Journal of Computational and Applied Mathematics 482 (2026) 117325

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Stability and convergence of a variable step generalized BDF2 scheme

Jie Shen [a], Jiwei Zhang [b], Chengchao Zhao [a,*]

[a] School of Mathematical Science, Eastern Institute of Technology, Ningbo, Zhejiang, 315200, PR China
[b] School of Mathematics and Statistics, and Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan, 430072, China

ARTICLE INFO

AMS subject classifications:
65M12
65M50
35K20

Keywords:
Generalized BDF2 scheme
Adaptive time step
Energy stability
Optimal error estimate

ABSTRACT

We consider in this paper a variable-step generalized second-order BDF scheme (VS-gBDF2) based on the Taylor expansion at time $t_{n-1}+\tau_n\xi$ ($\xi \geq 1$). Using the linear reaction-diffusion equation as an example, we investigate the numerical stability, convergence and computational efficiency of the proposed scheme for any $\xi \geq 1$. We provide an asymptotically compatible analysis on the discrete energy dispassion law and establish stability and second-order convergence of the VS-gBDF2 scheme for any $\xi \geq 1$. These results are proved under a restriction on the ratio of adjacent steps $0 \leq \underline{r}_\xi < r_k < \bar{r}_\xi$, where the upper bound $\bar{r}_\xi$ and the lower bound $\underline{r}_\xi$ are dependent on the parameter ξ. The asymptotic compatibility means that the required ratio restriction will reduce to $0 < r_k < r_{\max} \approx 4.8645$ as $\xi \to 1$, which is the ratio restriction for the classical variable-step BDF2 (VS-BDF2) given in (J. Math., 2021, 06:471-488). Numerical examples are provided to substantiate our theoretical analysis and validate the effectiveness of the adaptive time-step strategy. In particular, the proposed adaptive VS-gBDF2 schemes are shown to have stronger stability and higher efficiency than the classical ($\xi = 1$) VS-BDF2 schemes when an appropriate value of ξ is chosen.

1. Introduction

The adaptive time-stepping strategy is a heuristic method that can significantly improve computational efficiency without sacrificing accuracy in the realm of numerical analysis and scientific computing. It demonstrates practical utility, especially in capturing the multi-scale behaviors, i.e., a system is evolving quickly in some regions of time while slowly changing in other regions [1–3].

The rigorous numerical analysis of such adaptive strategies is relatively easy for one-step methods such as the backward Euler and Crank-Nicolson schemes since they involve only one degree (the current step size τ_n) of freedom. But, the analysis of variable-step multistep methods (two or more steps that involve multiple degrees of freedom) is more challenging. For instance, the typical variable-step second-order backward differentation formula (VS-BDF2) is known for its robust stability, but its analysis, as documented in Thomée's classical book [4], is already highly nontrivial even for linear parabolic equations. Becker [5] presented a stability and convergence analysis for VS-BDF2 by introducing the adjacent time-step ratio restriction $0 < r_k := \tau_k/\tau_{k-1} \leq (2+\sqrt{13})/3 \approx 1.868$. Such a ratio restriction is subsequently relaxed by Emmrich [6] to $0 < r_k < 1.91$, by Wang et. al [7]. to $0 < r_k < 1+\sqrt{2}$, and by Chen et al. [8] to $0 < r_k < 3.561$ using a novel technique of positive-definiteness of BDF2 kernels. Under the same condition, Liao and Zhang[9] carried out the stability and convergence by introducing a discrete orthogonal convolution (DOC) kernel, and Zhang and Zhao [10] further relaxed the ratio restriction to $0 < r_k \leq 4.8645$.

* Corresponding author.
E-mail addresses: jshen@eitech.edu.cn (J. Shen), jiweizhang@whu.edu.cn (J. Zhang), cczhao@eitech.edu.cn (C. Zhao).

https://doi.org/10.1016/j.cam.2025.117325
Received 24 October 2025; Received in revised form 23 December 2025
Available online 29 December 2025

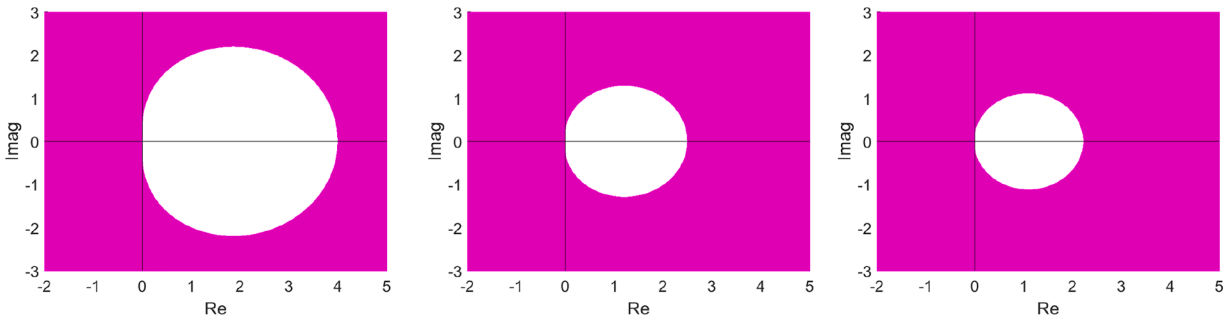

Fig. 1.1. The pink parts show the region of absolute stability of the gBDF2 method with Taylor expansion at $t_{n+\xi-1}, \xi = 1, 2.5, 5$.

Recently, a generalized BDF2 (gBDF2) formula based on Taylor expansion at time $t_{n+\xi-1}$ is constructed in [11] as

$$v_t(t_n) \approx \bar{D}_2^{\xi} v^n = \frac{1}{2\tau}\Big((2\xi+1)v^n - 4\xi v^{n-1} + (2\xi-1)v^{n-2}\Big), \tag{1.1}$$

where ξ is a free parameter and τ is the time-step size. If setting $\xi = 1$, (1.1) reduces to the classical BDF2 scheme. It offers better stability than the classical BDF2 scheme as illustrated in Fig. 1.1 (c.f. [11, Fig. 1] for more details), which shows the stability region increases as ξ increases. Due to its stronger stability, the gBDF2 has been successfully applied to the Navier-Stokes equations in [11] to establish the unconditional stability, for the first time, of a second-order consisting splitting scheme.

In this paper, we propose a variable-step generalized second-order BDF scheme (VS-gBDF2) based on the Taylor expansions at time $t_{n-1} + \tau_n\xi$ ($\xi \geq 1$) and introduce an adaptive time-step strategy in (6.1). Taking the Cahn-Hilliard (CH) equation as an example, as we show in Fig. 6.2, the proposed adaptive VS-gBDF2 scheme exhibits stronger numerical stability with increasing ξ, consistent with the gradually expanding stability region depicted in Fig. 1.1. Due to its strong stability, the adaptive VS-gBDF2 scheme also achieves higher computational efficiency when an appropriate ξ is chosen, see Fig. 6.3.

We also provide a numerical analysis for the VS-gBDF2 scheme. Without loss of generality, we apply the VS-gBDF2 scheme to the linear reaction-diffusion equations, and provide an asymptotically compatible analysis on the mesh-robust energy dissipation law, stability and second-order convergence of the VS-gBDF2 scheme (2.5) for any $\xi \geq 1$. The mesh-robustness means that the proposed results hold for any adjacent time-step ratio r_k satisfying restriction **C1** introduced in Section 2.2, namely, $0 \leq \underline{r}_{\xi} < r_k \leq \bar{r}_{\xi}$. Here the upper bound $\bar{r}_{\xi}$ and the lower bound $\underline{r}_{\xi}$ are dependent on ξ. Especially, as $\xi \to 1$, **C1** will reduce to $0 < r_k \leq 4.8645$ and the VS-gBDF2 method (2.2) will reduce to the VS-BDF2 method (2.3). The compatibility, as $\xi \to 1$, on these results with the classical variable-step BDF2 method in [10] is termed asymptotically compatible.

The analysis introduced in this paper can be in principle extended to (i) various nonlinear equations: phase field models [3,12–16], the jump-diffusion option pricing model in finance [17–19], the Navier-Stokes equation [11,20,21] and so on; (ii) various schemes incorporating VS-gBDF2: the (generalized) scalar auxiliary variable (SAV) approach [14,22–25], Newton linearization method [26], convex splitting method [27–29] and so on.

The rest of the paper is organized as follows. In Section 2, we show the positive semi-definiteness of the VS-gBDF2 kernels under the ratio restriction **C1**. Based on the positive semi-definiteness, we prove in Section 3 that the VS-gBDF2 scheme (1.1) admits a discrete energy dispassion law. Then, we develop some new properties of the DOC and DCC kernels in Section 4, and provide in Section 5 an asymptotically compatible analysis, and establish the stability and for the VS-gBDF2 scheme (1.1). In Section 6, we provide some numerical examples to validate our theoretical results.

2. The VS-gBDF2 scheme and its kernels

We start by describing the VS-gBDF2 scheme, and then proceed to show that its kernels are positive semi-definite.

2.1. The VS-gBDF2 scheme

Set nonuniform time levels $0 = t_0 < t_1 < \cdots < t_N = T$ with the time-step size $\tau_k := t_k - t_{k-1}$, the maximum step size $\tau := \max_{1\leq k\leq N} \tau_k$ and the adjacent time-step ratio $r_k = \tau_k/\tau_{k-1} (2 \leq k \leq N)$. Denote by $v^n := v(t_n)$ and $\nabla_\tau u^n := u^n - u^{n-1}$. Then, the VS-gBDF2 method based on Taylor expansion at $t_{n-1} + \tau_n\xi$ is given as

$$D_2^{\xi} v^n = \frac{1+2r_n\xi}{(1+r_n)\tau_n} v^n - \frac{1+(2\xi-1)r_n}{\tau_n} v^{n-1} + \frac{(2\xi-1)r_n^2}{(1+r_n)\tau_n} v^{n-2} \tag{2.1}$$

$$= \frac{1+2r_n\xi}{(1+r_n)\tau_n} \nabla_\tau v^n - \frac{r_n^2(2\xi-1)}{(1+r_n)\tau_n} \nabla_\tau v^{n-1}, \tag{2.2}$$

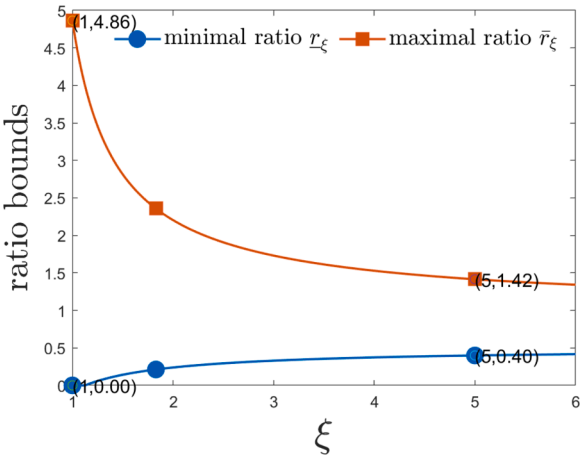

Fig. 2.1. The red line and blue line represent the parametric curves of the upper bound $\bar{r}_\xi$ and lower bound $\underline{r}_\xi$, respectively, as functions of the parameter ξ.

where $\xi \geq 1$ is a constant. If $r_n \equiv 1$, then (2.2) reduces to the gBDF2 method (1.1). If $\xi = 1$, (2.2) reduces to the classical VS-BDF2 scheme [8–10]

$$D_2^1 v^n = \frac{1+2r_n}{(1+r_n)\tau_n}\nabla_\tau v^n - \frac{r_n^2}{(1+r_n)\tau_n}\nabla_\tau v^{n-1}. \tag{2.3}$$

Without loss of generality, we consider the linear reaction-diffusion equation:

$$\begin{aligned} &u_t(\boldsymbol{x},t) = \Delta u(\boldsymbol{x},t) + \kappa u(\boldsymbol{x},t) + f(\boldsymbol{x},t), && \boldsymbol{x}\in\Omega\subset\mathbb{R}^d, t\in(0,T],\\ &u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), && \boldsymbol{x}\in\bar{\Omega},\\ &u(\boldsymbol{x},t) = 0, && \boldsymbol{x}\in\partial\Omega, t\in[0,T], \end{aligned} \tag{2.4}$$

where the reaction coefficient $\kappa \in \mathbb{R}$ is a given constant, and Ω is a bounded domain. Set $E_\xi v^n := \xi v^n - (\xi-1)v^{n-1}$, $n \geq 1$. Then, the VS-gBDF2 scheme for (2.4) is

$$D_2^\xi u^n = \Delta E_\xi u^n + \kappa E_\xi u^n + E_\xi f^n, \quad 2 \leq n \leq N, \tag{2.5}$$

where $f^n := f(\boldsymbol{x}, t_n)$. For the method (2.5), we need the starting values u^0 and u^1. We set $u^0 = u(0)$ and compute u^1 using other numerical methods, such as the backward Euler method.

Define the VS-gBDF2 kernels as

$$A_0^{(n,\xi)} := \frac{1+2r_n\xi}{(1+r_n)\tau_n}, \quad A_1^{(n,\xi)} := -\frac{r_n^2(2\xi-1)}{(1+r_n)\tau_n}, \ A_j^{(n,\xi)} = 0\ (n \geq j \geq 2), \tag{2.6}$$

the VS-gBDF2 (2.2) can be written as a convolution form of

$$D_2^\xi u^n := \sum_{k=1}^{n} A_{n-k}^{(n,\xi)}\nabla_\tau u^k, \quad \text{for all } n \geq 2. \tag{2.7}$$

We point out that the positive definiteness of VS-gBDF2 kernels plays a pivotal role in the stability and convergence analysis of the VS-gBDF2 scheme (2.2), and this property is valid under the condition **C1** presented in the next subsection.

2.2. *The ratio condition* **C1** *for the positive definiteness*

The restriction on the ratio of adjacent time steps is given by

C1: $\underline{r}_\xi < r_k \leq \bar{r}_\xi, \quad 2 \leq k \leq N,$

where $\bar{r}_\xi$ is the maximal real root of the following equation

$$(1+2x\xi) = (2\xi-1)x^{3/2}, \tag{2.8}$$

and the lower bound $\underline{r}_\xi$ is given by

$$\underline{r}_\xi = \begin{cases} 0, & \text{if} \quad 1 \le \xi \le \frac{9}{8}, \\ \frac{2\sqrt{\bar{r}_\xi}-\bar{r}_\xi}{1+\bar{r}_\xi} & \text{if} \quad \xi > \frac{9}{8}. \end{cases} \tag{2.9}$$

The relationship between $\bar{r}_\xi, \underline{r}_\xi$ and ξ is illustrated in Fig. 2.1. Fig. 2.1 shows that the ratio condition becomes more severe as ξ increases. This indicates that while the VS-gBDF2 method exhibits increased stability with larger ξ, its mesh robustness diminishes. Therefore, in practical applications, ξ should not be set too large.

Lemma 1. *Let $\bar{r}_\xi$ and $\underline{r}_\xi$ be defined by (2.8) and (2.9) respectively. Then $\bar{r}_\xi$ and $\underline{r}_\xi$ exist uniquely and satisfy*

1. *The upper bound $\bar{r}_\xi$ is monotonically non-increasing with respect to ξ and tends to 1^+ as $\xi \to \infty$.*
2. *The lower bound $\underline{r}_\xi$ is monotonically non-decreasing with respect to ξ. Moreover, $\underline{r}_\xi \ge 0$ for any $\xi \ge 1$ and tends to $1/2$ as $\xi \to \infty$.*

Proof. Let $r_{\max} \approx 4.8645$ be the real solution of (2.8) with $\xi = 1$. We define $g_\xi(x) := (2\xi-1)x^{3/2} - 2x\xi - 1$. For $\xi \ge 1$, we observe that

$$g_\xi(1) = -2 < 0, \quad g_\xi(r_{\max}) = 2(r_{\max}+1)(\xi-1) \ge 0,$$

which implies that $\bar{r}_\xi$ exists in the interval $(1, r_{\max})$ for any $\xi \ge 1$. From (2.8), ξ can be represented in terms of $\bar{r}_\xi$ as follows:

$$\xi = \frac{1+\bar{r}_\xi^{3/2}}{2\bar{r}_\xi(\sqrt{\bar{r}_\xi}-1)}. \tag{2.10}$$

Let $h(x) = \frac{1+x^{3/2}}{2x(\sqrt{x}-1)}$, $x > 1$. Its derivative is given by:

$$\frac{\mathrm{d}}{\mathrm{d}x}h(x) = \frac{1 - \frac{1}{2}x^{3/2} - \frac{3}{2}\sqrt{x}}{2(x^{3/2}-x)^2} < 0,$$

which implies that $h(x)$ is non-increasing on $(1,\infty)$. Hence, $\bar{r}_\xi$ is monotonically non-increasing with respect to ξ and tends to 1^+ as $\xi \to \infty$. Similarly, it can be verified that $\underline{r}_\xi$ is non-increasing with respect to $\bar{r}_\xi$. Thus, $\underline{r}_\xi$ exists uniquely and is non-decreasing with respect to ξ. Simple calculations indicate that $\bar{r}_\xi = 4$ and $\underline{r}_\xi = 0$ if $\xi = \frac{9}{8}$ and $\bar{r}_\xi \to 1^+$ as $\xi \to \infty$, which together with the monotonicity of $\bar{r}_\xi$ and $\underline{r}_\xi$ imply that $\underline{r}_\xi \ge 0$ for any $\xi \ge 1$ and $\underline{r}_\xi$ tends to 0.5 as $\xi \to \infty$. The proof is completed. □

Remark 1. If $\xi \to 1$, as we know, VS-gBDF2 method (2.2) reduces to VS-BDF2 method (2.3). In the situation of $\xi = 1$, we have the lower bound $\underline{r}_\xi = 0$ and the upper bound $\bar{r}_\xi = r_{\max} \approx 4.8645$, where $r_{\max}$ is the positive root of equation $x^3 = (1+2x)^2$, which is consistent with the positive definiteness presented in [10] for VS-BDF2 kernels. In this sense, we say that the ratio restriction is asymptotically compatible.

Remark 2.

1. In [30], the authors introduce a weighted and shifted BDF2 (WSBDF2) method with variable time steps, namely,

$$D_{ws}v^n := \frac{1+2\theta r_n}{\tau_n(1+r_n)}\nabla_\tau v^n + \frac{(1-2\theta)r_n^2}{\tau_n(1+r_n)}\nabla_\tau v^{n-1}, \tag{2.11}$$

   where the parameter $\theta \in (\frac{1}{2}, 1]$. In fact, the WSBDF2 method is a special case of the VS-gBDF2 method (2.2) with $\xi \in (\frac{1}{2}, 1]$.
2. When the WSBDF2 method is applied to the linear reaction-diffusion equation (2.4) with $\kappa = 0$, stability and second-order convergence are proven in [30] under the ratio condition $0 < r_n \le r_s$. The upper bound r_s is non-increasing with respect to θ, with $r_s \to r_{\max} \approx 4.8645$ as $\theta \to 1^-$, and $r_s \to +\infty$ as $\theta \to \frac{1}{2}^+$. In this sense, the decreasing property of our upper bound $\bar{r}_\xi$ is consistent with behaviors in [30].

2.3. The positive semi-definiteness of VS-gBDF2 kernels

We now consider the positive-definiteness of VS-gBDF2 kernels. In the beginning we define

$$\mathfrak{T}_\xi(x,y) = \frac{2+4x\xi-(2\xi-1)x^2/\sqrt{\bar{r}_\xi}}{1+x} - \frac{(2\xi-1)\sqrt{\bar{r}_\xi}y}{1+y}, \tag{2.12}$$

where $\xi \ge 1$ is a constant and $\bar{r}_\xi$ is defined by (2.8). The proof of the positive semi-definiteness relies on the assistance of the lemma below.

Lemma 2. *Let $\bar{r}_\xi$ and $\underline{r}_\xi$ be defined by (2.8) and (2.9) respectively. Then it holds that*

$$\mathfrak{T}_\xi(x,y) \ge 0, \quad \forall \underline{r}_\xi \le x, y \le \bar{r}_\xi, \quad \xi \ge 1. \tag{2.13}$$

Proof. It is straightforward to verify that $\mathfrak{T}_\xi(x,y)$ is decreasing with respect to y. Then, we have for any $\underline{r}_\xi \le y \le \bar{r}_\xi$

$$\mathfrak{T}_\xi(x,y) \ge \frac{2+4x\xi-(2\xi-1)x^2/\sqrt{\bar{r}_\xi}}{1+x} - \frac{(2\xi-1)\bar{r}_\xi^{3/2}}{1+\bar{r}_\xi}.$$

Inserting (2.10) into the above inequality, one has

$$\begin{aligned}\mathfrak{T}_\xi(x,y) &\ge \frac{2\sqrt{\bar{r}_\xi}(\bar{r}_\xi^{3/2}-\bar{r}_\xi+(1+\bar{r}_\xi^{3/2})x)-(1+\bar{r}_\xi)x^2-\bar{r}_\xi^2(1+x)}{(1+x)(\bar{r}_\xi^{3/2}-\bar{r}_\xi)\sqrt{\bar{r}_\xi}}\\ &:= \frac{-\big((1+\bar{r}_\xi)x-(2\sqrt{\bar{r}_\xi}-\bar{r}_\xi)\big)(x-\bar{r}_\xi)}{(1+x)(\bar{r}_\xi^{3/2}-\bar{r}_\xi)\sqrt{\bar{r}_\xi}}.\end{aligned} \tag{2.14}$$

If $1 \le \xi \le \frac{9}{8}$, it follows from (2.9) and (2.10) that $\underline{r}_\xi = 0, \bar{r}_\xi \ge 4$. Then one has for $\underline{r}_\xi \le x \le \bar{r}_\xi$ that $(1+\bar{r}_\xi)x-(2\sqrt{\bar{r}_\xi}-\bar{r}_\xi) \ge \bar{r}_\xi - 2\sqrt{\bar{r}_\xi} \ge 0$. If $\xi > \frac{9}{8}$, it follows from (2.10) that $\bar{r}_\xi < 4$. It is easy to check by (2.9) that $(1+\bar{r}_\xi)x-(2\sqrt{\bar{r}_\xi}-\bar{r}_\xi) \ge 0$ for any $\underline{r}_\xi \le x \le \bar{r}_\xi$. Thus, we have for $\underline{r}_\xi \le x, y \le \bar{r}_\xi$ that

$$\mathfrak{T}_\xi(x,y) \ge 0. \tag{2.15}$$

The proof is completed. □

Lemma 3. *Assume* **C1** *is satisfied, then it holds for any real sequence* $\{w_k\}_{k=1}^n$ *that*

$$2w_k\sum_{j=1}^k A_{k-j}^{(k,\xi)}w_j \ge \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_{k+1}}{1+r_{k+1}}\frac{w_k^2}{\tau_k} - \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_k}{1+r_k}\frac{w_{k-1}^2}{\tau_{k-1}}, \tag{2.16}$$

$$2\sum_{k=2}^n w_k\sum_{j=2}^k A_{k-j}^{(k,\xi)}w_j \ge 0, \quad \text{for } n \ge 2. \tag{2.17}$$

Proof. Using the Young inequality $2ab \le \epsilon a^2 + b^2/\epsilon$ and taking $\epsilon = 1/\sqrt{\bar{r}_\xi}$, one has

$$\begin{aligned}w_k\sum_{j=1}^k A_{k-j}^{(k,\xi)}w_j &= \frac{2+4r_k\xi}{(1+r_k)\tau_k}w_k^2 - \frac{2r_k^2(2\xi-1)}{(1+r_k)\tau_k}w_kw_{k-1}\\ &\ge \frac{2+4r_k\xi-(2\xi-1)r_k^2/\sqrt{\bar{r}_\xi}}{1+r_k}\frac{w_k^2}{\tau_k} - \frac{r_k\sqrt{\bar{r}_\xi}(2\xi-1)}{1+r_k}\frac{w_{k-1}^2}{\tau_{k-1}}\\ &= \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_{k+1}}{1+r_{k+1}}\frac{w_k^2}{\tau_k} - \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_k}{1+r_k}\frac{w_{k-1}^2}{\tau_{k-1}} + \mathfrak{T}(r_k,r_{k+1})\frac{w_k^2}{\tau_k}.\end{aligned} \tag{2.18}$$

Then (2.16) is achieved by applying Lemma 2 to (2.18).

It follows from the definition (2.6) that $\sum_{j=2}^k A_{k-j}^{(k,\xi)}w_j = \sum_{j=1}^k A_{k-j}^{(k,\xi)}w_j$ if $k \ge 3$. By the inequality (2.16), the direct calculation yields

$$\begin{aligned}2\sum_{k=2}^n w_k\sum_{j=2}^k A_{k-j}^{(k,\xi)}w_j &= 2A_0^{(2,\xi)}w_2^2 + 2\sum_{k=3}^n w_k\sum_{j=2}^k A_{k-j}^{(k,\xi)}w_j\\ &\ge \frac{2(1+2r_2\xi)}{(1+r_2)}\frac{w_2^2}{\tau_2} - \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_3}{1+r_3}\frac{w_2^2}{\tau_2}\\ &\ge \frac{(2\xi-1)r_2^2}{(1+r_2)\sqrt{\bar{r}_\xi}}\frac{w_2^2}{\tau_2} \ge 0,\end{aligned}$$

where the last inequality uses the inequality (2.13). The proof is completed. □

Remark 3. As pointed out in Remark 1, the positive semi-definiteness presented in Lemma 3 is asymptotically compatible. The subsequent analysis of discrete energy dispassion in Section 3 and stability and convergence in Section 5 is also asymptotically compatible in the same sense.

3. Discrete energy dissipation

We now consider the energy (H^1 semi-norm) stability of VS-gBDF2 scheme (2.5) by defining a modified discrete energy as

$$\mathcal{E}^k := \frac{(2\xi-1)\sqrt{\bar{r}_\xi}r_{k+1}}{2(1+r_{k+1})}\tau_k\|\partial_\tau u^k\|^2 + \bar{\mathcal{E}}^k \quad \text{for } \kappa \le 0 \text{ and } k \ge 1, \tag{3.1}$$

with $\partial_\tau u^k := \nabla_\tau u^k/\tau_k$, where $\xi \ge 1$ is a constant and the original energy

$$\bar{\mathcal{E}}^k := \frac{1}{2}\|\nabla u^k\|^2 - \frac{\kappa}{2}\|u^k\|^2. \tag{3.2}$$

Theorem 1. *Assume* **C1** *holds and* $\kappa \le 0$. *Then the discrete energy* (3.1) *holds*

$$\partial_\tau \mathcal{E}^k \le (E_\xi f^k, \partial_\tau u^k), \quad 2 \le k \le n. \tag{3.3}$$

And the discrete solution u^n *exhibits unconditional stability in the energy norm as*

$$\sqrt{\mathcal{E}^n} \le \sqrt{\mathcal{E}^1} + 2C_\Omega\Big(\sum_{k=2}^{n} \|\nabla_\tau E_\xi f^k\| + \|E_\xi f^1\|\Big). \tag{3.4}$$

Proof. Set $n = k$ in (2.5) and take inner products with $\nabla_\tau u^k$, one has

$$(D_2^\xi u^k, \nabla_\tau u^k) = (\Delta E_\xi u^k + \kappa E_\xi u^k + E_\xi f^k, \nabla_\tau u^k), \quad 2 \le n \le N. \tag{3.5}$$

For the first term in (3.5), we apply (2.16) to have

$$(D_2^\xi u^k, \nabla_\tau u^k) \ge \frac{(2\xi-1)\sqrt{\bar{r}_\xi} r_{k+1}}{2(1+r_{k+1})} \frac{\|\nabla_\tau u^k\|^2}{\tau_k} - \frac{(2\xi-1)\sqrt{\bar{r}_\xi} r_k}{2(1+r_k)} \frac{\|\nabla_\tau u^{k-1}\|^2}{\tau_{k-1}}. \tag{3.6}$$

For the second and third terms in (3.5), we use the identity $2(a + c(a-b))(a-b) = a^2 - b^2 + (1+2c)(a-b)^2$ to have

$$2(\Delta E_\xi u^k, \nabla_\tau u^k) = -\nabla_\tau \|\nabla u^k\|^2 - (2\xi-1)\|\nabla_\tau \nabla u^k\|^2, \tag{3.7}$$

$$2(\kappa E_\xi u^k, \nabla_\tau u^k) = \kappa \nabla_\tau \|u^k\|^2 + (2\xi-1)\kappa \|\nabla_\tau u^k\|^2, \tag{3.8}$$

where (3.7) uses the formula of integration by parts. Inserting (3.7) and (3.8) into (3.5), and then using the definition (3.1), one produces (3.3).

Multiplying τ_k by (3.3) and summing the resulting inequality from $k = 2$ to n, we have

$$\mathcal{E}^n \le \mathcal{E}^1 + \sum_{k=2}^{n} (E_\xi f^k, \nabla_\tau u^k). \tag{3.9}$$

Noting that

$$\sum_{k=2}^{n} (E_\xi f^k, \nabla_\tau u^k) = (E_\xi f^n, u^n) - \sum_{k=2}^{n-1} (\nabla_\tau E_\xi f^{k+1}, u^k) - (E_\xi f^2, u^1). \tag{3.10}$$

Then we use the Cauchy-Schwarz inequality to have

$$\begin{aligned} \sum_{k=2}^{n} (E_\xi f^k, \nabla_\tau u^k) &\le \|E_\xi f^n\| \|u^n\| + \sum_{k=2}^{n-1} \|\nabla_\tau E_\xi f^{k+1}\| \|u^k\| + \|E_\xi f^2\| \|u^1\| \\ &\le C_\Omega\Big(\sqrt{\mathcal{E}^n}\|E_\xi f^n\| + \sum_{k=2}^{n-1} \sqrt{\mathcal{E}^k}\|\nabla_\tau E_\xi f^{k+1}\| + \sqrt{\mathcal{E}^1}\|E_\xi f^2\|\Big), \end{aligned} \tag{3.11}$$

where the last inequality uses the Poincaré inequality. Together with (3.9), one has

$$\mathcal{E}^n \le \mathcal{E}^1 + C_\Omega\Big(\sqrt{\mathcal{E}^n}\|E_\xi f^n\| + \sum_{k=2}^{n-1} \sqrt{\mathcal{E}^k}\|\nabla_\tau E_\xi f^{k+1}\| + \sqrt{\mathcal{E}^1}\|E_\xi f^2\|\Big). \tag{3.12}$$

Choosing n_0 such that $\mathcal{E}^{n_0} = \max_{1\le k\le n} \mathcal{E}^k$. We take $n = n_0$ in (3.12) and have

$$\mathcal{E}^{n_0} \le \sqrt{\mathcal{E}^{n_0}}\Big(\sqrt{\mathcal{E}^1} + C_\Omega\Big(\|E_\xi f^{n_0}\| + \sum_{k=2}^{n_0-1} \|\nabla_\tau E_\xi f^{k+1}\| + \|E_\xi f^2\|\Big)\Big). \tag{3.13}$$

Then we use the facts $n_0 \le n$ and $\mathcal{E}^n \le \mathcal{E}^{n_0}$ to obtain

$$\sqrt{\mathcal{E}^n} \le \sqrt{\mathcal{E}^{n_0}} \le \sqrt{\mathcal{E}^1} + C_\Omega\Big(\|E_\xi f^n\| + \sum_{k=3}^{n} \|\nabla_\tau E_\xi f^k\| + \|E_\xi f^2\|\Big). \tag{3.14}$$

From the equality $f^n = \sum_{k=2}^{n} \nabla_\tau f^k + f^1$, one has

$$\|E_\xi f^n\| + \sum_{k=3}^{n} \|\nabla_\tau E_\xi f^k\| + \|E_\xi f^2\| \le 2\sum_{k=2}^{n} \|\nabla_\tau E_\xi f^k\| + 2\|E_\xi f^1\|. \tag{3.15}$$

Inserting (3.15) into (3.14), one obtains (3.4). The proof is completed. □

If $f(x,t) = 0$, the inequality (3.9) implies the discrete energy dissipation law

$$\mathcal{E}^k \le \mathcal{E}^{k-1}, \quad \text{for } k \ge 2. \tag{3.16}$$

This property is crucial for numerical schemes in simulating the gradient flow problems, cf [3,8,11]. and the references therein.

4. The DOC and DCC kernels

To handle the difficulty arising from variable-step sizes, we will introduce the techniques of DOC [9] and DCC [10] kernels. They will play a pivotal role in bridging the positive-definiteness of VS-gBDF2 kernels and the stability and convergence of VS-gBDF2 scheme. For the VS-gBDF2 kernels defined in (2.7), the corresponding DOC and DCC kernels also depend on the parameter ξ. Specifically, the DOC kernels $\theta_{n-j}^{(n,\xi)}$ are defined by

$$\sum_{j=k}^{n} \theta_{n-j}^{(n,\xi)} A_{j-k}^{(j,\xi)} = \delta_{nk}, \quad \text{for all } 1 \le k \le n, \tag{4.1}$$

where δ_{nk} represents the Kronecker delta symbol. And DCC kernels $p_{n-j}^{(n,\xi)}$ are defined by

$$\sum_{j=k}^{n} p_{n-j}^{(n,\xi)} A_{j-k}^{(j,\xi)} \equiv 1, \quad \forall 1 \le k \le n,\ 1 \le n \le N. \tag{4.2}$$

The definition (4.1) of DOC kernels entails that

$$\begin{aligned} \sum_{j=2}^{n} \theta_{n-j}^{(n,\xi)} D_2 u^j &= \sum_{l=2}^{n} \nabla_\tau u^l \sum_{j=l}^{n} \theta_{n-j}^{(n,\xi)} A_{j-l}^{(j,\xi)} + \theta_{n-2}^{(n,\xi)} A_1^{(2,\xi)} \nabla_\tau u^1 \\ &= \nabla_\tau u^n + \theta_{n-2}^{(n,\xi)} A_1^{(2,\xi)} \nabla_\tau u^1, \quad 2 \le n \le N, \end{aligned} \tag{4.3}$$

where we have changed the order of summation. Note that we recall that $A_{j-1}^{(j,\xi)} \equiv 0$ for all $j \ge 3$. Applying the techniques in [9,10], we can similarly derive the following properties.

Proposition 1. *The DOC and the DCC kernels have the relationships of*

$$p_{n-j}^{(n,\xi)} = \sum_{l=j}^{n} \theta_{l-j}^{(l,\xi)}, \quad \forall 1 \le j \le n, \tag{4.4}$$

$$\theta_{n-j}^{(n,\xi)} = p_{n-j}^{(n,\xi)} - p_{n-1-j}^{(n-1,\xi)}, \quad \forall 1 \le j \le n, \tag{4.5}$$

where $p_{-1}^{(n,\xi)} := 0$ $(\forall n \ge 0)$ *is defined.*

Proof. Set $q_{n-j}^{(n,\xi)} = \sum_{l=j}^{n} \theta_{l-j}^{(l,\xi)}$. It follows from the definition (4.1) that

$$\sum_{j=k}^{n} q_{n-j}^{(n,\xi)} A_{j-k}^{(j,\xi)} = \sum_{j=k}^{n} \sum_{l=j}^{n} \theta_{l-j}^{(l,\xi)} A_{j-k}^{(j,\xi)} = \sum_{l=k}^{n} \sum_{j=k}^{l} \theta_{l-j}^{(l,\xi)} A_{j-k}^{(j,\xi)} = \sum_{l=k}^{n} \delta_{lk} = 1.$$

Hence, $q_{n-j}^{(n,\xi)} = \sum_{l=j}^{n} \theta_{l-j}^{(l,\xi)}$ $(1 \le j \le n)$ are solutions to (4.2). Noting the DCC kernels uniquely exist due to the definition (4.2). Thus, we have $p_{n-j}^{(n,\xi)} = q_{n-j}^{(n,\xi)} = \sum_{l=j}^{n} \theta_{l-j}^{(l,\xi)}$. The equality (4.5) follows from (4.4), so the proof is completed. □

Lemma 4 ([9]). *If the VS-gBDF2 kernels* $A_{n-k}^{(n,\xi)}$ *defined in (2.7) are positive semi-definite, then the DOC kernels* $\theta_{n-k}^{(n,\xi)}$ *defined in (4.1) are also positive semi-definite in the sense that for any real sequence* $\{\omega_j\}_{j=2}^{n}$,

$$\sum_{k=2}^{n} \omega_k \sum_{j=2}^{k} \theta_{k-j}^{(k,\xi)} \omega_j \ge 0, \quad \forall n \ge 1.$$

Lemma 5. *The DOC kernels* $\theta_{n-k}^{(n,\xi)}$ *in (4.1) fulfill* $\sum_{j=1}^{n} \theta_{n-j}^{(n,\xi)} \equiv \tau_n$.

Proof. If follows from the definition (2.6) that $\sum_{l=1}^{j} A_{j-l}^{(j,\xi)} \tau_l \equiv 1$. Together with exchanging the order of summation, one can obtain

$$\sum_{j=1}^{n} \theta_{n-j}^{(n,\xi)} = \sum_{j=1}^{n} \theta_{n-j}^{(n,\xi)} \sum_{l=1}^{j} A_{j-l}^{(j,\xi)} \tau_l = \sum_{l=1}^{n} \tau_l \sum_{j=l}^{n} \theta_{n-j}^{(n,\xi)} A_{j-l}^{(j,\xi)} = \tau_n.$$

The proof is completed. □

Lemma 6. *The DOC kernels* $\theta_{n-k}^{(n,\xi)}$ *in (4.1) exhibit an explicit formula*

$$\theta_{n-k}^{(n,\xi)} = \frac{(1+r_k)\tau_n}{1+2r_k\xi} \prod_{i=k+1}^{n} \frac{(2\xi-1)r_i}{1+2r_i\xi}, \quad n \ge 2. \tag{4.6}$$

Proof. By the definition (4.1) and the VS-gBDF2 kernels in (2.7), one has

$$\theta_0^{(n,\xi)} A_0^{(n,\xi)} = 1 \quad \text{and} \quad \theta_{n-k}^{(n,\xi)} A_0^{(k,\xi)} = -\frac{A_1^{(k+1,\xi)}}{A_0^{(k+1,\xi)}} \theta_{n-k-1}^{(n,\xi)} A_0^{(k+1,\xi)}.$$

Thus, a simple induction yields

$$\theta_{n-k}^{(n,\xi)} A_0^{(k,\xi)} = \prod_{i=k+1}^{n} \frac{(2\xi-1)r_i^2}{1+2r_i\xi} = \frac{\tau_n}{\tau_k} \prod_{i=k+1}^{n} \frac{(2\xi-1)r_i}{1+2r_i\xi}.$$

It yields the claimed formula and completes the proof. □

Proposition 2. *Let τ be the maximum time step and $r_k \le r_* < \bar{r}_\xi$ for some $r^* > 0$. Then the DCC kernels $p_{n-k}^{(n)}$ defined in (4.2) satisfy*

$$p_{n-j}^{(n,\xi)} = \frac{(1+r_j)}{1+2r_j\xi} \sum_{k=j}^{n} \tau_k \prod_{i=j+1}^{k} \frac{(2\xi-1)r_i}{1+2r_i\xi}, \quad 2 \le j \le n, \tag{4.7}$$

$$\sum_{j=1}^{n} p_{n-j}^{(n,\xi)} = t_n, \quad p_{n-j}^{(n,\xi)} \le \frac{(1+r_j)}{1+2r_j\xi} \cdot \frac{1}{1-R_*} \sqrt{\tau\tau_j}, \tag{4.8}$$

where $\prod_{i=j+1}^{k}(\cdot) = 1$ for $j \ge k$ and

$$R_* := \frac{(2\xi-1)r_*^{\frac{3}{2}}}{1+2r_*\xi} < 1. \tag{4.9}$$

Proof. The first claim (4.7) and the left equality in (4.8) can be verified by (4.4) and (4.6). Noting that the function $g_\xi(x) := \frac{(2\xi-1)x^{3/2}}{1+2\xi x}$ is increasing with respect to x for any $\xi \ge 1$, we then have

$$p_{n-j}^{(n,\xi)} \le \frac{(1+r_j)\sqrt{\tau_j}}{1+2r_j\xi} \sum_{k=j}^{n} \sqrt{\tau_k} \prod_{i=j+1}^{k} R_* = \frac{(1+r_j)\sqrt{\tau\tau_j}}{1+2r_j\xi} \sum_{k=0}^{n-j} R_*^k. \tag{4.10}$$

It follows from the monotonicity of the function $g_\xi(x)$ and the definition (2.8) that $R_* < 1$. Hence, the proof is completed by applying (4.9) to (4.10). □

5. Stability and convergence analysis

We now consider the stability and convergence analysis for VS-gBDF2 scheme (2.7).

Lemma 7 (A discrete Grönwall inequality). *Assume $\lambda > 0$ and the sequences $\{v_j\}_{j=1}^N$ and $\{\eta_j\}_{j=0}^N$ are nonnegative. If*

$$v_n \le \lambda \sum_{j=1}^{n-1} \tau_j v_j + \sum_{j=0}^{n} \eta_j, \quad \text{for} \quad 1 \le n \le N,$$

then it holds

$$v_n \le \exp\left(\lambda t_{n-1}\right) \sum_{j=0}^{n} \eta^j, \quad \text{for} \quad 1 \le n \le N.$$

The standard induction hypothesis can prove lemma 7 and the proof is omitted here.

5.1. Stability analysis

We hereby present the stability of VS-gBDF2 scheme (2.7), followed by the ensuing theorem.

Theorem 2. *Let $\underline{r}_\xi$ and $\bar{r}_\xi$ be defined in condition* **C1**. *Assume the time-step ratio satisfies $\underline{r}_\xi \le r_k \le r_*$ for any $r_* < \bar{r}_\xi$. Then the solution u^n of VS-gBDF2 scheme (2.5) is unconditionally stable in the L^2-norm.*

If $\kappa > 0$ and the maximum time-step size $\tau \le 1/(4\kappa(2\xi-1)\xi)$, it holds

$$\|u^n\| \le 2\exp\left(4(2\xi-1)(\xi+\bar{r}_\xi(\xi-1))t_{n-1}\right) \cdot \left(\|u^1\| + 2(2\xi-1)(t_n \max_{2\le k\le n} \|E_\xi f^k\| + \frac{1}{1-R_*}\sqrt{\tau\tau_1}\|\partial_\tau u^1\|)\right), \tag{5.1}$$

where $\partial_\tau u^1 := \nabla_\tau u^1/\tau_1$ and R_* is defined by (4.9).

If $\kappa \le 0$, it holds

$$\|u^n\| \le \|u^1\| + 2(2\xi-1)\Big(t_n \max_{2\le k\le n} \|E_\xi f^k\| + \frac{1}{1-R_*}\sqrt{\tau\tau_1}\|\partial_\tau u^1\|\Big). \tag{5.2}$$

Proof. Applying the property (4.3) of DOC kernels to (2.5), we have for $k \ge 2$

$$\nabla_\tau u^k = \sum_{j=2}^{k} \theta_{k-j}^{(k,\xi)}(\Delta E_\xi u^j + \kappa E_\xi u^j + E_\xi f^j) - \theta_{k-2}^{(k,\xi)} A_1^{(2,\xi)} \nabla_\tau u^1. \tag{5.3}$$

Noting the positive semi-definiteness of the DOC kernels in Lemma 4, we have

$$\sum_{k=2}^{n}\sum_{j=2}^{k}\langle E_\xi u^k, \theta_{k-j}^{(k,\xi)}\Delta E_\xi u^j\rangle = -\sum_{k=2}^{n}\sum_{j=2}^{k}\langle E_\xi \nabla u^k, \theta_{k-j}^{(k,\xi)} E_\xi \nabla u^j\rangle \le 0. \tag{5.4}$$

Taking the inner product with $2E_\xi u^k$ on both sides of (5.3), summing the resulting from 2 to n and using (5.4), we have for $2 \le n \le N$

$$2\sum_{k=2}^{n}\langle \nabla_\tau u^k, E_\xi u^k\rangle \le 2\sum_{k=2}^{n}\sum_{j=2}^{k}\theta_{k-j}^{(k,\xi)}\langle E_\xi u^k, \kappa E_\xi u^j + E_\xi f^j\rangle - 2\langle A_1^{(2,\xi)}\nabla_\tau u^1, \sum_{k=2}^{n}\theta_{k-2}^{(k,\xi)} E_\xi u^k\rangle. \tag{5.5}$$

Applying the identity

$$2\langle \nabla_\tau u^k, E_\xi u^k\rangle = \|u^k\|^2 - \|u^{k-1}\|^2 + (1+2\xi)\|\nabla_\tau u^k\|^2$$

to (5.5), we have

$$\|u^n\|^2 \le \|u^1\|^2 + 2\sum_{k=2}^{n}\sum_{j=2}^{k}\theta_{k-j}^{(k,\xi)}\langle E_\xi u^k, \kappa E_\xi u^j + E_\xi f^j\rangle - 2\langle A_1^{(2,\xi)}\nabla_\tau u^1, \sum_{k=2}^{n}\theta_{k-2}^{(k,\xi)} E_\xi u^k\rangle. \tag{5.6}$$

If $\kappa \le 0$ in (5.6), we set $u_\xi^k = \xi\|u^k\| + (\xi-1))\|u^{k-1}\|$. Then applying the Cauchy-Schwarz inequality and Lemma 4 to (5.6), one has

$$\|u^n\|^2 \le \|u^1\|^2 + 2\sum_{k=2}^{n}\sum_{j=2}^{k}\theta_{k-j}^{(k,\xi)}u_\xi^k\|E_\xi f^j\| - 2A_1^{(2,\xi)}\|\nabla_\tau u^1\|\sum_{k=2}^{n}\theta_{k-2}^{(k,\xi)}u_\xi^k \tag{5.7}$$

Set $\|u^{n_0}\| = \max\limits_{1\le k\le n}\|u^k\|$, it follows from (5.7) that

$$\|u^{n_0}\|^2 \le \|u^1\|\|u^{n_0}\| + 2(2\xi-1)\|u^{n_0}\|\left(\sum_{k=2}^{n}p_{n-k}^{(n,\xi)}\|E_\xi f^k\| - A_1^{(2,\xi)}p_{n-2}^{(n,\xi)}\|\nabla_\tau u^1\|\right),$$

where one exchanges the order of summation and uses property (4.4). Eliminating a $\|u^{n_0}\|$ on both sides and noting $n_0 \le n$, we have

$$\|u^n\| \le \|u^{n_0}\| \le \|u^1\| + 2(2\xi-1)\left(\sum_{k=2}^{n}p_{n-k}^{(n,\xi)}\|E_\xi f^k\| - A_1^{(2,\xi)}p_{n-2}^{(n,\xi)}\|\nabla_\tau u^1\|\right). \tag{5.8}$$

By (2.6) and (4.8), one has

$$-A_1^{(2,\xi)}p_{n-2}^{(n,\xi)} \;\le\; \frac{r_2^{\frac{3}{2}}(2\xi-1)}{1+2r_2\xi}\frac{1}{1-R_*}\sqrt{\frac{\tau}{\tau_1}} \;\le\; \frac{1}{1-R_*}\sqrt{\frac{\tau}{\tau_1}}. \tag{5.9}$$

Inserting (5.9) into (5.8), using the property (4.8), one can obtain (5.2).

If $\kappa > 0$ in (5.6), we apply the Cauchy-Schwarz inequality to have

$$\|u^n\|^2 \le \|u^1\|^2 + 2\kappa\sum_{k=2}^{n}\sum_{j=2}^{k}\theta_{k-j}^{(k,\xi)}u_\xi^j\|E_\xi u^k\| + 2\sum_{k=2}^{n}\sum_{j=2}^{k}\theta_{k-j}^{(k,\xi)}u_\xi^k\|E_\xi f^j\| - 2A_1^{(2,\xi)}\|\nabla_\tau u^1\|\sum_{k=2}^{n}\theta_{k-2}^{(k,\xi)}u_\xi^k. \tag{5.10}$$

Selecting n_0 such that $\|u^{n_0}\| = \max_{1\le k\le n}\|u^k\|$, one has

$$\|u^{n_0}\|^2 \le \|u^1\|\|u^{n_0}\| + 2(2\xi-1)\|u^{n_0}\|\sum_{k=2}^{n}\left(\kappa\tau_k\|E_\xi u^k\| + p_{n-k}^{(n,\xi)}\|E_\xi f^k\|\right) - 2(2\xi-1)A_1^{(2,\xi)}p_{n-2}^{(n,\xi)}\|\nabla_\tau u^1\|\|u^{n_0}\|, \tag{5.11}$$

where one uses Lemma 5 and the property (4.4). Eliminating a $\|u^{n_0}\|$ for both sides of (5.11), we further have

$$\|u^n\| \le \|u^{n_0}\| \le \|u^1\| + 2(2\xi-1)\sum_{k=2}^{n}\left(\kappa\tau_k\|E_\xi u^k\| + p_{n-k}^{(n,\xi)}\|E_\xi f^k\|\right) + 2(2\xi-1)\frac{1}{1-R_*}\sqrt{\tau\tau_1}\|\partial_\tau u^1\|, \tag{5.12}$$

where the inequality (5.9) is used. Taking the maximum time-step size $\tau \le \frac{1}{4\kappa(2\xi-1)\xi}$, one has

$$\begin{aligned}\sum_{k=2}^{n}\tau_k\|E_\xi u^k\| &\le \xi\sum_{k=2}^{n}\tau_k\|u^k\| + \bar{r}_\xi(\xi-1)\sum_{k=1}^{n-1}\tau_k\|u^k\| \\ &\le (\xi+\bar{r}_\xi(\xi-1))\sum_{k=1}^{n-1}\tau_k\|u^k\| + \xi\tau_k\|u^n\| \\ &\le (\xi+\bar{r}_\xi(\xi-1))\sum_{k=1}^{n-1}\tau_k\|u^k\| + \frac{1}{4(2\xi-1)\kappa}\|u^n\|.\end{aligned} \tag{5.13}$$

Inserting (5.13) into (5.12), we have

$$\|u^n\| \le 2\|u^1\| + 4(2\xi-1)(\xi+\bar{r}_\xi(\xi-1))\sum_{k=2}^{n-1}\tau_k\|u^k\| + 4(2\xi-1)(\sum_{k=2}^{n}p_{n-k}^{(n,\xi)}\|E_\xi f^k\| + \frac{1}{1-R_*}\sqrt{\tau\tau_1}\|\partial_\tau u^1\|).$$

Then together with the Grönwall inequality in Lemma 7 and (4.8), one can derive the result (5.1). The proof is completed. □

5.2. Convergence analysis

Set $e^n := u(t_n, x) - u^n(x)$ $(n \geq 1)$. It follows from (2.5) that the error function $e^n (2 \leq n \leq N)$ satisfies

$$D_2^{\xi} e^n = \Delta E_{\xi} e^n + \kappa E_{\xi} e^n + \eta^n + R\Big(\Delta u(t_n) + \kappa u(t_n) + f^n\Big), \tag{5.14}$$

with the truncation errors $\eta^n = D_2^{\xi} u(t_n) - u_t(t_{n-1} + \tau_n \xi)$ and $R(v) = v(t_{n-1} + \tau_n \xi)) - E_{\xi} v(t_n)$.

We now present the estimates of the truncation errors η^n and $R(\Delta u(t_n) + \kappa u(t_n) + f^n)$ in the following two lemmas, their proofs are left to Appendix A for brevity.

Lemma 8. *The truncation error $\eta^n := D_2 u(t_n) - u_t(t_n)$ can be bounded by*

$$\|\eta^n\| \leq c_1 \tau^2, \quad 2 \leq n \leq N, \tag{5.15}$$

where the constant c_1 is given by

$$c_1 := \Big(\frac{1}{2}(1 + (2\xi - 1)\bar{r}_{\xi})(1 + \xi(\xi - 1)) + (\xi - \frac{1}{2})((1 + \bar{r}_{\xi})^2 + (\xi - 1)(1 + \bar{r}_{\xi}\xi))\Big)\|u_{ttt}\|_{L^{\infty}(0,T;L^2(\Omega))}. \tag{5.16}$$

Lemma 9. *Assume that $v \in C^2([0,T])$, then the truncation error $R(v(t_n)) := v(t_{n-1} + \tau_n \xi) - E_{\xi} v(t_n)$ $(1 \leq j \leq N)$ can be bounded by*

$$\|R(v(t_n))\| \leq \Big((\xi - 1)^2 + 1\Big)\frac{\tau_n^2}{2}\|v\|_{L^{\infty}(0,T;L^2(\Omega))}. \tag{5.17}$$

Furthermore, we have

$$\|R\Big(\Delta u(t_n) + \kappa u(t_n) + f^n\Big)\| \leq c_2 \tau_n^2, \tag{5.18}$$

where

$$c_2 := ((\xi - 1)^2 + 1)\frac{(|\kappa| + 1)}{2}(\|u_{tt}\|_{L^{\infty}(0,T;H^2(\Omega))} + \|f_{tt}\|_{L^{\infty}(0,T;L^2(\Omega))}). \tag{5.19}$$

Theorem 3. *Let $\underline{r}_{\xi}$ and $\bar{r}_{\xi}$ be defined by condition* **C1**. *Assume $\underline{r}_{\xi} \leq r_k \leq r_*$ for any $r_* < \bar{r}_{\xi}$. Then the discrete solution u^n to VS-gBDF2 scheme (2.5) has the second-order convergence in the L^2-norm.*

If $\kappa > 0$ and the maximum time-step size $\tau \leq 1/(4\kappa(2\xi - 1)\xi)$, it holds

$$\|e^n\| \leq 2\exp\Big(4(2\xi - 1)(\xi + \bar{r}_{\xi}(\xi - 1))t_{n-1}\Big) \cdot \Big(\|e^1\| + 2(2\xi - 1)\Big(t_n(c_1 + c_2)\tau^2 + \frac{1}{1 - R_*}\sqrt{\tau\tau_1}\|\partial_{\tau} e^1\|\Big)\Big), \tag{5.20}$$

where R_*, c_1, c_2 are defined by (4.9), (5.16) and (5.19) respectively.

If $\kappa \leq 0$, it holds

$$\|e^n\| \leq \|e^1\| + 2(2\xi - 1)\Big(t_n(c_1 + c_2)\tau^2 + \frac{1}{1 - R_*}\sqrt{\tau\tau_1}\|\partial_{\tau} e^1\|\Big). \tag{5.21}$$

Theorem 3 can be derived immediately by using Theorem 2, (5.14) and Lemmas 8 and 9, we omit its proof here.

5.3. The backward Euler method for the start value

In this subsection, we consider the stability and convergence of VS-gBDF2 method if the backward Euler method

$$\partial_{\tau} u^1 = \Delta u^1 + \kappa u^1 + f^1 \tag{5.22}$$

is used to calculate u^1 by beginning with the following lemma.

Lemma 10. *Let $a, b, c > 0$. If $a^2 \leq b^2 + ac$, then we have $a \leq b + c$.*

Proof. By simple calculation, one has

$$(a - \frac{c}{2})^2 - (\frac{c}{2} + b)^2 = a^2 - ac - b^2 - bc \leq -bc < 0,$$

which implies

$$|a - \frac{c}{2}| < |\frac{c}{2} + b|.$$

Hence, we infer that $a \leq b + c$ for $a > c/2$. Since the claim is immediately proved if $a \leq c/2$, the proof is completed. □

Theorem 4. *Let the start value u^1 be calculated by the backward Euler method (5.22). Let $\underline{r}_{\xi}$ and $\bar{r}_{\xi}$ be defined in condition* **C1**. *Assume the time-step ratio satisfies $\underline{r}_{\xi} \leq r_k \leq r_*$ for any $r_* < \bar{r}_{\xi}$. Then the solution u^n of VS-gBDF2 scheme (2.5) is unconditionally stable in the L^2-norm. If $\kappa > 0$ and the maximum time-step size $\tau \leq \min\{1, 1/(4\kappa(2\xi - 1)\xi)\}$, it holds*

$$\|u^n\| \leq 4\exp\Big(4(2\xi - 1)(\xi + \bar{r}_{\xi}(\xi - 1))t_{n-1}\Big)C^* \cdot \Big(\|u^0\| + \sqrt{\tau}\|\nabla u^0\| + \tau_1\|f^1\| + t_n \max_{2 \leq k \leq n}\|E_{\xi} f^k\|\Big), \tag{5.23}$$

where $\partial_{\tau} u^1 := \nabla_{\tau} u^1/\tau_1$, R_* is defined by (4.9) and $C^* := (2\xi - 1)(2 + (1 + |\kappa|)/(1 - R^*))$. If $\kappa \leq 0$ and $\tau \leq 1$, it holds

$$\|u^n\| \leq 2C^*\Big(\|u^0\| + \sqrt{\tau}\|\nabla u^0\| + \tau_1\|f^1\| + t_n \max_{2 \leq k \leq n}\|E_{\xi} f^k\|\Big). \tag{5.24}$$

Proof. We shall estimate $\|u^1\|$ and $\|\partial_\tau u\|$ respectively. On the one hand, taking the inner products with $2u^1$ on both size of (5.22) and using the equality $2a(a-b) = a^2 - b^2 + (a-b)^2$, one has

$$\|u^1\|^2 - \|u^0\|^2 + \|\nabla_\tau u^1\|^2 + 2\tau_1\|\nabla u^1\|^2 = 2\kappa\tau_1\|u^1\|^2 + 2\tau_1(f^1, u^1). \tag{5.25}$$

If $\kappa \le 0$, the identity (5.25) leads to

$$\|u^1\|^2 + \|\nabla_\tau u^1\|^2 \le \|u^0\|^2 + 2\tau_1\|f^1\|\|u^1\|. \tag{5.26}$$

Applying Lemma 10 to (5.26), one has

$$\|u^1\| \le \|u^0\| + 2\tau_1\|f^1\|. \tag{5.27}$$

If $\kappa > 0$, set $\tau \le 1/(4\kappa(2\xi-1)\xi) \le 1/(4\kappa)$ in (5.25), one produces

$$\|u^1\|^2 \le 2\|u^0\|^2 + 4\tau_1\|f^1\|\|u^1\|, \tag{5.28}$$

which together with Lemma 10 yields

$$\|u^1\| \le 2\|u^0\| + 4\tau_1\|f^1\|. \tag{5.29}$$

On the other hand, taking the inner products with $2\nabla_\tau u^1$ on both size of (5.22) and using the equality $2a(a-b) = a^2 - b^2 + (a-b)^2$, one has

$$\frac{2}{\tau_1}\|\nabla_\tau u^1\|^2 + \|\nabla u^1\|^2 + \|\nabla_\tau \nabla u^1\|^2 = \|\nabla u^0\|^2 + 2\kappa\|\nabla_\tau u^1\|^2 + 2(f^1 + \kappa u^0, \nabla_\tau u^1). \tag{5.30}$$

If $\kappa \le 0$, the identity (5.30) produces

$$\frac{2}{\tau_1}\|\nabla_\tau u^1\|^2 \le \|\nabla u^0\|^2 - 2\kappa\|u^0\|\|\nabla_\tau u^1\| + 2\|f^1\|\|\nabla_\tau u^1\|. \tag{5.31}$$

Applying Lemma 10 to (5.31), one has

$$\|\partial_\tau u^1\| \le \frac{1}{\sqrt{\tau_1}}\|\nabla u^0\| - \kappa\|u^0\| + \|f^1\|. \tag{5.32}$$

If $\kappa > 0$, from (5.30), we have

$$\frac{(2 - 2\kappa\tau_1)}{\tau_1}\|\nabla_\tau u^1\|^2 \le \|\nabla u^0\|^2 + 2\kappa\|u^0\|\|\nabla_\tau u^1\| + 2\|f^1\|\|\nabla_\tau u^1\|.$$

Since $\tau \le 1/(4\kappa(2\xi-1)\xi) \le 1/(4\kappa)$, one can find $2 - 2\kappa\tau_1 \ge 1$. Thus, we arrive at

$$\frac{\|\nabla_\tau u^1\|^2}{\tau_1} \le \|\nabla u^0\|^2 + 2\kappa\|u^0\|\|\nabla_\tau u^1\| + 2\|f^1\|\|\nabla_\tau u^1\|,$$

which together with Lemma 10 yields

$$\|\partial_\tau u^1\| \le \frac{1}{\sqrt{\tau_1}}\|\nabla u^0\| + 2\kappa\|u^0\| + 2\|f^1\|. \tag{5.33}$$

Then inserting (5.29) and (5.33) into (5.1), one obtains (5.23). And inserting (5.27) and (5.32) into (5.2), one has (5.24). The proof is completed. □

From (5.22), we have

$$\partial_\tau e^1 = \Delta e^1 + \kappa e^1 + \eta^1 \tag{5.34}$$

with the truncation error $\eta^1 = u(t_1) - u_t(t_1)$. The Taylor expansion produces

$$\|\eta^1\| = \|\int_0^{t_1} u_{tt}(s)s\mathrm{d}s\| \le \|u_{tt}\|_{L^\infty(0,T;L^2(\Omega))}\tau_1. \tag{5.35}$$

Then we have the following convergence theorem.

Theorem 5. *Let the start value u^1 be calculated by the backward Euler method (5.22). Let $\underline{r}_\xi$ and $\bar{r}_\xi$ be defined by condition* **C1**. *Assume the time-step ratio satisfies $\underline{r}_\xi \le r_k \le r_*$ for any $r_* < \bar{r}_\xi$. If $\kappa > 0$ and the maximum time-step size satisfies $\tau \le \min\{1, \frac{1}{4\kappa(2\xi-1)\xi}\}$, it holds*

$$\|e^n\| \le 4\exp\Big(4(2\xi-1)(\xi + \bar{r}_\xi(\xi-1))t_{n-1}\Big)C^*\Big(\|e^0\| + \sqrt{\tau}\|\nabla e^0\| + C_1\tau^2\Big),$$

where $C_1 := t_n(c_1 + c_2) + \|u_{tt}\|_{L^\infty(0,T;L^2(\Omega))}$ and R_*, c_1, c_2 are defined by (4.9), (5.16) and (5.19) respectively. If $\kappa < 0$ and $\tau \le 1$, it holds

$$\|e^n\| \le 2C^*\Big(\|e^0\| + \sqrt{\tau}\|\nabla e^0\| + C_1\tau^2\Big).$$

6. Numerical examples

We present in this section several numerical examples to validate our analysis and to show the effectiveness of the VS-gBDF2 schemes.

Table 1
Errors and time convergence orders for $\xi = 2.5$ ($\bar{r}_\xi = 1.91, \underline{r}_\xi = 0.29$).

N	e(N)	order	$r_{\max}$	$r_{\min}$
160	2.796e-04	–	1.91	0.29
320	6.112e-05	2.19	1.91	0.29
640	1.607e-05	1.93	1.91	0.29
1280	3.926e-06	2.03	1.91	0.29

Table 2
Errors and time convergence orders for $\xi = 5$ ($\bar{r}_\xi = 1.42, \underline{r}_\xi = 0.$).

N	error	order	$r_{\max}$	
160	4.116e-04	–	1.42	0.40
320	9.254e-05	2.15	1.42	0.40
640	2.486e-05	1.90	1.42	0.40
1280	5.423e-06	2.20	1.42	0.40

6.1. *The accuracy and efficiency of scheme (2.5)*

We now present two examples to demonstrate the accuracy, energy decay and efficiency of the proposed scheme (2.5) for (2.4). Set $\Omega = (0, 1)$, $\kappa = 0$ and $T = 2$. Take the spatial length $h = 1/M$ for a positive integer M, the discrete grids $\Omega_h = \{x_i = ih | 0 \le i \le M\}$. In the simulations, we use the standard finite difference method for spatial discretization and choose the number of spatial mesh M by $M = N$.

Example 1. In this example, we investigate the quantitative accuracy of fully discrete scheme. To obtain the variable time steps, the time steps are generated by $\tau_k = T\lambda_k/\Lambda$ for $1 \le k \le N$, where $\Lambda = \sum_{k=1}^{N} \lambda_k$ and λ_k is drawn as follows. For each $k = 1, 2, \cdots, N$, randomly choose a number from a uniform distribution on $(0, 1)$, denoted as $\hat{\lambda}_k$ and then truncate the value by $\lambda_k = \min\{\max\{\underline{r}_\xi \lambda_{k-1}, \hat{\lambda}_k\}, \bar{r}_\xi \lambda_{k-1}\}$ if $k \ge 2$. Thus, the ratios of adjacent time steps satisfy **C1**, i.e., $\underline{r}_\xi \le r_k \le \bar{r}_\xi$.

To investigate the convergence order of the fully discrete scheme, we take $f = (\pi^2 - \kappa + 1)\exp(t)\sin(\pi x)$ in (2.4) so that the exact solution is $u = \exp(t)\sin(\pi x)$. The discrete L^2-norm at the final time T and the convergence rate are calculated respectively by

$$e(N) = h\sqrt{\sum_{1\le i\le M} (u(x_i, T) - u_h^N(x_i))^2}, \quad Order = \log_2(e(N)/e(2N)).$$

In each run, the discrete L^2-norm, convergence rates, maximal and minimal adjacent time-step ratios are listed in Tables 1–2 for different ξ ($\xi = 2.5$ and $\xi = 5$). Tables 1 and 2 show the second order convergence rates, which agrees with the results in Theorem 4.

Example 2. We now consider the evolution of modified energy and numerical solutions by taking $\kappa = 0$ and the initial values as $u_0(x) = \sin(\pi x)$ in (2.4). We introduce the adaptive time-stepping strategy [12,13]

$$\tau_{n+1} := \min\left\{ \max\left\{ \underline{r}_\xi \tau_n, \tau_{\min}, \frac{\tau_{\max}}{\sqrt{1 + \alpha|\partial_\tau \mathcal{E}^n|^2}} \right\}, \bar{r}_\xi \tau_n \right\}, \tag{6.1}$$

where $\mathcal{E}^n$ denotes the discrete energy defined in (3.2), $\bar{r}_\xi$ and $\underline{r}_\xi$ are defined by **C1**, $\tau_{\min}$ and $\tau_{\max}$ are the minimum and maximum time steps respectively and α is a tunable parameter related to the level of the adaptivity. Set $\tau_{\min} = 1e-5$, $\tau_{\max} = 1e-3$, $\alpha = 0.5$, $T = 0.5$ and $N = 1000$.

We compare the evolution of the modified energy and the maximum norm of solutions $\|U_h^n\|_\infty$ under different time steps (uniform and adaptive steps) and different values of ξ. We also choose the classical uniform-step BDF2 scheme (i.e., $\xi = 1$ and $r_n \equiv 1$) with $\tau = 1e-5$ as the reference solution. Fig. 6.1a shows that the modified energy $\mathcal{E}^n$ is decreasing all the time, which is consistent with Theorem 1. From Figs. 6.1a and 6.1b, the modified energy $\mathcal{E}^n$ and maximum norm solution $\|U_h^N\|_\infty$ appear to be closer to the reference values when employing the adaptive time-stepping strategy (6.1). It indicates that the adaptive time-stepping strategy (6.1) can significantly improve the accuracy.

We now consider the CPU time for different adaptive and fixed step sizes. Fig. 6.1c plots the adaptive time steps, which shows the adaptive step is relatively large when energy evolves slowly while relatively small when energy evolves quickly. Figs. 6.1d shows that the CPU time for the adaptive scheme is similar to that of the large time step $\tau = 1e-3$, and significantly less than the CPU time with the small time step $\tau = 1e-5$. And as shown in Figs. 6.1a and 6.1b, the accuracy of the adaptive time-step strategy is comparable to that of the small time step $\tau = 1e-5$, indicating that adaptive strategy can significantly reduce the CPU time without sacrificing accuracy.

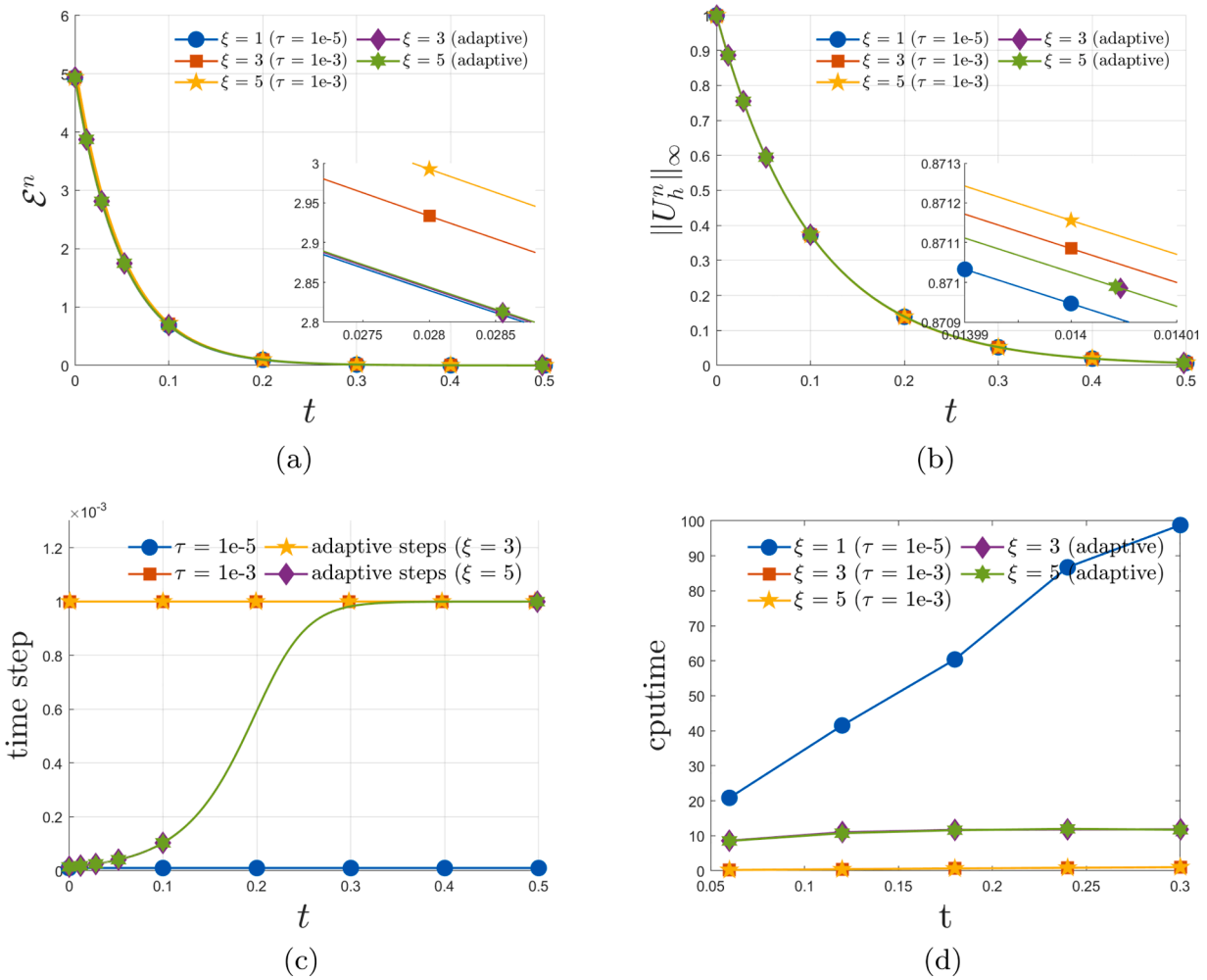

Fig. 6.1. (a): Evolution of modified discrete energy; (b): Evolution of maximum norm of solutions; (c): adaptive time-step size; (d): CPU time contrast.

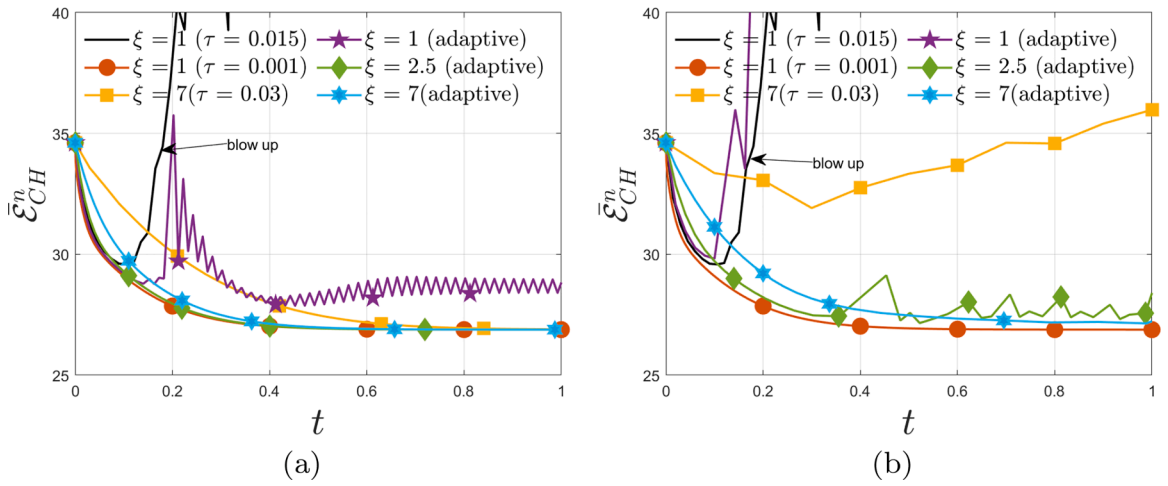

Fig. 6.2. Evolution of original discrete energy:(a) $\tau_{\min} = 0.01, \tau_{\max} = 0.03$ in strategy (6.1);(b) $\tau_{\min} = 0.02, \tau_{\max} = 0.1$ in strategy (6.1).

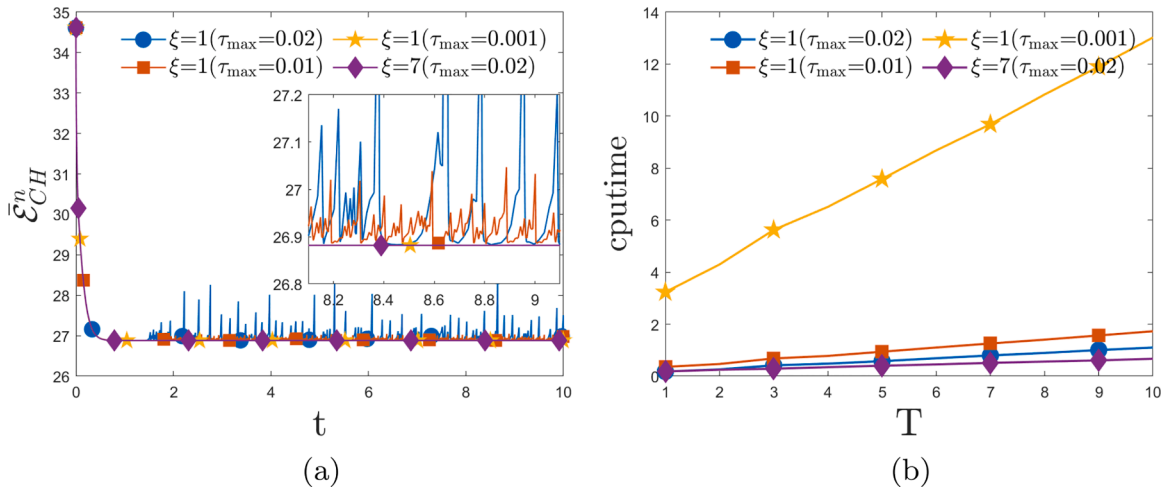

Fig. 6.3. Evolution of original discrete energy and CPU times with a minimal time step $\tau_{\min} = 10^{-5}$ and varying maximal time steps $\tau_{\max}$ in the adaptive strategy (6.1) for $\xi = 1$ and $\xi = 7$.

6.2. An application to Cahn-hilliard equation

Example 3. As illustrated in Fig 1.1 or [11, Figure 1], the gBDF2 scheme preforms better stability as ξ increases. We now present an example to investigate the stability of VS-gBDF2 scheme by consider the Cahn-Hilliard equation

$$\partial_t u = \Delta \mu \quad \text{with} \quad \mu = -\Delta u + \frac{1}{\varepsilon^2}(u^3 - u), \quad (\boldsymbol{x}, t) \in \Omega \times (0, T], \tag{6.2}$$

with periodic boundary conditions and the initial condition $u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x})$. The system (6.2) can be viewed as the H^{-1} gradient flow of the Ginzburg-Landau type energy functional

$$E[u] = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + \frac{(u^2-1)^2}{4\varepsilon^2} \right) \mathrm{d}\boldsymbol{x}, \tag{6.3}$$

which admits the following energy dissipation law

$$\frac{\mathrm{d}}{\mathrm{d}t} E[u(t)] = \int_\Omega \frac{\delta E}{\delta u} \partial_t u \mathrm{d}\boldsymbol{x} = -\int_\Omega |\nabla \mu|^2 \mathrm{d}\boldsymbol{x}. \tag{6.4}$$

By adopting the Fourier-spectral method to discretize the space with N Fourier modes for each directions, the VS-gBDF2 scheme for (6.2) is given by

$$D_2^\xi U_h^n = -\Delta^2 E_\xi U_h^n + \frac{1}{\varepsilon^2}\Delta((\hat{U}_h^n)^3 - \hat{U}_h^n), \quad 1 \le n \le N, \tag{6.5}$$

where

$$\hat{U}_h^n = \begin{cases} (1 + r_n\xi)U_h^{n-1} - r_n\xi U_h^{n-2}, & n \ge 2, \\ U_h^{n-1}, & n = 1. \end{cases} \tag{6.6}$$

We now consider the coalescence of two kissing bubbles by taking $\Omega = (0, 2\pi)^2$, $\varepsilon^2 = 0.1$ and the initial value as

$$u_0(\boldsymbol{x}) = \max_{1 \le i \le 2} \tanh\left(R - \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{4\varepsilon^2} \right),$$

where $x_1 = \pi - 1, y_1 = \pi, x_2 = \pi + 1, y_2 = \pi, R = 1$. The adaptive time-stepping strategy (6.1) is employed by defining the original discrete energy

$$\bar{\mathcal{E}}^n_{\mathrm{CH}} = \int_\Omega \left(\frac{1}{2}|\nabla U_h^n|^2 + \frac{((U_h^n)^2 - 1)^2}{4\varepsilon^2} \right) \mathrm{d}\boldsymbol{x}. \tag{6.7}$$

Set $\alpha = 0.1$, the final time $T = 1$ and the number of Fourier modes $N = 128$.

We compare the evolution of the original discrete energy under different choices of $\tau_{\min}$ and $\tau_{\max}$ in the adaptive time strategy (6.1). As documented in [31], smaller time steps are usually required to guarantee accuracy and energy stability for semi-implicit scheme (6.5). When we adopt a large uniform time step $\tau = 0.015$ in Fig. 6.2, the original discrete energy blows up at $t = 0.2$. This phenomenon can be significantly improved when ξ increases as we show in Fig. 6.2. Even when the minimum step size $\tau_{\min} > 0.015$, as we show in Fig. 6.2b, the original discrete energy can still ensure accuracy when $\xi = 7$. Thus, the VS-gBDF2 scheme performs better

stability as ξ increases. This result is consistent with the gradually increasing absolute stability region illustrated in Fig. 1.1 as beta increases.

In Fig. 6.2, we also show the energy evolution curves computed by VS-gBDF2 scheme with adaptive time-stepping strategy (6.1) and gBDF2 scheme with uniform time step ($\tau = \tau_{\max}$) when $\xi = 7$. It can be seen that the VS-gBDF2 scheme significantly outperforms the gBDF2, highlighting the superiority of VS-gBDF2 scheme and adaptive time-stepping strategy (6.1).

To demonstrate the advantage of our VS-gBDF2 (i.e., $\xi > 1$), we further investigate the energy stability and the corresponding CPU times by comparing with the classical VS-BDF2 scheme. In the simulations, we take a minimal time step $\tau_{\min} = 10^{-5}$ and vary the maximal time steps $\tau_{\max}$ in the adaptive strategy (6.1). Fig. 6.3a shows that the adaptive VS-gBDF2 scheme with $\xi = 7$ remains energy-stable even for a larger size $\tau_{\max} = 0.02$, but the energy of the classical adaptive VS-BDF2 scheme becomes oscillating for both $\tau_{\max} = 0.02$ and $\tau_{\max} = 0.01$, and become stable for a smaller $\tau_{\max} = 0.001$. Fig. 6.3b shows that the corresponding CPU times. It follows from Figs. 6.3a and 6.3b that our VS-gBDF2 scheme for $\xi = 7$ requires much less CPU time than the classical VS-BDF2 scheme to remain the energy stable.

7. Concluding remarks

We considered in this work a variable step generalized BDF2 scheme based on the Taylor expansions at time $t_{n-1} + \tau_n\xi$ ($\xi \geq 1$), and provided an asymptotically compatible analysis on the discrete energy dispassion law, stability and convergence of the VS-gBDF2 scheme for any $\xi \geq 1$ under the ratio restriction **C1**. We also introduce an adaptive time-step strategy and provide numerical examples to demonstrate our theoretical analysis and validate the effectiveness of the adaptive time-step strategy.

Our analysis on VS-gBDF2 scheme can be extended to a family of nonlinear parabolic equations and combined with other spatial discretizations. In the future, we will consider the VS-gBDF2 scheme for nonlinear problems.

Data availability

No data was used for the research described in the article.

Declaration of competing interest

The authors certify that they have no competing interest in this paper.

Acknowledgments

The work of J. Shen is supported in part by NSFC 12371409, 12531015. The work of J. Zhang is supported by NSFC 12171376 and the Fundamental Research Funds for the Central Universities 2042021kf0050 and WHU-2022-SYJS-0002.

Appendix A. The proof of Lemmas 8 and 9

Proof. (The proof of Lemma 8) For notational simplicity, we denote $t_{n,\xi} := t_{n-1} + \xi\tau_n$. By using the Taylor's expansion, one has

$$u(t_n) = u(t_{n,\xi}) - (\xi-1)\tau_n u_t(t_{n,\xi}) + \frac{1}{2}(\xi-1)^2\tau_n^2 u_{tt}(t_{n,\xi}) - \frac{1}{2}\int_{t_n}^{t_{n,\xi}} (r-t_n)^2 u_{ttt}(r)\mathrm{d}r, \tag{A.1}$$

$$u(t_{n-1}) = u(t_{n,\xi}) - \xi\tau_n u_t(t_{n,\xi}) + \frac{1}{2}\xi^2\tau_n^2 u_{tt}(t_{n,\xi}) - \frac{1}{2}\int_{t_{n-1}}^{t_{n,\xi}} (r-t_{n-1})^2 u_{ttt}(r)\mathrm{d}r, \tag{A.2}$$

$$u(t_{n-2}) = u(t_{n,\xi}) - (\xi+1/r_n)\tau_n u_t(t_{n,\xi}) + \frac{1}{2}(\xi+1/r_n)^2\tau_n^2 u_{tt}(t_{n,\xi}) - \frac{1}{2}\int_{t_{n-2}}^{t_{n,\xi}} (r-t_{n-2})^2 u_{ttt}(r)\mathrm{d}r. \tag{A.3}$$

Combining (2.1) with (A.1)-(A.3), the truncation error can be expressed by

$$\eta^n := -\frac{1+2r_n\xi}{2(1+r_n)\tau_n}\int_{t_n}^{t_{n,\xi}} (r-t_n)^2 u_{ttt}(r)\mathrm{d}r + \frac{1-r_n+2r_n\xi}{2\tau_n}\int_{t_{n-1}}^{t_{n,\xi}} (r-t_{n-1})^2 u_{ttt}(r)\mathrm{d}r - \frac{(2\xi-1)r_n^2}{2(1+r_n)\tau_n}\int_{t_{n-2}}^{t_{n,\xi}} (r-t_{n-2})^2 u_{ttt}(r)\mathrm{d}r. \tag{A.4}$$

Noting that

$$\begin{aligned}
&\Big\| \int_{t_{n-1}}^{t_{n,\xi}} (r-t_{n-1})^2 u_{ttt}(r)\mathrm{d}r - \int_{t_n}^{t_{n,\xi}} (r-t_n)^2 u_{ttt}(r)\mathrm{d}r \Big\| \\
=&\Big\| \int_{t_{n-1}}^{t_n} (r-t_{n-1})^2 u_{ttt}(r)\mathrm{d}r + \int_{t_n}^{t_{n,\xi}} [2(r-t_n)\tau_n + \tau_n^2] u_{ttt}(r)\mathrm{d}r \Big\| \\
\leq& \left(\frac{1}{3} + \xi(\xi-1)\right)\tau_n^3 \|u_{ttt}\|_{L^\infty(0,T;L^2(\Omega))},
\end{aligned} \tag{A.5}$$

and

$$\Big\| \int_{t_{n-2}}^{t_{n,\xi}} (r-t_{n-2})^2 u_{ttt}(r)\mathrm{d}r - \int_{t_n}^{t_{n,\xi}} (r-t_n)^2 u_{ttt}(r)\mathrm{d}r \Big\| \tag{A.6}$$

$$=\| \int_{t_{n-2}}^{t_n} (r-t_{n-2})^2 u_{ttt}(r)\mathrm{d}r + \int_{t_n}^{t_{n,\xi}} [2(r-t_n)(\tau_n+\tau_{n-1}) + (\tau_n+\tau_{n-1})^2] u_{ttt}(r)\mathrm{d}r\|$$

$$\geq -\Big(\frac{(1+\frac{1}{r_n})^3}{3} + (1+\frac{1}{r_n})(\xi-1)^2 + (1+\frac{1}{r_n})^2(\xi-1) \Big) \tau_n^3 \|u_{ttt}\|_{L^\infty(0,T;L^2(\Omega))}. \tag{A.7}$$

Then inserting the bounds (A.5) and (A.7) into (A.4), η^n can be bounded by

$$\|\eta^n\| \leq \frac{1-r_n+2r_n\xi}{2} \Big(\frac{1}{3} + \xi(\xi-1) \Big) \tau_n^2 \|u_{ttt}\|_{L^\infty(0,T;L^2(\Omega))}$$

$$+(\xi-\frac{1}{2}) \Big(\frac{(1+r_n)^2\tau_{n-1}}{3} + (\xi-1)(1+r_n\xi)\tau_n \Big) \tau_n \|u_{ttt}\|_{L^\infty(0,T;L^2(\Omega))}. \tag{A.8}$$

Hence, the proof is completed by applying $r_n \leq \bar{r}_\xi$ to (A.8). □

Proof. (The proof of Lemma 9) The Taylor's expansion produces

$$v(t_n) = v(t_{n,\xi}) - v_t(t_{n,\xi})(\xi-1)\tau_n + \int_{t_n}^{t_{n,\xi}} v_{tt}(r)(r-t_n)\mathrm{d}r,$$

$$v(t_{n-1}) = v(t_{n,\xi}) - v_t(t_{n,\xi})\xi\tau_n + \int_{t_{n-1}}^{t_{n,\xi}} v_{tt}(r)(r-t_{n-1})\mathrm{d}r.$$

Hence, we have

$$\begin{aligned} \|R(v(t_n))\| &= \|\xi \int_{t_n}^{t_{n,\xi}} v_{tt}(r)(r-t_n)\mathrm{d}r - (\xi-1)\int_{t_{n-1}}^{t_{n,\xi}} v_{tt}(r)(r-t_{n-1})\mathrm{d}r\| \\ &= \| \int_{t_n}^{t_{n,\xi}} v_{tt}(r)(r-t_{n-1}-\xi\tau_n)\mathrm{d}r - (\xi-1)\int_{t_{n-1}}^{t_n} v_{tt}(r)(r-t_{n-1})\mathrm{d}r\| \\ &\leq \Big((\xi-1)^2+1 \Big) \frac{\tau_n^2}{2} \|v\|_{L^\infty(0,T;L^2(\Omega))}, \end{aligned} \tag{A.9}$$

which derives (5.17). (5.18) can be immediately obtained by (5.17) and the proof is completed. □

References

[1] Y. He, Y. Liu, T. Tang, On large time-stepping methods for the Cahn–Hilliard equation, Appl. Numer. Math. 57 (5–7) (2007) 616–628.
[2] F. Lou, T. Tang, H. Xie, Parameter-free time adaptivity based on energy evolution for the Cahn-Hilliard equation, Commun. Comput. Phys 19 (2016) 1542.
[3] Z. Zhang, Z. Qiao, An adaptive time-stepping strategy for the Cahn-Hilliard equation, Commun. Comput. Phys. 11 (4) (2012) 1261–1278.
[4] V. Thomée, Galerkin finite element methods for parabolic problems, second edition, 1, Springer-Verlag, 2006.
[5] J. Becker, A second order backward difference method with variable steps for a parabolic problem, BIT Numer. Math. 38 (1998) 644–662.
[6] E. Emmrich, Stability and error of the variable two-step BDF for semilinear parabolic problems, J. Appl. Math. Comput. 19 (2005) 33–55.
[7] W. Wang, M. Mao, Z. Wang, Stability and error estimates for the variable stepsize BDF2 method for linear and semilinear parabolic equations, Adv. Comput. Math. 47 (2021) 1–28.
[8] W. Chen, X. Wang, Y. Yan, 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.
[9] H. Liao, Z. Zhang, Analysis of adaptive BDF2 scheme for diffusion equations, Math. Comp. 90 (329) (2021) 1207–1226.
[10] J. Zhang, C. Zhao, Sharp error estimate of BDF2 scheme with variable time steps for linear reaction-diffusion equations, J. Math. (06) (2021) 471–488.
[11] F. Huang, J. Shen, Stability and error analysis of a second-order consistent splitting scheme for the Navier–Stokes equations, SIAM J. Numer. Anal. 61 (5) (2023) 2408–2433.
[12] Z. Qiao, Z. Zhang, T. Tang, An adaptive time-stepping strategy for the molecular beam epitaxy models, SIAM J. Sci. Comput. 33 (3) (2011) 1395-1414.
[13] Y. Di, Y. Wei, J. Zhang, C. Zhao, Sharp error estimate of an implicit BDF2 scheme with variable time steps for the phase field crystal model, J. Sci. Comput. 92 (65) (2022).
[14] D. Hou, Z. Qiao, An implicit–explicit second-order BDF numerical scheme with variable steps for gradient flows, J. Sci. Comput. 94 (2) (2023) 39.
[15] H. Ceniceros, R. Nos, A. Roma, Three-dimensional, fully adaptive simulations of phase-field fluid models, J. Comput. Phys. 229 (17) (2010) 6135–6155.
[16] C. Elliott, S. Larsson, Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation, Math. Comp. 58 (198) (1992) 603–630.
[17] C. Zhao, R. Yang, Y. Di, J. Zhang, Sharp error estimate of variable time-step IMEX BDF2 scheme for parabolic integro-differential equations with initial singularity arising in finance, J. Comput. Math. 43 (5) (2025) 1118-1140.
[18] A. Matache, C. Schwab, T. Wihler, Fast numerical solution of parabolic integrodifferential equations with applications in finance, SIAM J. Sci. Comput. 27 (2) (2005) 369–393.
[19] P. Tankov, Financial modelling with jump processes, Chapman and Hall/CRC, 2003.
[20] J. Guermond, J. Shen, A new class of truly consistent splitting schemes for incompressible flows, J. Comput. Phys. 192 (1) (2003) 262–276.
[21] H. Johnston, J. Liu, Accurate, stable and efficient Navier–Stokes solvers based on explicit treatment of the pressure term, J. Comput. Phys. 199 (1) (2004) 221–259.
[22] Y. Wei, J. Zhang, C. Zhao, Y. Zhao, An unconditionally energy dissipative, adaptive IMEX BDF2 scheme and its error estimates for Cahn–Hilliard equation on generalized SAV approach, IMA J. Numer. Anal. 45 (4) (2025) 2171–2203.
[23] Y. Di, Y. Ma, J. Shen, J. Zhang, A variable time-step IMEX-BDF2 SAV scheme and its sharp error estimate for the Navier–Stokes equations, ESIAM Math. Model. Numer. Anal. 57 (3) (2023) 1143–1170.
[24] F. Huang, J. Shen, Z. Yang, A highly efficient and accurate new scalar auxiliary variable approach for gradient flows, SIAM J. Sci. Comput. 42 (4) (2020) A2514-A2536.
[25] F. Huang, J. Shen, A new class of implicit-explicit BDFk SAV schemes for general dissipative systems and their error analysis, Comput. Methods Appl. Mech. Engrg. 392 (2022) 114718.
[26] C. Zhao, N. Liu, Y. Ma, J. Zhang, Unconditionally optimal error estimate of a linearized variable-time-step BDF2 scheme for nonlinear parabolic equations, Commun. Math. Sci. 21 (3) (2023) 775–794.

[27] Z. Guan, J. Lowengrub, C. Wang, S. Wise, Second order convex splitting schemes for periodic nonlocal Cahn–Hilliard and Allen–Cahn equations, J. Comput. Phys. 277 (2014) 48–71.
[28] D. Eyre, Unconditionally gradient stable time marching the Cahn-Hilliard equation, MRS Online Proceed. Library (OPL) 529 (1998) 39.
[29] A. Christlieb, J. Jones, K. Promislow, B. Wetton, M. Willoughby, High accuracy solutions to energy gradient flows from material science models, J. Comput. Phys. 257 (2014) 193–215.
[30] M. Chen, F. Yu, Q. Zhang, Weighted and shifted BDF2 methods on variable grids, arXiv preprint arXiv:2108.02910. (2021).
[31] D. Li, Z. Qiao, T. Tang, Characterizing the stabilization size for semi-implicit Fourier-spectral method to phase field equations, SIAM J. Numer. Anal. 54 (3) (2016) 1653–1681.