

Unconditionally optimal convergence of an energy-conserving and linearly implicit scheme for nonlinear wave equations

Waixiang Cao¹, Dongfang Li^{2,3*} & Zhimin Zhang^{4,5}

¹*School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China;*

²*School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China;*

³*Hubei Key Laboratory of Engineering Modeling and Scientific Computing, Huazhong University of Science and Technology, Wuhan 430074, China;*

⁴*Beijing Computational Science Research Center, Beijing 100193, China;*

⁵*Department of Mathematics, Wayne State University, Detroit, MI 48202, USA*

Email: caowx@bnu.edu.cn, dfl@hust.edu.cn, z Zhang@math.wayne.edu

Received March 30, 2020; accepted March 5, 2021; published online September 6, 2021

Abstract In this paper, we present and analyze an energy-conserving and linearly implicit scheme for solving the nonlinear wave equations. Optimal error estimates in time and superconvergent error estimates in space are established without certain time-step restrictions. The key is to estimate directly the solution bounds in the H^2 -norm for both the nonlinear wave equation and the corresponding fully discrete scheme, while the previous investigations rely on the temporal-spatial error splitting approach. Numerical examples are presented to confirm energy-conserving properties, unconditional convergence and optimal error estimates, respectively, of the proposed fully discrete schemes.

Keywords scalar auxiliary variable, wave equations, stability, error estimate, superconvergence

MSC(2020) 65M60, 65M12, 65N12

Citation: Cao W X, Li D F, Zhang Z M. Unconditionally optimal convergence of an energy-conserving and linearly implicit scheme for nonlinear wave equations. *Sci China Math*, 2022, 65: 1731–1748, <https://doi.org/10.1007/s11425-020-1857-5>

1 Introduction

We present an energy-conserving and linearly implicit scheme as well as the unconditionally optimal error estimates for solving the following wave equation:

$$\begin{aligned} u_{tt} &= \Delta u - \lambda u - F'(u), & (\mathbf{x}, t) &\in \Omega \times (0, T], \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \quad u_t(\mathbf{x}, 0) = u_1(\mathbf{x}), & \mathbf{x} &\in \Omega \end{aligned} \quad (1.1)$$

with the periodic boundary condition, where $\lambda \geq 0$ is a constant, Ω is a polygonal or polyhedral domain in \mathbb{R}^d ($d = 2, 3$), u_0 and u_1 are sufficiently smooth, and $F \in C^2(\mathbb{R})$ is the nonlinear potential. For

* Corresponding author

simplicity, we assume that Ω is a rectangular or cubic domain. Nonlinear wave equations are widely used to model plenty of complicated natural phenomena in a variety of scientific fields [11, 12, 28, 34]. In the past several decades, it has been one of the hot spots in the numerical analysis of different schemes for the equations [3, 8, 15, 33, 35].

There are many papers that consider error analysis of the fully discrete schemes for the nonlinear problems under the following assumption (see, e.g., [7, 9, 10, 16]):

$$|F'(u^n) - F'(u_h^n)| \leq L|u^n - u_h^n|, \quad (1.2)$$

where u^n and u_h^n are, respectively, the theoretical and numerical solutions, and $L > 0$ is the Lipschitz coefficient. A classical model satisfying (1.2) is the sine-Gordon equation, whose nonlinear term is $\sin(u)$. However, as pointed in [15], the assumption (1.2) is not the typical behavior of the general nonlinear wave equations and thus its range of the actual applicability is limited.

In order to deal with the non-Lipschitz nonlinearity, one common way is to impose *a priori* boundedness of the numerical approximations u_h^n . In classical finite element analysis, the numerical solutions in the maximum norm are usually estimated by

$$\begin{aligned} \|u_h^n\|_{L^\infty} &\leq \|R_h u^n\|_{L^\infty} + \|R_h u^n - u_h^n\|_{L^\infty} \\ &\leq \|R_h u^n\|_{L^\infty} + Ch^{-d/2} \|R_h u^n - u_h^n\|_{L^2} \\ &\leq \|R_h u^n\|_{L^\infty} + Ch^{-d/2} (\tau^p + h^{r+1}), \end{aligned} \quad (1.3)$$

where R_h is the projection operator, and $r + 1$ and p are convergence orders in spatial and temporal directions, respectively. Consequently, a time-step restriction $\tau = O(h^{\frac{d}{2p}})$ is needed in (1.3) (see, e.g., [6, 13, 14, 25, 26]). Such a time-step restriction appears widely in the numerical analysis but is not always necessary in actual applications.

Unconditional convergence means that the established error bound is valid without the above-mentioned time-step restriction. To achieve the unconditional convergence, a temporal-spatial error splitting approach is presented recently [18–20, 22, 23]. The key idea of the approach is to introduce a time discrete system, whose solution is denoted by U^n . Then one can obtain the following error estimates:

$$\|U^n\|_{H^2} \leq C \quad \text{and} \quad \|R_h U^n - u_h^n\|_{L^2} \leq Ch^2.$$

The boundedness of the numerical solutions is obtained by

$$\begin{aligned} \|u_h^n\|_{L^\infty} &\leq \|R_h U^n - u_h^n\|_{L^\infty} + \|R_h U^n - \mathcal{I}_h U^n\|_{L^\infty} + \|\mathcal{I}_h U^n\|_{L^\infty} \\ &\leq Ch^{-d/2} (\|R_h U^n - u_h^n\|_{L^2} + \|R_h U^n - \mathcal{I}_h U^n\|_{L^2}) + \|U^n\|_{L^\infty} \\ &\leq Ch^{2-d/2} + \|U^n\|_{H^2}. \end{aligned} \quad (1.4)$$

Here, $\mathcal{I}_h U^n$ denotes the interpolation function of U^n . It implies that the numerical solutions is bounded if the temporal and spatial step-sizes are sufficiently small, respectively. Then the error estimates can be proved by following the usual way. In spite of the interesting and instructive work, additional error estimates in different norms are required in the proof, and so far, most unconditional convergence results are focused on nonlinear parabolic problems.

The nonlinear wave equations (1.1) have several remarkable features. First, the models (1.1) are energy-conserving, i.e.,

$$E(t) = \int_{\Omega} (u_t^2 + |\nabla u|^2 + \lambda|u|^2 + 2F(u)) dx = E(0).$$

Second, the typical nonlinear terms are non-Lipschitz continuous. Third, the solutions have different regularities. A natural question is whether we can develop some effective unconditional convergence numerical schemes for nonlinear wave equations, taking all the remarkable features into account.

In the present paper, we present an energy-conserving and linearly implicit scheme for solving the nonlinear wave equations (1.1). The scheme is of order 2 in the temporal directions and no additional

initial iterations are required. The construction of the scheme is based on the recently-developed scalar auxiliary variable (SAV) approach combined with the finite element methods, classical Crank-Nicolson methods and extrapolation approximation. We show that our fully-discrete schemes conserve the energy and are convergent without certain time-step restrictions. Unlike the previous temporal-spatial error splitting approach, we estimate the solution directly in the following procedure: (1) obtain the bounds in the H^2 -norm of the solutions for both the nonlinear wave equations and the corresponding fully discrete schemes; (2) establish the bound for numerical solutions by applying the embedding inequality; (3) obtain the unconditionally optimal error estimates in time and superconvergent error estimates in space.

We remark that the key to constructing the energy-conserving schemes is the SAV idea, which has been applied successfully to the gradient flows [1, 29–31]. Very recently, the idea was introduced to develop energy-conserving schemes for the conservative laws [4, 5, 21, 24]. However, much attention has been paid to the stability and energy-conserving properties, and no unconditional convergence results of fully discrete SAV schemes for nonlinear wave equations are found in the literature. This is the main motivation and contribution of the present study.

The rest of this paper is organized as follows. In Section 2, we propose a fully discrete scheme for the nonlinear wave equations (1.1). In Section 3, we present a detailed proof to show the energy-conserving properties and unconditional convergence for the temporal discretization. Error estimates for the fully discrete solution is established in Section 4, where we prove that the approximation error is unconditionally optimal in time and superconvergent in space (under the H^1 -norm). In Section 5, we present several numerical examples to confirm the theoretical results. Finally, conclusions are presented in Section 6.

2 The linearly implicit method

In this section, we present a fully discrete numerical scheme, which preserves the discrete energy.

Suppose

$$E_1(u) = \int_{\Omega} F(u) d\mathbf{x} \geq -c_0$$

for some $c_0 > 0$, i.e., it is bounded from below, and let $C_0 > c_0$ so that

$$E_1(u) + C_0 > 0.$$

We introduce the following scalar auxiliary variable (SAV):

$$r(t) = \sqrt{E(u)}, \quad E(u) = \int_{\Omega} F(u) d\mathbf{x} + C_0,$$

and rewrite (1.1) as

$$\begin{aligned} u_t &= v, \\ v_t &= \Delta u - \lambda u - \frac{r}{\sqrt{E(u)}} f(u), \\ r_t &= \frac{1}{2\sqrt{E(u)}} \int_{\Omega} f(u) u_t d\mathbf{x}, \end{aligned} \tag{2.1}$$

where $f(u) = F'(u)$.

Let \mathcal{T}_h be the usual regular triangulation of the polygonal domain Ω . Denote by h_T the mesh size of \mathcal{T}_h , where h_T is the diameter of the element $T \in \mathcal{T}_h$, and $h = \max_{T \in \mathcal{T}_h} h_T$. Let V_h be the classical finite-dimensional subspace of $H^1(\Omega)$, which consists of the usual continuous piecewise polynomials of degree k ($k \geq 1$) on \mathcal{T}_h , i.e.,

$$V_h = \{v \in C^0(\Omega) : v|_T \in \mathbb{P}_k, \forall T \in \mathcal{T}_h\},$$

where \mathbb{P}_k denotes the space of polynomials of degree no more than k .

Let $\tau = \frac{T}{N}$ with N being a given integer and $t_n = n\tau, n = 0, 1, \dots, N$. Define

$$u^n = u(\mathbf{x}, t_n), \quad v^n = v(\mathbf{x}, t_n), \quad r^n = r(t_n).$$

For any sequence of the functions $\{f^n\}_{n=0}^N$, we define for all $n (n = 0, \dots, N - 1)$,

$$D_\tau f^{n+1} := \frac{Df^{n+1}}{\tau} = \frac{f^{n+1} - f^n}{\tau}, \quad \tilde{f}^{n+\frac{1}{2}} := \frac{1}{2}(3f^n - f^{n-1}), \quad \hat{f}^{n+\frac{1}{2}} := \frac{f^{n+1} + f^n}{2}.$$

Note that for $n = 0$, we define $\tilde{f}^{\frac{1}{2}} = f^0$.

To design an energy-conserving and linearly implicit numerical scheme, which is easy to implement and efficient, we consider the following fully discrete Crank-Nicolson Galerkin SAV method: find $u_h^{n+1} \in V_h, v_h^{n+1} \in V_h$ and $r_h^{n+1} \in \mathbb{R}$ for $n = 0, \dots, N - 1$ such that for all $(w_h, \zeta_h) \in V_h \times V_h$,

$$\begin{aligned} (D_\tau u_h^{n+1}, w_h) &= (\hat{v}_h^{n+\frac{1}{2}}, w_h), \\ (D_\tau v_h^{n+1}, \zeta_h) &= -(\nabla \hat{u}_h^{n+\frac{1}{2}}, \nabla \zeta_h) - (\lambda \hat{u}_h^{n+\frac{1}{2}}, \zeta_h) - \left(\frac{\hat{r}_h^{n+\frac{1}{2}}}{\sqrt{E(\hat{u}_h^{n+\frac{1}{2}})}} f(\hat{u}_h^{n+\frac{1}{2}}), \zeta_h \right), \\ r_h^{n+1} - r_h^n &= \frac{1}{2\sqrt{E(\hat{u}_h^{n+\frac{1}{2}})}} \int_\Omega f(\hat{u}_h^{n+\frac{1}{2}})(u_h^{n+1} - u_h^n) d\mathbf{x}, \end{aligned} \tag{2.2}$$

where

$$(u, v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})d\mathbf{x}, \quad f(\hat{u}_h^{\frac{1}{2}}) = f(u_h^0),$$

and initial values are chosen as

$$(u_h^0, v_h^0, r_h^0) = (R_h u_0, R_h u_1, \sqrt{E(u_0)}).$$

Here, $R_h u_0$ is the Ritz projection of u_0 , which will be defined later.

Equivalently, we rewrite the above scheme (2.2) into the following linear form:

$$((4I - \tau^2 \Delta_h + \tau^2 \lambda I)u_h^{n+1}, w_h) + \frac{\tau^2}{2}(u_h^{n+1}, b_1)(b_1, w_h) = (g, w_h) + \frac{\tau^2}{2}(u_h^n, b_1)(b_1, w_h) \tag{2.3}$$

for all $w_h \in V_h$, where

$$(\Delta_h u_h, v_h) := -(\nabla u_h, \nabla v_h)$$

and

$$b_1 = \frac{f(\hat{u}_h^{n+\frac{1}{2}})}{\sqrt{E(\hat{u}_h^{n+\frac{1}{2}})}}, \quad g = (4I + \tau^2 \Delta_h - \tau^2 \lambda I)u_h^n + 4\tau v_h^n - 2\tau^2 r_h^n b_1.$$

Choosing w_h in (2.3) to be the basis function of V_h leads to a linear equation of the form

$$A\mathbf{u}^{n+1} + (\mathbf{u}^{n+1}, \mathbf{b}_1)\mathbf{b}_2 = \mathbf{g}$$

for some matrix A and vectors $\mathbf{b}_1, \mathbf{b}_2$ and \mathbf{g} . By taking the inner product with \mathbf{b}_1 in the above equation, we obtain $(\mathbf{u}^{n+1}, \mathbf{b}_1)$ and then derive \mathbf{u}^{n+1} . Hence, the scheme is easy to implement and very efficient. We also refer to [30, 31] for more detailed information.

3 Unconditional energy preservation and convergence for the temporal discretization

In this section, we prove that the Galerkin SAV method (2.2) preserves the energy unconditionally. Moreover, we establish the convergence analysis of the SAV approach with minimum assumptions.

We begin with the energy preservation property of the Galerkin SAV approach. Define the energy

$$E^n = \sqrt{\frac{1}{2}(\|v_h^n\|^2 + \|\nabla u_h^n\|^2 + \lambda\|u_h^n\|^2) + (r_h^n)^2}, \quad 1 \leq n \leq N.$$

Here, $\|u\|^2 = (u, u) = \|u\|_{L^2}^2$. Taking

$$(w_h, \zeta_h) = (v_h^{n+1} - v_h^n, u_h^{n+1} - u_h^n)$$

and multiplying the third equation of (2.2) by $r_h^{n+1} + r_h^n$, we derive

$$\frac{1}{2}(\|v_h^{n+1}\|^2 - \|v_h^n\|^2 + \|\nabla u_h^{n+1}\|^2 - \|\nabla u_h^n\|^2 + \lambda\|u_h^{n+1}\|^2 - \lambda\|u_h^n\|^2) + (r_h^{n+1})^2 - (r_h^n)^2 = 0.$$

Consequently,

$$E^{n+1} = E^n = E^0, \quad \forall n \geq 1.$$

Now we consider a time-discrete system of the equations

$$\begin{aligned} D_\tau U^{n+1} &= \hat{V}^{n+\frac{1}{2}}, \\ D_\tau V^{n+1} &= \Delta \hat{U}^{n+\frac{1}{2}} - \lambda \hat{U}^{n+\frac{1}{2}} - \frac{\hat{R}^{n+\frac{1}{2}}}{\sqrt{E(\tilde{U}^{n+\frac{1}{2}})}} f(\tilde{U}^{n+\frac{1}{2}}), \\ R^{n+1} - R^n &= \frac{1}{2\sqrt{E(\tilde{U}^{n+\frac{1}{2}})}} \int_\Omega f(\tilde{U}^{n+\frac{1}{2}})(U^{n+1} - U^n) d\mathbf{x} \end{aligned} \tag{3.1}$$

subject to the periodic boundary condition and the following initial conditions:

$$U^0(\mathbf{x}) = u_0(\mathbf{x}), \quad V^0(\mathbf{x}) = u_1(\mathbf{x}).$$

As we may observe, the numerical solution (u_h^n, v_h^n, r_h^n) can be viewed as the Galerkin approximation of the above time-discrete system of the equation. To study the convergence of the temporal discretization (3.1), we need some preliminaries.

First, for the simplicity of notations, throughout this paper, we denote by C a generic positive constant, which depends solely upon the physical parameters of the problem and is independent of τ, h and n , and it is not necessary to be the same at every appearance. We adopt the usual notations for Sobolev spaces, e.g., $W^{m,p}(I)$ on the sub-domain $I \in \Omega$ equipped with the norm $\|\cdot\|_{W^{m,p},I}$ and the semi-norm $|\cdot|_{W^{m,p},I}$. We omit the index I when $I = \Omega$. Especially, when $p = 2$, we set $W^{m,p}(I) = H^m(I)$, $\|\cdot\|_{W^{m,p},I} = \|\cdot\|_{H^m,I}$ and $|\cdot|_{W^{m,p},I} = |\cdot|_{H^m,I}$. The notation $\alpha \lesssim \beta$ implies that α is bounded by β multiplied by a constant independent of τ, h and n .

Second, we present the Grönwall inequality, which plays an important role in our later convergence analysis and error estimates.

Lemma 3.1 (See [17]). *Let τ, B and a_k, b_k, c_k, γ_k for integers $k > 0$ be nonnegative numbers such that*

$$a_n + \tau \sum_{k=0}^n b_k \leq \tau \sum_{k=0}^n \gamma_k a_k + \tau \sum_{k=0}^n c_k + B \quad \text{for } n \geq 0.$$

Suppose that $\tau\gamma_k < 1$ for all k , and set $\sigma_k = (1 - \tau\gamma_k)^{-1}$. Then

$$a_n + \tau \sum_{k=0}^n b_k \leq \left(\tau \sum_{k=0}^n c_k + B \right) \exp \left(\tau \sum_{k=0}^n \gamma_k \sigma_k \right).$$

Lemma 3.2 (See [2]). *Let $I = [a, b]$ and $\alpha(t), \beta(t), u(t) \in C^0(I)$. Suppose $\beta(t) \geq 0$ and*

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)ds, \quad \forall t \in I.$$

Then

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)e^{\int_s^t \beta(r)dr} ds, \quad \forall t \in I.$$

Now we are ready to study the convergence of the solution of (3.1). Taking the inner product of the first two equations with $V^{n+1} - V^n$, $U^{n+1} - U^n$ and multiplying the third equation of (3.1) by $R^{n+1} + R^n$, we derive

$$\frac{1}{2}(\|V^{n+1}\|^2 - \|V^n\|^2 + \|\nabla U^{n+1}\|^2 - \|\nabla U^n\|^2 + \lambda\|U^{n+1}\|^2 - \lambda\|U^n\|^2) + (R^{n+1})^2 - (R^n)^2 = 0,$$

which indicates that

$$\|V^n\| + \|U^n\|_{H^1} + |R^n| \lesssim 1, \quad \forall 1 \leq n \leq N.$$

As pointed out in [29], the energy stable scheme is not sufficient for the convergence which typically needs bounds in higher norms. Following the idea in [29], our convergence analysis is along this line: we first start from the energy preservation to derive the error bounds in higher norms (i.e., the H^2 estimates) for the solution U^n , and thus obtain L^∞ for U^n thanks to the embedding theory, and then we use the bounds in H^2 -norms to show that the numerical solution U^n converges to the exact solution u^n in some suitable norms as τ tends to zero. To this end, we need the bounds in the H^2 -norm of the PDE system (1.1). The error bounds for the solution of (3.1) are similar to those of the PDE system.

Note that most of the convergence and error analysis for linearly implicit scheme are based on the so-called Lipschitz assumption, i.e.,

$$|F'(u_1) - F'(u_2)| \leq L|u_1 - u_2|, \quad \forall u_1, u_2. \quad (3.2)$$

The above assumption greatly limits its range of applicability. Following the basic idea of [29], we adopt the following assumption instead of the Lipschitz assumption in our convergence analysis:

$$|f'(x)| < C(|x|^p + 1), \quad p \geq 0, \quad \text{if } n = 1, 2; \quad 0 < p < 4, \quad \text{if } n = 3; \quad (3.3)$$

$$|f''(x)| < C(|x|^p + 1), \quad p \geq 0, \quad \text{if } n = 1, 2; \quad 0 < p < 3, \quad \text{if } n = 3. \quad (3.4)$$

It has been proved in [29] that if $f(u)$ satisfies the conditions (3.3)–(3.4), then for some σ ($0 \leq \sigma < 1$), it holds that

$$\|f''(u)\|_{L^\infty} + \|f'(u)\|_{L^\infty} \leq C(1 + \|\nabla \Delta u\|^\sigma) \leq \epsilon \|\nabla \Delta u\| + C_\epsilon \quad (3.5)$$

for any $\epsilon > 0$ with C_ϵ being a constant depending on ϵ .

We present the following estimates for the exact solution of (1.1).

Proposition 3.3. *Assume that u is the solution of (1.1), $u_0 \in H^3$, $u_1 \in H^2$ and (3.3)–(3.4) hold. Then for any $T > 0$,*

$$(\|\Delta u\| + \|\Delta u_t\| + \|\nabla \Delta u\|)(T) \lesssim 1.$$

Proof. First, multiplying u_t on both sides of (1.1) and using the integration by parts yield

$$\frac{d}{dt} \left(\|u_t\|^2 + \|\nabla u\|^2 + \lambda\|u\|^2 + 2 \int_{\Omega} F(u) \right) = 0,$$

which indicates that

$$\|u_t\|^2 + \|u\|_{H^1}^2 + \int_{\Omega} F(u) \lesssim 1. \quad (3.6)$$

On the other hand, we multiply $\Delta^2 u_t$ on both sides of (1.1) and again use the integration by parts to obtain

$$\frac{1}{2} \frac{d}{dt} (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2) = -(\Delta f(u), \Delta u_t). \quad (3.7)$$

Integrating with respect to time from 0 to t and using the Cauchy-Schwarz inequality yield

$$\begin{aligned} & (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(t) \\ & \leq (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(0) + \int_0^t (\|\Delta u_t\|^2 + \|\Delta f(u)\|^2)(t) dt. \end{aligned} \quad (3.8)$$

By (3.5) and the identity

$$\Delta f(u) = f'(u)\Delta u + f''(u)|\nabla u|^2,$$

we have that for all $0 \leq \delta < 1$,

$$\begin{aligned} \|\Delta f(u)\| &\leq \|f''(u)\|_{L^\infty} \|\nabla u\|_{L^4}^2 + \|f'(u)\|_{L^\infty} \|\Delta u\| \\ &\leq C(\|f''(u)\|_{L^\infty} + \|f'(u)\|_{L^\infty})(\|\nabla u\|_{L^4}^2 + \|\Delta u\|) \\ &\leq C(1 + \|\nabla \Delta u\|^\delta)(\|\nabla u\|_{L^4}^2 + \|\Delta u\|). \end{aligned}$$

As for the term $\|\nabla u\|_{L^4}$, we use the Sobolev embedding theory and the interpolation inequality about the spaces H^s (see, e.g., [27]) and then obtain

$$\|\nabla u\|_{L^4} \leq C\|\nabla u\|_{H^{d/4}} \leq C\|\nabla u\|^{1-d/8}\|\nabla \Delta u\|^{d/8} \leq C\|\nabla \Delta u\|^{d/8}.$$

Moreover, by using the integration by parts and (3.6), we have

$$\|\Delta u\|^2 \leq C\|\nabla \Delta u\|\|\nabla u\| \leq C\|\nabla \Delta u\|.$$

Consequently,

$$\|\Delta f(u)\|^2 \leq C(1 + \|\nabla \Delta u\|^{2\delta})(\|\nabla u\|_{L^4}^4 + \|\Delta u\|^2) \leq C(1 + \|\nabla \Delta u\|^2).$$

Substituting the above inequality into (3.8) gives

$$\begin{aligned} &(\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(t) \\ &\lesssim (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(0) + 1 + \int_0^t (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2)(t) dt. \end{aligned} \tag{3.9}$$

By the Grönwall inequality in Lemma 3.2, it holds that

$$(\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(t) \lesssim (\|\Delta u_t\|^2 + \|\nabla \Delta u\|^2 + \lambda\|\Delta u\|^2)(0) + 1 \lesssim 1.$$

This finishes our proof. □

Similar to the proof in the above proposition, we also have the following H^2 estimates for the solution of (3.1).

Proposition 3.4. *Assume that (U^n, V^n, R^n) are the solutions of (3.1) and (3.3)–(3.4) hold. Then*

$$\|\Delta U^n\| + \|\Delta V^n\|^2 + \|\nabla \Delta U^n\|^2 \lesssim 1. \tag{3.10}$$

Proof. First, we have that from the first equation of (3.1),

$$D_\tau \nabla \Delta U^{n+1} = \nabla \Delta \hat{V}^{n+\frac{1}{2}}.$$

Multiplying the above equation with $\nabla(V^{n+1} - V^n)$ and the second equation of (3.1) with $\Delta^2(U^{n+1} - U^n)$, and then using the integration by parts, we obtain

$$\begin{aligned} &\|\nabla \Delta U^{n+1}\|^2 - \|\nabla \Delta U^n\|^2 + \|\Delta V^{n+1}\|^2 - \|\Delta V^n\|^2 + \lambda\|\Delta U^{n+1}\|^2 - \lambda\|\Delta U^n\|^2 \\ &= \frac{2\hat{R}^{n+\frac{1}{2}}}{\sqrt{E(\tilde{U}^{n+\frac{1}{2}})}} (\nabla f(\tilde{U}^{n+\frac{1}{2}}), \nabla \Delta(U^{n+1} - U^n)). \end{aligned}$$

Writing

$$f_1(\tilde{U}^{n+\frac{1}{2}}) = \frac{2\hat{R}^{n+\frac{1}{2}} f(\tilde{U}^{n+\frac{1}{2}})}{\sqrt{E(\tilde{U}^{n+\frac{1}{2}})}}$$

and summing up the above equation for all n from 0 to m yield

$$\begin{aligned} & \|\nabla\Delta U^{m+1}\|^2 - \|\nabla\Delta U^0\|^2 + \|\Delta V^{m+1}\|^2 - \|\Delta V^0\|^2 + \lambda\|\Delta U^{m+1}\|^2 - \lambda\|\Delta U^0\|^2 \\ &= (\nabla f_1(\tilde{U}^{m+\frac{1}{2}}), \nabla\Delta U^{m+1}) - (\nabla f_1(\tilde{U}^{\frac{1}{2}}), \nabla\Delta U^0) + \sum_{n=1}^m (\nabla(f_1(\tilde{U}^{n-\frac{1}{2}}) - f_1(\tilde{U}^{n+\frac{1}{2}})), \nabla\Delta U^n) \\ &= (\nabla f_1(U^{m+1}), \nabla\Delta U^{m+1}) - (\nabla f_1(\tilde{U}^{\frac{1}{2}}), \nabla\Delta U^0) + I_1 - I, \end{aligned} \tag{3.11}$$

where

$$f_1(U^{m+1}) = \frac{2\hat{R}^{m+\frac{1}{2}}f(U^{m+1})}{\sqrt{E(\tilde{U}^{m+\frac{1}{2}})}}$$

and

$$\begin{aligned} I &= (\nabla f_1(U^{m+1}) - \nabla f_1(\tilde{U}^{m+\frac{1}{2}}), \nabla\Delta U^{m+1}), \\ I_1 &= \sum_{n=1}^m (\nabla(f_1(\tilde{U}^{n-\frac{1}{2}}) - f_1(\tilde{U}^{n+\frac{1}{2}})), \nabla\Delta U^n). \end{aligned}$$

Since R^n is bounded and $E(U)$ is bounded from below, we have that from (3.5),

$$\|\nabla f_1(U^{m+1})\| \leq C(1 + \|f'(U^{m+1})\|_{L^\infty}) \leq \epsilon\|\nabla\Delta U^{m+1}\| + C_\epsilon. \tag{3.12}$$

Consequently,

$$|(\nabla f_1(U^{m+1}), \nabla\Delta U^{m+1}) - (\nabla f_1(\tilde{U}^{\frac{1}{2}}), \nabla\Delta U^0)| \leq \frac{1}{4}\|\nabla\Delta U^{m+1}\|^2 + \frac{1}{4}\|\nabla\Delta U^0\|^2 + C. \tag{3.13}$$

On the other hand, we note that

$$\begin{aligned} & \nabla f_1(\tilde{U}^{n-\frac{1}{2}}) - \nabla f_1(\tilde{U}^{n+\frac{1}{2}}) \\ &= f'_1(\tilde{U}^{n-\frac{1}{2}})\nabla\tilde{U}^{n-\frac{1}{2}} - f'_1(\tilde{U}^{n+\frac{1}{2}})\nabla\tilde{U}^{n+\frac{1}{2}} \\ &= f'_1(\tilde{U}^{n-\frac{1}{2}})(\nabla\tilde{U}^{n-\frac{1}{2}} - \nabla\tilde{U}^{n+\frac{1}{2}}) + (f'_1(\tilde{U}^{n-\frac{1}{2}}) - f'_1(\tilde{U}^{n+\frac{1}{2}}))\nabla\tilde{U}^{n+\frac{1}{2}} \\ &= \frac{\tau}{2}f'_1(\tilde{U}^{n-\frac{1}{2}})\nabla(\hat{V}^{n-\frac{3}{2}} - 3\hat{V}^{n-\frac{1}{2}}) + \frac{\tau}{2}f''_1(\theta\tilde{U}^{n+\frac{1}{2}} + (1-\theta)\tilde{U}^{n-\frac{1}{2}})(\hat{V}^{n-\frac{3}{2}} - 3\hat{V}^{n-\frac{1}{2}})\nabla\tilde{U}^{n+\frac{1}{2}} \end{aligned}$$

for some $\theta \in (0, 1)$, where in the last step, we have used the first equation of (3.1), which yields

$$\nabla\tilde{U}^{n-\frac{1}{2}} - \nabla\tilde{U}^{n+\frac{1}{2}} = \frac{1}{2}\nabla(4U^{n-1} - 3U^n - U^{n-2}) = \frac{\tau}{2}\nabla(\hat{V}^{n-\frac{3}{2}} - 3\hat{V}^{n-\frac{1}{2}}).$$

By (3.12) and the fact that

$$\|V^n\| + \|\nabla U^n\| \lesssim 1,$$

we have

$$\|\nabla f_1(\tilde{U}^{n-\frac{1}{2}}) - \nabla f_1(\tilde{U}^{n+\frac{1}{2}})\|^2 \leq \tau^2(\epsilon\|\nabla\Delta\tilde{U}^{n-\frac{1}{2}}\|^2 + \epsilon\|\nabla\Delta\tilde{U}^{n+\frac{1}{2}}\|^2 + \|\nabla(\hat{V}^{n-\frac{3}{2}} - 3\hat{V}^{n-\frac{1}{2}})\|^2 + C^2),$$

and thus

$$|I_1| \leq C\tau \sum_{n=1}^m (\|\nabla\Delta U^n\|^2 + \|\nabla V^n\|^2) + C\tau.$$

Similarly, it holds that

$$\|\nabla f_1(\tilde{U}^{m+\frac{1}{2}}) - \nabla f_1(U^{m+1})\|^2 \leq \tau^2\epsilon \sum_{n=m-1}^{m+1} (\|\nabla V^n\|^2 + \|\nabla\Delta U^n\|^2) + C_\epsilon\tau^2.$$

Then

$$|I| \leq \tau^2 \sum_{n=m-1}^{m+1} (\|\nabla V^n\|^2 + \epsilon \|\nabla \Delta U^n\|^2) + C_\epsilon \tau^2 + \frac{1}{4} \|\nabla \Delta U^{m+1}\|^2.$$

Substituting (3.13), we see that the estimates of I_1 and I into (3.11) yield

$$\begin{aligned} & \frac{1}{2} \|\nabla \Delta U^{m+1}\|^2 + \|\Delta V^{m+1}\|^2 + \lambda \|\Delta U^{m+1}\|^2 \\ & \leq \|\Delta V^0\|^2 + \|\nabla \Delta U^0\|^2 + \lambda \|\Delta U^0\|^2 + C\tau \sum_{n=1}^{m+1} (\|\nabla \Delta U^n\|^2 + \|\Delta V^n\|^2) + C. \end{aligned}$$

By the Grönwall inequality, we have

$$\|\nabla \Delta U^{m+1}\|^2 + \|\Delta V^{m+1}\|^2 + \lambda \|\Delta U^{m+1}\|^2 \leq \|\Delta V^0\|^2 + \|\nabla \Delta U^0\|^2 + \lambda \|\Delta U^0\|^2 \leq C.$$

In the case $\lambda = 0$, we note that

$$\Delta U^{m+1} = \Delta U^m + \frac{\tau}{2} (\Delta V^{m+1} + \Delta V^m).$$

Then

$$\|\Delta U^{m+1}\| \leq \|\Delta U^m\| + \frac{\tau}{2} \|\Delta V^{m+1} + \Delta V^m\|,$$

which yields

$$\|\Delta U^{m+1}\| \leq \|\Delta U^0\| + C \leq C.$$

This finishes the proof of (3.10). The proof is completed. \square

As a direct consequence of (3.10) and the embedding inequality, we have

$$\|U^n\|_\infty \leq C \|\Delta U^n\| \leq C, \quad \forall n. \tag{3.14}$$

Remark 3.5. Following the same argument as in [29], we conclude that: assume $u_0 \in H^3$; when τ tends to zero, we have $U^n \rightarrow u^n$ strongly in $L^\infty(0, T; H^{3-\epsilon})$, $\forall \epsilon > 0$, weak-star in $L^\infty(0, T; H^3)$, $V^n \rightarrow v^n$ weak-star in $L^\infty(0, T; H^2)$, and $R^n \rightarrow r^n$ weak-star in $L^\infty(0, T)$.

Theorem 3.6. Suppose that u is the solution of (1.1) satisfying

$$\|u_0\|_{H^2} + \|u\|_{L^\infty((0,T),H^2)} + \|u_t\|_{L^2((0,T),H^2)} + \|u_{tt}\|_{L^2((0,T),H^2)} \lesssim 1.$$

Then (3.1) admits a unique solution (U^n, V^n, R^n) such that

$$\|u^n - U^n\|_{H^1} + \|v^n - V^n\| + |r^n - R^n| \lesssim \tau^2.$$

Proof. First, we define

$$E_u^n = u^n - U^n, \quad E_v^n = v^n - V^n, \quad E_r^n = r^n - R^n, \quad H(u) = \frac{f(u)}{\sqrt{E(u)}}.$$

By taking $t = t_{n+\frac{1}{2}}$ in (2.1) and using (3.1), we obtain

$$\begin{aligned} D_\tau E_u^{n+1} &= \hat{E}_v^{n+\frac{1}{2}} + T_1, \\ D_\tau E_v^{n+1} &= \Delta \hat{E}_u^{n+\frac{1}{2}} - \lambda \hat{E}_u^{n+\frac{1}{2}} - \hat{r}^{n+\frac{1}{2}} H(u^{n+\frac{1}{2}}) + \hat{R}^{n+\frac{1}{2}} H(\tilde{U}^{n+\frac{1}{2}}) + T_2, \\ E_r^{n+1} - E_r^n &= \frac{1}{2} \int_\Omega (H(u^{n+\frac{1}{2}})(u^{n+1} - u^n) - H(\tilde{U}^{n+\frac{1}{2}})(U^{n+1} - U^n)) d\mathbf{x} + T_3, \end{aligned} \tag{3.15}$$

where T_1, T_2 and T_3 denote the truncation errors, i.e.,

$$T_1 = D_\tau u^{n+1} - u_t^{n+\frac{1}{2}} + v^{n+\frac{1}{2}} - \hat{v}^{n+\frac{1}{2}},$$

$$\begin{aligned}
 T_2 &= \Delta(u^{n+\frac{1}{2}} - \hat{u}^{n+\frac{1}{2}}) + D_\tau v^{n+1} - v_t^{n+\frac{1}{2}} - \lambda u^{n+\frac{1}{2}} + \lambda \hat{u}^{n+\frac{1}{2}} + (\hat{r}^{n+\frac{1}{2}} - r^{n+\frac{1}{2}})H(u^{n+\frac{1}{2}}), \\
 T_3 &= \tau(D_\tau r^{n+1} - r_t^{n+\frac{1}{2}}) - \frac{\tau}{2} \int_\Omega H(u^{n+\frac{1}{2}})(D_\tau u^{n+1} - u_t^{n+\frac{1}{2}})d\mathbf{x}.
 \end{aligned}$$

Multiplying the first equation with $E_v^{n+1} - E_v^n$, the second equation with $E_u^{n+1} - E_u^n$, and the third equation with $E_r^{n+1} + E_r^n$ in (3.15), and then summing up three equalities, we obtain

$$\begin{aligned}
 &\frac{1}{2}(\|E_v^{n+1}\|^2 - \|E_v^n\|^2 + \|\nabla E_u^{n+1}\|^2 - \|\nabla E_u^n\|^2 + \lambda\|E_u^{n+1}\|^2 - \lambda\|E_u^n\|^2) + |E_r^{n+1}|^2 - |E_r^n|^2 \\
 &= (I_2, DE_u^{n+1}) + (T_2, DE_u^{n+1}) - (T_1, DE_v^{n+1}) + 2(T_3 + I_3)\hat{E}_r^{n+\frac{1}{2}},
 \end{aligned} \tag{3.16}$$

where

$$I_2 = \hat{r}^{n+\frac{1}{2}}(H(\tilde{U}^{n+\frac{1}{2}}) - H(u^{n+\frac{1}{2}})), \tag{3.17}$$

$$I_3 = \frac{1}{2} \int_\Omega (H(u^{n+\frac{1}{2}}) - H(\tilde{U}^{n+\frac{1}{2}}))(u^{n+1} - u^n)d\mathbf{x}. \tag{3.18}$$

We next estimate the terms T_i and I_i ($i \leq 3$), respectively. By the Taylor expansion, it holds that

$$\|T_1\| \lesssim \tau^2, \quad \|T_2\| \lesssim \tau^2, \quad |T_3| \lesssim \tau^3.$$

By (3.14) and the fact that $f \in C^2(\mathbb{R})$, we have

$$|H(U^n)| + |H'(U^n)| + |f'(U^n)| + |f''(U^n)| \lesssim 1, \quad \forall n.$$

Then there exists some $\theta \in (0, 1)$ such that

$$\begin{aligned}
 \|H(\tilde{U}^{n+\frac{1}{2}}) - H(u^{n+\frac{1}{2}})\| &= \|H'(\theta\tilde{U}^{n+\frac{1}{2}} + (1-\theta)u^{n+\frac{1}{2}})(\tilde{E}_u^{n+\frac{1}{2}} + \tilde{u}^{n+\frac{1}{2}} - u^{n+\frac{1}{2}})\| \\
 &\lesssim \tau^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|,
 \end{aligned}$$

and thus

$$\|I_2\| \lesssim \tau^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|, \quad |I_3| \lesssim \tau(\tau^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|).$$

Consequently,

$$|(T_3 + I_3)\hat{E}_r^{n+\frac{1}{2}}| \lesssim \tau|\hat{E}_r^{n+\frac{1}{2}}|^2 + \tau^{-1}(|I_3|^2 + |T_3|^2) \lesssim \tau^5 + \tau(|\hat{E}_r^{n+\frac{1}{2}}|^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|^2). \tag{3.19}$$

Note that

$$\|DE_u^{n+1}\| = \tau\|\hat{E}_v^{n+\frac{1}{2}} + T_1\| \lesssim \tau^3 + \tau\|\hat{E}_v^{n+\frac{1}{2}}\|.$$

Then

$$|(I_2 + T_2, DE_u^{n+1})| \lesssim \tau\|I_2\|^2 + \tau\|T_2\|^2 + \tau^{-1}\|DE_u^{n+1}\|^2 \lesssim \tau^5 + \tau(\|\hat{E}_v^{n+\frac{1}{2}}\|^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|^2). \tag{3.20}$$

On the other hand, in light of the second equation of (3.15), we have

$$\begin{aligned}
 (T_1, DE_v^{n+1}) &= \tau(T_1, \Delta\hat{E}_u^{n+\frac{1}{2}} - \lambda\hat{E}_u^{n+\frac{1}{2}} - \hat{r}^{n+\frac{1}{2}}H(u^{n+\frac{1}{2}}) + \hat{R}^{n+\frac{1}{2}}H(\tilde{U}^{n+\frac{1}{2}}) + T_2) \\
 &= -\tau(\nabla T_1, \nabla\hat{E}_u^{n+\frac{1}{2}}) + \tau(T_1, -\lambda\hat{E}_u^{n+\frac{1}{2}} - \hat{r}^{n+\frac{1}{2}}H(u^{n+\frac{1}{2}}) + \hat{R}^{n+\frac{1}{2}}H(\tilde{U}^{n+\frac{1}{2}}) + T_2).
 \end{aligned}$$

Noticing that

$$\begin{aligned}
 \|\hat{r}^{n+\frac{1}{2}}H(u^{n+\frac{1}{2}}) - \hat{R}^{n+\frac{1}{2}}H(\tilde{U}^{n+\frac{1}{2}})\| &= \|\hat{E}_r^{n+\frac{1}{2}}H(u^{n+\frac{1}{2}}) + \hat{R}^{n+\frac{1}{2}}(H(u^{n+\frac{1}{2}}) - H(\tilde{U}^{n+\frac{1}{2}}))\| \\
 &\lesssim |\hat{E}_r^{n+\frac{1}{2}}| + \tau^2 + \|\tilde{E}_u^{n+\frac{1}{2}}\|,
 \end{aligned}$$

we have

$$|(T_1, DE_v^{n+1})| \lesssim \tau^3(\|\nabla \hat{E}_u^{n+\frac{1}{2}}\| + \|\hat{E}_u^{n+\frac{1}{2}}\|^2 + \tau^2 + |E_r^{n+\frac{1}{2}}| + \|\tilde{E}_u^{n+\frac{1}{2}}\|). \tag{3.21}$$

Plugging (3.19)–(3.21) into (3.16) yields

$$\begin{aligned} & \frac{1}{2}(\|E_v^{n+1}\|^2 - \|E_v^n\|^2 + \|\nabla E_u^{n+1}\|^2 - \|\nabla E_u^n\|^2 + \lambda\|E_u^{n+1}\|^2 - \lambda\|E_u^n\|^2) + (E_r^{n+1})^2 - (E_r^n)^2 \\ & \lesssim \tau^5 + \tau(\|\tilde{E}_u^{n+\frac{1}{2}}\|^2 + \|\hat{E}_v^{n+\frac{1}{2}}\|^2 + \|\nabla \hat{E}_u^{n+\frac{1}{2}}\|^2 + \|\hat{E}_u^{n+\frac{1}{2}}\|^2 + |\hat{E}_r^{n+\frac{1}{2}}|^2). \end{aligned}$$

Summing up all n from 0 to m and using the initial values, we obtain

$$\begin{aligned} & \|E_v^{m+1}\|^2 + \|\nabla E_u^{m+1}\|^2 + \lambda\|E_u^{m+1}\|^2 + |E_r^{m+1}|^2 \\ & \leq C\tau^4 + C\tau \sum_{n=0}^{m+1} (\|E_v^n\|^2 + \|\nabla E_u^n\|^2 + \lambda\|E_u^n\|^2 + |E_r^n|^2). \end{aligned}$$

Then the desired result follows from the conclusion in Lemma 3.1. □

4 Error estimates for the fully discrete solution

In this section, we establish the error estimates for the solution of (2.2). By the error decomposition, we have

$$w^n - w_h^n = w^n - W^n + W^n - w_h^n = w^n - W^n + e_w^n, \quad w = u, v, r.$$

In light of the conclusion in Theorem 3.6, we only need to estimate the term e_w^n ($w = u, v, r$). To this end, we first define the Ritz projection operator $R_h : H_0^1(\Omega) \rightarrow V_h$ by

$$(\nabla(v - R_h v), \nabla \omega) = 0, \quad \forall \omega \in V_h.$$

Then e_u^n (similar for e_v^n) can be decomposed into

$$e_u^n = U^n - u_h^n = \xi_u^n + \eta_u^n, \quad \xi_u^n = R_h U^n - u_h^n, \quad \eta_u^n = U^n - R_h U^n.$$

According to the standard finite element method (FEM) theory [32], it holds that

$$\|v - R_h v\|_{L^2} + h\|\nabla(v - R_h v)\|_{L^2} \leq Ch^s \|v\|_{H^s}, \quad \forall v \in H^s(\Omega) \cap H_0^1(\Omega) \tag{4.1}$$

for $1 \leq s \leq k + 1$.

Note that the exact solutions of (3.1) satisfy

$$\begin{aligned} (D_\tau U^{n+1}, w_h) &= (\hat{V}^{n+\frac{1}{2}}, w_h), \\ (D_\tau V^{n+1}, \zeta_h) &= -(\nabla \hat{U}^{n+\frac{1}{2}}, \nabla \zeta_h) - \lambda(\hat{U}^{n+\frac{1}{2}}, \zeta_h) - (\hat{R}^{n+\frac{1}{2}} H(\tilde{U}^{n+\frac{1}{2}}), \zeta_h), \\ R^{n+1} - R^n &= \frac{1}{2} \int_\Omega H(\tilde{U}^{n+\frac{1}{2}})(U^{n+1} - U^n) d\mathbf{x}. \end{aligned} \tag{4.2}$$

Subtracting (2.2) from (4.2) gives the following error equation:

$$\begin{aligned} (D_\tau \xi_u^{n+1}, w_h) &= (\xi_v^{n+\frac{1}{2}}, w_h) + R_1(w_h), \\ (D_\tau \xi_v^{n+1}, \zeta_h) &= -(\nabla \hat{\xi}_u^{n+\frac{1}{2}}, \nabla \zeta_h) - \lambda(\hat{\xi}_u^{n+\frac{1}{2}}, \zeta_h) - (\hat{e}_r^{n+\frac{1}{2}} H(\tilde{u}_h^{n+\frac{1}{2}}), \zeta_h) + R_2(\zeta_h) + I_2(\zeta_h), \\ e_r^{n+1} - e_r^n &= \frac{1}{2} \int_\Omega H(\tilde{u}_h^{n+\frac{1}{2}})(\xi_u^{n+1} - \xi_u^n) d\mathbf{x} + R_3 + I_3, \end{aligned} \tag{4.3}$$

where

$$H(u) = \frac{f(u)}{\sqrt{E(u)}}$$

and

$$\begin{aligned}
 R_1(w_h) &= (\hat{\eta}_v^{n+\frac{1}{2}}, w_h) - (D_\tau \eta_u^{n+1}, w_h), \\
 R_2(\zeta_h) &= -(\nabla \hat{\eta}_u^{n+\frac{1}{2}}, \nabla \zeta_h) - \lambda(\hat{\eta}_u^{n+\frac{1}{2}}, \zeta_h) - (D_\tau \eta_v^{n+1}, \zeta_h), \\
 I_2(\zeta_h) &= (\hat{R}^{n+\frac{1}{2}}(H(\tilde{u}_h^{n+\frac{1}{2}}) - H(\tilde{U}^{n+\frac{1}{2}})), \zeta_h), \\
 R_3 &= \frac{1}{2} \int_\Omega H(\tilde{u}_h^{n+\frac{1}{2}})(\eta_u^{n+1} - \eta_u^n) d\mathbf{x}, \\
 I_3 &= \frac{1}{2} \int_\Omega (H(\tilde{U}^{n+\frac{1}{2}}) - H(\tilde{u}_h^{n+\frac{1}{2}}))(U^{n+1} - U^n) d\mathbf{x}.
 \end{aligned}
 \tag{4.4}$$

Theorem 4.1. Suppose that u is the solution of (1.1) satisfying

$$\|u_0\|_{H^{k+1}} + \|u\|_{L^\infty((0,T),H^{k+1})} + \|u_t\|_{L^2((0,T),H^{k+1})} + \|u_{tt}\|_{L^2((0,T),H^{k+1})} \lesssim 1$$

and (u_h^n, v_h^n, r_h^n) is the solution of (2.2) with

$$(u_h^0, v_h^0, r_h^0) = (R_h u_0, R_h u_1, \sqrt{E(u_0)}).$$

Then

$$\|R_h v^n - v_h^n\| + \|R_h u^n - u_h^n\|_{H^1} + |r^n - r_h^n| \lesssim h^{k+1} + \tau^2.
 \tag{4.5}$$

Proof. We first estimate the terms ξ_u^n and ξ_v^n . By taking

$$(w_h, \zeta_h) = (D\xi_v^{n+1}, D\xi_u^{n+1})$$

and multiplying the third equation of (4.3) by $2\hat{e}_r^{n+\frac{1}{2}} = e_r^{n+1} + e_r^n$, we derive

$$\begin{aligned}
 &\frac{1}{2}(\|\xi_v^{n+1}\|^2 - \|\xi_v^n\|^2 + \|\nabla \xi_u^{n+1}\|^2 - \|\nabla \xi_u^n\|^2 + \lambda\|\xi_u^{n+1}\|^2 - \lambda\|\xi_u^n\|^2) + (e_r^{n+1})^2 - (e_r^n)^2 \\
 &= R_2(D\xi_u^{n+1}) + I_2(D\xi_u^{n+1}) - R_1(D\xi_v^{n+1}) + 2(R_3 + I_3)\hat{e}_r^{n+\frac{1}{2}},
 \end{aligned}$$

where R_i and I_i ($i \leq 3$) are given in (4.4). Summing up all n from 0 to m and using the initial error

$$\xi_u^0 = \xi_v^0 = 0,$$

we obtain

$$\begin{aligned}
 &\frac{1}{2}(\|\xi_v^{m+1}\|^2 + \|\nabla \xi_u^{m+1}\|^2 + \lambda\|\xi_u^{m+1}\|^2) + |e_r^{m+1}|^2 \\
 &= \sum_{n=0}^m (R_2(D\xi_u^{n+1}) + I_2(D\xi_u^{n+1}) - R_1(D\xi_v^{n+1}) + 2(R_3 + I_3)\hat{e}_r^{n+\frac{1}{2}}).
 \end{aligned}
 \tag{4.6}$$

To estimate the terms on the right-hand side of (4.6), we shall first make the hypothesis that there exists a positive constant C_* such that

$$\|u_h^n\|_{L^\infty} \leq C_*.
 \tag{4.7}$$

This hypothesis will be verified later by using the method of mathematical induction.

Due to (4.7) and the fact that $f \in C^2(\mathbb{R})$, we have

$$|H(u_h^n)| + |H'(u_h^n)| + |f'(u_h^n)| + |f''(u_h^n)| \lesssim 1, \quad \forall n.$$

Then

$$|R_3| \lesssim \tau h^{k+1}, \quad |R_2(D\xi_u^{n+1})| \lesssim h^{k+1} \|D\xi_u^{n+1}\|.$$

By the Taylor expansion, there exists a $\theta \in (0, 1)$ such that

$$\|H(\tilde{u}_h^{n+\frac{1}{2}}) - H(\tilde{U}^{n+\frac{1}{2}})\| = \|H'(\theta\tilde{u}_h^{n+\frac{1}{2}} + (1-\theta)\tilde{U}^{n+\frac{1}{2}})(\tilde{u}_h^{n+\frac{1}{2}} - \tilde{U}^{n+\frac{1}{2}})\|$$

$$\lesssim h^{k+1} + \|\tilde{\xi}_u^{n+\frac{1}{2}}\|. \tag{4.8}$$

Then

$$|I_2(D\xi_u^{n+1})| \lesssim (h^{k+1} + \|\tilde{\xi}_u^{n+\frac{1}{2}}\|)\|D\xi_u^{n+1}\|, \quad |I_3| \lesssim \tau(h^{k+1} + \|\tilde{\xi}_u^{n+\frac{1}{2}}\|),$$

and thus

$$\left| \sum_{n=0}^m 2(R_3 + I_3)\hat{e}_r^{n+\frac{1}{2}} \right| \leq Ch^{2(k+1)} + \tau \sum_{n=1}^m (\|\tilde{\xi}_u^{n+\frac{1}{2}}\|^2 + |\hat{e}_r^{n+\frac{1}{2}}|^2). \tag{4.9}$$

On the other hand, we choose

$$w_h = D\xi_u^{n+1}$$

in (4.3) to obtain

$$\|D\xi_u^{n+1}\| \lesssim \tau(\|\hat{\xi}_v^{n+\frac{1}{2}}\| + h^{k+1}),$$

which yields, together with the Cauchy-Schwarz inequality,

$$|I_2(D\xi_u^{n+1})| + |R_2(D\xi_u^{n+1})| \lesssim \tau h^{2(k+1)} + \tau(\|\hat{\xi}_v^{n+\frac{1}{2}}\|^2 + \|\tilde{\xi}_u^{n+\frac{1}{2}}\|^2).$$

Consequently,

$$\sum_{n=0}^m |I_2(D\xi_u^{n+1})| + |R_2(D\xi_u^{n+1})| \lesssim h^{2(k+1)} + \tau \sum_{n=0}^m (\|\hat{\xi}_v^{n+\frac{1}{2}}\|^2 + \|\tilde{\xi}_u^{n+\frac{1}{2}}\|^2). \tag{4.10}$$

As for the term

$$\sum_{n=0}^m R_1(D\xi_v^{n+1})$$

in (4.6), we recall the definition of R_1 in (4.4) to obtain

$$\begin{aligned} \left| \sum_{n=0}^m R_1(D\xi_v^{n+1}) \right| &= \left| (\hat{\eta}_v^{m+\frac{1}{2}} - D_\tau \eta_u^{m+1}, \xi_v^{m+1}) + \sum_{n=1}^m (\hat{\eta}_v^{n-\frac{1}{2}} - \hat{\eta}_v^{n+\frac{1}{2}} + D_\tau \eta_u^{n+1} - D_\tau \eta_u^n, \xi_v^n) \right| \\ &\leq Ch^{k+1} \|\xi_v^{m+1}\| + C\tau h^{k+1} \sum_{n=1}^m \|\xi_v^n\| \\ &\leq Ch^{2(k+1)} + C\tau \sum_{n=1}^m \|\xi_v^n\|^2 + \frac{1}{4} \|\xi_v^{m+1}\|^2. \end{aligned}$$

Substituting the above inequality and (4.9)–(4.10) into (4.6), we obtain

$$\|\xi_v^{m+1}\|^2 + \|\nabla \xi_u^{m+1}\|^2 + \|\xi_u^{m+1}\|^2 + |e_r^{m+1}|^2 \lesssim h^{2(k+1)} + \tau \sum_{n=0}^{m+1} (\|\xi_v^n\|^2 + |e_r^n|^2 + \|\xi_u^n\|^2).$$

By the Grönwall inequality given in Lemma 3.1,

$$\|\xi_v^n\| + \|\xi_u^n\|_{H^1} + |e_r^n| \lesssim h^{k+1}, \quad \forall n \geq 1. \tag{4.11}$$

Then from the triangle inequality and the conclusion in Theorem 3.6,

$$\begin{aligned} &\|R_h v^n - v_h^n\| + \|R_h u^n - u_h^n\|_{H^1} + |r^n - \tau_h^n| \\ &\leq \|\xi_v^n\| + \|\xi_u^n\|_{H^1} + |e_r^n| + \|R_h(v^n - V^n)\| + \|R_h(u^n - U^n)\| + |r^n - R^n| \lesssim h^{k+1} + \tau^2. \end{aligned}$$

The proof is completed. □

Remark 4.2. Note that the optimal convergence rate for the H^1 -error approximation is $O(h^k)$. The error estimate in (4.5) indicates that the Galerkin SAV solution u_h^n is superclose to the Ritz projection of the exact solution $R_h u^n$ under the H^1 -norm, which is one order higher than the counterpart optimal convergence rate. As a direct consequence of (4.5), we have the following optimal error estimates:

$$\begin{aligned} \|v^n - v_h^n\|_{H^1} + \|u^n - u_h^n\|_{H^1} &\lesssim h^k + \tau^2, \\ \|v^n - v_h^n\| + \|u^n - u_h^n\| + |r^n - r_h^n| &\lesssim h^{k+1} + \tau^2. \end{aligned}$$

To end this section, we prove the inequality (4.7).

Lemma 4.3. Under the conditions of Theorem 4.1, it holds that

$$\|u_h^n\|_{L^\infty} \leq C_*, \quad \forall n \geq 1,$$

where the constant C_* is independent of τ and h .

Proof. We will show the above inequality by induction. To this end, we first denote by $\mathcal{I}_h U \in V_h$ the interpolation function of U . By the approximation theory, we have

$$\|\mathcal{I}_h U - U\| + \|\mathcal{I}_h U - R_h U\| \lesssim h^2 \|U\|_{H^2}, \quad \|\mathcal{I}_h U\|_{L^\infty} \lesssim \|U\|_{L^\infty}.$$

Note that

$$\|u_h^0\|_{L^\infty} = \|R_h u^0\|_{L^\infty} \leq C.$$

By (4.11) and the inverse inequality $\|v_h\|_{L^\infty} \lesssim h^{-\frac{d}{2}} \|v_h\|$ for all $v_h \in V_h$, we obtain

$$\begin{aligned} \|u_h^1\|_{L^\infty} &\leq \|u_h^1 - R_h U^1\|_{L^\infty} + \|R_h U^1 - \mathcal{I}_h U^1\|_{L^\infty} + \|\mathcal{I}_h U^1\|_{L^\infty} \\ &\leq C(h^{2-\frac{d}{2}} + \|U^1\|_{L^\infty}) \leq C_1(h^{2-\frac{d}{2}} + \|U^1\|_{H^2}). \end{aligned}$$

Now we choose a positive constant h_1 which is small enough to satisfy

$$C_1 h_1^{\frac{1}{2}} \leq C.$$

Then for $h \in (0, h_1]$, we derive that

$$\|u_h^1\|_{L^\infty} \leq C + \|U^1\|_{H^2} \leq C_3.$$

Therefore, we can choose the positive constant C_* independent of h and τ such that

$$C_* \geq \max\{2\|U^n\|_{H^2}, \|u_h^1\|_{L^\infty}\}.$$

Then (4.7) is valid for $n = 1$. Next, suppose that (4.7) holds for all $l \leq n - 1$. We will show that it is also valid for n . Thanks to (4.11), we have

$$\|u_h^n - R_h U^n\|_{H^1} \leq C h^{k+1}.$$

Then

$$\begin{aligned} \|u_h^n\|_{L^\infty} &\leq \|u_h^n - R_h U^n\|_{L^\infty} + \|R_h U^n - \mathcal{I}_h U^n\|_{L^\infty} + \|\mathcal{I}_h U^n\|_{L^\infty} \\ &\leq C h^{-\frac{d}{2}} (\|u_h^n - R_h U^n\| + \|R_h U^n - \mathcal{I}_h U^n\|) + \|U^n\|_{L^\infty} \\ &\leq C_1 h^{\frac{1}{2}} + \frac{C_*}{2}. \end{aligned}$$

Let h_1 be small enough to satisfy

$$C_1 h_1^{\frac{1}{2}} \leq \frac{C_*}{2}.$$

Then for $h \in (0, h_1]$, we derive that

$$\|u_h^n\|_{L^\infty} \leq C_1 h^{\frac{1}{2}} + \frac{C_*}{2} \leq C_*.$$

This completes the induction. □

5 Numerical simulations

We present several numerical results to confirm our theoretical findings in this section.

Example 5.1. Consider the following Klein-Gordon equation:

$$u_{tt} = u_{xx} + u_{yy} + u - u^3 + g_1(x, y, t), \quad (x, y, t) \in [0, 1]^2 \times [0, T], \tag{5.1}$$

where $u(x, y, 0)$, $u_t(x, y, 0)$ and $g_1(x, y, t)$ are given by the exact solution

$$u(x, t) = \exp(-t)x^2(1 - x)^2y^2(1 - y)^2. \tag{5.2}$$

We test convergence orders of the fully discrete scheme using uniform triangulation with $M + 1$ nodes in each spatial direction, and take

$$N = M \quad \text{and} \quad N = M^{\frac{3}{2}}$$

for the linear finite element method (L-FEM) and the quadratic finite element method (Q-FEM), respectively. We list the errors at time $T = 1$ as well as the convergence rates in Table 1. Here and below, we define

$$\|e\|_0 = \|u^N - u_h^N\|, \quad \|e\|_1 = \|R_h u^N - u_h^N\|_{H^1}.$$

These results indicate that the fully discrete scheme is convergent and has order $\mathcal{O}(\tau^2 + h^{r+1})$. We also test the unconditional convergence of the fully discrete scheme with different spatial step-sizes for every fixed τ . The l_2 -errors at time $T = 1$ are shown in Figure 1. When the temporal stepsize is fixed, the L_2 -errors tend to a constant. They imply that the error estimates hold without certain time-space grid restrictions.

Table 1 Errors and convergence orders for 2D problems

M	L-FEM				Q-FEM			
	$\ e\ _0$	Order	$\ e\ _1$	Order	$\ e\ _0$	Order	$\ e\ _1$	Order
8	4.54E-4	–	5.56E-4	–	2.47E-6	–	4.60E-5	–
16	1.11E-4	2.03	1.61E-4	1.83	2.71E-7	3.19	6.85E-6	2.74
24	4.90E-5	2.02	7.40E-5	1.96	7.66E-8	3.12	2.41E-6	2.57
32	2.74E-5	2.02	4.24E-5	1.96	3.18E-8	3.05	9.81E-7	3.10
40	1.75E-5	2.01	2.73E-5	1.97	1.62E-8	3.02	4.88E-7	3.12

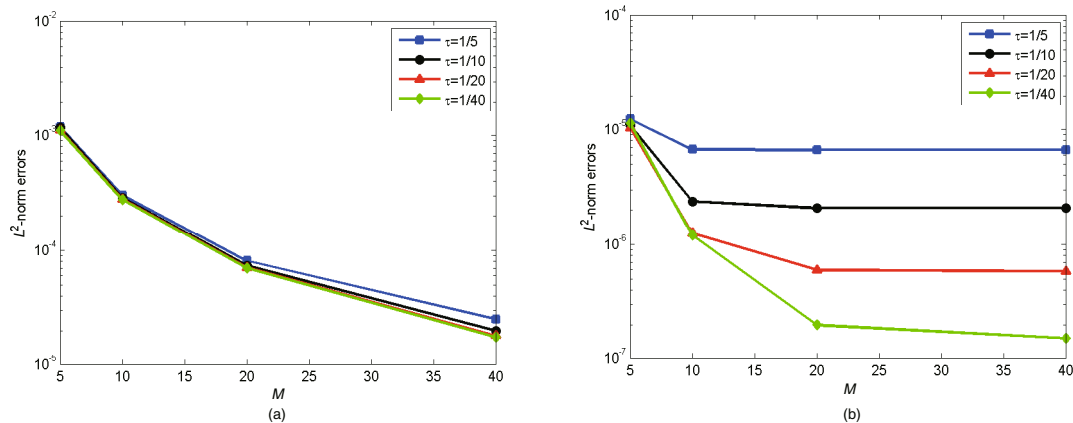


Figure 1 (Color online) L^2 -errors of the linear and quadratic finite element approximation

Then we set $g_1(x, y, t) = 0$, $T = 100$, $N = 10$ and $M = 10$, and solve the problem using the L-FEM. The time discretization is achieved by the linearized Crank-Nicolson (LCN) method and by the proposed SAV LCN method, respectively. The evolutions of the discrete energies are shown in Figure 2. Clearly, the energies obtained by the LCN finite element method increase as time goes on, while the one obtained by our method remains the same. It implies that numerical solutions by the SAV approach conserve the energy.

Example 5.2. Consider the following sine-Gordon equation:

$$u_{tt} = u_{xx} + u_{yy} + u_{zz} + \sin(u) + g_2(x, y, z, t), \quad (x, y, z, t) \in [0, 1]^3 \times [0, 1], \tag{5.3}$$

where the initial conditions and $g_2(x, y, z, t)$ are produced from the exact solution

$$u(x, t) = (1 + t^3) \sin(2\pi x) \sin(2\pi y) \sin(2\pi z). \tag{5.4}$$

We still take $N = M$ and $N = M^{\frac{3}{2}}$ for the linear and quadratic finite element approximations, respectively. The numerical errors at time $T = 1$ as well as the convergence rates are presented in Table 2. The given results indicate that the fully discrete scheme has order $\mathcal{O}(\tau^2 + h^{r+1})$.

Next, we set $g_2(x, y, z, t) = 0$, $T = 20$, $N = 10$ and $M = 10$, and solve the problem by the linear finite element method. The evolutions of the discrete energies for the 3D problems are displayed in Figure 3. Clearly, the discrete energies by the SAV approach remain unchanged, while the ones obtained by the LCN finite element method increase as time goes on. They further confirm the findings in this study.

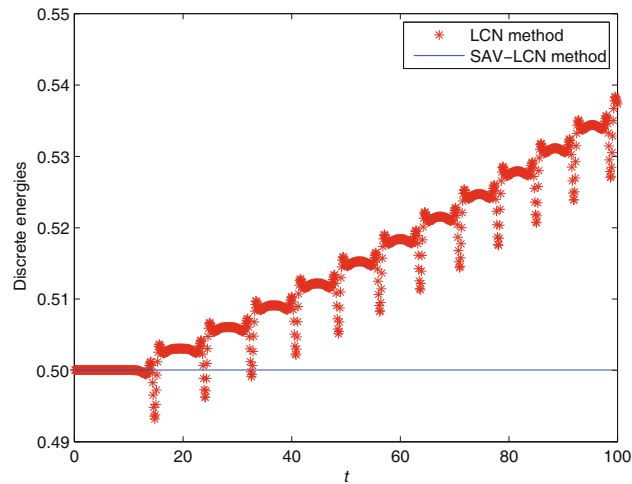


Figure 2 (Color online) The evolutions of the discrete energies

Table 2 Errors and convergence orders for 3D problems

M	L-FEM				M	Q-FEM			
	$\ e\ _0$	Order	$\ e\ _1$	Order		$\ e\ _0$	Order	$\ e\ _1$	Order
16	6.36E-2	—	4.43E-1	—	10	5.69E-3	—	1.29E-1	—
20	4.17E-2	1.89	2.94E-1	1.83	12	3.24E-3	3.09	7.41E-2	2.96
24	2.94E-2	1.92	2.08E-1	1.91	14	1.99E-3	3.16	4.91E-2	2.67
28	2.18E-2	1.94	1.55E-1	1.90	16	1.31E-3	3.13	3.40E-2	2.75

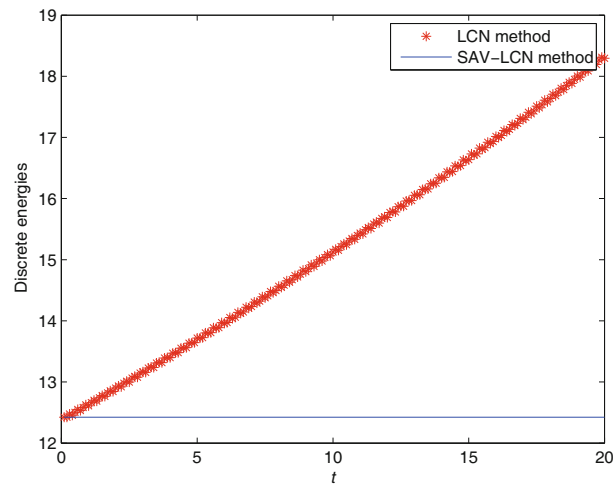


Figure 3 (Color online) The evolutions of the discrete energies

6 Conclusion

In this study, we present a linearly implicit numerical scheme for solving the nonlinear wave equations (1.1). The scheme is developed by combining the SAV approach with the finite element methods, classical Crank-Nicolson methods and extrapolation approximation. The fully discrete scheme is proved to be unconditionally convergent and energy-conserving. Numerical illustrations are presented to confirm the theoretical findings.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 11771162, 11771128, 11871106, 11871092 and 11926356) and National Safety Administration Fund (Grant No. U1930402).

References

- 1 Akrivis G, Li B, Li D. Energy-decaying extrapolated RK-SAV methods for the Allen-Cahn and Cahn-Hilliard equations. *SIAM J Sci Comput*, 2019, 41: A3703–A3727
- 2 Bellman R. The stability of solutions of linear differential equations. *Duke Math J*, 1943, 10: 643–647
- 3 Brugnano L, Caccia G F, Iavernaro F. Energy conservation issues in the numerical solution of the semilinear wave equation. *Appl Math Comput*, 2015, 270: 842–870
- 4 Cai J, Shen J. Two classes of linearly implicit local energy-preserving approach for general multi-symplectic Hamiltonian PDEs. *J Comput Phys*, 2020, 401: 108975
- 5 Cai W, Jiang C, Wang Y, et al. Structure-preserving algorithms for the two-dimensional sine-Gordon equation with Neumann boundary conditions. *J Comput Phys*, 2019, 395: 166–185
- 6 Cannon J R, Lin Y. Non-classical H^1 projection and Galerkin methods for non-linear parabolic integro-differential equations. *Calcolo*, 1988, 25: 187–201
- 7 Cano B, Moreta M J. Multistep cosine methods for second-order partial differential systems. *IMA J Numer Anal*, 2010, 30: 431–461
- 8 Cao W, Li D, Zhang Z. Optimal superconvergence of energy conserving local discontinuous Galerkin methods for wave equations. *Commun Comput Phys*, 2017, 21: 211–236
- 9 Carstensen C, Dolzmann G. Time-Space discretization of the nonlinear hyperbolic system $u_{tt} = \text{div}(\sigma(DU)) + DU_t$. *SIAM J Numer Anal*, 2004, 42: 75–89
- 10 Dendy J E. Galerkin's method for some highly nonlinear problems. *SIAM J Numer Anal*, 1977, 14: 327–347
- 11 Dodd R K, Eilbeck I C, Gibbon J D, et al. *Solitons and Nonlinear Wave Equations*. London-New York: Academic Press, 1982
- 12 Drazin P J, Johnson R S. *Solitons: An introduction*. *Physics Today*, 1990, 43: 70–71
- 13 Farago I. Finite element method for solving nonlinear parabolic equations. *Comput Math Appl*, 1991, 21: 59–69
- 14 Garcia S. Improved error estimates for mixed finite-element approximations for nonlinear parabolic equations: The

- discrete-time case. *Numer Methods Partial Differential Equations*, 1994, 10: 149–169
- 15 Gauckler L. Error analysis of trigonometric integrators for semilinear wave equations. *SIAM J Numer Anal*, 2015, 53: 1082–1106
 - 16 Grimm V. A note on the Gautschi-type method for oscillatory second-order differential equations. *Numer Math*, 2005, 102: 61–66
 - 17 Heywood J G, Rannacher R. Finite-element approximation of the nonstationary Navier-Stokes problem. Part IV: Error analysis for second-order time discretization. *SIAM J Numer Anal*, 1990, 27: 353–384
 - 18 Li B, Gao H, Sun W. Unconditionally optimal error estimates of a Crank-Nicolson Galerkin method for the nonlinear thermistor equations. *SIAM J Numer Anal*, 2014, 52: 933–954
 - 19 Li B, Sun W. Error analysis of linearized semi-implicit Galerkin finite element methods for nonlinear parabolic equations. *Int J Numer Anal Model*, 2013, 10: 622–633
 - 20 Li B, Sun W. Unconditional convergence and optimal error estimates of a Galerkin-mixed FEM for incompressible miscible flow in porous media. *SIAM J Numer Anal*, 2013, 51: 1959–1977
 - 21 Li D, Sun W. Linearly implicit and high-order energy-conserving schemes for nonlinear wave equations. *J Sci Comput*, 2020, 83: 65
 - 22 Li D, Wang J, Zhang J. Unconditionally convergent L_1 -Galerkin FEMs for nonlinear time-fractional Schrödinger equations. *SIAM J Sci Comput*, 2017, 39: A3067–A3088
 - 23 Li D, Wu C, Zhang Z. Linearized Galerkin FEMs for nonlinear time fractional parabolic problems with non-smooth solutions in time direction. *J Sci Comput*, 2019, 80: 403–419
 - 24 Li X, Wen J, Li D. Mass- and energy-conserving difference schemes for nonlinear fractional Schrödinger equations. *Appl Math Lett*, 2021, 111: 106686
 - 25 Luskin M. A Galerkin method for nonlinear parabolic equations with nonlinear boundary conditions. *SIAM J Numer Anal*, 1979, 16: 284–299
 - 26 Rachford Jr H H. Two-level discrete-time Galerkin approximations for second order nonlinear parabolic partial differential equations. *SIAM J Numer Anal*, 1973, 10: 1010–1026
 - 27 Roger T. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, 2nd ed. New York: Springer, 1997
 - 28 Sassaman R, Biswas A. Soliton perturbation theory for phi-four model and nonlinear Klein-Gordon equations. *Commun Nonlinear Sci Numer Simul*, 2009, 14: 3239–3249
 - 29 Shen J, Xu J. Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows. *SIAM J Numer Anal*, 2018, 56: 2895–2912
 - 30 Shen J, Xu J, Yang J. The scalar auxiliary variable (SAV) approach for gradient flows. *J Comput Phys*, 2018, 353: 407–416
 - 31 Shen J, Xu J, Yang J. A new class of efficient and robust energy stable schemes for gradient flows. *SIAM Rev*, 2019, 61: 474–506
 - 32 Thomée V. *Galerkin Finite Element Methods for Parabolic Problems*. Berlin-Heidelberg: Springer-Verlag, 1997
 - 33 Wang B, Wu X. The formulation and analysis of energy-preserving schemes for solving high-dimensional nonlinear Klein-Gordon equations. *IMA J Numer Anal*, 2019, 39: 2016–2044
 - 34 Wazwaz A M. New travelling wave solutions to the Boussinesq and the Klein-Gordon equations. *Commun Nonlinear Sci Numer Simul*, 2008, 13: 889–901
 - 35 Wu X, Wang B, Shi W. Efficient energy preserving integrators for oscillatory Hamiltonian systems. *J Comput Phys*, 2013, 235: 587–605