Discrete energy analysis of the third-order variable-step BDF time-stepping for diffusion equations

Hong-lin Liao∗ Tao Tang† Tao Zhou‡

Abstract

This is one of our series works on discrete energy analysis of the variable-step BDF schemes. In this part, we present stability and convergence analysis of the third-order BDF (BDF3) schemes with variable steps for linear diffusion equations, see e.g. [SIAM J. Numer. Anal., 58:2294-2314] and [Math. Comp., 90: 1207-1226] for our previous works on the BDF2 scheme. To this aim, we first build up a discrete gradient structure of the variable-step BDF3 formula under the condition that the adjacent step ratios are less than 1.4877, by which we can establish a discrete energy dissipation law. Mesh-robust stability and convergence analysis in the $L^2$ norm are then obtained. Here the mesh robustness means that the solution errors are well controlled by the maximum time-step size but independent of the adjacent time-step ratios. We also present numerical tests to support our theoretical results.

Keywords: diffusion equations, variable-step third-order BDF scheme, discrete gradient structure, discrete orthogonal convolution kernels, stability and convergence

AMS subject classifications: 65M06, 65M12

1 Introduction

In this paper, we aim to develop a discrete energy technique for the stability and convergence of three-step backward differentiation formula (BDF3) with variable time-steps. To this end, we consider the linear reaction-diffusion problem in a bounded convex domain $\Omega$,

$$\partial_t u - \varepsilon \Delta u = \kappa(x)u + f(t, x) \quad \text{for } x \in \Omega \text{ and } 0 < t < T,$$

subject to the Dirichlet boundary condition $u = 0$ on a smooth boundary $\partial\Omega$, and the initial data $u(0, x) = u_0$ for $x \in \Omega$. We assume that the diffusive coefficient $\varepsilon > 0$ is a constant and the reaction coefficient $\kappa(x)$ is smooth and bounded by $\kappa^* > 0$. 

∗ORCID 0000-0003-0777-6832. School of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; Key Laboratory of Mathematical Modelling and High Performance Computing of Air Vehicles (NUAA), MIIT, Nanjing 211106, China. Emails: liaohl@nuaa.edu.cn and liaohl@src.ac.cn. This author’s work is supported by NSF of China under grant number 12071216.

†Division of Science and Technology, BNU-HKBU United International College, Zhuhai, Guangdong Province, China. Email: tangt@sustech.edu.cn. This author’s work is partially supported by NSF of China under grant numbers 11731006 and NNW2018-ZT4A06 project.

‡Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China. Email: tzhou@lsec.cc.ac.cn. This author’s work is partially supported by NSF of China (under grant numbers 11822111, 11688101, and 11731006).
The BDF schemes are widely used for stiff or differential-algebraic problems [7,8]. Recently, they were also applied for simulating hyperbolic systems with multiscale relaxation [1] and stiff kinetic equations [5]. For such applications, BDF schemes with variable steps are shown to be computationally efficient in capturing the multi-scale time behaviors [1,2,4,6,8,10,16]. However, rigorous theoretically analysis (stability and convergence) for variable-step BDF schemes is challenging. This motivates our serious works on this topic, and one can find our previous works on the variable-step BDF2 scheme [11,13].

To begin, we consider the temporal mesh

\[ 0 = t_0 < t_1 < t_2 < \cdots < t_N = T \]

with the variable time-step

\[ \tau_k := t_k - t_{k-1}, \quad 1 \leq k \leq N. \]

The maximum step size and the adjacent time-step ratios are defined respectively as

\[ \tau := \max_{1 \leq k \leq N} \tau_k, \quad r_k := \frac{\tau_k}{\tau_{k-1}} \quad \text{for} \ 2 \leq k \leq N. \]

For any sequences \( \{v^n\}_{n=0}^N \), we denote \( \partial_\tau v^n := (v^n - v^{n-1})/\tau_n \). Then by taking \( v^n = v(t_n) \), the variable-step BDF3 formula [3] yields

\[ D_3 v^n = d_0(r_n, r_{n-1}) \partial_\tau v^n + d_1(r_n, r_{n-1}) \partial_\tau v^{n-1} + d_2(r_n, r_{n-1}) \partial_\tau v^{n-2}, \quad (1.2) \]

where

\[ d_0(x, y) := \frac{1 + 2x}{1 + x} + \frac{xy}{1 + y + xy}, \quad (1.3) \]

\[ d_1(x, y) := -\frac{x}{1 + x} - \frac{xy}{1 + y + xy} - \frac{xy^2}{1 + y + xy}, \quad (1.4) \]

\[ d_2(x, y) := \frac{xy^2}{1 + y + xy}, \quad \text{for} \ x, y \geq 0. \quad (1.5) \]

Without losing the generality, we assume that the discrete solution \( u^1 \) and \( u^2 \) are given. We now consider a time-discrete solution for the diffusion equations, \( u^k(x) \approx u(t_k, x) \) for \( x \in \Omega \), by the following variable-step BDF3 time-stepping scheme

\[ D_3 u^k = \varepsilon \Delta u^k + \kappa u^k + f^k, \quad 3 \leq k \leq N, \quad (1.6) \]

where \( f^k(x) = f(t_k, x) \).

To the best of our knowledge, there are very few theoretical results on the variable-step BDF3 scheme in literature. For linear diffusion problems, Calvo and Grigorieff [3] established the following \( L^2 \) norm stability estimate under the step-ratio condition \( r_k < 1.199 \),

\[ \| u^n \| \leq C \exp (C T_n) \left( \| u_0 \| + \sum_{j=1}^n \tau_j \| f^j \| \right) \quad \text{for} \ n \geq 1, \]

where the ratio-dependent \emph{prefactor} \( \Gamma_n \) is given by

\[ \Gamma_n := \sum_{k=2}^n \left| r_k - r_{k-1} \right|. \]
This means that any mixture of the $k$-step BDF ($k \in \{1, 2, 3\}$) is stable provided the number of changes between increasing and decreasing mesh-sizes is uniformly bounded. Nonetheless, the quantity $\exp(CT_n)$ can be unbounded since $\Gamma_n$ could blow up for certain time-step series at vanishing time-steps. To see this, we consider a specific class of time-steps $\{\ldots, \tau_1, \mu \tau_1, \tau_1, \mu \tau_1, \ldots\}$ with a positive constant $\mu \neq 1$. Then for a finite time $T = \frac{M}{2}(1+\mu)\tau_1$, one has

$$\Gamma_M = \sum_{k=2}^{M} |\mu - \mu^{-1}| = (M - 1)|\mu - \mu^{-1}| \to \infty \quad \text{as } \tau_1 \to 0.$$ 

To remedy this issue, we present shall in this work a new framework for analyzing the variable step BDF3 scheme. Our contribution is two folds:

- We build up a discrete gradient structure of BDF3 formula under the step-ratio condition
  $$0 < r_k < R_e,$$
  where $R_e \approx 1.4877$ is the unique positive root of
  $$d_1(R_e, 0) + \frac{7}{10} \sqrt{R_e} d_2(R_e, R_e) = 0.$$ 

  Based on this, we present the discrete energy stability for the variable-step BDF3 scheme.

- We further present the stability and convergence analysis in the $L^2$ norm for the BDF3 scheme under the adjacent step ratios $0 < r_k < R_e$. We also show that these analysis results are mesh-robustly, which means that the associated prefactors in our analysis are independent of the time-step ratios $r_k$. That is, unbounded quantities such as $\Gamma_n$ in $[3]$ are removed.

The rest of this work is organized as following. In the next section, we provide with some preliminary tools, e.g., discrete orthogonal convolution kernels, for our analysis. In section 3, we present the energy stability of the variable-step BDF3 scheme. The stability and convergence analysis of the variable-step BDF3 scheme is then presented in section 4. This is followed by some numerical examples in section 5.

2 Discrete orthogonal convolution kernels

In this section, we present some preliminary tools for our analysis. We remark that the mesh-robust stability and convergence of the variable-step BDF2 scheme have been established in our previous works $[10, 13]$. Nonetheless, the extension to BDF3 formula is theoretically challenging due to the additional degrees of freedom. This work is builds upon our previous work $[9]$, where the stability of variable-step BDF3 method for nonlinear ODE problems was verified under the step-ratio condition $r_k < R_3 \approx 2.553$.

To begin, we write the BDF3 formula $[1, 2]$ as a convolution of local difference quotients

$$D_3v^n := \sum_{j=1}^{n} d_{n-j}^{(n)} \partial_j v^j \quad \text{for } n \geq 3, \quad (2.7)$$

where the associated discrete BDF3 kernels

$$d_{j}^{(n)} := d_j(r_n, r_{n-1}) \quad \text{for } j = 0, 1, 2, \quad \text{and} \quad d_{j}^{(n)} := 0 \quad \text{for } n \geq j + 1 \geq 4. \quad (2.8)$$
Assume always that the summation \( \sum_{k=i}^{j} \cdot \) to be zero and the product \( \prod_{k=i}^{j} \cdot \) to be one if the index \( i > j \). As for the BDF3 kernels \( \vartheta_{n-j}^{(n)} \) with any fixed indexes \( n \), we recall a class of discrete orthogonal convolution (DOC) kernels \( \{ \vartheta_{n-j}^{(n)} \}_{j=3}^{n} \) by a recursive procedure, also see [13],

\[
\vartheta_{0}^{(n)} := \frac{1}{d_{0}^{(n)}} \quad \text{and} \quad \vartheta_{n-j}^{(n)} := -\frac{1}{d_{0}^{(n)}} \sum_{i=j+1}^{n} \vartheta_{n-i}^{(n)} \vartheta_{i-j}^{(n)} \quad \text{for} \quad 3 \leq j \leq n-1. \tag{2.9}
\]

Obviously, the DOC kernels \( \vartheta_{n-j}^{(n)} \) satisfy the following discrete orthogonality identity

\[
\sum_{i=j}^{n} \vartheta_{n-i}^{(n)} \vartheta_{i-j}^{(n)} = \delta_{nj} \quad \text{for any} \quad 3 \leq j \leq n, \tag{2.10}
\]

where \( \delta_{nj} \) is the Kronecker delta symbol with \( \delta_{nj} = 0 \) if \( j \neq n \). Furthermore, with the identity matrix \( I_{m \times m} \) \( (m := n-2) \), the above discrete orthogonality identity (2.10) also implies

\[
\Theta_{3} D_{3} = I_{m \times m},
\]

where the two \( m \times m \) matrices \( D_{3} \) and \( \Theta_{3} \) are defined by

\[
D_{3} := \begin{pmatrix}
    d_{0}^{(3)} & d_{1}^{(4)} & \ldots & d_{n}^{(n)} \\
    \vdots & \vdots & \ddots & \vdots \\
    d_{n-3}^{(n)} & d_{n-2}^{(n)} & \ldots & d_{0}^{(n)}
\end{pmatrix} \quad \text{and} \quad \Theta_{3} := \begin{pmatrix}
    \vartheta_{0}^{(3)} & \vartheta_{1}^{(4)} & \vartheta_{2}^{(4)} \\
    \vdots & \vdots & \ddots \\
    \vartheta_{n-3}^{(n)} & \vartheta_{n-2}^{(n)} & \vartheta_{0}^{(n)}
\end{pmatrix}.
\]

Obviously, one has \( D_{3} \Theta_{3} = I_{m \times m} \), which implies the following mutual orthogonality identity

\[
\sum_{i=j}^{n} d_{n-i}^{(n)} \vartheta_{i-j}^{(n)} = \delta_{nj} \quad \text{for any} \quad 3 \leq j \leq n. \tag{2.11}
\]

Note that, this identity (2.11) will be used to study the discrete property of DOC kernels; while the above identity (2.10) will be used to reformulate the discrete scheme (1.6). By exchanging the summation order and using (2.10), one has

\[
\sum_{i=3}^{n} \vartheta_{n-i}^{(n)} D_{3} v^{j} = \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} \sum_{j=1}^{2} d_{i-j}^{(i)} \partial_{\tau} v^{j} + \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} \sum_{j=3}^{i} d_{i-j}^{(i)} \partial_{\tau} v^{j} = \sum_{j=1}^{2} \partial_{\tau} v^{j} \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} d_{i-j}^{(i)} + \sum_{j=3}^{n} \partial_{\tau} v^{j} \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} d_{i-j}^{(i)} = T_{3}^{n} [v] + \partial_{\tau} v^{n} \quad \text{for} \quad n \geq 3, \tag{2.12}
\]

where \( T_{3}^{n} [v] \) represents the starting effect on the numerical solution at \( t_{n} \), or

\[
T_{3}^{n} [v] := \sum_{j=1}^{2} \partial_{\tau} v^{j} \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} d_{i-j}^{(i)} = \partial_{\tau} v^{2} \sum_{i=3}^{n} \vartheta_{n-i}^{(n)} d_{i-2}^{(i)} + \vartheta_{n-3}^{(n)} d_{2}^{(3)} \partial_{\tau} v^{1} \quad \text{for} \quad n \geq 3. \tag{2.13}
\]
Multiplying both sides of the equation (1.6) by the DOC kernels $\vartheta^{(n)}_{n-k}$, and summing $k$ from 3 to $n$, we apply (2.12) to get the following equivalent form

$$\partial_t u^n = -\mathcal{T}_3^n[u] + \sum_{k=3}^{n} \vartheta^{(n)}_{n-k} (\varepsilon \Delta u^k + \kappa u^k) + \sum_{k=3}^{n} \vartheta^{(n)}_{n-k} f^k \quad \text{for } 3 \leq n \leq N. \tag{2.14}$$

This formulation expresses the BDF3 solution at time $t_n$ as a (global) convolution summation of all previous solutions with DOC kernels $\vartheta^{(n)}_{n-k}$ as the convolutional weights.

Next section constructs a discrete gradient structure of variable-step BDF3 formula and derives the energy stability of BDF3 scheme via the original form (1.6). Section 3 addresses the $L^2$ norm stability and convergence analysis via the discrete convolution form (2.14). Some numerical examples are presented in the last section to support our theoretical results.

### 3 Positive definiteness and energy stability

To show the energy stability of the variable-step BDF3 scheme, we first investigate sufficient conditions on the adjacent time-step ratios $r_k$ so that the discrete kernels $\{\tau_n d^{(n)}_{n-k}\}$ are positive definite.

For certain adaptive time-stepping process, one may choose the step size $\tau_{m+1}$ (or the step ratio $r_{m+1}$) properly according to the information from previous time-step ratios $\{r_k\}_{k=2}^m$. Actually, the positive definiteness should be determined by the eigenvalues of the pentadiagonal symmetric matrix $B_3 = B_L + B_L^T$, where

$$B_L := \begin{pmatrix}
\tau_3 d_0^{(3)} \\
\tau_4 d_1^{(4)} \\
\tau_5 d_2^{(5)} \\
\vdots \\
\tau_n d_{n-1}^{(n)}
\end{pmatrix}.$$

A sufficient and necessary condition for the positive definiteness of $B_3$ would be a certain combination involving all time-step ratios; however, it remains open at this moment. We consider only certain restriction of each step ratio, like $0 < r_k < R_e$ for a fixed positive constant $R_e < R_3 \approx 2.553$, the recent stability constraint \cite{9} for the ODE problems.

#### Table 1: Minimum eigenvalue of step-rescaled matrix $\tilde{B}_3$ on random meshes.

| $n$ | $R_e = 1.20$ | $R_e = 1.50$ | $R_e = 1.69$ | $R_e = 1.70$ |
|-----|-------------|-------------|-------------|-------------|
| 50  | 1.12        | 5.08e-01    | 6.12e-02    | -4.55e-02   |
| 100 | 1.07        | 4.35e-01    | 4.58e-02    | -5.29e-02   |
| 200 | 1.08        | 4.18e-01    | -2.06e-02   | -8.49e-02   |

Numerical tests on random meshes are performed to examine the positive definiteness of $B_3$ via the step-rescaled matrix $\tilde{B}_3 := \Lambda_\sigma^{-1}(B_L + B_L^T)\Lambda_\sigma^{-1}$, where $\Lambda_\sigma = \text{diag}(\sqrt{\tau_3}, \sqrt{\tau_4}, \ldots, \sqrt{\tau_n})$. 

We take a finite time $T = 1$ with $n$ grid points and let $r_k$ ($2 \leq k \leq n$) be uniformly distributed random numbers over $(0, R_e)$. Table I lists the minimum eigenvalue (each data is the minimum value of 200 runs on different random meshes) of $B_3$ for the fixed step-ratio limits $R_e = 1.20, 1.50, 1.69$ and 1.70. To ensure the positive definiteness, Table I suggests that the maximum step-ratio limit $R_e < 1.69$ is necessary, while we will prove theoretically that $R_e < 1.4877$ is sufficient in the next subsection.

### 3.1 Discrete gradient structure

To derive the energy stability of numerical scheme, we need a discrete gradient structure of variable-step BDF3 formula. In the following, we will seek two nonnegative quadratic functionals $G$ and $F$ such that

$$J_n := 2 v_n \tau_n \sum_{j=3}^{n} d_{n-j}^{(n)} v_j = G[v_n, v_{n-1}] - G[v_{n-1}, v_{n-2}] + F[v_n, v_{n-1}, v_{n-2}], \quad (3.2)$$

for $n \geq 3$, where the discrete BDF3 kernels $d_{j}^{(n)}$ are defined by (2.8), such that the associated quadratic form is positive definite

$$2 \sum_{k=3}^{n} v_k \tau_k \sum_{j=3}^{k} d_{k-j}^{(k)} v_j > 0 \quad \text{if } v_k \neq 0.$$  

This seems to be a difficult task due to the presence of variable kernels $d_{j}^{(n)}$, refer to the recent comments in [6 section 3.4]. For the uniform case with $r_k \equiv 1$, it has been shown in [7] that the uniform BDF3 formula admits the following discrete gradient structure

$$2 v_n \left( \frac{11}{6} v_n - \frac{7}{6} v_{n-1} + \frac{1}{3} v_{n-2} \right) = \frac{1}{3} v_n^2 + \frac{1}{3} \left( \frac{7}{4} v_n - v_{n-1} \right)^2 - \left[ \frac{1}{3} v_n^2 + \frac{1}{3} \left( \frac{7}{4} v_n - v_{n-1} \right)^2 \right] + \frac{95}{48} v_n^2 + \frac{1}{3} \left( v_n - \frac{7}{4} v_{n-1} + v_{n-2} \right)^2.$$  

This decomposition is optimal in the sense that the minimum eigenvalue bound $\frac{95}{38}$ of the associated quadratic form is sharp, see [12, Lemma 2.4], due to the Grenander-Szegö theorem.

To deal with the variable-step case, our first task is to introduce a step-rescale transform $v_j = w_j/\sqrt{\tau_j}$ for $j \geq 1$, cf. the above step-rescaled matrix $B_3$, to remove the time-step factor $\tau_n$ in the discrete kernels of (3.2). One can get

$$J_n = 2w_n \sum_{j=3}^{n} \tilde{d}_{n-j}^{(n)} w_j \quad \text{with} \quad \tilde{d}_{n-j}^{(n)} := \frac{\sqrt{\tau_n}}{\sqrt{\tau_j}} d_{n-j}^{(n)} \quad \text{for } n \geq j + 2 \geq 3. \quad (3.3)$$

It is reasonable to assume that

$$\tilde{G}[w_n, w_{n-1}] := a_{n+1} w_n^2 + b_{n+1} (\gamma w_n - w_{n-1})^2,$$

$$\tilde{F}[w_n, w_{n-1}, w_{n-2}] := p_{n+1} w_n^2 + q_{n+1} (\gamma w_n - w_{n-1})^2 + c_n (w_n - \gamma w_{n-1} + w_{n-2})^2,$$

where the nonnegative variable coefficients $a_n, b_n, c_n, p_n, q_n$ and the real parameter $\gamma$ (for which $\gamma = 7/4$ would not necessarily be optimal) is determined such that

$$J_n = \tilde{G}[w_n, w_{n-1}] - \tilde{G}[w_{n-1}, w_{n-2}] + \tilde{F}[w_n, w_{n-1}, w_{n-2}]. \quad (3.4)$$
The principle of identity gives the following relationships for the undetermined coefficients:

coefficients of $w_n w_{n-2}$: $2c_n = 2\tilde{a}_2^{(n)}$;

coefficients of $w_n^2 (w_{n-1} w_{n-2})$: $-b_n + c_n = 0$;

coefficients of $w_n w_{n-1}$: $-2\gamma(b_{n+1} + q_{n+1}) - 2\gamma c_n = 2\tilde{a}_1^{(n)}$;

coefficients of $w_n^2$: $-a_n + (b_{n+1} + q_{n+1}) - \gamma^2 b_n + \gamma^2 c_n = 0$;

coefficients of $w_n^2$: $a_{n+1} + p_{n+1} + \gamma^2 (b_{n+1} + q_{n+1}) + c_n = 2\tilde{a}_0^{(n)}$.

They yield that

$$a_n = -\frac{1}{\gamma} \tilde{a}_1^{(n)} - \tilde{a}_2^{(n)}, \quad b_n = \tilde{a}_2^{(n)}, \quad c_n = \tilde{a}_2^{(n)},$$

$$p_{n+1} = 2\tilde{a}_0^{(n)} + \gamma \tilde{a}_1^{(n)} + (\gamma^2 - 1) \tilde{a}_2^{(n)} + \frac{1}{\gamma} \tilde{a}_1^{(n+1)} + \tilde{a}_2^{(n+1)},$$

$$q_{n+1} = a_n - b_{n+1} = -\frac{1}{\gamma} \tilde{a}_1^{(n)} - \tilde{a}_2^{(n)} - \tilde{a}_2^{(n+1)}.$$

According to the definitions in (3.3) and (2.8), the coefficients $b_n$ and $c_n$ are always positive, while $a_n$ is also positive if $q_{n+1} \geq 0$. Thus the above assumption (3.4) requires

$$q_{n+1} = -\frac{1}{\gamma} \tilde{a}_1^{(n)} - \tilde{a}_2^{(n)} - \tilde{a}_2^{(n+1)} \geq 0, \quad (3.5)$$

$$p_{n+1} = 2\tilde{a}_0^{(n)} - \tilde{a}_2^{(n)} + \gamma^2 \left( \frac{1}{\gamma} \tilde{a}_1^{(n)} + \tilde{a}_2^{(n)} \right) + \frac{1}{\gamma} \tilde{a}_1^{(n+1)} + \tilde{a}_2^{(n+1)} > 0 \quad \text{for all } r_k < R_e. \quad (3.6)$$

The inequality system (3.5)-(3.6) involves five independent variables $r_{n+1}$, $r_n$, $r_{n-1}$, $\gamma$ and the step-ratio limit $R_e$. In general, we are not able to solve it exactly to determine the optimal values of $\gamma$ so that the resulting step-ratio limit $R_e$ is as large as possible. As done in previous studies [2,9], we consider a specific grid with constant step-ratio $r_k = r$ for a rough estimate of $\gamma$. In such case the first condition (3.5) becomes

$$\gamma \leq \frac{-d_1(r, r)}{2\sqrt{r} d_2(r, r)} \quad \text{for all } r < R_e.$$

An obvious choice is $\tilde{\gamma} = \frac{-d_1(\bar{R}_e, \bar{R}_e)}{2\sqrt{\bar{R}_e} d_2(\bar{R}_e, \bar{R}_e)}$ ($\tilde{\gamma}$ is decreasing with respect to $\bar{R}_e$), but the parameter $\tilde{\gamma}$ is not sufficient to ensure (3.5). Actually, we take $r_{n-1} = 0$ and $r_{n+1} = r_n = \bar{R}_e$ and get

$$-\frac{\sqrt{\bar{R}_e}}{\tilde{\gamma}} d_1(\bar{R}_e, 0) - \bar{R}_e d_2(\bar{R}_e, \bar{R}_e) = \frac{-\bar{R}_e^4(\bar{R}_e^3 + \bar{R}_e^2 - 1)}{(\bar{R}_e^3 + \bar{R}_e + 1)(\bar{R}_e^3 + 3\bar{R}_e^2 + 2\bar{R}_e + 1)} < 0$$

if the step-ratio limit $\bar{R}_e \geq 1$, so that the desired condition (3.5) is invalid. In turn, taking $r_{n-1} = 0$ and $r_{n+1} = r_n = r$ in (3.5) yields

$$\gamma \leq \frac{-d_1(r, 0)}{\sqrt{r} d_2(r, r)} \quad \text{for all } r < R_e.$$

7
It introduces a reliable choice ($\bar{\gamma}$ is also decreasing with respect to $\bar{R}_e$)

$$\bar{\gamma} := \frac{-d_1(\bar{R}_e, 0)}{\sqrt{\bar{R}_e d_2(\bar{R}_e, \bar{R}_e)}},$$

Consider the second condition (3.6) with $\gamma = \bar{\gamma}$ and $r_{n+1} = r_n = r_{n-1} = \bar{R}_e$. Thus we can determine the value of $\bar{R}_e$ by assuming that

$$2d_0(\bar{R}_e, \bar{R}_e) - R_ed_2(\bar{R}_e, \bar{R}_e) + \bar{\gamma}^2 \left( \frac{1}{\bar{\gamma}} \sqrt{\bar{R}_e d_1(\bar{R}_e, \bar{R}_e)} + R_ed_2(\bar{R}_e, \bar{R}_e) \right) + \frac{1}{\bar{\gamma}} \sqrt{\bar{R}_e d_1(\bar{R}_e, \bar{R}_e)} + R_ed_2(\bar{R}_e, \bar{R}_e) = 0.$$  

We solve this equation numerically and find the unique positive roots $\bar{R}_e \approx 1.4965$ with the corresponding parameter $\bar{\gamma} \approx 0.6924$.

To simplify the subsequent mathematical derivations, we fix the parameter $\gamma = 7/10$ which is very close to $\bar{\gamma} \approx 0.6924$. The corresponding maximum step-ratio $R_e$ is determined by

$$d_1(R_e, 0) + \frac{7}{10} \sqrt{R_e d_2(R_e, R_e)} = 0 \quad \text{or} \quad \frac{10}{7}(R_e + 1) - \frac{R_e^2 \sqrt{R_e}}{R_e^2 + R_e + 1} = 0.$$  

We solve this equation numerically and find the unique positive root $R_e \approx 1.4877$. This choice is mainly because the restriction (3.5) should be necessary and sharp, while the inequality (3.6) can be relaxed appropriately.

We are now ready to prove the following lemma, which gives the discrete gradient structure. Note that, the complex conditions (3.5)-(3.6) make the proof tedious and lengthy and some technical lemmas are included in Appendix A.

**Lemma 3.1.** Define the following functions

$$d_*(x, y) := -\frac{10}{7} \sqrt{x} d_1(x, y) - \sqrt{y} d_2(x, y),$$

$$p(x, y, z) := 2d_0(y, z) - \sqrt{y} d_2(y, z) - \frac{49}{100} d_*(y, z) - d_*(x, y),$$

$$q(x, y, z) := d_*(y, z) - \sqrt{y} d_2(x, y), \quad \text{for } 0 < x, y, z < R_e$$

If the step-ratios $0 < r_k < R_e$, there exist two nonnegative functionals $G$ and $F$ such that

$$2v_n \tau_n \sum_{j=3}^{n} d_{n-j}^{(n)} v_j = G[v_n, v_{n-1}] - G[v_{n-1}, v_{n-2}] + F[v_n, v_{n-1}, v_{n-2}] \quad \text{for } n \geq 3,$$

where the Lyapunov-type functional

$$G[v_n, v_{n-1}] := d_*(r_{n+1}, r_n) \tau_n v_n^2 + \sqrt{r_{n+1} r_n} d_2(r_{n+1}, r_n) \left( \frac{7}{10} \sqrt{\tau_n} v_n - \sqrt{\tau_{n-1}} v_{n-1} \right)^2,$$

and the remainder term

$$F[v_n, v_{n-1}, v_{n-2}] := p(r_{n+1}, r_n, r_{n-1}) \tau_n v_n^2 + q(r_{n+1}, r_n, r_{n-1}) \left( \frac{7}{10} \sqrt{\tau_n} v_n - \frac{7}{10} \sqrt{\tau_{n-1}} v_{n-1} + \sqrt{\tau_{n-2}} v_{n-2} \right)^2 \geq \frac{\tau_n v_n^2}{50}.$$
Proof. According to Lemma \ref{lem:A.1} with \( x := r_{n+1}, y := r_n \) and \( z := r_{n-1} \), the condition \eqref{eq:3.4} holds for \( \gamma = 7/10 \), that is,
\[
q_{n+1} = -\frac{10}{7} \sqrt{r_n} d_1^{(n)} - \sqrt{r_n r_{n-1}} d_2^{(n)} - \sqrt{r_{n+1} r_n} d_2^{(n+1)} = q(r_{n+1}, r_n, r_{n-1}) \geq 0, \quad n \geq 3.
\]
Lemma \ref{lem:A.1} also implies that \( a(r_n, r_{n-1}) \geq b(r_{n+1}, r_n) > 0 \). Applying Lemma \ref{lem:A.2} with \( x := r_{n+1}, y := r_n \) and \( z := r_{n-1} \), one has
\[
p_{n+1} = 2d_0^{(n)} + \frac{7}{10} \sqrt{r_n} d_1^{(n)} + \frac{51}{100} \sqrt{r_n r_{n-1}} d_2^{(n)} + \frac{10}{7} \sqrt{r_{n+1} r_n} d_2^{(n+1)} + \sqrt{r_{n+1} r_n} d_2^{(n+1)} = p(r_{n+1}, r_n, r_{n-1}) > \frac{1}{50}, \quad n \geq 3.
\]
Obviously, the condition \eqref{eq:3.5} holds for \( \gamma = 7/10 \). They imply that the discrete gradient structure \eqref{eq:3.4} holds, that is,
\[
2w_n \sum_{j=3}^{n} i(n) j w_j = \tilde{G}[w_n, w_{n-1}] - \tilde{G}[w_{n-1}, w_{n-2}] + \tilde{F}[w_n, w_{n-1}, w_{n-2}]
\]
for \( n \geq 3 \), where
\[
\tilde{G} = \left(-\frac{10}{7} \sqrt{r_n} d_1^{(n+1)} - \sqrt{r_{n+1} r_n} d_2^{(n+1)}\right) w_n^2 + \sqrt{r_{n+1} r_n} d_2^{(n+1)} \left(\frac{7}{10} w_n - w_{n-1}\right)^2
\]
and the remainder term
\[
\tilde{F} = p_{n+1} w_n^2 + q_{n+1} \left(\frac{7}{10} w_n - w_{n-1}\right)^2 + \sqrt{r_n r_{n-1}} d_2^{(n)} \left(w_n - \frac{7}{10} w_{n-1} + w_{n-2}\right)^2 \geq \frac{1}{50} w_n^2.
\]
The claimed result follows from \eqref{eq:3.11} immediately by taking \( w_k := \sqrt{r_k} v_k \) and
\[
G[v_n, v_{n-1}] := \tilde{G} \left[\sqrt{r_n} v_n, \sqrt{r_{n-1}} v_{n-1}\right],
\]
\[
F[v_n, v_{n-1}, v_{n-2}] := \tilde{F} \left[\sqrt{r_n} v_n, \sqrt{r_{n-1}} v_{n-1}, \sqrt{r_{n-2}} v_{n-2}\right].
\]
It completes the proof. \( \square \)

On the uniform mesh with \( d_0 = 11/6, d_1 = -7/6 \) and \( d_2 = 1/3 \), Lemma \ref{lem:3.1} gives
\[
2v_n \left(\frac{11}{6} v_n - \frac{7}{6} v_{n-1} + \frac{1}{3} v_{n-2}\right) = \frac{4}{3} v_n^2 + \frac{1}{3} \left(\frac{7}{10} v_n - v_{n-1}\right)^2 - \left[\frac{4}{3} v_{n-1}^2 + \frac{1}{3} \left(\frac{7}{10} v_{n-1} - v_{n-2}\right)^2\right] + \frac{101}{75} v_n^2 + \left(\frac{7}{10} v_n - v_{n-1}\right)^2 + \frac{1}{3} \left(v_n - \frac{7}{10} v_{n-1} + v_{n-2}\right)^2.
\]
This decomposition arrives at a smaller bound \( \frac{101}{75} \approx 1.347 \) than the optimal bound \( \frac{56}{45} \approx 1.244 \) in \cite[Lemma 2.4]{12} for the minimum eigenvalue of the associated quadratic form. Actually, a sharp estimate of the minimum eigenvalue is not the main purpose here. As seen, the main goal of Lemma \ref{lem:3.1} is to make the step-ratio limit \( R_e \) as large as possible on the basis of realizing a discrete gradient structure \eqref{eq:3.2}. According to the numerical tests at the beginning of this section, the step-ratio limit \( R_e \approx 1.4877 \) would be nearly optimal for having a discrete gradient structure \eqref{eq:3.11}, which implies the positive definiteness of discrete kernels \( \{\tau_n d_n^{(n)}\}_{n=3}^{\infty} \).
Lemma 3.2. If the step-ratios $0 < r_k < R_e \approx 1.4877$, the discrete kernels $\{\tau_n d_{n-k}^{(k)}\}_{k=3}^n$ are positive definite. In the sense that

$$2 \sum_{k=3}^n \xi_k \sum_{j=3}^k \tau_k d_{k-j}^{(k)} \xi_j \geq \frac{1}{50} \sum_{k=3}^n \tau_k \xi_k^2 \text{ for } n \geq 3.$$ 

Proof. By taking $v_1 = v_2 = 0$ and $v_j = \xi_j$ for $3 \leq j \leq n$ in (3.10) and summing the resulting equalities form $n = 3$ to $n = m$, one obtains the claimed inequality by replacing $m$ by $n$. □

3.2 Energy dissipation law

We now prove the energy ($H^1$ seminorm) stability of BDF3 scheme (1.6) for the dissipative case. This property would be practically important when the variable-step BDF3 scheme is applied to the gradient flow problems, cf. the discussions [4,10] on variable-step BDF2 method.

Theorem 3.1. Assume that $\kappa \leq 0$ and $f \equiv 0$. If $0 < r_k < R_e \approx 1.4877$ for $k \geq 2$, the BDF3 scheme (1.6) is unconditionally energy stable in the sense that

$$E^n \leq E^{n-1}, \text{ for } n \geq 3, \quad (3.12)$$

where the (modified) discrete energy $E^n$ is defined by

$$E^n := \varepsilon \| \nabla u^n \|^2 + \langle -\kappa u^n, u^n \rangle + \langle 1, G[\partial_r u^n, \partial_r u^{n-1}] \rangle \text{ for } n \geq 2. \quad (3.13)$$

Proof. Making the inner product of (1.6) with $2\tau_n \partial_r u^n$, one obtains

$$2 \langle D_3 u^n, \tau_n \partial_r u^n \rangle + 2\varepsilon \langle \nabla u^n, \nabla u^n - \nabla u^{n-1} \rangle + 2 \langle -\kappa u^n, u^n - u^{n-1} \rangle = 0, \text{ for } n \geq 3.$$ 

Taking $v_j = \partial_r u^j$ in Lemma 3.1 gives

$$2 \langle D_3 u^n, \tau_n \partial_r u^n \rangle \geq \langle 1, G[\partial_r u^n, \partial_r u^{n-1}] \rangle - \langle 1, G[\partial_r u^{n-1}, \partial_r u^{n-2}] \rangle \text{ for } n \geq 3.$$ 

With the help of the inequality $2a(a - b) \geq a^2 - b^2$, it is easy to obtain that

$$E^n - E^{n-1} \leq 0, \text{ for } n \geq 3.$$ 

The proof is complete. □

4 Stability and convergence analysis

In this section, we shall show the stability and convergence analysis of the variable-step BDF3 scheme.

4.1 Properties of DOC kernels

We first present the following lemma that shows the DOC-type kernels $\{\tau_n d_{n-k}^{(k)}\}_{k=3}^n$ are positive definite.
Lemma 4.1. If the step-ratios $0 < r_k < R_e \approx 1.4877$, the discrete kernels $\{\tau_n \vartheta_{n-k}^{(n)}\}_{k=3}^{n}$ are positive definite in the sense that, for any nonzero sequences $\{\xi_k\}$,

$$2 \sum_{k=3}^{n} \xi_k \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \xi_j > 0, \quad n \geq 3.$$  

Proof. For any nonzero sequences $\{\xi_k\}$, let $\eta_k := \sum_{j=3}^{k} \vartheta_{k-j}^{(k)} \xi_j$ for $k \geq 3$. Multiplying both sides of this equality by the BDF3 kernels $d_{n-k}^{(n)}$ and summing the index $k$ from $k = 3$ to $n$, we apply the orthogonality identity (2.11) to get

$$\sum_{k=3}^{n} d_{n-k}^{(n)} \eta_k = \sum_{k=3}^{n} d_{n-k}^{(n)} \sum_{j=3}^{k} \vartheta_{k-j}^{(k)} \xi_j = \sum_{k=3}^{n} \xi_j \sum_{j=3}^{n} d_{n-k}^{(n)} \vartheta_{k-j}^{(k)} \eta_j = \xi_n \quad \text{for } n \geq 3,$$

where the summation order has been exchanged in the second equality. Since the sequence $\{\eta_k\}$ is also nonzero, Lemma 3.2 gives

$$2 \sum_{k=3}^{n} \xi_k \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \xi_j = 2 \sum_{k=3}^{n} \eta_k \tau_k \sum_{j=3}^{k} d_{n-k}^{(n)} \vartheta_{k-j}^{(k)} \eta_j > 0 \quad \text{for } n \geq 3.$$

The claimed inequality is verified. \qed

As noted in [3], the DOC kernels $\vartheta_{n-k}^{(n)}$ in (2.9) are not always positive, but they decay exponentially such that the absolute summations of DOC kernels are uniformly bounded, as stated in the following lemma.

Lemma 4.2. [3] If the step ratios $r_k \leq R_e$ for $k \geq 2$, there exists a positive constant $K_3$ such that the DOC kernels $\vartheta_{n-j}^{(n)}$ in (2.9) satisfy

$$\sum_{j=3}^{n} |\vartheta_{n-j}^{(n)}| \leq K_3 \quad \text{and} \quad \sum_{j=i}^{n} |\vartheta_{j-i}^{(j)}| \leq K_3 \quad \text{for } n \geq 3 \ (i \geq 3),$$

where $K_3$ is independent of the time $t_n$ and the step ratios $r_k \in (0, R_e] \ (k \geq 2)$.

4.2 $L^2$ norm stability

We are now ready to present the $L^2$ norm stability.

Theorem 4.1. If the step ratios $0 < r_k < R_e$ for $k \geq 2$, the BDF3 solution of (1.6) with $\kappa < 0$ is mesh-robustly stable in the $L^2$ norm, that is,

$$\|u^n\| \leq \|u^2\| + K_3 \tau \|\partial_r u^1\| + 4K_3 \tau \|\partial_r u^2\| + 2 \sum_{k=3}^{n} \tau_k \sum_{j=3}^{k} |\vartheta_{k-j}^{(k)}| \|f^j\|$$

$$\leq \|u^2\| + K_3 \tau \|\partial_r u^1\| + 4K_3 \tau \|\partial_r u^2\| + 2K_3 t_n \max_{3 \leq k \leq n} \|f^k\| \quad \text{for } n \geq 3.$$
Proof. Thanks to Lemma 4.2, it remains to verify the first estimate. Making the inner product of the equation (2.14) with $2\tau_n u^n$, and summing the resulting equality from $n = 3$ to $m$, one has
\[
2 \sum_{k=3}^{m} \tau_k \langle u^k, \partial_r u^k \rangle = -2 \sum_{k=3}^{m} \tau_k \langle u^k, I_3^k[u] \rangle - 2\varepsilon \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \langle \nabla u^k, \nabla u^j \rangle
\]
\[
- 2 \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \langle u^k, (-\kappa) u^j \rangle + 2 \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \langle u^k, \theta_{k-j}^{(k)} f^j \rangle
\]
\[
\text{for } m \geq 3. \text{ Lemma 4.1 leads to}
\]
\[
2 \sum_{k=3}^{m} \tau_k \langle u^k, \partial_r u^k \rangle \leq -2 \sum_{k=3}^{m} \tau_k \langle u^k, I_3^k[u] \rangle + 2 \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \langle u^k, f^j \rangle \quad \text{for } m \geq 3.
\]
Note that, $2\tau_k \langle u^k, \partial_r u^k \rangle \geq \|u^k\|^2 - \|u^{k-1}\|^2$. Then applying the Cauchy-Schwarz inequality and Lemma 4.2, we get
\[
\|u^m\|^2 \leq \|u^2\|^2 + 2 \sum_{k=3}^{m} \tau_k \|u^k\| \|I_3^k[u]\| + 2 \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \|u^k\| \|f^j\| \quad \text{for } m \geq 3.
\]
Taking some integer $m_0$ ($2 \leq m_0 \leq m$) such that $\|u^{m_0}\| = \max_{2 \leq k \leq m} \|u^k\|$. Taking $m := m_0$ in the above inequality, one gets
\[
\|u^{m_0}\|^2 \leq \|u^2\| \|u^{m_0}\| + 2 \|u^{m_0}\| \sum_{k=3}^{m_0} \tau_k \|I_3^k[u]\| + 2 \|u^{m_0}\| \sum_{k=3}^{m_0} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \|f^j\|
\]
and thus
\[
\|u^m\| \leq \|u^{m_0}\| \leq \|u^2\| + 2\tau \sum_{k=3}^{m} \|I_3^k[u]\| + 2 \sum_{k=3}^{m} \sum_{j=3}^{k} \tau_k \vartheta_{k-j}^{(k)} \|f^j\| \quad \text{for } m \geq 3. \quad (4.2)
\]

Now we evaluate the term $\sum_{k=3}^{m} \|I_3^k[u]\|$ stemmed from the starting values. Recalling the definition (2.8) of discrete BDF3 kernels with the increasing property (A.1) of $|d_1(x, y)|$ and $d_2(x, y)$, it is easy to check that
\[
d_2^{(3)} \leq d_2(R_e, R_e) \leq \frac{1}{2} \quad \text{and} \quad |d_1^{(3)}| + d_2^{(4)} \leq -d_1(R_e, R_e) + d_2(R_e, R_e) \leq 2.
\]
Thus we apply the formula (2.13) and Lemma 4.2 to get
\[
\sum_{k=3}^{m} \|I_3^k[u]\| \leq \|\partial_r u^2\| \sum_{k=3}^{m} \sum_{i=3}^{k} \|\vartheta_{k-i}^{(k)}|d_{i-2}^{(i)}| + \|\partial_r u^1\| \sum_{k=3}^{m} \|\vartheta_{k-3}^{(k)}|d_{2}^{(3)}|
\]
\[
= \|\partial_r u^2\| \sum_{i=3}^{m} \|d_{i-2}^{(i)}\| \sum_{k=i}^{m} \|\vartheta_{k-i}^{(k)}| + \|\partial_r u^1\| d_{2}^{(3)} \sum_{k=3}^{m} \|\vartheta_{k-3}^{(k)}|
\]
\[
\leq 2K_3 \|\partial_r u^2\| + \frac{1}{2} K_3 \|\partial_r u^1| \quad \text{for } m \geq 3. \quad (4.3)
\]
Inserting this estimate (4.3) into (4.2), one obtains the first result and completes the proof. \qed
Then the standard Grönwall inequality completes the proof.

Applying Lemma 4.2 and the initial estimate (4.3), we can derive that

\[ m \]

Taking some integer \( m \) (\( 2 \leq m \leq m \)) such that \( \| u^m \| = \max_{2 \leq k \leq m} \| u^k \| \). Taking \( m := m_1 \) in the above inequality, one gets

\[ \| u^m \| \leq \| u^2 \| + 2 \sum_{k=3}^{m_1} \tau_k \| T^k_3 [u] \| + 2 \kappa^* \sum_{k=3}^{m_1} \tau_k \| \theta^{(k)}_{k-1} \| \| u^k \| \| u^j \| \]

Applying Lemma 4.2 and the initial estimate (4.3), we can derive that

\[ \| u^m \| \leq \| u^2 \| + K_3 \| \partial_r u^1 \| + 4 K_3 \| \partial_r u^2 \| + 2 K_3 \kappa^* \sum_{k=3}^{m_1} \tau_k \| u^k \| + 2 K_3 t_n \max_{3 \leq k \leq m} \| f^k \| \]

for \( 3 \leq m \leq N \). With the time-step condition \( \tau \leq 1/(4 K_3 \kappa^*) \), it arrives at

\[ \| u^n \| \leq 2 \| u^2 \| + 2 K_3 \| \partial_r u^1 \| + 8 K_3 \| \partial_r u^2 \| + 4 K_3 \kappa^* \sum_{k=3}^{n-1} \tau_k \| u^k \| + 4 K_3 t_n \max_{3 \leq k \leq n} \| f^k \| \]

Then the standard Grönwall inequality completes the proof.

The above theorems remove the unbounded quantity \( \Gamma_n \) in (3) completely. They show that the variable-step BDF3 scheme is surprisingly stable with respect to the changes of time-steps if the step ratios satisfy a sufficient condition \( r_k < R_e \). Extensive tests on random time meshes in Section 4 suggest that this step-ratio constraint is far from necessary for the stability.

\[ \]

4.3 \( L^2 \) norm convergence

We finally present the \( L^2 \) norm convergence. For the auxiliary functions (1.3)-(1.5), it is easy to check that

\[ d_0(x, y) + d_1(x, y) + d_2(x, y) = 1, \]

\[ d_0(x, y) + \frac{1 + 2x}{x} d_1(x, y) + \frac{1 + 2y + 2xy}{xy} d_2(x, y) = 0, \]

\[ d_0(x, y) + \frac{1 + 3x + 3x^2}{x^2} d_1(x, y) + \frac{1 + 3y + 3xy + 3y^2 + 6xy^2 + 3x^2y^2}{x^2y^2} d_2(x, y) = 0. \]
Consider a smooth function $v$ and let $\zeta^j[v] := D_3 v(t_j) - v'(t_j)$ be the truncation error at $t_j$ ($j \geq 3$) of variable-step BDF3 formula. By applying the Taylor’s expansion with integral remainder, one can apply the identities (4.4)-(4.6) to find that

$$\zeta^j[v] = \sum_{i=j-2}^{j} \frac{1}{6\tau_i} \int_{t_{i-1}}^{t_i} K_{j,j-i}(t)v^{(4)}(t)\,dt,$$  

(4.7)

where the involved integral kernels read

$$K_{j,0}(t) = (d_0^{(j)} - r_j d_1^{(j)})(t_{j-1} - t)^3 + r_j (d_1^{(j)} - r_{j-1} d_2^{(j)})(t_{j-2} - t)^3 + r_j r_{j-1} d_2^{(j)}(t_{j-3} - t)^3,$$

$$K_{j,1}(t) = (d_1^{(j)} - r_{j-1} d_2^{(j)})(t_{j-2} - t)^3 + r_{j-1} d_2^{(j)}(t_{j-3} - t)^3,$$

$$K_{j,2}(t) = d_2^{(j)}(t_{j-3} - t)^3.$$

Reminding the increasing property (A.1), it is not difficult to prove that there exists a bounded constant $K_v$ such that

$$|\zeta^j[v]| \leq K_v \tau^3 \quad \text{for } 3 \leq j \leq N,$$  

(4.8)

where $K_v$ is always dependent on the function $v$, but independent of the time $t_n$, the step sizes $\tau_n$ and the step ratios $r_n$ (even when $r_n$ approaches the limit $R_e$).

Let $\tilde{u}^n := u(t_n) - u^n$ be the solution error of the variable-step BDF3 scheme (1.6). We have the error equation

$$D_3 \tilde{u}^k = \varepsilon \Delta \tilde{u}^k + \kappa \tilde{u}^k + \zeta^k[u], \quad 1 \leq k \leq N$$  

(4.9)

together with the initial conditions $\tilde{u}^0$, $\tilde{u}^1$ and $\tilde{u}^2$. Then applying the priori stability estimate in Theorem 4.2 to the error equation (4.9), we can use the error bound (4.8) to verify the following convergence result.

**Theorem 4.3.** Assume that the solution of (1.1) is smooth enough in time. If the time-step ratios $0 < r_k < R_e$ ($k \geq 2$) with the maximum time-step $\tau \leq 1/(4K_3\kappa^*)$, the solution error $\tilde{u}^n = u(t_n) - u^n$ of the variable-step BDF3 scheme (1.6) satisfies

$$\|\tilde{u}^n\| \leq 2 \exp(4K_3\kappa^* t_{n-1}) \left(\|\tilde{u}^2\| + K_3\tau\|\partial_r \tilde{u}^1\| + 4K_3\tau\|\partial_r \tilde{u}^2\| + 2K_3K_u t_n \tau^3\right)$$

for $3 \leq n \leq N$. Here the constants $K_3$ and $K_u$ are independent of the time $t_n$, the step sizes $\tau_n$ and the step ratios $r_n$ (even when $r_n$ approaches the limit $R_e$). Thus the BDF3 scheme is mesh-robustly convergent in the $L^2$ norm.

To start the third-order stiff solver, one can apply a third-order Runge-Kutta method to compute the starting solutions $u^1$ and $u^2$. Our error estimate in Theorem 4.3 also implies that a second-order starting scheme for computing $u^1$ and $u^2$ would be adequate to achieve the overall third-order accuracy since it can generate third-order accurate solutions at the first two levels.
5 Numerical experiments

We shall present in this section some numerical examples to support our theoretical findings. To this end, we consider the heat equation $\partial_t u - \varepsilon \Delta u = f$ on the square domain $\Omega = (0, 2\pi)^2$ with periodic boundary conditions. We choose the exterior force $f$ and the diffusive coefficient $\varepsilon = 0.1$ such that the equation yields a smooth solution $u = \cos(t) \sin(x) \sin(y)$.

To start the three-step stiff solver, we use the variable-step BDF2 method and a two-stage third-order singly diagonally implicit Runge-Kutta method in our numerical implementations. The numerical stability and convergence are tested until time $T = 1$ in two scenarios:

(a) The periodic time steps $\{\tau_1, \mu \tau_1, \ldots, \tau_1, \mu \tau_1, \ldots, \tau_1, \mu \tau_1\}$ with a constant $\mu > 1$, where $\tau_1 = 2/(N(1 + \mu))$ and the maximum step-ratio $r_{\text{max}} = r_{2j} = \mu$ ($j = 1, 2, \ldots, N/2$).

(b) The random time steps $\tau_k = \epsilon_k / \sum_{k=1}^{N} \epsilon_k$, where $\epsilon_k \in (0, 1)$ are uniformly distributed random numbers.

Table 2: The BDF3 solutions on periodic meshes with $\mu = 2R_e$ starting by a third-order RK

| $N$ | $\tau(N)$ | $e(N)$ | Order | $N_1$ |
|-----|------------|--------|-------|-------|
| 80  | 1.87e-02   | 1.12e-06 | –     | 40    |
| 160 | 9.36e-03   | 1.42e-07 | 2.98  | 80    |
| 320 | 4.68e-03   | 1.78e-08 | 2.99  | 160   |
| 640 | 2.34e-03   | 2.23e-09 | 3.00  | 320   |
| 1280| 1.17e-03   | 2.80e-10 | 3.00  | 640   |

Table 3: The BDF3 solutions on periodic meshes with $\mu = 4R_e$ starting by the BDF2 method

| $N$ | $\tau(N)$ | $e(N)$ | Order | $N_1$ |
|-----|------------|--------|-------|-------|
| 80  | 2.14e-02   | 1.10e-06 | –     | 40    |
| 160 | 1.07e-02   | 1.39e-07 | 2.98  | 80    |
| 320 | 5.35e-03   | 1.74e-08 | 2.99  | 160   |
| 640 | 2.68e-03   | 2.19e-09 | 3.00  | 320   |
| 1280| 1.34e-03   | 2.74e-10 | 2.99  | 640   |

We record the $L^2$ norm error $e(N) := \max_{1 \leq n \leq N} \|v(t_n) - v^n\|$ in each run and compute the numerical order of convergence by

$$\text{Order} \approx \log \left( \frac{e(N)}{e(2N)} \right) \log \left( \frac{\tau(N)}{\tau(2N)} \right)$$

where $\tau(N)$ denotes the maximum time-step size for total $N$ subintervals.

Numerical results on the periodic time meshes are listed in Tables 2 and 3 in which we also record the number (denote by $N_1$) of time levels with the step ratios $r_k \geq R_e \approx 1.4877$. We
Table 4: The BDF3 solutions on random meshes starting by a third-order RK

| N  | τ(N)       | e(N)       | Order | r_{max}  | N_1  |
|----|------------|------------|-------|----------|------|
| 80 | 2.51e-02   | 1.61e-06   | –     | 28.32    | 29   |
| 160| 1.29e-02   | 2.18e-07   | 2.88  | 167.21   | 53   |
| 320| 6.32e-03   | 2.75e-08   | 2.99  | 401.76   | 110  |
| 640| 3.18e-03   | 3.53e-09   | 2.96  | 1656.74  | 206  |
| 1280| 1.53e-03  | 4.36e-10   | 3.02  | 1584.01  | 420  |

Table 5: The BDF3 solutions on random meshes starting by the BDF2 method

| N  | τ(N)       | e(N)       | Order | r_{max}  | N_1  |
|----|------------|------------|-------|----------|------|
| 80 | 2.42e-02   | 1.71e-06   | –     | 746.55   | 13   |
| 160| 1.22e-02   | 2.42e-07   | 2.82  | 110.90   | 26   |
| 320| 6.54e-03   | 2.90e-08   | 3.06  | 79.85    | 70   |
| 640| 3.18e-03   | 3.31e-09   | 3.13  | 371.22   | 125  |
| 1280| 1.57e-03  | 4.44e-10   | 2.90  | 1321.80  | 254  |

observe that (i) the numerical solution is stable even if there are 50% of step-ratios greater than our theoretical restriction; (ii) both a third-order SDIRK method and the second-order BDF2 method are enough to achieve the third-order accuracy, as predicted in Theorem 4.3. Table 4-5 record the numerical results on random time meshes. We see that variable-step BDF3 method is mesh-robust with a desired convergence rate, even if many of step-ratios are much greater than our theoretical limit, and this well be further investigated in our future studies.

Acknowledgements

The authors would like to thank Dr. Ji Bingquan and Dr. Wang Jindi for their help on extensive numerical tests on random time meshes.

A Technical results for Lemma 3.1

By using the definitions (1.3)-(1.5), it is easy to check that

\[
\frac{\partial}{\partial x} |d_\nu(x, y)| > 0 \quad \text{and} \quad \frac{\partial}{\partial y} |d_\nu(x, y)| > 0 \quad \text{for } \nu = 0, 1, 2 \text{ and } x, y > 0. \tag{A.1}
\]

That is, the functions \(d_0(x, y), -d_1(x, y)\) and \(d_2(x, y)\) are increasing with respect to \(x, y > 0\).

This appendix presents some technical lemmas related to the functions \(d_0, d_1\) and \(d_2\) defined by (1.3)-(1.5), respectively. The subsequent analysis is somewhat technically complex and the mathematical derivations have been checked carefully by a symbolic calculation software.
Lemma A.1. For the function \( q \) defined by (3.8), it holds that
\[
q = -\frac{10}{7} \sqrt{y} d_1(y, z) - \sqrt{yz} d_2(y, z) - \sqrt{xy} d_2(x, y) > 0 \quad \text{for } 0 < x, y, z < R_e.
\]
Proof. We consider the auxiliary function
\[
\eta(x, y, z) := \frac{1}{y^3} \left( -\frac{10}{7} y d_1(y^2, z^2) - y z d_2(y^2, z^2) - x y d_2(x^2, y^2) \right)
= \frac{10}{7} \frac{x^3 (x^2 + 1) y^2}{(y^2 + 1) (x^2 y^2 + y^2 + 1)} + \frac{10 z^2 (y^2 z^2 + 2 z^2 + 1)}{7 (z^2 + 1) (y^2 z^2 + z^2 + 1)} - \frac{(y^2 + 1) z^5}{(z^2 + 1) (y^2 z^2 + z^2 + 1)}.
\]
It remains to verify \( \eta(x, y, z) > 0 \) for \( 0 < x, y, z < \sqrt{R_e} \). Simple calculations give
\[
\frac{\partial \eta}{\partial x} = -\frac{x^2 y^2 \left[ 3 x^4 y^2 + x^2 (6 y^2 + 5) + 3 (y^2 + 1) \right]}{(y^2 + 1) (x^2 y^2 + y^2 + 1)^2} < 0 \quad \text{(A.2)}
\]
such that \( \eta(x, y, z) > \eta(\sqrt{R_e}, y, z) \) for \( 0 < x, y, z < \sqrt{R_e} \). Moreover, we get
\[
\frac{\partial \eta}{\partial y} = -\frac{2 y}{7 (y^2 + 1)^2} \left[ -\eta_1(x, y) + \eta_2(y, z) \right], \quad \text{(A.3)}
\]
where \( \eta_1 \) and \( \eta_2 \) are defined by
\[
\eta_1(x, y) := \frac{7 x^3 (x^2 + 1) \left( (x^2 + 1) y^4 - 1 \right)}{(x^2 y^2 + y^2 + 1)^2},
\]
\[
\eta_2(y, z) := \frac{20 (y^2 + 1)^2 z^6 + 7 (y^2 + 1)^2 z^5 + 10 (y^4 + 4 y^2 + 3) z^4 + 10 (2 y^2 + 3) z^2 + 10}{(z^2 + 1) (y^2 z^2 + z^2 + 1)^2}.
\]
For the function \( \eta_1 \), one has
\[
\frac{\partial \eta_1}{\partial y} = \frac{28 x^3 (x^2 + 1)^2 (y^3 + y)}{(x^2 y^2 + y^2 + 1)^3} > 0 \quad \text{for } 0 < x, y < \sqrt{R_e}
\]
and
\[
\frac{\partial \eta_1(x, \sqrt{R_e})}{\partial x} = \frac{7 x^2}{(R_e x^2 + R_e + 1)^3} \left[ 3 R_e^3 x^6 + (9 R_e^2 + 7 R_e - 1) R_e x^4 
+ (9 R_e^3 + 10 R_e^2 - 4 R_e - 5) x^2 + 3 (R_e - 1) (R_e + 1)^2 \right] > 0
\]
such that
\[
\eta_1(x, y) < \eta_1(x, \sqrt{R_e}) < \eta_1(\sqrt{R_e}, \sqrt{R_e}) < 7 \quad \text{for } 0 < x, y < \sqrt{R_e}. \quad \text{(A.4)}
\]
For the function \( \eta_2 \), one has
\[
\frac{\partial \eta_2}{\partial y} = \frac{4 y (y^2 + 1) z^5 (10 z + 7)}{(z^2 + 1) (y^2 z^2 + z^2 + 1)^3} > 0 \quad \text{for } 0 < y, z < \sqrt{R_e}
and
\[ \frac{\partial \eta_2(0, z)}{\partial z} = \frac{z^4(7z - 30 + \sqrt{1145})(30 + \sqrt{1145} - 7z)}{7(z^2 + 1)^4} > 0 \quad \text{for } 0 < z < \sqrt{R_e} \]
such that
\[ \eta_2(y, z) > \eta_2(0, z) > \eta_2(0, 0) = 10 \quad \text{for } 0 < y, z < \sqrt{R_e}. \]

Thanks to (A.4), one has
\[ \eta_1(x, y) < \eta_2(y, z) \quad \text{for } 0 < x, y, z < \sqrt{R_e}. \]

Thus the formula (A.3) shows that
\[ \frac{\partial \eta}{\partial y} < 0 \quad \text{for } 0 < x, y, z < \sqrt{R_e}, \text{ see Figure 1 (a), and then} \]
\[ \eta(x, y, z) > \eta(\sqrt{R_e}, y, z) > \eta(\sqrt{R_e}, \sqrt{R_e}, z) \quad \text{for } 0 < x, y, z < \sqrt{R_e}. \quad (A.5) \]

It remains to examine the following function with respect to \( 0 < z < \sqrt{R_e} \),
\[ \eta(\sqrt{R_e}, \sqrt{R_e}, z) = \frac{10}{7(R_e + 1)} - \frac{R_e^2\sqrt{R_e}}{R_e^2 + R_e + 1} - \frac{7(R_e + 1)z^5 - 10(R_e + 2)z^4 - 10z^2}{7(z^2 + 1)((R_e + 1)z^2 + 1)}. \]

We solve the equation \( \frac{d}{dz}\eta(\sqrt{R_e}, \sqrt{R_e}, z) = 0 \) numerically and find two real solutions \( z_0^* = 0 \) and \( z_1^* = 1.07229 \). Thus
\[ \eta(\sqrt{R_e}, \sqrt{R_e}, z) > \min \left\{ \eta(\sqrt{R_e}, \sqrt{R_e}, 0), \eta(\sqrt{R_e}, \sqrt{R_e}, z_0^*), \eta(\sqrt{R_e}, \sqrt{R_e}, z_1^*) \right\} \]
\[ = \eta(\sqrt{R_e}, \sqrt{R_e}, 0) \approx 0.0000017 > 0. \]

Thus the fact (A.5) arrives at the claimed inequality and completes the proof. \( \square \)

Figure 1: Surfaces of \( \eta(\sqrt{R_e}, y, z) \) and \( \zeta(\sqrt{R_e}, y, z) \) over the domain \((0, \sqrt{R_e})^2\).

**Lemma A.2.** For the function \( p \) defined by (3.9) for \( 0 < x, y, z < R_e \), it holds that
\[ p = 2d_0(y, z) + \frac{7}{10} \sqrt{y}d_1(y, z) - \frac{51}{100} \sqrt{yz}d_2(y, z) + \frac{10}{7} \sqrt{x}d_1(x, y) + \sqrt{xy}d_2(x, y) > \frac{1}{50}. \]
Proof. Reminding the function \( d_* \) defined by (3.7), we consider an auxiliary function

\[
\zeta(x, y, z) := \frac{\zeta_1(y, z) + \zeta_2(x, y)}{1 + y^2} \quad \text{for } 0 < x, y, z < \sqrt{R_e},
\]

where

\[
\zeta_1(y, z) := 2d_0(y^2, z^2) + \frac{7}{10}yd_1(y^2, z^2) - \frac{51}{100}yzd_2(y^2, z^2),
\]

\[
\zeta_2(x, y) := -d_*(x^2, y^2) = \frac{10}{7}xd_1(x^2, y^2) + xyd_2(x^2, y^2).
\]

Obviously, one has

\[
p(x^2, y^2, z^2) = \zeta_1(y, z) + \zeta_2(x, y) = \zeta(x, y, z)(1 + y^2) \quad \text{for } 0 < x, y, z < \sqrt{R_e}. \quad (A.6)
\]

At first we examine the function \( \zeta(x, y, z) \). It is not difficult to show that

\[
\zeta_2(x, y) = -\frac{10x^3}{7(x^2 + 1)} - \frac{x^3y^2(10 - 7y)}{7(y^2 + 1)} - \frac{(10y + 7)y^3}{7(y^2 + 1)} \frac{x^3}{1 + y^2 + y^2x^2}.
\]

Since \( x^3/(c_1 + c_2x^2) \) (\( c_1 > 0 \) and \( c_2 \geq 0 \) are constants) is increasing with respect to \( x \), the function \( \zeta_2 \) and \( \zeta \) are decreasing with respect to \( x \) so that

\[
\zeta(x, y, z) > \zeta(\sqrt{R_e}, y, z) \quad \text{for } 0 < x, y, z < \sqrt{R_e}. \quad (A.7)
\]

By following the elementary arguments in (A.3)-(A.5), one can verify that \( \frac{\partial}{\partial y}\zeta(\sqrt{R_e}, y, z) < 0 \); but the tediously long details are omitted here. As depicted in Figure 1 (b), \( \zeta(\sqrt{R_e}, y, z) \) has no any extreme points in \((0, R_3)^2\). We consider the minimum value along the four boundaries:

(i) Along the side \( y = 0 \), we have \( \zeta(\sqrt{R_e}, 0, z) = 2 - 10R_e\sqrt{R_e}/(7 + 7R_e) \approx 0.9579 \).

(ii) Along the side \( y = \sqrt{R_e} \),

\[
\zeta(\sqrt{R_e}, \sqrt{R_e}, z) \approx \frac{-0.372002z^5 + 0.0240772z^4 + 0.638833z^2 + 0.104165}{z^4 + 1.40198z^2 + 0.401978}.
\]

It has a unique maximum point at \( z \approx 0.448068 \) and then

\[
\zeta(\sqrt{R_e}, \sqrt{R_e}, z) > \min\left\{ \zeta(\sqrt{R_e}, \sqrt{R_e}, 0), \zeta(\sqrt{R_e}, \sqrt{R_e}, z_0), \zeta(\sqrt{R_e}, \sqrt{R_e}, \sqrt{R_e}) \right\} = \zeta(\sqrt{R_e}, \sqrt{R_e}, \sqrt{R_e}) > \frac{1}{50} \quad \text{for } 0 < z < \sqrt{R_e}.
\]

(iii) Along the side \( z = 0 \), we have

\[
\zeta(\sqrt{R_e}, y, 0) \approx \frac{1.11457y^5 - 0.676283y^4 - 0.281384y^3 + 1.105y^2 + 0.385086}{(y^2 + 1)^2(y^2 + 0.401978)}.
\]

It is decreasing such that \( \zeta(\sqrt{R_e}, y, 0) > \zeta(\sqrt{R_e}, \sqrt{R_e}, 0) \approx 0.259131 \) for \( 0 < y < \sqrt{R_e} \).
Along the side $z = \sqrt{Re}$, one has

$$
\zeta(\sqrt{Re}, y, \sqrt{Re}) \approx \frac{(y^2 + 1)^{-2}}{y^4 + 2.07416y^2 + 0.672179} \left( -0.790618y^9 - 1.48448y^7 + 1.32372y^6 
- 0.825252y^5 + 2.77809y^4 - 1.06972y^3 + 3.03679y^2 + 0.643932 \right).
$$

It is also decreasing such that

$$
\zeta(\sqrt{Re}, y, \sqrt{Re}) > \zeta(\sqrt{Re}, \sqrt{Re}, \sqrt{Re}) > \frac{1}{50} \quad \text{for } 0 < y < \sqrt{Re}.
$$

It follows from (A.7) that $\zeta(x, y, z) > \frac{1}{50}$ for $0 < x, y, z < \sqrt{Re}$. According to (A.6), we have $p(x^2, y^2, z^2) > \frac{1+y^2}{50}$ for $0 < x, y, z < \sqrt{Re}$. This completes the proof.

References

[1] G. Albi, G. Dimarco and L. Pareschi, Implicit-explicit multistep methods for hyperbolic systems with multiscale relaxation, SIAM J. Sci. Comput., 42(4), A2402–A2435 (2020).

[2] M. Calvo, T. Grande and R. D. Grigorieff, On the zero stability of the variable order variable stepsize BDF-formulas, Numer. Math., 57, 39–50 (1990).

[3] M. Calvo and R. D. Grigorieff, Time discretisation of parabolic problems with the variable 3-step BDF, BIT, 42, 689-701 (2002).

[4] 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 (1) (2019), 495–525.

[5] G. Dimarco and L. Pareschi, Implicit-explicit linear multistep methods for stiff kinetic equations, SIAM J. Numer. Anal., 55, 664–690 (2017).

[6] V. DeCaria, A. Guzel, W. Layton and Y. Li, A variable stepsize, variable order family of low complexity, SIAM J. Sci. Comput., 43 (3), A2130-A2160 (2021).

[7] E. Hairer, S. P. Norsett and G. Wanner, Solving Ordinary Differential Equations I: Nonstiff Problems, Volume 8 of Springer Series in Computational Mathematics, Second Edition, Springer-Verlag, 1992.

[8] E. Hairer and G. Wanner, Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, Springer Series in Computational Mathematics Volume 14, Second Edition, Springer-Verlag, 2002.

[9] Z. Li and H.-L. Liao, Stability of variable-step BDF2 and BDF3 methods, arXiv:2201.00527, 2021, submitted.

[10] H.-L. Liao, B. Ji and L. Zhang, An adaptive BDF2 implicit time-stepping method for the phase field crystal model, IMA J. Numer. Anal., 42(1), 649–679 (2022).
[11] H.-L. Liao, T. Tang and T. Zhou, *On energy stable, maximum-bound preserving, second-order BDF scheme with variable steps for the Allen-Cahn equation*, SIAM J. Numer. Anal., 58(4), 2294-2314 (2020).

[12] H.-L. Liao, T. Tang and T. Zhou, *A new discrete energy technique for multi-step backward difference formulas*, CSIAM Trans. Appl. Math., 2022, doi: 10.4208/csiam-am.SO-2021-0032.

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

[14] M. Pierre, *Maximum time step for the BDF3 scheme applied to gradient flows*, Calcolo, 58 (2021), doi:10.1007/s10092-020-00393-3.

[15] A.M. Stuart and A.R. Humphries, *Dynamical systems and numerical analysis*, Cambridge University Press, New York, 1998.

[16] D. Wang and S. T. Ruuth, *Variable step-size implicit-explicit linear multistep methods for time-dependent partial differential equations*, J. Comput. Math., 26(6), 838–855 (2008).