.00	3.02	2.00	3.02

Table 6 Temporal discretization errors of the proposed method and linearized method

	SAV scheme (3.6) with $h = 2^{-6}$		Scheme (6.10) with $h = 2^{-9}$	
τ	$P1$ element	$P2$ element	$P1$ element	$P2$ element
1/16	6.392E-04	1.918E-06	2.367E-02	2.365E-02
1/32	6.392E-04	1.085E-06	6.196E-03	6.083E-03
1/64	6.392E-04	1.077E-06	1.572E-04	1.503E-03
rate	*	*	2.00	3.02

7 Conclusions

In this article, we have proposed a Gauss collocation time-stepping method, with finite element method spatial discretization, for the NLSW equation in a d-dimensional torus with $d \in \{1, 2, 3\}$. The scalar auxiliary variable approach is used to preserves the mass and energy conservations simultaneously at the discrete level. We have proved the existence, uniqueness and convergence of the numerical solution to the exact solution with order $O(h^p + \tau^{k+1})$ in the $L^\infty(0, T; H^1)$ norm, where (p, k) is the degree of the space-time finite elements. The proof is presented for the NLSW equation in a d-dimensional torus with $d \in \{1, 2, 3\}$, but it can be naturally extended to other boundary conditions, such as the homogeneous Dirichlet boundary condition.

The proof of (5.28), though omitted in this paper, actually uses the same technique of [37] which controls the L^∞ norm of e_h^u by using the inverse inequality of finite element space plus the L^2 norm of $\Delta_h e_h^u$. This requires the dimension of the space to be smaller than 3. This approach does not require any grid-ratio condition between τ and h. It is worth noting that our analysis method can be similarly extended to higher spatial dimensions. For higher-dimensional problems, we cannot use the L^2 norm of $\Delta_h e_h^u$ to control the L^∞ norm of e_h^u. In this case, we need to control the L^∞ norm of e_h^u by only using the inverse inequality of finite element space. This will lead to a grid-ratio condition $\tau^{k+1} = o(h^{\frac{d}{2}-1})$ and requires the degree p of finite elements to satisfy that $p \geq \frac{d}{2} - 1$.

Appendix: Consistency Analysis

Since the spatial Ritz projection R_h maps $H^1(\Omega)$ into $\mathcal{S}_h \subset H^1(\Omega)$, and the temporal Ritz projection R_τ^n maps $W^{1,\infty}(I_n; H^1(\Omega))$ into $\mathbb{P}^k \otimes H^1(\Omega)$, we can show that every term of d_u^n and d_v^n defined in (3.22a) and (3.22b) is in $C(I_n; H^1(\Omega))$. This implies $d_u^n \in C(I_n; H^1(\Omega))$ and $d_v^n \in C(I_n; H^1(\Omega))$.

By using the triangle inequality, from (3.22a) we have

$$
\begin{aligned}
\max_{t\in I_n}\|d_u^n\|_{H^1} &\leq \max_{t\in I_n}\|\partial_t R_\tau^n(R_h u - u)\|_{H^1} + \max_{t\in I_n}\|v - R_\tau^n v\|_{H^1} + \max_{t\in I_n}\|R_\tau^n v - R_\tau^n R_h v\|_{H^1} \\
&=: D_1^u + D_2^u + D_3^u. \qquad (A.1)
\end{aligned}
$$

Choosing $m = 0$ in (2.21), we obtain the following stability result:

$$\|R_\tau^n u\|_{W^{1,\infty}(I_n;H^s)} \le C\|u\|_{W^{1,\infty}(I_n;H^s)}, \tag{A.2}$$

which together with (2.6) implies

$$D_1^u \le C\|R_h u - u\|_{W^{1,\infty}(I_n;H^1)} \le Ch^p\|u\|_{W^{1,\infty}(I_n;H^{p+1})}.$$

With the help of (2.6) , (2.21) and (2.22), we have

$$\begin{aligned}
&D_2^u \le C\tau^{k+1}\|v\|_{W^{k+1,\infty}(I_n;H^1)} \le C\tau^{k+1}\|u\|_{W^{k+2,\infty}(I_n;H^1)},\\
&D_3^u \le Ch^p \max_{t\in I_n}\|R_\tau^n v\|_{H^{p+1}} \le Ch^p \max_{t\in I_n}\|v\|_{H^{p+1}} \le Ch^p\|u\|_{W^{1,\infty}(I_n;H^{p+1})}.
\end{aligned}$$

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

For the estimate of d_v^n, we use the triangle inequality to derive

$$\begin{aligned}
&\max_{t\in I_n}\|d_v^n\|_{L^2}\\
&\le \max_{t\in I_n}\|\partial_t R_\tau^n(R_h v - v)\|_{L^2} + \max_{t\in I_n}\|v - R_\tau^n R_h v\|_{L^2} + \max_{t\in I_n}\|\Delta_h R_h(u - R_\tau^n u)\|_{L^2}\\
&+ \max_{t\in I_n}\|rg(u)u - I_\tau^n[rg(u)u]\|_{L^2} + \max_{t\in I_n}\|rg(u)u - r_h^* g(u_h^*)u_h^*\|_{L^2}\\
&=: D_1^v + D_2^v + D_3^v + D_4^v + D_5^v.
\end{aligned} \tag{A.3}$$

Similar to the estimate of d_u^n, one can show

$$\begin{aligned}
&D_1^v \le Ch^p\|v\|_{W^{1,\infty}(I_n;H^p)} \le Ch^p\|u\|_{W^{2,\infty}(I_n;H^p)},\\
&D_2^v \le C\tau^{k+1}\|u\|_{W^{k+2,\infty}(I_n;L^2)} + Ch^p\|u\|_{W^{1,\infty}(I_n;H^p)}.
\end{aligned}$$

Using identity (2.5) and inequality (2.21), we have

$$\begin{aligned}
D_3^v &= \max_{t\in I_n}\|\Delta_h R_h(u - R_\tau^n u)\|_{L^2} = \max_{t\in I_n}\|P_h\Delta(u - R_\tau^n u)\|_{L^2}\\
&\le \max_{t\in I_n}\|u - R_\tau^n u\|_{H^2} \le C\tau^{k+1}\|u\|_{W^{k+1,\infty}(I_n,H^2)},\\
D_4^v &= \max_{t\in I_n}\|rg(u)u - I_\tau^n[rg(u)u]\|_{L^2} \le C\tau^{k+1}.
\end{aligned}$$

By using the triangle inequality, we decompose D_6^v into two parts

$$\begin{aligned}
D_5^v &\le \max_{t\in I_n}\|rg(u)u - rg(R_h u)R_h u\|_{L^2} + \max_{t\in I_n}\|rg(R_h u)R_h u\\
&\quad - R_\tau^n rg(R_\tau^n R_h u)R_\tau^n R_h u\|_{L^2}\\
&\le C(h^p + \tau^{k+1}).
\end{aligned}$$

Plugging the estimates of $D_j^v,\ 1 \le j \le 6$ into (A.3) leads to the desired estimate for $\|d_v^n\|$.

Next, we rewrite d_r^n as

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

and test this expression by $P_\tau^n w \in \mathbb{P}^{k-1}$ in the time interval I_n with $w \in L^2(I_n)$. This yields

$$\begin{aligned} \int_{I_n} P_\tau^n d_r^n w \,\mathrm{d}t &= \int_{I_n} d_r^n P_\tau^n w \,\mathrm{d}t \\ &\le \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, v - v_h^*\big) P_\tau^n w \,\mathrm{d}t \\ &\quad + C\tau^{\frac{1}{2}} \big\|\big(g(u)u - g(u_h^*)u_h^*, v_h^*\big)\big\|_{L^\infty(I_n)} \|w\|_{L^2(I_n)} \\ &\quad + C\tau^{k+\frac{3}{2}} \big\|\partial_t^{k+1}\big(g(u_h^*)u_h^*, v_h^*\big)\big\|_{L^\infty(I_n)} \|w\|_{L^2(I_n)} \\ &\le \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, v - v_h^*\big) P_\tau^n w \,\mathrm{d}t + C\tau^{\frac{1}{2}}(h^p + \tau^{k+1})\|w\|_{L^2(I_n)}. \end{aligned} \tag{A.4}$$

The first term on the right-hand side of the above inequality can be estimated as follows

$$\begin{aligned} \frac{1}{2}\mathrm{Re}\int_{I_n} \big(g(u)u, v - v_h^*\big) P_\tau^n w \,\mathrm{d}t &= \int_{I_n} \big(g(u)u, v - R_\tau^n v\big) P_\tau^n w \,\mathrm{d}t \\ &\quad + \int_{I_n} \big(g(u)u, R_\tau^n(v - R_h v)\big) P_\tau^n w \,\mathrm{d}t =: D_1^r + D_2^r, \end{aligned} \tag{A.5}$$

where

$$\begin{aligned} D_1^r &= \int_{I_n} \big(g(u)u P_\tau^n w - P_\tau^n(g(u)u P_\tau^n w), v - R_\tau^n v\big) \,\mathrm{d}t \\ &\le C\tau^{\frac{1}{2}} \|g(u)u P_\tau^n w - P_\tau^n[g(u)u P_\tau^n w]\|_{L^2(I_n;L^2)} \|v - R_\tau^n v\|_{L^\infty(I_n;L^2)} \\ &\le C\tau^{\frac{3}{2}} \|P_\tau^n w\|_{L^2(I_n;L^2)} \|v - R_\tau^n v\|_{L^\infty(I_n;L^2)} \quad \text{(here (2.13) is used)} \\ &\le C\tau^{k+\frac{3}{2}} \|w\|_{L^2(I_n)} \|\partial_t^{k+1} u\|_{L^\infty(I_n;L^2)} \quad \text{(here (2.21) is used),} \\ D_2^r &\le C\tau^{\frac{1}{2}} \|g(u)u\|_{L^\infty(I_n;L^2)} \|v - R_h v\|_{L^\infty(I_n,L^2)} \|w\|_{L^2(I_n)} \\ &\le C\tau^{\frac{1}{2}} h^p \|w\|_{L^2(I_n)} \|u\|_{W^{1,\infty}(I_n;H^{p+1})}. \end{aligned}$$

Substituting these estimates into (A.4), we get

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

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

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

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

Funding The work of the authors was partially supported by grants from National Natural Science Foundation of China (grants no. 12071020, 12131005, and U1930402) and the Research Grants Council of the Hong Kong Special Administrative Region, China. (Project Nos. CityU 11302718, CityU 11300621)

Declarations

Conflicts of interest The authors declare no competing interests.

References

1. Berge, L., Colin, T.: Un problème de perturbation singulière pour une équation d'eveloppe en physique des plasmas. Comptes rendus de l'Academie des sciences. Série 1, Mathématique. **320**(1), 31–34 (1999)
2. Bao, W., Dong, X., Xin, J.: Comparisons between sine-Gordon and perturbed nonlinear Schrödinger equations for modeling light bullets beyond critical collapse. Physica D **239**(13), 1120–1134 (2010)
3. Tsutsumi, M.: Nonrelativistic approximation of nonlinear Klein-Gordon equations in two space dimensions. Nonlinear Analysis: Theory, Methods & Applications. **8**(6), 637–643 (1984)
4. Fei, Z., Pérez-García, V.M., Vázquez, L.: Numerical simulation of nonlinear Schrödinger systems: a new conservative scheme. Appl. Math. Comput. **71**(2–3), 165–177 (1995)
5. Delfour, M., Fortin, M., Payr, G.: Finite-difference solutions of a non-linear Schrödinger equation. J. Comput. Phys. **44**(2), 277–288 (1981)
6. Akrivis, G.D.: Finite difference discretization of the cubic Schrödinger equation. IMA J. Numer. Anal. **13**(1), 115–124 (1993)
7. Akrivis, G.D., Dougalis, V.A., Karakashian, O.A.: On fully discrete Galerkin methods of second-order temporal accuracy for the nonlinear Schrödinger equation. Numer. Math. **59**(1), 31–53 (1991)
8. Antoine, X., Bao, W., Besse, C.: Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations. Computer Physics Communications. 184(12), 2621–2633 (2013)
9. Bao, W., Tang, Q., Xu, Z.: Numerical methods and comparison for computing dark and bright solitons in the nonlinear Schrödinger equation. J. Comput. Phys. **235**, 423–445 (2013)
10. Besse, C.: A relaxation scheme for the nonlinear Schrödinger equation. SIAM J. Numer. Anal. **42**(3), 934–952 (2004)
11. Feng, X., Liu, H., Ma, S.: Mass- and energy-conserved numerical schemes for nonlinear Schrödinger equations. Communications in Computational Physics. **26**(5), 1365–1396 (2019)
12. Gao, Z., Xie, S.: Fourth-order alternating direction implicit compact finite difference schemes for two-dimensional Schrödinger equations. Appl. Numer. Math. **61**(4), 593–614 (2011)
13. Sanz-Serna, J.: Methods for the numerical solution of the nonlinear Schrödinger equation. mathematics of computation. **43**(167), 21–27 (1984)
14. Hong, J., Liu, Y., Munthe-Kaas, H., Zanna, A.: Globally conservative properties and error estimation of a multi-symplectic scheme for Schrödinger equations with variable coefficients. Appl. Numer. Math. **56**(6), 814–843 (2006)
15. Liu, H., Huang, Y., Lu, W., Yi, N.: On accuracy of the mass-preserving dg method to multi-dimensional Schrödinger equations. IMA Journal of Numerical Analysis. 39(2), 760–791 (2019)
16. Wang, J.: A new error analysis of Crank–Nicolson Galerkin FEMs for a generalized nonlinear Schrödinger equation. Journal of Scientific Computing. 60(2), 390–407 (2014)

17. Xu, Y., Shu, C.-W.: Local discontinuous Galerkin methods for nonlinear Schrödinger equations. J. Comput. Phys. **205**(1), 72–97 (2005)
18. Bao, W., Cai, Y.: Uniform error estimates of finite difference methods for the nonlinear Schrödinger equation with wave operator. SIAM J. Numer. Anal. **50**(2), 492–521 (2012)
19. Guo, L., Xu, Y.: Energy conserving local discontinuous Galerkin methods for the nonlinear Schrödinger equation with wave operator. J. Sci. Comput. **65**(2), 622–647 (2015)
20. Hu, H., Chen, Y.: A conservative difference scheme for two-dimensional nonlinear Schrödinger equation with wave operator. Numer. Methods Partial Differential Equations. **32**(3), 862–876 (2016)
21. Wang, T.-c., Zhang, L.-m.: Analysis of some new conservative schemes for nonlinear Schrödinger equation with wave operator. Applied Mathematics and Computation. **182**(2), 1780–1794 (2006)
22. Zhang, L., Chang, Q.: A conservative numerical scheme for a class of nonlinear Schrödinger equation with wave operator. Appl. Math. Comput. **145**(2–3), 603–612 (2003)
23. Li, X., Zhang, L., Wang, S.: A compact finite difference scheme for the nonlinear Schrödinger equation with wave operator. Appl. Math. Comput. **219**(6), 3187–3197 (2012)
24. Li, X., Zhang, L., Zhang, T.: A new numerical scheme for the nonlinear Schrödinger equation with wave operator. J. Appl. Math. Comput. **54**(1), 109–125 (2017)
25. Cheng, X., Wu, F.: Several conservative compact schemes for a class of nonlinear Schrödinger equations with wave operator. Boundary Value Problems. 2018(1), 1–17 (2018)
26. Cai, W., He, D., Pan, K.: A linearized energy–conservative finite element method for the nonlinear Schrödinger equation with wave operator. Applied Numerical Mathematics. 140, 183–198 (2019)
27. Cheng, X., Qin, H., Zhang, J.: Convergence of an energy-conserving scheme for nonlinear space fractional Schrödinger equations with wave operator. J. Comput. Appl. Math. **400**, 113–762 (2022)
28. Cheng, X., Yan, X., Qin, H., Wang, H.: Optimal l^∞ error estimates of the conservative scheme for two-dimensional Schrödinger equations with wave operator. Computers & Mathematics with Applications 100, 74–82 (2021)
29. Shen, J., Xu, J., Yang, J.: The scalar auxiliary variable (SAV) approach for gradient flows. J. Comput. Phys. **353**, 407–416 (2018)
30. Shen, J., Xu, J., Yang, J.: A new class of efficient and robust energy stable schemes for gradient flows. SIAM Rev. **61**(3), 474–506 (2019)
31. Yang, X.: Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends. J. Comput. Phys. **327**, 294–316 (2016)
32. Yang, X., Ju, L.: Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model. Comput. Methods Appl. Mech. Eng. **315**, 691–712 (2017)
33. Yang, X., Ju, L.: Linear and unconditionally energy stable schemes for the binary fluid–surfactant phase field model. Computer Methods in Applied Mechanics and Engineering. 318, 1005–1029 (2017)
34. Yang, X., Zhao, J., Wang, Q., Shen, J.: Numerical approximations for a three component Cahn–Hilliard phase-field model based on the invariant energy quadratization method. Mathematical Models and Methods in Applied Sciences. 27(11), 1993–2030 (2017)
35. Brugnano, L., Zhang, C., Li, D.: A class of energy-conserving Hamiltonian boundary value methods for nonlinear Schrödinger equation with wave operator. Commun. Nonlinear Sci. Numer. Simul. **60**, 33–49 (2018)
36. Li, X., Gong, Y., Zhang, L.: Linear high-order energy-preserving schemes for the nonlinear Schrödinger equation with wave operator using the scalar auxiliary variable approach. J. Sci. Comput. **88**(1), 1–25 (2021)
37. Feng, X., Li, B., Ma, S.: High-order mass-and energy-conserving SAV-Gauss collocation finite element methods for the nonlinear Schroödinger equation. SIAM J. Numer. Anal. **59**(3), 1566–1591 (2021)
38. Evans, L.C.: Partial differential equations. American Mathematical Soc. (2010)
39. Brenner, S.C., Scott, L.R.: The mathematical theory of finite element methods. Springer (2008)
40. Shen, J., Tang, T., Wang, L.-L.: Spectral methods, volume 41 of Springer Series in Computational Mathematics. Springer (2011)
41. Canuto, C., Hussaini, M.Y., Quarteroni, A., Zang, T.A.: Spectral methods. Springer. Fundamentals in single domains (2006)
42. Golub, G.H., Welsch, J.H.: Calculation of Gauss quadrature rules. Mathematics of computation 23(106), 221–230 (1969)
43. Kopriva, D.A.: Implementing spectral methods for partial differential equations: algorithms for scientists and engineers. Springer (2009)

Authors and Affiliations

Shu Ma[1] · Jilu Wang[2] · Mingyan Zhang[3] · Zhimin Zhang[4]

Shu Ma
shuma2@cityu.edu.hk

Jilu Wang
wangjilu03@gmail.com; wangjilu@hit.edu.cn

Zhimin Zhang
ag7761@wayne.edu

1 Department of Mathematics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong

2 School of Science, Harbin Institute of Technology, Shenzhen 518055, China

3 Beijing Computational Science Research Center, Beijing 100193, China

4 Department of Mathematics, Wayne State University, Detroit, MI 48202, USA