Fourth-Order Structure-Preserving Method for the Conservative Allen-Cahn Equation

Xiaowei Chen, Xu Qian and Songhe Song*

Department of Mathematics, National University of Defense Technology, Changsha, Hunan 410073, China

Received 24 October 2021; Accepted (in revised version) 22 February 2022

Abstract. We propose a class of up to fourth-order maximum-principle-preserving and mass-conserving schemes for the conservative Allen-Cahn equation equipped with a non-local Lagrange multiplier. Based on the second-order finite-difference semi-discretization in the spatial direction, the integrating factor Runge-Kutta schemes are applied in the temporal direction. Theoretical analysis indicates that the proposed schemes conserve mass and preserve the maximum principle under reasonable time step-size restriction, which is independent of the space step size. Finally, the theoretical analysis is verified by several numerical examples.

AMS subject classifications: 65N06, 65N12

Key words: Maximum-principle-preserving, mass-conserving scheme, the conservative Allen-Cahn equation.

1 Introduction

The classical Allen-Cahn (AC) equation was proposed by Allen and Cahn [1] in 1979 to describe the phenomenological model of the inverse phase boundary motion in crystals. As an important class of phase field models, the AC equation has been widely applied in image processing [2], mean curvature motion, materials science [3, 4], and so on. In recent years, many studies have been conducted on the classical AC equation [5–8].

The classical AC equation is considered as a well-known prototypical gradient flow

$$\frac{\partial u(x,t)}{\partial t} = \epsilon^2 \Delta u(x,t) + f(u(x,t)), \quad x \in \Omega, \quad t > 0,$$

(1.1)

where $\Omega = [a,b] \subseteq \mathbb{R}$ is the bounded domain. The parameter $\epsilon > 0$ and $u$ usually represent the interfacial width and the difference between the concentrations of two mixtures’ components, respectively. The symbol $\Delta$ denotes the usual Laplacian operator and $f(u)$ is the...
negative derivative of a polynomial double-well potential, i.e., \( f(u) = -F'(u) \). Consider the initial and periodic boundary conditions

\[
\begin{align*}
  u(x,0) &= u_0(x), \quad x \in \Omega, \\
  u(a,t) &= u(b,t), \quad t \geq 0.
\end{align*}
\]

The \( L^2 \) inner product and norm are denoted as

\[
\langle f, g \rangle = \int_{\Omega} fg \, dx, \quad \| f \| = \left( \int_{\Omega} |f|^2 \, dx \right)^{\frac{1}{2}},
\]

respectively. The \( L^\infty \) norm is defined as

\[
\| f \|_{L^\infty} = \max_{x \in \Omega} |f(x)|.
\]

The energy functional of the classical AC equation is defined as

\[
E[u] = \frac{\epsilon^2}{2} \langle \nabla u, \nabla u \rangle + \langle F(u), 1 \rangle = \int_{\Omega} \left( \frac{\epsilon^2}{2} |\nabla u(x,t)|^2 + F(u(x,t)) \right) \, dx,
\]

where

\[
F(u) = \frac{1}{4} (u^2 - 1)^2, \quad f(u) = -F'(u) = u - u^3.
\]

By taking the \( L^2 \) inner product of Eq. (1.1) with \( \partial_t u(x,t) \), we obtain

\[
\frac{d}{dt} \int_{\Omega} u(x,t) \, dx = -\int_{\Omega} |\partial_t u(x,t)|^2 \, dx \leq 0, \quad \forall t > 0.
\]

Thus, the classical AC equation satisfies the energy dissipation law. By taking the \( L^2 \) inner product of Eq. (1.1) with 1, we have

\[
\frac{d}{dt} \int_{\Omega} u(x,t) \, dx = \epsilon^2 \int_{\Omega} \Delta u(x,t) \, dx + \int_{\Omega} f(u(x,t)) \, dx, \quad \forall t > 0.
\]

It can be proven that the classical AC equation can not conserve the mass unless

\[
\int_{\Omega} f(u) \, dx = 0.
\]

In this paper, by introducing a Lagrange multiplier

\[
\lambda = \frac{1}{|\Omega|} \int_{\Omega} f(u(x,t)) \, dx,
\]

the conservative modification of the classical AC equation is expressed as [30]

\[
\partial_t u(x,t) = \epsilon^2 \Delta u(x,t) + \tilde{f}(u(x,t)), \quad x \in \Omega, \quad t > 0,
\]
where

\[ \tilde{f}(u(x,t)) = f(u(x,t)) - \frac{1}{|\Omega|} \int_{\Omega} f(u(x,t)) \, dx = f(u(x,t)) - \lambda(t). \]  \tag{1.7} \]

It is well-known that the classical AC equation satisfies the maximum-principle-preserving (MPP) property, i.e., if the initial value satisfies \( \|u_0\|_{L^\infty} \leq \beta \), then the solution satisfies \( \|u(t)\|_{L^\infty} \leq \beta \) \cite{9}. In recent years, the development of high-order accurate structure-preserving algorithms has been a hot topic, including the dissipation of energy, conservation of mass, and preservation of maximum principle. The first- and second-order MPP scheme are constructed by using the exponential time difference (ETD) method in the temporal direction. By introducing a Lagrange multiplier, Li \cite{10} obtained the AC equation with mass conservation. Then, the second-order stable difference scheme of the mass-conserved AC equation is constructed by using the exponential difference method in the temporal direction. Besides, MPP can be obtained unconditionally. Du \cite{11, 12} exploited the ETD method to construct the first- and second-order discrete MPP schemes. According to the nonlinear parabolic equation, an MPP scheme that can effectively solve the radiation diffusion and nonlinear heat wave problem was constructed by Peng \cite{13}. Liao \cite{14} presented a second-order MPP time-stepping scheme for the time-fractional AC equation with nonuniform time steps. Moreover, a space-time related Lagrange multiplier was first introduced by Kim \cite{15} to the AC equation to strengthen the conservation of mass. Then, Kim applied an operator splitting method in the spatial direction to obtain the semi-discrete scheme. Meanwhile, by employing the Crank-Nicolson (CN) method, the second-order full-discrete form was obtained. In order to construct the MPP discrete scheme for the general AC equation with a nonlinear mobility, Shen \cite{16} first combined the central finite difference method for approximating the diffusion term with the upwind scheme for the advection term in the spatial direction. Then, they adopted the standard semi-implicit scheme in the temporal direction. Besides, a second-order scheme \cite{17} that can both conserve mass and boundary for the Keller-Segel equation was obtained. Furthermore, there are lots of other studies of MPP \cite{18–22}. In addition to the MPP, the energy dissipation is also an important physical property for the conservative AC equation. There exists many works denoted to the energy stable numerical scheme, for example, implicit method \cite{23}, the convex splitting method \cite{24}, linear stabilized method \cite{25}, invariant energy quadratization method \cite{26}, scalar auxiliary variable method \cite{27}, Lagrange multiplier \cite{28} and cut-off post-processing method \cite{29}.

The exponential time integrators have been widely investigated recently \cite{30–33}. To handle the stiff nonlinear part well regardless of homogeneous or inhomogeneous boundary conditions, Ju \cite{34} combined the stabilized ETD method with the fast Fourier transform (FFT) in the spatial direction. Kassam \cite{35} proposed an improved fourth-order Runge-Kutta (RK) method for solving stiff nonlinear partial differential equations. Meanwhile, the application and comparison of the KdV equation and the classical AC equation were conducted by using the time-splitting method and other factors. After this, the strong stability preserving (SSP) integrating factor Runge-Kutta (IFRK) method was first
proposed by Iserwood [36]. The problem that implicit and implicit-explicit methods have strict requirements on the time step was solved, and the calculation efficiency is greatly improved. Then, Zhang [37] used the IFRK method to propose a class of high order maximum principle preserving schemes for solving the AC equation and Ju [38] studied the fully-discrete maximum bound principle (MBP)-preserving IFRK method instead of the SSP property for a class of semilinear parabolic equations. Nevertheless, the current studies of the AC equation with mass conservation or MPP can achieve third-order accuracy in the temporal direction by applying the stabilized IFRK method [39]. Aiming at this problem, a scheme with unconditional mass conservation and MPP is proposed in this paper, and the proposed scheme can achieve a high degree of convergence in the temporal direction.

The rest of this paper is arranged as follows. In Section 2, the conservative AC equation which is obtained by adding a Lagrange multiplier can achieve MPP and conserve mass unconditionally. In Section 3, the finite difference method is exploited in the spatial direction and the IFRK method is applied in the temporal direction to construct the full-discrete scheme of the conservative AC equation. It is proved that the high-order full-discrete scheme can achieve MPP and conserve mass unconditionally. In Section 4, the correctness of the full-discrete scheme is verified by 1D, 2D, and 3D experiments. Some concluding remarks are presented in Section 5.

2 Preliminaries

Different from the classical AC equation (1.1), the conservative AC equation (1.6) can not only keep energy dissipation and satisfy MPP, but also preserve the conservation law of mass. In this section, we introduce the properties of the conservative AC equation.

**Theorem 2.1.** The conservative AC equation (1.6) satisfies the energy dissipation law

\[ \frac{d}{dt} E[u(x,t)] \leq 0, \quad \forall t > 0. \] (2.1)

**Proof.** By taking \( L^2 \) inner product of Eq. (1.6) with \( \partial_t u(x,t) \), the following energy dissipation law is obtained

\[ \frac{d}{dt} E[u(x,t)] = \left\langle \frac{\delta E}{\delta u} \partial_t u(x,t) \right\rangle = - \int_{\Omega} |\partial_t u(x,t)|^2 dx \leq 0, \quad \forall t > 0. \] (2.2)

Thus, we complete the proof. \( \square \)

**Theorem 2.2.** The conservative AC equation (1.6) preserves the mass conservation law

\[ \frac{d}{dt} M = 0, \quad \forall t > 0, \] (2.3)

where \( M := \langle u, 1 \rangle \).
Proof. Taking the $L^2$ inner product with 1 on both sides of Eq. (1.6), we have

$$
\frac{d}{dt} \int_{\Omega} u(x,t)dx = \epsilon^2 \int_{\Omega} \Delta u(x,t)dx + \int_{\Omega} \bar{f}(u(x,t))dx,
$$

(2.4)

where

$$
\int_{\Omega} \Delta u(x,t)dx = 0
$$

(2.5)

can be deduced based on the periodic boundary conditions. Meanwhile,

$$
\int_{\Omega} \bar{f}(u(x,t))dx = 0,
$$

(2.6)

so we can get

$$
\frac{d}{dt} \int_{\Omega} u(x,t)dx = 0,
$$

(2.7)

which is equivalent to

$$
\frac{d}{dt} M = 0, \ \forall t > 0.
$$

It indicates that the conservative AC equation can conserve mass unconditionally. \qed

Assumption 2.1 ([40]). There exists a constant $\beta > 0$ such that

$$
\forall \omega \in [-\beta, \beta], \ f(\beta) \leq f(\omega) \leq f(-\beta).
$$

(2.8)

Theorem 2.3 ([40]). For the conservative AC equation (1.6), if Assumption 2.1 holds and the initial value satisfies $\|u(x,0)\|_{L^\infty} \leq \beta$ for any $x \in \Omega$, we have $\|u(x,t)\|_{L^\infty} \leq \beta$ for any $x \in \Omega$.

We choose the polynomial double-well potential function

$$
F(u) = \frac{1}{4}(u^2 - 1)^2, \ f(u) = -F'(u) = u - u^3.
$$

(2.9)

Let $f'(u) \geq 0$, we can deduce that $f(u)$ is monotonically nondecreasing function while $u \in [-\sqrt{3}^3, \sqrt{3}^3]$. Therefore

$$
f\left(-\frac{\sqrt{3}}{3}\right) \leq f(u) \leq f\left(\frac{\sqrt{3}}{3}\right), \ u \in \left[-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right].
$$

(2.10)

By calculation and the plot of $f(u)$ shown in [10], we can obtain

$$
f\left(2\frac{\sqrt{3}}{3}\right) \leq f(u) \leq f\left(-2\frac{\sqrt{3}}{3}\right), \ u \in \left[-\frac{2\sqrt{3}}{3}, \frac{2\sqrt{3}}{3}\right].
$$

(2.11)

In combination with the Assumption 2.1, we can know that $f$ satisfies $\beta \in [2\sqrt{3}/3, +\infty)$. 
3 High order maximum-principle-preserving and mass-conserving schemes

Many methods have been proposed in previous studies to solve the conservative AC equation, such as the ETD1, ETDRK2, and CN schemes. However, these methods can only preserve the maximum principle up to second-order accuracy in the temporal direction. Thus the calculation efficiency is reduced. In this section, by using the finite difference method in the spatial direction and the IFRK method in the temporal direction, we propose a class of high-order scheme for the conservative AC equation. It is proved that the full-discrete scheme allows the AC equation to conserve mass and admit maximum principle. Before discretizing the conservative AC equation, let us introduce some preliminaries.

Lemma 3.1 ([10]). Under Assumption 2.1, when \( u(x,t) \in [-\beta,\beta] \), \( x \in \Omega \), we have

\[
f(\beta) \leq \lambda(t) \leq f(-\beta),
\]

where integral term \( \lambda(t) \) is independent of \( x \).

Then, set the constant \( \bar{\tau} \) as

\[
\bar{\tau} \leq \left( \max_{|\xi| \leq \beta} |f'(\xi)| \right)^{-1}. \tag{3.2}
\]

Lemma 3.2. Under Assumption 2.1 and the choice of a constant \( \bar{\tau} \), we have

\[
\|\xi + \bar{\tau} f(\xi)\|_{L^\infty} \leq \beta, \quad \bar{\tau} \leq \left( \max_{|\xi| \leq \beta} |f'(\xi)| \right)^{-1}, \tag{3.3}
\]

for any \( \xi(x) \in C(\Omega) \) with \( \|\xi\|_{L^\infty} \leq \beta \).

Proof. Above all, let’s define

\[
g(\xi) = \frac{1}{\bar{\tau}} \xi(x) + f(\xi).
\]

It can be deduced that for any \( \xi(x) \in C(\Omega) \) with \( \|\xi\|_{L^\infty} \leq \beta \), it holds that

\[
0 \leq g'(\xi) = \frac{1}{\bar{\tau}} + f'(\xi(x)) \leq \frac{2}{\bar{\tau}}, \quad \forall x \in \Omega,
\]

which indicates that the function \( g(\xi) \) is a monotonically nondecreasing function. Thus,

\[
g(-\beta) \leq g(\xi) \leq g(\beta), \tag{3.5}
\]

which is equivalent to

\[
-\frac{\beta}{\bar{\tau}} + f(-\beta) \leq \frac{1}{\bar{\tau}} \xi(x) + f(\xi) \leq \frac{\beta}{\bar{\tau}} + f(\beta), \quad \forall x \in \Omega. \tag{3.6}
\]
From Lemma 3.1, we have
\[
-f(-\beta) \leq -\lambda(t) \leq -f(\beta).
\] (3.7)

Therefore,
\[
-\frac{\beta}{\bar{\tau}} \leq -\frac{\beta}{\bar{\tau}} + f(-\beta) - \lambda(t) \leq N(\xi) \leq -\frac{\beta}{\bar{\tau}} + f(\beta) - \lambda(t) \leq \frac{\beta}{\bar{\tau}},
\] (3.8)

where
\[
N(\xi) = \frac{1}{\bar{\tau}} \xi(x) + f(\xi) - \lambda(t) = \frac{1}{\bar{\tau}} \xi(x) + \xi(x) - \xi(x)^3 - \lambda(t),
\]

so that
\[
\left\| \frac{1}{\bar{\tau}} \xi(x) + \xi(x) - \xi(x)^3 - \lambda(t) \right\|_{L^\infty} \leq \frac{\beta}{\bar{\tau}}.
\] (3.9)

Then
\[
\|\xi(x) + \bar{\tau} f(\xi)\|_{L^\infty} \leq \beta,
\] (3.10)

which is equivalent to
\[
\|\xi(x) + \bar{\tau} f(\xi)\|_{L^\infty} \leq \beta,
\] (3.11)

where
\[
\bar{\tau} \leq \left( \max_{|\xi| \leq \beta} |f'(\xi)| \right)^{-1}.
\]

This completes the proof. \qed

3.1 Finite difference semi-discretization

Let the mesh size \( h = \frac{b-a}{N} \), the grid points \( \Omega_h = \{ x_j | x_j = a + jh, j = 0, 1, \cdots, N-1 \} \), \( V_N = \{ v \in V_h, x_j \in \Omega_h \} \subset \mathbb{R}^N \), equipped with discrete \( L^2 \) inner product, \( L^2 \) norm, and \( L^\infty \) norm which is defined as
\[
\langle u, v \rangle_h = h \sum_{j=0}^{N-1} u_j v_j, \quad \| u \|^2_h = \langle u, u \rangle_h, \quad \| u \|_{L^\infty} = \max_{i=0, \cdots, N-1} |u_i|, \quad \forall u, v \in V_N.
\] (3.12)

Then, the central finite difference discretization of \( \partial_{xx} \) is denoted as
\[
\begin{bmatrix}
-2 & 1 & & & \\
1 & -2 & 1 & & \\
& & \ddots & \ddots & \\
& & 1 & -2 & 1 \\
1 & & & 1 & -2
\end{bmatrix}_{N \times N}
\]

Define \( L = \epsilon^2 D_2, I \in \mathbb{R}^{N \times N} \) is the identity matrix and \( 1 = [1, 1, \cdots, 1]^T \in \mathbb{R}^N \).
Lemma 3.3 ([39]). For any \( \tau > 0 \), it holds that \( \| e^{\tau L} \|_{\infty} = 1 \).

Lemma 3.4. For the matrix \( L \), the following equality holds:

\[
\langle e^{\tau L} u^n, 1 \rangle = \langle u^n, 1 \rangle, \quad \forall \tau \geq 0.
\] (3.13)

Proof. Using the symmetry of \( L \) that

\[
\langle e^{\tau L} u^n, 1 \rangle = 1^T e^{\tau L} u^n = 1^T (e^{\tau L})^T u^n = \langle u^n, e^{\tau L} 1 \rangle = \langle u^n, 1 \rangle, \quad \forall \tau \geq 0.
\] (3.14)

This completes the proof. \( \square \)

Then, the spatial semi-discretization in 1D is expressed as

\[
u_t = \epsilon_2 D_2 u + \bar{f}(u), \quad x \in \Omega_h, \quad t > 0.
\] (3.15)

Theorem 3.1. The semi-discrete conservative AC Eq. (3.15) preserves the mass conservation law, i.e.,

\[
d \frac{d}{dt} M = 0, \quad \forall t > 0,
\] (3.16)

where \( M := \langle u, 1 \rangle \).

Proof. By taking the \( L^2 \) inner product with 1 on both side of the Eq. (3.15) and using the periodic boundary condition, the semi-discrete mass conservation law is obtained

\[
d \frac{d}{dt} \langle u, 1 \rangle = \langle \epsilon^2 D_2 u, 1 \rangle + \langle \bar{f}(u), 1 \rangle,
\] (3.17)

where

\[
\langle \epsilon^2 D_2 u, 1 \rangle = 0
\] (3.18)

can be deduced based on the periodic boundary conditions. Meanwhile,

\[
\langle \bar{f}(u), 1 \rangle = 0,
\] (3.19)

so we can obtain

\[
d \frac{d}{dt} \langle u, 1 \rangle = 0,
\] (3.20)

which is equivalent to

\[
d \frac{d}{dt} M = 0, \quad \forall t > 0.
\]

It shows that the semi-discrete conservative AC Eq. (3.15) satisfies the mass conservation law. \( \square \)

Theorem 3.2 ([10]). It shows that for the semi-discrete conservative AC Eq. (3.15) with the initial value \( u_0 = u_0(x) \) and \( \| u_0 \|_{\infty} \leq \beta \), the solution \( u \) to the semi-discrete system (3.15) satisfies the maximum principle

\[
\| u(t) \|_{\infty} \leq \beta, \quad \forall t \geq 0.
\] (3.21)
3.2 Integrating factor Runge-Kutta time integration

Consider the Lawson transformation of unknown

\[ v(t) = e^{-Lt} u(t). \]  

Through this transformation, the equivalent semi-discrete system can be obtained if and only if \( v \) solves

\[ v_t = e^{-Lt} \tilde{f}(e^{-Lt} v), \quad \forall t \geq 0, \]  

where \( v(0) = u(0) = u_0 \). Think about the \( s \)-stage and \( p \)-th order explicit RK scheme defined by the Butcher table

\[
\begin{array}{c|cccc}
  c & A & b \\
  \hline
  0 & a_{1,0} & 0 & \cdots \\
  \vdots & \vdots & \vdots & \ddots \\
  \vdots & \vdots & \vdots & \ddots \\
  c_{s-1} & a_{s-1,0} & \cdots & 0 \\
  \hline
  b_0 & b_1 & \cdots & b_{s-1}
\end{array}
\]

where

\[ a_{i,j} = 0, \quad (i \leq j), \quad \sum_{i=0}^{s-1} b_i = 1, \quad c_i = \sum_{i=0}^{s-1} a_{i,j}, \quad (i = 0, 1, \cdots, s-1), \]

and the Butcher coefficients are constrained by certain accuracy and stability requirements. Then, applying the RK method to problem (3.23) yields the IFRK scheme:

\[ u_{n,i} = e^{c_i \tau L} u^n + \tau \sum_{j=0}^{i-1} a_{i,j} e^{(c_i - c_j) \tau L} \tilde{f}_{n,j}, \quad i = 0, 1, \cdots, s-1, \]

\[ u^{n+1} = e^{\tau L} u^n + \tau \sum_{i=0}^{s-1} b_i e^{(1-c_i) \tau L} \tilde{f}_{n,i}, \]

with \( u_{n,0} = u^n, \quad \tilde{f}_{n,i} = \tilde{f}(u_{n,i}). \)

**Theorem 3.3.** The full-discrete conservative AC Eq. (3.24b) preserves the mass conservation law, i.e.,

\[ M_t = M = \cdots = M^0, \quad \forall t \geq 0, \]  

where \( M = \langle u', 1 \rangle \) is the discrete global mass at time \( t \) with \( i = 1, \cdots, s. \)

**Proof.** Above all, the properties of \( \tilde{f}_{n,i} \) holds

\[ \langle \tilde{f}_{n,i}, 1 \rangle = 0. \]
After that, taking $l^2$ inner product with 1 on both sides of the Eq. (3.24a) and combining with Lemma 3.4, we have

$$\langle u_n, i_1, 1 \rangle = \langle e^{c_i \tau L} u^n, 1 \rangle + \tau \sum_{j=0}^{i-1} d_{ij} e^{(c_i-c_j) \tau L} \tilde{f}_{n,j}, 1 \rangle$$

$$= \langle u^n, 1 \rangle + \tau \sum_{j=0}^{i-1} d_{ij} \tilde{f}_{n,j}, 1 \rangle$$

$$= \langle u^n, 1 \rangle, \quad i = 1, \cdots, s. \quad (3.27)$$

Thus,

$$M_{n+1} = M_n = \cdots = M_0, \quad \forall t \geq 0.$$

This completes the proof.

Next, we will prove that the discrete maximum principle holds for a special class of IFRK schemes under a certain time step constraint. For

$$\alpha_{i,j} \geq 0, \quad \sum_{j=0}^{i-1} a_{i,j} = 1, \quad \forall i = 1, \cdots, s,$$

we present the IFRK scheme in the Shu-Osher form [42]

$$\begin{cases}
    u_{n,0} = u^n, \\
    u_{n,j} = \sum_{j=0}^{i-1} e^{(c_i-c_j) \tau L} (\alpha_{i,j} u_{n,j} + \beta_{i,j} \tau \tilde{f}_{n,j}), \quad i = 1, \cdots, s, \\
    u^{n+1} = u_{n,s},
\end{cases} \quad (3.28)$$

where

$$a_{s,j} = b_i, \quad j = 0, \cdots, s-1, \quad (3.29a)$$

$$\beta_{i,j} = a_{i,j} - \sum_{k=j+1}^{i-1} \alpha_{i,k} a_{k,j}, \quad i = 1, \cdots, s, \quad j = 0, \cdots, i-1. \quad (3.29b)$$

**Theorem 3.4.** In fact, when the coefficients $\alpha_{i,j}$ and $\beta_{i,j}$ satisfy

$$\alpha_{i,j} \geq 0, \quad \sum_{j=0}^{i-1} \alpha_{i,j} = 1, \quad (3.30a)$$

$$\beta_{i,j} \geq 0, \quad \beta_{i,j} = 0 \quad \text{if} \quad \alpha_{i,j} = 0, \quad (3.30b)$$
with the monotonic increasing abscissas \(c_{i,j}\), i.e., \(0 = c_0 \leq c_1 \leq \cdots \leq c_n = 1\), this class of IFRK method is also called the SSP-IFRK method, which satisfies the MPP under the SSP condition on the time step size \([36]\). So if the explicit full-discrete IFRK system in the Shu-Osher form (3.28) with the initial value satisfying \(\|u^0\|_{\infty} \leq \beta\), then the numerical solution \(u^n\) satisfies the maximum principle when the time step size \(\tau\) satisfies

\[
\tau \leq \min_{i,j} \left| \frac{\alpha_{i,j}}{\beta_{i,j}} \right| \bar{\tau}, \quad \bar{\tau} \leq \left( \max_{|\xi| \leq \beta} |f'(\xi)| \right)^{-1}
\]

(3.31) for \(i = 1, \cdots, s, j = 0, \cdots, i - 1\).

**Proof.** We prove this by mathematical induction. Assuming \(\|u^a\|_{\infty} \leq \beta\), we will show that \(\|u^{a+1}\|_{\infty} \leq \beta\). When \(\tau\) satisfies the condition (3.31), we can prove the following conclusion by using Lemma 3.2 and combining Eq. (3.13) with Eq. (3.26)

\[
\|u_{n+1}\|_{\infty} = \left\| \sum_{j=0}^{i-1} e^{(c_i - c_j)\tau L} (\alpha_{i,j} u_{n,j} + \beta_{i,j} \bar{f}_{n,j}) \right\|_{\infty}.
\]

\[
= \left\| \sum_{j=0}^{i-1} (\alpha_{i,j} u_{n,j} + \beta_{i,j} \bar{f}_{n,j}) \right\|_{\infty}
\]

\[
= \sum_{j=0}^{i-1} \left| \alpha_{i,j} \right| u_{n,j} + \left| \beta_{i,j} \right| \bar{f}_{n,j} \leq \beta, \quad i = 1, \cdots, s.
\]

(3.32)

This completes the proof. \(\Box\)

### 3.3 Error estimate

Consider the 1D problem, by using the definition of the \(p\)-th order \((1 \leq p \leq s)\) RK scheme, we can derive the following convergence result.

**Theorem 3.5.** Given \(T > 0\), assume that \(u(t) \in C^p[0, T]\) is the exact solution of the semi-discrete scheme (3.15) and \(u^n\) is the numerical solution of the full-discrete IFRK system (3.24a)-(3.24b), respectively. Suppose the initial value is smooth in \([0, T]\) and satisfies \(\|u_0\|_{\infty} \leq \beta\), under the condition of Theorem 3.4, the error estimate can be written as

\[
\|u(t_n) - u^n\|_{\infty} \leq C_1 (e^{A\tau t_n} - 1) \tau^p \quad \text{for} \quad t_n \leq T,
\]

(3.33)

for any \(\tau > 0\), where the constant \(C_1 > 0\) is independence of \(\tau\) and

\[A = \max_{|\xi| \leq \beta} |f'(\xi)|.\]
Proof. Denote the reference functions $U_{n,i}$, which possess $U_{n,0} = u(t_n)$ and $U_{n,s} = u(t_{n+1})$, thus we can obtain

$$
U_{n,i} = e^{c_i \tau L} u(t_n) + \tau \sum_{j=0}^{i-1} a_{ij} e^{(c_j-c_i) \tau L} f(U_{n,j}), \quad i = 0, 1, \ldots, s-1, \quad (3.34a)
$$

$$
U_{n,s} = e^{\tau L} u(t_n) + \tau \sum_{i=0}^{s-1} b_i e^{(1-c_i) \tau L} f(U_{n,i}) + R_s, \quad (3.34b)
$$

where the truncation error $R_s$ satisfies

$$
\|R_s\|_{l_\infty} \leq C_s \tau^{p+1},
$$

with the constant $C_s$ depends on the $C^p[0,T]$-norm of the $u$, the $C^p[-\beta,\beta]$-norm of $f$, $s$, $T$, $\|L\|_{l_\infty}$ and $p$, but is independent of $\tau$.

Let $e^n = u(t_n) - u^n$ and $e_{n,i} = U_{n,i} - u_{n,i}$ for $i = 0, 1, \ldots, s-1$, hence the following equations can be obtained by subtracting Eq. (3.24a) from Eq. (3.34a) and subtracting Eq. (3.24b) from Eq. (3.34b), respectively,

$$
e_{n,i} = e^{c_i \tau L} e^n + \tau \sum_{j=0}^{i-1} a_{ij} e^{(c_j-c_i) \tau L} (\bar{f}(U_{n,j}) - \bar{f}(u_{n,j})), \quad i = 0, 1, \ldots, s-1, \quad (3.35a)
$$

$$
e^{n+1} = e^{\tau L} e^n + \tau \sum_{i=0}^{s-1} b_i e^{(1-c_i) \tau L} (\bar{f}(U_{n,i}) - \bar{f}(u_{n,i})) + R_s. \quad (3.35b)
$$

By noting $a_{ij} \leq c_i \leq 1$ and using Lemma 3.3, we can derive

$$
\|e_{n,i}\|_{l_\infty} \leq \|e^n\|_{l_\infty} + \tau \sum_{j=0}^{i-1} \|\bar{f}(U_{n,j}) - \bar{f}(u_{n,j})\|_{l_\infty}
$$

$$
\leq \|e^n\|_{l_\infty} + \tau A \sum_{j=0}^{i-1} \|e_{n,j}\|_{l_\infty}
$$

$$
\leq \|e^n\|_{l_\infty} + \tau A \sum_{j=0}^{i-1} (1 + \tau A)^j \|e^n\|_{l_\infty}
$$

$$
= (1 + \tau A)^i \|e^n\|_{l_\infty}, \quad i = 0, 1, \ldots, s-1, \quad (3.36a)
$$

$$
\|e^{n+1}\|_{l_\infty} \leq \|e^n\|_{l_\infty} + \tau \sum_{i=0}^{s-1} \|\bar{f}(U_{n,i}) - \bar{f}(u_{n,i})\|_{l_\infty} + \|R_s\|_{l_\infty}
$$

$$
\leq \|e^n\|_{l_\infty} + \tau A \sum_{i=0}^{s-1} \|e_{n,i}\|_{l_\infty} + C_s \tau^{p+1}
$$

$$
\leq \|e^n\|_{l_\infty} + \tau A \sum_{i=0}^{s-1} (1 + \tau A)^i \|e^n\|_{l_\infty} + C_s \tau^{p+1}
$$

$$
= (1 + \tau A)^s \|e^n\|_{l_\infty} + C_s \tau^{p+1}. \quad (3.36b)
$$
By induction, it can be arranged as

\[ \| e^n \|_{l}\infty \leq (1 + \tau A)^{sn} \| e^0 \|_{l}\infty + C_s \tau^{p+1} \sum_{k=0}^{n-1} (1 + \tau A)^{sn}. \]

(3.37)

Thanks to \( \| e^0 \|_{l}\infty = 0 \), therefore we use the geometric sequence summation formula to calculate \( \sum_{k=0}^{n-1} (1 + \tau A)^{sn} \), and then utilize Taylor expansion to get

\[ \| e^n \|_{l}\infty \leq C_1 (e^{Ast} - 1) \tau^p, \]

(3.38)

where \( C_1 = C_s / As \), \( t_n = \tau n \). This completes the proof.

Noting that the second-order finite discretization for 2D and 3D Laplace operators can be directly obtained by using Kronecker products, and the resulting differentiation matrices satisfy Lemma 3.2 and 3.3. Thus the preservation of maximum principle, conservation of mass and error estimate can be similarly proved. To save space, we omit them.

4 Numerical simulations

In this section, we carry out some numerical experiments to demonstrate the performance of the IFRK scheme presented in Section 3. First, the convergence rates of the proposed full-discrete scheme (3.28) in both temporal and spatial directions are verified using a 1D example. Afterwards, we verify the preservation of maximum principle and conservation of mass by three examples. Finally, we present a 3D simulation example to show effectiveness of the proposed IFRK method. In this paper, the following RK schemes [36,42] are used

\[
\begin{align*}
\text{RK}(1,1): \quad & \begin{bmatrix} 0 & 0 \\ 1 \end{bmatrix}, \\
\text{RK}(2,2): \quad & \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 & 2 \end{bmatrix}, \\
\text{RK}(3,3): \quad & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 4 & 9 & 9 & 16 \end{bmatrix}, \\
\text{RK}(4,4): \quad & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}.
\end{align*}
\]

According to the condition (3.31), the maximum time steps for different IFRK methods to satisfy MPP are calculated, and the results are listed in Table 1.

In addition, the mass errors are computed by using

\[ \text{Mass Error} = |M^t - M^0|. \]
Table 1: The maximum time steps for different IFRK schemes.

<table>
<thead>
<tr>
<th>RK schemes</th>
<th>RK(1, 1)</th>
<th>RK(2, 2)</th>
<th>RK(3, 3)</th>
<th>RK(4, 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \min \frac{\tau}{\tau_{mp}} )</td>
<td>1</td>
<td>1</td>
<td>( \frac{3}{5} )</td>
<td>( \frac{7}{5} )</td>
</tr>
<tr>
<td>( \tau_{mp} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{5} )</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{7}{5} )</td>
</tr>
</tbody>
</table>

4.1 Example 1

Let us consider the following initial value

\[ u(x,0) = 0.05 \sin(x), \quad x \in [0, 2\pi), \quad t > 0. \tag{4.1} \]

First, by setting \( \epsilon = 0.01 \) and choosing \( T = 1, \) \( NX = 2^8, \tau = 2^{-10} \) as the reference solution, the convergence order of different IFRK methods listed in Table 2 shows that fourth-order accuracy can be obtained in the temporal direction through selecting different \( \tau \) values as \( dt, \frac{dt}{2}, \frac{dt}{4} \), where \( dt = 2^{-3} \). Then, choosing \( T = 1, \) \( NX = 2^{12}, \tau = 2^{-10} \) as the reference solution, the spatial precision is obtained and listed in Table 3. Besides, it can be seen from Table 3 that second-order accuracy in the spatial direction is maintained by using the finite difference method with the spatial grid from being refined from \( 2^6 \) to \( 2^8 \) uniformly.

Let \( T = 100 \) and set the time step \( \tau \) as \( \frac{2}{5} \) and 1.5. From left to right, Fig. 1 shows numerical solutions of \( u \), evolutions of infinite norms \( \| u \|_\infty \), and mass errors by using RK(1, 1), RK(2, 2), RK(3, 3) and RK(4, 4), respectively. It can be seen from the left column of Fig. 1 that the numerical solutions are relatively stable when \( \tau \) takes a small time step as \( \frac{2}{5} \). However, when \( \tau \) takes a time step of 1.5 which is larger than \( \tau_{mp} \), solutions obtained

Table 2: Temporal accuracy test of the IFRK, \( dt = 2^{-3} \).

<table>
<thead>
<tr>
<th>( dt )</th>
<th>IFRK(1, 1)</th>
<th>IFRK(2, 2)</th>
<th>IFRK(3, 3)</th>
<th>IFRK(4, 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( dt/2 )</td>
<td>5.216e-03</td>
<td>7.295e-03</td>
<td>2.194e-04</td>
<td>3.064e-04</td>
</tr>
<tr>
<td>( dt/4 )</td>
<td>2.743e-03</td>
<td>3.833e-03</td>
<td>1.967e-03</td>
<td>8.031e-05</td>
</tr>
<tr>
<td>( \frac{dt}{2} )</td>
<td>1.408e-03</td>
<td>9.037e-03</td>
<td>1.472e-05</td>
<td>2.056e-05</td>
</tr>
<tr>
<td>( \frac{dt}{4} )</td>
<td>6.560e-06</td>
<td>1.967e-03</td>
<td>1.472e-05</td>
<td>2.056e-05</td>
</tr>
<tr>
<td>( dt/2 )</td>
<td>8.614e-07</td>
<td>2.929</td>
<td>1.121e-08</td>
<td>1.568e-08</td>
</tr>
<tr>
<td>( dt/4 )</td>
<td>1.397e-08</td>
<td>2.964</td>
<td>7.158e-10</td>
<td>3.970</td>
</tr>
</tbody>
</table>

Table 3: Spatial accuracy test of the finite difference method.

| \( NX \) | RK Order | \( L^2 \) Error | Order | \( L^\infty \) Error | Order | RK Order | \( L^2 \) Error | Order | \( L^\infty \) Error | Order | RK Order | \( L^2 \) Error | Order | \( L^\infty \) Error | Order |
|----------|----------|----------------|-------|---------------------|-------|----------|----------------|-------|---------------------|-------|----------|----------------|-------|---------------------|-------|----------|----------------|
| \( 2^6 \) | IFRK(1, 1) | 1.078e-08 | - | 1.013e-08 | - | IFRK(2, 2) | 1.078e-08 | - | 1.013e-08 | - |
| \( 2^7 \) | 2.694e-09 | 2.001 | 2.532e-09 | 2.001 | 2.694e-09 | 2.001 | 2.532e-09 | 2.001 |
| \( 2^8 \) | 6.716e-10 | 2.004 | 6.311e-10 | 2.004 | 6.716e-10 | 2.004 | 6.311e-10 | 2.004 |
| \( 2^6 \) | IFRK(3, 3) | 1.078e-08 | - | 1.013e-08 | - | IFRK(4, 4) | 1.078e-08 | - | 1.013e-08 | - |
| \( 2^7 \) | 2.694e-09 | 2.001 | 2.532e-09 | 2.001 | 2.694e-09 | 2.001 | 2.532e-09 | 2.001 |
| \( 2^8 \) | 6.716e-10 | 2.004 | 6.311e-10 | 2.004 | 6.716e-10 | 2.004 | 6.311e-10 | 2.004 |
Figure 1: Solutions of $u$ (left column), evolutions of $|u|_{L^\infty}$ (middle column) and mass errors (right column) by using different RK methods. Parameters: $\epsilon = 0.01$, $NX = 128$, $T = 100$. 
by RK(1, 1), RK(2, 2), and RK(3, 3) show a large oscillation, while the solution obtained by RK(4, 4) has better stability. Meanwhile, the middle column of Fig. 1 shows that when \( \tau \) is set as \( \frac{2}{9} \), all the RK methods achieve MPP; when \( \tau \) is set as 1.5, neither RK(1, 1) nor RK(3, 3) can achieve MPP. Although RK(2, 2) can satisfy MPP, it fluctuates greatly. By contrast, RK(4, 4) also satisfies MPP, but it is more stable. So, it can be seen from the right column of Fig. 1 that regardless of the size of the time step and the selection of the RK method, the mass error can reach machine accuracy, which verifies the conservation of mass. Hence, the fourth-order IFRK scheme can be considered as a good attempt to preserve the MPP and mass conservation law.

4.2 Example 2

In two dimensions, the following initial value is set for the conservative AC equation

\[
u(x, y, 0) = \cos(2\pi x)\cos(2\pi y), \quad (x, y) \in [-0.5, 0.5]^2, \quad t > 0.
\]  

(4.2)

Let \( \epsilon = 0.01 \) and \( NX = NY = 128 \). It can be seen from Fig. 2 that the solution is mixed

\[\text{Figure 2: Solutions computed by using RK(4, 4) scheme at } T = 0, 100, 200, 600. \text{ Parameters: } \epsilon = 0.01, NX = NY = 128, \tau = \frac{2}{9}.\]
together at $T = 0$, and the numerical solution becomes clearer over the time. Finally, the solution reaches a stable state when $T = 600$. Evolutions of $\|u\|_\infty$ and mass errors under different time steps by employing RK(1, 1), RK(2, 2), RK(3, 3), and RK(4, 4) are demonstrated in Fig. 3 and Fig. 4, respectively. The Fig. 3 presents evolutions of $\|u\|_\infty$ under different time steps. It can be seen that when $\tau \leq \frac{2}{9}$, the scheme satisfies MPP. However, it is clear that RK(1, 1) cannot achieve MPP when the time step increases to 1.5. Even though infinite norms of $\|u\|_\infty$ by using RK(2, 2) and RK(3, 3) remain within the black line, the infinite norm of RK(3, 3) oscillates significantly. In comparison, RK(4, 4) is a more stable scheme that satisfies MPP, but the the infinite norm is not as good as that under small time steps. Therefore, it is verified in this example that the RK(4, 4) scheme is the most stable scheme that can preserve the maximum value, and the MPP is achieved at certain step sizes. Besides, the Fig. 4 shows the mass conservation property can be preserved with different time steps. Thus, it can be seen that the stability of mass is not affected by time steps. That is, the full-discrete scheme of the constructed conservative
AC equation can satisfies the conservation law of mass.

4.3 Example 3

We consider the following three-dimensional initial value

\[
\begin{align*}
    u(x,y,z,0) &= \cos(2\pi x) \cos(4\pi y) \cos(6\pi z), \quad (x,y,z) \in [0,1]^3, \quad t > 0.
\end{align*}
\]  

Set \( \epsilon = 0.05, \, NX = NY = NZ = 64 \). Numerical solutions of \( u \) and mass errors by employing RK(4, 4) under different time steps for a long period of time are presented in Fig. 5 and Fig. 6, respectively. First, the snapshot of the numerical solution shown in Fig. 5 is blurry at first, and the solution are mixed. Then, the solution becomes clearly separated until \( T = 100 \). As shown in Fig. 6, the evolution of \( \|u\|_{L^\infty} \) within the range of the black line is almost smooth, and the mass error still reaches the machine precision. As a result, the proposed scheme by using RK(4, 4) for the conservative AC equation can describe the
Figure 5: Solutions computed by using RK(4, 4) scheme at $T=0, 50, 80, 100$. Parameters: $\epsilon=0.05$, $N_X=N_Y=N_Z=64$, $\tau=\frac{1}{5}$.

evolution of the numerical solution well. Besides, it can achieve MPP and conserve mass.

5 Conclusions

In this paper, we have developed a stable and high order MPP scheme for solving the conservative AC equation. The mass-conserving AC equation is constructed by introducing a Lagrange multiplier. And then, the finite difference method and the IFRK method are respectively applied in the spatial direction and the temporal direction. The proposed scheme (3.28) has three main advantages. Above all, it is well-known that the explicit method enjoys less computation for solving partial differential equations, so the new scheme (3.28) has high calculation efficiency due to the application of the explicit IFRK method. The second point, the proposed scheme of the conservative AC equation can sat-
satisfy MPP and conserve mass simultaneously, and it can be calculated steadily over a long period. Last but not least, the proposed scheme can reach up to fourth-order accuracy in the temporal direction. Moreover, the theoretical analysis result is verified by 1D, 2D and 3D experiments. Compared with different IFRK schemes, the fourth-order scheme can be generalized to construct a higher order mass conservative scheme steadily.

An immediate future work could be consider the Allen-Cahn equation with Peng-Robinson equation of state. This problem was first studied in [43] and has been recently studied by Huang [44]. So we want to apply IFRK method to this equation to obtain the high order mass-preserving and MPP scheme.

Acknowledgements

The work is supported by the National Key R&D Program of China (No. 2020YFA070980 0), the National Key Project (No. GJXM92579), the National Natural Science Foundation of China (No. 12071481).

References


[43] Z. H. Qiao and S. Y. Sun, Two-phase fluid simulation using a diffuse interface model with Peng-