ESAIM: M2AN 58 (2024) 421–455
https://doi.org/10.1051/m2an/2024008

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

AN EFFICIENT TWO-GRID HIGH-ORDER COMPACT DIFFERENCE SCHEME WITH VARIABLE-STEP BDF2 METHOD FOR THE SEMILINEAR PARABOLIC EQUATION

Bingyin Zhang[1] and Hongfei Fu[1,2,*]

Abstract. Due to the lack of corresponding analysis on appropriate mapping operator between two grids, high-order two-grid difference algorithms are rarely studied. In this paper, we firstly discuss the boundedness of a local bi-cubic Lagrange interpolation operator. And then, taking the semilinear parabolic equation as an example, we first construct a variable-step high-order nonlinear difference algorithm using compact difference technique in space and the second-order backward differentiation formula with variable temporal stepsize in time. With the help of discrete orthogonal convolution kernels, temporal-spatial error splitting idea and a cut-off numerical technique, the unique solvability, maximum-norm stability and corresponding error estimate of the high-order nonlinear difference scheme are established under assumption that the temporal stepsize ratio satisfies $r_k := \tau_k/\tau_{k-1} < 4.8645$. Then, an efficient two-grid high-order difference algorithm is developed by combining a small-scale variable-step high-order nonlinear difference algorithm on the coarse grid and a large-scale variable-step high-order linearized difference algorithm on the fine grid, in which the constructed piecewise bi-cubic Lagrange interpolation mapping operator is adopted to project the coarse-grid solution to the fine grid. Under the same temporal stepsize ratio restriction $r_k < 4.8645$ on the variable temporal stepsize, unconditional and optimal fourth-order in space and second-order in time maximum-norm error estimates of the two-grid difference scheme is established. Finally, several numerical experiments are carried out to demonstrate the effectiveness and efficiency of the proposed scheme.

Mathematics Subject Classification. 65M06, 65M12, 65M15, 65M50.

Received June 23, 2023. Accepted January 28, 2024.

1. Introduction

It is well known that the analytical solutions of nonlinear parabolic PDEs arising from a variety of physical and engineering applications are not available in most cases. Thus, numerous efforts have been devoted to the development of efficient numerical schemes, see [3, 9, 10, 17, 24, 36]. Generally speaking, fully-implicit numerical schemes are usually proved to be unconditionally stable. Unfortunately, at each time step, one has to solve a system of nonlinear equations [17, 21], in which an extra iterative process must be imposed, and this in

Keywords and phrases. High-order two-grid difference algorithm, variable-step BDF2 method, piecewise bi-cubic Lagrange interpolation, DOC kernels, unique solvability, maximum-norm error estimate.

[1] School of Mathematical Sciences, Ocean University of China, Qingdao, Shandong 266100, P.R. China.
[2] Laboratory of Marine Mathematics, Ocean University of China, Qingdao, Shandong 266100, P.R. China.
*Corresponding author: fhf@ouc.edu.cn

turn may cause severe computational costs. Instead, a very popular and widely-used approach is the so-called implicit-explicit scheme, which treats the linear term implicitly and the nonlinear term explicitly. However, if the corresponding globally continuous condition of the nonlinear term (*e.g.*, $|f''(w)| \leq K_f$ for $w \in \mathbb{R}$) cannot be imposed or the boundedness of numerical solution in L^∞ norm cannot be obtained, this method usually suffers from a very restrictive temporal stepsize condition caused by use of inverse inequality for convergence, *e.g.*, $\tau = \mathcal{O}(h^{d/2p})$, where h is the spatial mesh size, d is the space dimension and p is the accuracy of the temporal discretization. Therefore, such restrictions may lead to the use of a small temporal stepsize, and thus much computational time may be consumed. Here we refer to [2, 3, 43, 51] for an incomplete list of references.

Another efficient and powerful strategy is the two-grid method which is proposed by Xu *et al.* [36, 48]. The basic idea of this kind of method is to reduce the solution of a large-scale nonlinear problem on the fine grid to a small-scale nonlinear problem on the coarse grid plus a large-scale linear problem on the fine grid. Hence, basically it includes two solution steps: First, one solve a nonlinear system on the coarse grid to obtain a rough approximation, and then solve a linearized system resulting from the rough solution to derive a corrected solution on the fine grid. Up to now, this technique has been widely applied to numerically solve many types of nonlinear PDEs, *e.g.*, [9, 18, 46] for parabolic equations, [34, 39] for Darcy-Forchheimer equations, [14] for Navier-Stokes equations, [13, 22, 25] for time-fractional equations and [7, 49] for other nonlinear equations. Compared to traditional implicit-explicit scheme, the main advantages of the two-grid scheme are twofold: (i) if the boundedness of numerical solution in L^∞ norm is not obtained, the temporal stepsize restriction of the method under the local Lipschitz continuous condition on nonlinear term usually is only related to H instead of h, which is much more weaker, see [9, 49]; (ii) the two-grid method which treats the nonlinearity on the coarse grid and solving the linear system on the fine grid [36], is much more stable and accurate than the implicit-explicit one when solving the nonlinear PDEs whose solutions change rapidly with respect to time. In this paper, we will investigate the numerical stability of these two methods for general semilinear parabolic PDEs by carrying out representative numerical examples.

Unlike the finite element method, which generates pointwise solution in space and thus is easy to develop two-grid finite element method [6, 14, 48], the solution yielded by finite difference method is only on discrete grid points, and therefore, an appropriate accuracy-preserving mapping operator from the coarse-grid function space to the fine-grid function space is required to construct and analyze the two-grid difference method, *e.g.*, piecewise linear/bilinear interpolation for second-order two-grid difference schemes [9, 22, 39, 49, 50]. However, due to the lack of corresponding analysis on appropriate high-order mapping operator, the work about high-order two-grid difference method is meager and, in fact, numerical analysis is also lack. This motivates us to develop high-order two-grid difference scheme for semilinear parabolic PDEs with general boundary condition, *e.g.*, Dirichlet or periodic boundary condition, by introducing and analyzing appropriate mapping operator.

In this paper, to illustrate the application of the proposed high-order two-grid difference method, we focus on the following semilinear parabolic equation

$$u_t(x,y,t) - c\Delta u(x,y,t) = f(u(x,y,t)) + g(x,y,t), \quad (x,y) \in \Omega,\ t \in (0,T], \tag{1.1}$$

where $\Omega = (0, L_x) \times (0, L_y)$, subject to the initial condition

$$u(x,y,0) = \varphi(x,y), \quad (x,y) \in \Omega, \tag{1.2}$$

and periodic boundary condition or Dirichlet boundary condition

$$u(x,y,t) = \psi(x,y,t), \quad (x,y) \in \partial\Omega,\ t \in [0,T], \tag{1.3}$$

where $c > 0$ and $\Delta = \partial_x^2 + \partial_y^2$ is the Laplacian operator, $f(u)$ is the nonlinear term, φ, ψ and g are given smooth functions.

For many time-dependent PDEs, *e.g.*, Allen-Cahn equation [30], whose solution admitting multiple time scales, adaptive temporal stepsize strategies [23, 37] are heuristic and available approaches to improve accuracy and/or efficiency. Due to its strong stability, variable-step BDF2 method is practically valuable for stiff or

differential-algebraic problems [15, 40]. However, compared to those one-step methods, such as the backward Euler and Crank-Nicolson schemes, numerical analysis of the nonuniform BDF2 method would be challenging. In particular, for a linear parabolic problem, [5] proved that, if $0 < r_k := \tau_k/\tau_{k-1} \leq 1.868$ with $\tau_k := t_k - t_{k-1}$ the kth temporal stepsize, the variable-step BDF2 scheme is zero-stable and second-order convergence containing a prefactor $\exp(C\Gamma_n)$, where $\Gamma_n := \sum_k^{n-2}[r_k - r_{k+2}]_+$ with $[x]_+$ the positive part of x. Recently, by introducing a generalized discrete Grönwall inequality, Chen *et al.* [8] circumvented such a prefactor in error analysis under a little stronger stepsize ratio restriction $0 < r_k \leq 1.53$. In [45], the authors developed an implicit-explicit BDF2 method with variable stepsize for the parabolic partial integro-differential equations and proved its stability and convergence with $0 < r_k \leq 1.91$. The authors in [30] considered the fully-implicit BDF2 scheme for the Allen–Cahn equation and established the maximum-norm stability under $r_k < 1 + \sqrt{2}$ by developing a novel kernel recombination and complementary technique. To analyze the variable-step BDF2 scheme for linear reaction-diffusion equations, Liao and Zhang [28] introduced a new concept, namely, discrete orthogonal convolution (DOC) kernels, and they improved the unconditional stability in the L^2 norm to $r_k \leq 3.561$. Subsequently, with the help of DOC kernels and corresponding convolution inequalities, there is a great progress on the stability and error estimates of variable-step BDF2 method for nonlinear PDEs under $r_k < 3.561$ [31, 33, 42] and the further improved stepsize ratio restriction $r_k < 4.8645$ [11, 12, 26, 32], respectively.

Among all the variable-step BDF2 methods for nonlinear PDEs in the literature mentioned above, they treat the nonlinear terms fully or partially implicit, in which a nonlinear iteration must be implemented at each time step. Very recently, the authors of [19] proposed a linear second-order difference method for solving the Allen–Cahn equation, and they derived the L^∞ error estimate of the proposed scheme under $r_k < 1+\sqrt{2}$. In [52], Zhao *et al.* presented a linearized variable-step BDF2 scheme for solving nonlinear parabolic equation, and they proved the unconditional error estimate under $r_k < 4.8645$ and the maximum temporal stepsize $\tau \leq C\frac{1}{\sqrt{N}}$ by adopting the error splitting approach. After then, Li *et al.* extended this method to solve a nonlinear Ginzburg-Landau equation [44] and coupled Ginzburg-Landau equations [27] under the same conditions. In [35], a positivity-preserving and energy stable BDF2 scheme with variable stepsize was developed for the Cahn–Hilliard equation with nonlinear logarithmic potential, and convergence analysis in L^2 norm was established under $\tau \leq Ch$. For solving gradient flow problems, Hou and Qiao [16] proposed an unconditionally energy stable implicit-explicit BDF2 scheme with variable temporal stepsize using the SAV method, and derived the error estimates under the mild restriction on the adjacent temporal stepsize ratio $r_k < 4.8645$. Our goal is to construct and analyze an efficient high-order two-grid difference algorithm with nonuniform BDF2 method for the nonlinear parabolic equations (1.1)–(1.3). Compared to the existing literature, our contributions are mainly fourfold:

- An efficient variable-step two-grid fourth-order compact difference method is proposed, by using compact difference scheme, two-grid method, variable-step BDF2 formula as well as a developed piecewise bi-cubic Lagrange interpolation operator.
- We first prove the unique solvability, unconditional maximum-norm stability and error estimate for the nonlinear variable-step compact difference scheme under the temporal stepsize ratio restriction $r_k < 4.8645$, where only local continuous condition imposed on the nonlinear reaction term, see (3.17). The main numerical techniques including the applications of DOC kernels and an error splitting approach for a cut-off auxiliary scheme, which avoids the use of inverse estimate for the maximum-norm error and thus conquers the limitation on the maximum temporal stepsize, *i.e.*, no limited condition is required for the temporal-spatial stepsize ratio.
- Furthermore, by discussing the boundedness of the proposed piecewise bi-cubic Lagrange interpolation operator with periodic or Dirichlet boundary condition, unconditional and optimal-order maximum-norm error estimate of the variable-step BDF2 two-grid compact difference scheme is established under optimal coarse-fine grid condition $H = \mathcal{O}(h^{1/2})$.
- Several numerical experiments are presented to illustrate the effectiveness and efficiency of the proposed variable-step (adaptive) two-grid compact difference method, and comparisons of computational efficiency and stability with other schemes are also tested.

The remainder of this paper is organized as follows. In Section 2, we introduce and analyze the high-order mapping operator between the coarse-grid function space and fine-grid function space. In Section 3, we first propose a nonlinear compact difference scheme with variable-step BDF2 method for the semilinear parabolic equation subject to Dirichlet boundary condition, and then rigorously prove the unique solvability, unconditional stability and convergence, based on which we construct an efficient variable-step two-grid compact difference scheme in Section 4, and optimal-order maximum-norm error analysis is derived unconditionally. Moreover, in Section 5, the developed methods and techniques are extended to the context of periodic boundary condition. Several numerical experiments are presented to demonstrate the accuracy and efficiency of the proposed method in Section 6. Finally, some concluding remarks are drawn in the last section. Throughout this paper, we use C to denote a generic positive constant that may depend on the given data, but is independent of the mesh parameters.

2. High-order mapping operator between two grids

In this section, we first propose and analyze a high-order mapping operator between two grids based on Lagrange interpolation, which plays a significant role in the construction and numerical analysis of high-order two-grid difference method in the subsequent sections.

2.1. Some notations

Given two positive integers N_x^H and N_y^H, we define a uniform coarse grid $(x_i^H, y_j^H) := (iH_x, jH_y)$ for $0 \le i \le N_x^H$ and $0 \le j \le N_y^H$, with corresponding coarse mesh sizes $H_x := L_x/N_x^H$ and $H_y := L_y/N_y^H$. Moreover, for fixed positive integers $M_x, M_y \ge 2$, denote $N_x^h := M_x N_x^H$, $N_y^h := M_y N_y^H$ and define a uniform fine grid $(x_i^h, y_j^h) := (ih_x, jh_y)$ for $0 \le i \le N_x^h$ and $0 \le j \le N_y^h$, with corresponding fine mesh sizes $h_x := H_x/M_x$ and $h_y := H_y/M_y$. Denote $H := \max\{H_x, H_y\}$ and $h := \max\{h_x, h_y\}$.

Let $\bar{\omega}_\kappa := \left\{(i,j) \mid 0 \le i \le N_x^\kappa,\ 0 \le j \le N_y^\kappa\right\}$, $\omega_\kappa := \bar{\omega}_\kappa \cap \Omega$ and $\partial\omega_\kappa := \bar{\omega}_\kappa \cap \partial\Omega$ denote the sets of spatial grids, where $\kappa = H$ or h. Accordingly, we define the following discrete spaces of grid functions

$$\mathcal{V}_\kappa = \{v = \{v_{i,j}\} \mid (i,j) \in \bar{\omega}_\kappa\} \quad \text{and} \quad \mathcal{V}_\kappa^0 = \{v \mid v \in \mathcal{V}_\kappa \ \text{ and } \ v_{i,j} = 0 \text{ if } (i,j) \in \partial\omega_\kappa\}.$$

For any $w, q \in \mathcal{V}_\kappa$, we introduce the following notations

$$d_{\kappa,x} w_{i+\frac{1}{2},j} := \frac{w_{i+1,j} - w_{i,j}}{\kappa_x}, \quad d_{\kappa,x}^2 w_{i,j} := \frac{d_{\kappa,x} w_{i+\frac{1}{2},j} - d_{\kappa,x} w_{i-\frac{1}{2},j}}{\kappa_x},$$

$$\mathcal{A}_{\kappa,x} w_{i,j} := \begin{cases} w_{i,j} + \dfrac{\kappa_x^2}{12} d_{\kappa,x}^2 w_{i,j} = \dfrac{1}{12}(w_{i-1,j} + 10 w_{i,j} + w_{i+1,j}), & 1 \le i \le N_x^\kappa - 1, \\ w_{i,j}, & i = 0, N_x^\kappa. \end{cases}$$

Similarly, the notations $d_{\kappa,y} w_{i,j+\frac{1}{2}}$, $d_{\kappa,y}^2 w_{i,j}$ and $\mathcal{A}_{\kappa,y} w_{i,j}$ can be defined. Furthermore, we denote $\Delta_\kappa := d_{\kappa,x}^2 + d_{\kappa,y}^2$, $\Lambda_\kappa := \mathcal{A}_{\kappa,x} d_{\kappa,y}^2 + \mathcal{A}_{\kappa,y} d_{\kappa,x}^2$ and $\mathcal{A}_\kappa := \mathcal{A}_{\kappa,x}\mathcal{A}_{\kappa,y}$.

Besides, we also introduce the discrete inner products

$$(w,q)_\kappa = \kappa_x \kappa_y \sum_{i=1}^{N_x^\kappa - 1} \sum_{j=1}^{N_y^\kappa - 1} w_{i,j} q_{i,j},$$

$$(w,q)_{\kappa,x} = \kappa_x \kappa_y \sum_{i=0}^{N_x^\kappa - 1} \sum_{j=1}^{N_y^\kappa - 1} w_{i+\frac{1}{2},j} q_{i+\frac{1}{2},j}, \quad (w,q)_{\kappa,y} = \kappa_x \kappa_y \sum_{i=1}^{N_x^\kappa - 1} \sum_{j=0}^{N_y^\kappa - 1} w_{i,j+\frac{1}{2}} q_{i,j+\frac{1}{2}},$$

$$< w, q >_\kappa = \kappa_x \kappa_y \left[\frac{1}{4} \sum_{i=\{0,N_\kappa^x\}} \sum_{j=\{0,N_\kappa^y\}} + \frac{1}{2} \sum_{i=1}^{N_\kappa^x - 1} \sum_{j=\{0,N_\kappa^y\}} + \frac{1}{2} \sum_{i=\{0,N_\kappa^x\}} \sum_{j=1}^{N_\kappa^y - 1} + \sum_{i=1}^{N_\kappa^x - 1} \sum_{j=1}^{N_\kappa^y - 1} \right] w_{i,j} q_{i,j},$$

and corresponding discrete L^2 and L^∞ norms

$$\|w\|_\kappa = \sqrt{(w,w)_\kappa}, \quad |||w|||_\kappa = \sqrt{< w, w >_\kappa}, \quad \|w\|_{\mathcal{A},\kappa} = \sqrt{(\mathcal{A}_\kappa w, w)_\kappa}, \quad \|w\|_{\kappa,\infty} = \max_{(i,j)\in\bar{\omega}_\kappa} |w_{i,j}|.$$

It is easy to check that $\|\mathcal{A}_\kappa w\|_\kappa \le |||w|||_\kappa$ for any $w \in \mathcal{V}_\kappa$. Moreover, two well-known and useful lemmas (*cf.* [29, 47]) are listed below.

Lemma 2.1. *For any $w \in \mathcal{V}_\kappa^0$, we have $\frac{4}{9}\|w\|_\kappa \le \|\mathcal{A}_\kappa w\|_\kappa \le \|w\|_\kappa$ and $\frac{2}{3}\|w\|_\kappa \le \|w\|_{\mathcal{A},\kappa} \le \|w\|_\kappa$.*

Lemma 2.2. *For any $w \in \mathcal{V}_\kappa^0$, there exists a positive constant C_0 such that $\|w\|_{\kappa,\infty} \le C_0\left(\|w\|_{\mathcal{A},\kappa} + \|\Lambda_\kappa w\|_\kappa\right)$.*

2.2. Piecewise bi-cubic Lagrange interpolation

An important tool used in the construction of high-order two-grid method is the local high-order Lagrange interpolation from coarse-grid space to fine-grid space. We shall present and discuss its properties in this subsection.

We first define the one-dimensional piecewise cubic Lagrange interpolation along x-direction. For each $x \in (x_i^H, x_{i+1}^H)$ with $0 \le i \le N_x^H - 1$, we use $\left\{\phi_{i,s}^x(x)\right\}_{s=0}^3$ to represent the cubic Lagrange interpolation basis functions. For $1 \le i \le N_x^H - 2$, $\phi_{i,s}^x(x)$ is defined as

$$\phi_{i,s}^x(x) := \begin{cases} -\dfrac{(x-x_i^H)(x-x_{i+1}^H)(x-x_{i+2}^H)}{6H_x^3}, & s=0, \\ \dfrac{(x-x_{i-1}^H)(x-x_{i+1}^H)(x-x_{i+2}^H)}{2H_x^3}, & s=1, \\ -\dfrac{(x-x_{i-1}^H)(x-x_i^H)(x-x_{i+2}^H)}{2H_x^3}, & s=2, \\ \dfrac{(x-x_{i-1}^H)(x-x_i^H)(x-x_{i+1}^H)}{6H_x^3}, & s=3. \end{cases} \tag{2.1}$$

For $i=0$, *i.e.*, $x \in (x_0^H, x_1^H)$, we define $\phi_{0,s}^x(x) := \phi_{1,s}^x(x)$; and for $i = N_x^H - 1$, *i.e.*, $x \in (x_{N_x^H-1}^H, x_{N_x^H}^H)$, we define $\phi_{N_x^H-1,s}^x(x) := \phi_{N_x^H-2,s}^x(x)$. Then, for any continuous function $w(x)$, the piecewise cubic Lagrange interpolation operator $\Pi_{H,x}$ along x-direction is defined as

$$\Pi_{H,x} w(x) := \begin{cases} \displaystyle\sum_{s=0}^3 w_s\ \phi_{0,s}^x(x), & x \in (x_0^H, x_1^H), \quad i=0, \\ \displaystyle\sum_{s=0}^3 w_{i-1+s}\ \phi_{i,s}^x(x), & x \in (x_i^H, x_{i+1}^H), \quad 1 \le i \le N_x^H - 2, \\ \displaystyle\sum_{s=0}^3 w_{N_x^H-3+s}\ \phi_{N_x^H-1,s}^x(x), & x \in (x_{N_x^H-1}^H, x_{N_x^H}^H), \quad i = N_x^H - 1, \end{cases} \tag{2.2}$$

where $w_i = w(x_i^H)$ for $0 \le i \le N_x^H$.

Similarly, we can define the cubic Lagrange interpolation basis functions $\{\phi_{j,s}^y(y)\}_{s=0}^3$ and corresponding piecewise cubic Lagrange interpolation operator $\Pi_{H,y}$ along y-direction. Therefore, the piecewise bi-cubic Lagrange interpolation operator Π_H can be defined as the tensor product of the one-dimensional piecewise cubic Lagrange interpolation operators in two directions, that is, $\Pi_H := \Pi_{H,y}\Pi_{H,x}$. When no confusion caused, below we denote $(\Pi_H w)_{i,j} = \Pi_H w_{i,j}$ for $(i,j) \in \bar{\omega}_h$ and $w \in \mathcal{V}_H$.

Lemma 2.3 ([38]). *Assume that $w \in W^{4,\infty}(\Omega)$, there exists positive constants C_1 and C_2, independent of H, such that*

$$\|w - \Pi_H w\|_{h,\infty} \le C_1 H^4 \quad \text{and} \quad \|w - \Pi_H w\|_h \le C_2 H^4.$$

Next, the bounds for $\{\phi^x_{i,s}(x)\}_{s=0}^3$ are given in the following lemma in order to support the proof of boundedness conclusions for the operator Π_H.

Lemma 2.4. *The cubic Lagrange interpolation basis functions $\{\phi^x_{i,s}(x)\}_{s=0}^3$ defined in (2.1) are bounded, i.e., for $1 \le i \le N_x^H - 2$, we have*

$$\left|\phi^x_{i,0}(x)\right| \le \begin{cases} 1, & x \in (x^H_{i-1}, x^H_i), \\ \dfrac{\sqrt{3}}{27}, & x \in (x^H_i, x^H_{i+1}), \\ \dfrac{\sqrt{3}}{27}, & x \in (x^H_{i+1}, x^H_{i+2}), \end{cases} \qquad \left|\phi^x_{i,1}(x)\right| \le \begin{cases} \dfrac{7\sqrt{7}+10}{27}, & x \in (x^H_{i-1}, x^H_i), \\ 1, & x \in (x^H_i, x^H_{i+1}), \\ \dfrac{7\sqrt{7}-10}{27}, & x \in (x^H_{i+1}, x^H_{i+2}), \end{cases}$$

$$\left|\phi^x_{i,2}(x)\right| \le \begin{cases} \dfrac{7\sqrt{7}-10}{27}, & x \in (x^H_{i-1}, x^H_i), \\ 1, & x \in (x^H_i, x^H_{i+1}), \\ \dfrac{7\sqrt{7}+10}{27}, & x \in (x^H_{i+1}, x^H_{i+2}), \end{cases} \qquad \left|\phi^x_{i,3}(x)\right| \le \begin{cases} \dfrac{\sqrt{3}}{27}, & x \in (x^H_{i-1}, x^H_i), \\ \dfrac{\sqrt{3}}{27}, & x \in (x^H_i, x^H_{i+1}), \\ 1, & x \in (x^H_{i+1}, x^H_{i+2}). \end{cases}$$

Proof. Consider auxiliary function

$$Z(x) = \left(x - x^H_i\right)\left(x - x^H_{i+1}\right)\left(x - x^H_{i+2}\right) = (x - iH_x)(x - (i+1)H_x)(x - (i+2)H_x).$$

Note that its first derivative $Z'(x) = 3x^2 - 6(i+1)H_x x + (3i^2 + 6i + 2)H_x^2$ has two zero-points $x_- = \left(i + 1 - \frac{\sqrt{3}}{3}\right) H_x$ and $x_+ = \left(i + 1 + \frac{\sqrt{3}}{3}\right) H_x$. Thus, we have

$$\max_{x \in (x^H_{i-1}, x^H_i)} |Z(x)| = \max\left\{Z\left((i-1)H_x\right), Z\left(iH_x\right)\right\} = 6H_x^3,$$

$$\max_{x \in (x^H_i, x^H_{i+1})} |Z(x)| = \max\left\{Z\left(iH_x\right), Z\left(\left(i + 1 - \frac{\sqrt{3}}{3}\right)H_x\right), Z\left((i+1)H_x\right)\right\} = \frac{2\sqrt{3}}{9}H_x^3,$$

$$\max_{x \in (x^H_{i+1}, x^H_{i+2})} |Z(x)| = \max\left\{Z\left((i+1)H_x\right), Z\left(\left(i + 1 + \frac{\sqrt{3}}{3}\right)H_x\right), Z\left((i+2)H_x\right)\right\} = \frac{2\sqrt{3}}{9}H_x^3,$$

which, together with the definition of $\phi^x_{i,0}(x)$, leads to the first conclusion. The remaining conclusions can be similarly proved. □

Next, we apply Lemma 2.4 to establish the boundedness result for the piecewise bi-cubic Lagrange interpolation operator Π_H in the discrete L^2 and L^∞ norms.

Lemma 2.5. *For any $w \in \mathcal{V}_H^0$, the following estimate holds*

$$\|\Pi_H w\|_h \le 4\left(\frac{3 + 27^2 + (10 + 7\sqrt{7})^2}{27^2}\right)\|w\|_H =: C_3\|w\|_H.$$

Proof. Denote $\xi_{i,j} = \Pi_{H,x} w_{i,j}$, and then

$$\|\Pi_H w\|_h^2 = h_x h_y \sum_{i=1}^{N_x^h-1} \sum_{j=1}^{N_y^h-1} (\Pi_H w_{i,j})^2 = h_x \sum_{i=1}^{N_x^h-1} \left(h_y \sum_{j=1}^{N_y^h-1} (\Pi_{H,y} \xi_{i,j})^2 \right).$$

For each fixed $1 \le i \le N_x^h - 1$, denote $\eta_{i,j} = \Pi_{H,y} \xi_{i,j}$ and we see

$$\begin{aligned} h_y \sum_{j=1}^{N_y^h-1} (\Pi_{H,y} \xi_{i,j})^2 &= h_y \sum_{k=1}^{N_y^H} \sum_{j=(k-1)M_y+1}^{kM_y} \eta_{i,j}^2 \\ &= h_y \sum_{j=1}^{M_y} \eta_{i,j}^2 + h_y \sum_{k=2}^{N_y^H-1} \sum_{j=(k-1)M_y+1}^{kM_y} \eta_{i,j}^2 + h_y \sum_{j=(N_y^H-1)M_y+1}^{N_y^h-1} \eta_{i,j}^2 =: I_1 + I_2 + I_3. \end{aligned} \tag{2.3}$$

We use the identity $\left(\sum_{i=1}^4 a_i\right)^2 \le 4 \sum_{i=1}^4 a_i^2$ and the definition of cubic Lagrange interpolation operator $\Pi_{H,y}$ (*cf.* (2.2)) to obtain

$$I_1 = h_y \sum_{j=1}^{M_y} \left(\sum_{s=0}^{3} \phi_{0,s}^y(y_j^h) \xi_{i,s} \right)^2 \le 4 h_y \sum_{j=1}^{M_y} \sum_{s=0}^{3} \left(\phi_{0,s}^y(y_j^h) \right)^2 (\xi_{i,s})^2, \tag{2.4}$$

which, together with Lemma 2.4, yields

$$\begin{aligned} I_1 &\le 4 h_y \sum_{j=1}^{M_y} \left(\xi_{i,0}^2 + \frac{(7\sqrt{7}+10)^2}{27^2} \xi_{i,1}^2 + \frac{(7\sqrt{7}-10)^2}{27^2} \xi_{i,2}^2 + \frac{3}{27^2} \xi_{i,3}^2 \right) \\ &= 4 H_y \left(\xi_{i,0}^2 + \frac{(7\sqrt{7}+10)^2}{27^2} \xi_{i,1}^2 + \frac{(7\sqrt{7}-10)^2}{27^2} \xi_{i,2}^2 + \frac{3}{27^2} \xi_{i,3}^2 \right). \end{aligned} \tag{2.5}$$

Analogous to the process (2.4)–(2.5), we can also show

$$\begin{aligned} I_2 &\le 4 H_y \sum_{k=2}^{N_y^H-1} \left(\frac{3}{27^2} \xi_{i,k-2}^2 + \xi_{i,k-1}^2 + \xi_{i,k}^2 + \frac{3}{27^2} \xi_{i,k+1}^2 \right), \\ I_3 &\le 4 H_y \left(\frac{3}{27^2} \xi_{i,N_y^H-3}^2 + \frac{(7\sqrt{7}-10)^2}{27^2} \xi_{i,N_y^H-2}^2 + \frac{(7\sqrt{7}+10)^2}{27^2} \xi_{i,N_y^H-1}^2 + \xi_{i,N_y^H}^2 \right). \end{aligned} \tag{2.6}$$

Now, inserting (2.5)–(2.6) into (2.3) gives us

$$\begin{aligned} h_y \sum_{j=1}^{N_y^h-1} (\Pi_{H,y} \xi_{i,j})^2 &\le 4 H_y \Bigg(\frac{3 + 27^2 + (7\sqrt{7}+10)^2}{27^2} \xi_{i,1}^2 + \frac{3 + 2 \times 27^2 + (7\sqrt{7}-10)^2}{27^2} \xi_{i,2}^2 \\ &\quad + \frac{9 + 2 \times 27^2}{27^2} \xi_{i,3}^2 + \frac{6 + 2 \times 27^2}{27^2} \sum_{j=4}^{N_y^H-4} \xi_{i,j}^2 + \frac{9 + 2 \times 27^2}{27^2} \xi_{i,N_y^H-3}^2 \\ &\quad + \frac{3 + 2 \times 27^2 + (7\sqrt{7}-10)^2}{27^2} \xi_{i,N_y^H-2}^2 + \frac{3 + 27^2 + (7\sqrt{7}+10)^2}{27^2} \xi_{i,N_y^H-1}^2 \Bigg) \\ &\le C_3 H_y \sum_{j=1}^{N_y^H-1} \xi_{i,j}^2, \end{aligned}$$

which further implies

$$\|\Pi_H w\|_h^2 \le h_x \sum_{i=1}^{N_x^h-1}\left(C_3 H_y \sum_{j=1}^{N_y^H-1} \xi_{i,j}^2\right) = C_3 H_y \sum_{j=1}^{N_y^H-1}\left(h_x \sum_{i=1}^{N_x^h-1} (\Pi_{H,x} w_{i,j})^2\right). \tag{2.7}$$

Next, for each fixed $1 \le j \le N_y^H - 1$, we replace $\{\xi, y, j\}$ with $\{w, x, i\}$ in (2.3)–(2.7) to similarly obtain

$$h_x \sum_{i=1}^{N_x^h-1} (\Pi_{H,x} w_{i,j})^2 \le C_3 H_x \sum_{i=1}^{N_x^H-1} w_{i,j}^2.$$

Consequently, we have

$$\|\Pi_H w\|_h^2 \le C_3^2 H_x H_y \sum_{i=1}^{N_x^H-1}\sum_{j=1}^{N_y^H-1} w_{i,j}^2 = C_3^2 \|w\|_H^2.$$

The proof is completed. □

Lemma 2.6. *For any $w \in \mathcal{V}_H^0$, the following estimate holds*

$$\|\Pi_H w\|_{h,\infty} \le \left(\frac{54+2\sqrt{3}}{27}\right)^2 \|w\|_{H,\infty} =: C_4 \|w\|_{H,\infty}.$$

Proof. Suppose $\|\Pi_H w\|_{h,\infty} = |\Pi_H w_{i^*,j^*}|$ where $iM_x \le i^* \le (i+1)M_x$ and $jM_y \le j^* \le (j+1)M_y$. Denote $\xi_{i^*,j} = \Pi_{H,x} w_{i^*,j}$ for $0 \le j \le N_y^H$, then the triangle inequality and Lemma 2.4 give us

$$\begin{aligned}
\|\Pi_H w\|_{h,\infty} = |\Pi_{H,y}\xi_{i^*,j^*}| &= \left|\sum_{s=0}^{3} \phi_{j,s}^y(y_{j^*}^h)\xi_{i^*,j-1+s}\right| \\
&\le \frac{\sqrt{3}}{27}|\xi_{i^*,j-1}| + |\xi_{i^*,j}| + |\xi_{i^*,j+1}| + \frac{\sqrt{3}}{27}|\xi_{i^*,j+2}| \\
&\le \frac{54+2\sqrt{3}}{27} \max_{1\le k\le N_y^H-1} |\xi_{i^*,k}|, \qquad \text{if } M_y \le j^* \le (N_y^H-1)M_y,
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
\|\Pi_H w\|_{h,\infty} &\le \frac{7\sqrt{7}+10}{27}|\xi_{i^*,1}| + \frac{7\sqrt{7}-10}{27}|\xi_{i^*,2}| + \frac{\sqrt{3}}{27}|\xi_{i^*,3}| \\
&\le \frac{14\sqrt{7}+\sqrt{3}}{27} \max_{1\le k\le N_y^H-1} |\xi_{i^*,k}|, \qquad \text{if } 1 \le j^* < M_y,
\end{aligned} \tag{2.9}$$

$$\|\Pi_H w\|_{h,\infty} \le \frac{14\sqrt{7}+\sqrt{3}}{27} \max_{1\le k\le N_y^H-1} |\xi_{i^*,k}|, \qquad \text{if } (N_y^H-1)M_y < j^* \le N_y^h - 1. \tag{2.10}$$

Next, it is necessary to estimate $\max_{1\le k\le N_y^H-1}|\xi_{i^*,k}|$. For fixed index $1 \le k \le N_y^H - 1$, let $|\xi_{i^{**},k}| = \max_{1\le \ell\le N_x^h-1}|\xi_{\ell,k}|$ where $iM_x \le i^{**} \le (i+1)M_x$. Thus, with a similar treatment to the above estimates, we derive

$$|\xi_{i^*,k}| = |\Pi_{H,x} w_{i^*,k}| \le |\Pi_{H,x} w_{i^{**},k}| \le \frac{54+2\sqrt{3}}{27} \max_{1\le \ell\le N_x^H-1} |w_{\ell,k}|,$$

which, together with (2.8)–(2.10), completes the proof. □

3. A nonlinear variable-step compact difference scheme

In this section, combined with the variable-step BDF2 method, we are committed to establishing a nonlinear compact difference scheme for the semilinear parabolic equations (1.1)–(1.3) enclosed with Dirichlet boundary condition. Meanwhile, we shall develop the corresponding unconditional stability and optimal-order error estimate by imposing the following regularity assumption

$$u \in C^1(0,T;H^6(\Omega)) \cap C^3(0,T;H^2(\Omega)). \tag{3.1}$$

By the Sobolev embedding theorem, there exist two constants m and M such that the solution of problem (1.1)–(1.3) is bounded, *i.e.*, $u \in \mathcal{B} := [m, M]$. Most previous works always require that the nonlinear function $f(u)$ is continuously differentiable with uniformly bounded derivative on $\mathbb{R}$. However, in this paper we only assume that the nonlinear term $f \in C^3(\mathcal{B}_\delta)$ holds locally on $\mathcal{B}_\delta := [m-\delta, M+\delta]$ for a fixed small $\delta > 0$.

3.1. Nonlinear compact difference scheme

To construct variable temporal stepsize numerical schemes, we consider a nonuniform temporal partition $0 = t_0 < t_1 < \cdots < t_N = T$ with temporal stepsize $\tau_k := t_k - t_{k-1}$ for $1 \le k \le N$. Denote the maximum temporal stepsize $\tau := \max_{1\le k\le N} \tau_k$ and adjacent temporal stepsize ratio $r_k := \tau_k/\tau_{k-1}$ $(k \ge 2)$. For any real sequence $\{w^k\}_{k=0}^N$, set $\nabla_\tau w^k = w^k - w^{k-1}$. The variable-step BDF2 formula is defined as

$$\mathcal{D}_2 w^n := \frac{1+2r_n}{\tau_n(1+r_n)}\nabla_\tau w^n - \frac{r_n^2}{\tau_n(1+r_n)}\nabla_\tau w^{n-1} \quad \text{for} \quad n \ge 2.$$

In particular, when $n = 1$, we use the BDF1 (*i.e.*, backward Euler) formula $\mathcal{D}_1 w^1 = \frac{1}{\tau_1}\nabla_\tau w^1$ for the first time level discretization. Let $r_1 = 0$, we then rewrite the above variable-step BDF formula as a unified discrete convolution summation

$$\mathcal{D}_2 w^n = \sum_{k=1}^n b_{n-k}^{(n)} \nabla_\tau w^k \quad \text{for} \quad 1 \le n \le N,$$

where the discrete convolution kernels $\{b_{n-k}^{(n)}\}_{k=1}^n$ are defined by $b_0^{(1)} := 1/\tau_1$, and when $n \ge 2$,

$$b_0^{(n)} := \frac{1+2r_n}{\tau_n(1+r_n)}, \quad b_1^{(n)} := -\frac{r_n^2}{\tau_n(1+r_n)}, \quad \text{and} \quad b_j^{(n)} := 0 \ \text{ for } 2 \le j \le n-1.$$

Next, we introduce the so-called DOC kernels $\{\theta_{n-m}^{(n)}\}_{m=1}^n$ by [11,28,32]

$$\sum_{m=k}^n \theta_{n-m}^{(n)} b_{m-k}^{(m)} = \delta_{n,k}, \quad 1 \le k \le n, \tag{3.2}$$

for $1 \le n \le N$, where $\delta_{n,k}$ is the Kronecker delta symbol with $\delta_{n,k} = 1$ if $n = k$ and $\delta_{n,k} = 0$ if $n \ne k$. By exchanging the summation order and using the definition (3.2), it is easy to check that

$$\sum_{m=1}^n \theta_{n-m}^{(n)} \mathcal{D}_2 w^m = \sum_{k=1}^n \nabla_\tau w^k \sum_{m=k}^n \theta_{n-m}^{(n)} b_{m-k}^{(m)} = \nabla_\tau w^n, \quad 1 \le n \le N. \tag{3.3}$$

By introducing an auxiliary variable $q(x,y,t) = u_t(x,y,t)$, model (1.1) can be rewritten as

$$q(x,y,t) = u_t(x,y,t), \quad (x,y) \in \Omega,\ t \in (0,T], \tag{3.4}$$

$$q(x,y,t) = c\Delta u(x,y,t) + f(u(x,y,t)) + g(x,y,t), \quad (x,y) \in \Omega,\ t \in (0,T]. \tag{3.5}$$

In the following, we shall present the variable-step compact difference approximation of model (1.1) through the equivalent system (3.4)–(3.5).

Let $U_{i,j}^n := u(x_i^h, y_j^h, t_n)$ and $Q_{i,j}^n := q(x_i^h, y_j^h, t_n)$ be the exact nodal solutions. Then, we apply the compact operator $\mathcal{A}_h$ in space and variable-step BDF formula in time to (3.4)–(3.5) to get the following equations

$$Q_{i,j}^n = \mathcal{D}_2 U_{i,j}^n + (R_t^n)_{i,j}, \quad (i,j) \in \omega_h,\ 1 \le n \le N, \tag{3.6}$$

$$\mathcal{A}_h Q_{i,j}^n = c\Lambda_h U_{i,j}^n + \mathcal{A}_h f(U_{i,j}^n) + \mathcal{A}_h g_{i,j}^n + (R_s^n)_{i,j}, \quad (i,j) \in \omega_h,\ 0 \le n \le N, \tag{3.7}$$

where $(R_t^n)_{i,j}$ and $(R_s^n)_{i,j}$ respectively represent the local temporal and spatial truncation errors at point (x_i^h, y_j^h, t_n) that

$$(R_t^n)_{i,j} = u_t(x_i^h, y_j^h, t_n) - \mathcal{D}_2 U_{i,j}^n \quad \text{and} \quad (R_s^n)_{i,j} = c\mathcal{A}_h \Delta u(x_i^h, y_j^h, t_n) - c\Lambda_h U_{i,j}^n.$$

Moreover, under the regularity condition (3.1), by using the Taylor expansion and well-known Bramble-Hilbert Lemma, it is easy to see

$$\|R_t^1\|_h + \|\Lambda_h R_t^1\|_h \le C_5 \tau_1, \quad \|R_t^n\|_h + \|\Lambda_h R_t^n\|_h \le C_6 \tau_n \tau \quad \text{for } 2 \le n \le N, \tag{3.8}$$

$$\|R_s^0\|_h + \|R_s^n\|_h + \|\mathcal{D}_2 R_s^n\|_h \le C_7 h^4 \quad \text{for } 1 \le n \le N. \tag{3.9}$$

Let $\{u_{i,j}^n, q_{i,j}^n\}$ be the finite difference approximations to $\{U_{i,j}^n, Q_{i,j}^n\}$. Then, by dropping the local truncation errors in (3.6)–(3.7), we obtain the following nonlinear variable-step BDF2 compact difference scheme

$$q_{i,j}^n = \mathcal{D}_2 u_{i,j}^n, \quad (i,j) \in \omega_h, \tag{3.10}$$

$$\mathcal{A}_h q_{i,j}^n = c\Lambda_h u_{i,j}^n + \mathcal{A}_h f(u_{i,j}^n) + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h, \tag{3.11}$$

for $1 \le n \le N$, subject to the initial and boundary conditions

$$u_{i,j}^0 = \varphi(x_i^h, y_j^h),\ \mathcal{A}_h q_{i,j}^0 = c\Lambda_h u_{i,j}^0 + \mathcal{A}_h f(u_{i,j}^0) + \mathcal{A}_h g_{i,j}^0, \quad (i,j) \in \omega_h, \tag{3.12}$$

$$u_{i,j}^n = \psi(x_i^h, y_j^h, t_n), \quad (i,j) \in \partial\omega_h,\ 0 \le n \le N. \tag{3.13}$$

Remark 3.1. From the computational viewpoint, by eliminating the auxiliary variable $q_{i,j}^n$, we get an equivalent computable scheme for the primal variable $u_{i,j}^n$, that is

$$\mathcal{D}_2 \mathcal{A}_h u_{i,j}^n - c\Lambda_h u_{i,j}^n = \mathcal{A}_h f(u_{i,j}^n) + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h,\ 1 \le n \le N. \tag{3.14}$$

However, for the maximum-norm error analysis below, we prefer to consider the error splitting scheme (3.10)–(3.11), in which unconditional maximum-norm error estimate can be proved, see Theorems 3.9 and 3.10.

In the following, we aim to prove the unique solvability and error analysis for the nonlinear variable-step compact difference scheme (3.10)–(3.13). The main difficulties in deriving error estimate for the nonlinear difference scheme are two-folds. One is the maximum-bound of the numerical solution, and the other is an unconditional error estimate without any temporal-spatial stepsize ratio condition. On one hand, the temporal-spatial error splitting approach (3.10)–(3.11) itself plays an important role in the unconditional maximum-norm error analysis. On the other hand, using a cut-off approach, an auxiliary nonlinear compact difference scheme is to be introduced below for the purpose of unique solvability and maximum-bound of the numerical solution to (3.10)–(3.13).

Let $(\bar{u}^n, \bar{q}^n) = \{\bar{u}_{i,j}^n, \bar{q}_{i,j}^n\}$ be solutions of the following nonlinear compact difference scheme

$$\bar{q}_{i,j}^n = \mathcal{D}_2 \bar{u}_{i,j}^n, \quad (i,j) \in \omega_h, \tag{3.15}$$

$$\mathcal{A}_h \bar{q}_{i,j}^n - c\Lambda_h \bar{u}_{i,j}^n = \mathcal{A}_h \bar{f}(\bar{u}_{i,j}^n) + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h, \tag{3.16}$$

for $1 \le n \le N$, subject to the same initial and boundary conditions. Here $\bar{f}(w) \in C^2(\mathbb{R})$ is chosen as a smooth extension function of f from $\mathcal{B}_\delta$ to $\mathbb{R}$, and $f(w)$ is viewed as a cut-off function of $\bar{f}(w)$ (*cf.* Refs. [4, 49]). For example,

$$\bar{f}(w) := \begin{cases} 0, & w < m - 2\delta, \\ \in C^2[m-2\delta, m-\delta], & w \in [m-2\delta, m-\delta], \\ f(w), & w \in \mathcal{B}_\delta, \\ \in C^2[M+\delta, M+2\delta], & w \in [M+\delta, M+2\delta], \\ 0, & w > M + 2\delta. \end{cases}$$

It is obvious that such defined $\bar{f}(\cdot)$ is globally twice continuously differentiable with uniformly bounded derivative on $\mathbb{R}$, *i.e.*,

$$|\bar{f}(w)| + |\bar{f}'(w)| + |\bar{f}''(w)| \le K_f \qquad \text{for} \quad w \in \mathbb{R}. \tag{3.17}$$

Below we shall first show the unique solvability and error analysis for the nonlinear auxiliary scheme (3.15)–(3.16) instead of (3.10)–(3.11). Several useful lemmas are presented for the subsequent theoretical analysis.

Lemma 3.2 ([11, 32])**.** *Assume that the adjacent temporal stepsize ratio r_k satisfy $0 < r_k < 4.8645$, then the DOC kernels $\{\theta_{n-k}^{(n)}\}$ defined in (3.2) are positive semi-definite, i.e., for any real sequence $\{w^k\}_{k=1}^n$, it holds that*

$$\sum_{k=1}^{n} w^k \sum_{m=1}^{k} \theta_{k-m}^{(k)} w^m \ge 0 \quad \text{for } n \ge 1.$$

Lemma 3.3 ([11, 28])**.** *The DOC kernels $\{\theta_{n-k}^{(n)}\}$ defined in (3.2) have the following properties*

$$\theta_{n-m}^{(n)} > 0 \quad \text{for } 1 \le m \le n, \qquad \sum_{m=1}^{n} \theta_{n-m}^{(n)} = \tau_n \quad \text{for } n \ge 1,$$

$$\sum_{m=1}^{n} \sum_{k=m}^{n} \theta_{k-m}^{(k)} = t_n, \quad \text{and} \quad \sum_{k=m}^{n} \theta_{k-m}^{(k)} \le 2\tau \quad \text{for } 1 \le m \le n.$$

Lemma 3.4. *Assume that the adjacent temporal stepsize ratio r_k satisfy $0 < r_k < 4.8645$, then for any real sequence $\{w^k\}_{k=1}^n \in \mathcal{V}_\kappa^0$, it holds that*

$$\sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\Lambda_\kappa w^m, w^k)_\kappa \le 0.$$

Proof. Due to $\mathcal{A}_{\kappa,x}$ and $\mathcal{A}_{\kappa,y}$ are self-adjoint and positive definite operators, there exist $\eta_{\kappa,x}$ and $\eta_{\kappa,y}$ such that $\mathcal{A}_{\kappa,x} = \eta_{\kappa,x}^2$ and $\mathcal{A}_{\kappa,y} = \eta_{\kappa,y}^2$. Thus, by summation by parts and homogeneous boundary conditions, we have

$$\begin{aligned} &\sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\Lambda_\kappa w^m, w^k)_\kappa \\ &= -\sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\mathcal{A}_{\kappa,x} d_{\kappa,y} w^m, d_{\kappa,y} w^k)_{\kappa,y} - \sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\mathcal{A}_{\kappa,y} d_{\kappa,x} w^m, d_{\kappa,x} w^k)_{\kappa,x} \\ &= -\sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\eta_{\kappa,x} d_{\kappa,y} w^m, \eta_{\kappa,x} d_{\kappa,y} w^k)_{\kappa,y} - \sum_{k=1}^{n} \sum_{m=1}^{k} \theta_{k-m}^{(k)} (\eta_{\kappa,y} d_{\kappa,x} w^m, \eta_{\kappa,y} d_{\kappa,x} w^k)_{\kappa,x} \le 0, \end{aligned}$$

where Lemma 3.2 has been applied in the last step. □

3.2. Unique solvability

It is noticed that in [30], the authors proposed a nonlinear variable-step BDF2 scheme for the Allen–Cahn equation with a polynomial type double-well nonlinear potential, and they proved its unique solvability by showing that the solution of the nonuniform BDF2 scheme is equivalent to a minimization problem with strictly convex energy functional. In this subsection, the unique solvability of the auxiliary BDF2 scheme (3.15)–(3.16) for semilinear parabolic equations with general nonlinearity will be discussed by the Browder's fixed point theorem (see *e.g.* [1]). For simplicity of presentation, below we assume $c = 1$. By homogenization treatment, it suffices to consider the corresponding homogeneous case.

Theorem 3.5. *The auxiliary nonlinear compact difference scheme* (3.15)–(3.16) *is solvable if the maximum temporal stepsize satisfies* $\tau \le \frac{4}{13K_f}$.

Proof. Note that (3.15)–(3.16) is equivalent to

$$b_0^{(n)}\mathcal{A}_h\bar{u}_{i,j}^n - G_{i,j}^n - \Lambda_h\bar{u}_{i,j}^n - \mathcal{A}_h\bar{f}(\bar{u}_{i,j}^n) - \mathcal{A}_h g_{i,j}^n = 0, \quad (i,j)\in\omega_h, \tag{3.18}$$

where $G_{i,j}^n$ is defined as

$$G_{i,j}^n := (b_0^{(n)} - b_1^{(n)})\mathcal{A}_h\bar{u}_{i,j}^{n-1} + b_1^{(n)}\mathcal{A}_h\bar{u}_{i,j}^{n-2}.$$

Denote the mapping

$$[T\bar{u}]_{i,j}^n := b_0^{(n)}\mathcal{A}_h\bar{u}_{i,j}^n - G_{i,j}^n - \Lambda_h\bar{u}_{i,j}^n - \mathcal{A}_h\bar{f}(\bar{u}_{i,j}^n) - \mathcal{A}_h g_{i,j}^n, \quad (i,j)\in\omega_h.$$

Then, taking the inner product of $T\bar{u}^n$ with $\bar{u}^n$ in the sense of $(\cdot,\cdot)_h$ gives us

$$\begin{aligned}(T\bar{u}^n,\bar{u}^n)_h &= b_0^{(n)}\left(\mathcal{A}_h\bar{u}^n,\bar{u}^n\right)_h - (G^n,\bar{u}^n)_h - (\Lambda_h\bar{u}^n,\bar{u}^n)_h - \left(\mathcal{A}_h\bar{f}(\bar{u}^n),\bar{u}^n\right)_h - (\mathcal{A}_h g^n,\bar{u}^n)_h \\ &:= \sum_{k=1}^5 S_k.\end{aligned} \tag{3.19}$$

Next, we estimate (3.19) term-by-term. For S_2, utilizing Cauchy-Schwarz inequality yields

$$S_2 \ge -\left((b_0^{(n)} - b_1^{(n)})\|\bar{u}^{n-1}\|_{\mathcal{A},h} - b_1^{(n)}\|\bar{u}^{n-2}\|_{\mathcal{A},h}\right)\|\bar{u}^n\|_{\mathcal{A},h}.$$

Moreover, by summation by parts and homogeneous boundary conditions, and noting that $\mathcal{A}_{h,x}$ and $\mathcal{A}_{h,y}$ are self-adjoint and positive definite operators, we have

$$S_3 = (\mathcal{A}_{h,y}d_{h,x}\bar{u}^n, d_{h,x}\bar{u}^n)_{h,x} + (\mathcal{A}_{h,x}d_{h,y}\bar{u}^n, d_{h,y}\bar{u}^n)_{h,y} \ge 0.$$

Besides, we use Cauchy-Schwarz inequality and triangle inequality to get

$$\begin{aligned}S_4 + S_5 &= -\left(\mathcal{A}_h\left(\bar{f}(\bar{u}^n) + g^n\right),\bar{u}^n\right)_h \ge -|||\,\bar{f}(\bar{u}^n) + g^n\,|||_h\,\|\bar{u}^n\|_h \\ &\ge -|||\,\bar{f}(\bar{u}^n) - \bar{f}(\varphi)\,|||_h\,\|\bar{u}^n\|_h - |||\bar{f}(\varphi) + g^n\,|||_h\,\|\bar{u}^n\|_h,\end{aligned}$$

which, together with the globally Lipschitz continuous of $\bar{f}$ and Lemma 2.1, further leads to

$$\begin{aligned}S_4 + S_5 &\ge -K_f\left(\|\bar{u}^n\|_h + \|\varphi\|_h\right)\|\bar{u}^n\|_h - |||\bar{f}(\varphi) + g^n\,|||_h\,\|\bar{u}^n\|_h \\ &\ge -\frac{9}{4}K_f\|\bar{u}^n\|_{\mathcal{A},h}^2 - \frac{3}{2}\left(K_f\|\varphi\|_h + |||f(\varphi) + g^n|||_h\right)\|\bar{u}^n\|_{\mathcal{A},h}.\end{aligned}$$

Denote

$$Y := (b_0^{(n)} - b_1^{(n)})\|\bar{u}^{n-1}\|_{\mathcal{A},h} - b_1^{(n)}\|\bar{u}^{n-2}\|_{\mathcal{A},h} + \frac{3}{2}\left(K_f\|\varphi\|_h + |||f(\varphi) + g^n|||_h\right).$$

Thus, inserting the above estimates into (3.19), if $\tau_n \le \frac{4}{13K_f}$, we conclude

$$(T\bar{u}^n, \bar{u}^n)_h \ge \left(b_0^{(n)} - \frac{9}{4}K_f\right)\|\bar{u}^n\|_{\mathcal{A},h}^2 - Y\|\bar{u}^n\|_{\mathcal{A},h} \ge (K_f\|\bar{u}^n\|_{\mathcal{A},h} - Y)\|\bar{u}^n\|_{\mathcal{A},h} \ge 0,$$

when $\|\bar{u}^n\|_{\mathcal{A},h} = Y/K_f$. Consequently, Browder's fixed point theorem shows there exists a $\bar{u}^n \in \mathcal{V}_h^0$ such that $T\bar{u}^n = 0$, which implies the solvability of (3.18), and thus the auxiliary scheme (3.15)–(3.16). □

Theorem 3.6. *The solution of the auxiliary nonlinear compact difference scheme* (3.15)–(3.16) *is unique if the adjacent temporal stepsize ratios r_k satisfy $0 < r_k < 4.8645$ and the maximum temporal stepsize $\tau \le \frac{1}{6K_f}$.*

Proof. The argument is by contradiction. Suppose that there are two solutions $(\bar{u}_1^n, \bar{q}_1^n)$ and $(\bar{u}_2^n, \bar{q}_2^n)$ satisfying (3.15)–(3.16) with the same initial and boundary conditions. Denote $v^n := \bar{u}_1^n - \bar{u}_2^n$, It is easy to check that

$$\mathcal{D}_2\mathcal{A}_h v_{i,j}^n - \Lambda_h v_{i,j}^n = \mathcal{A}_h\bar{f}([\bar{u}_1]_{i,j}^n) - \mathcal{A}_h\bar{f}([\bar{u}_2]_{i,j}^n). \tag{3.20}$$

Let $n = m$ in (3.20), and then, multiply it by the DOC kernels $\theta_{k-m}^{(k)}$ and sum over m from 1 to k yields

$$\nabla_\tau\mathcal{A}_h v_{i,j}^k - \sum_{m=1}^{k}\theta_{k-m}^{(k)}\Lambda_h v_{i,j}^m = \sum_{m=1}^{k}\theta_{k-m}^{(k)}\mathcal{A}_h\left(\bar{f}([\bar{u}_1]_{i,j}^m) - \bar{f}([\bar{u}_2]_{i,j}^m)\right), \tag{3.21}$$

where we have used the identity (3.3).

Furthermore, taking the inner product of the above equation with $2v^k$, and summing the resulting equation from $k = 1$ to n gives

$$\|v^n\|_{\mathcal{A},h}^2 - 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}(\Lambda_h v^m, v^k)_h \le 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\left(\mathcal{A}_h\bar{f}(\bar{u}_1^n) - \mathcal{A}_h\bar{f}(\bar{u}_2^n), v^k\right)_h, \tag{3.22}$$

where $v^0 = 0$ has been used.

Using Cauchy-Schwarz inequality and Lemma 2.1 we see

$$\begin{aligned}
2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\left(\mathcal{A}_h\bar{f}(\bar{u}_1^n) - \mathcal{A}_h\bar{f}(\bar{u}_2^n), v^k\right)_h &\le 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|\bar{f}(\bar{u}_1^n) - \bar{f}(\bar{u}_2^n)\|_{\mathcal{A},h}\|v^k\|_{\mathcal{A},h}\\
&\le 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|\bar{f}(\bar{u}_1^n) - \bar{f}(\bar{u}_2^n)\|_h\|v^k\|_{\mathcal{A},h},
\end{aligned}$$

which, together with the global Lipschitz continuous property of $\bar{f}$, gives

$$\begin{aligned}
2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\left(\mathcal{A}_h\bar{f}(\bar{u}_1^n) - \mathcal{A}_h\bar{f}(\bar{u}_2^n), v^k\right)_h &\le 2K_f\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|v^m\|_h\|v^k\|_{\mathcal{A},h}\\
&\le 3K_f\sum_{k=1}^{n}\|v^k\|_{\mathcal{A},h}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|v^m\|_{\mathcal{A},h}.
\end{aligned} \tag{3.23}$$

Now, inserting (3.23) into (3.22), and noting that Lemma 3.4 shows that the second left-hand side term of (3.22) is positive, we obtain

$$\|v^n\|_{\mathcal{A},h}^2 \le 3K_f\sum_{k=1}^{n}\|v^k\|_{\mathcal{A},h}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|v^m\|_{\mathcal{A},h}.$$

Choosing n^* $(1 \le n^* \le n)$ such that $\|v^{n^*}\|_{\mathcal{A},h} = \max_{1\le k\le n} \|v^n\|_{\mathcal{A},h}$. Then, the above inequality yields

$$\|v^{n^*}\|_{\mathcal{A},h}^2 \le 3K_f \sum_{k=1}^{n^*} \|v^k\|_{\mathcal{A},h} \sum_{m=1}^{k} \theta_{k-m}^{(k)} \|v^{n^*}\|_{\mathcal{A},h},$$

which, eliminating $\|v^{n^*}\|_{\mathcal{A},h}$ from both sides and using Lemma 3.3, implies

$$\|v^n\|_{\mathcal{A},h} \le \|v^{n^*}\|_{\mathcal{A},h} \le 3K_f \sum_{k=1}^{n^*} \|v^k\|_{\mathcal{A},h} \sum_{m=1}^{k} \theta_{k-m}^{(k)} \le 3K_f \sum_{k=1}^{n} \tau_k \|v^k\|_{\mathcal{A},h}.$$

Then, for $\tau_n \le 1/(6K_f)$, an application of the discrete Grönwall inequality gives

$$\|v^n\|_{\mathcal{A},h} \le 6K_f \sum_{k=1}^{n-1} \tau_k \|v^k\|_{\mathcal{A},h} \le 0,$$

which proves the uniqueness of the solution to (3.15) and thus to (3.16). □

3.3. Stability analysis

In this subsection, we discuss the unconditional stability of the auxiliary nonlinear variable-step compact difference scheme (3.15)–(3.16) in the sense of maximum-norm.

Theorem 3.7. *If the adjacent temporal stepsize ratio $0 < r_k < 4.8645$ and the maximum temporal stepsize $\tau \le \frac{1}{6K_f}$, then there exists a positive constant $K_1 = K_1(T, K_f, C_0)$, such that the auxiliary nonlinear scheme (3.15)–(3.16) is unconditionally stable in the sense that*

$$\|\bar{u}^n\|_{h,\infty} \le K_1 \left(\|\varphi\|_h + \|\Lambda_h \varphi\|_h + \|f(\varphi)\|_h + \max_{0\le k\le n} \|g^k\|_h + \max_{1\le k\le n} \|\mathcal{D}_2 g^k\|_h \right).$$

Proof. By Lemma 2.2, it needs to prove $\|\bar{u}^n\|_{\mathcal{A},h}$ and $\|\Lambda_h \bar{u}^n\|_h$. While (3.16) shows that $\|\Lambda_h \bar{u}^n\|_h$ can be controlled by $\|\bar{u}^n\|_{\mathcal{A},h}$ and $\|\bar{q}^n\|_{\mathcal{A},h}$, *i.e.*,

$$\begin{aligned} \|\Lambda_h \bar{u}^n\|_h &\le \|\mathcal{A}_h \bar{f}(\bar{u}^n)\|_h + \|\mathcal{A}_h \bar{q}^n\|_h + \|g^n\|_h \\ &\le \frac{3}{2}(K_f \|\bar{u}^n\|_{\mathcal{A},h} + \|\bar{q}^n\|_{\mathcal{A},h}) + K_f \|\varphi\|_h + \|f(\varphi)\|_h + \|g^n\|_h. \end{aligned} \tag{3.24}$$

In the following, we split the proof of Theorem 3.7 into three main steps.

Step I. Estimate for $\|\bar{u}^n\|_{\mathcal{A},h}$. Similar as (3.14), inserting (3.15) into (3.16) gives us

$$\mathcal{D}_2 \mathcal{A}_h \bar{u}_{i,j}^n = \Lambda_h \bar{u}_{i,j}^n + \mathcal{A}_h \bar{f}(\bar{u}_{i,j}^n) + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h. \tag{3.25}$$

Analogous to the proof of (3.20)–(3.21), it follows from (3.3) that

$$\nabla_\tau \mathcal{A}_h \bar{u}_{i,j}^k - \sum_{m=1}^{k} \theta_{k-m}^{(k)} \Lambda_h \bar{u}_{i,j}^m = \sum_{m=1}^{k} \theta_{k-m}^{(k)} \mathcal{A}_h \bar{f}(\bar{u}_{i,j}^m) + \sum_{m=1}^{k} \theta_{k-m}^{(k)} \mathcal{A}_h g_{i,j}^m, \tag{3.26}$$

which, by taking the inner product with $2\bar{u}^k$ and summing the resulting equation from $k = 1$ to n, gives

$$\begin{aligned} \|\bar{u}^n\|_{\mathcal{A},h}^2 - \|\varphi\|_{\mathcal{A},h}^2 - 2\sum_{k=1}^{n}\sum_{m=1}^{k} \theta_{k-m}^{(k)} (\Lambda_h \bar{u}^m, \bar{u}^k)_h \le{}& 2\sum_{k=1}^{n}\sum_{m=1}^{k} \theta_{k-m}^{(k)} \left(\mathcal{A}_h \bar{f}(\bar{u}^m), \bar{u}^k\right)_h \\ &+ 2\sum_{k=1}^{n}\sum_{m=1}^{k} \theta_{k-m}^{(k)} \left(\mathcal{A}_h g^m, \bar{u}^k\right)_h. \end{aligned} \tag{3.27}$$

With a similar treatment to (3.23) and using Lemma 3.4, the above inequality yields

$$\|\bar{u}^n\|_{\mathcal{A},h}^2 \le \|\varphi\|_{\mathcal{A},h}^2 + \sum_{k=1}^{n}\left(3K_f\|\bar{u}^k\|_{\mathcal{A},h} + 2K_f\|\varphi\|_h + 2\|\bar{f}(\varphi)\|_h\right)\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|\bar{u}^m\|_{\mathcal{A},h} + 2\sum_{k=1}^{n}\|\bar{u}^k\|_{\mathcal{A},h}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|g^m\|_h.$$

Similarly, if $\|\bar{u}^{n^*}\|_{\mathcal{A},h} = \max_{1\le k\le n}\|\bar{u}^n\|_{\mathcal{A},h}$, using Lemma 3.3 we have

$$\begin{aligned}\|\bar{u}^n\|_{\mathcal{A},h} &\le \|\bar{u}^{n^*}\|_{\mathcal{A},h}\\ &\le \|\varphi\|_{\mathcal{A},h} + \sum_{k=1}^{n^*}\left(3K_f\|\bar{u}^k\|_{\mathcal{A},h} + 2K_f\|\varphi\|_h + 2\|\bar{f}(\varphi)\|_h\right)\sum_{m=1}^{k}\theta_{k-m}^{(k)} + 2\sum_{k=1}^{n^*}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|g^m\|_h\\ &\le \|\varphi\|_h + \sum_{k=1}^{n}\tau_k\left(3K_f\|\bar{u}^k\|_{\mathcal{A},h} + 2K_f\|\varphi\|_h + 2\|\bar{f}(\varphi)\|_h\right) + 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|g^m\|_h.\end{aligned} \tag{3.28}$$

For the last term of (3.28), we exchange the order of summation and utilize Lemma 3.3 to obtain

$$\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|g^m\|_h = \sum_{m=1}^{n}\|g^m\|_h\sum_{k=m}^{n}\theta_{k-m}^{(k)} \le t_n\max_{1\le k\le n}\|g^k\|_h.$$

Now, insert the above inequality into (3.28), for $\tau \le 1/(6K_f)$, we have

$$\|\bar{u}^n\|_{\mathcal{A},h} \le (2+4K_f t_n)\|\varphi\|_h + 4t_n\|\bar{f}(\varphi)\|_h + 4t_n\max_{1\le k\le n}\|g^k\|_h + 6K_f\sum_{k=1}^{n-1}\tau_k\|\bar{u}^k\|_{\mathcal{A},h}. \tag{3.29}$$

Then, applying the discrete Grönwall inequality to (3.29) yields

$$\|\bar{u}^n\|_{\mathcal{A},h} \le (2+4K_fT)\exp(6K_fT)\left(\|\varphi\|_h + \|f(\varphi)\|_h + \max_{1\le k\le n}\|g^k\|_h\right). \tag{3.30}$$

Step II. Estimate for $\|\bar{q}^n\|_{\mathcal{A},h}$. Acting the difference operator $\mathcal{D}_2$ on (3.16), and then inserting (3.15) into the resulting equation, we obtain

$$\mathcal{D}_2\mathcal{A}_h\bar{q}_{i,j}^n = \Lambda_h\bar{q}_{i,j}^n + \mathcal{D}_2\mathcal{A}_h\bar{f}(\bar{u}_{i,j}^n) + \mathcal{D}_2\mathcal{A}_h g_{i,j}^n. \tag{3.31}$$

Similarly, applying the DOC kernel relation (3.3) to (3.15) and (3.31), respectively, we arrive at

$$\nabla_\tau\bar{u}_{i,j}^k = \sum_{m=1}^{k}\theta_{k-m}^{(k)}\bar{q}_{i,j}^m, \tag{3.32}$$

$$\nabla_\tau\mathcal{A}_h\bar{q}_{i,j}^k = \sum_{m=1}^{k}\theta_{k-m}^{(k)}\Lambda_h\bar{q}_{i,j}^m + \nabla_\tau\mathcal{A}_h\bar{f}(\bar{u}_{i,j}^k) + \sum_{m=1}^{k}\theta_{k-m}^{(k)}\mathcal{D}_2\mathcal{A}_h g_{i,j}^m. \tag{3.33}$$

Now, taking the inner product of (3.33) with $2\bar{q}^k$ and summing the resulting equation from $k=1$ to n, we get

$$\|\bar{q}^n\|_{\mathcal{A},h}^2 - \|q^0\|_{\mathcal{A},h}^2 \le 2\sum_{k=1}^{n}\|\nabla_\tau\bar{f}(\bar{u}^k)\|_h\|\bar{q}^k\|_{\mathcal{A},h} + 2\sum_{k=1}^{n}\sum_{m=1}^{k}\theta_{k-m}^{(k)}\|\mathcal{D}_2 g^m\|_h\|\bar{q}^k\|_{\mathcal{A},h}. \tag{3.34}$$

Due to the global Lipschitz continuity of $\bar{f}$ and norm equivalence in Lemma 2.1, we obtain

$$\|\nabla_\tau \bar{f}(\bar{u}^k)\|_h \le K_f \|\nabla_\tau \bar{u}^k\|_h \le \frac{3}{2} K_f \sum_{m=1}^{k} \theta_{k-m}^{(k)} \|\bar{q}^m\|_{\mathcal{A},h}, \tag{3.35}$$

where (3.32) has been used in tha last inequality.

Then, inserting (3.35) into (3.34), we have

$$\|\bar{q}^n\|_{\mathcal{A},h}^2 \le \|q^0\|_{\mathcal{A},h}^2 + 3K_f \sum_{k=1}^{n} \|\bar{q}^k\|_{\mathcal{A},h} \sum_{m=1}^{k} \theta_{k-m}^{(k)} \|\bar{q}^m\|_{\mathcal{A},h} + 2\sum_{k=1}^{n} \|\bar{q}^k\|_{\mathcal{A},h} \sum_{m=1}^{k} \theta_{k-m}^{(k)} \|\mathcal{D}_2 g^m\|_h.$$

Similarly, if $\|\bar{q}^{n^*}\|_{\mathcal{A},h} = \max_{1\le k\le n} \|\bar{q}^k\|_{\mathcal{A},h}$, using Lemma 3.3, we have

$$\begin{aligned}\|\bar{q}^n\|_{\mathcal{A},h} \le \|\bar{q}^{n^*}\|_{\mathcal{A},h} &\le \|q^0\|_{\mathcal{A},h} + 3K_f \sum_{k=1}^{n^*} \|\bar{q}^k\|_{\mathcal{A},h} \sum_{m=1}^{k} \theta_{k-m}^{(k)} + 2\sum_{k=1}^{n^*}\sum_{m=1}^{k} \theta_{k-m}^{(k)} \|\mathcal{D}_2 g^m\|_h \\ &\le \|q^0\|_{\mathcal{A},h} + 2t_n \max_{1\le k\le n} \|\mathcal{D}_2 g^k\|_h + 3K_f \sum_{k=1}^{n} \tau_k \|\bar{q}^k\|_{\mathcal{A},h},\end{aligned}$$

from which, for $\tau \le 1/(6K_f)$, the discrete Grönwall inequality further means

$$\|\bar{q}^n\|_{\mathcal{A},h} \le (2+4T)\exp(6K_f T)\left(\|q^0\|_{\mathcal{A},h} + \max_{1\le k\le n} \|\mathcal{D}_2 g^k\|_h\right). \tag{3.36}$$

Step III. Estimate for $\|\bar{u}^n\|_{h,\infty}$. Inserting (3.30) and (3.36) into (3.24), we obtain

$$\|\Lambda_h \bar{u}^n\|_h \le C\left(\|\varphi\|_h + \|f(\varphi)\|_h + \|q^0\|_{\mathcal{A},h} + \max_{1\le k\le n}\left(\|g^k\|_h + \|\mathcal{D}_2 g^k\|_h\right)\right). \tag{3.37}$$

Finally, combinations of (3.30), (3.37) and (3.12) based on Lemma 2.2 proves the conclusion. □

3.4. Convergence analysis

Denote $\bar{e}_u^n := U^n - \bar{u}^n$, $\bar{e}_q^n := Q^n - \bar{q}^n$, $e_u^n := U^n - u^n$ and $e_q^n := Q^n - q^n$ for $0 \le n \le N$. It is easy to see that $\bar{e}_u^0 = e_u^0 = 0$ and $\|\mathcal{A}_h \bar{e}_q^0\|_h = \|\mathcal{A}_h e_q^0\|_h = \|R_s^0\|_h \le C_7 h^4$.

Next, we shall first give an error estimate between the exact solution and the numerical solution yielded by the auxiliary scheme (3.15)–(3.16).

Remark 3.8. Theorem 3.7 indicates that the maximum-norm of $\bar{u}^n$ can be controlled by the discrete norms of g^k and φ. Furthermore, by the triangle inequality, it yields

$$\|\bar{e}_u^n\|_{h,\infty} \le \|U^n\|_{h,\infty} + \|\bar{u}^n\|_{h,\infty} \le C_{\bar{u}} \quad \text{for } 1 \le n \le N,$$

where $C_{\bar{u}} := \|U^n\|_{h,\infty} + K_1\left(\|\varphi\|_h + \|\Lambda_h\varphi\|_h + \|f(\varphi)\|_h + \max_{0\le k\le n}\|g^k\|_h + \max_{1\le k\le n}\|\mathcal{D}_2 g^k\|_h\right)$. In the following, we will provide further estimate of $\|\bar{e}_u^n\|_{h,\infty}$ to ensure $\|\bar{e}_u^n\|_{h,\infty} \le \delta$, which implies $\bar{u}^n \in \mathcal{B}_\delta$ and thus $\bar{f} \equiv f$.

Theorem 3.9. *Assume that the solution of* (1.1)–(1.3) *satisfy the regularity assumption* (3.1). *If the adjacent temporal stepsize ratio* $0 < r_k < 4.8645$ *and the maximum temporal stepsize* $\tau \le 1/(6(C_{\bar{u}}+1)K_f)$, *then there exists a positive constant* $K_2 = K_2(T, K_f, C_{\bar{u}}, C_0, C_5 - C_7)$, *such that the following estimates hold for the auxiliary nonlinear scheme* (3.15)–(3.16)

$$\|U^n - \bar{u}^n\|_{h,\infty} + \|Q^n - \bar{q}^n\|_h \le K_2\left(\tau^2 + h^4\right) \quad \text{for } 1 \le n \le N.$$

Proof. Subtracting (3.15)–(3.16) from (3.6)–(3.7), respectively, we get the following error equations

$$\bar{e}^n_{q,i,j} = \mathcal{D}_2\bar{e}^n_{u,i,j} + (R_t)^n_{i,j}, \quad (i,j) \in \omega_h, \tag{3.38}$$

$$\mathcal{A}_h\bar{e}^n_{q,i,j} - \Lambda_h\bar{e}^n_{u,i,j} = \mathcal{A}_h\left(f(U^n_{i,j}) - \bar{f}(\bar{u}^n_{i,j})\right) + (R_s)^n_{i,j}, \quad (i,j) \in \omega_h, \tag{3.39}$$

for $1 \le n \le N$. Note that (3.38)–(3.39) have basically the same form as (3.15)–(3.16). The main difference lies in the error term $\mathcal{A}_h\left(f(U^n_{i,j}) - \bar{f}(\bar{u}^n_{i,j})\right)$, which remains to be estimated.

Step I. Estimate for $\|\bar{e}^n_u\|_{\mathcal{A},h}$. Let $R^n = \mathcal{A}_h R_t + R_s$. We insert (3.38) into (3.39) to get

$$\mathcal{D}_2\mathcal{A}_h\bar{e}^n_{u,i,j} - \Lambda_h\bar{e}^n_{u,i,j} = \mathcal{A}_h\left(\bar{f}(U^n_{i,j}) - \bar{f}(\bar{u}^n_{i,j})\right) + R^n_{i,j}, \quad (i,j) \in \omega_h,$$

where we have used the fact that $\bar{f}(U^n_{i,j}) = f(U^n_{i,j})$ as the true solution $U^n_{i,j} \in \mathcal{B} \subset \mathcal{B}_\delta$. It follows from the global Lipschitz continuous property of $\bar{f}$ that

$$\|\bar{f}(U^n) - \bar{f}(\bar{u}^n)\|_h \le \frac{3}{2}K_f\|\bar{e}^n_u\|_{\mathcal{A},h}.$$

Consequently, analogous to the proof of (3.26)–(3.28) for the analysis of (3.25), we have

$$\|\bar{e}^n_u\|_{\mathcal{A},h} \le 3K_f\sum_{k=1}^n \tau_k\|\bar{e}^k_u\|_{\mathcal{A},h} + 2\sum_{k=1}^n\sum_{m=1}^k \theta^{(k)}_{k-m}\|R^m\|_h. \tag{3.40}$$

While, for the last term of (3.40), some special attention should be paid. We exchange the order of summation and utilize Lemma 3.3 to obtain

$$\begin{aligned}\sum_{k=1}^n\sum_{m=1}^k \theta^{(k)}_{k-m}\|R^m\|_h &= \sum_{m=1}^n \|R^m\|_h\sum_{k=m}^n \theta^{(k)}_{k-m} = \|R^1\|_h\sum_{k=1}^n\theta^{(k)}_{k-1} + \sum_{m=2}^n\|R^m\|_h\sum_{k=m}^n\theta^{(k)}_{k-m}\\ &\le 2\tau\|R^1\|_h + \max_{2\le m\le n}\|R^m\|_h\sum_{m=2}^n\sum_{k=m}^n\theta^{(k)}_{k-m} \le 2\tau\|R^1\|_h + t_n\max_{2\le m\le n}\|R^m\|_h,\end{aligned}$$

which, together with the triangle inequality and (3.8)–(3.9), gives

$$2\sum_{k=1}^n\sum_{m=1}^k\theta^{(k)}_{k-m}\|R^m\|_h \le C\left(\tau^2 + h^4\right).$$

Now, insert the above inequality into (3.40), and for $\tau \le 1/(6K_f)$, the application of discrete Grönwall inequality yields

$$\|\bar{e}^n_u\|_{\mathcal{A},h} \le C\left(\tau^2 + h^4\right). \tag{3.41}$$

Step II. Estimate for $\|\bar{e}^n_q\|_{\mathcal{A},h}$. We conclude from (3.38) *via* using the DOC kernel relation (3.3) that

$$\|\nabla_\tau\bar{e}^k_u\|_h \le \sum_{m=1}^k\theta^{(k)}_{k-m}\|\bar{e}^m_q\|_h + \sum_{m=1}^k\theta^{(k)}_{k-m}\|R^m_t\|_h \le \frac{3}{2}\sum_{m=1}^k\theta^{(k)}_{k-m}\|\bar{e}^m_q\|_{\mathcal{A},h} + \sum_{m=1}^k\theta^{(k)}_{k-m}\|R^m_t\|_h, \tag{3.42}$$

and analogous to the proof of (3.31)–(3.34), it follows from (3.39) that

$$\|\bar{e}^n_q\|^2_{\mathcal{A},h} \le \|\bar{e}^0_q\|^2_{\mathcal{A},h} + 2\sum_{k=1}^n\|\nabla_\tau\left(\bar{f}(U^k) - \bar{f}(\bar{u}^k)\right)\|_h\|\bar{e}^k_q\|_{\mathcal{A},h} + 3\sum_{k=1}^n\sum_{m=1}^k\theta^{(k)}_{k-m}\|\mathcal{D}_2R^m_s - \Lambda_hR^m_t\|_h\|\bar{e}^k_q\|_{\mathcal{A},h}. \tag{3.43}$$

Due to $\bar{f} \in C^2(\mathbb{R})$, we see

$$\begin{aligned}\nabla_\tau\left(\bar{f}(U^k)-\bar{f}(\bar{u}^k)\right) &= \bar{e}_u^k\int_0^1 \bar{f}'(U^k-s\bar{e}_u^k)ds - \bar{e}_u^{k-1}\int_0^1 \bar{f}'(U^{k-1}-s\bar{e}_u^{k-1})ds \\ &= \nabla_\tau\bar{e}_u^k\int_0^1 \bar{f}'(U^k-s\bar{e}_u^k)ds + \bar{e}_u^{k-1}\int_0^1 \left[\bar{f}'(U^k-s\bar{e}_u^k)-\bar{f}'(U^{k-1}-s\bar{e}_u^{k-1})\right]ds \\ &= \nabla_\tau\bar{e}_u^k\int_0^1 \bar{f}'(U^k-s\bar{e}_u^k)ds + \bar{e}_u^{k-1}\int_0^1 \bar{f}''(\mu(s))\left(\nabla_\tau U^k - s\nabla_\tau\bar{e}_u^k\right)ds,\end{aligned} \tag{3.44}$$

where $\mu(s)$ between $U^{k-1}-s\bar{e}_u^{k-1}$ and $U^k-s\bar{e}_u^k$. Thus, considering (3.17) and Remark 3.8, it means

$$\|\nabla_\tau\left(\bar{f}(U^k)-\bar{f}(\bar{u}^k)\right)\|_h \le C\tau_k\|\bar{e}_u^{k-1}\|_h + (C_{\bar{u}}+1)K_f\|\nabla_\tau\bar{e}_u^k\|_h.$$

Therefore, substituting (3.42) into the above inequality, and then inserting the conclusion into (3.43), we get

$$\begin{aligned}\|\bar{e}_q^n\|_{\mathcal{A},h}^2 \le \|\bar{e}_q^0\|_{\mathcal{A},h}^2 &+ 3(C_{\bar{u}}+1)K_f\sum_{k=1}^n\|\bar{e}_q^k\|_{\mathcal{A},h}\sum_{m=1}^k\theta_{k-m}^{(k)}\|\bar{e}_q^m\|_{\mathcal{A},h} + C\sum_{k=1}^{n-1}\tau_k\|\bar{e}_u^k\|_{\mathcal{A},h}\|\bar{e}_q^k\|_{\mathcal{A},h} \\ &+ C\sum_{k=1}^n\sum_{m=1}^k\theta_{k-m}^{(k)}\left(\|\mathcal{D}_2R_s^m\|_h + \|\Lambda_hR_t^m\|_h + \|R_t^m\|_h\right)\|\|\bar{e}_q^k\|_{\mathcal{A},h}.\end{aligned}$$

Next, similar to the derivation of (3.36), utilizing (3.8)–(3.9) and the result (3.41), we have for $\tau \le 1/(6(C_{\bar{u}}+1)K_f)$,

$$\|\bar{e}_q^n\|_{\mathcal{A},h} \le 2\|\bar{e}_q^0\|_{\mathcal{A},h} + 6(C_{\bar{u}}+1)K_f\sum_{k=1}^{n-1}\tau_k\|\bar{e}_q^k\|_{\mathcal{A},h} + C\left(\tau^2+h^4\right).$$

Consequently, the application of discrete Grönwall inequality yields

$$\|\bar{e}_q^n\|_{\mathcal{A},h} \le C\left(\tau^2+h^4\right). \tag{3.45}$$

Step III. Estimate for $\|\bar{e}_u^n\|_{h,\infty}$. It follows from (3.39), (3.41), (3.45) and the triangle inequality that

$$\|\Lambda_h\bar{e}_u^n\|_h \le C\left(\|\bar{e}_q^n\|_{\mathcal{A},h} + \|\bar{e}_u^n\|_{\mathcal{A},h} + \|R_s^n\|_h\right) \le C\left(\tau^2+h^4\right),$$

which, together with (3.41) and Lemma 2.2, completes the proof. □

Finally, we would like to give the L^∞ norm error estimate between the exact solution and the numerical solution yielded by the variable-step BDF2 compact difference scheme (3.10)–(3.13). Note that Theorem 3.9 shows that $\|\bar{e}_u^n\|_{h,\infty} \le K_2(\tau^2+h^4) \le \delta/2$, for τ and h sufficiently small. Thus, it holds that $\bar{u}_{i,j}^n \in \mathcal{B}_\delta$, which further implies $\bar{f}(\bar{u}_{i,j}^n) = f(\bar{u}_{i,j}^n)$ and thus in this context $\bar{u}_{i,j}^n \equiv u_{i,j}^n$. At this moment, the auxiliary nonlinear scheme (3.15)–(3.16) reduces to (3.10)–(3.11). Therefore, we summarize those conclusions given by Theorems 3.5, 3.6, 3.7 and 3.9 in the following theorem.

Theorem 3.10. *Under the conditions in Theorem 3.9 and if the stepsizes τ and h are sufficiently small, the nonlinear variable-step BDF2 compact difference scheme (3.10)–(3.13) admits a unique solution satisfying*

$$\begin{aligned}\|u^n\|_{h,\infty} &\le K_1\left(\|\varphi\|_h + \|\Lambda_h\varphi\|_h + \|f(\varphi)\|_h + \max_{0\le k\le n}\|g^k\|_h + \max_{1\le k\le n}\|\mathcal{D}_2g^k\|_h\right), \\ \|U^n-u^n\|_{h,\infty} &+ \|Q^n-q^n\|_h \le K_2\left(\tau^2+h^4\right),\end{aligned}$$

for $1 \le n \le N$.

4. An efficient variable-step two-grid compact difference scheme

In order to solve the nonlinear parabolic equation (1.1)–(1.3) efficiently, we shall propose an efficient two-grid compact difference algorithm based on the variable-step BDF2 formula and the piecewise bi-cubic Lagrange interpolation developed in Section 2. The new algorithm, which shall reduce the modeling of a large-scale nonlinear system (3.10)–(3.13) to a small-scale nonlinear system on the coarse grid and a large-scale linearized system on the fine grid, is defined as follows.

Step 1. On the coarse grid, solve a small-scale nonlinear compact difference scheme to find rough solutions $(u_H^n, q_H^n) = \{u_{H,i,j}^n, q_{H,i,j}^n\}$ by

$$q_{H,i,j}^n = \mathcal{D}_2 u_{H,i,j}^n, \quad (i,j) \in \omega_H, \tag{4.1}$$

$$\mathcal{A}_H q_{H,i,j}^n - c\Lambda_H u_{H,i,j}^n = \mathcal{A}_H f(u_{H,i,j}^n) + \mathcal{A}_H g_{i,j}^n, \quad (i,j) \in \omega_H, \tag{4.2}$$

subject to the initial and boundary conditions (3.12)–(3.13) defined on coarse grid.

Step 2. On the fine grid, solve a large-scale linearized compact difference scheme to produce corrected solutions $(u_h^n, q_h^n) = \{u_{h,i,j}^n, q_{h,i,j}^n\}$ based on the rough solutions (u_H^n, q_H^n) in Step 1 by

$$q_{h,i,j}^n = \mathcal{D}_2 u_{h,i,j}^n, \quad (i,j) \in \omega_h, \tag{4.3}$$

$$\mathcal{A}_h q_{h,i,j}^n - c\Lambda_h u_{h,i,j}^n = \mathcal{A}_h \mathcal{F}_{i,j}^n + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h, \tag{4.4}$$

subject to the initial and boundary conditions (3.12)–(3.13) defined on fine grid, where $\mathcal{F}_{i,j}^n$ represents a Newton linearization from coarse grid to fine grid defined as

$$\mathcal{F}_{i,j}^n := f\left(\Pi_H u_{H,i,j}^n\right) + f'\left(\Pi_H u_{H,i,j}^n\right)\left(u_{h,i,j}^n - \Pi_H u_{H,i,j}^n\right). \tag{4.5}$$

Remark 4.1. In fact, by eliminating the auxiliary variables $q_{H,i,j}^n$ and $q_{h,i,j}^n$, we can get an equivalent scheme to (4.1)–(4.4) for the primal variables $u_{H,i,j}^n$ and $u_{h,i,j}^n$ that

$$\mathcal{D}_2 \mathcal{A}_H u_{H,i,j}^n - c\Lambda_H u_{H,i,j}^n = \mathcal{A}_H f(u_{H,i,j}^n) + \mathcal{A}_H g_{i,j}^n, \quad (i,j) \in \omega_H, \tag{4.6}$$

$$\mathcal{D}_2 \mathcal{A}_h u_{h,i,j}^n - c\Lambda_h u_{h,i,j}^n = \mathcal{A}_h \mathcal{F}_{i,j}^n + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \omega_h. \tag{4.7}$$

In practical computation, we first solve a small-scale nonlinear equation (4.6) to obtain a rough solution u_H^n, and then solve a large-scale linear equation (4.7) to get an updated solution u_h^n. Meanwhile, q_H^n and q_h^n can be explicitly derived *via* (4.1) and (4.3), respectively.

On the coarse grid, denote $e_{u,H}^n = U^n - u_H^n$ and $e_{q,H}^n = Q^n - q_H^n$ for $0 \le n \le N$. Analogous to Theorem 3.10, we can immediately reach the following conclusion.

Theorem 4.2. *Under the conditions in Theorem 3.9 and if the stepsizes τ and H are sufficiently small, the nonlinear compact difference scheme (4.1)–(4.2) defined on the coarse grid admits a unique solution satisfying*

$$\|U^n - u_H^n\|_{H,\infty} + \|Q^n - q_H^n\|_H \le K_2\left(\tau^2 + H^4\right) \quad \text{for } 1 \le n \le N.$$

Based on Theorem 4.2 and Lemma 2.6, the following corollary can be derived.

Corollary 4.3. *Under the conditions in Theorem 4.2, the numerical solution u_H^n of the nonlinear compact difference scheme (4.1)–(4.2) satisfies*

$$\|U^n - \Pi_H u_H^n\|_{h,\infty} \le K_3\left(\tau^2 + H^4\right) \quad \text{for } 1 \le n \le N, \tag{4.8}$$

and consequently, $\|U^n - \Pi_H u_H^n\|_{h,\infty} \le \delta$ for sufficiently small τ and H, i.e., the interpolation solution $\Pi_H u_H^n$ is bounded on the fine grid and satisfies

$$\Pi_H u_H^n \in B_\delta. \tag{4.9}$$

Proof. It is clear that (4.9) is valid if (4.8) holds. Now, we perform the splitting

$$\|U^n - \Pi_H u_H^n\|_{h,\infty} \le \|U^n - \Pi_H U^n\|_{h,\infty} + \|\Pi_H U^n - \Pi_H u_H^n\|_{h,\infty},$$

where the first right-hand side term could be bounded as $\|U^n - \Pi_H U^n\|_{h,\infty} \le C_1 H^4$ *via* Lemma 2.3, and due to the linear property of the piecewise bi-cubic Lagrange interpolation operator with respect to the interpolated function, Lemma 2.6 and Theorem 4.2 give us

$$\|\Pi_H U^n - \Pi_H u_H^n\|_{h,\infty} = \|\Pi_H e_{u,H}^n\|_{h,\infty} \le C_4 \|e_H^n\|_{H,\infty} \le C_4 K_2 \left(\tau^2 + H^4\right),$$

which completes the proof of estimate (4.8) by the triangle inequality with $K_3 := C_1 + C_4 K_2$. □

At last, we shall give an error estimate for $e_{u,h}^n = U^n - u_h^n$ and $e_{q,h}^n = Q^n - q_h^n$ on the fine grid for the linearized compact difference scheme (4.3)–(4.4). For the sake of simplicity, below we denote $\eta^n := U^n - \Pi_H u_H^n$.

Lemma 4.4. *Suppose the conditions in Theorem 4.2 hold. Then there exist positive constants* $K_4 = K_4(K_f, K_3, |\Omega|)$, $K_5 = K_5(K_f, K_3)$ *and* $K_6 = K_6(K_f, K_3)$, *such that*

$$\|f(U^m) - \mathcal{F}^m\|_h \le K_f \|e_{u,h}^m\|_h + K_4 \left(\tau^2 + H^8\right), \tag{4.10}$$

$$\begin{aligned} \|\nabla_\tau \left(f(U^m) - \mathcal{F}^m\right)\|_h \le{}& K_f \|\nabla_\tau e_{u,h}^m\|_h + K_f(1+\delta)\|\nabla_\tau \eta^m\|_h \|e_{u,h}^m\|_{h,\infty} \\ &+ K_5 \left(\tau^2 + H^4\right) \|\nabla_\tau \eta^m\|_h + K_6 \tau_m \left(\|e_{u,h}^m\|_h + \tau^2 + H^8\right), \end{aligned} \tag{4.11}$$

for $1 \le m \le N$.

Proof. First, we apply Taylor expansion of $f(U^m)$ at $\Pi_H u_H^m$ to obtain

$$f(U^m) = f\left([\Pi_H u_H^m]\right) + f'\left(\Pi_H u_H^m\right) \eta^m + \int_{\Pi_H u_H^m}^{U^m} f''(s)\left(U^m - s\right) ds.$$

Then, subtract $\mathcal{F}^m$ in (4.5) from this equation, we have

$$f(U^m) - \mathcal{F}^m = f'\left(\Pi_H u_H^m\right) e_{u,h}^m + \int_{\Pi_H u_H^m}^{U^m} f''(s)\left(U^m - s\right) ds. \tag{4.12}$$

Therefore, we apply Corollary 4.3 to derive

$$\|f(U^m) - \mathcal{F}^m\|_h \le K_f \left(\|e_{u,h}^m\|_h + \|\eta^m\|_{h,\infty}\|\eta^m\|_h\right) \le K_f \|e_{u,h}^m\|_h + K_4 \left(\tau^4 + H^8\right).$$

Next, we further conclude from (4.12) that

$$\nabla_\tau \left(f(U^m) - \mathcal{F}^m\right) = \nabla_\tau \left(f'\left(\Pi_H u_H^m\right) e_{u,h}^m\right) + \nabla_\tau \int_{\Pi_H u_H^m}^{U^m} f''(s)\left(U^m - s\right) ds := I_1 + I_2. \tag{4.13}$$

For the first term of (4.13), by the Taylor expansion at U^m we get

$$\begin{aligned} I_1 &= \nabla_\tau f'\left(\Pi_H u_H^m\right) e_{u,h}^m + f'\left(\Pi_H u_H^{m-1}\right) \nabla_\tau e_{u,h}^m \\ &= \left(\nabla_\tau f'\left(U^m\right) - \nabla_\tau \int_{\Pi_H u_H^m}^{U^m} f''\left(s\right) ds\right) e_{u,h}^m + f'\left(\Pi_H u_H^{m-1}\right) \nabla_\tau e_{u,h}^m. \end{aligned}$$

Similar to (3.44), there exists $\hat{\mu}(s)$ between $U^{m-1} - s\eta^{m-1}$ and $U^m - s\eta^m$ such that

$$\nabla_\tau \int_{\Pi_H u_H^m}^{U^m} f''(s)\, ds = \nabla_\tau \eta^m \int_0^1 f''(U^m - s\eta^m) ds + \eta^{m-1} \int_0^1 f'''(\hat{\mu}(s)) \left(\nabla_\tau U^m - s\nabla_\tau \eta^m\right) ds,$$

which implies

$$\|I_1\|_h \le K_f \|\nabla_\tau e_{u,h}^m\|_h + K_f(1+\delta)\|\nabla_\tau \eta^m\|_h \|e_{u,h}^m\|_{h,\infty} + C\tau_m \|e_{u,h}^m\|_h. \tag{4.14}$$

Similarly, for I_2, there exists $\check{\mu}(s)$ between $U^{m-1} - s\eta^{m-1}$ and $U^m - s\eta^m$ such that

$$I_2 = \nabla_\tau \left(\eta^m\right)^2 \int_0^1 s f''\left(U^m - s\eta^m\right) ds + \left(\eta^{m-1}\right)^2 \int_0^1 f'''\left(\check{\mu}(s)\right) \left(s\nabla_\tau U^m - s^2 \nabla_\tau \eta^m\right) ds.$$

Consequently, we have

$$\|I_2\|_h \le K_f \|\nabla_\tau \eta^m\|_h \|\eta^m + \eta^{m-1}\|_{h,\infty} + C\left(\tau_m + \|\nabla_\tau \eta^m\|_h\right) \|\eta^{m-1}\|_{h,\infty}^2. \tag{4.15}$$

Therefore, inserting the estimates (4.14)–(4.15) into (4.13), and Corollary 4.3 implies the conclusion (4.11). □

Theorem 4.5. *Suppose the conditions in Theorem 4.2 hold. Furthermore, if the stepsizes $\tau \le \tau_0$ and $h < H$ are sufficiently small, then there exists a positive constant $K_7 = K_7(T, K_f, C_{\bar{u}}, C_0 - C_7)$, such that the following error estimate holds for the two-grid compact difference scheme (4.1)–(4.4)*

$$\|U^n - u_h^n\|_{h,\infty} + \|Q^n - q_h^n\|_h \le K_7\left(\tau^2 + h^4 + H^8\right) \quad \text{for } 1 \le n \le N.$$

Proof. From (3.6) and (4.1), we can easily get the following error equation

$$e_{q,H,i,j}^n = \mathcal{D}_2 e_{u,H,i,j}^n + (R_t)_{i,j}^n, \quad (i,j) \in \omega_H,$$

which, with a similar treatment to (3.42), implies

$$\|\nabla_\tau e_{u,H}^n\|_H \le \sum_{m=1}^n \theta_{n-m}^{(n)} \left(\|e_{q,H}^m\|_H + \|R_t^m\|_H\right) \le K_2 \tau_n \left(\tau^2 + H^4\right) + \sum_{m=1}^n \theta_{n-m}^{(n)} \|R_t^m\|_H, \tag{4.16}$$

for $1 \le n \le N$, where Lemma 3.3 and the estimate for $\|e_{q,H}^m\|_H$ in Theorem 4.2 have been used in the last inequailty. Similarly, it follows from (3.6) and (4.3) that

$$\|\nabla_\tau e_{u,h}^n\|_h \le \sum_{m=1}^n \theta_{n-m}^{(n)} \left(\|e_{q,h}^m\|_h + \|R_t^m\|_h\right) \le \frac{3}{2} \sum_{m=1}^n \theta_{n-m}^{(n)} \left(\|e_{q,h}^m\|_{\mathcal{A},h} + \|R_t^m\|_h\right). \tag{4.17}$$

Step I. Estimate for $\|e_{u,h}^n\|_{\mathcal{A},h}$. For the linearized scheme (4.3)–(4.4) on the fine grid, we can get two very similar error equations

$$e_{q,h,i,j}^n = \mathcal{D}_2 e_{u,h,i,j}^n + (R_t)_{i,j}^n, \quad (i,j) \in \omega_h, \tag{4.18}$$

$$\mathcal{A}_h e_{q,h,i,j}^n - \Lambda_h e_{u,h,i,j}^n = \mathcal{A}_h f(U_{i,j}^n) - \mathcal{A}_h \mathcal{F}_{i,j}^n + (R_s)_{i,j}^n, \quad (i,j) \in \omega_h. \tag{4.19}$$

Substituting (4.18) into (4.19) gives us

$$\mathcal{D}_2 \mathcal{A}_h e_{u,h,i,j}^n - \Lambda_h e_{u,h,i,j}^n = \mathcal{A}_h f(U_{i,j}^n) - \mathcal{A}_h \mathcal{F}_{i,j}^n + R_{i,j}^n, \quad (i,j) \in \omega_h,$$

which, together with a similar treatment to (3.25)–(3.27), leads to

$$\|e_{u,h}^n\|_{\mathcal{A},h}^2 \le 2\sum_{k=1}^n \sum_{m=1}^k \theta_{k-m}^{(k)} \left(\mathcal{A}_h f(U^m) - \mathcal{A}_h \mathcal{F}^m, e_{u,h}^k\right)_h + 2\sum_{k=1}^n \sum_{m=1}^k \theta_{k-m}^{(k)} \left(R^m, e_{u,h}^k\right)_h.$$

Furthermore, we apply Lemmas 2.1, 4.4 and use a similar approach as Step I of Theorem 3.9 to derive

$$\|e_{u,h}^n\|_{\mathcal{A},h} \le 6K_f \sum_{k=1}^{n-1} \tau_k \|e_{u,h}^k\|_{\mathcal{A},h} + C\left(\tau^2 + h^4 + H^8\right), \quad \text{for } \tau \le 1/(6K_f),$$

from which an application of the discrete Grönwall inequality yields the estimate

$$\|e_{u,h}^n\|_{\mathcal{A},h} \le C(\tau^2 + h^4 + H^8). \tag{4.20}$$

Step II. Estimate for $\|e_{q,h}^n\|_{\mathcal{A},h}$. Analogous to the proof of Step II in Theorem 3.9, it follows from (4.19) that

$$\begin{aligned} &\|e_{q,h}^n\|_{\mathcal{A},h}^2 \\ &\le \|e_{q,h}^0\|_{\mathcal{A},h}^2 + 2\sum_{k=1}^{n} \|\nabla_\tau \left(f(U^k) - \mathcal{F}^k\right)\|_h \|e_{q,h}^k\|_{\mathcal{A},h} + 3\sum_{k=1}^{n}\sum_{m=1}^{k} \theta_{k-m}^{(k)} \|\mathcal{D}_2 R_s^m - \Lambda_h R_t^m\|_h \|e_{q,h}^k\|_{\mathcal{A},h}. \end{aligned} \tag{4.21}$$

Compared with (3.43), the only difference lies in the nonlinear term $\|\nabla_\tau \left(f(U^k) - \mathcal{F}^k\right)\|_h$, which is estimated in Lemma 4.4. While the term $\|e_{u,h}^k\|_{h,\infty}$ in Lemma 4.4 remains to be analyzed. Actually, Lemma 2.2, (3.9), (4.10), (4.19) and (4.20) lead to

$$\begin{aligned} \|e_{u,h}^k\|_{h,\infty} &\le C_0\left(\|e_{u,h}^k\|_{\mathcal{A},h} + \|\Lambda_h e_{u,h}^k\|_h\right) \le C_0\left(\|\mathcal{A}_h e_{q,h}^k\|_h + \|e_{u,h}^k\|_{\mathcal{A},h} + \|R_s^k\|_h + \|f(U^k) - \mathcal{F}^k\|_h\right) \\ &\le \frac{3C_0}{2}\|e_{q,h}^k\|_{\mathcal{A},h} + C\left(\tau^2 + h^4 + H^8\right). \end{aligned} \tag{4.22}$$

Now, let $\|e_{q,h}^{n^*}\|_{\mathcal{A},h} = \max_{1\le k\le n} \|e_{q,h}^k\|_{\mathcal{A},h}$. By inserting (4.17) and (4.22) into (4.11) for the estimate $\|\nabla_\tau \left(f(U^k) - \mathcal{F}^k\right)\|_h$ in (4.21), and applying Lemma 3.3 we derive

$$\begin{aligned} \|e_{q,h}^n\|_{\mathcal{A},h} \le \|e_{q,h}^{n^*}\|_{\mathcal{A},h} &\le \|e_{q,h}^0\|_{\mathcal{A},h} + 3K_f \sum_{k=1}^{n} \tau_k \|e_{q,h}^k\|_{\mathcal{A},h} + 3\left(1+\delta\right) C_0 K_f \sum_{k=1}^{n} \|\nabla_\tau \eta^k\|_h \|e_{q,h}^k\|_{\mathcal{A},h} \\ &\quad + C\left(\tau^2 + H^4\right)\sum_{k=1}^{n} \|\nabla_\tau \eta^k\|_h + C\sum_{k=1}^{n} \tau_k \left(\|e_{u,h}^k\|_{\mathcal{A},h} + \tau^4 + H^8\right) \\ &\quad + C\sum_{k=1}^{n}\sum_{m=1}^{k} \theta_{k-m}^{(k)} \left(\|\mathcal{D}_2 R_s^m\|_h + \|\Lambda_h R_t^m\|_h + \|R_t^m\|_h\right) \\ &\le \|e_{q,h}^0\|_{\mathcal{A},h} + 3K_f \sum_{k=1}^{n} \tau_k \|e_{q,h}^k\|_{\mathcal{A},h} + 3\left(1+\delta\right) C_0 K_f \sum_{k=1}^{n} \|\nabla_\tau \eta^k\|_h \|e_{q,h}^k\|_{\mathcal{A},h} \\ &\quad + C\left(\tau^2 + H^4\right)\sum_{k=1}^{n} \|\nabla_\tau \eta^k\|_h + C\left(\tau^2 + h^4 + H^8\right), \end{aligned} \tag{4.23}$$

in which (3.8)–(3.9) and (4.20) have been used in the last step.

Next, we pay special attention on the estimate of $\|\nabla_\tau \eta^k\|_h$ in (4.23). By the triangle inequality, Lemmas 2.3, 2.5 and (4.16) we get

$$\begin{aligned} \|\nabla_\tau \eta^k\|_h &\le \|\nabla_\tau \left(U^k - \Pi_H U^k\right)\|_h + \|\nabla_\tau \Pi_H e_{u,H}^k\|_h \le C_2 \tau_k H^4 + C_3 \|\nabla_\tau e_{u,H}^k\|_H \\ &\le C_2 \tau_k H^4 + C_3 \sum_{m=1}^{k} \theta_{k-m}^{(k)} \left(\|e_{q,H}^m\|_H + \|R_t^m\|_H\right), \end{aligned}$$

which, by Lemma 3.3, Theorem 4.2 and (3.8), further implies

$$\|e_{q,h}^n\|_{\mathcal{A},h} \le \|e_{q,h}^0\|_{\mathcal{A},h} + 3K_f\left(1+(1+\delta)C_0C_{11}\right)\sum_{k=1}^{n}\tau_k\|e_{q,h}^k\|_{\mathcal{A},h} + C\left(\tau^2+h^4+H^8\right),$$

with $C_{11} := C_2 + C_3(2K_2 + C_5 + C_6)$. Then, an application of the discrete Grönwall inequality for $\tau \le \tau_0 := 1/\left(6K_f\left(1+(1+\delta)C_0C_{11}\right)\right)$, gives us

$$\|e_{q,h}^n\|_{\mathcal{A},h} \le C\left(\tau^2+h^4+H^8\right). \tag{4.24}$$

Step III. Estimate for $\|e_{u,h}^n\|_{h,\infty}$. It follows from (4.24) and (4.22) that $\|e_{u,h}^n\|_{h,\infty} \le C\left(\tau^2+h^4+H^8\right)$. □

5. Extension to periodic boundary condition

In this section, we extend the ideas and derivations in previous sections to the semilinear parabolic equation (1.1)–(1.3) with periodic boundary condition. Firstly, we denote the following periodic spaces of grid functions on grids $\bar{\omega}_\kappa$

$$\mathcal{V}_\kappa^p = \{v | v \in \mathcal{V}_\kappa \ \text{ and } \ v \text{ is periodic}\}.$$

Furthermore, for any grid functions $w, q \in \mathcal{V}_\kappa^p$, the discrete inner product and corresponding norms are redefined as

$$(w,q)_\kappa = \kappa_x\kappa_y\sum_{i=1}^{N_x^\kappa}\sum_{j=1}^{N_y^\kappa} w_{i,j}q_{i,j}, \quad \|w\|_\kappa = \sqrt{(w,w)_\kappa}, \quad \|w\|_{\mathcal{A},\kappa} = \sqrt{(\mathcal{A}_\kappa w, w)_\kappa},$$

and the following lemma also holds.

Lemma 5.1 ([29]). *For any $w \in \mathcal{V}_\kappa^p$, we have $\frac{4}{9}\|w\|_\kappa \le \|\mathcal{A}_\kappa w\|_\kappa \le \|w\|_\kappa$. and $\frac{2}{3}\|w\|_\kappa \le \|w\|_{\mathcal{A},\kappa} \le \|w\|_\kappa$.*

In the context of periodic boundary case, we still adopt the piecewise bi-cubic Lagrange interpolation operator defined in Section 2 to construct the high-order two-grid difference scheme. At this moment, Lemmas 2.5–2.6 are still valid with small modifications of the proof. However, for simplicity of presentation, below we only show the conclusions without proof.

Lemma 5.2. *For any $w \in \mathcal{V}_H^p$, the following estimates hold*

$$\|\Pi_H w\|_h \le C_3\|w\|_H, \quad \|\Pi_H w\|_{h,\infty} \le C_4\|w\|_{H,\infty}.$$

Now, an efficient two-grid fourth-order compact difference scheme for model (1.1)–(1.3) under periodic boundary condition is proposed similarly as follows.

Step 1. On the coarse grid, solve a small-scale nonlinear compact finite difference scheme to find rough solutions $(u_H^n, q_H^n) = \{u_{H,i,j}^n, q_{H,i,j}^n\}$ by

$$q_{H,i,j}^n = \mathcal{D}_2 u_{H,i,j}^n, \quad (i,j) \in \bar{\omega}_H, \tag{5.1}$$

$$\mathcal{A}_H q_{H,i,j}^n - c\Lambda_H u_{H,i,j}^n = \mathcal{A}_H f(u_{H,i,j}^n) + \mathcal{A}_H g_{i,j}^n, \quad (i,j) \in \bar{\omega}_H, \tag{5.2}$$

subject to the initial condition (3.12) and periodic boundary condition defined on coarse grid.

Step 2. On the fine grid, solve a large-scale linearized compact difference scheme to produce corrected solutions $(u_h^n, q_h^n) = \{u_{h,i,j}^n, q_{h,i,j}^n\}$ based on the rough solutions (u_H^n, q_H^n) in Step 1 by

$$q_{h,i,j}^n = \mathcal{D}_2 u_{h,i,j}^n, \quad (i,j) \in \bar{\omega}_h, \tag{5.3}$$

$$\mathcal{A}_h q_{h,i,j}^n - c\Lambda_h u_{h,i,j}^n = \mathcal{A}_h F_{i,j}^n + \mathcal{A}_h g_{i,j}^n, \quad (i,j) \in \bar{\omega}_h, \tag{5.4}$$

subject to the initial condition (3.12) and periodic boundary condition defined on fine grid.

Following the proofs of Corollary 4.3, Theorems 4.2 and 4.5, together with Lemmas 5.1–5.2, the unique solvability and error estimates for the two-grid algorithm (5.1)–(5.4) can be proved very similarly, and we skip the detailed proof here.

Theorem 5.3. *Assume that the solution of* (1.1)–(1.3) *satisfy the regularity assumption* (3.1)*. If the adjacent temporal stepsize ratio* $0 < r_k < 4.8645$ *and the stepsizes* τ *and* $h < H$ *are sufficiently small, the two-grid compact difference scheme* (5.1)–(5.4) *admits a unique solution satisfying*

$$\|U^n - u_h^n\|_{h,\infty} + \|Q^n - q_h^n\|_h \le K_7 \left(\tau^2 + h^4 + H^8\right) \quad \text{for } 1 \le n \le N.$$

6. Numerical examples

In this section, we shall present several numerical experiments to test the effectiveness and efficiency of the variable-step two-grid compact difference scheme. In the computation, a Newton-type iterative procedure with tolerance error 1.0×10^{-13} is performed to solve the nonlinear algebra systems at each time level.

6.1. Accuracy and numerical stability tests on uniform temporal grids

In this subsection, we shall compare the numerical accuracy of the two-grid compact difference scheme (4.1)–(4.4) (or (4.6)–(4.7)) with the standard nonlinear scheme (3.14) as well as the following implicit-explicit scheme

$$\mathcal{D}_2 \mathcal{A}_h u_{i,j}^n - c\Lambda_h u_{i,j}^n = \mathcal{A}_h f(u_{i,j}^{n,*}) + \mathcal{A}_h g_{i,j}^n, \tag{6.1}$$

where

$$u^{n,*} := \begin{cases} 2u^{n-1} - u^{n-2}, & n \ge 2, \\ u^0, & n = 1. \end{cases}$$

For this purpose, we consider the following semilinear parabolic equation

$$u_t - \Delta u = u - u^3 + g, \quad (x,y) \in (0,1)^2, \ t \in (0,\pi],$$

where the linear part $g(x,y,t)$ is given such that the exact solution is one of the following three types, *i.e.*,

- Case I: $u(x,y,t) = [5\sin(t) + 2\sin(5t)]\sin(2\pi x)\sin(2\pi y)$;
- Case II: $u(x,y,t) = [10\sin(t) + 5\sin(2t) + 2\sin(5t) + \sin(10t)]\sin(2\pi x)\sin(2\pi y)$;
- Case III: $u(x,y,t) = [10\sin(t) + 50\sin(2t) + 30\sin(5t) + 10\sin(10t)]\sin(2\pi x)\sin(2\pi y)$.

Figure 1 displays the evolution of these solutions with respect to (w.r.t.) time, in which it can be clearly observed that the solution in Case I changes most smoothly while the solution in Case III changes most sharply w.r.t. time.

Firstly, we set $N_x^h = N_y^h$ and $M_x = M_y = 10$ and adjust N and N_x^h to investigate the spatial convergence of these three methods. Numerical results for Cases I–III are listed in Tables 1–3 respectively, which indicates the fourth-order spatial accuracy for both the nonlinear method and the two-grid method. But the implicit-explicit method is only successfully implemented for Case I and fails for the other two cases, which may be caused by improper treatment of the time dependence and nonlinearity.

Secondly, to test the temporal convergence rates, we fix $N_x^h = N_y^h = 10N_x^H = 10N_y^H = 300$ and present the numerical results w.r.t. N in Table 4 for Case I. We can observe that these three methods all have second-order temporal accuracy as proved. However, when the solution changes dramatically over time (*e.g.* Case II or III), the implicit-explicit method becomes unstable ($|u| \to \infty$), while both the nonlinear method and two-grid method can generate the desired numerical solutions with the same magnitude accuracy, as seen in Tables 5–6.

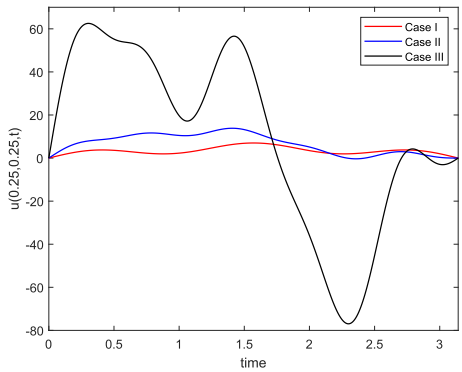


FIGURE 1. Evolution of exact solution w.r.t. time at point $(0.25, 0.25)$.

TABLE 1. Spatial convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case I on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
(N, N_x^h)	Error	Order	Error	Order	Error	Order
$(80, 100)$	1.63×10^{-3}	–	1.63×10^{-3}	–	1.88×10^{-3}	3
$(180, 150)$	3.28×10^{-4}	3.96	3.28×10^{-4}	3.96	3.54×10^{-4}	4.12
$(320, 200)$	1.04×10^{-4}	3.99	1.04×10^{-4}	3.99	1.10×10^{-4}	4.07
$(500, 250)$	4.27×10^{-5}	3.99	4.27×10^{-5}	3.99	4.45×10^{-5}	4.04

TABLE 2. Spatial convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case II on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
(N, N_x^h)	Error	Order	Error	Order	Error	Order
$(80, 100)$	4.47×10^{-3}	–	4.47×10^{-3}	–	Inf	–
$(180, 150)$	9.62×10^{-4}	3.79	9.62×10^{-4}	3.79	Inf	–
$(320, 200)$	3.10×10^{-4}	3.93	3.10×10^{-4}	3.93	Inf	–
$(500, 250)$	1.28×10^{-4}	3.97	1.28×10^{-4}	3.97	1.35×10^{-4}	–

TABLE 3. Spatial convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case III on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
(N, N_x^h)	Error	Order	Error	Order	Error	Order
$(80, 100)$	3.68×10^{-2}	–	3.68×10^{-2}	–	Inf	–
$(180, 150)$	8.05×10^{-3}	3.75	8.05×10^{-3}	3.75	Inf	–
$(320, 200)$	2.61×10^{-3}	3.92	2.61×10^{-3}	3.92	Inf	–
$(500, 250)$	1.08×10^{-3}	3.96	1.08×10^{-3}	3.96	Inf	–

TABLE 4. Temporal convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case I on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
N	Error	Order	Error	Order	Error	Order
32	1.55×10^{-2}	–	1.55×10^{-2}	–	1.05×10^{-2}	–
64	4.04×10^{-3}	1.98	4.04×10^{-3}	1.98	2.97×10^{-3}	1.82
128	1.03×10^{-3}	1.93	1.03×10^{-3}	1.93	7.12×10^{-4}	2.06
256	2.58×10^{-4}	1.99	2.58×10^{-4}	1.99	1.73×10^{-4}	2.05

TABLE 5. Temporal convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case II on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
N	Error	Order	Error	Order	Error	Order
128	1.86×10^{-3}	–	1.86×10^{-3}	–	Inf	–
256	4.82×10^{-4}	1.95	4.82×10^{-4}	1.95	Inf	–
512	1.22×10^{-4}	1.98	1.22×10^{-4}	1.98	1.29×10^{-4}	–
1024	3.07×10^{-5}	1.99	3.07×10^{-5}	1.99	3.21×10^{-5}	2.01

TABLE 6. Temporal convergence of nonlinear scheme (3.14), two-grid scheme (4.6)–(4.7) and implicit-explicit scheme (6.1) for Case III on uniform temporal grids.

	Nonlinear scheme		Two-grid scheme		Implicit-explicit scheme	
N	Error	Order	Error	Order	Error	Order
128	9.29×10^{-3}	–	9.29×10^{-3}	–	Inf	–
256	2.52×10^{-3}	1.88	2.52×10^{-3}	1.88	Inf	–
512	6.46×10^{-4}	1.97	6.46×10^{-4}	1.97	Inf	–
1024	1.63×10^{-4}	1.99	1.63×10^{-4}	1.99	Inf	–

Furthermore, the numerical results show that if we further refine the temporal grids (*e.g.*, changing N from 256 to 512 in Tab. 5), the implicit-explicit method may produce a correct result. But as shown in Table 6, its stability requirement for the temporal grid is quite related to the smoothness of the solution w.r.t. time. It is seen that the implicit-explicit discretization has much more strict stability condition compared to the other two methods when the solution u changes sharply w.r.t. time, for which a convincing explanation is that approximating the nonlinear term *via* solutions at previous time levels may leads to inaccuracy in this context.

6.2. Accuracy and efficiency tests on variable-step temporal grids

To check the accuracy and efficiency on variable-step temporal grids, we consider model (1.1)–(1.3) on $(0,1)^2 \times (0,1]$ with $c = \frac{1}{8\pi^2}$ and $f(u) = u - u^3$. The linear part g is determined such that the exact solution is $u(x,y,t) = \sin(t)\,(\sin(2\pi x) + 0.5\sin(6\pi x))\,(\sin(2\pi y) + 0.5\sin(6\pi y))$. The variable-step temporal grids are generated randomly by

$$\tau_k := T\frac{\theta_k}{S}, \quad \text{with } S = \sum_{k=1}^{N} \theta_k,$$

TABLE 7. Spatial convergence of nonlinear scheme (3.14) and two-grid scheme (4.6)–(4.7) on variable-step temporal grids.

		Nonlinear scheme		Two-grid scheme	
(N, N_x^h)	$\max r_k$	Error	Order	Error	Order
$(80, 100)$	4.3125	3.94×10^{-5}	–	3.13×10^{-5}	–
$(180, 150)$	4.4777	8.22×10^{-6}	3.87	7.38×10^{-6}	3.57
$(320, 200)$	4.5252	2.40×10^{-6}	4.28	2.38×10^{-6}	3.93
$(500, 250)$	4.5481	9.61×10^{-7}	4.10	9.57×10^{-7}	4.08

TABLE 8. Temporal convergence of nonlinear scheme (3.14) and two-grid scheme (4.6)–(4.7) on variable-step temporal grids.

		Nonlinear scheme		Two-grid scheme	
(N, N_x^h)	$\max r_k$	Error	Order	Error	Order
$(40, 40)$	3.9128	4.20×10^{-5}	–	1.03×10^{-4}	–
$(80, 80)$	3.9673	9.38×10^{-6}	2.16	9.27×10^{-6}	3.47
$(160, 160)$	4.4221	4.04×10^{-6}	2.08	4.04×10^{-6}	2.05
$(320, 320)$	4.5826	2.27×10^{-6}	2.01	2.27×10^{-6}	2.01

TABLE 9. Errors and CPU times of nonlinear scheme (3.14) and two-grid scheme (4.6)–(4.7) on variable-step temporal grids.

	Nonlinear scheme		Two-grid scheme $M_x = M_y = 8$		Two-grid scheme $M_x = M_y = 16$	
(N, N_x^h)	Error	CPU times	Error	CPU times	Error	CPU times
(240, 240)	4.03×10^{-6}	12 m 19 s	3.84×10^{-6}	3 m 48 s	1.89×10^{-4}	2 m 56 s
(320, 320)	2.26×10^{-6}	27 m 3 s	2.24×10^{-6}	9 m 51 s	1.57×10^{-5}	8 m 5 s
(400, 400)	1.46×10^{-6}	48 m 48 s	1.46×10^{-7}	19 m 30 s	2.51×10^{-6}	16 m 39 s
(480, 480)	9.95×10^{-7}	1 h 18 m 18 s	9.95×10^{-7}	32 m 31 s	9.98×10^{-7}	28 m 21 s

where θ_k is randomly drawn from the uniform distribution on the interval $(1/4.8645, 1)$ such that the adjacent temporal stepsize ratio $r_k < 4.8645$.

As the implicit-explicit scheme (6.1) may generate wrong results, we only test the nonlinear scheme (3.14) and two-grid scheme (4.6)–(4.7). We firstly test the errors and convergence rates in spatial and temporal directions for both methods with $N_x^h = N_y^h$ and $M_x = M_y = 5$. The corresponding numerical results are listed in Tables 7–8 respectively, which indicates the fourth-order accuracy in space and second-order accuracy in time as proved in Theorems 3.10 and 4.5. Moreover, we compare the CPU times consumed by the two methods in Table 9 for various $M_x = M_y$, in which we can clearly observe that (i) the proposed two-grid method has significantly improved the computational efficiency, for example, it takes more than one hour for the implementation of the nonlinear scheme when $N = N_x^h = 480$, while the two-grid scheme with $M_x = 8$ consumes only about half an hour to desire the same error; (ii) when increasing the number M_x (*i.e.*, reducing the scale of nonlinear systems on coarse grid), the computational efficiency can be further improved but correspondingly brings larger errors.

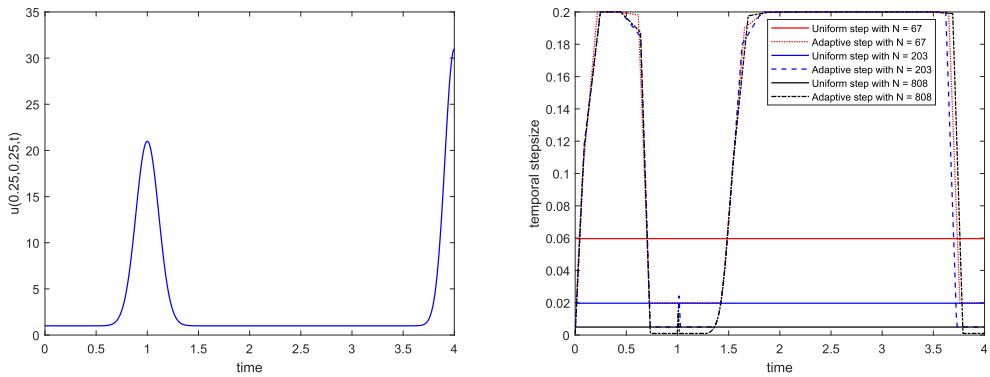


FIGURE 2. Evolution of exact solution w.r.t. time at $(0.25, 0.25)$ (left) and temporal stepsizes of the two-grid scheme and adaptive two-grid scheme (right).

6.3. Effectiveness of adaptive temporal stepsize strategy

In this test, we consider model (1.1)–(1.3) on $(0,1)^2 \times (0,4]$ with $c = 1$ and $f(u) = \sin u$, and the exact solution is chosen as $u(x,y,t) = \left[1 + 20e^{-40(t-1)^2} + 30e^{-60(t-4)^2}\right] \sin(2\pi x)\sin(2\pi y)$. Figure 2 (left) depicts the evolution of solution w.r.t. time at fixed point $(0.25, 0.25)$, which consists of two peaks and admits multiple time scales. Therefore, the variable-step two-grid scheme based on the adaptive temporal stepsize strategy [20, 32] will be adopted to improve the temporal accuracy

$$\tau_{n+1} = \min\left\{\max\left\{\tau_{\min}, \frac{\tau_{\max}}{\sqrt{1+\eta\|\partial_\tau u^n\|_h^2}}\right\}, r_{\max}\tau_n\right\}, \tag{6.2}$$

where $\partial_\tau u^n := \nabla_\tau u^n/\tau_n$ and $r_{\max} = 4.8$ which satisfies the restrictions in Theorems 3.10 and 4.5. Here $\tau_{\min}$ and $\tau_{\max}$ are the pre-determined minimum and maximum temporal stepsize and η is a pre-chosen parameter. In this example, we uniformly set $\tau_{\max} = 0.2$, $\eta = 500$ and gradually reduce $\tau_{\min}$ to generate grids with distinct stepsizes.

We select $N_x^h = N_y^h = 250$ to test the temporal convergence rates yielded by the two-grid scheme on uniform and adaptive temporal grids with $M_x = M_y = 10$. The numerical results are presented in Table 10, which shows that the standard two-grid scheme is second-order accurate in time, while the variable-step two-grid scheme based on the adaptive temporal stepsize strategy (6.2) has better convergence and much smaller errors with the same number of grids. In other words, the standard two-grid scheme on uniform temporal grids requires more temporal steps, and of course more CPU times, to generate numerical solutions with the same magnitude accuracy as adaptive method which uses less temporal steps. The reason can be more intuitively observed from Figure 2 (right), which shows that when the solution varies sharply, small temporal stepsizes are adaptively created to capture the fast evolution process, while otherwise large temporal stepsizes are generated to accelerate the time integration.

6.4. Application to phase-field Allen–Cahn equation

In this subsection, we consider the following Allen–Cahn equation with a polynomial double-well potential, subject to periodic boundary condition

$$u_t - \varepsilon^2\Delta u = u - u^3, \quad (x,y) \in \Omega,\ t \in (0,T],$$

TABLE 10. Temporal convergence of two-grid scheme and adaptive two-grid scheme.

		Two-grid scheme			Adaptive two-grid scheme		
$\tau_{\min}$	N	CPU times	Error	Order	CPU times	Error	Order
0.02	67	27.96 s	6.12×10^{-1}	–	28.73 s	4.36×10^{-2}	–
0.01	114	45.63 s	1.98×10^{-1}	2.12	48.43 s	9.27×10^{-3}	2.91
0.005	203	1 m 21 s	5.05×10^{-2}	2.37	1 m 22 s	2.02×10^{-3}	2.64
0.002	489	3 m 13 s	6.25×10^{-3}	2.38	3 m 19 s	2.82×10^{-4}	2.24
0.001	808	5 m 30 s	2.00×10^{-3}	2.27	5 m 28 s	6.91×10^{-5}	2.80

where ε is the interaction length that describes the thickness of the transition boundary between materials. It is well known that the energy dissipation law [16, 41]

$$E[u](t) \leq E[u](s), \quad \forall t > s$$

holds for the Allen–Cahn equation, where $E[u](t)$ represents the Lyapunov energy functional, namely

$$E[u](t) = \int_\Omega \frac{\varepsilon^2}{2}|\nabla u|^2 + F(u)d\mathbf{x} \ \text{ with } \ F(u) = \frac{1}{4}(1-u^2)^2.$$

Moreover, it has been observed that the evolution of the energy $E[u](t)$ usually involves both fast and slow stages of change in the long time simulation. Thus, it is highly desirable for numerical methods to preserve the discrete energy dissipation law on the nonuniform temporal grids. Define the discrete energy functional $\mathcal{E}[u^n]$ as

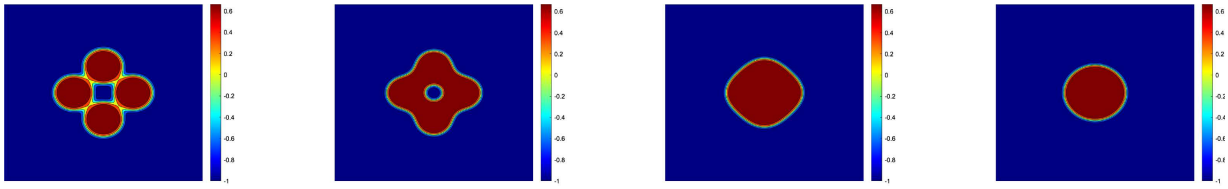
$$\mathcal{E}[u^n] := -\frac{\varepsilon^2}{2}h_x h_y \sum_{i=1}^{N_x^h}\sum_{j=1}^{N_y^h} u_{i,j}^n \Delta\Theta_h u_{i,j}^n + \frac{1}{4}h_x h_y \sum_{i=1}^{N_x^h}\sum_{j=1}^{N_y^h} \left(1-(u_{i,j}^n)^2\right)^2.$$

In this test, we would also compare effectiveness and efficiency of the high-order nonlinear difference scheme (3.14), the two-grid difference scheme (4.6)–(4.7) and the adaptive two-grid difference scheme using a similar adaptive temporal stepsize strategy (6.2) by replacing $\partial_\tau u^n$ with $\partial_\tau \mathcal{E}[u^n]$.

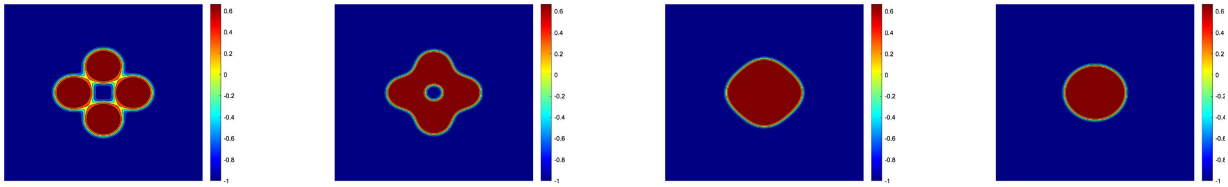
Example 6.1. In this example, we set $\varepsilon = 0.02$ and apply these three methods to simulate the merging of four bubbles with an initial condition

$$\begin{aligned}\varphi(x,y) = &-\tanh\left(\left((x-0.3)^2+y^2-0.2^2\right)/\varepsilon\right)\tanh\left(\left((x+0.3)^2+y^2-0.2^2\right)/\varepsilon\right)\\ &\times\tanh\left(\left(x^2+(y-0.3)^2-0.2^2\right)/\varepsilon\right)\tanh\left(\left(x^2+(y+0.3)^2-0.2^2\right)/\varepsilon\right).\end{aligned}$$

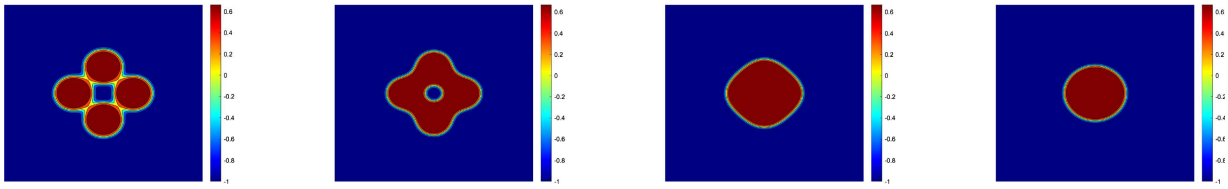
In this simulation, a 384×384 uniform mesh is taken to discretize the spatial domain $\Omega = (-1,1)^2$ and the ratio of coarse-fine grids are setted as $M_x = M_y = 3$. We start with the modeling of the solution by the nonlinear scheme and two-grid scheme with a constant temporal stepsize $\tau = 0.1$ until time $T = 100$. Parameters in the adaptive temporal stepsize strategy (6.2) are selected as $\tau_{\min} = 0.1$, $\tau_{\max} = 1$ and $\eta = 3200$. In Figure 3, it displays a comparison on the evolution of solution snapshots among the three methods, in which the gradually merging and shrinking process of the initial four-drops over time can be clearly observed. As can be seen in the figures, there seems no distinguishable differences among these methods. Next, we investigate the efficiency of the two-grid method on uniform grid and on adaptive grid for long time modeling. As seen in Figure 4 (left), the evolution of the free energy w.r.t. time for these three methods coincide, which consists very well with the corresponding results in [30]. Moreover, Table 11 indicates that the adaptive two-grid scheme has significant advantage in computational efficiency over the other two schemes. For example, it takes about 10 hours for the



A. Nonlinear scheme with fixed temporal stepsize $\tau = 0.1$



B. Two-grid scheme with fixed temporal stepsize $\tau = 0.1$



C. Adaptive two-grid scheme with $\tau_{\min} = 0.1$, $\tau_{\max} = 1$ and $\eta = 3200$

FIGURE 3. Solution snapshots of Allen–Cahn equation at $t = 1, 10, 50, 100$ (from left to right) yielded by nonlinear scheme, two-grid scheme and adaptive two-grid scheme.

TABLE 11. CPU times and the total number of temporal steps for three schemes.

	Nonlinear scheme		Two-grid scheme		Adaptive two-grid scheme	
T	N	CPU times	N	CPU times	N	CPU times
10	100	57 m 42 s	100	26 m 10 s	34	9 m 31 s
30	300	3 h 50 m 55 s	300	1 h 47 m 24 s	71	20 m 2 s
50	500	5 h 58 m 39 s	500	2 h 50 m 56 s	96	37 m 25 s
100	1000	9 h 43 m 51 s	1000	5 h 9 m 43 s	156	1 h 2 m 22 s

implementation of the nonlinear scheme up to $T = 100$, while the two-grid method using uniform temporal grid consumes about 5 hours. What is even more amazing is that the developed adaptive two-grid method using variable-step temporal grid takes only about one hour! In fact, the total number of adaptive temporal steps is only 156, while it takes 1000 steps for the uniform grid. Finally, the adaptive temporal stepsize curve of the adaptive two-grid method is plotted in Figure 4 (right), which also demonstrates the superiority of the variable-step two-grid compact difference scheme.

Example 6.2. In this example, we consider the coarsening process governed by the Allen–Cahn equation with the model parameter $\varepsilon = 0.01$ and computational domain $\Omega = (0, 1)^2$. Here we choose a random initial condition $\varphi(x, y) = -0.05 + 0.1 \times rand(x, y)$.

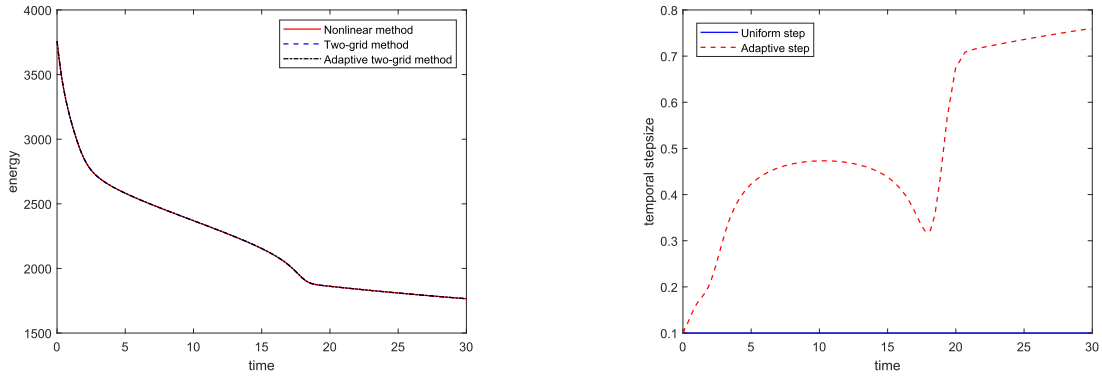
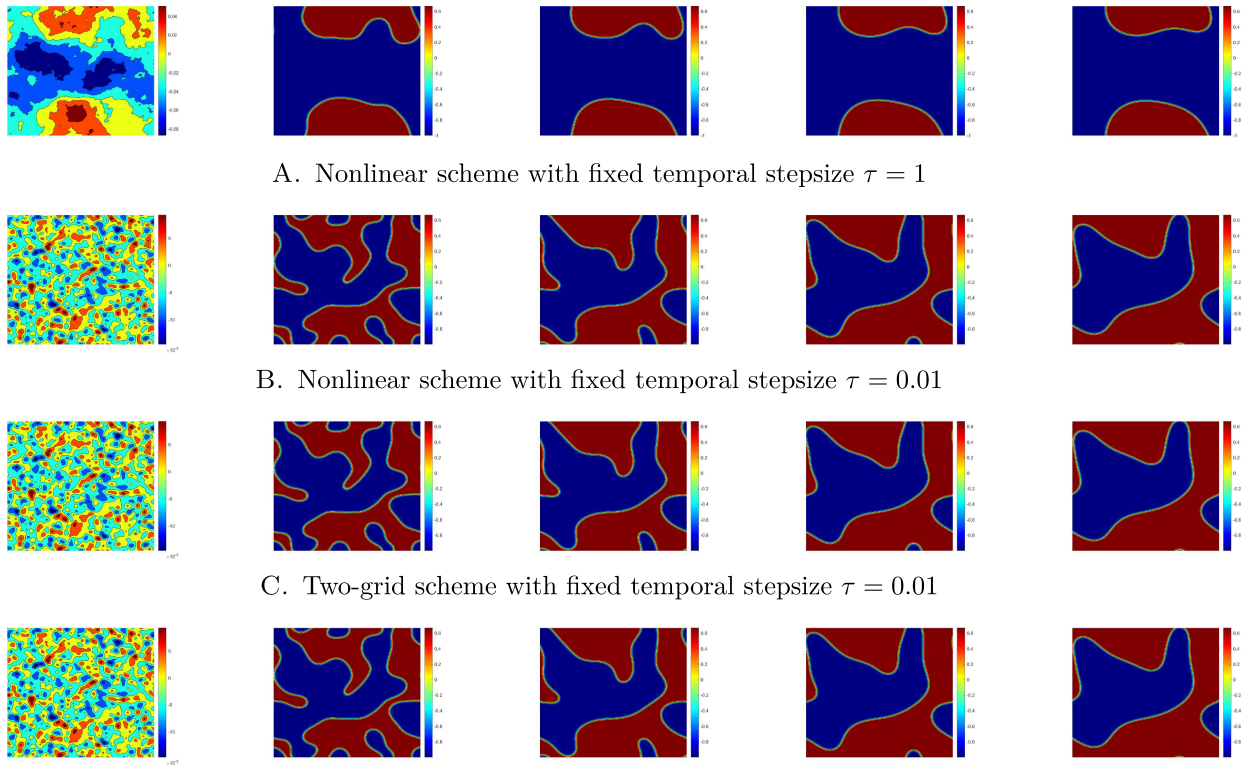


FIGURE 4. Evolutions of energy (left) and temporal stepsizes (right) for the nonlinear scheme, two-grid scheme and adaptive two-grid scheme until time $T = 30$.

This simulation is performed under $N_x^h = N_y^h = 384$ and $M_x = M_y = 3$. Due to the fact that the initial values are randomly given, we provide the startup values on coarse grid for the two-grid scheme by implementing the nonlinear algorithm up to $T = 0.5$. In Figure 5, it displays the evolution of the coarsening dynamic for nonlinear and two-grid compact schemes with different time strategies. It is observed that the nonlinear scheme with large uniform temporal stepsize $\tau = 1$ yields inaccurate solution u, while the adaptive two-grid scheme gives the correct coarsening pattern which is consistent with the results obtained by the nonlinear and two-grid methods with small uniform temporal stepsize $\tau = 0.01$. In Figure 6, we depict the evolution of discrete energies and temporal stepsizes w.r.t. time, which shows that the energy dissipation of the adaptive two-grid method agrees very well with the the nonlinear and two-grid methods using small uniform temporal stepsize. Moreover, the efficiency of the proposed two-grid scheme using the adaptive temporal stepsize strategy can also be seen from Figure 6 (right) and Table 12. For example, the adaptive two-grid method using variable-step temporal grid costs only 33 minutes for time marching to $T = 100$, while the two-grid method using uniform temporal stepsize $\tau = 0.01$ consumes more than 8 hours, even worse the implementation of the nonlinear scheme runs nearly 19 hours.

7. Concluding remarks

High-order two-grid difference scheme for nonlinear PDEs are rarely studied in existing literature due to, *e.g.*, the lack of the appropriate accuracy-preserving mapping operator. To address this issue, we introduce a piecewise bi-cubic Lagrange interpolation operator between two grids, and discuss its boundedness under L^2 and L^∞ norms. Moreover, to effectively solve the nonlinear PDEs whose solutions may admit multiple time scales, variable-step temporal discretization methods, in particular, the variable-step multistep methods are naturally and valuable to improve accuracy for stiff problems. However, its numerical analysis is much more challenging than the single-step methods. As an illustration, combined with the variable-step BDF2 formula, an efficient high-order two-grid difference method is developed for the semilinear parabolic equation with Dirichlet or periodic boundary conditions. The unique solvability of the nonlinear problem on coarse grid is shown by Browder's fixed point theorem. Moreover, with the help of DOC kernels, the boundedness of the high-order mapping operator, temporal-spatial error splitting method and a cut-off approach, unconditional and optimal-order error estimates for the two-grid method on both coarse and fine grids are rigourously proved under $r_k := \tau_k/\tau_{k-1} < 4.8645$, where the cut-off technique is used to reduce the regularity requirement on the nonlinear term f to the local Lipschitz continuous condition. Several numerical examples are carried out to confirm the theoretical findings.



A. Nonlinear scheme with fixed temporal stepsize $\tau = 1$

B. Nonlinear scheme with fixed temporal stepsize $\tau = 0.01$

C. Two-grid scheme with fixed temporal stepsize $\tau = 0.01$

D. Adaptive two-grid scheme with $\tau_{\min} = 0.01$, $\tau_{\max} = 1$ and $\eta = 8 \times 10^4$

FIGURE 5. Solution snapshots of coarsening dynamics for Allen–Cahn equation at $t = 1, 20, 50, 80, 100$ (from left to right) yielded by nonlinear scheme, two-grid scheme and adaptive two-grid scheme.

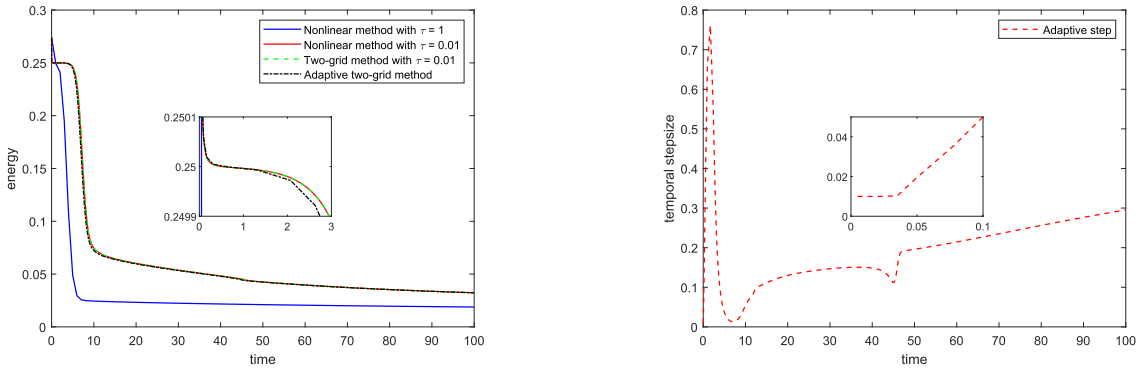


FIGURE 6. Evolutions of energy (left) and time steps (right) for the nonlinear scheme, two-grid scheme and adaptive two-grid scheme until time $T = 100$.

TABLE 12. CPU times and the total number of temporal steps for three schemes.

	Nonlinear scheme with $\tau = 0.01$		Two-grid scheme with $\tau = 0.01$		Adaptive two-grid scheme	
T	N	CPU times	N	CPU times	N	CPU times
20	2000	4 h 29 m 59 s	2000	2 h 12 m 22 s	384	19 m 53 s
50	5000	10 h 33 m 53 s	5000	4 h 38 m 8 s	588	27 m 40 s
80	8000	15 h 41 m 54 s	8000	6 h 41 m 51 s	722	31 m 35 s
100	10000	18 h 43 m 29 s	10000	8 h 4 m 26 s	795	32 m 57 s

Acknowledgements

The authors would like to express their most sincere thanks to the referees for their very helpful comments and suggestions, which greatly improved the quality of this paper. This work was supported in part by the National Natural Science Foundation of China (Nos. 11971482,12131014), by the Fundamental Research Funds for the Central Universities (No. 202264006) and by the OUC Scientific Research Program for Young Talented Professionals.

REFERENCES

[1] G. Akrivis, Finite difference discretization of the cubic Schrödinger equation. *IMA J. Numer. Anal.* **13** (1993) 115–124.

[2] G. Akrivis and S. Larsson, Linearly implicit finite element methods for the time-dependent Joule heating problem. *BIT Numer. Math.* **45** (2005) 429–442.

[3] G. Akrivis, M. Crouzeix and C. Makridakis, Implicit-explicit multistep methods for quasilinear parabolic equations. *Numer. Math.* **82** (1999) 521–541.

[4] W. Bao and Y. Cai, Uniform error estimates of finite difference methods for the nonlinear schrödinger equation with wave operator. *SIAM J. Numer. Anal.* **50** (2012) 492–521.

[5] J. Becker, A second order backward difference method with variable steps for a parabolic problem. *BIT Numer. Math.* **38** (1998) 644–662.

[6] Y. Chen, Y. Huang and D. Yu, A two-grid method for expanded mixed finite-element solution of semilinear reaction-diffusion equations. *Int. J. Numer. Meth. Eng.* **57** (2003) 193–209.

[7] C. Chen, K. Li, Y. Chen and Y. Huang, Two-grid finite element methods combined with Crank-Nicolson scheme for nonlinear Sobolev equations. *Adv. Comput. Math.* **45** (2019) 611–630.

[8] 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** (2019) 495–525.

[9] C.N. Dawson, M.F. Wheeler and C.S. Woodward, A two-grid finite difference scheme for nonlinear parabolic equations. *SIAM J. Numer. Anal.* **35** (1998) 435–452.

[10] J. de Frutos and J. Novo, A posteriori error estimation with the p-version of the finite element method for nonlinear parabolic differential equations. *Comput. Methods Appl. Mech. Eng.* **191** (2002) 4893–4904.

[11] Y. Di, Y. Wei, J. Zhang and C. Zhao, Sharp error estimate of an implicit BDF2 scheme with variable time steps for the phase field crystal model. *J. Sci. Comput.* **92** (2022) 65.

[12] Y. Di, Y. Ma, J. Shen and J. Zhang, A variable time-step IMEX-BDF2 SAV scheme and its sharp error estimate for the Navier-Stokes equations. *ESAIM: M2AN* **57** (2023) 1143–1170.

[13] Y. Du, Y. Liu, H. Li, Z. Fang and S. He, Local discontinuous Galerkin method for a nonlinear time-fractional fourth-order partial differential equation. *J. Comput. Phys.* **344** (2017) 108–126.

[14] F. Durango and J. Novo, Two-grid mixed finite-element approximations to the Navier-Stokes equations based on a Newton-type step. *J. Sci. Comput.* **74** (2018) 456–473.

[15] C. Gear and K. Tu, The effect of variable mesh size on the stability of multistep methods. *SIAM J. Numer. Anal.* **11** (1974) 1025–1043.

[16] D. Hou and Z. Qiao, An implicit-explicit second-order BDF numerical scheme with variable steps for gradient flows. *J. Sci. Comput.* **94** (2023) 39.

[17] T. Hou, T. Tang and J. Yang, Numerical analysis of fully discretized Crank-Nicolson scheme for fractional-in-space Allen–Cahn equations. *J. Sci. Comput.* **72** (2017) 1214–1231.

[18] T. Hou, W. Jiang, X. Yang and H. Leng, Two-grid P_1^2–P_1 mixed finite element methods combined with Crank-Nicolson scheme for a class of nonlinear parabolic equations. *Appl. Numer. Math.* **137** (2019) 136–150.

[19] 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. Comput.* **92** (2023) 2515–2542.

[20] J. Huang, C. Yang and Y. Wei, Parallel energy-stable solver for a coupled Allen–Cahn and Cahn–Hilliard system. *SIAM J. Sci. Comput.* **42** (2020) C294–C312.

[21] B. Li and W. Sun, Error analysis of linearized semi-implicit Galerkin finite element methods for nonlinear parabolic equations. *Int. J. Numer. Anal. Mod.* **10** (2013) 622–633.

[22] X. Li and H. Rui, A two-grid block-centered finite difference method for the nonlinear time-fractional parabolic equation. *J. Sci. Comput.* **72** (2017) 863–891.

[23] X. Li, J. Shen and H. Rui, Energy stability and convergence of SAV block-centered finite difference method for gradient flows. *Math. Comput.* **88** (2019) 2047–2068.

[24] B. Li, Y. Ueda and G. Zhou, A second-order stabilization method for linearizing and decoupling nonlinear parabolic systems. *SIAM J. Numer. Anal.* **58** (2020) 2736–2763.

[25] X. Li, Y. Chen and C. Chen, An improved two-grid technique for the nonlinear time-fractional parabolic equation based on the block-centered finite difference method. *J. Comput. Math.* **40** (2021) 455–473.

[26] D. Li, X. Li, M. Mei and W. Yuan, A structure-preserving and variable-step BDF2 Fourier pseudo-spectral method for the two-mode phase field crystal model. *Math. Comput. Simulat.* **205** (2023) 483–506.

[27] M. Li, L. Wang and N. Wang, Variable-time-step BDF2 nonconforming VEM for coupled Ginzburg-Landau equations. *Appl. Numer. Math.* **186** (2023) 378–410.

[28] H. Liao and Z. Zhang, Analysis of adaptive BDF2 scheme for diffusion equations. *Math. Comput.* **90** (2021) 1207–1226.

[29] H. Liao, Z. Sun and H. Shi, Error estimate of fourth-order compact scheme for linear Schrödinger equations. *SIAM J. Numer. Anal.* **47** (2010) 4381–4401.

[30] H. Liao, T. Tang and T. Zhou, On energy stable, maximum-principle preserving, second-order BDF scheme with variable steps for the Allen–Cahn equation. *SIAM J. Numer. Anal.* **58** (2020) 2294–2314.

[31] H. Liao, X. Song, T. Tang and T. Zhou, Analysis of the second-order BDF scheme with variable steps for the molecular beam epitaxial model without slope selection. *Sci. China. Math.* **63** (2021) 887–902.

[32] H. Liao, B. Ji, L. Wang and Z. Zhang, Mesh-robustness of an energy stable BDF2 scheme with variable steps for the Cahn–Hilliard model. *J. Sci. Comput.* **92** (2022) 52.

[33] H. Liao, B. Ji and L. Zhang, An adaptive BDF2 implicit time-stepping method for the phase field crystal model. *IMA J. Numer. Anal.* **42** (2022) 649–679.

[34] W. Liu and J. Cui, A two-grid block-centered finite difference algorithm for nonlinear compressible Darcy-Forchheimer model in porous media. *J. Sci. Comput.* **74** (2018) 1786–1815.

[35] Q. Liu, J. Jing, M. Yuan and W. Chen, A positivity-preserving, energy stable BDF2 scheme with variable steps for the Cahn–Hilliard equation with logarithmic potential. *J. Sci. Comput.* **95** (2023) 37.

[36] M. Marion and J. Xu, Error estimates on a new nonlinear Galerkin method based on two-grid finite elements. *SIAM J. Numer. Anal.* **32** (1995) 1170–1184.

[37] Z. Qiao, Z. Zhang and T. Tang, An adaptive time-stepping strategy for the molecular beam epitaxy models. *SIAM J. Sci. Comput.* **33** (2011) 1395–1414.

[38] A. Quarteroni, R. Sacco and F. Saleri, Numerical Mathematics. Springer (2007).

[39] H. Rui and W. Liu, A two-grid block-centered finite difference method for Darcy-Forchheimer flow in porous media. *SIAM J. Numer. Anal.* **53** (2015) 1941–1962.

[40] L. Shampine and M. Reichelt, The MATLAB ODE suite. *SIAM J. Sci. Comput.* **18** (1997) 1–22.

[41] J. Shen, J. Xu and J. Yang, A new class of efficient and robust energy stable schemes for gradient flows. *SIAM Rev.* **61** (2019) 474–506.

[42] H. Sun, X. Zhao, H. Cao, R. Yang and M. Zhang, Stability and convergence analysis of adaptive BDF2 scheme for the Swift-Hohenberg equation. *Commun. Nonlinear. Sci.* **111** (2022) 106412.

[43] H. Wang, An optimal-order error estimate for a family of ELLAM-MFEM approximations to porous medium flow. *SIAM J. Numer. Anal.* **46** (2008) 2133–2152.

[44] N. Wang and M. Li, Unconditional error analysis of a linearized BDF2 virtual element method for nonlinear Ginzburg-Landau equation with variable time step. *Commun. Nonlinear. Sci.* **116** (2023) 106889.

[45] W. Wang, Y. Chen and H. Fang, On the variable two-step IMEX BDF method for parabolic integro-differential equations with nonsmooth initial data arising in finance. *SIAM J. Numer. Anal.* **57** (2019) 1289–1317.

[46] Y. Wang, Y. Chen, Y. Huang and H. Yi, A family of two-grid partially penalized immersed finite element methods for semi-linear parabolic interface problems. *J. Sci. Comput.* **88** (2021) 80.

[47] S. Xie, S. Yi and T. Kwon, Fourth-order compact difference and alternating direction implicit schemes for telegraph equations. *Comput. Phys. Commun.* **183** (2012) 552–569.

[48] J. Xu, Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.* **33** (1996) 1759–1777.

[49] J. Xu, S. Xie and H. Fu, A two-grid block-centered finite difference method for the nonlinear regularized long wave equation. *Appl. Numer. Math.* **172** (2022) 128–148.

[50] J. Zhang and H. Rui, Numerical analysis of two-grid block-centered finite difference method for two-phase flow in porous medium. *Adv. Appl. Math. Mech.* **14** (2022) 1433–1455.

[51] X. Zhao, Z. Sun and Z. Hao, A fourth-order compact ADI scheme for two-dimensional nonlinear space fractional Schrödinger equation. *SIAM J. Sci. Comput.* **36** (2014) A2865–A2886.

[52] C. Zhao, N. Liu, Y. Ma and J. Zhang, Unconditionally optimal error estimate of a linearized variable-time-step BDF2 scheme for nonlinear parabolic equations. *Commun. Math. Sci.* **21** (2023) 775–794.