CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2024-0032

Vol. **6**, No. 1, pp. 106-147
March 2025

A Novel up to Fourth-Order Equilibria-Preserving and Energy-Stable Exponential Runge-Kutta Framework for Gradient Flows

Haifeng Wang, Jingwei Sun, Hong Zhang* and Xu Qian

Department of Mathematics, College of Science, National University of Defense Technology, Changsha 410073, P.R. China

Received 25 July 2024 ; Accepted 7 November 2024

Abstract. In this work, we develop and analyze a family of up to fourth-order, unconditionally energy-stable, single-step schemes for solving gradient flows with global Lipschitz continuity. To address the exponential damping/growth behavior observed in Lawson's integrating factor Runge-Kutta approach, we propose a novel strategy to maintain the original system's steady state, leading to the construction of an exponential Runge-Kutta (ERK) framework. By integrating the linear stabilization technique, we provide a unified framework for examining the energy stability of the ERK method. Moreover, we show that certain specific ERK schemes achieve unconditional energy stability when a sufficiently large stabilization parameter is utilized. As a case study, using the no-slope-selection thin film growth equation, we conduct an optimal rate convergence analysis and error estimate for a particular three-stage, third-order ERK scheme coupled with Fourier pseudo-spectral discretization. This is accomplished through rigorous eigenvalue estimation and nonlinear analysis. Numerical experiments are presented to confirm the high-order accuracy and energy stability of the proposed schemes.

AMS subject classifications: 65M12, 65M15, 65M70
Key words: Gradient flows, exponential Runge-Kutta method, unconditional energy stability, error estimate.

1 Introduction

Gradient flows represent a significant class of physical models driven by free energy, characterized by a specific dissipation mechanism. Numerous challenges in fluid dynamics and material science can be effectively represented through gradient flow equations. Specifically, one main characteristic of this problem from a physical aspect is that the energy functional is decreasing all the time. Therefore, it is essential to develop ef-

*Corresponding author. *Email addresses:* hf_wang1031@163.com (H. Wang), sjw@nudt.edu.cn (J. Sun), zhanghnudt@163.com (H. Zhang), qianxu@nudt.edu.cn (X. Qian)

ficient and accurate numerical schemes that preserve the energy dissipation property at the discrete level. The primary objective of this paper is to develop a class of high order, unconditionally energy stable schemes for gradient flows. To illustrate the main idea, we consider a general gradient flow model with total free energy in the following form:

$$E(u)=\int_\Omega \frac{1}{2}u\cdot\mathcal{L}u+F(u)\mathrm{dx}, \tag{1.1}$$

where Ω is a bounded domain, $\mathcal{L}$ is a symmetric non-negative operator and $F(u)$ is a nonlinear potential function. With a specific symmetric non-negative operator $\mathcal{G}$ that commutes with $\mathcal{L}$, the gradient flow can be expressed with respect to the aforementioned energy as follows:

$$u_t=-\mathcal{G}\frac{\delta E(u)}{\delta u}=-\mathcal{G}\big(\mathcal{L}u+f(u)\big), \tag{1.2}$$

where $f(u):=F'(u)$. A common choice for the operator $\mathcal{G}$ is 1 or $-\Delta$, corresponding to the L^2 gradient flow and H^{-1} gradient flow, respectively.

Gradient flow models often exhibit strong stiffness and nonlinearity, and their steady states, which typically require long-time simulations to reach, are of considerable interest in practice. Therefore, there is a significant demand for numerical algorithms that are both efficient and accurate. Additionally, to avoid nonphysical phenomena appearing in the numerical solution, it is essential to preserve the energy dissipation during the numerical simulation. Consequently, various strategies have been designed to solve gradient flows, such as the nonlinear convex splitting scheme [8, 36, 43, 44], linear splitting (stabilization) technique [11, 34, 39], invariant energy quadratization (IEQ) method [45], and scalar auxiliary variable (SAV) algorithm [37, 38]. Among these, the nonlinear convex splitting method is known for its capacity to ensure unconditional energy stability and unique solvability. However, these methods often require solving nonlinear systems at each time step. By reformulating the energy functionals, Yang, Shen, and co-authors [37, 38, 45, 46] successively proposed the IEQ method and the SAV method. These methods can be conveniently utilized to construct schemes that are linearly solvable and unconditionally energy stable, albeit with the energy in these schemes being in a modified form. In recent years, the linear splitting (stabilization) technique has received much attention because of its capability to enlarge the stability region of linear implicit schemes. As a result, a large number of linearly implicit schemes, including the implicit-explicit Runge-Kutta (IMEX RK), exponential differencing (ETD) multi-step (MS), and ETD Runge-Kutta (ETDRK) schemes, have attracted considerable interest for solving gradient flow problems [6,21,40]. For general gradient flow equations with global Lipschitz assumptions, Fu and Yang [10] established the unconditional energy stability for a second-order stabilization ETDRK scheme with respect to the original energy. Introducing a third-order accurate Douglas-Dupont stabilization term, Cheng *et al.* [4] derived energy stability for a third-order ETDMS scheme regarding a modified energy with a few numerical correction terms. Recently, based on the global Lipschitz condition and the use of linear stabilization terms, Fu *et al.* [9] developed a four-stage, third-order stabilization

implicit-explicit Runge-Kutta (IMEX RK) scheme with unconditional energy stability for gradient flow models. Nevertheless, constructing a fourth-order energy-stable scheme remains a challenging task.

This study aims to construct a novel energy-stable exponential RK framework with high-order accuracy while preserving the energy dissipation property for gradient flow equations. In recent years, Lawson's integrating factor Runge-Kutta (IFRK) framework [24] has gained popularity because of its unconditional contractivity and asymptotic stability for semi-linear problems [33]. Additionally, it is known for its capability to preserve the strong stability of hyperbolic conservation laws [18] and the maximum principles of Allen-Cahn type equations [20, 41, 47, 50, 52–54]. However, the drawbacks of the IFRK method are also quite apparent, with the notable one being the occurrence of an exponential damping/growth phenomenon [16,25]. Although this phenomenon can suppress numerical instability, it tends to oversmooth the numerical solution, resulting in deviations from the solution of the original problem and rendering it unreliable. In this work, we develop a new strategy to tackle this issue and construct a class of unconditionally energy-stable schemes for general gradient flows with improved accuracy. Specifically, we take the no-slope-selection (NSS) thin film equation as an example to carry out rigorous error analysis. The main contributions of this work include:

- By carefully combining exponential and linear functions to approximate exponential functions in each IFRK stage, we develop a novel exponential Runge-Kutta framework. This framework eliminates the exponential damping/growth phenomenon in the IFRK method while maintaining the original steady states (equilibria).
- A unified framework is presented to examine the energy stability of the proposed ERK schemes. We demonstrate that the novel ERK schemes, incorporating a series of up to fourth-order RK coefficients, ensure the dissipation of the original energy for the gradient flow without any time-step restrictions.

The rest of this paper is organized as follows. In Section 2, we provide a succinct review of the linear splitting (stabilization) method for general gradient flows. Following this, an analysis of the exponential effects stemming from Lawson's IFRK method is conducted, leading to the construction of the ERK framework. In section 3, we explore the energy stability of the ERK scheme within a unified framework. Section 4 presents generic order conditions to determine the truncation error of ERK schemes and subsequently provides the convergence analysis for the NSS equation. In Section 5, various numerical experiments are presented to verify the high-order accuracy, efficiency, and unconditional energy stability. Finally, some concluding remarks are given in Section 6.

2 Lawson's IFRK method and improvements

In this section, we provide a brief overview of Lawson's IFRK method and proceed to analyze the exponential damping/growth phenomenon in IFRK schemes, aiming to pin-

point its fundamental cause. Following this analysis, we propose a novel strategy to eliminate the exponential effect by constructing a set of nonlinear approximations to substitute the exponential functions used in the IFRK methods. The implementation of this innovative approach paves the way for the development of the equilibrium-preserving ERK framework.

2.1 The linear splitting (stabilization) method

After implementing an appropriate spatial discretization (e.g. the Fourier collocation discretization as outlined in Section 4.1) of the gradient flow equation (1.2), we present a semi-discrete system in the following form:

$$\boldsymbol{u}_t = -G_N\big(L_N\boldsymbol{u}+f(\boldsymbol{u})\big), \quad t\in(0,T], \tag{2.1}$$

where $\boldsymbol{u}$ is the semi-discrete solution, the symbols G_N and L_N represent symmetric positive (semi)-definite matrices that commute and arise from the discretization of $\mathcal{G}$ and $\mathcal{L}$, respectively. Consequently, the discrete total energy with respect to (2.1) can be written as

$$E_N(\boldsymbol{u}) = \frac{1}{2}\langle \boldsymbol{u}, L_N\boldsymbol{u}\rangle + \langle F(\boldsymbol{u}),1\rangle, \tag{2.2}$$

where $\langle\cdot,\cdot\rangle$ denotes the discrete L^2 inner product. To simplify the argument, we assume that the function f is globally Lipschitz continuous with a Lipschitz constant ℓ, i.e.

$$F(a)-F(b)\le f(b)(a-b)+\frac{\ell}{2}(a-b)^2, \quad \forall a,b\in\mathbb{R}. \tag{2.3}$$

By introducing $\kappa G_N(\boldsymbol{u}-\boldsymbol{u})$ with $\kappa\ge 0$, we consider the following stabilization semi-discretization:

$$\boldsymbol{u}_t = -L_\kappa\boldsymbol{u}+N_\kappa(\boldsymbol{u}), \tag{2.4}$$

where $L_\kappa = G_N(L_N+\kappa I)$ and $N_\kappa(\boldsymbol{u}) = G_N(\kappa\boldsymbol{u}-f(\boldsymbol{u}))$. During computation, since L_κ is linear and stiff, $-L_\kappa\boldsymbol{u}$ can be integrated exactly, and the nonlinear part $N_\kappa(\boldsymbol{u})$ is typically treated explicitly for efficiency. Therefore, the linear splitting method allows us to use exponential integrators to solve the problem. In addition, κ acts as a stabilization factor to enhance the dissipation of the linear part so as to bound the Lipschitz growth of nonlinear terms in the analysis and practice. In fact, stabilization is widely used to construct energy-stable schemes, and numerical experiments demonstrate that stabilization will significantly improve numerical performance, which we will illustrate in the following section in detail.

Remark 2.1. In this study, we assume that the nonlinear term f is globally Lipschitz continuous, which is crucial for the subsequent energy stability analysis. However, in the majority of gradient flow models, this criterion is not naturally satisfied. Instead, since for most classical gradient flow equations, the nonlinear potential function satisfies

local Lipschitz continuity, therefore, deriving an appropriate maximum norm bound for the numerical solution could effectively justify the assumption.

A considerable amount of research has been devoted to evaluating the maximum norm bound of numerical solutions or, more broadly, to conducting global-in-time energy stability for some specific gradient flow equations. The Allen-Cahn (AC) type equation, which inherently possesses the maximum bound principle (MBP), has spurred considerable effort in constructing numerical schemes that preserve both MBP as well as energy stability [6, 15, 19, 20, 48]. For Cahn-Hilliard (CH) equation, Li *et al.* [27–29] utilized harmonic analysis in borderline spaces to deduce a maximum bound for the numerical solution in a first-order scheme and hence establish the unconditional energy stability and characterize the stabilization parameter. More recently, by fully utilizing the H^2 regularity assumption brought by the free energy and Sobolev embedding inequality, Li *et al.* [30] and Zhang *et al.* [51] established the uniform-in-time energy stability for the exponential time difference (ETD) numerical scheme and IMEX RK scheme for the phase field crystal equation (PFC) [30]. Moreover, different from the aforementioned model, the no-slope-selection (NSS) thin film model [21] meets the Lipschitz assumption with respect to ∇u, that is

$$\langle F(\nabla u)-F(\nabla v),1\rangle \le \langle u-v, f(\nabla v)\rangle + \frac{\ell}{2}\|\nabla(u-v)\|^2,$$

where $\ell = 1/8$. Given that the nonlinear energy potential function for the NSS equation is a bit special, we intend to provide a more detailed discussion of the NSS equation in Section 4.

2.2 Exponential effect of integrating factor method

We divide the time interval into a series of nodes $\{t_n | t_n = n\tau, n=0,1,\dots,[T/\tau]\}$, where $\tau>0$ represents the time step size. Denoting $\boldsymbol{v}(t) = \mathrm{e}^{tL_\kappa}\boldsymbol{u}(t)$ in (2.4), we obtain an equivalent system in the following form by the variation-of-constants method:

$$\begin{cases} \boldsymbol{v}_t = \mathrm{e}^{tL_\kappa} N_\kappa(\mathrm{e}^{-tL_\kappa}\boldsymbol{v}), \quad t\in(0,T], \\ \boldsymbol{v}(0) = \boldsymbol{u}^0. \end{cases} \tag{2.5}$$

Then, solving system (2.5) using the standard explicit s-stage, p-th-order RK method and transforming the solution back via the inverse relation $\boldsymbol{u}(t) = \mathrm{e}^{-tL_\kappa}\boldsymbol{v}(t)$, we obtain

$$\begin{cases} \boldsymbol{u}_{n,0} = \boldsymbol{u}^n, \\ \boldsymbol{u}_{n,i} = \mathrm{e}^{-c_i\tau L_\kappa}\left(\boldsymbol{u}^n + \tau\sum\limits_{j=0}^{i-1} a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} N_\kappa(\boldsymbol{u}_{n,j})\right), \quad i=1,\dots,s, \\ \boldsymbol{u}^{n+1} = \boldsymbol{u}_{n,s}. \end{cases} \tag{2.6}$$

Here, the underlying RK coefficients $a_{i,j}\ge 0$ are constrained by certain accuracy requirements. The intermediate solution $\boldsymbol{u}_{n,i}$ are approximations to $\boldsymbol{u}(t_{n,i})$ with $t_{n,i} = t_n + c_i\tau$,

and the abscissas are defined as follows: $c_0=0, c_i=\sum_{j=0}^{i-1}a_{i,j}$ for $1\le i\le s$, and $c_s=1$ for consistency. The formulation (2.6) is denoted as Lawson's IFRK method [24].

Although IFRK schemes offer advantages in dealing with the stiffness [14, 23, 24], they fail to preserve the equilibria of the semi-discrete equation. To illustrate this issue, we denote $\boldsymbol{u}^*$ as a nontrivial equilibrium state to (2.5), i.e. $L_\kappa \boldsymbol{u}^* = N_\kappa(\boldsymbol{u}^*)$. Denoting the discrete Fourier transform of $\boldsymbol{u}^*$ as $\hat{\boldsymbol{u}}^*$ and assuming that $\hat{\boldsymbol{u}}_{n,j}=\hat{\boldsymbol{u}}^*$ for $j=0,1,\ldots,i-1$, we apply the discrete Fourier transform to the formulation (2.6) and obtain

$$\begin{aligned}(\hat{\boldsymbol{u}}_{n,i})_{k,l} &= \mathrm{e}^{-c_i\tau\Lambda_{k,l}}\left((\hat{\boldsymbol{u}}^n)_{k,l}+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau\Lambda_{k,l}}\Big[\underbrace{\widehat{\big(N_\kappa(\boldsymbol{u}_{n,j})\big)}-\Lambda_{k,l}(\hat{\boldsymbol{u}}_{n,j})_{k,l}}_{=0}+\Lambda_{k,l}(\hat{\boldsymbol{u}}_{n,j})_{k,l}\Big]\right)\\ &= \mathrm{e}^{-c_i\tau\Lambda_{k,l}}\left(1+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau\Lambda_{k,l}}\Lambda_{k,l}\right)(\hat{\boldsymbol{u}}^*)_{k,l},\quad i\le s,\end{aligned}$$

where $\{\Lambda_{k,l}\}$ are the eigenvalues of L_κ. The nonlinear term is evaluated in the physical space and subsequently transformed into the Fourier space.

Because

$$\mathrm{e}^{-c_i\tau\Lambda_{k,l}}\left(1+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau\Lambda_{k,l}}\Lambda_{k,l}\right)\neq 1$$

unless $c_i\tau\Lambda_{k,l}=0$, the IFRK method (2.6) cannot preserve the equilibrium state and will introduce unwanted exponential damping effect into the Fourier modes of the state when $c_i>c_j$ for $i>j$ (or growth if $c_i<c_j$), especially when $\tau\Lambda_{k,l}$ is large. Although the exponential damping effect could help to suppress numerical instabilities such as those arising from the aliasing error or under-resolution [16], it would lead to a time-step-dependent smoothing of the high gradient associated with the interfacial layers [25]. This results in a discrepancy between numerical and analytical solutions at high wavenumbers.

Example 2.1. To demonstrate the exponential effect of IFRK schemes, we consider the 1D Cahn-Hillard equation, which is a classical gradient flow equation with

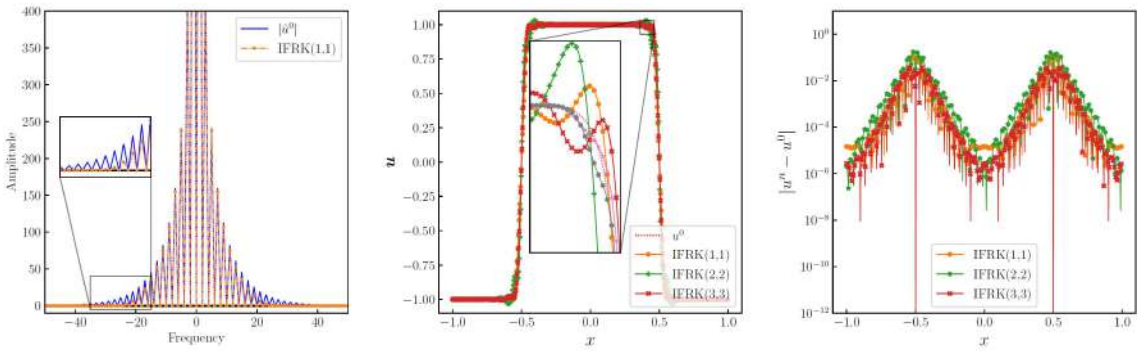
$$\mathcal{G}=-\Delta,\quad \mathcal{L}=-\varepsilon^2\Delta,\quad F(u)=\frac{1}{4}(1-u^2)^2$$

in (1.1)-(1.2). The domain is set as $\Omega=(-1,1)$ with periodic boundary conditions, and the initial condition is chosen as

$$u^0(x)=\tanh\left(\frac{0.5-|x|}{\sqrt{2}\varepsilon}\right),$$

which is a local minimum of the energy functional and the profile exhibits two equilibrium interfaces at $x=\pm 0.5$.

We selected the following parameters: $\varepsilon=0.02, \tau=0.0001, \kappa=0$, and final time $T=0.2$. The spatial discretization is carried out using the Fourier pseudo-spectral method in



(a) Amplitude spectrum of the solutions (b) Solutions at $t=0.2$ (c) $|\boldsymbol{u}^n-\boldsymbol{u}^0|$ at $t=0.2$

Figure 1: (a): Plot of the amplitude spectrum with respect to the initial condition and the solutions obtained by IFRK(1,1) method. (b): Profiles of the solution obtained by IFRK schemes for $\varepsilon=0.02$. (c): Absolute errors of the obtained solutions versus the initial condition.

Section 4.1 with $N=2048$. Fig. 1 presents the results computed using some first- to third-order IFRK schemes (with underlying explicit Euler coefficient, Heun's second- and third-order RK coefficients ([12, p. 135])). In Fig. 1(a), we present the amplitude spectrum of the initial condition and the solutions obtained from the IFRK(1,1) method at $t=0.2$. It is evident that the amplitude of the solution obtained from the IFRK(1,1) method is attenuated at high frequencies due to the exponential effect. Consequently, the IFRK schemes fail to preserve the equilibrium state, even in the absence of a stabilization term, as shown in Fig. 1(b). Although applying higher-order IFRK schemes could improve the accuracy, significant errors caused by exponential damping persisted near the interfacial regions ($x\approx\pm0.5$).

2.3 A novel exponential Runge-Kutta framework

In this section, our objective is to construct a framework that removes the exponential effect from the IFRK approach. Previous studies [17,23,25,49] have made efforts to address the exponential effect. To address this issue, we denote $\psi_i(\tau L_\kappa)$ as an approximation to $e^{c_i\tau L_\kappa}$ in (2.6), and ensure that it satisfies the following conditions for the Fourier modes:

$$\psi_i^{-1}(\tau\Lambda_{k,l})\left(1+\tau\sum_{j=0}^{i-1}a_{i,j}e^{c_j\tau\Lambda_{k,l}}\Lambda_{k,l}\right)\equiv 1,\quad i=1,\ldots,s.$$

Hence, $\psi_i(\tau L_\kappa)$ can be directly formulated using a combination of exponential and linear functions in the following manner:

$$\psi_i(\tau L_\kappa)=I+\tau\sum_{j=0}^{i-1}a_{i,j}e^{c_j\tau L_\kappa}L_\kappa,\quad i=1,\ldots,s. \tag{2.7}$$

In turn, by substituting the function $\psi_i(\tau L_\kappa)$ for $\mathrm{e}^{c_i\tau L_\kappa}$ in the IFRK formulation (2.6), we obtain the following modification:

$$\boldsymbol{u}_{n,i}=\psi_i^{-1}(\tau L_\kappa)\left(\boldsymbol{u}^n+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,j})\right),\quad i=1,\ldots,s. \tag{2.8}$$

To confirm that (2.8) preserves the equilibrium as well as eliminates the exponential effect in the IFRK framework, we assume $\boldsymbol{u}_{n,j}=\boldsymbol{u}^*$ for $j=0,1,\ldots,i-1$. This yields

$$\begin{aligned}\boldsymbol{u}_{n,i}&=\psi_i^{-1}(\tau L_\kappa)\left(\boldsymbol{u}^*+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau L_\kappa}N_\kappa(\boldsymbol{u}^*)\right)\\&=\underbrace{\psi_i^{-1}(\tau L_\kappa)\left(I+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau L_\kappa}L_\kappa\right)}_{\equiv I}\boldsymbol{u}^*=\boldsymbol{u}^*,\quad i=1,\ldots,s.\end{aligned}$$

Thus, the approximations $\psi_i(\tau L_\kappa)$ in (2.7) effectively address the issue outlined in the previous subsection. For brevity, we denote the improved formulation (2.8) as the exponential Runge-Kutta (ERK) method, as the approximations $\psi_i(\tau L_\kappa)$ are constructed using exponential functions.

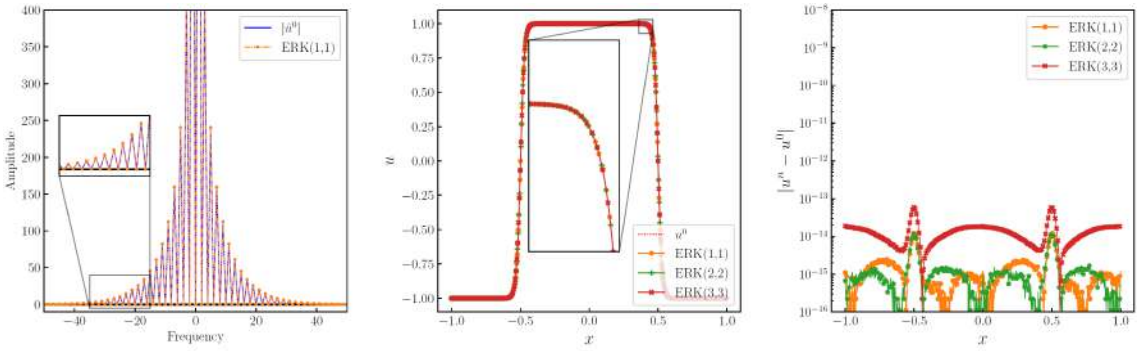
Remark 2.2. It is worth noting that the formulation constructed in (2.7) is indeed an approximation to $\mathrm{e}^{c_i\tau L_\kappa}$. This can be illustrated by considering the following differential equation:

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}\Psi(t)=L_\kappa\mathrm{e}^{tL_\kappa},\quad t\in(0,\tau],\\ \Psi(0)=I\end{cases}$$

with the exact solution $\Psi(t)=\mathrm{e}^{tL_\kappa}$. It is clear that $\psi_i(\tau L_\kappa)$ is the corresponding i-th stage approximation obtained by discretizing the above equation using the underlying explicit Runge-Kutta method.

Remark 2.3. In the existing literature, various approaches have been proposed to alleviate or eliminate exponential effects by approximating exponential functions within the integrating factor framework. For further details, interested readers may consult the works of [5, 17, 25, 53]. Our approach is distinctive in that it replaces only one exponential function in each IFRK stage, enabling the derivation of up to fourth-order unconditionally energy-stable schemes. Moreover, the accuracy tests in Section 5.1 will illustrate that our ERK schemes demonstrate superior accuracy compared to the original IFRK schemes.

We now revisit Example 2.1 and adopt first- to third-order ERK schemes to solve the problem. By employing the same parameter settings, we present the numerical result in Fig. 2. The amplitude spectrum of the solutions obtained from the ERK(1,1) scheme overlapped with those of the initial condition. Furthermore, all ERK schemes are able to effectively preserve the initial equilibrium state, with the absolute error staying within machine accuracy.



(a) Amplitude spectrum of the solutions (b) Solutions at $t = 0.2$ (c) $|\boldsymbol{u}^n - \boldsymbol{u}^0|$ at $t = 0.2$

Figure 2: (a): Plot of the amplitude spectrum with respect to the initial condition and the solutions obtained by ERK(1,1). (b): Profiles of the solution obtained by the ERK schemes for $\varepsilon = 0.02$. (c): Absolute errors of the obtained solutions versus the initial condition.

3 Energy stability of ERK schemes

In this section, we aim to present a theoretical framework to examine that a specific class of the RK coefficients can enable the proposed ERK scheme to unconditionally preserve the original energy dissipation law for gradient flows, provided the stabilization parameter κ is chosen appropriately.

First, we present some preliminary results of ERK schemes, which will facilitate the construction of the main result.

Lemma 3.1. *Let $u_{n,i}$ be the i-th-stage solution obtained by ERK schemes. It holds that*

$$-L_\kappa \boldsymbol{u}_{n,i} + N_\kappa(\boldsymbol{u}_{n,i-1}) = \tau^{-1} \sum_{k=0}^{i} \omega_{i,k}(\tau L_\kappa) \boldsymbol{u}_{n,k} \tag{3.1}$$

with

$$\sum_{k=0}^{i} \omega_{i,k}(\tau L_\kappa) = 0, \quad i = 1, \ldots, s,$$

where $\omega_{i,k}(\tau L_\kappa)$ depend on the RK coefficients $a_{i,j}$ and τL_κ.

Proof. We prove this theorem by induction. When $i = 1$, we have

$$\begin{aligned} -L_\kappa \boldsymbol{u}_{n,1} + N_\kappa(\boldsymbol{u}_{n,0}) &= -L_\kappa \boldsymbol{u}_{n,1} + (\tau a_{1,0} I)^{-1} \big(\psi_1(\tau L_\kappa) \boldsymbol{u}_{n,1} - \boldsymbol{u}_{n,0}\big) \\ &= (\tau a_{1,0} I)^{-1} \big(\big(-a_{1,0} \tau L_\kappa + \psi_1(\tau L_\kappa)\big) \boldsymbol{u}_{n,1} - \boldsymbol{u}_{n,0} \big) \\ &= \tau^{-1} \big(\omega_{1,1}(\tau L_\kappa) \boldsymbol{u}_{n,1} + \omega_{1,0}(\tau L_\kappa) \boldsymbol{u}_{n,0}\big), \end{aligned}$$

where

$$\omega_{1,1}(\tau L_\kappa) = (a_{1,0} I)^{-1} \big(-a_{1,0} \tau L_\kappa + \psi_1(\tau L_\kappa)\big), \quad \omega_{1,0}(\tau L_\kappa) = -(a_{1,0} I)^{-1}.$$

By recalling the definition of $\psi_1(\tau L_\kappa)$, it is clear that $\omega_{1,1}(\tau L_\kappa) + \omega_{1,0}(\tau L_\kappa) = 0$.

We assume the result holds for $i=1,\dots,\ell-1$, $\ell\leq s$, then

$$N_\kappa(\boldsymbol{u}_{n,\ell-1})=\tau^{-1}\sum_{k=0}^{\ell}\omega_{\ell,k}(\tau L_\kappa)\boldsymbol{u}_{n,k}+L_\kappa\boldsymbol{u}_{n,j}.$$

For case $i=\ell$, we obtain

$$\begin{aligned}
&-L_\kappa\boldsymbol{u}_{n,\ell}+N_\kappa(\boldsymbol{u}_{n,\ell-1})\\
=&-L_\kappa\boldsymbol{u}_{n,\ell}+\left(a_{\ell,\ell-1}\mathrm{e}^{c_{\ell-1}\tau L_\kappa}\tau\right)^{-1}\left[\psi_\ell(\tau L_\kappa)\boldsymbol{u}_{n,\ell}-\boldsymbol{u}_{n,0}-\tau\sum_{j=0}^{\ell-2}a_{\ell,j}\mathrm{e}^{c_j\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,j})\right]\\
=&-L_\kappa\boldsymbol{u}_{n,\ell}+\left(a_{\ell,\ell-1}\mathrm{e}^{c_{\ell-1}\tau L_\kappa}\tau\right)^{-1}\Bigg[\psi_\ell(\tau L_\kappa)\boldsymbol{u}_{n,\ell}-\boldsymbol{u}_{n,0}\\
&\qquad-\tau\sum_{j=0}^{\ell-2}a_{\ell,j}\mathrm{e}^{c_j\tau L_\kappa}\left(\tau^{-1}\sum_{k=0}^{j}\omega_{j,k}(\tau L_\kappa)\boldsymbol{u}_{n,k}+L_\kappa\boldsymbol{u}_{n,j}\right)\Bigg].
\end{aligned}$$

Thus, $L_\kappa\boldsymbol{u}_{n,\ell}+N_\kappa(\boldsymbol{u}_{n,\ell-1})$ is decomposed into a combination of $u_{n,k}$ with $k=0,\dots,\ell$, and the summation of coefficients satisfies

$$\begin{aligned}
\tau^{-1}\sum_{k=0}^{\ell}\omega_{\ell,k}&=-L_\kappa+\left(a_{\ell,\ell-1}\mathrm{e}^{c_{\ell-1}\tau L_\kappa}\tau\right)^{-1}\left[\psi_\ell(\tau L_\kappa)-I-\tau\sum_{j=0}^{\ell-2}a_{\ell,j}\mathrm{e}^{c_j\tau L_\kappa}L_\kappa\right]\\
&=\left(a_{\ell,\ell-1}\mathrm{e}^{-c_{\ell-1}\tau L_\kappa}\tau\right)^{-1}\left[\psi_\ell(\tau L_\kappa)-I-\tau\sum_{j=0}^{\ell-1}a_{\ell,j}\mathrm{e}^{-c_j\tau L_\kappa}L_\kappa\right]=0,
\end{aligned}$$

where the last equality is obtained using the definition of $\psi_\ell(\tau L_\kappa)$. This completes the proof. □

Lemma 3.2. *The i-th stage solution $\boldsymbol{u}_{n,i}$ obtained by the ERK method satisfies the following relationship:*

$$-L_\kappa\boldsymbol{u}_{n,i}+N_\kappa(\boldsymbol{u}_{n,i-1})=\tau^{-1}\sum_{j=1}^{i}\Delta_{i,j}(\tau L_\kappa)(\boldsymbol{u}_{n,j}-\boldsymbol{u}_{n,j-1}) \tag{3.2}$$

with

$$\Delta_{i,j}(\tau L_\kappa)=\sum_{k=j}^{i}\omega_{i,k}(\tau L_\kappa),\quad j=1,\dots,i,$$

where $\omega_{i,k}$ are the expansion coefficients of $L_\kappa\boldsymbol{u}_{n,i}+N_\kappa(\boldsymbol{u}_{n,i-1})$ in (3.1).

Proof. By utilizing the Lemma 3.1 and noting that $\omega_{i,k}$ satisfy $\sum_{k=0}^{i}\omega_{i,k}=0$, we can directly obtain

$$\begin{aligned}
&-L_\kappa\boldsymbol{u}_{n,i}+N_\kappa(\boldsymbol{u}_{n,i-1})\\
=&\tau^{-1}\sum_{k=0}^{i}\omega_{i,k}(\tau L_\kappa)\boldsymbol{u}_{n,k}=\tau^{-1}\sum_{k=1}^{i}\omega_{i,k}(\tau L_\kappa)\boldsymbol{u}_{n,k}-\tau^{-1}\sum_{k=1}^{i}\omega_{i,k}(\tau L_\kappa)\boldsymbol{u}_{n,0}
\end{aligned}$$

$$
\begin{aligned}
&=\tau^{-1}\sum_{k=2}^{i}\omega_{i,k}(\tau L_\kappa)\boldsymbol{u}_{n,k}-\tau^{-1}\sum_{k=2}^{i}\omega_{i,k}(\tau L_\kappa)\boldsymbol{u}_{n,1}+\tau^{-1}\sum_{k=1}^{i}\omega_{i,k}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\
&\quad\cdots\\
&=\tau^{-1}\omega_{i,i}(\tau L_\kappa)(\boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1})+\cdots+\tau^{-1}\sum_{k=1}^{i}\omega_{i,k}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\
&=\tau^{-1}\sum_{j=1}^{i}\Delta_{i,j}(\tau L_\kappa)(\boldsymbol{u}_{n,j}-\boldsymbol{u}_{n,j-1}),
\end{aligned}
$$

and the proof is complete. □

Lemma 3.3. *Consider the semi-discrete system* (2.1) *of gradient flows, assuming that the stabilization parameter satisfies* $\kappa\geq\ell/2$, *where* ℓ *is defined in* (2.3), *then the two adjacent stage solutions obtained by the ERK method satisfy the following relationship:*

$$
\begin{aligned}
&E_N(\boldsymbol{u}_{n,i})-E_N(\boldsymbol{u}_{n,i-1})\\
&\leq-\big\langle\boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1},G_N^{-1}\big(-L_\kappa\boldsymbol{u}_{n,i}+N_\kappa(\boldsymbol{u}_{n,i-1})\big)\big\rangle\\
&=-\left\langle\boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1},\tau^{-1}G_N^{-1}\sum_{j=1}^{i}\Delta_{i,j}(\tau L_\kappa)(\boldsymbol{u}_{n,j}-\boldsymbol{u}_{n,j-1})\right\rangle,\quad i=1,\ldots,s,
\end{aligned}
$$

where the coefficients $\Delta_{i,j}$ *are defined in* (3.2).

Proof. According to the Lipschitz continuity assumption for f, given any $\boldsymbol{u},\boldsymbol{v}\in\mathcal{M}_h$, we have

$$
\begin{aligned}
\langle F(\boldsymbol{v})-F(\boldsymbol{u}),1\rangle&\leq\langle\boldsymbol{v}-\boldsymbol{u},f(\boldsymbol{u})\rangle+\frac{\ell}{2}\langle\boldsymbol{v}-\boldsymbol{u},\boldsymbol{v}-\boldsymbol{u}\rangle\\
&=-\big\langle G_N^{-1}N_\kappa(\boldsymbol{u}),\boldsymbol{v}-\boldsymbol{u}\big\rangle+\kappa\langle\boldsymbol{u},\boldsymbol{v}-\boldsymbol{u}\rangle+\frac{\ell}{2}\big\langle\boldsymbol{v}-\boldsymbol{u},\boldsymbol{v}-\boldsymbol{u}\big\rangle.
\end{aligned}\tag{3.3}
$$

On the other hand, for the difference of linear parts, it holds that

$$
\begin{aligned}
&\frac{1}{2}(\langle\boldsymbol{v},L_N\boldsymbol{v}\rangle-\langle\boldsymbol{u},L_N\boldsymbol{u}\rangle)\\
&=\langle\boldsymbol{v}-\boldsymbol{u},L_N\boldsymbol{v}\rangle-\frac{1}{2}\langle\boldsymbol{v}-\boldsymbol{u},L_N(\boldsymbol{v}-\boldsymbol{u})\rangle\\
&=-\big\langle\boldsymbol{v}-\boldsymbol{u},-G_N^{-1}L_\kappa\boldsymbol{v}\big\rangle-\kappa\langle\boldsymbol{v},\boldsymbol{v}-\boldsymbol{u}\rangle-\frac{1}{2}\langle\boldsymbol{v}-\boldsymbol{u},L_N(\boldsymbol{v}-\boldsymbol{u})\rangle.
\end{aligned}\tag{3.4}
$$

By assuming that $\kappa\geq\ell/2$ and then adding (3.3) and (3.4), we obtain following result:

$$
\begin{aligned}
E_N(\boldsymbol{v})-E_N(\boldsymbol{u})\leq&-\big\langle\boldsymbol{v}-\boldsymbol{u},G_N^{-1}(-L_\kappa\boldsymbol{v}+N_\kappa\boldsymbol{u})\big\rangle-\left(\kappa-\frac{\ell}{2}\right)\|\boldsymbol{v}-\boldsymbol{u}\|^2\\
&-\frac{1}{2}\langle\boldsymbol{v}-\boldsymbol{u},L_N(\boldsymbol{v}-\boldsymbol{u})\rangle.
\end{aligned}\tag{3.5}
$$

Letting $\boldsymbol{v}=\boldsymbol{u}_{n,i}$, $\boldsymbol{u}=\boldsymbol{u}_{n,i-1}$, and substituting Eq. (3.2) into (3.5) gives

$$\begin{aligned}
&E_N(\boldsymbol{u}_{n,i})-E_N(\boldsymbol{u}_{n,i-1})\\
\leq&-\big\langle \boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1},G_N^{-1}\big(-L_\kappa \boldsymbol{u}_{n,i}+N_\kappa(\boldsymbol{u}_{n,i-1})\big)\big\rangle\\
=&-\left\langle \boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1},\tau^{-1}G_N^{-1}\sum_{j=1}^{i}\Delta_{i,j}(\tau L_\kappa)(\boldsymbol{u}_{n,j}-\boldsymbol{u}_{n,j-1})\right\rangle.
\end{aligned}$$

This completes the proof. □

To facilitate the derivation of the main energy stability result, we define the notation $\delta\boldsymbol{U}=[\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0},\dots,\boldsymbol{u}_{n,s}-\boldsymbol{u}_{n,s-1}]^\top$ and $\boldsymbol{G}_N=I_s\otimes G_N$, where I_s is the identity matrix in $\mathbb{R}^{s\times s}$. In addition, we introduce the lower triangular matrix $M(z)=(\Delta_{i,j}(z))_{1\leq j\leq i,1\leq i\leq s}$, which we refer to as the energy stability matrix, and denote its associate symmetric matrix by $S(z)=(1/2)(M(z)+M(z)^\top)$.

Theorem 3.1. *Under the assumption that the stabilization parameter satisfies $\kappa\geq\ell/2$ and the matrix $S(z)$ is positive (semi)-definite for $z\geq 0$, the stabilization ERK(s,p) method is energy-stable without any time step constrain.*

Proof. According to the assumption on κ and applying the Lemma 3.3, the energy difference can be bounded by a quadratic form

$$\begin{aligned}
E_N(\boldsymbol{u}_{n,s})-E_N(\boldsymbol{u}_{n,0})=&\sum_{i=1}^{s}E(\boldsymbol{u}_{n,i})-E(\boldsymbol{u}_{n,i-1})\\
\leq&-\sum_{i=1}^{s}\left\langle \boldsymbol{u}_{n,i}-\boldsymbol{u}_{n,i-1},\tau^{-1}G_N^{-1}\sum_{j=1}^{i}\Delta_{i,j}(\tau L_\kappa)(\boldsymbol{u}_{n,j}-\boldsymbol{u}_{n,j-1})\right\rangle\\
=&-\big\langle \delta\boldsymbol{U},(\tau\boldsymbol{G}_N)^{-1}M(\tau L_\kappa)\delta\boldsymbol{U}\big\rangle\\
=&-\big\langle \delta\boldsymbol{U},(\tau\boldsymbol{G}_N)^{-1}S(\tau L_\kappa)\delta\boldsymbol{U}\big\rangle.
\end{aligned}$$

The last equality hold since $\boldsymbol{G}_N^{-1}$ is symmetric, and commutes with $M(\tau L_\kappa)$. Since $G_N,\tau L_\kappa$ are positive (semi)-definite, and the matrix $S(z)$ is assumed to be positive (semi)-definite for $z\geq 0$, we complete the proof. □

Based on the aforementioned results, for any ERK scheme with specified RK coefficients, it is sufficient to verify the positive definiteness of the corresponding energy stability matrix $M(z)$ and its associate symmetric matrix $S(z)$ to examine the energy stability of the ERK scheme. In the subsequent section, we will introduce first- to fourth-order ERK schemes and conduct a case-by-case energy stability analysis.

Example 3.1 (ERK(1,1) Scheme)**.** According to (2.8), we reformulate the first stage of the ERK scheme as

$$N_\kappa(\boldsymbol{u}_{n,0}=(a_{1,0}\tau I)^{-1}\big(\psi_1(\tau L_\kappa)\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}\big).$$

Upon performing a calculation, we obtain

$$\begin{aligned}-L_\kappa \boldsymbol{u}_{n,1}+N_\kappa(\boldsymbol{u}_{n,0}) &= (a_{1,0}\tau I)^{-1}(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\ &= \tau^{-1}\Delta_{1,1}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}),\end{aligned}$$

where $\Delta_{1,1}(\tau L_\kappa)=(a_{1,0}I)^{-1}$. Consequently, the energy stability matrix of ERK(1,1) is $M(z)=(a_{1,0})^{-1}$. Given that $a_{1,0}=1$ for the underlying RK(1,1) (explicit Euler) coefficient, $S(z)=1$ is obviously positive definite. Therefore, it follows that the ERK(1,1) scheme is unconditionally energy stable.

Example 3.2 (ERK(2,2) Scheme)**.** According to the second stage of the ERK scheme, we have

$$N_\kappa(\boldsymbol{u}_{n,1})=\left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}\left(\psi_2(\tau L_\kappa)\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,0}-a_{2,0}\tau N_\kappa(\boldsymbol{u}_{n,0})\right).$$

In turn, a careful calculation yields

$$\begin{aligned}-L_\kappa \boldsymbol{u}_{n,2}+N_\kappa(\boldsymbol{u}_{n,1}) &= -L_\kappa \boldsymbol{u}_{n,2}+\left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}\left(\psi_2(\tau L_\kappa)\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,0}-a_{2,0}\tau N_\kappa(\boldsymbol{u}_{n,0})\right)\\ &= \left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}(I+a_{2,0}\tau L_\kappa)\boldsymbol{u}_{n,2}-\left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}\boldsymbol{u}_{n,0}\\ &\quad -\left(a_{2,0}a_{1,0}^{-1}I\right)\left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}\left(\psi_1(\tau L_\kappa)\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}\right)\\ &= \left(a_{2,1}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}(I+a_{2,0}\tau L_\kappa)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})\\ &\quad +\left(a_{2,1}a_{1,0}\tau \mathrm{e}^{c_1\tau L_\kappa}\right)^{-1}(a_{1,0}-a_{2,0})I(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\ &= \tau^{-1}\Delta_{2,2}(\tau L_\kappa)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})+\tau^{-1}\Delta_{2,1}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}),\end{aligned}$$

where we have

$$\begin{aligned}\Delta_{2,1}(z) &= \left(a_{2,1}a_{1,0}\mathrm{e}^{c_1 z}\right)^{-1}(a_{1,0}-a_{2,0}),\\ \Delta_{2,2}(z) &= \left(a_{2,1}\mathrm{e}^{c_1 z}\right)^{-1}(a_{2,0}z+1).\end{aligned}$$

As a result, the energy stability matrix of ERK(2,2) are presented as follows:

$$M(z)=\begin{bmatrix}\Delta_{1,1}(z) & 0\\ \Delta_{2,1}(z) & \Delta_{2,2}(z)\end{bmatrix}=\begin{bmatrix}\dfrac{1}{a_{1,0}} & 0\\ \dfrac{a_{1,0}-a_{2,0}}{a_{2,1}a_{1,0}e^{c_1 z}} & \dfrac{1+a_{2,0}z}{a_{2,1}e^{c_1 z}}\end{bmatrix}.$$

Consider the following one parameter two-stage, second-order RK coefficients:

$$\begin{array}{c|cc} 0 & 0 & 0\\ \alpha & \alpha & 0\\ \hline & 1-\dfrac{1}{2\alpha} & \dfrac{1}{2\alpha}\end{array}, \tag{3.6}$$

and substituting these coefficients into the energy stability matrix, we compute the sequential principal minor of the associate symmetric matrix $S(z)$, which are

$$D_1=\frac{1}{\alpha},\quad D_2=\frac{-\alpha^4+2\alpha^3-2\alpha^2+\alpha(2\alpha z+2\alpha-z)e^{\alpha z}+\alpha-1/4}{\alpha^2e^{2\alpha z}}.$$

After careful verification, we conclude that the second leading principal minor, D_2, is nonnegative and monotonically increasing with respect to z for $1/2\leq\alpha\leq 2.185425$. Thus, the first leading principle minor, D_1, is naturally nonnegative. The ERK(2,2) is unconditional energy stable within the aforementioned range of α, which includes some well-known schemes such as the midpoint scheme ($\alpha=1/2$), Ralston's scheme ($\alpha=3/4$), and Heun's second-order scheme ($\alpha=1$).

Example 3.3 (ERK(3,3) Scheme)**.** For a three-stage ERK scheme, we can reformulate the nonlinear term $N_\kappa(\boldsymbol{u}_{n,2})$ as

$$N_\kappa(\boldsymbol{u}_{n,2})=\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left(\psi_3(\tau L_\kappa)\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,0}-a_{3,0}\tau N_\kappa(\boldsymbol{u}_{n,0})-a_{3,1}\tau e^{c_1\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,1})\right),$$

which will result in the following formulation:

$$\begin{aligned}
&-L_\kappa\boldsymbol{u}_{n,3}+N_\kappa(\boldsymbol{u}_{n,2})\\
=&-L_\kappa\boldsymbol{u}_{n,3}+\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left(\psi_3(\tau L_\kappa)\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,0}-a_{3,0}\tau N_\kappa(\boldsymbol{u}_{n,0})-a_{3,1}\tau e^{c_1\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,1})\right)\\
=&\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left(I+a_{3,0}\tau L_\kappa+a_{3,1}\tau e^{c_1\tau L_\kappa}L_\kappa\right)\boldsymbol{u}_{n,3}-\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\boldsymbol{u}_{n,0}\\
&-\left(a_{3,0}a_{1,0}^{-1}I\right)\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left(\psi_1(\tau L_\kappa)\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}\right)\\
&-\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}a_{3,1}\tau e^{c_1\tau L_\kappa}\left(a_{2,1}\tau e^{c_1\tau L_\kappa}\right)^{-1}\\
&\times\left(\psi_2(\tau L_\kappa)\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,0}-a_{2,0}\tau(a_{1,0}\tau)^{-1}\left(\psi_1(\tau L_\kappa)\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}\right)\right)\\
=&\left(a_{3,2}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left(I+a_{3,0}\tau L_\kappa+a_{3,1}\tau e^{c_1\tau L_\kappa}L_\kappa\right)(\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,2})\\
&+\left(a_{3,2}a_{2,1}\tau e^{c_2\tau L_\kappa}\right)^{-1}\left((a_{2,1}-a_{3,1})I+(a_{3,0}a_{2,1}-a_{3,1}a_{2,0})\tau L_\kappa\right)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})\\
&+\left(a_{3,2}a_{2,1}a_{1,0}\tau e^{c_2\tau L_\kappa}\right)^{-1}(a_{3,1}a_{2,0}+a_{2,1}a_{1,0}-a_{3,1}a_{1,0}-a_{3,0}a_{2,1})I(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\
=&\tau^{-1}\Delta_{3,3}(\tau L_\kappa)(\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,2})+\tau^{-1}\Delta_{3,2}(\tau L_\kappa)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})+\tau^{-1}\Delta_{3,1}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0}).
\end{aligned}$$

Thus, the following entries are obtained:

$$\begin{aligned}
\Delta_{3,1}(z)&=\left(a_{3,2}a_{2,1}a_{1,0}e^{c_2z}\right)^{-1}(a_{3,1}a_{2,0}+a_{2,1}a_{1,0}-a_{3,1}a_{1,0}-a_{3,0}a_{2,1}),\\
\Delta_{3,2}(z)&=\left(a_{3,2}a_{2,1}e^{c_2z}\right)^{-1}\left((a_{3,0}a_{2,1}-a_{3,1}a_{2,0})z+a_{2,1}-a_{3,1}\right),\\
\Delta_{3,3}(z)&=\left(a_{3,2}e^{c_2z}\right)^{-1}\left(a_{3,1}e^{c_1z}z+a_{3,0}z+1\right),
\end{aligned}$$

and the energy stability matrix corresponding to ERK(3,3) is

$$M(z)=\begin{bmatrix} \dfrac{1}{a_{1,0}} & 0 & 0 \\ \dfrac{a_{1,0}-a_{2,0}}{a_{2,1}a_{1,0}e^{c_1z}} & \dfrac{1+a_{2,0}z}{a_{2,1}e^{c_1z}} & 0 \\ \dfrac{m_1}{a_{3,2}a_{2,1}a_{1,0}\mathrm{e}^{c_2z}} & \dfrac{m_2}{a_{3,2}a_{2,1}\mathrm{e}^{c_2z}} & \dfrac{m_3}{a_{3,2}\mathrm{e}^{c_2z}} \end{bmatrix},$$

where

$$\begin{aligned} m_1&=a_{3,1}a_{2,0}+a_{2,1}a_{1,0}-a_{3,1}a_{1,0}-a_{3,0}a_{2,1},\\ m_2&=(a_{2,1}-a_{3,1})+(a_{3,0}a_{2,1}-a_{3,1}a_{2,0})z,\\ m_3&=1+a_{3,0}z+a_{3,1}\mathrm{e}^{c_1z}z. \end{aligned}$$

Consider the following two classical three-stage, third-order RK coefficients:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ \hline 1 & \frac{1}{4} & 0 & \frac{3}{4} \end{array}, \qquad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{8}{15} & \frac{8}{15} & 0 & 0 \\ \frac{2}{3} & \frac{1}{4} & \frac{5}{12} & 0 \\ \hline 1 & \frac{1}{4} & 0 & \frac{3}{4} \end{array}, \tag{3.7}$$

Heun's third-order scheme Wray's third-order scheme

and we present eigenvalues of the corresponding symmetric matrices $S(z)$ separately in Fig. 3. It is evident that all eigenvalues are nonnegative, confirming that the ERK(3,3) scheme under the aforementioned RK coefficients is unconditionally energy stable.

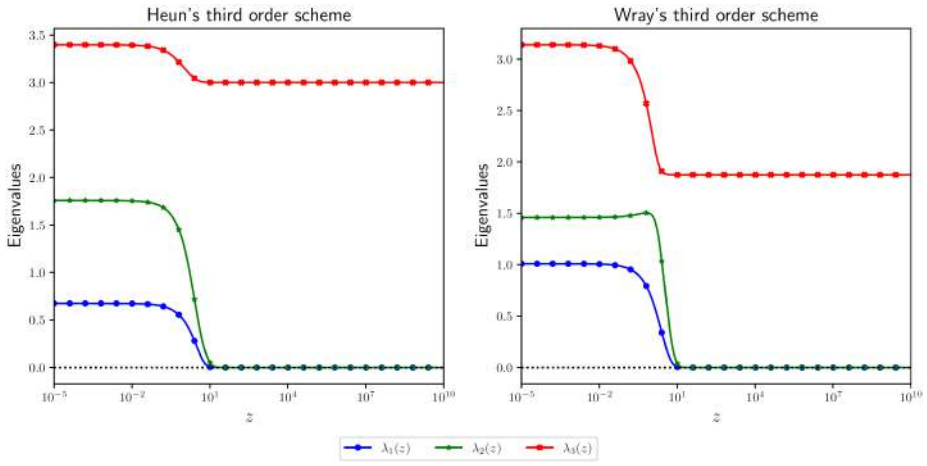


Figure 3: Eigenvalues of $S(s)$ corresponding to ERK(3,3).

Example 3.4 (ERK(4,4) Scheme)**.** For a four-stage ERK scheme, we can similarly obtain the expansion of the nonlinear term $N_\kappa(\boldsymbol{u}_{n,3})$ as

$$\begin{aligned}N_\kappa(\boldsymbol{u}_{n,3}) = \big(a_{4,3}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big(&\psi_4(\tau L_\kappa)\boldsymbol{u}_{n,4}-\boldsymbol{u}_{n,0}-a_{4,0}\tau N_\kappa(\boldsymbol{u}_{n,0})\\ &-a_{4,1}\tau \mathrm{e}^{c_1\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,1})-a_{4,2}\tau \mathrm{e}^{c_2\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,2})\big),\end{aligned}$$

which will yield the following formulation after careful calculation:

$$\begin{aligned}&-L_\kappa\boldsymbol{u}_{n,4}+N_\kappa(\boldsymbol{u}_{n,3})\\ =&-L_\kappa\boldsymbol{u}_{n,4}+\big(a_{4,3}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big(\psi_4(\tau L_\kappa)\boldsymbol{u}_{n,4}-\boldsymbol{u}_{n,0}-a_{4,0}\tau N_\kappa(\boldsymbol{u}_{n,0})\\ &\qquad-a_{4,1}\tau \mathrm{e}^{c_1\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,1})-a_{4,2}\tau \mathrm{e}^{c_2\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,2})\big)\\ =&\big(a_{4,3}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big(I+a_{4,0}\tau L_\kappa+a_{4,1}\tau e^{c_1\tau L_\kappa}L_\kappa+a_{4,2}\tau e^{c_2\tau L_\kappa}L_\kappa\big)(\boldsymbol{u}_{n,4}-\boldsymbol{u}_{n,3})\\ &+\big(a_{4,3}a_{3,2}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big((a_{3,2}-a_{4,2})I+(a_{3,2}a_{4,0}-a_{4,2}a_{3,0})\tau L_\kappa\\ &\qquad+(a_{4,1}a_{3,2}-a_{4,2}a_{3,1})\tau e^{c_1\tau L_\kappa}L_\kappa\big)(\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,2})\\ &+\big(a_{4,3}a_{3,2}a_{2,1}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big((a_{42}a_{31}-a_{4,2}a_{2,1}-a_{4,1}a_{3,2}+a_{3,2}a_{2,1})I\\ &\qquad+(a_{4,2}a_{3,1}a_{2,0}-a_{4,1}a_{3,2}a_{2,0}-a_{4,2}a_{3,0}a_{2,1}\\ &\qquad\quad+a_{4,0}a_{3,2}a_{2,1})\tau L_\kappa\big)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})\\ &+\big(a_{4,3}a_{3,2}a_{2,1}a_{1,0}\tau e^{c_3\tau L_\kappa}\big)^{-1}\big((a_{3,2}a_{2,1}a_{1,0}-a_{4,2}a_{2,1}a_{1,0}+a_{4,2}a_{3,1}a_{1,0}\\ &\qquad-a_{1,0}a_{3,2}a_{4,1}-a_{2,0}a_{3,1}a_{4,2}+a_{2,0}a_{3,2}a_{4,1}\\ &\qquad+a_{2,1}a_{3,0}a_{4,2}-a_{2,1}a_{3,2}a_{4,0})I\big)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\\ =&\tau^{-1}\big(\Delta_{4,4}(\tau L_\kappa)(\boldsymbol{u}_{n,4}-\boldsymbol{u}_{n,3})+\Delta_{4,3}(\tau L_\kappa)(\boldsymbol{u}_{n,3}-\boldsymbol{u}_{n,2})\\ &\quad+\Delta_{4,2}(\tau L_\kappa)(\boldsymbol{u}_{n,2}-\boldsymbol{u}_{n,1})+\Delta_{4,1}(\tau L_\kappa)(\boldsymbol{u}_{n,1}-\boldsymbol{u}_{n,0})\big).\end{aligned}$$

Consequently, the following entries are established for the fourth row of the energy stability matrix:

$$\begin{aligned}\Delta_{4,1}(z) =& \big(a_{4,3}a_{3,2}a_{2,1}a_{1,0}e^{c_3z}\big)^{-1}\big(a_{3,2}a_{2,1}a_{1,0}-a_{4,2}a_{2,1}a_{1,0}+a_{4,2}a_{3,1}a_{1,0}\\ &-a_{1,0}a_{3,2}a_{4,1}-a_{2,0}a_{3,1}a_{4,2}+a_{2,0}a_{3,2}a_{4,1}\\ &+a_{2,1}a_{3,0}a_{4,2}-a_{2,1}a_{3,2}a_{4,0}\big),\\ \Delta_{4,2}(z) =& (a_{4,3}a_{3,2}a_{2,1}e^{c_3z})^{-1}\big((a_{4,2}a_{3,1}a_{2,0}-a_{4,1}a_{3,2}a_{2,0}-a_{4,2}a_{3,0}a_{2,1}\\ &+a_{4,0}a_{3,2}a_{2,1})z+a_{42}a_{31}-a_{4,2}a_{2,1}\\ &-a_{4,1}a_{3,2}+a_{3,2}a_{2,1}\big),\\ \Delta_{4,3}(z) =& \big(a_{4,3}a_{3,2}e^{c_3z}\big)^{-1}\big((a_{4,1}a_{3,2}-a_{4,2}a_{3,1})e^{c_1z}z+(a_{3,2}a_{4,0}-a_{4,2}a_{3,0})z\\ &+a_{3,2}-a_{4,2}\big),\\ \Delta_{4,4}(z) =& \big(a_{4,3}e^{c_3z}\big)^{-1}\big(a_{4,1}e^{c_1z}z+a_{4,2}e^{c_2z}z+a_{4,0}z+1\big).\end{aligned}$$

Consider the classical RK(4,4) coefficients

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline 1 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array},$$

and its energy stability matrix

$$M(z)=\begin{bmatrix} 2 & 0 & 0 & 0 \\ 2e^{-\frac{z}{2}} & 2e^{-\frac{z}{2}} & 0 & 0 \\ e^{-\frac{z}{2}} & e^{-\frac{z}{2}} & e^{-\frac{z}{2}} & 0 \\ -2e^{-z} & ze^{-z} & \left(2e^{\frac{z}{2}}z+z+4\right)e^{-z} & \left(4e^{\frac{z}{2}}z+z+6\right)e^{-z} \end{bmatrix}.$$

We present the corresponding eigenvalue of $S(z)$ in Fig. 4. It is observed that the eigenvalue is not always nonnegative for $z \geq 0$. Therefore, the ERK(4,4) can not theoretically guarantee energy stability.

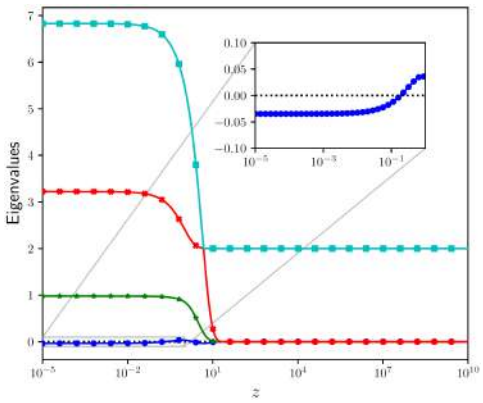


Figure 4: Eigenvalues of $S(z)$ corresponding to ERK$(4,4)$.

Example 3.5 (ERK(5,4) Scheme)**.** To avoid the tedious of presenting the formulation and to conserve space, we directly give the entries for the fifth row of the energy stability matrix as follows:

$$\begin{aligned} \Delta_{5,1}(z) = &\left(a_{5,4}a_{4,3}a_{3,2}a_{2,1}a_{1,0}e^{c_4 z}\right)^{-1} \\ &\times \left(a_{1,0}a_{2,1}a_{3,2}a_{4,3} - a_{1,0}a_{2,1}a_{3,2}a_{5,3} + a_{1,0}a_{2,1}a_{4,2}a_{5,3} - a_{1,0}a_{2,1}a_{4,3}a_{5,2}\right. \\ &\left. - a_{1,0}a_{3,1}a_{4,2}a_{5,3} + a_{1,0}a_{3,1}a_{4,3}a_{5,2} + a_{1,0}a_{3,2}a_{4,1}a_{5,3} - a_{1,0}a_{3,2}a_{4,3}a_{5,1}\right. \end{aligned}$$

$$
\begin{aligned}
&+a_{2,0}a_{3,1}a_{4,2}a_{5,3}-a_{2,0}a_{3,1}a_{4,3}a_{5,2}-a_{2,0}a_{3,2}a_{4,1}a_{5,3}+a_{2,0}a_{3,2}a_{4,3}a_{5,1}\\
&-a_{2,1}a_{3,0}a_{4,2}a_{5,3}+a_{2,1}a_{3,0}a_{4,3}a_{5,2}+a_{2,1}a_{3,2}a_{4,0}a_{5,3}-a_{2,1}a_{3,2}a_{4,3}a_{5,0}\big),\\
\Delta_{5,2}(z)=&\big(a_{5,4}a_{4,3}a_{3,2}a_{2,1}e^{c_4z}\big)^{-1}\\
&\times\big(a_{2,0}a_{3,1}a_{4,3}a_{5,2}z-a_{2,0}a_{3,1}a_{4,2}a_{5,3}z+a_{2,0}a_{3,2}a_{4,1}a_{5,3}z-a_{2,0}a_{3,2}a_{4,3}a_{5,1}z\\
&+a_{2,1}a_{3,0}a_{4,2}a_{5,3}z-a_{2,1}a_{3,0}a_{4,3}a_{5,2}z-a_{2,1}a_{3,2}a_{4,0}a_{5,3}z+a_{2,1}a_{3,2}a_{4,3}a_{5,0}z\\
&+a_{2,1}a_{3,2}a_{4,3}-a_{2,1}a_{3,2}a_{5,3}+a_{2,1}a_{4,2}a_{5,3}-a_{2,1}a_{4,3}a_{5,2}-a_{3,1}a_{4,2}a_{5,3}\\
&+a_{3,1}a_{4,3}a_{5,2}+a_{3,2}a_{4,1}a_{5,3}-a_{3,2}a_{4,3}a_{5,1}\big),\\
\Delta_{5,3}(z)=&\big(a_{5,4}a_{4,3}a_{3,2}e^{c_4z}\big)^{-1}\\
&\times\big((a_{3,1}a_{4,2}a_{5,3}-a_{3,1}a_{4,3}a_{5,2}-a_{3,2}a_{4,1}a_{5,3}+a_{3,2}a_{4,3}a_{5,1})e^{c_1z}z\\
&+(a_{3,0}a_{4,2}a_{5,3}-a_{3,0}a_{4,3}a_{5,2}-a_{3,2}a_{4,0}a_{5,3}+a_{3,2}a_{4,3}a_{5,0})z\\
&+a_{3,2}a_{4,3}-a_{3,2}a_{5,3}+a_{4,2}a_{5,3}-a_{4,3}a_{5,2}\big),\\
\Delta_{5,4}(z)=&\big(a_{5,4}a_{4,3}e^{c_4z}\big)^{-1}\\
&\times\big((a_{4,3}a_{5,1}-a_{4,1}a_{5,3})e^{c_1z}z+(a_{4,3}a_{5,2}-a_{4,2}a_{5,3})e^{c_2z}z\\
&+(a_{4,3}a_{5,0}-a_{4,0}a_{5,3})z+a_{4,3}-a_{5,3}\big),\\
\Delta_{5,5}(z)=&\big(a_{5,4}e^{c_4z}\big)^{-1}(a_{5,1}e^{c_1z}z+a_{5,2}e^{c_2z}z+a_{5,3}e^{c_3z}z+a_{5,0}z+1).
\end{aligned}
$$

We introduce a specific five-stage, fourth-order RK coefficients, commonly referred to as SSPRK$^+$(5,4) [18] with nondecreasing abscissas $c\approx[0,0.4549,0.5165,0.5165,0.9903]^\top$

c_0	0	0	0	0	0
c_1	0.454933915986784	0	0	0	0
c_2	0.196211867856647	0.320289519000399	0	0	0
c_3	0.080487729600967	0.131385407421748	0.30462824983433	0	0
c_4	0.063305675368379	0.103338011789111	0.23959797583628	0.584088613339939	0
1	0.163796836877076	0.231812658796517	0.12523964308588	0.305307460173213	0.173843401067318

In Fig. 5, we show the eigenvalues of the corresponding $S(z)$, in which the minimum eigenvalue is calculated as 0. It is evident that all eigenvalues are nonnegative, indicating that the ERK(5,4) scheme with the above RK coefficients is unconditionally energy stable.

Remark 3.1. In this section, we introduce a set of RK coefficients that could theoretically guarantee the unconditional energy stability of the proposed ERK scheme. To avoid confusion and to facilitate the analysis in the subsequent sections, unless stated otherwise, we will refer the ERK(2,2) schemes based on Heun's second-order RK coefficients (i.e. chosen $\alpha=1$ in (3.6)) and the ERK(3,3) scheme utilizing Heun's third-order RK coefficients (3.7).

Remark 3.2. As a classical numerical scheme, the multi-step method has been extensively studied and proven to possess certain discrete gradient structures [2, 3, 13, 31, 32], which

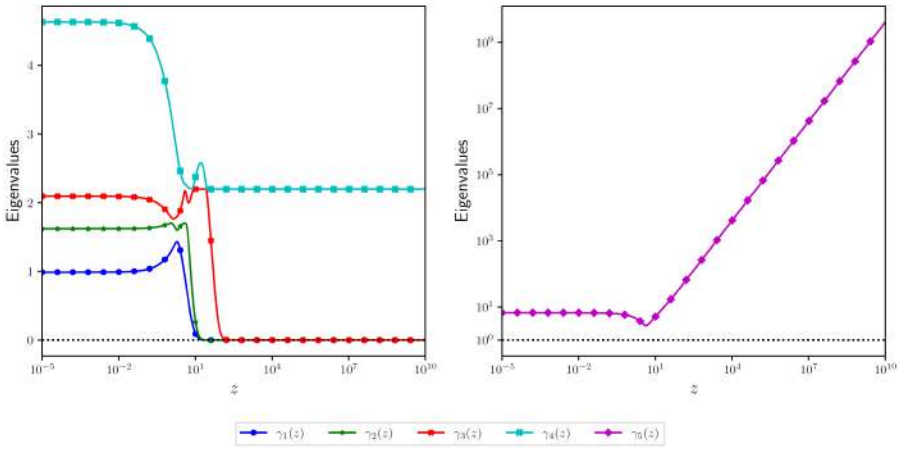


Figure 5: Eigenvalues of $S(z)$ corresponding to ERK(5,4).

preserves the energy dissipation for gradient flow equations with respect to the original energy added numerical correction terms. In contrast to the previous research that focused on multi-step schemes, our proposed method can maintain the original energy dissipation without incorporating additional terms. Furthermore, the multi-step scheme requires solving nonlinear equations at each time step since it is typically constructed based on the convex splitting scheme, and the ghost point problem at the initial stage has always been a difficulty that hinders the construction of a high-order scheme. Our proposed framework can easily construct linear, high-order, and unconditional energy-stable schemes. Therefore, the ERK schemes we constructed are superior to the traditional multi-step method when dealing with the gradient flow equation.

4 Convergence analysis

In this section, we give a rigorous convergence estimate for ERK schemes. Although the analysis is akin for different gradient flow equations, there are some differences in details. Therefore, we choose the no-slope selection (NSS) thin film growth equation as an example, whose nonlinear term naturally satisfies the Lipschitz condition, to conduct the estimate.

The NSS equation, as a class of L^2 gradient flow model of the energy functional (1.1) with

$$\mathcal{L}=\varepsilon^2\Delta^2,\quad F(u)=-\frac{1}{2}\ln(1+|\nabla u|^2),$$

and

$$f(u)=\nabla\cdot\left(\frac{\nabla u}{1+|\nabla u|^2}\right)$$

in (1.2), has a wide range of applications in the fields of physics and material science.

4.1 The Fourier pseudo-spectral discretization

Without loss of generality and to facilitate the subsequent analysis, we impose periodic boundary conditions for the NSS equation and employ the Fourier pseudo-spectral method to construct the spatial approximation on a 2D rectangular domain

$$\Omega=(x_0,x_0+L_x)\times(y_0,y_0+L_y).$$

The results for 3D could be obtained similarly without difficulties. To facilitate the representation of the collocation method, we denote N_x and N_y as two positive even integers and introduce the following index sets:

$$\begin{aligned}\mathcal{S}_N&=\{(p,q)\in\mathbb{Z}^2\,|\,0\le p\le N_x-1,0\le q\le N_y-1\},\\ \widehat{\mathcal{S}}_N&=\left\{(k,l)\in\mathbb{Z}^2\,\middle|\,-\frac{N_x}{2}\le k\le\frac{N_x}{2}-1,-\frac{N_y}{2}\le l\le\frac{N_y}{2}-1\right\}.\end{aligned}$$

The domain Ω is uniformly divided into an $N_x\times N_y$ mesh Ω_N, which consists of a set of nodes (x_p,y_q) with $x_p=x_0+ph_x, y_q=y_0+qh_y$, and $(p,q)\in\mathcal{S}_N$. Here, $h_x=L_x/N_x$ and $h_y=L_y/N_y$ represent the mesh size in each direction. In addition, the two-dimensional periodic grid function space is denoted as $\mathcal{M}_N$, i.e.

$$\mathcal{M}_N=\left\{\boldsymbol{v}:\Omega_N\to\mathbb{R}\,|\,v_{p+mN_x,q+nN_y}=v_{p,q},\forall(p,q)\in\mathcal{S}_N,(m,n)\in\mathbb{Z}^2\right\}.$$

For any $\boldsymbol{v},\boldsymbol{w}\in\mathcal{M}_N$, we denote the discrete L^2 inner product $\langle\cdot,\cdot\rangle$, discrete L^2 inner product $\|\cdot\|$, and the L^∞ norm as follows:

$$\langle\boldsymbol{v},\boldsymbol{w}\rangle=h_xh_y\sum_{(p,q)\in\mathcal{S}_N}v_{p,q}w_{p,q},\quad\|\boldsymbol{v}\|=\sqrt{\langle\boldsymbol{v},\boldsymbol{v}\rangle},\quad\|\boldsymbol{v}\|_\infty=\max_{(p,q)\in\mathcal{S}_N}|v_{p,q}|.$$

For a periodic grid function $\boldsymbol{g}\in\mathcal{M}_N$, we define its discrete Fourier expansion as

$$g_{p,q}=\sum_{(k,l)\in\widehat{\mathcal{S}}_N}\hat{g}_{k,l}\mathrm{e}^{\mathrm{i}\mu_xk(x_p-x_0)+\mathrm{i}\mu_yl(y_q-y_0)},\quad(p,q)\in\mathcal{S}_N,$$

where $\mu_x=2\pi/L_x$, and $\mu_y=2\pi/L_y$. The coefficients $\hat{g}_{k,l}$ are obtained using the discrete Fourier transform (DFT) as follows:

$$\hat{g}_{k,l}=\frac{1}{N_xN_y}\sum_{(p,q)\in\mathcal{S}_N}f_{p,q}\mathrm{e}^{-\mathrm{i}\mu_xk(x_p-x_0)-\mathrm{i}\mu_yl(y_q-y_0)},\quad(k,l)\in\widehat{\mathcal{S}}_N.$$

The Fourier collocation approximations to the first- and second-order partial derivatives in the x-direction, respectively, become

$$\begin{aligned}(D_x\boldsymbol{g})_{p,q}&=\sum_{(k,l)\in\widehat{\mathcal{S}}_N}(\mathrm{i}\mu_xk)\hat{g}_{k,l}\mathrm{e}^{\mathrm{i}\mu_xk(x_p-x_0)+\mathrm{i}\mu_yl(y_q-y_0)},\\ \left(D_x^2\boldsymbol{g}\right)_{p,q}&=\sum_{(k,l)\in\widehat{\mathcal{S}}_N}\left(-\mu_x^2k^2\right)\hat{g}_{k,l}\mathrm{e}^{\mathrm{i}\mu_xk(x_p-x_0)+\mathrm{i}\mu_yl(y_q-y_0)}.\end{aligned}$$

Naturally, the differential operator in the y-direction can be defined in the same way. Therefore, for any $\boldsymbol{g}\in\mathcal{M}_N,(\boldsymbol{g}_1,\boldsymbol{g}_2)^\top\in\mathcal{M}_N\times\mathcal{M}_N$, the discrete gradient, divergence, and Laplace operator are given respectively by

$$\nabla_N\boldsymbol{g}=\begin{pmatrix}D_x\boldsymbol{g}\\D_y\boldsymbol{g}\end{pmatrix},\quad \nabla_N\cdot\begin{pmatrix}\boldsymbol{g_1}\\\boldsymbol{g_2}\end{pmatrix}=D_xg_1+D_yg_2,\quad \Delta_N\boldsymbol{g}=\left(D_x^2\boldsymbol{g}+D_y^2\boldsymbol{g}\right).$$

The following proposition collects some useful corollaries without proof.

Proposition 4.1. *The grid functions and the discrete differential operators have the following properties:*

(i) $\langle\boldsymbol{v},\boldsymbol{w}\rangle=L_xL_y\sum_{(k,l)\in\widehat{\mathcal{S}}_N}\hat{v}_{k,l}\overline{\hat{w}}_{k,l}$ *for any* $\boldsymbol{v},\boldsymbol{w}\in\mathcal{M}_N$.

(ii) $\langle\Delta_N\boldsymbol{v},\boldsymbol{w}\rangle=-\langle\nabla_N\boldsymbol{v},\nabla_N\boldsymbol{w}\rangle=\langle\boldsymbol{v},\Delta_N\boldsymbol{w}\rangle$ *for any* $\boldsymbol{v},\boldsymbol{w}\in\mathcal{M}_N$.

(iii) $\langle\Delta_N\boldsymbol{v},\mathbf{1}\rangle=0$ *for any* $\boldsymbol{v}\in\mathcal{M}_N$.

(iv) Δ_N *is negative semi-definite on* $\mathcal{M}_N$.

By applying the above Fourier pseudo-spectral discretization, adding and subtracting the stabilization term $\kappa\Delta_N\boldsymbol{u}$, we obtain the corresponding semi-discrete system (2.1) for the NSS equation with

$$\begin{aligned}&L_\kappa:=\varepsilon^2\Delta_N^2-\kappa\Delta_N,\quad N_\kappa(\boldsymbol{u}):=-\nabla_N\cdot\boldsymbol{\beta}_\kappa(\nabla_N\boldsymbol{u}),\\&\boldsymbol{\beta}_\kappa(v)=\boldsymbol{\beta}(v)+\kappa\boldsymbol{v},\quad \boldsymbol{\beta}(\boldsymbol{v}):=\frac{\boldsymbol{v}}{1+|\boldsymbol{v}|^2}.\end{aligned}$$

Here, we denote the eigenvalues of $-\Delta_N$ and L_κ as

$$\lambda_{k,l}=\left(\mu_x^2k^2+\mu_y^2l^2\right)\quad\text{and}\quad\Lambda_{k,l}=\varepsilon^2\lambda_{k,l}^2+\kappa\lambda_{k,l},$$

respectively, for $(k,l)\in\widehat{S}_N$ in Fourier space.

Remark 4.1. In this study, we only consider the equation imposed with periodic boundary conditions and utilize the Fourier spectral collocation method for spatial discretization. This approach facilitates theoretical analysis and allows us to ignore the spatial discretization error relative to the temporal part, especially when a sufficiently fine mesh is employed in numerical experiments. For other boundary conditions, such as the Neumann or Dirichlet ones, similar theoretical techniques could be applied. For instance, when the homogeneous Neumann boundary condition is imposed, the solution could be approximated by the Fourier cosine transformation. The theoretical technique can be similarly applied, and spectral accuracy could also be achievable for spatial discretization in experiments. A similar conclusion is also applicable if the homogeneous Dirichlet boundary conditions are imposed, in which we will take the Fourier sine transformation to approximate the corresponding solution.

4.2 Truncation error of the ERK framework

To determine the accuracy of the ERK schemes, we first establish generic order conditions for an ERK(s,p) scheme. Subsequently, to avoid an excessive discussion length, we focus on the convergence analysis for the third-order ERK method underlying Heun's third-order coefficients in (3.7) for solving the NSS equation. To some extent, the analysis for other ERK schemes follows a similar approach, and we leave the details to the interested reader.

We begin by proposing the following lemma to illustrate that the ERK scheme can be reformulated as the standard explicit RK-type with parametric coefficients.

Lemma 4.1. *The ERK scheme* (2.8) *can be reformulated into the following standard explicit RK-type:*

$$\begin{cases} \boldsymbol{u}_{n,0}=\boldsymbol{u}^n, \\ \boldsymbol{u}_{n,i}=\boldsymbol{u}^n+\tau\sum\limits_{j=0}^{i-1}\hat{a}_{i,j}(\tau L_\kappa)\big(-L_\kappa\boldsymbol{u}_{n,j}+N_\kappa(\boldsymbol{u}_{n,j})\big), \quad i=1,\ldots,s, \\ \boldsymbol{u}^{n+1}=\boldsymbol{u}_{n,i}, \end{cases} \tag{4.1}$$

where the coefficients $\hat{a}_{i,j}(\tau L_\kappa)$ are explicitly formulated as

$$\hat{a}_{i,j}(\tau L_\kappa)=\psi_i^{-1}(\tau L_\kappa)\left[a_{i,j}\mathrm{e}^{c_j\tau L_\kappa}+\tau L_\kappa\sum_{k=j+1}^{i-1}a_{i,k}\mathrm{e}^{c_k\tau L_\kappa}\hat{a}_{k,j}(\tau L_\kappa)\right]. \tag{4.2}$$

Proof. We prove this by induction. First, we reformulate the first stage of the ERK scheme as follows:

$$\begin{aligned} \boldsymbol{u}_{n,1}&=\psi_1^{-1}(\tau L_\kappa)\big[\boldsymbol{u}^n+\tau a_{1,0}N_\kappa(\boldsymbol{u}^n)\big] \\ &=\boldsymbol{u}^n+\psi_1^{-1}(\tau L_\kappa)\big[\big(1-\psi_1(\tau L_\kappa)\big)\boldsymbol{u}^n+\tau a_{1,0}N_\kappa(\boldsymbol{u}^n)\big] \\ &=\boldsymbol{u}^n+\psi_1^{-1}(\tau L_\kappa)\big[-\tau a_{1,0}L_\kappa\boldsymbol{u}^n+\tau a_{1,0}N_\kappa(\boldsymbol{u}^n)\big] \\ &=\boldsymbol{u}^n+\psi_1^{-1}(\tau L_\kappa)\tau a_{1,0}\big(-L_\kappa\boldsymbol{u}^n+N_\kappa(\boldsymbol{u}^n)\big). \end{aligned}$$

Denoting $\hat{a}_{1,0}(\tau L_\kappa)=\psi_1^{-1}(\tau L_\kappa)a_{1,0}$, we obtain

$$\boldsymbol{u}_{n,1}=\boldsymbol{u}^n+\tau\hat{a}_{1,0}(\tau L_\kappa)\big(-L_\kappa\boldsymbol{u}^n+N_\kappa(\boldsymbol{u}^n)\big). \tag{4.3}$$

Assume that

$$\boldsymbol{u}_{n,j}=\boldsymbol{u}^n+\tau\sum_{k=0}^{j-1}\hat{a}_{j,k}(\tau L_\kappa)\big(-L_\kappa\boldsymbol{u}_{n,k}+N_\kappa(\boldsymbol{u}_{n,k})\big), \quad j=1,\ldots,i-1. \tag{4.4}$$

By using the induction assumption and definition (2.7) of the operator $\psi_i(\cdot)$, we can derive

$$\boldsymbol{u}_{n,i}=\psi_i^{-1}(\tau L_\kappa)\left[\boldsymbol{u}^n+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{c_j\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,j})\right]$$

$$
\begin{aligned}
&= \boldsymbol{u}^n + \psi_i^{-1}(\tau L_\kappa)\left[-\tau\sum_{j=0}^{i-1} a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} L_\kappa \boldsymbol{u}^n + \tau\sum_{j=0}^{i-1} a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} N_\kappa(\boldsymbol{u}_{n,j})\right] \\
&= \boldsymbol{u}^n + \psi_i^{-1}(\tau L_\kappa)\left[-\tau\sum_{j=0}^{i-1} a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} L_\kappa \left(\boldsymbol{u}_{n,j} - \tau\sum_{k=0}^{j-1}\hat{a}_{j,k}(\tau L_\kappa)\big(-L_\kappa \boldsymbol{u}_{n,k} + N_\kappa(\boldsymbol{u}_{n,k})\big)\right)\right. \\
&\qquad\qquad \left. + \tau\sum_{j=0}^{i-1} a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} N_\kappa(\boldsymbol{u}_{n,j})\right] \\
&= \boldsymbol{u}^n + \tau\psi_i^{-1}(\tau L_\kappa)\sum_{j=0}^{i-1}\left[a_{i,j}\mathrm{e}^{c_j\tau L_\kappa} + \tau L_\kappa \sum_{k=j+1}^{i-1} a_{i,k}\mathrm{e}^{c_k\tau L_\kappa}\hat{a}_{k,j}(\tau L_\kappa)\right]\big(-L_\kappa \boldsymbol{u}_{n,j} + N_\kappa(\boldsymbol{u}_{n,j})\big) \\
&= \boldsymbol{u}^n + \tau\sum_{j=0}^{i-1}\hat{a}_{i,j}(\tau L_\kappa)\big(-L_\kappa \boldsymbol{u}_{n,j} + N_\kappa(\boldsymbol{u}_{n,j})\big)
\end{aligned}
$$

(the summation $\sum_{k=j+1}^{i-1}$ is not performed for $j=i-1$). The result is proven. □

To analyze the order conditions of ERK schemes, we introduce the following autonomous ordinary differential equation (ODE) with an additional stabilization term $\kappa(u-u), \kappa>0$, i.e.

$$
\begin{cases} u_t = f(u) + \kappa u - \kappa u, \quad \forall t \in (0,T], \\ u(0) = u^0. \end{cases} \tag{4.5}
$$

Solving this equation using the ERK scheme and utilizing Lemma 4.1, we obtain the following discretization:

$$
\begin{cases} u_{n,0} = u^n, \\ u_{n,i} = u^n + \tau\sum_{j=0}^{i-1}\hat{a}_{i,j}(\tau\kappa) f(u_{n,j}), \quad i=1,\dots,s, \\ u^{n+1} = u_{n,s}, \end{cases} \tag{4.6}
$$

where

$$
\hat{a}_{i,j}(\tau\kappa) = \psi_i^{-1}(c_i\tau\kappa)\left[a_{i,j}\mathrm{e}^{c_j\tau\kappa} + \tau\kappa\sum_{k=j+1}^{i-1} a_{i,k}\mathrm{e}^{c_k\tau\kappa}\hat{a}_{k,j}(\tau\kappa)\right].
$$

The parametric Butcher tableau is given by

$$
\begin{array}{cc|c} c & \hat{c} & \hat{A} \\ \hline c_s & \hat{c}_s & \hat{b}^\top \end{array} = \begin{array}{cc|cccc} c_0 & \hat{c}_0 & 0 & 0 & 0 & 0 \\ c_1 & \hat{c}_1 & \hat{a}_{1,0} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ c_{s-1} & \hat{c}_{s-1} & \hat{a}_{s-1,0} & \hat{a}_{s-1,1} & \dots & 0 \\ \hline c_s & \hat{c}_s & \hat{a}_{s,0} & \hat{a}_{s,1} & \dots & \hat{a}_{s,s-1} \end{array}, \tag{4.7}
$$

where $\hat{a}_{i,j}$ represents the abbreviation of $\hat{a}_{i,j}(\tau\kappa)$, and the new parametric abscissas are obtained by $\hat{c}_i(\tau\kappa)=\sum_{j=0}^{i-1}\hat{a}_{ij}(\tau\kappa)$. Subsequently, we study the accuracy of the ERK scheme.

Lemma 4.2. *Assume the problem* (4.5) *has an exact solution* $u\in C^{p+1}[0,T]$ *and* $u^n=u(t_n)$. *If the s-stage ERK scheme meets the p-th-order condition* ($s=p, p\leq 3$) *in Table* 1, *then the ERK has a* $\mathcal{O}(\tau^{p+1})$ *local truncation error, which is*

$$u(t_{n+1})-u^{n+1}=\mathcal{O}(\tau^{p+1}).$$

Proof. The proof follows similar idea as that of [48, Lemma 3.1]. The details are skipped for the sake of brevity. □

Table 1: Order condition for ERK scheme.

Scheme	ERK(1,1)	ERK(2,2)	ERK(3,3)
Order condition	$1\quad -\hat{c}_1=\mathcal{O}(\tau)$	$1-\hat{c}_2=\mathcal{O}(\tau^2)$, $\frac{1}{2}-\hat{a}_{2,1}\hat{c}_1=\mathcal{O}(\tau)$	$1-\hat{c}_3=\mathcal{O}(\tau^3)$, $\frac{1}{2}-\hat{a}_{3,1}\hat{c}_1-\hat{a}_{3,2}\hat{c}_2=\mathcal{O}(\tau^2)$, $\frac{1}{3}-\hat{a}_{3,1}\hat{c}_1^2-\hat{a}_{3,2}\hat{c}_2^2=\mathcal{O}(\tau)$, $\frac{1}{6}-\hat{a}_{3,2}\hat{a}_{2,1}\hat{c}_1=\mathcal{O}(\tau)$

Remark 4.2. Through the use of symbolic mathematical tools, we have confirmed that all the first- to fourth-order RK coefficients mentioned in the preceding section, which adhere to the unconditional energy stability conditions of the ERK method, also fulfill the corresponding order conditions. For ease of reference, unless stated otherwise, in the subsequent sections of this paper, we denote by ERK(2,2), ERK(3,3) and ERK(5,4) the ERK schemes based on Heun's second- and third-order, and five stage, fourth-order RK coefficients, respectively.

4.3 Convergence estimate for the ERK(3,3) scheme

We first introduce a lemma concerning the nonlinear function in the NSS equation, which holds significant importance in the following analysis.

Lemma 4.3 ([21])**.** *For the mapping*

$$\boldsymbol{\beta}:\mathbb{R}^2\ \rightarrow\ \mathbb{R}^2,\quad \boldsymbol{\beta}(\boldsymbol{v})=\frac{\boldsymbol{v}}{1+|\boldsymbol{v}|^2},$$

the corresponding Jacobian matrix at $\boldsymbol{v}=(v_1,v_2)$ *is*

$$\nabla\boldsymbol{\beta}(\boldsymbol{v})=\frac{1}{(1+|\boldsymbol{v}|^2)^2}\begin{pmatrix}1-v_1^2+v_2^2 & -2v_1v_2\\ -2v_1v_2 & 1+v_1^2-v_2^2\end{pmatrix}.$$

The eigenvalues of $\nabla\boldsymbol{\beta}(\boldsymbol{v})$ *are*

$$\mu_1(\boldsymbol{v})=\frac{1-|\boldsymbol{v}|^2}{(1+|\boldsymbol{v}|^2)^2},\quad \mu_2(\boldsymbol{v})=\frac{1}{1+|\boldsymbol{v}|^2},$$

and they satisfy $-1/8\le\mu_1(\boldsymbol{v})\le\mu_2(\boldsymbol{v})\le 1$. *Then, we have*

$$|\boldsymbol{\beta}(\boldsymbol{v})-\boldsymbol{\beta}(\boldsymbol{w})|\le|\boldsymbol{v}-\boldsymbol{w}|,\quad \forall\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^2. \tag{4.8}$$

In the rest of this section, we will conduct the convergence estimate for the NSS equation in detail. With a sufficiently smooth initial data u^0, we assume the exact solution of the NSS equation preserves a regularity of class $\mathcal{R}$

$$u_e\in\mathcal{R}:=H^4(0,T;C^1)\cap H^3(0,T;H^{m_0+3})\cap L^\infty(0,T;H^{m_0+5}).$$

At the same time, we denote U^n as the interpolation values of the exact solution u_e at the mesh Ω_N at the fixed time t_n. This means $U^n_{p,q}:=u_e(x_p,y_q,t_n),(p,q)\in\mathcal{S}_N$. The initial data is given by

$$U^0_{p,q}=\boldsymbol{u}^0_{p,q}:=u_0(x_p,y_q). \tag{4.9}$$

For the exact solution u_e and its interpolation U, the ERK(3,3) scheme is considered, and a careful consistency analysis implies that

$$\begin{cases} U_{n,1}=\psi_1^{-1}(\tau L_\kappa)\left(U^n+\dfrac{1}{3}\tau N_\kappa(U^n)\right),\\ U_{n,2}=\psi_2^{-1}(\tau L_\kappa)\left(U^n+\dfrac{2}{3}\tau \mathrm{e}^{\frac{1}{3}\tau L_\kappa}N_\kappa(U_{n,1})\right),\\ U^{n+1}=\psi_3^{-1}(\tau L_\kappa)\left(U^n+\dfrac{1}{4}\tau N_\kappa(U^n)+\dfrac{3}{4}\tau \mathrm{e}^{\frac{2}{3}\tau L_\kappa}N_\kappa(U_{n,2})\right)+\tau R^n \end{cases} \tag{4.10}$$

with $\|\nabla_N R^n\|_2\le C(\tau^3+h^m)$. Here, according to (2.7), the functions $\psi_i(\cdot)$ have following form:

$$\psi_1(z)=1+\frac{1}{3}z,\quad \psi_2(z)=1+\frac{2}{3}\mathrm{e}^{\frac{1}{3}z}z,\quad \psi_3(z)=1+\frac{1}{4}z+\frac{3}{4}\mathrm{e}^{\frac{2}{3}z}z.$$

Notice that the profiles $U_{n,1}$ and $U_{n,2}$ are the reference solutions constructed based on the exact interpolation solution U^n.

To facilitate the convergence analysis for the ERK(3,3) scheme, we present the following preliminary estimates that are needed in the later analysis.

Proposition 4.2. *For any* $z\ge 0$ *and* $m=1,2,3$, $\psi_m(z)>0$ *are increasing and have following estimates:*

$$0\le\psi_1^{-\frac{1}{2}}(z)\left(\frac{z}{\psi_1(z)-1}\right)^{\frac{1}{2}}\le\sqrt{3}, \tag{4.11}$$

$$0\leq \mathrm{e}^{\frac{1}{3}z}\psi_2^{-\frac{1}{2}}(z)\left(\frac{z}{\psi_2(z)-1}\right)^{\frac{1}{2}}\leq\frac{\sqrt{6}}{2}, \tag{4.12}$$

$$0\leq \psi_3^{-\frac{1}{2}}(z)\left(\frac{z}{\psi_3(z)-1}\right)^{\frac{1}{2}}\leq 1, \tag{4.13}$$

$$0\leq \mathrm{e}^{\frac{2}{3}z}\psi_3^{-\frac{1}{2}}(z)\left(\frac{z}{\psi_3(z)-1}\right)^{\frac{1}{2}}\leq 1. \tag{4.14}$$

Lemma 4.4. *For any periodic grid function $\boldsymbol{v},\boldsymbol{w}\in\mathcal{M}_N$ and $m=1,2,3$, we have*

$$\langle G_m\boldsymbol{v},-\Delta_N\boldsymbol{v}\rangle=\varepsilon^2\tau\left\|(G_m^*)^{\frac{1}{2}}\nabla_N\Delta_N\boldsymbol{v}\right\|^2+\kappa\tau\left\|(G_m^*)^{\frac{1}{2}}\Delta_N\boldsymbol{v}\right\|^2, \tag{4.15}$$

$$\langle G_m\psi_m(\tau L_\kappa)\boldsymbol{v},-\Delta_N\boldsymbol{v}\rangle+\|\nabla_N(\boldsymbol{w}-\boldsymbol{v})\|^2\geq\langle G_m\boldsymbol{w},-\Delta_N\boldsymbol{w}\rangle, \tag{4.16}$$

where the operators G_m and G_m^* are defined as

$$G_m:=1-\psi_m^{-1}(\tau L_\kappa),\quad G_m^*:=\frac{1-\psi_m^{-1}(\tau L_\kappa)}{\tau L_\kappa}.$$

Proof. According to the definition, the discrete Fourier expansion of $-\psi_m^{-1}(\tau L_\kappa)\Delta_N v$ is

$$\left(-\psi_m^{-1}\Delta_N\boldsymbol{v}\right)_{i,j}=\sum_{k,l\in\hat{\mathcal{S}}_{\mathcal{N}}}\psi_m^{-1}(\tau\Lambda_{k,l})\lambda_{k,l}\hat{v}_{k,l}\mathrm{e}^{\mathrm{i}\mu_x k(x_i-x_0)+\mathrm{i}\mu_y(y_j-y_0)}.$$

Consequentially, for identity (4.15), the Parserval theorem yields

$$\begin{aligned}\langle G_m\boldsymbol{v},-\Delta_N\boldsymbol{v}\rangle&=\tau\left\langle\left(1-\psi_m^{-1}(\tau L_\kappa)\right)(\tau L_\kappa)^{-1}L_\kappa\boldsymbol{v},-\Delta_N\boldsymbol{v}\right\rangle\\&=\tau L_xL_y\sum_{k,l\in\hat{S}_h}\frac{1-\psi_m^{-1}(\tau\Lambda_{k,l})}{\tau\Lambda_{k,l}}\Lambda_{k,l}\lambda_{k,l}|\hat{v}_{k,l}|^2\\&=\tau L_xL_y\sum_{k,l\in\hat{S}_h}\frac{1-\psi_m^{-1}(\tau\Lambda_{k,l})}{\tau\Lambda_{k,l}}\left(\varepsilon^2\lambda_{k,l}^3+\kappa\lambda_{k,l}^2\right)|\hat{v}_{k,l}|^2\\&=\varepsilon^2\tau\left\|\left(\frac{1-\psi_m^{-1}(\tau L_\kappa)}{\tau L_\kappa}\right)^{\frac{1}{2}}\nabla_N\Delta_N\boldsymbol{v}\right\|^2+\kappa\tau\left\|\left(\frac{1-\psi_m^{-1}(\tau L_\kappa)}{\tau L_\kappa}\right)^{\frac{1}{2}}\Delta_N\boldsymbol{v}\right\|^2,\end{aligned}$$

where the last identity holds because all the eigenvalues of $(1-\psi_m^{-1}(\tau L_\kappa))/(\tau L_\kappa)$ are non-negative.

The inequality (4.16) can be derived in the same manner as follows:

$$\begin{aligned}&\langle G_m\psi_m(\tau L_\kappa)\boldsymbol{v},-\Delta_N\boldsymbol{v}\rangle+\|\nabla_N(\boldsymbol{w}-\boldsymbol{v})\|^2\\&=L_xL_y\sum_{k,l\in\hat{S}_h}\left(1-\psi_m^{-1}(\tau\Lambda_{k,l})\right)\psi_m(\tau\Lambda_{k,l})\lambda_{k,l}|\hat{v}_{k,l}|^2+L_xL_y\sum_{k,l\in\hat{S}_h}\lambda_{k,l}|\hat{w}_{k,l}-\hat{v}_{k,l}|^2\end{aligned}$$

$$
\begin{aligned}
&= L_x L_y \sum_{k,l\in\hat{S}_h} \big(1-\psi_m^{-1}(\tau\Lambda_{k,l})\big)\lambda_{k,l}\left[\psi_m(\tau\Lambda_{k,l})|\hat{v}_{k,l}|^2+\frac{1}{1-\psi_m^{-1}(\tau\Lambda_{k,l})}|\hat{w}_{k,l}-\hat{v}_{k,l}|^2\right]\\
&= L_x L_y \sum_{k,l\in\hat{S}_h} \big(1-\psi_m^{-1}(\tau\Lambda_{k,l})\big)\lambda_{k,l}\Bigg[|\hat{v}_{k,l}|^2+|\hat{w}_{k,l}-\hat{v}_{k,l}|^2\\
&\qquad\qquad \underbrace{+(\psi_m(\tau\Lambda_{k,l})-1)|\hat{v}_{k,l}|^2+\frac{1}{\psi_m(\tau\Lambda_{k,l})-1}|\hat{w}_{k,l}-\hat{v}_{k,l}|^2}_{\geq 2|\hat{v}_{k,l}||\hat{w}_{k,l}-\hat{v}_{k,l}|}\Bigg]\\
&\geq L_x L_y \sum_{k,l\in\hat{S}_h} \big(1-\psi_m^{-1}(\tau\Lambda_{k,l})\big)\lambda_{k,l}|\hat{w}_{k,l}|^2\\
&= \big\langle (1-\psi_m^{-1}(\tau L_\kappa))\boldsymbol{w}, -\Delta_N\boldsymbol{w}\big\rangle.
\end{aligned}
$$

This completes the proof. □

For notational convenience, we denote

$$
e^n = U^n-\boldsymbol{u}^n,\quad e_{n,i}=U_{n,i}-\boldsymbol{u}_{n,i},\quad \forall n\geq 0,\quad i=1,2. \tag{4.17}
$$

In turn, subtracting the numerical solution (2.8) from (4.10) leads to the error equation

$$
\begin{cases}
e_{n,1}=\psi_1^{-1}(\tau L_\kappa)\left(e^n+\dfrac{1}{3}\tau\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big)\right),\\
e_{n,2}=\psi_2^{-1}(\tau L_\kappa)\left(e^n+\dfrac{2}{3}\tau \mathrm{e}^{\frac{1}{3}\tau L_\kappa}\big(N_\kappa(U_{n,1})-N_\kappa(\boldsymbol{u}_{n,1})\big)\right),\\
e^{n+1}=\psi_3^{-1}(\tau L_\kappa)\Bigg(e^n+\dfrac{1}{4}\tau\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big)\\
\qquad\qquad +\dfrac{3}{4}\tau \mathrm{e}^{\frac{2}{3}\tau L_\kappa}\big(N_\kappa(U_{n,2})-N_\kappa(\boldsymbol{u}_{n,2})\big)\Bigg)+\tau R^n
\end{cases}
\tag{4.18}
$$

with $\|\nabla_N R^n\|\leq C(\tau^3+h^m)$.

Theorem 4.1. *Given initial data $u_0\in C_{\mathrm{per}}^{m_0+5}(\overline{\Omega})$, and suppose the unique solution for the NSS equation equipped with periodic boundary conditions is of regularity class $\mathcal{R}$. Let $\boldsymbol{u}^n$ denote the numerical solution computed by the ERK(3,3) scheme* (2.8)*. Then, if τ and h are sufficiently small, the following convergence estimate holds:*

$$
\big\|\nabla_N e^k\big\|+\left(\frac{1}{4}\varepsilon^2\tau\sum_{j=1}^{k}\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^j\big\|^2\right)^{\frac{1}{2}}\leq C^*(\tau^3+h^m),\quad \forall \tau k\leq T. \tag{4.19}
$$

Proof. For the first stage of the error equation (4.18), due to the current form not revealing a clear relationship between the linear and nonlinear parts, we denote $e_{n,1}^*=\psi_1^{-1}(\tau L_\kappa)e^n$

and rewrite the error equation as the following two-substage system:

$$e_{n,1}^*-e^n=\big(\psi_1^{-1}(\tau L_\kappa)-I\big)e^n, \tag{4.20}$$

$$e_{n,1}-e_{n,1}^*=\frac{1}{3}\tau\psi_1^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big). \tag{4.21}$$

Taking inner product of (4.20) with $-\Delta_N(e_{n,1}^*+e^n)$ indicates that

$$\big\langle e_{n,1}^*-e^n,-\Delta_N\big(e_{n,1}^*+e^n\big)\big\rangle+\big\langle G_1e^n,-\Delta_N\big(e_{n,1}^*+e^n\big)\big\rangle=0. \tag{4.22}$$

The first term can be analyzed with the help of the summation by parts formula

$$\big\langle e_{n,1}^*-e^n,-\Delta_N\big(e_{n,1}^*+e^n\big)\big\rangle=\big\|\nabla_Ne_{n,1}^*\big\|^2-\big\|\nabla_Ne^n\big\|^2. \tag{4.23}$$

For the second part, according to inequality (4.15), we have

$$\begin{aligned}&\big\langle G_1e^n,-\Delta_N\big(e_{n,1}^*+e^n\big)\big\rangle\\&=\big\langle G_1e^n,-\Delta_Ne_{n,1}^*\big\rangle+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_Ne^n\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_Ne^n\big\|^2.\end{aligned} \tag{4.24}$$

Therefore, substituting (4.23) and (4.24) into (4.22) implies that

$$\begin{aligned}&\big\|\nabla_Ne_{n,1}^*\big\|^2-\big\|\nabla_Ne^n\big\|^2+\big\langle G_1e^n,-\Delta_Ne_{n,1}^*\big\rangle\\&\quad+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_Ne^n\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_Ne^n\big\|^2=0.\end{aligned} \tag{4.25}$$

For the second substage (4.21), taking inner product with $-2\Delta_Ne_{n,1}$ gives

$$\big\langle e_{n,1}-e_{n,1}^*,-2\Delta_Ne_{n,1}\big\rangle=\frac{2}{3}\tau\big\langle\psi_1^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big),-\Delta_Ne_{n,1}\big\rangle. \tag{4.26}$$

Based on (4.16) and (4.25), the term on the left-hand side implies that

$$\begin{aligned}&\big\langle e_{n,1}-e_{n,1}^*,-2\Delta_Ne_{n,1}\big\rangle\\&=\|\nabla_Ne_{n,1}\|^2-\|\nabla_Ne_{n,1}^*\|^2+\|\nabla_N(e_{n,1}-e_{n.1}^*)\|^2\\&=\|\nabla_Ne_{n,1}\|^2-\big\|\nabla_Ne^n\big\|^2+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_Ne^n\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_Ne^n\big\|^2\\&\quad+\big\langle G_1e^n,-\Delta_Ne_{n,1}^*\big\rangle+\|\nabla_N(e_{n,1}-e_{n.1}^*)\|^2\\&\geq\|\nabla_Ne_{n,1}\|^2-\big\|\nabla_Ne^n\big\|^2+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_Ne^n\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_Ne^n\big\|^2\\&\quad+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_Ne_{n,1}\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_Ne_{n,1}\big\|^2.\end{aligned} \tag{4.27}$$

Moreover, for the nonlinear error inner product, applying the summation by parts formula and Lemma 4.3 indicates that

$$\frac{2}{3}\tau\big\langle\psi_1^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big),-\Delta_Ne_{n,1}\big\rangle$$

$$
\begin{aligned}
&= \frac{2}{3}\tau\big\langle \psi_1^{-1}(\tau L_\kappa)\nabla_N\cdot\big(\beta_\kappa(\nabla_N U^n)-\beta_\kappa(\nabla_N \boldsymbol{u}^n)\big), \Delta_N e_{n,1}\big\rangle \\
&= -\frac{2}{3}\tau\big\langle \psi_1^{-1}(\tau L_\kappa)\big(\beta_\kappa(\nabla_N U^n)-\beta_\kappa(\nabla_N \boldsymbol{u}^n)\big), \nabla_N\Delta_N e_{n,1}\big\rangle \\
&= -\frac{2}{3}\tau\left\langle \psi_1^{-\frac{1}{2}}(\tau L_\kappa)\left(\frac{\tau L_\kappa}{\psi_1(\tau L_\kappa)-1}\right)^{\frac{1}{2}}\big(\beta_\kappa(\nabla_N U^n)-\beta_\kappa(\nabla_N \boldsymbol{u}^n)\big), (G_1^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,1}\right\rangle \\
&\le \frac{2}{3}\sqrt{3}\tau(\kappa+1)\big\|\nabla_N e^n\big\|\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,1}\big\| \\
&\le \frac{2}{3}(\kappa+1)^2\varepsilon^{-2}\tau\big\|\nabla_N e^n\big\|^2+\frac{\varepsilon^2}{2}\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,1}\big\|^2. \qquad (4.28)
\end{aligned}
$$

Here, the second-to-last inequality utilizes the operator estimate (4.11) (in Proposition 4.2).

Combining (4.26) with (4.27), (4.28) and eliminating the term $\varepsilon^2\tau\|(G_1^*)^{1/2}\nabla_N\Delta_N e_{n,1}\|/2$ on the right hand leads to

$$
\begin{aligned}
&\|\nabla_N e_{n,1}\|^2-\big\|\nabla_N e^n\big\|^2+\varepsilon^2\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_N e^n\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_N e^n\big\|^2 \\
&\quad+\frac{\varepsilon^2}{2}\tau\big\|(G_1^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,1}\big\|^2+\kappa\tau\big\|(G_1^*)^{\frac{1}{2}}\Delta_N e_{n,1}\big\|^2 \\
&\le \frac{2}{3}(\kappa+1)^2\varepsilon^{-2}\tau\big\|\nabla_N e^n\big\|^2.
\end{aligned}
$$

As a direct result, we obtain a preliminary H^1 error estimate in the first ERK stage

$$
\|\nabla_N e_{n,1}\|^2 \le \left(1+\frac{2}{3}(\kappa+1)^2\varepsilon^{-2}\tau\right)\big\|\nabla_N e^n\big\|^2.
$$

By imposing the time step constraint $\tau \le (15/8)(\kappa+1)^{-2}\varepsilon^2$, we obtain

$$
\|\nabla_N e_{n,1}\| \le \left(1+\frac{2}{3}(\kappa+1)^2\varepsilon^{-2}\tau\right)^{\frac{1}{2}}\big\|\nabla_N e^n\big\| \le \frac{3}{2}\big\|\nabla_N e^n\big\|. \qquad (4.29)
$$

For the second stage of the error equation (4.18), we also decompose it into linear and nonlinear substages

$$
e_{n,2}^*-e^n = \big(\psi_2^{-1}(\tau L_\kappa)-I\big)e^n, \qquad (4.30)
$$

$$
e_{n,2}-e_{n,2}^* = \frac{2}{3}\tau\psi_2^{-1}(\tau L_\kappa)\mathrm{e}^{\frac{1}{3}\tau L_\kappa}\big(N_\kappa(U_{n,1})-N_\kappa(\boldsymbol{u}_{n,1})\big). \qquad (4.31)
$$

By employing the same procedure as in (4.25) and (4.26), we obtain

$$
\begin{aligned}
&\|\nabla_N e_{n,2}\|^2-\big\|\nabla_N e^n\big\|^2+\varepsilon^2\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e^n\big\|^2+\kappa\tau\big\|(G_2^*)^{\frac{1}{2}}\Delta_N e^n\big\|^2 \\
&\quad+\varepsilon^2\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\big\|^2+\kappa\tau\big\|(G_2^*)^{\frac{1}{2}}\Delta_N e_{n,2}\big\|^2 \\
&\le \frac{4}{3}\tau\big\langle \psi_2^{-1}(\tau L_\kappa)\mathrm{e}^{\frac{1}{3}\tau L_\kappa}\big(N_\kappa(U_{n,1})-N_\kappa(\boldsymbol{u}_{n,1})\big), -\Delta_N e_{n,2}\big\rangle. \qquad (4.32)
\end{aligned}
$$

For the nonlinear error inner product on the right-hand side, according to the operator estimate (4.12) and the preliminary error estimate (4.29), we obtain

$$\begin{aligned}
&\frac{4}{3}\tau\big\langle\psi_2^{-1}(\tau L_\kappa)\mathrm{e}^{\frac{1}{3}\tau L_\kappa}\big(N_\kappa(U_{n,1})-N_\kappa(\boldsymbol{u}_{n,1})\big),-\Delta_N e_{n,2}\big\rangle\\
=&-\frac{4}{3}\tau\big\langle\psi_2^{-1}(\tau L_\kappa)\mathrm{e}^{\frac{1}{3}\tau L_\kappa}\big(\beta_\kappa\big(\nabla_N u(t_{n,1})\big)-\beta_\kappa(\nabla_N u_{n,1})\big),\nabla_N\Delta_N e_{n,2}\big\rangle\\
=&-\frac{4}{3}\tau\left\langle\mathrm{e}^{\frac{1}{3}\tau L_\kappa}\psi_2^{-\frac{1}{2}}(\tau L_\kappa)\left(\frac{\tau L_\kappa}{\psi_2(\tau L_\kappa)-1}\right)^{\frac{1}{2}}\big(\beta_\kappa(\nabla_N U_{n,1})-\beta_\kappa(\nabla_N\boldsymbol{u}_{n,1})\big),(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\right\rangle\\
\le&\frac{2\sqrt{6}}{3}\tau(\kappa+1)\|\nabla_N e_{n,1}\|\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\big\|\\
\le&\frac{4}{3}(\kappa+1)^2\varepsilon^{-2}\tau\|\nabla_N e_{n,1}\|^2+\frac{\varepsilon^2}{2}\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\big\|^2\\
\le&3(\kappa+1)^2\varepsilon^{-2}\tau\big\|\nabla_N e^n\big\|^2+\frac{\varepsilon^2}{2}\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\big\|^2.
\end{aligned}$$

A substitution into (4.32) yields

$$\begin{aligned}
&\|\nabla_N e_{n,2}\|^2-\big\|\nabla_N e^n\big\|^2+\varepsilon^2\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e^n\big\|^2+\kappa\tau\big\|(G_2^*)^{\frac{1}{2}}\Delta_N e^n\big\|^2\\
&\quad+\frac{\varepsilon^2}{2}\tau\big\|(G_2^*)^{\frac{1}{2}}\nabla_N\Delta_N e_{n,2}\big\|^2+\kappa\tau\big\|(G_2^*)^{\frac{1}{2}}\Delta_N e_{n,2}\big\|^2\\
\le&3(\kappa+1)^2\varepsilon^{-2}\tau\big\|\nabla_N e^n\big\|^2.
\end{aligned}$$

Therefore, a preliminary H^1 error estimate has been derived in the second ERK stage,

$$\|\nabla_N e_{n,2}\|^2\le\big(1+3(\kappa+1)^2\varepsilon^{-2}\tau\big)\big\|\nabla_N e^n\big\|^2$$

such that

$$\|\nabla_N e_{n,2}\|\le\frac{2\sqrt{3}}{3}\big\|\nabla_N e^n\big\|,\tag{4.33}$$

provided that $\tau\le(1/9)(\kappa+1)^{-2}\varepsilon^2$.

The analysis in the third stage of the error equation (4.18) follows a similar idea. Again, we decompose the error equation into linear and nonlinear substages

$$e_{n,3}^*-e^n=\big(\psi_3^{-1}(\tau L_\kappa)-I\big)e_n\tag{4.34}$$

$$\begin{aligned}
e^{n+1}-e_{n,3}^*=&\frac{1}{4}\tau\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big)\\
&+\frac{3}{4}\tau\mathrm{e}^{\frac{2}{3}\tau L_\kappa}\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U_{n,2})-N_\kappa(\boldsymbol{u}_{n,2})\big).
\end{aligned}\tag{4.35}$$

For formulation (4.35), repeating the procedure in (4.25) and (4.26) yields

$$\big\|\nabla_N e^{n+1}\big\|^2-\big\|\nabla_N e^n\big\|^2+\varepsilon^2\tau\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^n\big\|^2+\kappa\tau\big\|(G_3^*)^{\frac{1}{2}}\Delta_N e^n\big\|^2$$

$$
\begin{aligned}
&+\varepsilon^2\tau\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|^2+\kappa\tau\big\|(G_3^*)^{\frac{1}{2}}\Delta_N e^{n+1}\big\|^2\\
\leq&\frac{1}{2}\tau\big\langle\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big),-\Delta_N e^{n+1}\big\rangle\\
&+\frac{3}{2}\tau\big\langle \mathrm{e}^{\frac{2}{3}\tau L_\kappa}\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U_{n,2})-N_\kappa(\boldsymbol{u}_{n,2})\big),-\Delta_N e^{n+1}\big\rangle\\
&+2\tau\big\langle R^n,-\Delta_N e^{n+1}\big\rangle.
\end{aligned}\tag{4.36}
$$

By combining the operator estimates (4.13) and (4.14) and the preliminary error estimate (4.33), the analysis for the two nonlinear error inner product terms can be similarly established

$$
\begin{aligned}
&\frac{1}{2}\tau\big\langle\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U^n)-N_\kappa(\boldsymbol{u}^n)\big),-\Delta_N e^{n+1}\big\rangle\\
=&-\frac{1}{2}\tau\left\langle\psi_3^{-\frac{1}{2}}(\tau L_\kappa)\left(\frac{\tau L_\kappa}{\psi_3(\tau L_\kappa)-1}\right)^{\frac{1}{2}}\big(\beta_\kappa(\nabla_N U^n)-\beta_\kappa(\nabla_N\boldsymbol{u}^n)\big),(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\right\rangle\\
\leq&\frac{1}{2}\tau(\kappa+1)\big\|\nabla_N e^n\big\|\big\|G_3^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|\\
\leq&\frac{1}{4}\varepsilon^{-2}(\kappa+1)^2\tau\big\|\nabla_N e^n\big\|^2+\frac{1}{4}\varepsilon^2\tau\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|^2,
\end{aligned}
$$

$$
\begin{aligned}
&\frac{3}{2}\tau\big\langle \mathrm{e}^{\frac{2}{3}\tau L_\kappa}\psi_3^{-1}(\tau L_\kappa)\big(N_\kappa(U_{n,2})-N_\kappa(\boldsymbol{u}_{n,2})\big),-\Delta_N e^{n+1}\big\rangle\\
=&-\frac{3}{2}\tau\left\langle \mathrm{e}^{\frac{2}{3}\tau L_\kappa}\psi_3^{-\frac{1}{2}}(\tau L_\kappa)\left(\frac{\tau L_\kappa}{\psi_3(\tau L_\kappa)-1}\right)^{\frac{1}{2}}\big(\beta_\kappa(\nabla_N U_{n,2})-\beta_\kappa(\nabla_N\boldsymbol{u}_{n,2})\big),(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\right\rangle\\
\leq&\frac{3}{2}(\kappa+1)\tau\|\nabla_N e_{n,2}\|\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|\\
\leq&\frac{9}{8}\varepsilon^{-2}(\kappa+1)^2\tau\|\nabla_N e_{n,2}\|^2+\frac{1}{2}\tau\varepsilon^2\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|^2\\
\leq&\frac{3}{2}\varepsilon^{-2}(\kappa+1)^2\tau\big\|\nabla_N e^n\big\|^2+\frac{1}{2}\tau\varepsilon^2\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|^2.
\end{aligned}
$$

The truncation error inner product term can be bound straightforwardly

$$
\tau\big\langle R^n,-2\Delta_N e^{n+1}\big\rangle\leq\tau\big\|\nabla_N R^n\big\|^2+\tau\big\|\nabla_N e^{n+1}\big\|^2.
$$

Hence, we derive the following estimate:

$$
\begin{aligned}
&\big\|\nabla_N e^{n+1}\big\|^2-\big\|\nabla_N e^n\big\|^2+\varepsilon^2\tau\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^n\big\|^2+\kappa\tau\big\|(G_3^*)^{\frac{1}{2}}\Delta_N e^n\big\|^2\\
&\quad+\frac{1}{4}\varepsilon^2\tau\big\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^{n+1}\big\|^2+\kappa\tau\big\|(G_3^*)^{\frac{1}{2}}\Delta_N e^{n+1}\big\|^2\\
\leq&\frac{7}{4}\varepsilon^{-2}(\kappa+1)^2\tau\big\|\nabla_N e^n\big\|^2+\tau\big\|\nabla_N R^n\big\|^2+\tau\big\|\nabla_N e^{n+1}\big\|^2.
\end{aligned}\tag{4.37}
$$

In turn, by using the discrete Grönwall inequality, we obtain the error estimate

$$\left\|\nabla_N e^{n+1}\right\|+\left(\frac{1}{4}\varepsilon^2\tau\sum_{j=1}^{n+1}\left\|(G_3^*)^{\frac{1}{2}}\nabla_N\Delta_N e^j\right\|^2\right)^{\frac{1}{2}}\leq C^*(\tau^3+h^m), \tag{4.38}$$

because $\|\nabla_N R^n\|\leq C(\tau^3+h^m)$. Thus, the proof is completed. □

5 Numerical experiments

This section presents a series of numerical experiments using the ERK schemes. Initially, we verify the convergence rates and subsequently conduct a long-time simulation for the NSS equation to confirm the energy dissipation characteristics of the proposed schemes. Furthermore, the long-time simulation also reveals other physical properties, such as the logarithmic decay for the energy, along with power laws governing the growth of surface roughness and mound width.

In our experiments, we use the fast Fourier transform (FFT) to ensure computational efficiency. It is noteworthy that the calculation of exponential functions $\mathrm{e}^{c_i\tau\Lambda_{k,l}}$ may exceed the effective range of the floating-point number calculations in a computer, especially considering that the eigenvalues $\Lambda_{k,l}$ of the operator L_κ are non-negative. To ensure that scheme (2.8) can be effectively implemented, assuming $c_i\leq 1, i=1,\ldots,s$, we multiplied both sides by $\mathrm{e}^{-\tau L_\kappa}$ in each stage of the scheme. This indicated that during the computation, we used its equivalent form

$$\boldsymbol{u}_{n,i}=\hat{\psi}_i^{-1}(\tau L_\kappa)\left(\mathrm{e}^{-\tau L_\kappa}\boldsymbol{u}^n+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{(c_j-1)\tau L_\kappa}N_\kappa(\boldsymbol{u}_{n,j})\right),\quad i=1,\ldots,s,$$

where

$$\hat{\psi}_i(\tau L_\kappa)=\mathrm{e}^{-\tau L_\kappa}+\tau\sum_{j=0}^{i-1}a_{i,j}\mathrm{e}^{(c_j-1)\tau L_\kappa}L_\kappa.$$

5.1 Accuracy tests

Example 5.1. In this example, consider the NSS equation on the domain $\Omega=(0,2\pi)^2$ with $\varepsilon^2=0.01$ and $T=0.05$. The initial condition is set as

$$u_0(x,y)=0.1\sin 2x\sin 3y \tag{5.1}$$

on the uniform mesh $N=N_x=N_y$, with $N=256$. To facilitate the evaluation of accuracy and efficiency, we regraded the result obtained by the ETDRK4 scheme [22] with $\kappa=0$ and $\tau=2^{-8}\delta/5$ as the reference solution. Subsequently, numerical simulations is carried out with $\kappa=1/16$ and various time step sizes $\tau=2^{-k}\delta$, $k=0,1,\ldots,8$, and the error with respect to the L^2 and H^1 errors is computed with respect to τ.

We conduct the numerical solution separately using the ERK and IFRK schemes, both based on the same underlying RK coefficients. For comparison, we also evaluated up-to-third order energy-stable ETDRK schemes [1,10] and the fourth-order ETDRK schemes [22].

The L^2 and H^1 errors are shown in Fig. 6. It is evident that all schemes achieve perfect first- to fourth-order convergence rates in the time direction. Notably, due to the preservation of equilibria, the ERK and ETDRK schemes present higher accuracy than the IFRK schemes. Additionally, although the ERK(3,3) schemes and ETDRK3 schemes demonstrate similar accuracy, the coefficients of the ERK(3,3) scheme are sparser than the ETDRK3 scheme. Thus, the ERK(3,3) scheme requires less CPU time than the IFRK(3,3) and ETDRK3 schemes to achieve the same level of accuracy as shown in the figure. Moreover, when comparing the ERK(5,4) and ETDRK4 schemes, the accuracy results are comparable. Although the ERK(5,4) scheme involves computing one additional stage compared to the ETDRK4 scheme, leading to slightly lower efficiency, it can theoretically guarantee energy stability, which the ETDRK4 cannot.

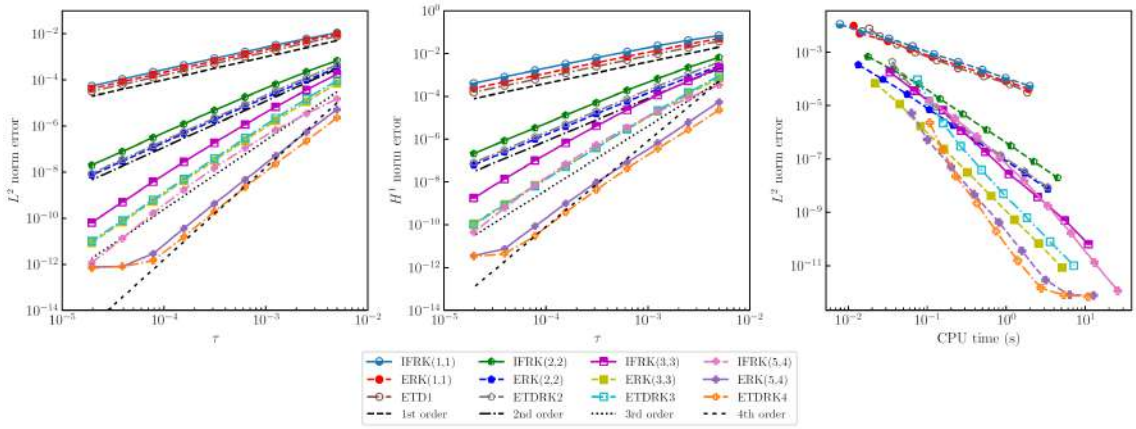


Figure 6: Discrete L^2 error (left) and H^1 error (right) computed by the ERK and ETDRK schemes at $T=0.05$ plotted versus the time step $\tau=2^{-k}\delta$ for $k=0,1,\ldots,8$ and $\delta=0.005$ with a spatial resolution of $N=256$.

5.2 Coarsening, energy dissipation, and other physical quantities

Example 5.2. For the NSS equation, we are not only interested in the property of energy dissipation, but also in the scaling laws [35, 42] that govern the energy, average surface roughness, and average slope, which have attracted significant interest from a physical perspective. In this example, we aim to present the numerical simulation using the ERK(3,3) scheme for the NSS equation to observe the long-term behavior of the solution. Concurrently, we want to demonstrate the unconditional energy stability of the ERK scheme in long-term simulations, while also emphasizing its ability to maintain other physical properties. The surface diffusion coefficient is set to be $\varepsilon=0.02$, the do-

main is taken to be $(0,L)^2$ with $L=12.8$, and the initial condition is set to be a random state by random numbers varying uniformly from -0.001 to 0.001. The uniform spatial resolution is given by $h=L/N$, where $N=512$, which is sufficient to capture the small structures in the solution for the given value of ε.

First, we consider the following important physical quantities in the NSS equation: (a) the energy $E(t)$, (b) the mass, (c) the average surface roughness

$$h(t)=|\Omega|^{-\frac{1}{2}}\|u(\cdot,t)-\bar{u}(t)\|,\quad \bar{u}(t)=|\Omega|^{-1}\int_\Omega u(\mathbf{x},t)\mathrm{d}\mathbf{x},$$

and (d) the average slope $m(t)=|\Omega|^{-1/2}\|\nabla u(\cdot,t)\|$. Here, the surface roughness and average slope were discretized as follows:

$$\begin{aligned}
h_N(t)&=\sqrt{\frac{1}{N_xN_y}\sum_{(i,j)\in\mathcal{S}_N}|u_{i,j}(t)-\bar{u}_N(t)|^2},\\
\bar{u}_N(t)&=\frac{1}{N_xN_y}\sum_{(i,j)\in\mathcal{S}_N}u_{i,j}(t),\\
m_N(t)&=\sqrt{\frac{1}{N^2}\sum_{(i,j)\in\mathcal{S}_N}\left|\left(\nabla_N u(t)\right)_{i,j}\right|^2}.
\end{aligned}$$

According to previous work [26], the no-slope-selection equation satisfies the following scaling laws:

$$E(t)\sim\mathcal{O}(-\ln t),\quad h(t)\sim\mathcal{O}\left(t^{\frac{1}{2}}\right),\quad m(t)\sim\mathcal{O}\left(t^{\frac{1}{4}}\right).$$

To facilitate the long-time simulation, we employed different time step sizes. Specifically, we set $\tau=0.004$ for the time interval $[0,400]$, $\tau=0.04$ for the time interval $[400,6000]$, $\tau=0.1$ for the time interval $[6000,100000]$, and $\tau=0.5$ for the time interval $[100000,300000]$. At each new time step, we initialized the numerical scheme by applying the final time output of the previous period as the initial data u^0. In Fig. 7, we display the snapshots of the height function u at different times, clearly illustrating the coarsening dynamics within the system. It can be seen in the earlier period that there were many hills (red) and valleys (blue) in the profile, while at the final time $t=300000$, the system reached saturation and a distinct one-hill-valley structure emerged, indicating that further coarsening was not possible.

The long-term characteristics of the solution, such as the rate of energy decay, the rate of average height growth, and the rate of mound width growth, offer valuable insights into the behavior of the system and are of great interest to the surface physics community. In Fig. 8, we present the semi-log plots of the energy data, the log-log plots of the average surface roughness data, and the log-log plot of the average mound width data, with the given physical parameter $\varepsilon=0.02$. The corresponding fitting lines were obtained by

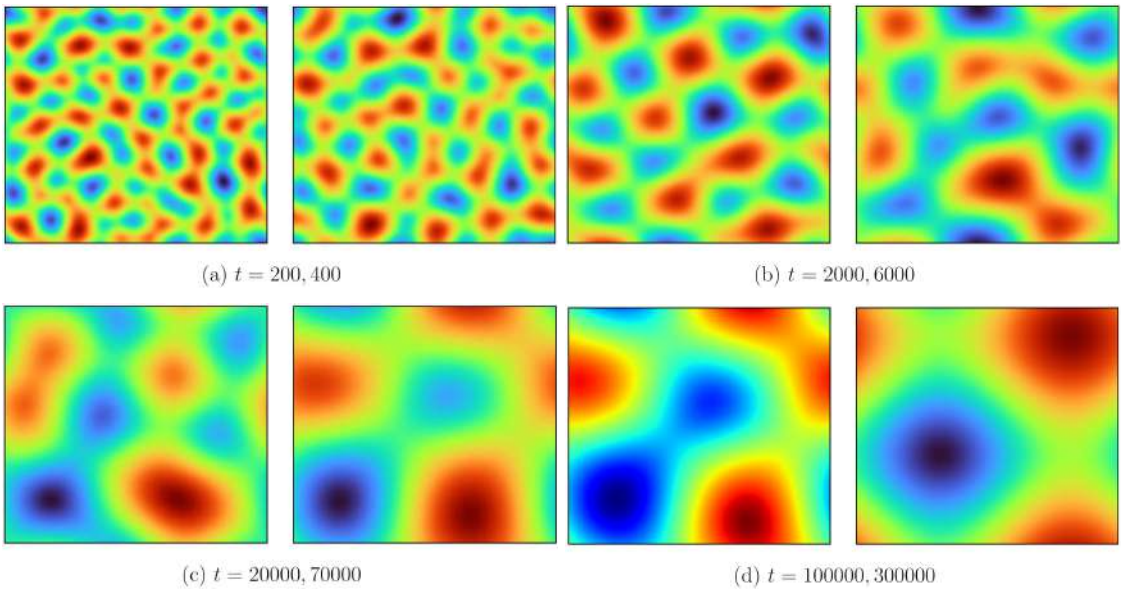


Figure 7: Snapshots of the numerical solution u at the indicated time with $\varepsilon = 0.02$.

sampling the computed data up to time $t = 400$ and performing the least squares fit. We utilized the following forms for the fitting line:

$$E(t) = a_e \ln t + b_e, \quad h(t) = a_h t^{b_h}, \quad m(t) = a_m t^{b_m}, \tag{5.2}$$

and the following parameters were obtained:

$$(a_e, b_e) = (-40.170, -151.974), \quad (a_h, b_h) = (0.411, 0.503), \quad (a_m, b_m) = (4.266, 0.251),$$

which clearly demonstrate scaling laws respectively. In addition, we present the evolution error of the mass in Fig. 8, which, despite the cumulative errors caused by long-term simulations, verified the mass conservation property of ERK schemes.

Finally, we recall the lower bound of the energy [2]:

$$E(t) \geq \frac{L^2}{2} \ln\left(\frac{4\varepsilon^2\pi^2}{L^2} - \frac{4\varepsilon^2\pi^2}{L^2} + 1\right) =: \gamma. \tag{5.3}$$

Since the energy is bounded below, it could not keep decreasing at the rate of $-\ln t$, and this fact can be found in the energy profile, where the decreasing rate begins to deviate significantly from the theoretical prediction. Regardless of the later time deviation from the accepted rates, the energy at which the system saturated is roughly consistent with the predicted minimum energy (5.3).

Example 5.3. In this example, we consider the phase field crystal (PFC) equation, which is a classical six-order in space model designed for studying the dynamics of atomic-scale crystal growth on a diffusive time scale. Introduced by Elder *et al.* [7], the PFC

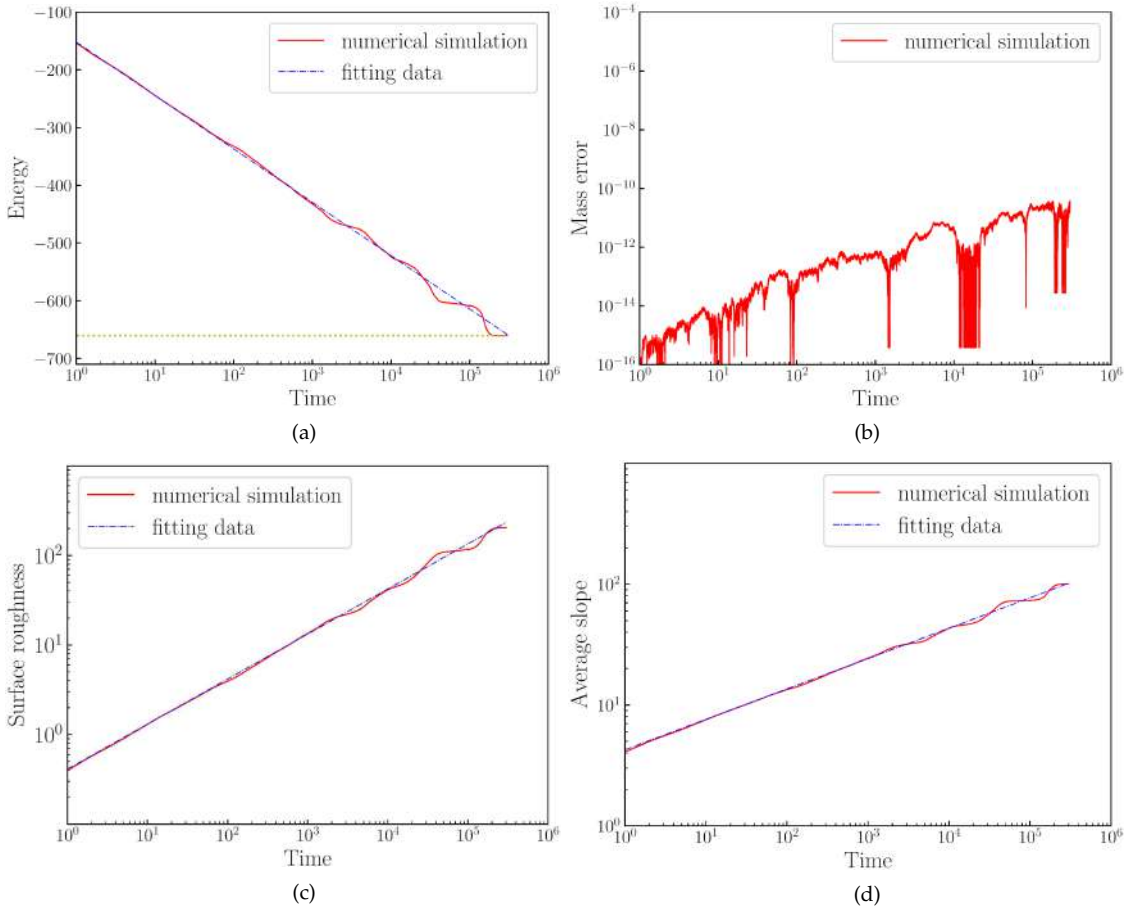


Figure 8: (a): Semi-log plot of the temporal evolution of the energy $E(t)$ for $\varepsilon=0.2$. The energy decreases like $-\ln t$ until saturation, and the dotted line denotes the minimum energy reached by the numerical simulation. (b): Log-log plot of the temporal evolution error of the mass. (c): Log-log plot of the average surface roughness u, denoted as $h(t)$, which grows like $t^{1/2}$. (d): Log-log plot of the average mound width, denoted as $m(t)$, which grows like $t^{1/4}$.

equation serves as an H^{-1} gradient flow model for the energy functional (1.1), with the linear operator $\mathcal{L}=(1+\Delta)^2$ and nonlinear potential $F(u)=u^4/4-\varepsilon u^2/2$, where $\varepsilon>0$ is a constant parameter. We intend to simulate the evolution of polycrystals by utilizing the PFC equation in a cubic domain. The specific procedures are detailed as follows. First, we construct three crystallites oriented with distinct directions by simulating the phase separation in a small domain $\Omega=(0,32)^3$, which is initiated with a random constant state $u_0(x,y)=0.285+0.1\times\text{rand}(x,y)$, and then we evolve it to a periodic body-centered-cubic (BCC) crystal lattice. Note that $\text{rand}(x,y)$ represents a uniformly distributed random variable within the range $(-1,1)$. Following this, the three BCC crystallites are superimposed into a uniform density field $u_0(x,y)=0.285$ within a larger domain $\Omega=(0,128)^3$ and simulate the growth and interactions of crystallites in a super-cooled liquid.

Initially, we simulate the phase separations using the ERK(5,4) scheme present in Example 3.5 with $\tau = 0.1$ from a random initial state in the domain $\Omega = (0,32)^3$, with $N = 64, \varepsilon = 0.25$, and terminal time $T = 1500$. Noting that the numerical solution consistently remains below 1, we set κ=2 for simulation. The snapshots of the crystallite growth are presented in Fig. 9, which clearly depicts the transformation of the system from a disordered state to a steady condition exhibiting a regular BCC pattern at $T = 1500$. This state is selected as the initial crystallite for the subsequent crystal growth simulation in a larger domain. Fig. 10 illustrates the dissipation of energy (top row) and conservation of mass (bottom row) as the dynamics evolve. The marked points on the energy curve correspond to the captured moments depicted in Fig. 9.

In addition, we simulate the interaction between the three crystallites within a larger domain $\Omega = (0,128)^3$ using the ERK(5, 4) scheme with $\tau = 1, N = 256$, and final time $T = 1000$. The three crystallites are injected into a uniform density field $u_0 = 0.285$ and positioned at $(32,96,96)$, $(96,96,96)$ and $(64,32,32)$, respectively. The evolution of crystal growth at time instants t=0,80,150, and 1000 is depicted in Fig. 11. The crystalline phases continue to expand and interact at approximately t=80, ultimately overspreading the entire domain and forming a prominent grain boundary. Fig. 10 demonstrates a monotone energy dissipation and the mass conservation to machine accuracy. We also note that due to the differing initial BCC patterns, the energy experiences a rapid decline in the initial stage. Consequently, this accelerates the process of reaching the equilibrium energy level in a large domain.

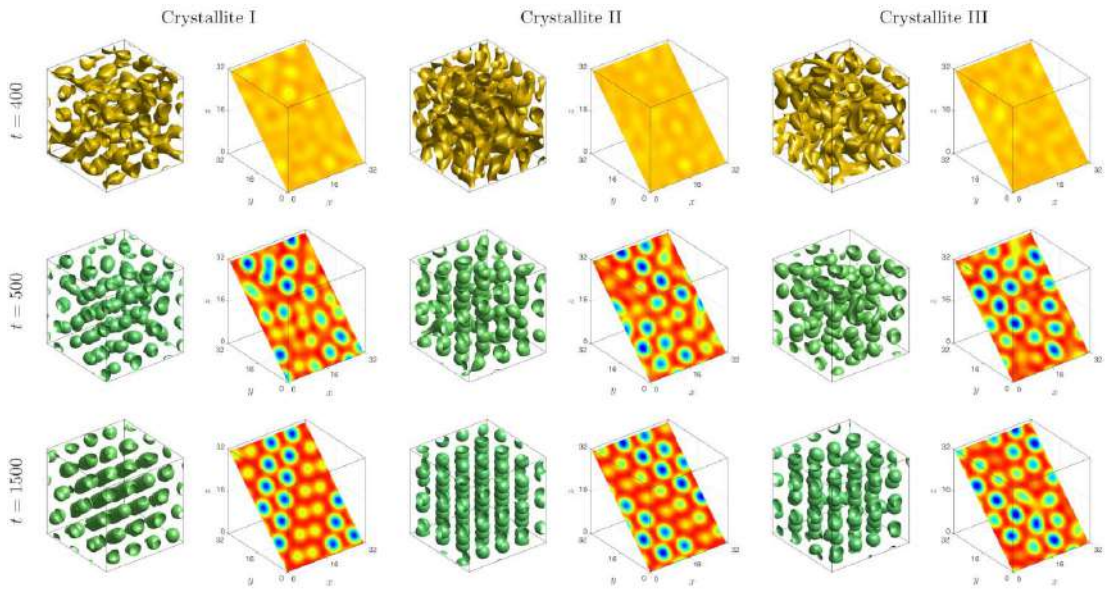


Figure 9: Snapshots of isosurfaces (first and third columns) and slice views (second and fourth columns) of 3D crystal growth, at $t = 400,500,1500$, computed using the ERK(5,4) scheme with random initial states. Parameters: $\varepsilon = 0.25, \kappa = 2$.

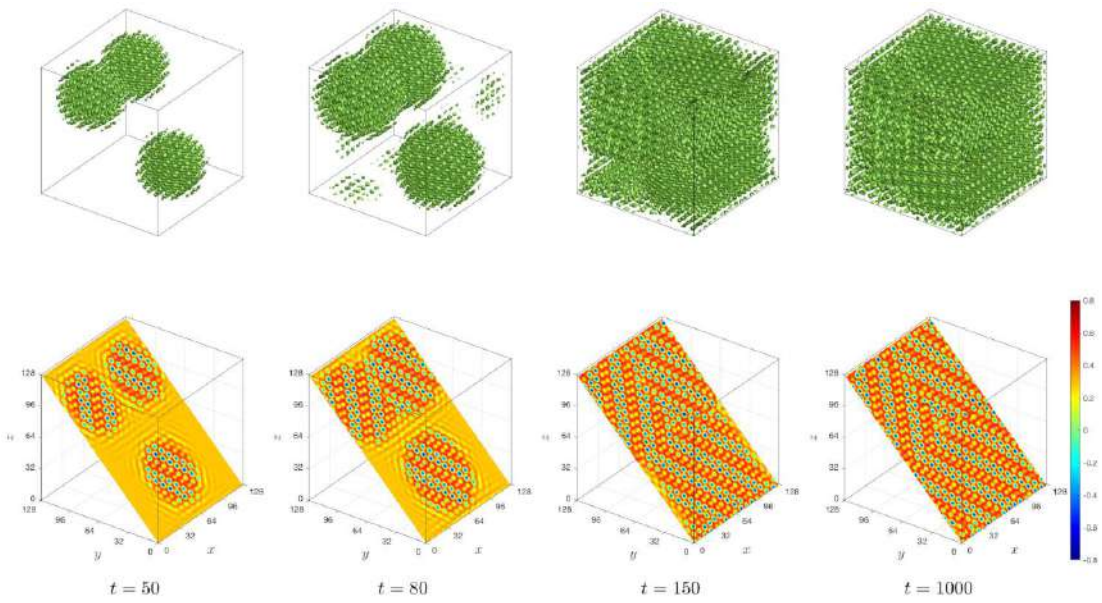


Figure 10: Evolution of the energy (top row) and relative mass error (bottom row) of the 3D crystal growth, computed by the ERK(5,4) scheme with random initial states in $\Omega = (0,32)^3$ (left and middle columns), and initial BCC patterns in $\Omega = (0,128)^3$ (right column). Parameters: $\varepsilon = 0.25, \kappa = 2$.

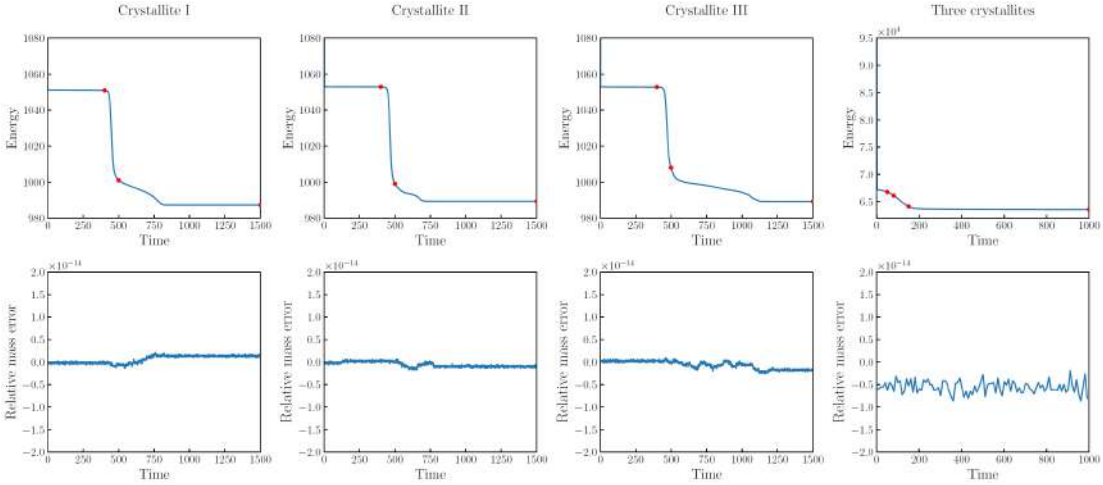


Figure 11: Snapshots of isosurfaces (top row) and slice views (bottom row) illustrating the 3D crystal growth at time instants $t = 0,80,150,1000$ computed by the ERK(5,4) scheme, starting with three initial BCC patterns. Parameters: $\varepsilon = 0.25, \kappa = 2$.

6 Conclusions

In this article, we have proposed and analyzed a novel class of up to fourth-order energy-stable ERK schemes for gradient flow equations, specifically targeting those nonlinear

terms that satisfy the global Lipschitz continuity. By examining the exponential decay/growth phenomenon within the IFRK framework, we devised novel approximations that combine exponential and linear functions to address this issue effectively. This led to the development of the ERK framework.

Furthermore, by utilizing the linear splitting (stabilization) technique, we presented a unified analytical framework to examine the energy stability for the ERK schemes underlying specific RK coefficients. Additionally, leveraging the rooted tree theory, we derived generic order conditions and determined the truncation error of ERK schemes with specific RK coefficients. By decomposing each stage of the ERK(3,3) scheme into a two-substage system and employing a meticulous Fourier eigenvalue estimate for these operators, we conducted an optimal rate error estimation for the NSS equation. Finally, we presented several numerical experiments to demonstrate the robustness and convergence of the proposed numerical scheme.

As a class of single-step methods, our approach holds a distinct advantage over high-order energy-stable multi-step methods. The multi-step methods have strict restrictions on the ratio of adjacent step sizes to maintain energy stability and the energy is a modified one. In contrast, our scheme does not restrict the time step size to preserve the original energy, making it more particularly well-suited for long-term simulations of gradient flow equations using adaptive time-stepping strategies. Additionally, the proposed ERK framework and the convergence analysis can be easily applied to general gradient flows. However, the global-in-time energy stability for models whose nonlinear terms do not satisfy the Lipschitz condition, such as the Allen-Cahn, Cahn-Hilliard, and phase field crystal equations, warrants further exploration. These technical details will be the focus of future work.

Acknowledgments

The authors would like to thank the anonymous referees for the valuable comments and constructive suggestions that have led to significant improvements in this work.

This work was supported by the National Key Research and Development Program of China (Grant No. 2020YFA0709803), by the National Natural Science Foundation of China (Grant Nos. 12271523, 12071481, 12371374), by the Science and Technology Innovation Program of Hunan Province (Grant Nos. 2022RC1192, 2021RC3082), by the Natural Science Foundation of Hunan (Grant No. 2021JJ20053), by the Defense Science Foundation of China (Grant Nos. 2021-JCJQ-JJ-0538, 2022-JCJQ-JJ-0879), and by the National University of Defense Technology (Grant No. 2023-LXY-FHJJ-002).

References

[1] W. Cao, H. Yang, and W. Chen, *An exponential time differencing Runge-Kutta method ETDRK32 for phase field models*, J. Sci. Comput., 99:6, 2024.

[2] W. Chen, S. Conde, C. Wang, X. Wang, and S. M. Wise, *A linear energy stable scheme for a thin film model without slope selection*, J. Sci. Comput., 52:546–562, 2012.

[3] W. Chen, X. Wang, Y. Yan, and Z. Zhang, *A second order BDF numerical scheme with variable steps for the Cahn-Hilliard equation*, SIAM J. Numer. Anal., 57:495–525, 2019.
[4] K. Cheng, Z. Qiao, and C. Wang, *A third order exponential time differencing numerical scheme for no-slope-selection epitaxial thin film model with energy stability*, J. Sci. Comput., 81:154–185, 2019.
[5] J. Du and Y. Yang, *Third-order conservative sign-preserving and steady-state-preserving time integrations and applications in stiff multispecies and multireaction detonations*, J. Comput. Phys., 395:489–510, 2019.
[6] Q. Du, L. Ju, X. Li, and Z. Qiao, *Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes*, SIAM Rev., 63:317–359, 2021.
[7] K. Elder, M. Katakowski, M. Haataja, and M. Grant, *Modeling elasticity in crystal growth*, Phys. Rev. Lett., 88:245701, 2002.
[8] D. J. Eyre, *Unconditionally gradient stable time marching the Cahn-Hilliard equation*, Mater. Res. Soc. Symp. Proc., 529:39–46, 1998
[9] Z. Fu, T. Tang, and J. Yang, *Energy diminishing implicit-explicit Runge-Kutta methods for gradient flows*, Math. Comp., 94:2745–2767, 2024.
[10] Z. Fu and J. Yang, *Energy-decreasing exponential time differencing Runge-Kutta methods for phase-field models*, J. Comput. Phys., 454:110943, 2022.
[11] K. Glasner and S. Orizaga, *Improving the accuracy of convexity splitting methods for gradient flow equations*, J. Comput. Phys., 315:52–64, 2016.
[12] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, in: Springer Series in Computational Mathematics, Vol. 8, Springer-Verlag, 1993.
[13] Y. Hao, Q. Huang, and C. Wang, *A third order BDF energy stable linear scheme for the no-slope-selection thin film model*, Commun. Comput. Phys., 29:905–929, 2021.
[14] M. Hochbruck, J. Leibold, and A. Ostermann, *On the convergence of Lawson methods for semilinear stiff problems*, Numer. Math., 145:553–580, 2020.
[15] D. Hou, L. Ju, and Z. Qiao, *A linear second-order maximum bound principle-preserving BDF scheme for the Allen-Cahn equation with a general mobility*, Math. Comp., 92:2515–2542, 2023.
[16] T. Y. Hou, J. S. Lowengrub, and M. J. Shelley, *Removing the stiffness from interfacial flows with surface tension*, J. Comput. Phys., 114:312–338, 1994.
[17] J. Huang and C.-W. Shu, *Bound-preserving modified exponential Runge-Kutta discontinuous Galerkin methods for scalar hyperbolic equations with stiff source terms*, J. Comput. Phys., 361:111–135, 2018.
[18] L. Isherwood, Z. J. Grant, and S. Gottlieb, *Strong stability preserving integrating factor Runge-Kutta methods*, SIAM J. Numer. Anal., 56:3276–3307, 2018.
[19] L. Ju, X. Li, and Z. Qiao, *Stabilized exponential-SAV schemes preserving energy dissipation law and maximum bound principle for the Allen-Cahn type equations*, J. Sci. Comput., 92:66, 2022.
[20] L. Ju, X. Li, Z. Qiao, and J. Yang, *Maximum bound principle preserving integrating factor Runge-Kutta methods for semilinear parabolic equations*, J. Comput. Phys., 439:110405, 2021.
[21] L. Ju, X. Li, Z. Qiao, and H. Zhang, *Energy stability and error estimates of exponential time differencing schemes for the epitaxial growth model without slope selection*, Math. Comp., 87:1859–1885, 2018.
[22] A.-K. Kassam and L. N. Trefethen, *Fourth-order time-stepping for stiff PDEs*, SIAM J. Sci. Comput., 26:1214–1233, 2005.
[23] S. Krogstad, *Generalized integrating factor methods for stiff PDEs*, J. Comput. Phys., 203:72–88, 2005.
[24] J. D. Lawson, *Generalized Runge-Kutta processes for stable systems with large Lipschitz constants*,

SIAM J. Numer. Anal., 4:372–380, 1967.
[25] P. H. Leo, J. S. Lowengrub, and H.-J. Jou, *A diffuse interface model for microstructural evolution in elastically stressed solids*, Acta Mater., 46:2113–2130, 1998.
[26] B. Li and J.-G. Liu, *Epitaxial growth without slope selection: Energetics, coarsening, and dynamic scaling*, J. Nonlinear Sci., 14:429–451, 2004.
[27] D. Li and Z. Qiao, *On second order semi-implicit Fourier spectral methods for 2D Cahn-Hilliard equations*, J. Sci. Comput., 70:301–341, 2017.
[28] D. Li and Z. Qiao, *On the stabilization size of semi-implicit Fourier-spectral methods for 3D Cahn-Hilliard equations*, Commun. Math. Sci., 15:1489–1506, 2017.
[29] D. Li, Z. Qiao, and T. Tang, *Characterizing the stabilization size for semi-implicit Fourier-spectral method to phase field equations*, SIAM J. Numer. Anal., 54:1653–1681, 2016.
[30] X. Li, Z. Qiao, C. Wang, and N. Zheng, *Global-in-time energy stability analysis for the exponential time differencing Runge-Kutta scheme for the phase field crystal equation*, arXiv:2406.06272, 2024.
[31] H.-l. Liao and Y. Kang, L^2 *norm error estimates of BDF methods up to fifth-order for the phase field crystal model*, IMA J. Numer. Anal., 44(4):2138–2164, 2024.
[32] H.-l. Liao, Y. Kang, and W. Han, *Discrete gradient structures of BDF methods up to fifth-order for the phase field crystal model*, arXiv:2201.00609, 2022.
[33] S. Maset and M. Zennaro, *Unconditional stability of explicit exponential Runge-Kutta methods for semi-linear ordinary differential equations*, Math. Comp., 78:957–967, 2009.
[34] S. Pei, Y. Hou, and B. You, *A linearly second-order energy stable scheme for the phase field crystal model*, Appl. Numer. Math., 140:134–164, 2019.
[35] L. Ratke and P. W. Voorhees, *Growth and Coarsening: Ostwald Ripening in Material Processing*, Springer, 2002.
[36] J. Shen, C. Wang, X. Wang, and S. M. Wise, *Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: Application to thin film epitaxy*, SIAM J. Numer. Anal., 50:105–125, 2012.
[37] J. Shen, J. Xu, and J. Yang, *The scalar auxiliary variable (SAV) approach for gradient flows*, J. Comput. Phys., 353:407–416, 2018.
[38] J. Shen, J. Xu, and J. Yang, *A new class of efficient and robust energy stable schemes for gradient flows*, SIAM Rev., 61:474–506, 2019.
[39] J. Shen and X. Yang, *Energy stable schemes for Cahn-Hilliard phase-field model of two-phase incompressible flows*, Chinese Ann. Math. Ser. B, 31:743–758, 2010.
[40] J. Shen and X. Yang, *Numerical approximations of Allen-Cahn and Cahn-Hilliard equations*, Discrete Contin. Dyn. Syst., 28:1669–1691, 2010.
[41] J. Sun, H. Zhang, X. Qian, and S. Song, *A family of structure-preserving exponential time differencing Runge-Kutta schemes for the viscous Cahn-Hilliard equation*, J. Comput. Phys., 492:112414, 2023.
[42] J. Y. Tsao, *Materials Fundamentals of Molecular Beam Epitaxy*, Academic Press, 2012.
[43] P. Vignal, L. Dalcin, D. L. Brown, N. Collier, and V. M. Calo, *An energy-stable convex splitting for the phase-field crystal equation*, Comput. Struct., 158:355–368, 2015.
[44] S. M. Wise, C. Wang, and J. S. Lowengrub, *An energy-stable and convergent finite-difference scheme for the phase field crystal equation*, SIAM J. Numer. Anal., 47:2269–2288, 2009.
[45] X. Yang, *Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends*, J. Comput. Phys., 327:294–316, 2016.
[46] X. Yang, J. Zhao, and Q. Wang, *Numerical approximations for the molecular beam epitaxial growth model based on the invariant energy quadratization method*, J. Comput. Phys., 333:104–127, 2017.
[47] H. Zhang, L. Liu, X. Qian, and S. Song, *Large time-stepping, delay-free, and invariant-set-*

preserving integrators for the viscous Cahn-Hilliard-Oono equation, J. Comput. Phys., 499:112708, 2024.

[48] H. Zhang, X. Qian, and S. Song, *Third-order accurate, large time-stepping and maximum-principle-preserving schemes for the Allen-Cahn equation*, Numer. Algorithms, 95:1213–1250, 2024.

[49] H. Zhang, X. Qian, J. Xia, and S. Song, *Efficient inequality-preserving integrators for differential equations satisfying forward Euler conditions*, ESAIM Math. Model. Numer. Anal., 57:1619–1655, 2023.

[50] H. Zhang, X. Qian, J. Xia, and S. Song, *Unconditionally maximum-principle-preserving parametric integrating factor two-step Runge-Kutta schemes for parabolic sine-gordon equations*, CSIAM Trans. Appl. Math., 4:177–224, 2023.

[51] H. Zhang, H. Wang, and X. Teng, *A second-order, global-in-time energy stable implicit-explicit Runge-Kutta scheme for the phase field crystal equation*, SIAM J. Numer. Anal., 62(6):2667–2697, 2024.

[52] H. Zhang, J. Yan, X. Qian, X. Chen, and S. Song, *Explicit third-order unconditionally structure-preserving schemes for conservative Allen-Cahn equations*, J. Sci. Comput., 90:1–29, 2022.

[53] H. Zhang, J. Yan, X. Qian, and S. Song, *Up to fourth-order unconditionally structure-preserving parametric single-step methods for semilinear parabolic equations*, Comput. Methods Appl. Mech. Engrg., 393:114817, 2022.

[54] H. Zhang, G. Zhang, Z. Liu, X. Qian, and S. Song, *On the maximum principle and high-order, delay-free integrators for the viscous Cahn-Hilliard equation*, Adv. Comput. Math., 50:1–46, 2024.