Journal of Computational and Applied Mathematics 367 (2020) 112455

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

An energy stable Runge–Kutta method for convex gradient problems

Jaemin Shin [a], June-Yub Lee [b,*]

[a] Institute of Mathematical Sciences, Ewha Womans University, Seoul 03760, Republic of Korea
[b] Department of Mathematics, Ewha Womans University, Seoul 03760, Republic of Korea

ARTICLE INFO

Article history:
Received 6 November 2018
Received in revised form 19 August 2019

Keywords:
Gradient flow
Convex problem
Stiffly accurate Runge–Kutta method
Positive definite condition
Unconditional energy stability
Unique solvability

ABSTRACT

We propose a class of Runge–Kutta methods which provide a simple unified framework to solve the gradient flow of a convex functional in an unconditionally energy stable manner. Stiffly accurate Runge–Kutta methods are high order accurate in terms of time and also assure the energy stability for any time step size when they satisfy the positive definite condition. We provide a detailed proof of the unconditional energy stability as well as unique solvability of the proposed scheme. We demonstrate the accuracy and stability of the proposed methods using numerical experiments for a specific example.

1. Introduction

The use of the gradient system has been fundamental in the development of many important concepts in dynamical systems [1] and it arises in a variety of applications, for example in modeling mesoscale morphological and microstructural evolution in materials [2–4]. The gradient flow can be represented as

$$\frac{\partial \Phi}{\partial t} = -\text{grad}\, F(\Phi), \tag{1}$$

where $F : \mathbb{R}^M \to \mathbb{R}$ is an energy functional, $\text{grad}\, F(\Phi)$ is a gradient of F, $\Phi(t) \in \mathbb{R}^M$, and M is the dimension of data space. Furthermore, the energy functional $F(\Phi)$ is non-increasing in time since (1) is of gradient type:

$$\frac{d}{dt}F(\Phi) = \left\langle \text{grad}\, F(\Phi), \frac{\partial \Phi}{\partial t} \right\rangle = -\|\text{grad}\, F(\Phi)\|^2 \leq 0, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^M$ with the corresponding norm $\|\Phi\|^2 = \langle \Phi, \Phi \rangle$. Here, the gradient operator "grad" depends on the given inner product space and (2) is called the energy dissipation property.

A convex (or contractive) problem is a subclass of the gradient flow problems, where $F(\Phi)$ is a convex functional as in the case of a linear diffusion equation, a porous medium equation [5], a nonlinear Schrödinger equation [6], and so on. Compared to the gradient flow (1) with a nonconvex functional which has multiple isolated equilibria, the development of a numerical method for the convex problem could be relatively easy since the equilibria form a convex set. However, finding an unconditionally energy stable high order method for the convex gradient flow continues to be an important question.

* Corresponding author.
E-mail address: jyllee@ewha.ac.kr (J.-Y. Lee).

https://doi.org/10.1016/j.cam.2019.112455
0377-0427/

Denoting Φ^n as an approximation of $\Phi(t^n)$, we aim to develop a numerical method for one-step evolution to obtain Φ^{n+1} at $t^n + \Delta t$. A numerical integration method for (1) is said to be gradient stable if a functional $F_{\Delta t} : \mathbb{R}^M \to \mathbb{R}$ with a critical time step size Δt_c exists such that:

$$\begin{aligned} &\text{(i) } F_{\Delta t} \text{ is bounded below,} \\ &\text{(ii) } F_{\Delta t}(\Phi) \to \infty \text{ as } \|\Phi\| \to \infty, \\ &\text{(iii) } F_{\Delta t}\left(\Phi^{n+1}\right) \le F_{\Delta t}\left(\Phi^n\right) \text{ for } \Phi^n \in \mathbb{R}^M, \end{aligned} \tag{3}$$

for all $0 < \Delta t \le \Delta t_c$. Such a definition is formulated in the work of [7]. In particular, these conditions indicate (i) boundness, (ii) coercivity, and (iii) energy decreasing for the discrete energy functional $F_{\Delta t}$. The concept of energy stability (iii) is an important part of the gradient stability (i)–(iii) for the gradient system [1,7]. Note that $F_{\Delta t}$ may or may not be identical to the original energy functional F. We are particularly interested in unconditional energy stability, which means that a method is energy stable no matter what time step $\Delta t > 0$ is taken.

High order time accuracy and unconditional energy stability are the two most important concepts to consider when developing a time-marching method for convex gradient flow, which has been a serious research objective in the field. A typical choice for the second order time-marching scheme for a parabolic equation might be the Crank–Nicolson method; however, it does not always guarantee the energy stability, as shown in Section 2.2. A few interesting alternative second-order energy stable methods have been proposed, like the discrete variational derivative method [8] and the discrete gradient method [9].

The Runge–Kutta (RK) methods are the most popular class of time-marching methods for improving time accuracy [7,10–15]. There have been related works [7,11] to achieve stability using the implicit RK method for the gradient flow, which provides sufficient conditions based on the *algebraic stability* for stiff equations. We propose a new criterion, namely *positive definite condition*, to ensure a stiffly accurate RK method to be energy stable for general convex gradient problems and will present examples of new RK methods not satisfying the algebraic stability. In this paper, we focus on two sufficient conditions to develop different numerical schemes for guaranteeing the energy stability of the convex problems.

In Section 2, we briefly review the Runge–Kutta methods and demonstrate that achieving unconditional energy stability for the convex gradient flow is not a trivial task. In Section 3, we prove that a stiffly accurate RK method with the positive definite condition ensures the unconditional energy stability. The unique solvability of the implicit equations for the scheme is proven in Section 4. In Section 5, we display a few methods up to the fourth-order accuracy, which guarantee the unconditional energy stability and unique solvability. Computational examples are also given in order to numerically demonstrate the accuracy and stability. Finally, conclusions are drawn in Section 6.

2. Preliminaries concerning Runge–Kutta methods

A continuous version of the gradient flow is

$$\frac{\partial \phi}{\partial t} = -\text{grad}\, \mathcal{F}(\phi), \tag{4}$$

where $\mathcal{F}$ is an energy functional and $\phi(\mathbf{x}, t)$ is a real function defined on a domain $\Omega \subset \mathbb{R}^d (d = 1, 2, 3)$. For a numerical approach based on the method of lines, $\phi(\mathbf{x}, t)$ will be discretized by $\Phi(t) \in \mathbb{R}^M$ whatever spatial discretizations are used. Although the gradient system (1) can be considered in infinite dimensional abstract vector space, we work with a finite dimensional system of ordinary differential equations as in (1).

2.1. The first order method

In order to numerically solve the gradient flow of the convex F with first-order accuracy, we simply consider the implicit Euler method

$$\frac{\Phi^{n+1} - \Phi^n}{\Delta t} = -\text{grad}\, F\left(\Phi^{n+1}\right). \tag{5}$$

It can be easily shown that the energy of the numerical solution of (5) monotonically decreases regardless of the time step size $\Delta t > 0$, by using the following lemma.

Lemma 1. *If $F(\Phi)$ is convex and the gradient of $F(\Phi)$ is well defined, then*

$$F(\Phi) - F(\Psi) \le \langle \text{grad} F(\Phi), \Phi - \Psi \rangle. \tag{6}$$

As a consequence of Lemma 1, we have

$$\begin{aligned} F\left(\Phi^{n+1}\right) - F\left(\Phi^n\right) &\le \left\langle \text{grad}\, F\left(\Phi^{n+1}\right), \Phi^{n+1} - \Phi^n \right\rangle \\ &= -\Delta t \left\| \text{grad}\, F\left(\Phi^{n+1}\right) \right\|^2 \le 0, \end{aligned} \tag{7}$$

and consequently, the implicit Euler method (5) is unconditionally energy stable.

One of the well-known convex problems is the heat equation (or the linear diffusion equation)

$$\frac{\partial \phi}{\partial t} = \Delta \phi, \tag{8}$$

where ϕ is the heat distribution on a domain Ω. In the aspect of the variational derivative, (8) can be described by the gradient flow for the energy functional

$$\mathcal{F}(\phi) = \int_{\Omega} \frac{1}{2} |\nabla \phi|^2 \, d\mathbf{x} \tag{9}$$

in the L_2 inner product space. Additionally, (8) can be described by the gradient flow for the energy functional

$$\mathcal{F}(\phi) = \int_{\Omega} \frac{1}{2} \phi^2 d\mathbf{x} \tag{10}$$

in the H^{-1} inner product space. Now, employing a semi-discrete representation by an approximation ϕ^n of $\phi(t^n)$, we have the first-order method

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = \Delta \phi^{n+1} \tag{11}$$

by the implicit Euler scheme (5) from the energy functionals (9) and (10) in the corresponding inner product spaces. Finally, the unconditional energy stability for (11) is surely guaranteed.

2.2. Crank–Nicolson method

A typical choice of numerical scheme for the second-order accuracy in time is the Crank–Nicolson method

$$\frac{\Phi^{n+1} - \Phi^n}{\Delta t} = -\frac{\operatorname{grad} F\left(\Phi^{n+1}\right) + \operatorname{grad} F\left(\Phi^n\right)}{2}. \tag{12}$$

With a semi-discrete form, the Crank–Nicolson method (12) for the heat equation (8) can be presented as

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = \Delta \left(\frac{\phi^{n+1} + \phi^n}{2} \right), \tag{13}$$

and it is easy to show that (13) is unconditionally energy stable, i.e., $\mathcal{F}\left(\phi^{n+1}\right) \le \mathcal{F}\left(\phi^n\right)$ for the energy functionals (9) or (10), regardless of the time step size $\Delta t > 0$. For example, with (10),

$$\begin{aligned} &\mathcal{F}\left(\phi^{n+1}\right) - \mathcal{F}\left(\phi^n\right) = \frac{1}{2} \left\langle \phi^{n+1} + \phi^n, \phi^{n+1} - \phi^n \right\rangle \\ &= \frac{\Delta t}{4} \left\langle \phi^{n+1} + \phi^n, \Delta\left(\phi^{n+1} + \phi^n\right) \right\rangle = -\frac{\Delta t}{4} \left\| \nabla\left(\phi^{n+1} + \phi^n\right) \right\|^2 \le 0, \end{aligned} \tag{14}$$

and the stability for (9) can be shown in a similar manner. However, this does not mean that applying the Crank–Nicolson method (12) to any convex problem guarantees the energy stability. For example, Fig. 1 shows the temporal accuracy and energy evolution for the numerical test in Section 5. The second-order time accuracy is numerically demonstrated, but the energy is fluctuating for larger time steps $\Delta t \ge 2$, which implies that the Crank–Nicolson method (12) does not always guarantee the energy stability.

2.3. Runge–Kutta methods

The Runge–Kutta (RK) method is very popular for obtaining a high order accuracy. For more information, we refer to [16,17]. In this section, we briefly review an s-stage RK method. The time-marching algorithm for the system of ordinary differential equations

$$\frac{\partial \Phi}{\partial t} = f(\Phi) \text{ and } \Phi(0) = \Phi^0 \tag{15}$$

can be written as follows with the Butcher table

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array} = \begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}, \tag{16}$$

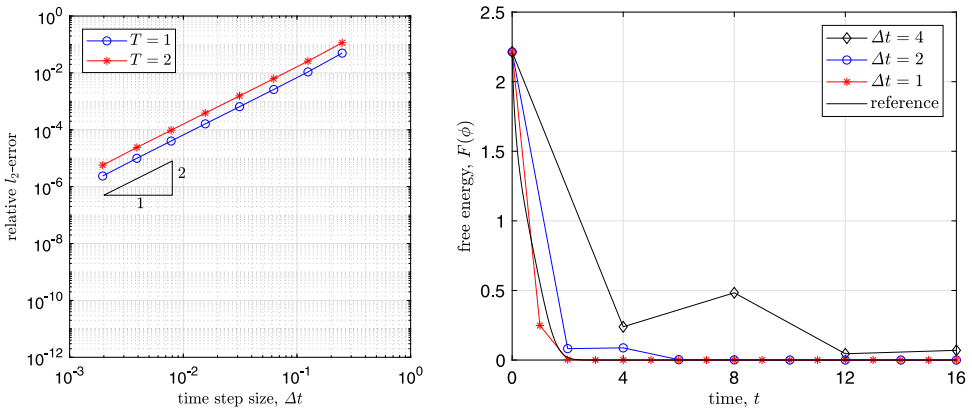

Fig. 1. Relative l_2-errors and energy evolutions of the example in Section 5 for various time steps Δt.

Table 1
Order conditions of RK methods up to the fourth-order accuracy.

Order	1	2	3	4
Condition	$\mathbf{b}\cdot\mathbf{1}=1$	$\mathbf{b}\cdot\mathbf{c}=1/2$	$\mathbf{b}\cdot\mathbf{c}^2=1/3$ $\mathbf{b}\cdot\mathbf{Ac}=1/6$	$\mathbf{b}\cdot\mathbf{c}^3=1/4$ $\mathbf{b}\cdot(\mathbf{c}\odot\mathbf{Ac})=1/8$ $\mathbf{b}\cdot\mathbf{Ac}^2=1/12$ $\mathbf{b}\cdot\mathbf{A}^2\mathbf{c}=1/24$

Here, we define the component-wise product as $\mathbf{x}\odot\mathbf{y}=(x_1y_1,x_2y_2,\ldots,x_sy_s)^T$ for $\mathbf{x}=(x_1,x_2,\ldots,x_s)^T$ and $\mathbf{y}=(y_1,y_2,\ldots,y_s)^T$, and $\mathbf{x}^m=\mathbf{x}\odot\mathbf{x}^{m-1}$ for $m>1$.

where $\mathbf{A}\in\mathbb{R}^{s\times s}$, $\mathbf{b}\in\mathbb{R}^s$, and the coefficients $\mathbf{c}$ are usually given by $\mathbf{c}=\mathbf{A1}$, with $\mathbf{1}=(1,1,\ldots,1)^T\in\mathbb{R}^s$. We denote Φ^n as an approximation of $\Phi(t^n)$. With a time step size $\Delta t>0$, the one-step evolution from t^n to $t^{n+1}=t^n+\Delta t$ of the RK method is given as follows. Given Φ^n, we calculate the next time approximation Φ^{n+1} with the intermediate stages, $\Phi_1,\Phi_2,\ldots,\Phi_s$. We set a zeroth-stage as $\Phi_0=\Phi^n$. For each stage $i=1,2,\ldots,s$, we calculate

$$\Phi_i=\Phi_0+\Delta t\sum_{j=1}^{s}a_{ij}f\left(\Phi_j\right). \tag{17}$$

Here, Φ_i is the approximation of $\Phi(t^n+c_i\Delta t)$ on the ith stage. In addition, we evaluate the next time approximation Φ^{n+1} as

$$\Phi^{n+1}=\Phi_0+\Delta t\sum_{j=1}^{s}b_jf\left(\Phi_j\right). \tag{18}$$

Table 1 shows the order conditions of the RK method up to fourth-order accuracy. These order conditions can be simply understood by considering the Taylor expansion shown in [14], particularly for relatively lower orders of accuracy. The same order conditions can be derived using the tree structure as in [16].

Remark 1. The implicit Euler method (5) and the Crank–Nicolson method (12) are special cases of RK methods, thus the result in Fig. 1 demonstrates that RK methods do not always guarantee energy stability.

3. Unconditionally energy stable Runge–Kutta methods

We start by considering two conditions which play key roles in the development of energy stable methods. First, we consider the *stiffly accurate condition* [17].

Condition SA (*Stiffly Accurate Condition*)**.** *The RK method (17) is said to be satisfying the stiffly accurate condition if* $\mathbf{A}$ *and* $\mathbf{b}$ *satisfy that* $b_j=a_{sj}$ *for all* j.

We now propose the s-stage RK methods with the stiffly accurate condition for the gradient flow (1), and we called it the stiffly accurate Runge–Kutta (SARK) method. Furthermore, we can explicitly define the Butcher table (16) using only a coefficient matrix $\mathbf{A}$ and the time evaluation (18) is reduced to $\Phi^{n+1}=\Phi_s$. Let us consider a general matrix $\mathbf{A}$ with s

stages. In order to evaluate one time step, we set $\Phi_0 = \Phi^n$ and evaluate

$$\Phi_i = \Phi_0 - \Delta t \sum_{j=1}^{s} a_{ij} \operatorname{grad} F\left(\Phi_j\right), \tag{19}$$

for $i = 1, 2, \ldots, s$ and we have $\Phi^{n+1} = \Phi_s$. Note that the implicit Euler method (5) can be represented as

$$\Phi_1 = \Phi_0 - \Delta t \operatorname{grad} F\left(\Phi_1\right), \tag{20}$$

where $\Phi_0 = \Phi^n$ and $\Phi_1 = \Phi^{n+1}$, so that (5) is one of the SARK methods with a trivial coefficient matrix $\mathbf{A} = (1)$.

Next, we present a condition to make the proposed scheme (19) unconditionally energy stable.

Condition PD (*Positive Definitness*)**.** *The RK method (17) is said to be satisfying the positive definite condition if a row difference matrix $\widetilde{\mathbf{A}}$ is positive definite after a symmetric transformation, where $\widetilde{\mathbf{A}} = \mathbf{A} - S\mathbf{A}$ is defined as follows with a shift matrix $S_{ij} = \delta_{i,j+1}$,*

$$\widetilde{\mathbf{A}} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1s} \\ \tilde{a}_{21} & \tilde{a}_{22} & \cdots & \tilde{a}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{s1} & \tilde{a}_{s2} & \cdots & \tilde{a}_{ss} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{pmatrix} - \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{11} & a_{12} & \cdots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s-1,1} & a_{s-1,2} & \cdots & a_{s-1,s} \end{pmatrix}. \tag{21}$$

Here, the positive definiteness of $\widetilde{\mathbf{A}}$ after a symmetric transformation means that the symmetric matrix $\left(\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^T\right)/2$ is positive definite where $\left(\Phi, \widetilde{\mathbf{A}}\Phi\right) = \frac{1}{2}\left(\Phi, \left(\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^T\right)\Phi\right) > 0$.

Remark 2. A RK method with the positive definite condition must be an implicit RK method which means that a_{ij} is a non-zero for some $i \le j$.

Remark 3. The implicit Euler method (5) shown in Section 2.1, as an example for the unconditional energy stability, can be described by the SARK method (19) with the coefficient matrix $\mathbf{A} = (1)$. The method satisfies the positive definite condition since $\widetilde{\mathbf{A}} = (1)$ is obviously positive definite.

Theorem 2 (*Unconditional Energy Stability*)**.** *A SARK method (19) satisfying the positive definiteness of $\widetilde{\mathbf{A}}$ (Condition PD), referred to as SARK-PD, is unconditionally energy stable, meaning that for any time step size $\Delta t > 0$,*

$$F\left(\Phi^{n+1}\right) \le F\left(\Phi^n\right). \tag{22}$$

Proof. Defining $[\![\Phi]\!]_m^n = \Phi_n - \Phi_m$ and $F_j = F\left(\Phi_j\right)$, we can rewrite (19) as

$$[\![\Phi]\!]_0^i = -\Delta t \sum_{j=1}^{s} a_{ij} \operatorname{grad} F_j, \tag{23}$$

for $i = 1, 2, \ldots, s$. The difference between the two adjacent stages can be calculated as

$$[\![\Phi]\!]_{i-1}^i = [\![\Phi]\!]_0^i - [\![\Phi]\!]_0^{i-1} = -\Delta t \sum_{j=1}^{s} \tilde{a}_{ij} \operatorname{grad} F_j. \tag{24}$$

According to Lemma 1, the energy difference between two adjacent stages is

$$[\![F(\Phi)]\!]_{i-1}^i \le \left\langle \operatorname{grad} F_i, [\![\Phi]\!]_{i-1}^i \right\rangle. \tag{25}$$

After summing up, we have

$$[\![F(\Phi)]\!]_0^s = \sum_{i=1}^{s} [\![F(\Phi)]\!]_{i-1}^i \le \sum_{i=1}^{s} \left\langle \operatorname{grad} F_i, [\![\Phi]\!]_{i-1}^i \right\rangle = -\Delta t \left\langle \operatorname{grad} \mathbf{F}, \widetilde{\mathbf{A}} \operatorname{grad} \mathbf{F} \right\rangle_s, \tag{26}$$

where $\operatorname{grad} \mathbf{F} = (\operatorname{grad} F_1, \ldots, \operatorname{grad} F_s)^T$ and we define an s-dimensional inner product as $\langle \mathbf{\Phi}, \mathbf{\Psi} \rangle_s = \sum_{i=1}^{s} \langle \Phi_i, \Psi_i \rangle$ for $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_s)^T$ and $\mathbf{\Psi} = (\Psi_1, \ldots, \Psi_s)^T$. Since $\widetilde{\mathbf{A}}$ is positive definite, $[\![F(\Phi)]\!]_0^s \le 0$. It follows that the energy dissipation $F\left(\Phi^{n+1}\right) = F(\Phi_s) \le F(\Phi_0) = F(\Phi^n)$. □

Therefore, we can design an energy stable method for the SARK methods using only the positive definiteness of $\widetilde{\mathbf{A}}$ and examples will be given in Sections 4 and 5. Stiffly accurate and positive definite conditions play important roles in developing an unconditionally energy stable RK method.

Remark 4. The Crank–Nicolson method (12) shown in Section 2.2 as a counterexample for energy stability can be described by the RK table

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}, \tag{27}$$

and it does not satisfy the positive definite condition. Furthermore, (27) is an example of the explicit first-stage RK methods in [12], which means that the first stage corresponds to the explicit stage, $a_{1j} = 0$ for all j. There is no explicit first-stage RK method satisfying the positive definite condition since the first row of the coefficient matrix is all zero.

Remark 5. The stiffly accurate condition is crucial to guaranteeing the energy stability in (26). Without assuming that $b_j = a_{sj}$ for $j = 1, 2, \ldots, s$, it can be easily derived from the convexity of $F(\phi)$ that

$$\begin{aligned} F\left(\Phi^{n+1}\right) - F\left(\Phi^{n}\right) &= F\left(\Phi^{n+1}\right) - F\left(\Phi_s\right) + \sum_{i=1}^{s} [\![F(\Phi)]\!]_{i-1}^{i} \\ &\le \left\langle \operatorname{grad} F\left(\Phi^{n+1}\right),\ \Phi^{n+1} - \Phi_s \right\rangle + \sum_{i=1}^{s} \left\langle \operatorname{grad} F_i,\ [\![\Phi]\!]_{i-1}^{i} \right\rangle \\ &= \sum_{j=1}^{s} \left\langle \operatorname{grad} F\left(\Phi^{n+1}\right),\ -\Delta t(b_j - a_{sj}) \operatorname{grad} F_j \right\rangle - \Delta t \left\langle \operatorname{grad} \mathbf{F},\ \widetilde{\mathbf{A}} \operatorname{grad} \mathbf{F} \right\rangle_s . \end{aligned} \tag{28}$$

The inequality $\frac{F\left(\Phi^{n+1}\right) - F\left(\Phi^{n}\right)}{\Delta t} \le -\left\langle \operatorname{grad} \mathbf{F}, \widetilde{\mathbf{A}} \operatorname{grad} \mathbf{F} \right\rangle_s$ in (26) holds only with the stiffly accurate condition, $b_j = a_{sj}$. Therefore, the unconditional energy stability can be achieved under the positive definiteness of $\widetilde{\mathbf{A}}$ with the stiffly accurate condition.

4. Unique solvability of implicit Runge–Kutta methods

Since the given equations (19) can be nonlinear in general, we need to consider the solvability issue. In this section, we introduce two kinds of requirements for the solvability of RK method; one is for the fully implicit RK method.

Theorem 3 (*Unique Solvability for Fully Implicit RK*)**.** *If the matrix* $\mathbf{A}$ *is symmetric and positive definite, then the RK scheme is uniquely solvable for any time step size* $\Delta t > 0$.

Proof. Considering all stages $i = 1, 2, \ldots, s$ of (19), we can rewrite as

$$\mathbf{\Phi} - \mathbf{\Phi}_0 = -\Delta t \mathbf{A} (\operatorname{grad} \mathbf{F}) \tag{29}$$

where $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_s)^T$, $\mathbf{\Phi}_0 = (\Phi_0, \ldots, \Phi_0)^T$, and $\operatorname{grad} \mathbf{F} = (\operatorname{grad} F_1, \ldots, \operatorname{grad} F_s)^T$. Since $\mathbf{A}^{-1}$ is positive definite symmetric and F is convex, the functional

$$Q(\mathbf{\Phi}) = \frac{1}{2} \left\langle \mathbf{\Phi} - \mathbf{\Phi}_0, \mathbf{A}^{-1} (\mathbf{\Phi} - \mathbf{\Phi}_0) \right\rangle + \Delta t \left\langle F(\mathbf{\Phi}), \mathbf{1} \right\rangle \tag{30}$$

has a unique minimizer for any time step size $\Delta t > 0$. Thus, (29) has a unique solution $\mathbf{\Phi}$ at the unique minimizer, $\operatorname{grad} Q(\mathbf{\Phi}) = 0$. □

Remark 6. We can specifically find a symmetric coefficient matrix $\mathbf{A}$ for the three-stage second order RK method

$$\mathbf{A} = \begin{pmatrix} \frac{59}{56} & -\frac{11}{7} & \frac{65}{112} \\ -\frac{11}{7} & \frac{45}{14} & -\frac{79}{112} \\ \frac{65}{112} & -\frac{79}{112} & \frac{9}{8} \end{pmatrix} \tag{31}$$

which is positive definite. Of course, it satisfies the positive definite condition in Theorem 2, and the coefficients $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$ made from the stiffly accurate condition and $\mathbf{c} = \mathbf{A1}$ satisfy the order conditions in Table 1 for the second-order accuracy.

Next, we discuss another requirement for the solvability. Methods with full matrices like (31) require the solution of s simultaneous implicit equations per step. Considering a lower triangular matrix to circumvent this difficulty, we simply need to solve a single implicit equation for each stage, which is motivated by the diagonally implicit RK method [10]. This means that, in implicit RK, we select a lower triangular type for the coefficient matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{pmatrix}. \tag{32}$$

With the lower triangular matrix (32), the implicit RK method can be rewritten as follows: First, we set a zeroth-stage as $\Phi_0 = \Phi^n$. Next, we calculate

$$\Phi_i = \Phi_0 - \Delta t \sum_{j=1}^{i} a_{ij} \operatorname{grad} F\left(\Phi_j\right), \tag{33}$$

for each stage $i = 1, 2, \ldots, s$. Finally, we set the next time approximation as $\Phi^{n+1} = \Phi_s$.

Theorem 4 (*Unique Solvability for Diagonally Implicit RK*)**.** *The implicit RK method with the diagonally implicit type (33) is uniquely solvable for any time step size $\Delta t > 0$, provided that $a_{ii} \geq 0$ for all i.*

Proof. For each stage $i = 1, 2, \ldots, s$ of (33), we need to solve

$$\Phi_i + a_{ii}\Delta t \operatorname{grad} F\left(\Phi_i\right) = S_i, \tag{34}$$

where

$$S_i = \Phi_0 - \Delta t \sum_{j=1}^{i-1} a_{ij} \operatorname{grad} F\left(\Phi_j\right). \tag{35}$$

Since F is convex and $a_{ii} \geq 0$, the functional

$$Q\left(\Phi\right) = \frac{1}{2}\left\|\Phi\right\|^2 + a_{ii}\Delta t F\left(\Phi\right) - \langle \Phi, S_i \rangle \tag{36}$$

has a unique minimizer. At the unique minimizer for ith stage, $\operatorname{grad} Q\left(\Phi\right) = 0$ and (34) has a unique solution Φ_i. Therefore, the proposed scheme (33) is shown to be uniquely solvable for any time step size $\Delta t > 0$. □

5. Numerical results

For the numerical test, we consider a simple convex gradient flow

$$\frac{\partial \phi}{\partial t} = -g\left(\phi\right) + \epsilon \Delta \phi, \tag{37}$$

where

$$g\left(\phi\right) = G'\left(\phi\right) \text{ and } G\left(\phi\right) = \frac{1}{4}\left(\phi^4 - 4\phi^3 + 6\phi^2\right), \tag{38}$$

and ϵ is a scale parameter between the reaction and diffusion terms. In this paper, we specifically choose $\epsilon = 10^{-4}$. In order to complete the system, we consider a finite domain $\Omega = [0, 1]$ with a zero Neumann boundary condition. Furthermore, (37) can be derived by the gradient flow for the energy functional

$$\mathcal{F}\left(\phi\right) = \int_{\Omega} \left(G\left(\phi\right) + \frac{\epsilon}{2}\left|\nabla \phi\right|^2\right) d\mathbf{x} \tag{39}$$

in the L_2 inner product space. Note that (39) is a convex functional, and thus (37) has only one equilibrium point.

Now, for the time integration, we introduce some coefficient matrices $\mathbf{A}$ up to the fourth-order accuracy, which satisfies the order conditions in Table 1 as well as the positive definite condition in Theorem 2. For practical reasons, we present coefficient matrices for diagonally implicit RK (DIRK) methods. For the sake of being self-consistent, we start with the first-order method $\mathbf{A}_1 = (1)$. The second- and third-order methods are given as follows

$$\mathbf{A}_2 = \begin{pmatrix} \frac{1}{6} & 0 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix} \tag{40}$$

and

$$\mathbf{A}_3 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}. \tag{41}$$

For the fourth-order method, it is relatively easy to find it in the singly diagonally implicit RK method (SDIRK), which means that all a_{ii} are equal in the DIRK method. Here, we introduce a five-stage method

$$\mathbf{A}_4 = \begin{pmatrix} 0.6000 & 0 & 0 & 0 & 0 \\ -0.3600 & 0.6000 & 0 & 0 & 0 \\ 0.2079 & 0.1921 & 0.6000 & 0 & 0 \\ -0.0291 & -0.0507 & 0.2337 & 0.6000 & 0 \\ -0.3678 & 0.5964 & -0.6164 & 0.7878 & 0.6000 \end{pmatrix}, \tag{42}$$

where the coefficients are rounded to four digits. We refer to chapter IV.6 in [17] for computing the coefficients of (42) from the coefficient set of $c_1 = 0.6$, $c_2 = 0.24$, and $c_3 = 1$.

The positive definiteness of $\widetilde{\mathbf{A}}$ is easily checked if the minimum of eigenvalues of $\left(\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^T\right)/2$ is positive. The minimum eigenvalues corresponding to $\mathbf{A}_2$, $\mathbf{A}_3$, and $\mathbf{A}_4$ are 0.0373, 0.0403, and 0.0032, respectively. Therefore, all listed tables assure the unconditional energy stability. In addition, all coefficients in the diagonal are positive values, indicating that those tables guarantee the unique solvability by Theorem 4.

Remark 7. The SARK-PD methods with the coefficient matrices (40)–(42) do not satisfy the algebraic stability [7]. This implies that the proposed SARK-PD method is a different type of energy stable method without using the algebraic stability.

5.1. Evolution of the solution for a one-dimensional example

We begin by showing the solution evolution of the example Eq. (37) with zero Neumann boundary condition and the initial condition

$$\phi(x, 0) = 2 + \cos(8\pi x)\cos(13\pi x) + \cos(4\pi x)\cos(13\pi x) \tag{43}$$

on a domain $\Omega = [0, 1]$. For the numerical simulations, the numerical solution is evolved to time $T_f = 8$.

After spatial discretization of (37), we apply RK methods to the resulting system of ordinary differential equations for the time integration. Since we focus only on the time-marching method with the zero Neumann boundary condition, we employ a discrete cosine transform [18] in MATLAB for spatial discretization,

$$\{x_i = (i - 0.5)\,\Delta x \text{ for } i = 1, \ldots, M_x\}, \tag{44}$$

where Δx is a space step size and M_x is a grid point number. The corresponding discrete energy is

$$F(\Phi) = (G(\Phi), \mathbf{1})_{l_2} - \frac{\epsilon}{2}(\Phi, \Delta_h \Phi)_{l_2}, \tag{45}$$

where $(\cdot, \cdot)_{l_2}$ is a discrete l_2 inner product and Δ_h is a discrete Laplacian corresponding to the discrete cosine transform. For a practical description, we refer to [19].

Fig. 2 shows the time evolution of the solution with a sufficiently small time step $\Delta t = T_f/2^{13}$ using the fourth-order table $\mathbf{A}_4$ with a fixed grid size $\Delta x = 1/128$. It changes to a plain profile and quickly goes down to the equilibrium solution $\phi = 0$.

We demonstrate the numerical convergence in space with the same conditions and parameters used at the beginning of this subsection. In order to estimate the convergence rate with respect to a grid size, simulations are performed by varying the grid points $M_x = 16, 24, 32, \ldots, 196, 256$. Fig. 3 shows the relative l_2-errors with respect to various grid size for various time steps $\Delta t = T_f/2^{11}, \ldots, T_f/2^5$. Here, the errors are computed through comparison with the reference numerical solution obtained using the fourth-order method with a time step $\Delta t = T_f/2^{13}$ and 384 grid points. As can be seen, the spatial convergence of the results under the grid refinements is evident. Furthermore, it shows that 128 grid points provide more than single precision sufficient spatial accuracy to estimate the numerical convergence with respect to the time steps.

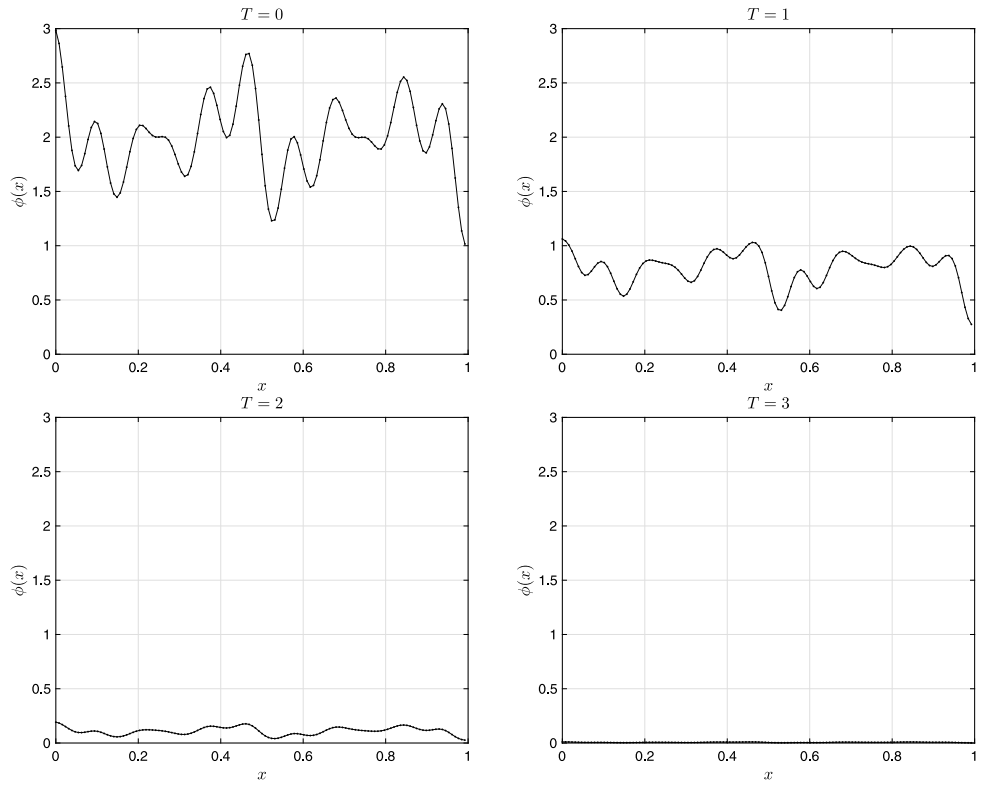

Fig. 2. Time evolution of solution in 1D.

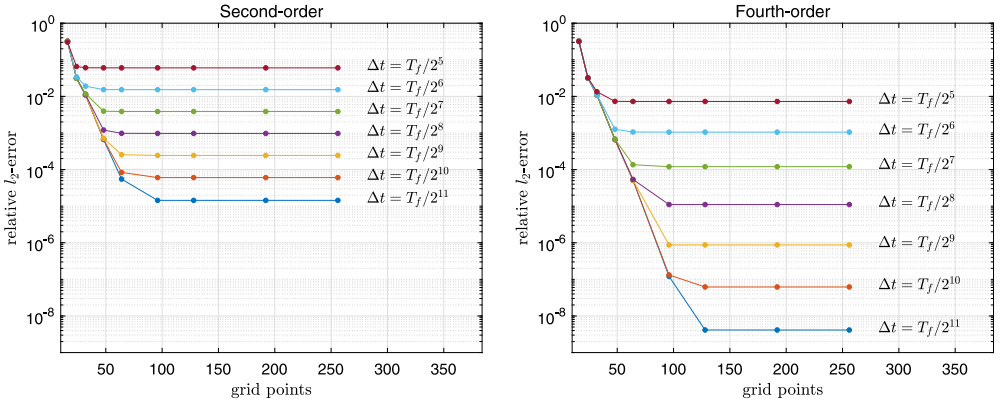

Fig. 3. Relative l_2-errors at $T = 2$ with respect to the grid points for various time steps Δt labeled on the right-end of the corresponding result.

5.2. Temporal accuracy test

We demonstrate the numerical convergence using the same conditions and parameters used in the previous subsection. In order to estimate the convergence rate with respect to the time step Δt, simulations are performed by varying the time steps $\Delta t = T_f/2^{11}, T_f/2^{10}, \ldots, T_f/2^4$.

Fig. 4 shows the relative l_2-errors of $\phi\,(x, t)$ with various time steps. Here, the errors are computed through comparison with the reference solution shown in the previous subsection. It is observed that the SARK-PD method with $\mathbf{A}_q$ gives the desired qth order accuracy in time.

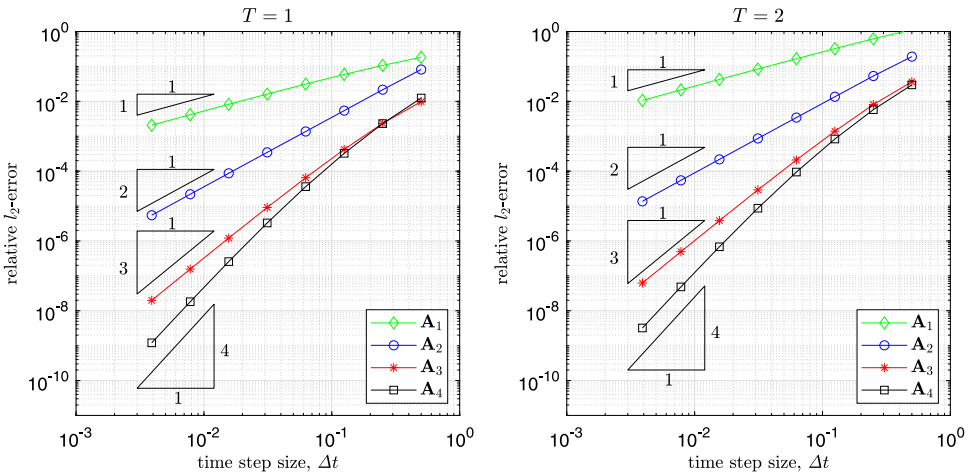

Fig. 4. Relative l_2-errors for various time steps Δt.

5.3. Energy stability test

We display the energy evolutions with the same conditions and parameters used in the previous subsection. Simulations are performed by relatively large time steps $\Delta t = 1, 2, 4$.

Fig. 5 shows the energy evolutions with large time steps, and energy dissipation is observed for all simulations. The solid line indicates the energy evolution of the reference solution shown in the previous subsection with a small time step $\Delta t = T_f/2^{13}$.

5.4. Comparison to algebraically stable methods

We compare the SARK-PD methods with the second- and fourth-order algebraically stable (AG-stable) RK methods introduced in [20],

$$\begin{array}{c|cc} \lambda & \lambda & 0 \\ 1-\lambda & 1-2\lambda & \lambda \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}, \tag{46}$$

where $\lambda \geq 1/4$ and

$$\begin{array}{c|ccc} \lambda & \lambda & 0 & 0 \\ \frac{1}{2} & \frac{1}{2}-\lambda & \lambda & 0 \\ 1-\lambda & 2\lambda & 1-4\lambda & \lambda \\ \hline & b_1 & b_2 & b_3 \end{array}, \tag{47}$$

where $b_1 = b_3 = 1/(6(1-2\lambda)^2)$, $b_2 = 1 - 2b_1$, and $\lambda = 1.06857902130163$. We refer to as $\mathbf{A}_2^{AG}$ and $\mathbf{A}_4^{AG}$ for (46) with $\lambda = 1/4$ and (47), respectively. Note that, unlike the proposed SARK-PD method, AG-stable RK methods (46) and (47) do not satisfy the stiffly accurate condition. Furthermore, as mentioned in Remark 7, the SARK-PD methods (40)–(42) are not AG-stable RK methods.

Since both RK methods (40)–(42) and (46)–(47) are diagonally implicit schemes, we employ the same Newton-type iteration with a bi-conjugate gradient method for the nonlinear equation (37)–(38). The computational cost is proportional to the number of RK stages s multiplied by the number of time step $T_f/\Delta t$. Fig. 6 shows the computational times for both methods with respect to the time steps. The computational times for the second-order methods are same since $s = 2$ for both methods, however, the time for the fourth-order AG-stable method is 3/5-times faster than the corresponding SARK-PD method.

Fig. 7 shows the relative l_2-errors of $\phi(x, t)$ with respect to the computational time corresponding to various time steps. The computational time increases as the error decreases and the ratio is approximately the order of convergence. The comparison shows that both of the SARK-PD method and the AG-stable RK method provide the desired order of accuracy as the computational time increases. The computational cost of DIRK methods is closely related to the number of stages and the relative l_2-error depends on the leading coefficients of the error expansions. As can be seen, computation

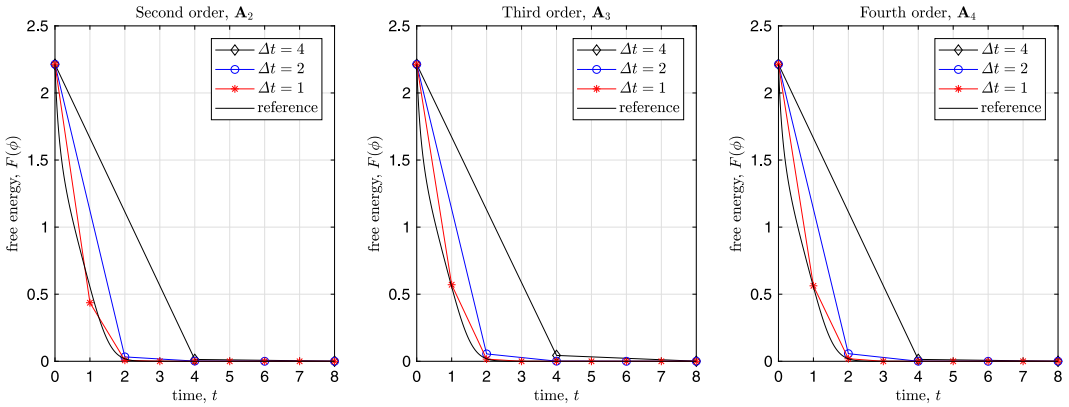

Fig. 5. Energy evolution for various time steps Δt.

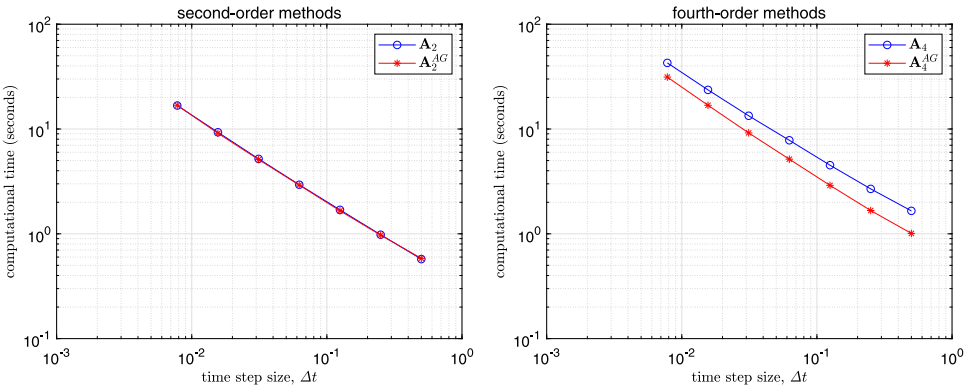

Fig. 6. Computational time for the solution at $T_f = 8$ with respect to time steps Δt.

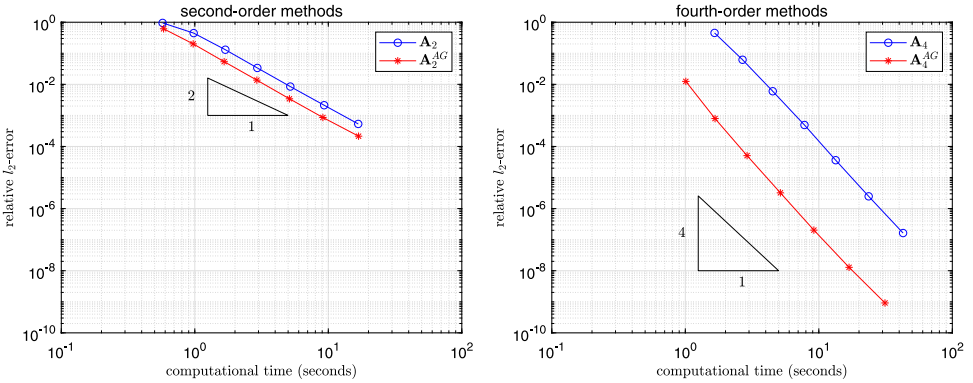

Fig. 7. Relative l_2-errors at $T_f = 8$ with respect to computational time.

efficiency of the AG-stable methods is superior than the proposed methods. It is problem specific and depends on the choice of the RK table so that we need further study for an optimized choice of the RK table. We leave it for future work since main contribution of this paper is to introduce another type of RK methods for guaranteeing the energy stability.

6. Conclusions

We proposed unconditionally energy stable methods to solve a convex gradient flow which is a fundamental problem in various dynamical systems. The proposed methods are high order accurate in time and assure the unconditional energy stability and unique solvability. Two additional conditions, stiffly accurate and positive definiteness, play important roles

in the energy stability of the Runge–Kutta methods. We demonstrated the existence of high order energy stable methods by presenting the specific RK tables up to fourth-order accuracy. One can request higher accuracy; however, we leave this for future work, because we think the fourth-order to be reasonably high accuracy for discussing a new possible class of RK methods for the unconditional energy stability. We presented numerical experiments to show high order accuracy and energy stability using a specific example. The simple comparison with algebraically stable RK methods do not show computational superiority of the proposed methods so we leave the optimizations of the RK tables with better performance for future work.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Korean Government MOE (2017R1D1A1B0-3032422, 2019R1A6A1A1-1051177) and MSIP, Korea (2017R1E1A1A0-3070161).

References

[1] J.K. Hale, Asymptotic Behavior of Dissipative Systems, American Mathematical Society, Providence, RI, 1988, p. ix + 198.
[2] J.W. Cahn, J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, J. Chem. Phys. 28 (2) (1958) 258–267.
[3] S.M. Allen, J.W. Cahn, A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening, Acta Metall. 27 (6) (1979) 1085–1095.
[4] K. Elder, M. Grant, Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals, Phys. Rev. E 70 (5) (2004) 051605.
[5] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations 26 (1–2) (2001) 101–174.
[6] W. Bao, Q. Du, Computing the ground state solution of Bose–Einstein condensates by a normalized gradient flow, SIAM J. Sci. Comput. 25 (5) (2004) 1674–1697.
[7] A.M. Stuart, A.R. Humphries, Model problems in numerical stability theory for initial value problems, SIAM Rev. 36 (2) (1994) 226–257.
[8] D. Furihata, T. Matsuo, Discrete Variational Derivative Method: a Structure-Preserving Numerical Method for Partial Differential Equations, CRC Press, 2010.
[9] R.I. McLachlan, G. Quispel, N. Robidoux, Geometric integration using discrete gradients, Phil. Trans. R. Soc. A 357 (1754) (1999) 1021–1045.
[10] R. Alexander, Diagonally implicit Runge–Kutta methods for stiff ODE's, SIAM J. Numer. Anal. 14 (6) (1977) 1006–1021.
[11] E. Hairer, C. Lubich, Energy-diminishing integration of gradient systems, IMA J. Numer. Anal. 34 (2) (2013) 452–461.
[12] L.M. Skvortsov, Diagonally implicit Runge–Kutta methods for stiff problems, Comput. Math. Math. Phys. 46 (12) (2006) 2110–2123.
[13] M. Günther, A. Sandu, Multirate generalized additive Runge Kutta methods, Numer. Math. 133 (3) (2016) 497–524.
[14] J. Shin, H.G. Lee, J.-Y. Lee, Convex splitting Runge–Kutta methods for phase-field models, Comput. Math. Appl. 73 (11) (2017) 2388–2403.
[15] J. Shin, H.G. Lee, J.-Y. Lee, Unconditionally stable methods for gradient flow using convex splitting Runge–Kutta scheme, J. Comput. Phys. 347 (2017) 367–381.
[16] E. Hairer, G. Wanner, Solving Ordinary Differential Equations. I, Vol. 8, Springer-Verlag, Berlin, 1993.
[17] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Vol. 14, Springer-Verlag, Berlin, 1996.
[18] N. Ahmed, T. Natarajan, K.R. Rao, Discrete cosine transform, IEEE Trans. Comput. 100 (1) (1974) 90–93.
[19] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes 3rd Edition: The Art of Scientific Computing, Cambridge university press, 2007.
[20] K. Burrage, J.C. Butcher, Stability criteria for implicit Runge–Kutta methods, SIAM J. Numer. Anal. 16 (1) (1979) 46–57.