

## Original articles

# Arbitrarily high-order explicit energy-conserving methods for the generalized nonlinear fractional Schrödinger wave equations

 Yang Liu<sup>a</sup>, Maohua Ran<sup>a,\*</sup>
<sup>a</sup> School of Mathematical Sciences and V.C. & V.R. Key Lab of Sichuan Province, Sichuan Normal University, Chengdu 610068, China

Received 4 June 2023; received in revised form 24 August 2023; accepted 31 August 2023

Available online 7 September 2023

## Abstract

A novel category of explicit conservative numerical methods with arbitrarily high-order is introduced for solving the nonlinear fractional Schrödinger wave equations in one and two dimensions. The proposed method is based on the scalar auxiliary variable approach. The equations studied is first transformed into an equivalent system by introducing a scalar auxiliary variable, and the energy is then reformulated as a sum of three quadratic terms. Applying the explicit relaxation Runge–Kutta methods in temporal and the Fourier pseudo-spectral discretization in spatial, the resulting time–space full discrete scheme is proved to preserve the reformulated energy in the discrete level to machine accuracy. The proposed methods improve the numerical stability during long-term computations, as demonstrated through numerical experiments. Also this idea can be easily extended to other similar equations, such as the nonlinear fractional wave equation and the fractional Klein–Gordon–Schrödinger equation.

© 2023 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

**Keywords:** Structure-preserving method; Fractional Schrödinger wave equations; Explicit relaxation Runge–Kutta method; Scalar auxiliary variable approach; Fourier pseudo-spectral method

## 1. Introduction

This paper focuses on the numerical solution of the initial boundary-value problems of the nonlinear fractional Schrödinger wave equations (NFSWEs) as follows

$$u_{tt} + (-\Delta)^{\alpha/2} u + i\kappa u_t + \beta g(|u|^2)u = 0, \quad \mathbf{x} \in \Omega, \quad t \in (0, T], \quad (1.1)$$

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad u_t(\mathbf{x}, 0) = u_1(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (1.2)$$

$$u(\mathbf{x} + \mathbf{L}, t) = u(\mathbf{x}, t), \quad t \in [0, T], \quad (1.3)$$

where,  $i = \sqrt{-1}$ ,  $1 < \alpha \leq 2$ , and  $\kappa, \beta (> 0)$  are two real constants,  $u(\mathbf{x}, t)$  is unknown complex valued function,  $u_0(\mathbf{x})$  and  $u_1(\mathbf{x})$  are known smooth functions, the nonlinear term  $g$  is a given real function,  $\mathbf{x} \in \Omega \subset \mathbb{R}^d$  ( $d = 1, 2$ ) and  $\mathbf{L}$  is the period. The fractional Laplacian operator can be expressed in terms of the Fourier transform, as

$$(-\Delta)^{\frac{\alpha}{2}} u(\mathbf{x}, t) = \mathcal{F}^{-1} [|\xi|^\alpha \mathcal{F}(u(\xi, t))], \quad (1.4)$$

\* Corresponding author at: School of Mathematical Sciences and V.C. & V.R. Key Lab of Sichuan Province, Sichuan Normal University, Chengdu 610068, China.

E-mail addresses: [liuyang@stu.sicnu.edu.cn](mailto:liuyang@stu.sicnu.edu.cn) (Y. Liu), [maohuaran@163.com](mailto:maohuaran@163.com) (M. Ran).

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  denote the Fourier transform and its inverse respectively, see [3]. The NFSWEs (1.1) can be considered as a generalization of the classical Schrödinger wave equations, the latter has numerous physical applications such as the Langmuir wave envelope approximation in plasma [6], and has been studied in depth, see e.g., [1,2,5,20].

Due to its importance in both theoretical and applied physics, there has been a great deal of interest in developing efficient and accurate numerical methods for solving the NFSWEs (1.1). Noticing that the NFSWEs has the following conservation law of energy

$$E(t) = \|u_t(\cdot, t)\|^2 + \left\| (-\Delta)^{\frac{\alpha}{4}} u(\cdot, t) \right\|^2 + \frac{\beta}{2} \int_{\mathbb{R}^d} G(|u|^2) dx = E(0), \quad 0 \leq t \leq T, \quad (1.5)$$

where  $G(s) = \int_0^s g(z) dz$ , see [1,13]. The scholars have a more significant interest in structure-preserving methods. This is because, for conservative problems, numerical methods that can preserve the underlying invariance are usually advantageous. Besides, conservative schemes can achieve good stability in long-time simulations. In the past few years, some structure-preserving methods for the NFSWEs have been proposed. For instance, Ran and Zhang [13] developed a three-level linearly implicit difference scheme that preserves a modified discrete mass and energy. Li and Zhao [12] proposed a conservative method by combining the Crank–Nicolson method and the Galerkin finite element method. Moreover, a fast Krylov subspace solver is introduced to reduce computational cost. Cheng and Qin [4] developed a linearly-implicit conservative scheme based on the scalar auxiliary variable (SAV) method, which preserves only a modified energy, but not mass. Hu et al. [9] proposed three energy-preserving spectral Galerkin methods by applying the Crank–Nicolson, SAV, and exponential-SAV (ESAV) methods in time, respectively. Zhang and Ran [21] proposed and analyzed the higher order energy-preserving difference scheme based on triangular-SAV (T-SAV) approach. Nevertheless, most of these methods proposed in the above paper are only focused on the one-dimensional case, have no higher than second-order accuracy in time, and/or are fully-implicit. It means that there are still many open questions, and efficient, accurate and explicit numerical methods would be desirable for future investigations, especially for high dimensions.

Among the various numerical methods with high-order accuracy, explicit Runge–Kutta (RK) methods are good candidates, because they belong to one-step methods, and have the characteristics of high order accuracy and simplicity of implementation. However, the standard RK methods do not necessarily preserve the desired physical properties of the system. To overcome this issue, Ketcheson [10] proposed the relaxation Runge–Kutta (RRK) method, which guarantees conservation or stability with respect to any inner product norm. Later, the RRK technique has been extended to general convex quantities in [15]. Hence, the conservation, dissipation, or other properties with respect to any convex functional are enforced by the addition of a relaxation parameter that multiplies the Runge–Kutta update at each step. The trade-off for these advantages is that a nonlinear algebraic system has to be solved to determine the relaxation parameter. Nevertheless, the construction of explicit conservative schemes for the case of non-quadratic invariants has not been considered by the authors. Fortunately, the invariant energy quadratization (IEQ) approach [18,19] and the SAV approach [4] can transform the non-quadratic energy into a quadratic form of a new variable via a change of variables, and the resulting new, equivalent system still retains a similar energy law in terms of the new variables. Motivated by these developments, this paper aims to develop numerical methods for the NFSWEs in one and two dimensions by combining the SAV method and the explicit RRK method. And the proposed method has the following advantages:

- explicit scheme;
- has accuracy of arbitrary high order in temporal direction;
- preserves the invariant quantity (1.5).

The remainder of this work is organized as follows. In Section 2, we reformulate the NFSWEs (1.1) into an equivalent system by introducing a scalar auxiliary variable. Section 3 obtains a semi-discrete conservative system by applying the Fourier pseudo-spectral approximation to the resulting reformulation. In Section 4, we present the explicit fully-discrete schemes for the reformulated semi-discrete systems by applying the relaxation Runge–Kutta method in time, and the proposed scalar auxiliary variable relaxation Runge–Kutta (SAV-RRK) schemes preserve the same convergence order as the standard scalar auxiliary variable Runge–Kutta (SAV-RK) method. We further estimate the relaxation coefficient introduced and then obtain the accuracy of the relaxation methods in Section 5. Numerical examples are presented in Section 6 to demonstrate the accuracy and conservation properties of the proposed schemes. A brief conclusion is drawn in Section 7.

## 2. SAV reformulations of the NFSWEs

It is well-known that all RK methods preserve arbitrary linear invariants, and only symplectic RK methods preserve arbitrary quadratic invariants. However, no RK methods can preserve arbitrary polynomial invariants of a degree higher than two or non-polynomial nonlinear invariants. To overcome this limitation and take advantage of the quadratic invariant-preserving property of relaxation RK methods, we adopt the newly developed SAV approach to rewrite high-degree energy functions as quadratic ones. It allows the NFSWEs (1.1) to be reformulated in equivalent forms that admit quadratic energy functions.

In order to maintain the positivity of the energy, we modify the energy in (1.5) by adding a positive constant  $C_0$  to  $\frac{\beta}{2} \int_{\mathbb{R}^d} G(|u|^2) dx$ . This modification has no essential effect on the energy invariance of the system described in (1.5). Thus we will continue to use  $E(t)$  for the modified energy, i.e.,

$$E(t) \stackrel{\text{def}}{=} \|u_t(\cdot, t)\|^2 + \left\| (-\Delta)^{\frac{\alpha}{4}} u(\cdot, t) \right\|^2 + \frac{\beta}{2} \int_{\mathbb{R}^d} G(|u|^2) dx + C_0. \quad (2.1)$$

In this, we consider a new scalar variable

$$w(t) \stackrel{\text{def}}{=} \sqrt{H(t) + C_0} \text{ with } H(t) \stackrel{\text{def}}{=} \frac{\beta}{2} \int_{\Omega} G(|u|^2) dx = \frac{\beta}{2} \int_{\Omega} \int_0^{|u|^2} g(z) dz dx. \quad (2.2)$$

The direct calculation produces that

$$\begin{aligned} w_t &= \frac{1}{2\sqrt{H(t) + C_0}} \frac{dH(t)}{dt} \\ &= \frac{1}{2\sqrt{H(t) + C_0}} \frac{d}{dt} \int_{\Omega} G(|u|^2) dx \\ &= \frac{\beta}{\sqrt{H(t) + C_0}} \int_{\Omega} g(|u|^2) \Re(u \bar{u}_t) dx \\ &= \Re \int_{\Omega} \frac{\beta g(|u|^2) u \bar{u}_t}{\sqrt{H(t) + C_0}} dx \\ &= \Re(b(u), u_t), \end{aligned} \quad (2.3)$$

where

$$b(u) \stackrel{\text{def}}{=} \beta g(|u|^2) u / \sqrt{H(t) + C_0}, \quad (2.4)$$

$\Re$  denotes the real part and  $(\cdot, \cdot)$  means the  $L^2$ -inner product over  $\Omega$ .

As a result, the NFSWEs (1.1)–(1.3) can be equivalently rewritten as

$$u_t = v, \quad (2.5)$$

$$v_t = -(-\Delta)^{\alpha/2} u - \text{ik} v - b(u)w, \quad (2.6)$$

$$w_t = \Re(b(u), u_t), \quad (2.7)$$

with the initial conditions

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad v(\mathbf{x}, 0) = u_1(\mathbf{x}), \quad w(0) = \sqrt{\frac{\beta}{2} \int_{\mathbb{R}^d} G(|u_0|^2) dx + C_0}. \quad (2.8)$$

Taking the inner product of systems (2.5)–(2.6) respectively with  $v_t$  and  $u_t$ , multiplying (2.7) by  $w(t)$ , summing the resulting equations, and finally taking an integration over the time interval  $[0, t]$  for the real part, one immediately has the modified energy conservation law

$$E(t) = \|v(\cdot, t)\|^2 + \left\| (-\Delta)^{\frac{\alpha}{4}} u(\cdot, t) \right\|^2 + w^2(t) = E(0), \quad (2.9)$$

which is essentially consistent with the original energy-conservation law (1.5) for NFSWEs (1.1)–(1.3).

### 3. Structure-preserving spatial discretization

Without loss of generality, we assume that the spatial dimension is two (i.e.,  $d = 2$ ). The Fourier pseudo-spectral method is a well-established technique for solving partial differential equations that can provide accurate and efficient approximations. This method involves approximating the solution in the spectral domain, where the derivatives are computed using fast Fourier transforms. Therefore, we utilize the Fourier pseudo-spectral method for the spatial discretization of the equivalent systems (2.5)–(2.8).

For positive integer  $M$  and even positive integers  $N_x$  and  $N_y$ , denote  $\tau = T/M$ ,  $h_x = L/N_x$ ,  $h_y = L/N_y$ . Define  $\Omega_h = \{(x_i, y_j) \mid 0 \leq i \leq N_x - 1; 0 \leq j \leq N_y - 1\}$ ,  $\Omega_\tau = \{t_m \mid 0 \leq m \leq M\}$  and  $\Omega_h^\tau = \Omega_h \times \Omega_\tau$ , with  $t_m = m\tau$ ,  $x_i = ih_x$  and  $y_j = jh_y$ . The vector forms at any time level  $t_m$  are given by

$$U^m = \left( u_{0,0}^m, \dots, u_{N_x-1,0}^m, u_{0,1}^m, \dots, u_{N_x-1,1}^m, \dots, u_{0,N_y-1}^m, \dots, u_{N_x-1,N_y-1}^m \right)^T, \quad (3.1)$$

$$V^m = \left( v_{0,0}^m, \dots, v_{N_x-1,0}^m, v_{0,1}^m, \dots, v_{N_x-1,1}^m, \dots, v_{0,N_y-1}^m, \dots, v_{N_x-1,N_y-1}^m \right)^T, \quad (3.2)$$

$$W^m = w^m. \quad (3.3)$$

For any grid functions  $u$  and  $v$  defined on  $\Omega_h$ , we define the discrete inner product and the associated discrete norms as

$$(u, v) = h_x h_y \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} u_{i,j} \bar{v}_{i,j}, \quad \|u\| = (u, u)^{\frac{1}{2}}, \quad \|u\|_\infty = \sup_{(x_i, y_j) \in \Omega_h} |u_{i,j}|. \quad (3.4)$$

Let  $(x_i, y_j) \in \Omega_h$  be the Fourier collocation points. Denote  $u_N(x, y)$  is the interpolation polynomial function of  $u(x, y)$ , then we have

$$u_N(x, y) = \sum_{k_1=-N_x/2}^{N_x/2} \sum_{k_2=-N_y/2}^{N_y/2} \tilde{u}_{k_1, k_2} e^{i\mu(k_1(x+L)+k_2(y+L))}, \quad (3.5)$$

in which  $\mu = \pi/L$ , and the coefficients

$$\tilde{u}_{k_1, k_2} = \frac{1}{N_x c_{k_1}} \frac{1}{N_y c_{k_2}} \sum_{l_1=0}^{N_x-1} \sum_{l_2=0}^{N_y-1} u(x_{l_1}, y_{l_2}) e^{-i\mu(k_1(x_{l_1}+L)+k_2(y_{l_2}+L))}, \quad (3.6)$$

where  $c_{k_1} = 1$  for  $|k_1| < N_x/2$ ,  $c_{k_2} = 1$  for  $|k_2| < N_y/2$ ,  $c_{k_1} = 2$  for  $k_1 = \pm N_x/2$ , and  $c_{k_2} = 2$  for  $k_2 = \pm N_y/2$ .

As a result, the fractional Laplacian  $(-\Delta)^{\frac{\alpha}{2}} u(x, y)$  can be approximated by

$$(-\Delta)^{\frac{\alpha}{2}} u_N(x, y) = \sum_{k_1=-N_x/2}^{N_x/2} \sum_{k_2=-N_y/2}^{N_y/2} |(k_1\mu)^2 + (k_2\mu)^2|^{\frac{\alpha}{2}} \tilde{u}_{k_1, k_2} e^{i\mu(k_1(x+L)+k_2(y+L))}. \quad (3.7)$$

Inserting (3.6) into (3.7), and considering the resulting equation at the point  $(x_i, y_j)$  gives that

$$\begin{aligned} & (-\Delta)^{\frac{\alpha}{2}} u_N(x_i, y_j) \\ &= \sum_{l_1=0}^{N_x-1} \sum_{l_2=0}^{N_y-1} u(x_{l_1}, y_{l_2}) \left( \sum_{k_1=-N_x/2}^{N_x/2} \sum_{k_2=-N_y/2}^{N_y/2} \frac{1}{N_x c_{k_1}} \frac{1}{N_y c_{k_2}} |\mu^2 \cdot \mathbf{k}^2|^{\frac{\alpha}{2}} e^{i\mu(k_1(x_i-x_{l_1})+k_2(y_j-y_{l_2}))} \right) \\ &= (D^\alpha U)_{i+jN_x}, \end{aligned} \quad (3.8)$$

where  $\mu^2 \cdot \mathbf{k}^2 = \mu^2 (k_1^2 + k_2^2)$ ,  $D^\alpha$  is spectral symmetric differential matrix with the elements

$$(D^\alpha)_{i+jN_x, l_1+l_2N_x} = \sum_{l_1=0}^{N_x-1} \sum_{l_2=0}^{N_y-1} \frac{1}{N_x c_{k_1}} \frac{1}{N_y c_{k_2}} |\mu^2 \cdot \mathbf{k}^2|^{\frac{\alpha}{2}} e^{i\mu(k_1(x_i-x_{l_1})+k_2(y_j-y_{l_2}))}. \quad (3.9)$$

Above all, applying the Fourier pseudo-spectral method to the previous equivalent system (2.5)–(2.7) in space gives the semi-discrete system as follows

$$U_t = V, \quad (3.10)$$

$$V_t = D^\alpha U - i\kappa V - b(U) \cdot W, \quad (3.11)$$

$$W_t = \Re(b(U), U_t), \quad (3.12)$$

with the initial conditions  $U^0, V^0, W^0$ , where  $\cdot$  means the point multiplication between vectors.

For the semi-discrete system (3.10)–(3.12), we have the following theorem.

**Theorem 3.1.** *The space semi-discrete system (3.10)–(3.12) admits the semi-discrete quadratic energy conservation law*

$$\frac{dE}{dt} = 0, \quad (3.13)$$

where

$$E(U, V, W) \stackrel{\text{def}}{=} \|V\|^2 + \|D^{\frac{\alpha}{2}} U\|^2 + (W)^2. \quad (3.14)$$

**Proof.** Taking the inner product of (3.10), (3.11) with  $V_t$  and  $U_t$  respectively, and multiplying (3.12) by  $W$ , we get

$$(V_t, U_t) = (V_t, V), \quad (3.15)$$

$$(V_t, U_t) = (D^\alpha U, U_t) - i\kappa (V, U_t) - W(b(U), U_t), \quad (3.16)$$

$$WW_t = W\Re(b(U), U_t). \quad (3.17)$$

Summing the resulting equations (3.15)–(3.17), we have

$$(V_t, V) + WW_t = (D^\alpha U, U_t) - i\kappa (V, U_t) - W(b(U), U_t) + W\Re(b(U), U_t). \quad (3.18)$$

Taking the real part in above equation gives

$$\Re(V_t, V) + \Re(WW_t) = \Re(D^\alpha U, U_t) - \Re(i\kappa (V, U_t)) - W\Re(b(U), U_t) + W\Re(b(U), U_t). \quad (3.19)$$

Thus, we know that

$$\Re(V_t, V) + \Re(WW_t) = \Re(D^\alpha U, U_t). \quad (3.20)$$

Noticing that

$$\Re(V_t, V) = \frac{1}{2} \frac{d}{dt} \|V\|^2, \quad \Re(WW_t) = \frac{1}{2} \frac{d}{dt} (W)^2, \quad \Re(D^\alpha U, U_t) = -\frac{1}{2} \frac{d}{dt} \|D^{\frac{\alpha}{2}} U\|^2, \quad (3.21)$$

and substituting (3.21) into (3.20) leads to

$$\frac{d}{dt} \left( \|V\|^2 + \|D^{\frac{\alpha}{2}} U\|^2 + (W)^2 \right) = 0. \quad (3.22)$$

This completes the proof.  $\square$

#### 4. Conservative explicit SAV-RRK methods

This section is devoted to present the invariant-conserving explicit RRK methods for the NFSWEs based on the SAV approach. We begin by recalling RRK methods and discussing their structure-preserving properties.

##### 4.1. Explicit SAV-RRK schemes

For brevity, we introduce following symbols

$$\mathbf{y} = (U, V, W)^T, \mathbf{y}_0 = (U^0, V^0, W^0)^T, \mathbf{f} = (f^1, f^2, f^3)^T \stackrel{\text{def}}{=} (V, D^\alpha U - i\kappa V - b(U) \cdot W, \Re(b(U), U_t))^T. \quad (4.1)$$

Then the semi-discrete system (3.10)–(3.12) based on SAV approach can be rewritten as

$$\begin{cases} \mathbf{y}_t = \mathbf{f}(\mathbf{y}), & t \in (0, T], \\ \mathbf{y}(0) = \mathbf{y}_0. \end{cases} \quad (4.2)$$

Let  $\mathbf{y}^m$  be the approximation to  $\mathbf{y}(t_m)$ . A  $s$ -stage explicit RK method [8] applied to (4.2) takes the form

$$\begin{cases} \mathbf{Y}_{mi} = \mathbf{y}^m + \tau \sum_{j=1}^{i-1} a_{ij} \mathbf{f}_{mj}, & i = 1, \dots, s, \\ \mathbf{y}^{m+1} = \mathbf{y}^m + \tau \sum_{i=1}^s b_i \mathbf{f}_{mi}, \end{cases} \quad (4.3)$$

where  $\mathbf{f}_{mj} = \mathbf{f}(\mathbf{Y}_{mj})$ ,  $j = 1, \dots, s$ . With the following notations

$$\begin{aligned} A &= (a_{ij})_{s \times s}, \quad a_{ij} = 0 \quad \text{for } j \geq i, \\ b &= (b_1, \dots, b_s)^T, \end{aligned} \quad (4.4)$$

the  $s$ -stage explicit SAV-RK methods can be represented by a Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \quad (4.5)$$

where the abscissa vectors  $c = (c_1, c_2, \dots, c_s)$  satisfy  $c_i = \sum_{j=1}^s a_{ij}$ ,  $i = 1, \dots, s$ .

However, it is well known that only special implicit RK methods may be symplectic or algebraically stable, and there are no symplectic or algebraically stable explicit RK methods. This leads to the fact that the explicit SAV-RK methods fail to preserve the underlying conservation property of the original problem (4.2). Thus, we introduce the explicit SAV-RRK method. Concretely, consider a single step over the interval  $[\hat{t}_m, \hat{t}_{m+1}]$ ,  $m \geq 0$ , and let  $\mathbf{y}_\gamma^m = (U_\gamma^m, V_\gamma^m, W_\gamma^m)^T$  be the numerical approximation to  $\mathbf{y}(\hat{t}_m)$ , then a  $s$ -stage explicit SAV-RRK method for (4.2) is defined as

$$\begin{cases} \mathbf{Y}_{mi} = \mathbf{y}_\gamma^m + \tau \sum_{j=1}^{i-1} a_{ij} \mathbf{f}_{mj}, & i = 1, \dots, s, \\ \mathbf{y}_\gamma^{m+1} = \mathbf{y}_\gamma^m + \gamma_m \tau \sum_{i=1}^s b_i \mathbf{f}_{mi}. \end{cases} \quad (4.6)$$

Similarly, the  $s$ -stage RRK methods (4.6) can be represented by a Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & \tilde{b}^T \end{array} \quad (4.7)$$

where  $\tilde{b} = \gamma_m b$  and  $\gamma_m \neq 0$  satisfies

$$\begin{cases} E_\gamma^{m+1} = E_\gamma^m, & \text{if } \sum_{i=1}^s b_i \mathbf{f}_{mi} \neq 0, \\ \gamma_m = 1, & \text{if } \sum_{i=1}^s b_i \mathbf{f}_{mi} = 0, \end{cases} \quad (4.8)$$

where

$$E_\gamma^m = \|V_\gamma^{m+1}\|^2 + \|D^{\frac{\alpha}{2}} U_\gamma^m\|^2 + (W_\gamma^m)^2. \quad (4.9)$$

An outstanding advantage of the explicit SAV-RRK method (4.6) is that one can calculate explicitly the parameter  $\gamma_m$ . In fact, one knows from (4.8) that  $\gamma_m = 1$  for  $\sum_{i=1}^s b_i \mathbf{f}_{mi} = 0$ . When  $\sum_{i=1}^s b_i \mathbf{f}_{mi} \neq 0$ , by a straightforward calculation one can get that

$$\begin{aligned} E_\gamma^{m+1} &= \|V_\gamma^{m+1}\|^2 + \|D^{\frac{\alpha}{2}} U_\gamma^{m+1}\|^2 + (W_\gamma^{m+1})^2 \\ &= \|V_\gamma^{m+1}\|^2 - (U_\gamma^{m+1}, D^\alpha U_\gamma^{m+1}) + (W_\gamma^{m+1})^2 \end{aligned}$$

$$\begin{aligned}
&= \|V_\gamma^m + \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^2\|^2 + (W_\gamma^m + \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^3)^2 \\
&\quad - (U_\gamma^m + \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^1, D^\alpha(U_\gamma^m + \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^1)) \\
&= \|V_\gamma^m\|^2 + \|D^{\frac{\alpha}{2}} U_\gamma^m\|^2 + (W_\gamma^m)^2 \\
&\quad + (V_\gamma^m, \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^2) + (\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^2, V_\gamma^m) + \|\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^2\|^2 \\
&\quad - (U_\gamma^m, D^\alpha(\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^1)) - (\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^1, D^\alpha U_\gamma^m) \\
&\quad + \|D^{\frac{\alpha}{2}}(\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^1)\|^2 + 2W_\gamma^m \gamma_m \tau \sum_{i=1}^s b_i f_{mi}^3 + (\gamma_m \tau \sum_{i=1}^s b_i f_{mi}^3)^2 \\
&= E_\gamma^m + \gamma_m \tau [(V_\gamma^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_\gamma^m) \\
&\quad - (U_\gamma^m, D^\alpha(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^\alpha U_\gamma^m) + 2W_\gamma^m \sum_{i=1}^s b_i f_{mi}^3] \\
&\quad + \gamma_m^2 \tau^2 [\|\sum_{i=1}^s b_i f_{mi}^2\|^2 + \|D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1)\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2]. \tag{4.10}
\end{aligned}$$

This together with (4.8) gives that

$$\begin{aligned}
&\gamma_m \tau [(V_\gamma^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_\gamma^m) - (U_\gamma^m, D^\alpha(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^\alpha U_\gamma^m) \\
&\quad + 2W_\gamma^m \sum_{i=1}^s b_i f_{mi}^3] + \gamma_m^2 \tau^2 [\|\sum_{i=1}^s b_i f_{mi}^2\|^2 + \|D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1)\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2] = 0,
\end{aligned}$$

which is a quadratic algebraic equation for  $\gamma_m$ . Noting that  $\gamma_m \neq 0$ , thus

$$\gamma_m = \frac{(V_\gamma^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_\gamma^m) - (U_\gamma^m, D^\alpha(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^\alpha U_\gamma^m) + 2W_\gamma^m \sum_{i=1}^s b_i f_{mi}^3}{-\tau [\|\sum_{i=1}^s b_i f_{mi}^2\|^2 + \|D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1)\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2]}. \tag{4.11}$$

This means that the value of  $\gamma_m$  can also be explicitly determined by (4.11) when  $\sum_{i=1}^s b_i f_{mi} \neq 0$ .

Actually, the value of  $\gamma_m$  defined in (4.8) is close to 1 as  $\tau \rightarrow 0$ , which will be further discussed in Section 5.1. Therefore, the explicit SAV-RRK methods (4.6) are well-defined. Moreover, it can be proved that the proposed methods preserve the energy invariance described in the following theorem.

**Theorem 4.1.** Suppose the order of explicit SAV-RK methods (4.3) is at least two, then the solution to explicit SAV-RRK methods (4.6) satisfies

$$E_\gamma^{m+1} = E_\gamma^m, \tag{4.12}$$

where  $E_\gamma^m$  is fined by (4.9).

**Proof.** When  $\sum_{i=1}^s b_i f_{mi} = 0$ , it follows from the second equation in (4.6) that  $y_\gamma^{m+1} \equiv y_\gamma^m$ , which automatically satisfies (4.12). For the case of  $\sum_{i=1}^s b_i f_{mi} \neq 0$ , we can also get (4.12) from the condition (4.8). This proof is completed.  $\square$

## 5. Accuracy of the explicit SAV-RRK methods

In this section, we will discuss the accuracy of the explicit SAV-RRK methods (4.6). To this end, we first estimate the relaxation coefficient  $\gamma_m$ , which plays a key role in the next analysis. Since the relaxation coefficient  $\gamma_m$  tends to vary from step to step, it can be estimated in the similar way over different steps. From this, here we just consider a single step over the interval  $[\hat{t}_m, \hat{t}_{m+1}]$ .

### 5.1. Estimate of the relaxation coefficient

We will only focus on the case of  $\sum_{i=1}^s b_i f_{mi} \neq 0$ , because the converse is simple. Let

$$\begin{aligned} S_m(\gamma) = & \|V_\gamma^m + \gamma\tau \sum_{i=1}^s b_i f_{mi}^2\|^2 + \|D^{\frac{\alpha}{2}}(U_\gamma^m + \gamma\tau \sum_{i=1}^s b_i f_{mi}^1)\|^2 \\ & + (W_\gamma^m + \gamma\tau \sum_{i=1}^s b_i f_{mi}^3)^2 - \|V_\gamma^m\|^2 - \|D^{\frac{\alpha}{2}}U_\gamma^m\|^2 - (W_\gamma^m)^2, \end{aligned} \quad (5.1)$$

then the value of  $\gamma_m$  defined in (4.8) is exactly the nonzero root of the function  $S_m(\gamma)$ .

Moreover, we have following results on  $S_m(\gamma)$ .

**Lemma 5.1.** Suppose the order of explicit SAV-RK methods (4.3) is  $p$ , then it holds that

$$S_m(1) = \mathcal{O}(\tau^{p+1}). \quad (5.2)$$

**Proof.** Similar to the proof of Lemma 3.1 in [11], we consider the initial value problem

$$\begin{cases} \tilde{y}' = f(\tilde{y}), & t \geq \hat{t}_m, \\ \tilde{y}(\hat{t}_m) = y_\gamma^m, \end{cases} \quad (5.3)$$

where  $y_\gamma^m$  is the solution to explicit SAV-RRK methods (4.6). By performing a single step using the explicit SAV-RRK methods (4.6), we arrive at the numerical solution  $y_\gamma^{m+1}$ . It follows from (5.1) that

$$\begin{aligned} S_m(\gamma) = & \|V_\gamma^{m+1}\|^2 + \|D^{\frac{\alpha}{2}}U_\gamma^{m+1}\|^2 + (W_\gamma^{m+1})^2 - \|V_\gamma^m\|^2 - \|D^{\frac{\alpha}{2}}U_\gamma^m\|^2 - (W_\gamma^m)^2 \\ = & \|V_\gamma^{m+1}\|^2 + \|D^{\frac{\alpha}{2}}U_\gamma^{m+1}\|^2 + (W_\gamma^{m+1})^2 - \|\tilde{V}(\hat{t}_m)\|^2 - \|D^{\frac{\alpha}{2}}\tilde{U}(\hat{t}_m)\|^2 - (\tilde{W}(\hat{t}_m))^2. \end{aligned} \quad (5.4)$$

From  $y_\gamma^m$ , the explicit SAV-RK methods (4.3) of order  $p$  generates the numerical solution  $\tilde{y}^{m+1}$ , which satisfies  $\tilde{y}^{m+1} = \tilde{y}(\hat{t}_m + \tau) + \mathcal{O}(\tau^{p+1})$  for sufficiently small  $\tau$ . It is crucial to observe that the explicit SAV-RRK methods defined by (4.6) will reduce to the explicit SAV-RK methods described by (4.3) when  $\gamma = 1$ . That is  $y_\gamma^{m+1} = \tilde{y}^{m+1}$ . Consequently, we can derive that

$$\begin{aligned} S_m(1) = & \|\tilde{V}^{m+1}\|^2 + \|D^{\frac{\alpha}{2}}\tilde{U}^{m+1}\|^2 + (\tilde{W}^{m+1})^2 - \|\tilde{V}(\hat{t}_m)\|^2 - \|D^{\frac{\alpha}{2}}\tilde{U}(\hat{t}_m)\|^2 - (\tilde{W}(\hat{t}_m))^2 \\ = & \|\tilde{V}^{m+1}\|^2 - \|\tilde{V}(\hat{t}_m + \tau)\|^2 + \|\tilde{V}(\hat{t}_m + \tau)\|^2 - \|\tilde{V}(\hat{t}_m)\|^2 \\ & + \|D^{\frac{\alpha}{2}}\tilde{U}^{m+1}\|^2 - \|D^{\frac{\alpha}{2}}\tilde{U}(\hat{t}_m + \tau)\|^2 + \|D^{\frac{\alpha}{2}}\tilde{U}(\hat{t}_m + \tau)\|^2 - \|D^{\frac{\alpha}{2}}\tilde{U}(\hat{t}_m)\|^2 \\ & + (\tilde{W}^{m+1})^2 - (\tilde{W}(\hat{t}_m + \tau))^2 + (\tilde{W}(\hat{t}_m + \tau))^2 - (\tilde{W}(\hat{t}_m))^2 \\ = & \mathcal{O}(\tau^{p+1}) + 2 \int_{\hat{t}_m}^{\hat{t}_m + \tau} [\Re(V_t, V) + \Re(W W_t) - \Re(D^\alpha U, U_t)] dt. \end{aligned} \quad (5.5)$$

By combining (5.5) and (3.20), this proof is complete.  $\square$

Based on Lemma 5.1, we can get the following estimation for the relaxation coefficient  $\gamma_m$ .

**Theorem 5.2.** Suppose the order of explicit SAV-RK methods (4.3) is  $p$  ( $p \geq 2$ ), then the relaxation coefficient  $\gamma_m$  defined in (4.8) satisfies

$$\gamma_m = 1 + \mathcal{O}(\tau^{p-1}). \quad (5.6)$$



**Proof.** Let us consider two cases based on the value of  $\sum_{i=1}^s b_i f_{mi}$ .

Case 1: According the definition of  $\gamma_m$  in (4.8), we have  $\gamma_m = 1$  when  $\sum_{i=1}^s b_i f_{mi} = 0$ , which automatically satisfies (5.6).

Case 2: When  $\sum_{i=1}^s b_i f_{mi} \neq 0$ , it follows from (4.10) that

$$\begin{aligned} S_m(\gamma) &= E_{\gamma}^{m+1} - E_{\gamma}^m \\ &= \gamma \tau \left[ (V_{\gamma}^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_{\gamma}^m) - (U_{\gamma}^m, D^{\alpha}(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^{\alpha} U_{\gamma}^m) \right. \\ &\quad \left. + 2W_{\gamma}^m \sum_{i=1}^s b_i f_{mi}^3 \right] + \gamma^2 \tau^2 \left[ \left\| \sum_{i=1}^s b_i f_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2 \right]. \end{aligned} \quad (5.7)$$

Noting that  $S_m(\gamma)$  is a quadratic function of  $\gamma$ , it has non-zero root

$$\gamma = \frac{\tau \left[ (V_{\gamma}^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_{\gamma}^m) - (U_{\gamma}^m, D^{\alpha}(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^{\alpha} U_{\gamma}^m) + 2W_{\gamma}^m \sum_{i=1}^s b_i f_{mi}^3 \right]}{-\tau^2 \left[ \left\| \sum_{i=1}^s b_i f_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2 \right]}. \quad (5.8)$$

Based on Lemma 5.1, we have

$$\begin{aligned} S_m(1) &= \tau \left[ (V_{\gamma}^m, \sum_{i=1}^s b_i f_{mi}^2) + (\sum_{i=1}^s b_i f_{mi}^2, V_{\gamma}^m) - (U_{\gamma}^m, D^{\alpha}(\sum_{i=1}^s b_i f_{mi}^1)) - (\sum_{i=1}^s b_i f_{mi}^1, D^{\alpha} U_{\gamma}^m) \right. \\ &\quad \left. + 2W_{\gamma}^m \sum_{i=1}^s b_i f_{mi}^3 \right] + \tau^2 \left[ \left\| \sum_{i=1}^s b_i f_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2 \right] \\ &= \mathcal{O}(\tau^{p+1}). \end{aligned} \quad (5.9)$$

Defining  $\tilde{f}_m = f(y(\tilde{t}_m))$ , then we have  $f_{mi} = \tilde{f}_m + \mathcal{O}(\tau)$  as  $\tau \rightarrow 0$ . Thus, it holds that

$$\begin{aligned} &\tau^2 \left[ \left\| \sum_{i=1}^s b_i f_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2 \right] \\ &= \tau^2 \left[ \left\| \sum_{i=1}^s b_i \tilde{f}_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i \tilde{f}_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i \tilde{f}_{mi}^3)^2 \right] + \mathcal{O}(\tau^3) \\ &= \mathcal{O}(\tau^2). \end{aligned} \quad (5.10)$$

Substituting (5.10) and (5.9) into (5.8) yields that

$$\begin{aligned} \gamma &= 1 + \frac{\mathcal{O}(\tau^{p+1})}{-\tau^2 \left[ \left\| \sum_{i=1}^s b_i f_{mi}^2 \right\|^2 + \left\| D^{\frac{\alpha}{2}}(\sum_{i=1}^s b_i f_{mi}^1) \right\|^2 + (\sum_{i=1}^s b_i f_{mi}^3)^2 \right]} \\ &= 1 + \frac{\mathcal{O}(\tau^{p+1})}{\mathcal{O}(\tau^2)} \\ &= 1 + \mathcal{O}(\tau^{p-1}). \end{aligned} \quad (5.11)$$

This proof is completed.  $\square$

**Remark 5.3.** Theorem 5.2 shows that the relaxation coefficient  $\gamma_m$  is close to 1 as  $\tau \rightarrow 0$ , which implies that the relaxation methods can be considered as a small perturbation of the standard methods.

## 5.2. Truncation errors of the explicit SAV-RRK methods

Now we further study the accuracy of the explicit SAV-RRK methods (4.6) based on the estimates on the relaxation coefficient  $\gamma_m$  given by Theorem 5.2.

Before analyzing the accuracy of the methods, we rewrite the explicit SAV-RRK methods (4.6) into

$$\begin{cases} Y_{mi} = y_{\gamma}^m + \tau \sum_{j=1}^{i-1} a_{ij} f_{mj}, & i = 1, \dots, s, \\ y^{m+1} = y_{\gamma}^m + \tau \sum_{i=1}^s b_i f_{mi}, \\ y_{\gamma}^{m+1} = y^{m+1} + (\gamma_m - 1) (y^{m+1} - y_{\gamma}^m). \end{cases} \quad (5.12)$$

The first two equations essentially form the standard explicit SAV-RK methods.

Similar to the discussion in [10], the numerical solution  $y_{\gamma}^{m+1}$  at  $\hat{t}_{m+1}$  can be interpreted in two ways, namely the Incremental Direction Technique (IDT) and the Relaxation Technique (RT).

1. IDT's angle:  $y_{\gamma}^{m+1}$  is considered as an approximation to  $y(\hat{t}_m + \tau)$ , where  $\hat{t}_m = t_m$  for  $m \geq 0$ .
2. RT's angle:  $y_{\gamma}^{m+1}$  is viewed as an approximation to  $y(\hat{t}_m + \gamma_m \tau)$ , and  $\hat{t}_m$  may not be equal to  $t_m$  when  $m > 0$ .

It is important to note that different interpretations of the numerical solution  $y_{\gamma}^{m+1}$  can yield different convergence results. Following the idea of the proof of Lemma 2.7 in [14], we have obtained the following results.

**Theorem 5.4.** Suppose the order of explicit SAV-RK methods (4.3) is  $p$  ( $p \geq 2$ ). Then the following statements hold:

- For IDT, the explicit SAV-RRK methods (4.6) have order  $p - 1$ .
- For RT, the explicit SAV-RRK methods (4.6) have order  $p$ .

**Proof.** Let  $\hat{t}_{mi} = \hat{t}_m + c_i \tau$ ,  $i = 1, \dots, s$ , then one has from the second equation in (5.12) that

$$y(\hat{t}_m + \tau) = y(\hat{t}_m) + \tau \sum_{i=1}^s b_i f(y(\hat{t}_{mi})) + \mathcal{O}(\tau^{p+1}). \quad (5.13)$$

Now we estimate the truncation error of the third equation in (5.12). Denote

$$\phi_m(y(t)) \stackrel{\text{def}}{=} y(\hat{t}_m) + \tau \sum_{i=1}^s b_i f(y(\hat{t}_{mi})),$$

and

$$y(\hat{t}_{m+1}) = \phi_m(y(t)) + (\gamma_m - 1) (\phi_m(y(t)) - y(\hat{t}_m)) + T_{m+1}. \quad (5.14)$$

Using the fact that  $\gamma_m = 1 + \mathcal{O}(\tau^{p-1})$  given in Theorem 5.2, one can obtain from (5.13) that

$$\begin{aligned} T_{m+1} &= y(\hat{t}_{m+1}) - \phi_m(y(t)) - (\gamma_m - 1) (\phi_m(y(t)) - y(\hat{t}_m)) \\ &= y(\hat{t}_{m+1}) - (y(\hat{t}_m + \tau) - \mathcal{O}(\tau^{p+1})) - (\gamma_m - 1) (y(\hat{t}_m + \tau) - \mathcal{O}(\tau^{p+1}) - y(\hat{t}_m)) \\ &= y(\hat{t}_{m+1}) - y(\hat{t}_m + \tau) - (\gamma_m - 1) (y(\hat{t}_m + \tau) - y(\hat{t}_m)) + \mathcal{O}(\tau^{p+1}) + \mathcal{O}((\gamma_m - 1)\tau^{p+1}) \\ &= y(\hat{t}_{m+1}) - y(\hat{t}_m + \tau) - (\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}((\gamma_m - 1)\tau^2) + \mathcal{O}(\tau^{p+1}) \\ &= y(\hat{t}_{m+1}) - y(\hat{t}_m + \tau) - (\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}(\tau^{p+1}). \end{aligned} \quad (5.15)$$

For IDT,  $y_{\gamma}^{m+1}$  is viewed as the approximation at  $\hat{t}_m + \tau$ , so  $\hat{t}_{m+1} = \hat{t}_m + \tau$ , which implies that  $T_{m+1} = -(\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}(\tau^{p+1}) = \mathcal{O}(\tau^p)$ . That is, this method (4.6) has order of  $p - 1$ .

For RT, since  $y_{\gamma}^{m+1}$  is viewed as the approximation at  $\hat{t}_m + \gamma_m \tau$ ,  $\hat{t}_{m+1} = \hat{t}_m + \gamma_m \tau$ . Using the Taylor expansion in the above estimate again, one gets

$$\begin{aligned} T_{m+1} &= y(\hat{t}_m + \tau + (\gamma_m - 1) \tau) - y(\hat{t}_m + \tau) - (\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}(\tau^{p+1}) \\ &= y(\hat{t}_m + \tau) + (\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}((\gamma_m - 1)^2 \tau^2) \\ &\quad - y(\hat{t}_m + \tau) - (\gamma_m - 1) y'(\hat{t}_m + \tau) \tau + \mathcal{O}(\tau^{p+1}) \\ &= \mathcal{O}(\tau^{p+1}), \end{aligned} \quad (5.16)$$

which implies that this method (4.6) has order  $p$ .  $\square$

**Remark 5.5.** In calculation, the order of IDT method might be higher than what we expect and even reaches the order of RT method, which can be observed in [Example 6.1](#).

**Remark 5.6.** The methods proposed in this paper can be easily applied to other similar fractional equations, such as the nonlinear fractional wave equation, and the fractional Klein–Gordon–Schrödinger equation.

## 6. Numerical examples

In this section, we carry out several numerical experiments to illustrate the accuracy and conservation property of the explicit SAV-RRK methods (4.6). Also some comparisons with other classical explicit SAV-RK methods are supported.

Without loss of generality, the following explicit RK methods will be used in calculation.

1. RK(2, 2) : two-stage, second order Heun method

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{4} & \frac{3}{4} \end{array} \quad (6.1)$$

2. RK(3, 3) : three-stage, third order Heun method

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ \hline \frac{2}{3} & \frac{1}{4} & 0 & \frac{3}{4} \end{array} \quad (6.2)$$

3. RK(4, 4) : four-stage, fourth order Gill method

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{2}-1}{2} & 1 - \frac{\sqrt{2}}{2} & 0 & 0 \\ 1 & 0 & -\frac{\sqrt{2}}{2} & 1 + \frac{\sqrt{2}}{2} & 0 \\ \hline & \frac{1}{6} & \frac{2-\sqrt{2}}{6} & \frac{2+\sqrt{2}}{6} & \frac{1}{6} \end{array} \quad (6.3)$$

It is worth noting that the main focus of this work is accuracy in temporal. Therefore, the spectral accuracy phenomenon presented in spatial is not shown. Additionally, given the inherent positivity of the energy in our numerical examples, we simply set the constant  $C_0 = 0$  in calculation.

The errors and the convergence order in temporal are calculated by

$$Error(\tau) = \|U_N^M - U_N^{2M}\|_{\infty}, \quad \text{order} = \log_2[Error(\tau)/Error(\tau/2)]. \quad (6.4)$$

In order to depict the conservation performance, the relative error of energy is calculated by

$$RE_{\gamma}^m = |(E_{\gamma}^m - E_{\gamma}^0)/E_{\gamma}^0|, \quad (6.5)$$

where  $E_{\gamma}^m$  denotes the discrete energy at  $t_m$ .

**Example 6.1** ([13]). First, we consider 1D NFSWEs (1.1)–(1.3) as

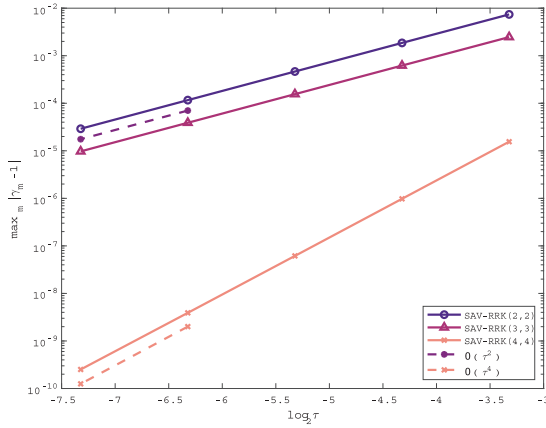
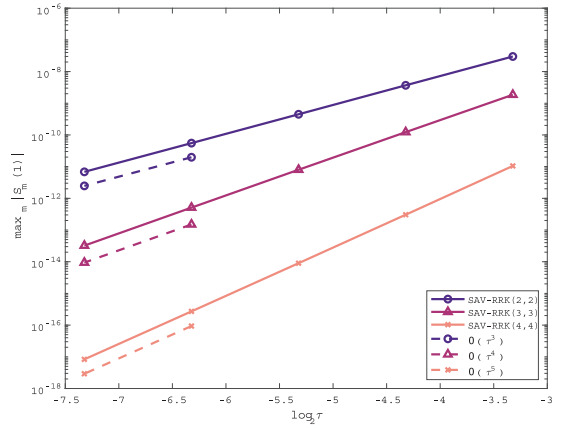
$$u_{tt} + (-\Delta)^{\alpha/2} u + iu_t + |u|^2 u = 0, \quad (x, t) \in (-25, 25) \times [0, T], \quad (6.6)$$

with  $u(x, 0) = (1 + i)xe^{-10(1-x)^2}$  and  $u_t(x, 0) = 0$ .

Without loss of generality, here we set  $\alpha = 1.5$  and  $T = 1$  to confirm the accuracy of the explicit SAV-RRK methods (4.6). The errors and convergence order in time for the standard SAV-RK, SAV-RRK(RT) and SAV-RRK(IDT) methods are listed in [Table 1](#). We can see that all SAV-RRK(RT) methods keep the same convergence order as the standard SAV-RK methods while the SAV-RRK(IDT) methods behave differently. The SAV-RRK(IDT)(2,2) and SAV-RRK(IDT)(4,4) give convergence order which are higher than the expected orders as indicated in [Theorem 5.4](#) while SAV-RRK(IDT)(3,3) gives convergence order reduced by one. This different behavior could be attributed to

**Table 1**Numerical errors and convergence order in time for [Example 6.1](#) when  $N = 32, T = 1$ .

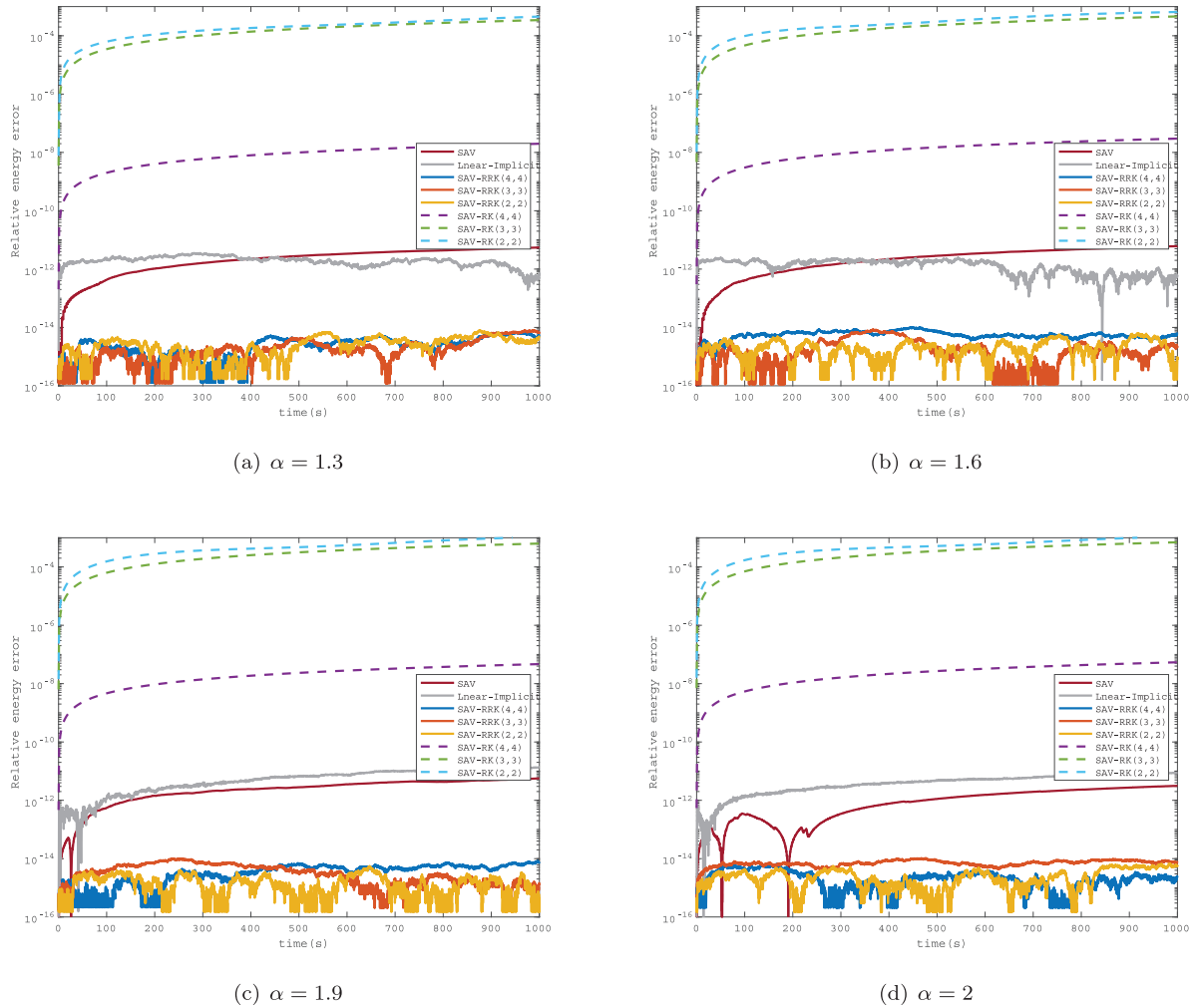
RK(Stage,Order)	$\tau$	SAV-RK		SAV-RRK(RT)		SAV-RRK(IDT)	
		Error ( $\tau$ )	Order	Error ( $\tau$ )	Order	Error ( $\tau$ )	Order
RK(2,2)	0.1	1.8552E-04	—	1.9063E-04	—	2.0325E-04	—
	0.05	4.6601E-05	1.9931	4.7240E-05	2.0126	5.0585E-05	2.0065
	0.025	1.1671E-05	1.9974	1.1750E-05	2.0074	1.2387E-05	2.0298
	0.0125	2.9200E-06	1.9990	2.9294E-06	2.0040	2.9549E-06	2.0677
	0.00625	7.3022E-07	1.9996	7.3130E-07	2.0021	6.6665E-07	2.1481
RK(3,3)	0.1	7.6862E-06	—	2.9857E-06	—	1.7245E-04	—
	0.05	9.5269E-07	3.0122	3.9037E-07	2.9352	4.3389E-05	1.9907
	0.025	1.1866E-07	3.0052	5.0077E-08	2.9626	1.0873E-05	1.9966
	0.0125	1.4809E-08	3.0023	6.3449E-09	2.9805	2.7208E-06	1.9986
	0.00625	1.8497E-09	3.0011	7.9858E-10	2.9901	6.8049E-07	1.9994
RK(4,4)	0.1	3.7701E-07	—	3.8894E-07	—	7.0733E-07	—
	0.05	2.3572E-08	3.9994	2.3939E-08	4.0221	4.4585E-08	3.9878
	0.025	1.4730E-09	4.0003	1.4843E-09	4.0115	2.8080E-09	3.9889
	0.0125	9.2041E-11	4.0003	9.2397E-11	4.0058	1.7743E-10	3.9842
	0.00625	5.7518E-12	4.0002	5.7630E-12	4.0029	1.1309E-11	3.9718

(a)  $\max_m |\gamma_m - 1|$ (b)  $\max_m |S_m(1)|$ **Fig. 1.**  $\max_m |\gamma_m - 1|$  and  $\max_m |S_m(1)|$  for some relaxation methods in [Example 6.1](#).

the convergence order of  $\max_m |\gamma_m - 1|$  get improved by one order for SAV-RRK(RT)(2, 2) and SAV-RRK(RT)(4, 4) as displayed in [Fig. 1](#).

Furthermore, we conduct an investigation into the relaxation coefficient  $\gamma_m$  at each step. We calculate  $\max_m |\gamma_m - 1|$  and  $\max_m |S_m(1)|$  with different time step  $\tau$ . The numerical results for some relaxation methods are displayed in [Fig. 1](#), from which one can see that the orders of the above two quantities are consistent with the theoretical results. This confirms the theoretical results in [Lemma 5.1](#) and [Theorem 5.2](#). It is worth noting that some similar results can be obtained when varying the value of  $\alpha$ , although these phenomena are not shown here.

We also run a long time simulation till  $T = 1000$  and plot the relative energy by the SAV-RRK(4,4) methods with different  $\alpha$  in [Fig. 2](#), which indicates that the proposed schemes can preserve the energy exactly in discrete level and the conservation performance is significantly better than the SAV method [\[4\]](#) and three-level linearly implicit difference method [\[16\]](#).



**Fig. 2.** Relative errors of energy with  $N = 32$ ,  $\tau = 0.01$  for different  $\alpha$  in Example 6.1.

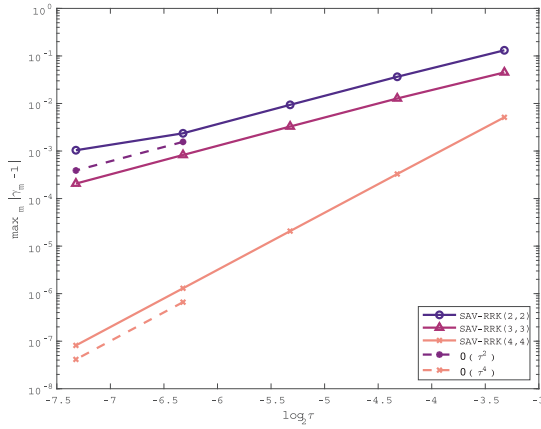
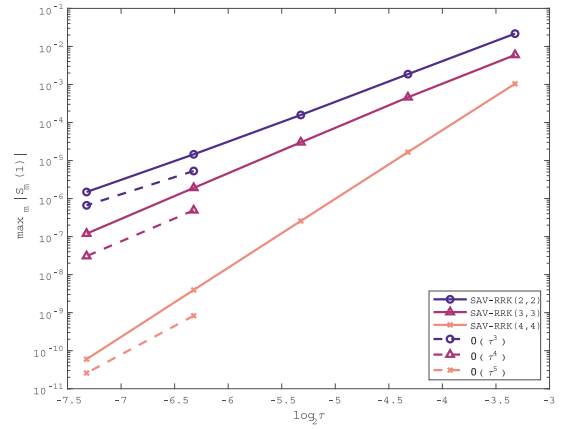
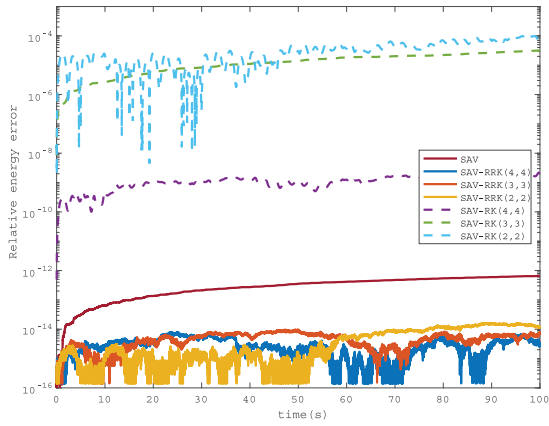
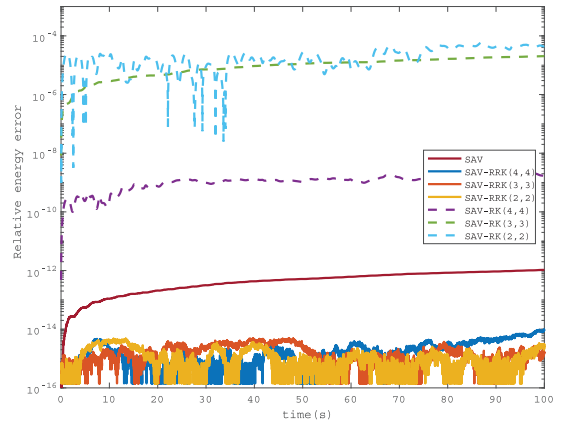
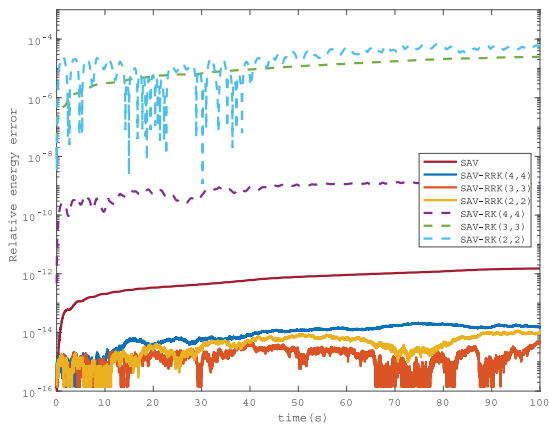
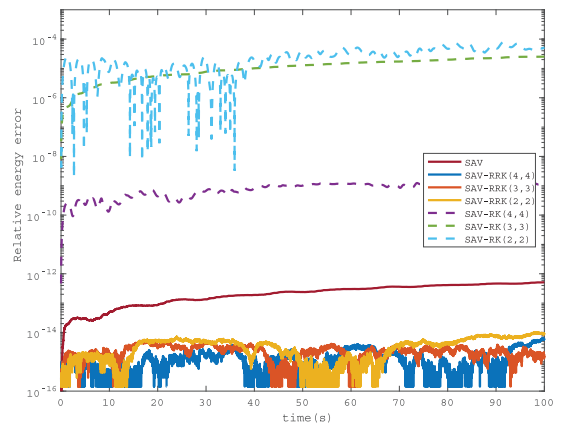
**Example 6.2.** Now we consider the 2D NFSWEs (1.1) with initial values

$$u(x, y, 0) = \text{sech}(x^2 + y^2), u_t(x, y, 0) = \sin(x + y) \text{sech}(-2(x^2 + y^2)), (x, y, t) \in \Omega \times [0, T],$$

where  $\Omega = [-5, 5] \times [-5, 5]$ .

In order to reconfirm the applicability of the theoretical results in 2D case, the same arguments as in Example 6.1 are considered in the calculation. That is,  $\alpha = 1.5$  and  $T = 1$ . The obtained convergence order in temporal and the relaxation coefficient  $\gamma$  are presented in Table 2 and Fig. 3, respectively. We can observe that the only difference from the 1D case is that the SAV-RRK(IDT) methods do not achieve a higher-than-expected order in the 2D case.

To demonstrate the effectiveness of the proposed methods in preserving energy, we perform long time simulations until  $T = 100$  and plot the relative energy errors using the SAV-RRK(4,4) methods with different  $\alpha$  in Fig. 4. The results reveal that the proposed methods can preserve energy exactly at the discrete level, and the conservation performance is significantly better than that of the SAV method [4].

(a)  $\max_m |\gamma_m - 1|$ (b)  $\max_m |S_m(1)|$ **Fig. 3.**  $\max_m |\gamma_m - 1|$  and  $\max_m |S_m(1)|$  for some relaxation (RT) methods in Example 6.2.(a)  $\alpha = 1.3$ (b)  $\alpha = 1.6$ (c)  $\alpha = 1.9$ (d)  $\alpha = 2$ **Fig. 4.** Relative errors of energy with  $N = 4$ ,  $\tau = 0.01$  for different  $\alpha$  in Example 6.2.

**Table 2**Numerical errors and convergence order in time for [Example 6.2](#) when  $N = 4, T = 1$ .

RK(Stage,Order)	$\tau$	SAV-RK		SAV-RRK		SAV-RRK(IDT)	
		Error ( $\tau$ )	Order	Error ( $\tau$ )	Order	Error ( $\tau$ )	Order
RK(2,2)	0.1	3.0217E-03	–	3.0102E-03	–	1.5692E-02	–
	0.05	7.4615E-04	2.0178	7.4702E-04	2.0106	9.6213E-03	0.7057
	0.025	1.8513E-04	2.0109	1.8587E-04	2.0069	5.2472E-03	0.8747
	0.0125	4.6090E-05	2.0060	4.6341E-05	2.0039	2.7312E-03	0.9420
	0.00625	1.1497E-05	2.0032	1.1569E-05	2.0021	1.3923E-03	0.9721
RK(3,3)	0.1	1.2581E-04	–	3.9379E-05	–	3.2535E-03	–
	0.05	1.6180E-05	2.9589	5.5726E-06	2.8210	7.9304E-04	2.0365
	0.025	2.0532E-06	2.9783	7.4443E-07	2.9041	1.9546E-04	2.0205
	0.0125	2.5863E-07	2.9889	9.6210E-08	2.9519	4.8500E-05	2.0108
	0.00625	3.2454E-08	2.9944	1.2228E-08	2.9760	1.2078E-05	2.0056
RK(4,4)	0.1	7.9185E-06	–	8.0508E-06	–	3.4013E-05	–
	0.05	4.9103E-07	4.0113	4.9644E-07	4.0194	3.3898E-06	3.3268
	0.025	3.0531E-08	4.0075	3.0805E-08	4.0104	3.6901E-07	3.1995
	0.0125	1.9026E-09	4.0042	1.9182E-09	4.0054	4.2681E-08	3.1120
	0.00625	1.1873E-10	4.0022	1.1966E-10	4.0027	5.1191E-09	3.0596

**Table 3**Numerical errors and convergence order in time for [Example 6.3](#) when  $N = 4, T = 1$ .

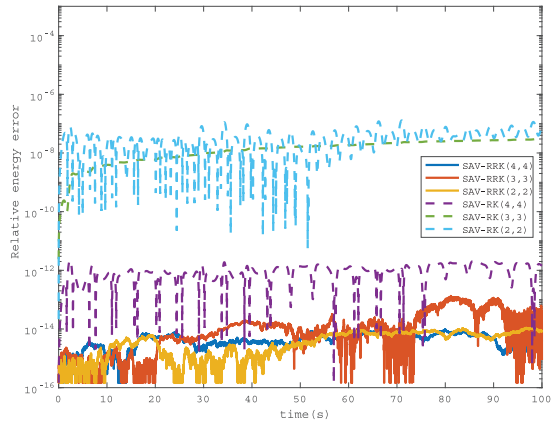
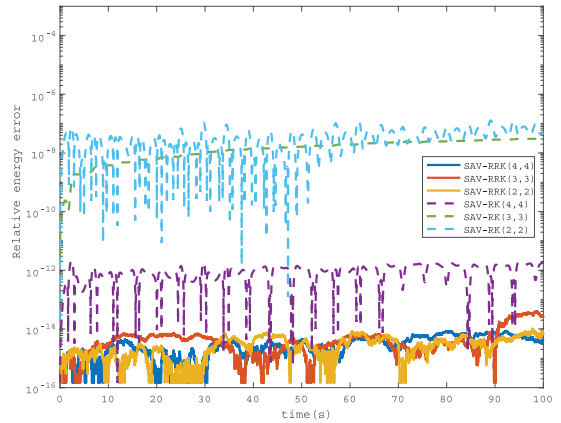
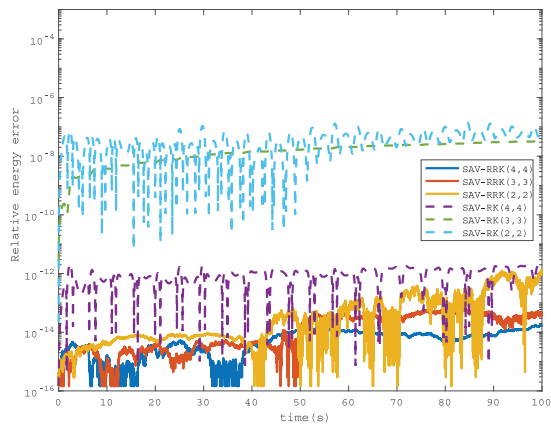
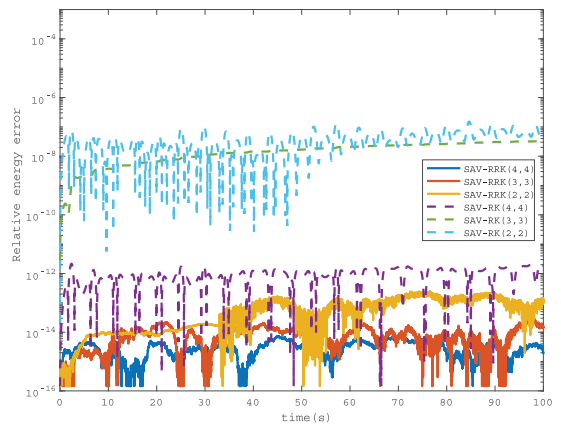
RK(Stage,Order)	$\tau$	SAV-RK		SAV-RRK(RT)		SAV-RRK(IDT)	
		Error ( $\tau$ )	Order	Error ( $\tau$ )	Order	Error ( $\tau$ )	Order
RK(2,2)	0.1	1.3395E-03	–	3.3870E-03	–	2.1470E-02	–
	0.05	3.4360E-04	1.9628	8.1480E-04	2.0555	1.0960E-02	0.9701
	0.025	8.6945E-05	1.9826	1.9951E-04	2.0300	5.5113E-03	0.9918
	0.0125	2.1865E-05	1.9915	4.9347E-05	2.0154	2.7600E-03	0.9977
	0.00625	5.4823E-06	1.9958	1.2270E-05	2.0078	1.3807E-03	0.9993
RK(3,3)	0.1	3.5168E-05	–	4.3927E-05	–	3.8213E-04	–
	0.05	4.3533E-06	3.0141	5.4825E-06	3.0022	8.0473E-05	2.2475
	0.025	5.3902E-07	3.0137	6.8378E-07	3.0032	1.8560E-05	2.1163
	0.0125	6.7058E-08	3.0068	8.5344E-08	3.0022	4.4475E-06	2.0611
	0.00625	8.3615E-09	3.0036	1.0659E-08	3.0012	1.0883E-06	2.0309
RK(4,4)	0.1	5.3561E-07	–	3.2716E-06	–	3.7745E-05	–
	0.05	3.6438E-08	3.8777	2.0654E-07	3.9855	4.8050E-06	2.9737
	0.025	2.3735E-09	3.9403	1.2898E-08	4.0012	6.0237E-07	2.9958
	0.0125	1.5146E-10	3.9700	8.0476E-10	4.0024	7.5303E-08	2.9999
	0.00625	9.5636E-12	3.9853	5.0241E-11	4.0016	9.4105E-09	3.0004

**Example 6.3** ([17]). Next, we consider the 2D nonlinear fractional wave equation

$$\begin{cases} u_{tt} + (-\Delta)^{\frac{\alpha}{2}} u + F'(u) = 0, (x, y, t) \in \Omega \times (0, T], \\ u(x, y, 0) = \frac{1}{2} \arctan \left( \exp \left( -\sqrt{x^2 + y^2} \right) \right), u_t(x, y, 0) = 0, \end{cases} \quad (6.7)$$

where  $\Omega = (-10, 10) \times (-10, 10)$ .

Taking the potential energy  $F(u) = u^2 \left( \frac{1}{4} u^2 - \frac{1}{2} \right)$ . The convergence order in time for  $\alpha = 1.5$  at  $T = 1$  is shown in [Table 3](#). Also, the evolution of discrete energy for a long-time simulation ( $T = 100$ ) is depicted in [Fig. 5](#). These data are calculated by applying the SAV-RRK(4,4) method. The results indicate that the proposed methods is also valid for the nonlinear fractional wave equations.

(a)  $\alpha = 1.3$ (b)  $\alpha = 1.6$ (c)  $\alpha = 1.9$ (d)  $\alpha = 2$ **Fig. 5.** Relative errors of energy with  $N = 4$ ,  $\tau = 0.01$  for different  $\alpha$  in [Example 6.3](#).

**Example 6.4** ([7]). Finally, we consider the 2D fractional Klein–Gordon–Schrödinger equation

$$\begin{cases} i\partial_t u - \frac{1}{2}(-\Delta)^{\frac{\alpha}{2}} u + u\phi = 0, (x, y, t) \in \Omega \times (0, T], \\ \partial_{tt}\phi + (-\Delta)^{\frac{\beta}{2}}\phi + \phi - |u|^2 = 0, (x, y, t) \in \Omega \times (0, T], \end{cases} \quad (6.8)$$

with the initial conditions

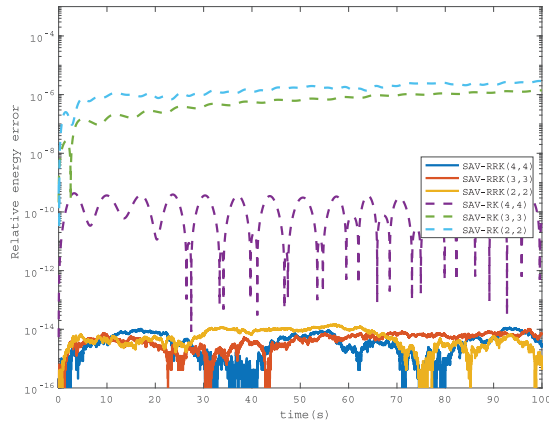
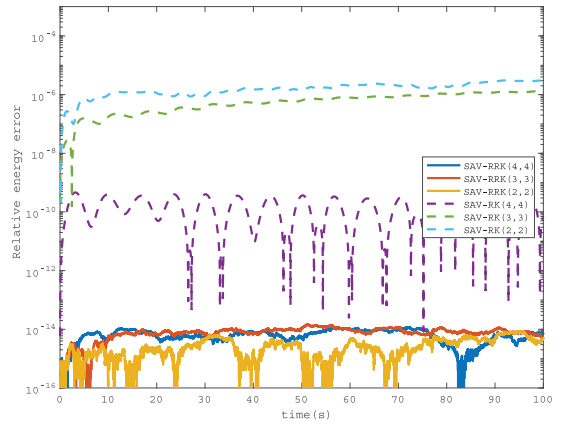
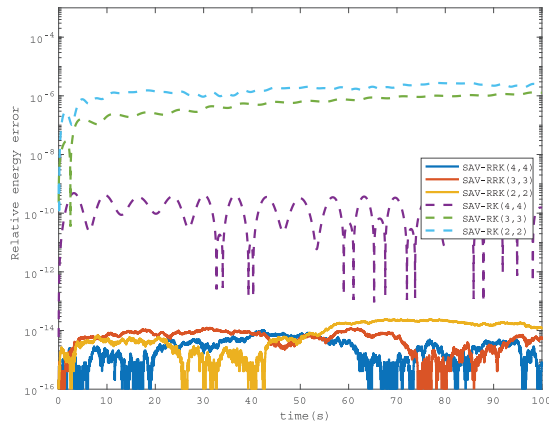
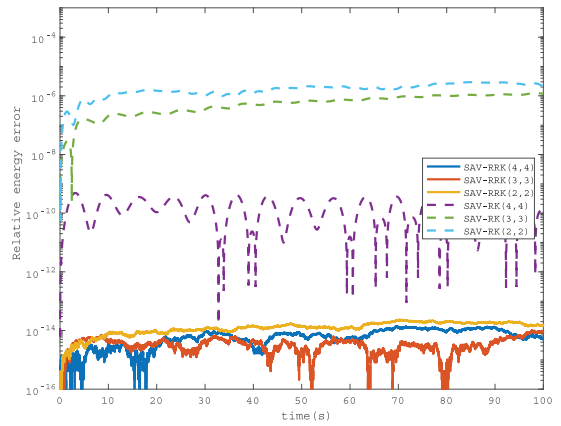
$$u(x, y, 0) = (1 + i)\exp(-|x|^2), \quad \phi(x, y, 0) = \operatorname{sech}(|x|^2), \quad \partial_t \phi(x, y, 0) = \sin(x + y)\operatorname{sech}(-2|x|^2), \quad (6.9)$$

where  $\Omega = [-10, 10] \times [-10, 10]$ .

Similar to the previous examples, some numerical errors and convergence order in time are listed in [Table 4](#) for  $\alpha = \beta = 1.5$ .

Furthermore, the relative energy with different  $\alpha$  and  $\beta$  in a long-time simulation up to  $T = 100$  is described in [Fig. 6](#). The phenomenon depicted demonstrates that the proposed methods accurately preserve energy at the discrete level for the fractional Klein–Gordon–Schrödinger equation.



(a)  $\alpha, \beta = 1.3$ (b)  $\alpha, \beta = 1.6$ (c)  $\alpha, \beta = 1.9$ (d)  $\alpha, \beta = 2$ **Fig. 6.** Relative errors of energy with  $N = 4$ ,  $\tau = 0.01$  for different  $\alpha$  in [Example 6.4](#).

## 7. Conclusions

In this paper, we present a new approach to develop high-order explicit energy-conserving numerical methods for the fractional nonlinear Schrödinger wave equations in two dimensions. The approach is proposed based on the recently introduced scalar auxiliary variable approach and the relaxation Runge–Kutta methods. The conservation properties of the proposed schemes are supported by theoretical analysis and numerical results. More numerical results show that the proposed methods are also effective for other similar fractional conservative problems, such as the nonlinear fractional wave equation and the fractional Klein–Gordon–Schrödinger equation, etc.

## Acknowledgments

This work is supported by the Sichuan Science and Technology Program (No. 2022JDTD0019) and the Laurent Mathematics Center of Sichuan Normal University and National-Local Joint Engineering Laboratory of System Credibility Automatic Verification (Grant No. ZD20220105).

**Table 4**Numerical errors and convergence order in time for [Example 6.4](#) when  $N = 4$ ,  $T = 1$ .

RK(Stage,Order)	$\tau$	SAV-RK		SAV-RRK(RT)		SAV-RRK(IDT)	
		Error ( $\tau$ )	Order	Error ( $\tau$ )	Order	Error ( $\tau$ )	Order
RK(2,2)	0.1	1.1875E−03	–	1.8837E−03	–	9.5325E−03	–
	0.05	2.7648E−04	2.1026	5.0394E−04	1.9023	6.7134E−03	0.5058
	0.025	6.6514E−05	2.0555	1.3036E−04	1.9508	3.8805E−03	0.7908
	0.0125	1.6300E−05	2.0288	3.3151E−05	1.9754	2.0757E−03	0.9026
	0.00625	4.0339E−06	2.0146	8.3587E−06	1.9877	1.0723E−03	0.9529
RK(3,3)	0.1	8.7748E−05	–	1.9567E−04	–	3.1789E−03	–
	0.05	1.1471E−05	2.9354	2.4630E−05	2.9900	8.2646E−04	1.9435
	0.025	1.4684E−06	2.9657	3.0916E−06	2.9940	2.1079E−04	1.9712
	0.0125	1.8580E−07	2.9824	3.8731E−07	2.9968	5.3231E−05	1.9854
	0.00625	2.3370E−08	2.9911	4.8471E−08	2.9983	1.3375E−05	1.9927
RK(4,4)	0.1	3.0741E−06	–	4.0627E−06	–	1.0278E−04	–
	0.05	1.9959E−07	3.9450	2.6020E−07	3.9647	1.3195E−05	2.9615
	0.025	1.2698E−08	3.9744	1.6461E−08	3.9825	1.6711E−06	2.9811
	0.0125	8.0044E−10	3.9876	1.0350E−09	3.9913	2.1024E−07	2.9906
	0.00625	5.0238E−11	3.9939	6.4882E−11	3.9957	2.6365E−08	2.9953

## References

- [1] W. Bao, Y. Cai, Uniform error estimates of finite difference methods for the nonlinear Schrödinger equation with wave operator, *SIAM J. Numer. Anal.* 50 (2) (2012) 492–521.
- [2] L. Brugnano, C. Zhang, D. Li, A class of energy-conserving Hamiltonian boundary value methods for nonlinear Schrödinger equation with wave operator, *Commun. Nonlinear Sci. Numer. Simul.* 60 (2018) 33–49.
- [3] L. Caffarelli, L. Silvestre, An extension problem related to the fractional Laplacian, *Comm. Partial Differential Equations* 32 (8) (2007) 1245–1260.
- [4] X. Cheng, H. Qin, J. Zhang, Convergence of an energy-conserving scheme for nonlinear space fractional Schrödinger equations with wave operator, *J. Comput. Appl. Math.* 400 (2022) 113762.
- [5] X. Cheng, F. Wu, Several conservative compact schemes for a class of nonlinear Schrödinger equations with wave operator, *Bound. Value Probl.* 2018 (1) (2018) 40.
- [6] T. Colin, P. Fabrie, Semidiscretization in time for nonlinear Schrödinger-waves equations, *Discr. Contin. Dynam. Syst.* 4 (4) (1998) 671–690.
- [7] Y. Fu, W. Cai, Y. Wang, Structure-preserving algorithms for the two-dimensional fractional klein–Gordon–Schrödinger equation, *Appl. Numer. Math.* 156 (2020) 77–93.
- [8] E. Hairer, G. Wanner, Runge–Kutta methods, explicit, implicit, *Encycl. Appl. Comput. Math.* (2015) 1282–1285.
- [9] D. Hu, W. Cai, X.-M. Gu, Y. Wang, Efficient energy preserving Galerkin–Legendre spectral methods for fractional nonlinear Schrödinger equation with wave operator, *Appl. Numer. Math.* 172 (2022) 608–628.
- [10] D.I. Ketcheson, Relaxation Runge–Kutta methods: Conservation and stability for inner-product norms, *SIAM J. Numer. Anal.* 57 (6) (2019) 2850–2870.
- [11] D. Li, X. Li, Z. Zhang, Implicit-explicit relaxation Runge–Kutta methods: Construction, analysis and applications to PDEs, *Math. Comp.* 92 (339) (2022) 117–146.
- [12] M. Li, Y.-L. Zhao, A fast energy conserving finite element method for the nonlinear fractional Schrödinger equation with wave operator, *Appl. Math. Comput.* 338 (2018) 758–773.
- [13] M. Ran, C. Zhang, A linearly implicit conservative scheme for the fractional nonlinear Schrödinger equation with wave operator, *Int. J. Comput. Math.* 93 (7) (2016) 1103–1118.
- [14] H. Ranocha, L. Lóczy, D.I. Ketcheson, General relaxation methods for initial-value problems with application to multistep schemes, *Numer. Math.* 146 (4) (2020a) 875–906.
- [15] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani, D.I. Ketcheson, Relaxation Runge–Kutta methods: Fully discrete explicit entropy-stable schemes for the compressible Euler and Navier–Stokes equations, *SIAM J. Sci. Comput.* 42 (2) (2020b) A612–A638.
- [16] P. Wang, C. Huang, A conservative linearized difference scheme for the nonlinear fractional Schrödinger equation, *Numer. Algorithms* 69 (3) (2015) 625–641.
- [17] N. Wang, M. Li, C. Huang, Unconditional energy dissipation and error estimates of the SAV Fourier spectral method for nonlinear fractional wave equation, *J. Sci. Comput.* 88 (1) (2021) 19.
- [18] X. Yang, L. Ju, Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model, *Comput. Methods Appl. Mech. Engrg.* 315 (2017a) 691–712.
- [19] X. Yang, L. Ju, Linear and unconditionally energy stable schemes for the binary fluid-urfactant phase field model, *Comput. Methods Appl. Mech. Engrg.* 318 (2017b) 1005–1029.

- [20] L. Zhang, Q. Chang, A conservative numerical scheme for a class of nonlinear Schrödinger equation with wave operator, *Appl. Math. Comput.* 145 (2) (2003) 603–612.
- [21] X. Zhang, M. Ran, Y. Liu, L. Zhang, A high-order structure-preserving difference scheme for generalized fractional Schrödinger equation with wave operator, *Math. Comput. Simulation* 210 (2023) 532–546.