Journal of Computational Physics 464 (2022) 111311

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

A generalized SAV approach with relaxation for dissipative systems ☆

Yanrong Zhang [a], Jie Shen [b,*]

[a] School of Mathematical Sciences and Fujian Provincial Key Laboratory on Mathematical Modeling and High Performance Scientific Computing, Xiamen University, Xiamen, Fujian, 361005, China

[b] Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 9 February 2022
Received in revised form 15 April 2022
Accepted 15 May 2022
Available online 20 May 2022

Keywords:
Dissipative system
Energy stability
Error estimate
Scalar auxiliary variable (SAV)
Gradient flows

ABSTRACT

The scalar auxiliary variable (SAV) approach [31] and its generalized version GSAV proposed in [20] are very popular methods to construct efficient and accurate energy stable schemes for nonlinear dissipative systems. However, the discrete value of the SAV is not directly linked to the free energy of the dissipative system, and may lead to inaccurate solutions if the time step is not sufficiently small. Inspired by the relaxed SAV method proposed in [21] for gradient flows, we propose in this paper a generalized SAV approach with relaxation (R-GSAV) for general dissipative systems. The R-GSAV approach preserves all the advantages of the GSAV approach, in addition, it dissipates a modified energy that is directly linked to the original free energy. We prove that the k-th order implicit-explicit (IMEX) schemes based on R-GSAV are unconditionally energy stable, and we carry out a rigorous error analysis for $k = 1, 2, 3, 4, 5$. We present ample numerical results to demonstrate the improved accuracy and effectiveness of the R-GSAV approach.

1. Introduction

When designing numerical schemes for nonlinear dissipative systems, it is crucial to preserve the energy dissipation law at the discrete level in order to eliminate non-physics numerical solutions. How to design efficient and accurate energy stable schemes for nonlinear dissipative systems has been a subject of extensive research in the past few decades. Existing popular approaches include, but not limited to: (i) Convex splitting approach [12,13,29,4]: it leads to unconditionally energy stable for a large class of gradient flows but it requires solving a nonlinear system at each time step; (ii) Stabilized linearly implicit approach [42,33]: It leads to unconditionally energy stable schemes for gradient flows with global Lipschitz conditions and only requires solving linear systems with constant coefficients at each time step; (iii) Exponential time differencing (ETD) approach [36,10,11]: It can lead to unconditionally energy stable schemes for certain mildly nonlinear systems and requires the diagonalization of the discrete Laplace operator; (iv) Invariant energy quadratization (IEQ) approach [37,39]: it leads to unconditionally energy stable schemes for a large class of gradient flows, but requires solving a coupled linear system with variable coefficients at each time step; (v) Scalar auxiliary variable (SAV) approach [31,32]: it leads to unconditionally energy stable schemes for a large class of gradient flows, but only requires solving two decoupled linear systems with constant coefficients at each time step. We refer to [32,9,34,11] and references therein for a more complete literature

☆ This work is supported in part by NSFC 11971407, NSF DMS-2012585 and AFOSR FA9550-20-1-0309.

* Corresponding author.
E-mail addresses: yanrongzhang@stu.xmu.edu.cn (Y. Zhang), shen7@purdue.edu (J. Shen).

https://doi.org/10.1016/j.jcp.2022.111311
0021-9991/

on this subject. Note that unconditional energy stable schemes constructed from the above approaches are mostly limited to first- or second-order (see, however, higher-order versions based on Runge-Kutta [1] or Gaussian collocation [16] which require solving coupled linear or nonlinear systems). In [20] (see also [19]), a generalized SAV approach is proposed for general dissipative systems. Its higher-order versions are also unconditionally energy stable and only requires solving one decoupled linear system with constant coefficients at each time step.

Thanks to their simplicity, efficiency, accuracy and generality, the SAV and GSAV approaches have received much attention recently, they, along with their various variations/extensions, have been used to construct unconditionally energy stable schemes for a large class of nonlinear systems, including various gradient flows (see, for instance, [43,5,6,28,24,34,41,40]), gradient flows with other global constraints (see, for instance, [8]), Navier-Stokes equations and related systems (see, for instance, [26,23,15,25]), time fractional PDEs [18,17], conservative or Hamiltonian systems (see, for instance, [3,2,14]), and many more. However, in the original SAV approach and its various variants, the discrete value of the SAV is not directly linked to its definition at the continuous level, and this may lead to a loss of accuracy when the time step is not sufficiently small. In order to see this more clearly, we briefly describe the original SAV approach below.

To fix the idea, we consider a free energy in the form

$$E_{tot}(\phi) = \frac{1}{2}(\mathcal{L}\phi, \phi) + \int_\Omega F(\phi)\mathrm{d}\boldsymbol{x}, \tag{1.1}$$

where $\mathcal{L}$ is a linear self-adjoint elliptic operator, $F(\phi)$ is a nonlinear potential function. Then, the gradient flow associated with the above free energy can be written as

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = -\mathcal{G}\mu, \\ \mu = \dfrac{\delta E_{tot}(\phi)}{\delta \phi} = \mathcal{L}\phi + F'(\phi), \end{cases} \tag{1.2}$$

with periodic or homogeneous Neumann boundary condition, and $\mathcal{G}$ is a positive definite operator. Let $E_1(\phi) = \int_\Omega F(\phi)\mathrm{d}\boldsymbol{x}$ and assume $E_1(\phi) + C_0 > 0$, where $C_0 > 0$ is a constant. The key idea of the original SAV approach [31,32] is to introduce a SAV $r(t) = \sqrt{E_1(\phi) + C_0} > 0$, and expand (1.2) into the following equivalent system

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = -\mathcal{G}\mu, \\ \mu = \mathcal{L}\phi + \dfrac{r(t)}{\sqrt{E_1(\phi) + C_0}} F'(\phi), \\ r_t = \dfrac{1}{2\sqrt{E_1(\phi) + C_0}} \displaystyle\int_\Omega F'(\phi)\phi_t \mathrm{d}\boldsymbol{x}. \end{cases} \tag{1.3}$$

Then, instead of discretizing (1.2), we can discretize the expanded system (1.3) which, with the additional SAV $r(t)$, allows us to construct efficient and unconditional energy stable schemes. For example, a first-order semi-discrete scheme for (1.3) can be constructed as follows

$$\frac{\phi^{n+1} - \phi^n}{\delta t} = -\mathcal{G}\mu^{n+1}, \tag{1.4}$$

$$\mu^{n+1} = \mathcal{L}\phi^{n+1} + \frac{r^{n+1}}{\sqrt{E_1(\phi^n) + C_0}} F'(\phi^n), \tag{1.5}$$

$$\frac{r^{n+1} - r^n}{\delta t} = \frac{1}{2\sqrt{E_1(\phi^n) + C_0}} \int_\Omega F'(\phi^n) \frac{\phi^{n+1} - \phi^n}{\delta t} \mathrm{d}\boldsymbol{x}. \tag{1.6}$$

It can be easily shown that the above scheme is unconditionally energy stable with a modified energy $\tilde{E}(\phi^n) = \frac{1}{2}(\mathcal{L}\phi^n, \phi^n) + |r^n|^2$, and only requires solving two linear systems with constant coefficients [31,32].

While the expanded system and the original system are equivalent at the continuous level, r^{n+1} is not directly linked to $\sqrt{\int_\Omega F(\phi^{n+1})\mathrm{d}\boldsymbol{x} + C_0}$ and may take very different values, if the time step is not sufficiently small, and lead to inaccurate solutions. In particular, for fixed δt, the ratio $\frac{r^{n+1}}{\sqrt{\int_\Omega F(\phi^{n+1})\mathrm{d}\boldsymbol{x} + C_0}}$ may converge to a value different from 1, leading to a wrong steady state solution (see Table 1 in [43]). One possible remedy for this is to monitor the ratio $\frac{r^{n+1}}{\sqrt{\int_\Omega F(\phi^{n+1})\mathrm{d}\boldsymbol{x} + C_0}}$ adaptively to ensure that it is always close to 1 at every time step. Another remedy is to use a Lagrange multiplier SAV approach [6] such that it dissipates the original energy, but it involves solving a nonlinear algebraic equation at each time step which may not admit a suitable solution when the time step is not sufficiently small.

Recently, an interesting relaxed SAV (R-SAV) approach was introduced in [21]. The idea is to add a relaxation step to the original SAV approach to link r^{n+1} with $\sqrt{\int_\Omega F(\phi^{n+1})\mathrm{d}\boldsymbol{x}+C_0}$ so that the updated sequence r^{n+1} is directly linked to $\sqrt{\int_\Omega F(\phi^{n+1})\mathrm{d}\boldsymbol{x}+C_0}$ in some way and is still dissipative. The cost of this relaxation step is negligible while numerical results in [21] show that the R-SAV approach can significantly improve the accuracy of the original SAV approach. However, this R-SAV approach is based on the original SAV approach which has two limitations/shortcomings: (i) it only applies to gradient flows; and (ii) it requires solving two linear systems at each time step. The generalized SAV (GSAV) approach proposed in [20,19] overcomes the above limitations/shortcomings while keeping the essential advantages of the original SAV approach, but it also suffers from the same problem that the computed SAV is not directly linked to the free energy and may lead to non accurate solutions as in the original SAV approach.

The main purpose of this paper is to construct a relaxed GSAV (R-GSAV) approach which links the SAV directly to the free energy, and enjoys the following advantages:

- It is unconditionally energy stable with respect to a modified energy that is closer to the original free energy, and provides a much improved accuracy when compared with the GSAV approach;
- it only requires solving one linear system with constant coefficients as opposed to the two linear systems by the R-SAV approach, so its computational cost is essentially half of the R-SAV approach;
- it can be applied to general dissipative systems, and its higher-order versions are shown to be unconditionally energy stable with optimal error estimates.

Moreover, our numerical results indicate that, for the ample numerical experiments that we tested, the modified energy of our R-GSAV schemes equals to the original free energy at almost all times.

The rest of this paper is organized as follows. In section 2, we provide a brief review of the original SAV approach and R-SAV approach for gradient flows, and the GSAV approach for general dissipative systems. In Section 3, we present the R-GSAV approach for general dissipative systems, and prove that the kth-order implicit-explicit (IMEX) schemes based on the R-GSAV approach is unconditionally stable. In Section 4, we carry out an error analysis for the kth order ($1 \le k \le 5$) IMEX schemes for Allen-Cahn type and Cahn-Hilliard type equations. In Section 5, we extend the R-GSAV approach to cases where multiple SAVs are used. We present in Section 6 comparisons of R-GSAV approach with original SAV, R-SAV and GSAV approaches, and provide ample numerical examples to validate its effectiveness.

2. A brief review of the original SAV, relaxed SAV and generalized SAV approaches

In order to motivate our new schemes, we briefly review below the original SAV, relaxed SAV and generalized SAV approaches.

2.1. The original SAV approach

A brief description of the original SAV approach for gradient flows is already provided in the introduction where a first-order scheme is introduced. Similarly, we can construct k-th order IMEX schemes for the expanded system (1.3) as follows:

$$\frac{\alpha_k\phi^{n+1}-A_k\left(\phi^n\right)}{\delta t}=-\mathcal{G}\mu^{n+1}, \tag{2.1}$$

$$\mu^{n+1}=\mathcal{L}\phi^{n+1}+\frac{r^{n+1}}{\sqrt{E_1(B_k(\phi^n))+C_0}}F'(B_k(\phi^n)), \tag{2.2}$$

$$\frac{\alpha_k r^{n+1}-A_k\left(r^n\right)}{\delta t}=\frac{1}{2\sqrt{E_1(B_k(\phi^n))+C_0}}\int\limits_\Omega F'(B_k(\phi^n))\frac{\alpha_k\phi^{n+1}-A_k\left(\phi^n\right)}{\delta t}\mathrm{d}\boldsymbol{x}, \tag{2.3}$$

where α_k, A_k and B_k can be derived by Taylor expansion. For the readers' convenience, we provide them for $k=1,2,3$ below:

First-order:

$$\alpha_1=1,\quad A_1\left(\phi^n\right)=\phi^n,\quad B_1\left(\phi^n\right)=\phi^n; \tag{2.4}$$

Second-order:

$$\alpha_2=\frac{3}{2},\quad A_2\left(\phi^n\right)=2\phi^n-\frac{1}{2}\phi^{n-1},\quad B_2\left(\phi^n\right)=2\phi^n-\phi^{n-1}; \tag{2.5}$$

Third-order:

$$\alpha_3=\frac{11}{6},\quad A_3\left(\phi^n\right)=3\phi^n-\frac{3}{2}\phi^{n-1}+\frac{1}{3}\phi^{n-2},\quad B_3\left(\phi^n\right)=3\phi^n-3\phi^{n-1}+\phi^{n-2}. \tag{2.6}$$

It has been shown that the scheme (2.1)-(2.3) for $k=1,2$ is unconditionally energy stable with a modified energy. For example, the modified energy is $\tilde{E}(\phi^{n+1})=\frac{1}{2}(\mathcal{L}\phi^{n+1},\phi^{n+1})+|r^{n+1}|^2$ for $k=1$, where r^{n+1} is only weakly linked to $\sqrt{E_1(B_k(\phi^{n+1}))+C_0}$. The consequence is that when the time step is not sufficiently small, the modified energy can deviate far away from the original energy, leading to inaccurate solutions.

2.2. The relaxed SAV approach

The relaxed SAV (R-SAV) approach proposed in [21] is as follows:
Step 1: Calculate the solution $\left(\phi^{n+1},\tilde{r}^{n+1}\right)$ based on original SAV approach:

$$\frac{\alpha_k\phi^{n+1}-A_k\left(\phi^n\right)}{\delta t}=-\mathcal{G}\mu^{n+1}, \tag{2.7}$$

$$\mu^{n+1}=\mathcal{L}\phi^{n+1}+\frac{\tilde{r}^{n+1}}{\sqrt{E_1(B_k(\phi^n))+C_0}}F'(B_k(\phi^n)), \tag{2.8}$$

$$\frac{\alpha_k\tilde{r}^{n+1}-A_k\left(r^n\right)}{\delta t}=\frac{1}{2\sqrt{E_1(B_k(\phi^n))+C_0}}\int_\Omega F'(B_k(\phi^n))\frac{\alpha_k\phi^{n+1}-A_k\left(\phi^n\right)}{\delta t}\mathrm{d}\boldsymbol{x}. \tag{2.9}$$

Step 2: Update the SAV r^{n+1} by a relaxation:

$$r^{n+1}=\zeta_0^{n+1}\tilde{r}^{n+1}+\left(1-\zeta_0^{n+1}\right)E_1\left(\phi^{n+1}\right),\quad \zeta_0^{n+1}\in\mathcal{V}. \tag{2.10}$$

Here, $\mathcal{V}$ is a set defined by

$$\mathcal{V}=\left\{\zeta\in[0,1]\,s.t.\ \left|r^{n+1}\right|^2-\left|\tilde{r}^{n+1}\right|^2\le\delta t\gamma\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)\right\} \tag{2.11}$$

for BDF1 scheme, and

$$\mathcal{V}=\left\{\zeta\in[0,1]\,s.t.\ \frac{1}{2}\left(\left|r^{n+1}\right|^2+\left|2r^{n+1}-r^n\right|^2-\left|\tilde{r}^{n+1}\right|^2-\left|2\tilde{r}^{n+1}-\tilde{r}^n\right|^2\right)\le\delta t\gamma\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)\right\} \tag{2.12}$$

for BDF2 scheme, where $\gamma\in[0,1]$ is a tunable parameter.

Remark 2.1 *(Optimal choice for ζ_0^{n+1}).* We describe below how to choose the relaxation parameter ζ_0^{n+1}. Taking BDF2 as an example, ζ_0^{n+1} can be chosen as a solution of the following optimization problem:

$$\zeta_0^{n+1}=\min_{\zeta\in[0,1]}\zeta,\quad \text{s.t. }\frac{1}{2}\left(\left|r^{n+1}\right|^2+\left|2r^{n+1}-r^n\right|^2-\left|\tilde{r}^{n+1}\right|^2-\left|2\tilde{r}^{n+1}-\tilde{r}^n\right|^2\right)\le\delta t\gamma\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right), \tag{2.13}$$

with $r^{n+1}=\zeta_0^{n+1}\tilde{r}^{n+1}+\left(1-\zeta_0^{n+1}\right)E_1\left(\phi^{n+1}\right)$. This can be simplified as

$$\zeta_0^{n+1}=\min_{\zeta\in[0,1]}\zeta,\quad \text{s.t. } f(\zeta)=a\zeta^2+b\zeta+c\le 0, \tag{2.14}$$

where the coefficients are

$$\begin{aligned}
a&=\tfrac{5}{2}\left(\tilde{r}^{n+1}-E_1\left(\phi^{n+1}\right)\right)^2,\\
b&=\left(\tilde{r}^{n+1}-E_1\left(\phi^{n+1}\right)\right)\left(5E_1\left(\phi^{n+1}\right)-2r^n\right),\\
c&=\tfrac{1}{2}\left(\left(E_1\left(\phi^{n+1}\right)\right)^2+\left(2E_1\left(\phi^{n+1}\right)-r^n\right)^2-\left(\tilde{r}^{n+1}\right)^2-\left(2\tilde{r}^{n+1}-r^n\right)^2\right)-\delta t\gamma\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right).
\end{aligned} \tag{2.15}$$

If $a=0$, we set $\zeta_0^{n+1}=0$. If $a>0$, notice that $f(1)=a+b+c=-\delta t\gamma\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)\le 0$, then we have $1\in\mathcal{V}$ i.e., $\mathcal{V}\neq\emptyset$, and the quadratic function $f(\zeta)$ has at least one real root. Then the solution of (2.13) is given by

$$\zeta_0^{n+1}=\max\left\{0,\frac{-b-\sqrt{b^2-4ac}}{2a}\right\}. \tag{2.16}$$

It can be shown that the scheme (2.7)-(2.10) with $k=1,2$ is unconditionally energy stable in the sense that
(i) For $k=1$, $R^{n+1}_{R-SAV-BDF1}-R^n_{R-SAV-BDF1}\le-\delta t(1-\gamma)\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)$, where

$$R^{n+1}_{R-SAV-BDF1}=\frac{1}{2}\left(\mathcal{L}\phi^{n+1},\phi^{n+1}\right)+\left|r^{n+1}\right|^2;$$

(ii) For $k=2$, $R^{n+1}_{R-SAV-BDF2} - R^{n}_{R-SAV-BDF2} \le -\delta t\,(1-\gamma)\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)$, where

$$R^{n+1}_{R-SAV-BDF2} = \frac{1}{4}\left(\left(\mathcal{L}\phi^{n+1},\phi^{n+1}\right) + \left(\mathcal{L}\left(2\phi^{n+1}-\phi^{n}\right), 2\phi^{n+1}-\phi^{n}\right)\right) + \frac{1}{2}\left(\left|r^{n+1}\right|^2 + \left|2r^{n+1}-r^{n}\right|^2\right).$$

Note that by the definition in (2.10), r^{n+1} is directly linked to $E_1(\phi^{n+1})$. Hence, the modified energy above is directly linked to the original free energy.

2.3. The generalized SAV (GSAV) approach

The GSAV approach was proposed in [19] for the following general dissipative systems:

$$\frac{\partial \phi}{\partial t} + \mathcal{A}\phi + g(\phi) = 0, \tag{2.17}$$

where $\mathcal{A}$ is a positive definite operator and $g(\phi)$ is a semi-linear or quasi-linear operator. We assume it satisfies an energy dissipation law as follows

$$\frac{\mathrm{d}E_{tot}(\phi)}{\mathrm{d}t} = -\mathcal{K}(\phi), \tag{2.18}$$

where $E_{tot}(\phi)$ is a free energy with lower bound $-C_0$ and $\mathcal{K}(\phi) \ge 0$ for all ϕ. Setting $E(\phi) = E_{tot}(\phi) + C_0$ and introducing a SAV $R(t) = E(\phi)$, we rewrite the equation (2.17) as the following system

$$\begin{cases} \dfrac{\partial \phi}{\partial t} + \mathcal{A}\phi + g(\phi) = 0, \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = -\dfrac{R(t)}{E(\phi)}\mathcal{K}(\phi). \end{cases} \tag{2.19}$$

Then, the BDFk GSAV schemes are as follows:

$$\frac{\alpha_k \bar{\phi}^{n+1} - A_k\left(\phi^n\right)}{\delta t} + \mathcal{A}\bar{\phi}^{n+1} + g\left[B_k\left(\bar{\phi}^n\right)\right] = 0, \tag{2.20}$$

$$\frac{1}{\delta t}\left(R^{n+1} - R^n\right) = -\frac{R^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\mathcal{K}(\bar{\phi}^{n+1}), \tag{2.21}$$

$$\xi_k^{n+1} = \frac{R^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}, \tag{2.22}$$

$$\phi^{n+1} = \eta_k^{n+1}\bar{\phi}^{n+1} \text{ with } \eta_k^{n+1} = 1 - \left(1 - \xi_k^{n+1}\right)^{k+1}, \tag{2.23}$$

where α_k, the operators A_k and B_k are as above.

It is shown in [19] that the above scheme is unconditional energy stable with a modified energy. However, as the original SAV approach, the dynamics of SAV R^{n+1} is only weakly linked to the energy $E(\phi^{n+1})$ and may deviate from it when the time step is not sufficiently small.

3. The relaxed generalized (R-GSAV) SAV approach

Inspired by the R-SAV approach described above, we construct below the relaxed GSAV approach, which not only inherits all the advantages of the GSAV approach, but can also significantly improve its accuracy.

Given $\phi^{n-k}, \ldots, \phi^n, R^{n-k}, \ldots, R^n$, we compute ϕ^{n+1}, R^{n+1} via the following two steps:

Step 1: Determine an intermediate solution $(\phi^{n+1}, \tilde{R}^{n+1})$ by using the GSAV method:

$$\frac{\alpha_k \bar{\phi}^{n+1} - A_k\left(\phi^n\right)}{\delta t} + \mathcal{A}\bar{\phi}^{n+1} + g\left[B_k\left(\bar{\phi}^n\right)\right] = 0, \tag{3.1}$$

$$\frac{1}{\delta t}\left(\tilde{R}^{n+1} - R^n\right) = -\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\mathcal{K}(\bar{\phi}^{n+1}), \tag{3.2}$$

$$\xi_k^{n+1} = \frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}, \tag{3.3}$$

$$\phi^{n+1} = \eta_k^{n+1}\bar{\phi}^{n+1} \text{ with } \eta_k^{n+1} = 1 - \left(1 - \xi_k^{n+1}\right)^{k+1}. \tag{3.4}$$

Step 2: Update the SAV R^{n+1} via the following relaxation:

$$R^{n+1}=\zeta_0^{n+1}\tilde{R}^{n+1}+(1-\zeta_0^{n+1})E(\phi^{n+1}),\quad \zeta_0^{n+1}\in\mathcal{V}, \tag{3.5}$$

where

$$\mathcal{V}=\left\{\zeta\in[0,1]\ s.t.\ \frac{R^{n+1}-\tilde{R}^{n+1}}{\delta t}=-\gamma^{n+1}\mathcal{K}(\phi^{n+1})+\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\mathcal{K}(\bar{\phi}^{n+1})\right\}, \tag{3.6}$$

with $\gamma^{n+1}\geq 0$ to be determined so that $\mathcal{V}$ is not empty.

We explain below how to choose ζ_0^{n+1} and γ^{n+1}. Plugging (3.5) into the equality of (3.6), we find that if we choose ζ_0^{n+1} and γ^{n+1} such that the following condition is satisfied:

$$(\tilde{R}^{n+1}-E(\phi^{n+1}))\zeta_0^{n+1}=\tilde{R}^{n+1}-E(\phi^{n+1})-\delta t\gamma^{n+1}\mathcal{K}(\phi^{n+1})+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\mathcal{K}(\bar{\phi}^{n+1}), \tag{3.7}$$

then, $\zeta_0^{n+1}\in\mathcal{V}$. The next theorem summarizes the choice of ζ_0^{n+1} and γ^{n+1}.

Theorem 3.1. *We choose ζ_0^{n+1} in (3.5) and γ^{n+1} in (3.6) as follows:*

1. *If $\tilde{R}^{n+1}=E(\phi^{n+1})$, we set $\zeta_0^{n+1}=0$ and $\gamma^{n+1}=\frac{\tilde{R}^{n+1}\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})\mathcal{K}(\phi^{n+1})}$.*
2. *If $\tilde{R}^{n+1}>E(\phi^{n+1})$, we set $\zeta_0^{n+1}=0$ and*

$$\gamma^{n+1}=\frac{\tilde{R}^{n+1}-E(\phi^{n+1})}{\delta t\mathcal{K}(\phi^{n+1})}+\frac{\tilde{R}^{n+1}\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})\mathcal{K}(\phi^{n+1})}. \tag{3.8}$$

3. *If $\tilde{R}^{n+1}<E(\phi^{n+1})$ and $\tilde{R}^{n+1}-E(\phi^{n+1})+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\mathcal{K}(\bar{\phi}^{n+1})\geq 0$, we set $\zeta_0^{n+1}=0$ and γ^{n+1} the same as (3.8).*
4. *If $\tilde{R}^{n+1}<E(\phi^{n+1})$ and $\tilde{R}^{n+1}-E(\phi^{n+1})+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\mathcal{K}(\bar{\phi}^{n+1})<0$, we set $\zeta_0^{n+1}=1-\frac{\delta t\tilde{R}^{n+1}\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})\left(E(\phi^{n+1})-\tilde{R}^{n+1}\right)}$ and $\gamma^{n+1}=0$.*

Then, (3.7) is satisfied in all cases above and $\zeta_0^{n+1}\in\mathcal{V}$. Furthermore, given $R^n\geq 0$, we have $R^{n+1}\geq 0$, $\xi_k^{n+1}\geq 0$, and the scheme (3.1)-(3.5) with the above choice of ζ_0^{n+1} and γ^{n+1} is unconditionally energy stable in the sense that

$$R^{n+1}-R^n=-\delta t\gamma^{n+1}\mathcal{K}(\phi^{n+1})\leq 0. \tag{3.9}$$

Furthermore, we have

$$R^{n+1}\leq E(\phi^{n+1})\quad \forall n\geq 0. \tag{3.10}$$

In addition, if $E(\phi)=\frac{1}{2}(\mathcal{L}\phi,\phi)+E_1(\phi)$ with $\mathcal{L}$ being a linear positive definite operator and $E_1(\phi)$ bounded from below, there exists $M_k>0$ such that

$$\left(\mathcal{L}\phi^n,\phi^n\right)\leq M_k^2,\quad \forall n. \tag{3.11}$$

Proof. It can be directly verified that, with the above choice of ζ_0^{n+1} and γ^{n+1}, (3.7) is satisfied in all cases so that $\zeta_0^{n+1}\in\mathcal{V}$.

Given $R^n\geq 0$. Since $E(\bar{\phi}^{n+1})>0$, it follows from (3.2) that

$$\tilde{R}^{n+1}=\frac{R^n}{1+\delta t\frac{\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})}}\geq 0. \tag{3.12}$$

Then we derive from (3.3) that $\xi_k^{n+1}\geq 0$, and we derive from (3.5) that $R^{n+1}\geq 0$.

Combining (3.2) and (3.6), we obtain (3.9).

For Cases 1-3, we have $\zeta_0^{n+1}=0$ so $R^{n+1}=E(\phi^{n+1})$. For Case 4, since $\zeta_0^{n+1}=1-\frac{\delta t\tilde{R}^{n+1}\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})\left(E(\phi^{n+1})-\tilde{R}^{n+1}\right)}\in[0,1]$ and $\tilde{R}^{n+1}<E(\phi^{n+1})$, we derive from (3.5) that $R^{n+1}\leq E(\phi^{n+1})$.

The proof of (3.11) is essentially the same as the proof of Theorem 1 of the GSAV scheme in [19]. For the readers' convenience, we provide the proof below.

Denote $M:=R^0=E[\phi(\cdot,0)]$, then we derive from (3.9) and (3.12) that $\tilde{R}^{n+1}\leq M$, $\forall n$.

Without loss of generality, we can assume $E_1(\phi)>1$ for all ϕ. It then follows from (3.3) that

$$|\xi_k^{n+1}|=\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\leq\frac{2M}{(\mathcal{L}\bar{\phi}^{n+1},\bar{\phi}^{n+1})+2}. \tag{3.13}$$

Then $\eta_k^{n+1} = 1 - \left(1-\xi_k^{n+1}\right)^{k+1} = \xi_k^{n+1} P_k(\xi_k^{n+1})$ with P_k being a polynomial of degree k. We derive from (3.13) that there exists $M_k > 0$ such that

$$|\eta_k^{n+1}| = |\xi_k^{n+1} P_k(\xi_k^{n+1})| \le \frac{M_k}{(\mathcal{L}\bar{\phi}^{n+1}, \bar{\phi}^{n+1}) + 2},$$

which, along with $\phi^{n+1} = \eta_k^{n+1}\bar{\phi}^{n+1}$, implies

$$\begin{aligned}(\mathcal{L}\phi^{n+1}, \phi^{n+1}) &= (\eta_k^{n+1})^2 (\mathcal{L}\bar{\phi}^{n+1}, \bar{\phi}^{n+1}) \\ &\le \Big(\frac{M_k}{(\mathcal{L}\bar{\phi}^{n+1}, \bar{\phi}^{n+1}) + 2}\Big)^2 (\mathcal{L}\bar{\phi}^{n+1}, \bar{\phi}^{n+1}) \le M_k^2.\end{aligned}$$

The proof is complete. □

Remark 3.1. From an energy approximation point of view, it is best that $R^{n+1} = E(\phi^{n+1})$. However, simply setting $R^{n+1} = E(\phi^{n+1})$ at each time step destroys the energy stability. We observe from the above statements that in most cases (Cases 1, 2, 3), we have $\zeta_0^{n+1} = 0$ which implies $R^{n+1} = E(\phi^{n+1})$, and then thanks to (3.9) and (3.10), we have

$$E(\phi^{n+1}) = R^{n+1} \le R^n \le E(\phi^n) \quad \text{(Cases 1, 2, 3)}. \tag{3.14}$$

Only in Case 4 where $\tilde{R}^{n+1}$ is significantly smaller than $E(\phi^{n+1})$, R^{n+1} is a value between $(\tilde{R}^{n+1}, E(\phi^{n+1}))$, and we can not prove $E(\phi^{n+1}) \le E(\phi^n)$. Hence, the original energy is proved to be dissipative in most situations, which is a significant improvement over the GSAV scheme.

Remark 3.2. The R-GSAV scheme (3.1)-(3.5) can be easily modified to handle an external force. Indeed, assuming the energy law for (2.17) with an external force is

$$\frac{\mathrm{d}E_{tot}(\phi)}{\mathrm{d}t} = -\mathcal{K}(\phi) + \mathcal{F}(\phi), \tag{3.15}$$

where $\mathcal{F}(\phi)$ is related to the external force f. Then, we only need to replace (3.2) by

$$\frac{1}{\delta t}\left(\tilde{R}^{n+1} - R^n\right) = -\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left(\mathcal{K}(\bar{\phi}^{n+1}) + \mathcal{F}(\bar{\phi}^{n+1})\right). \tag{3.16}$$

4. Error estimate

In this section, we will derive error estimates for the R-GSAV schemes applied to Allen-Cahn type equations and Cahn-Hilliard type equations by using the stability results in Theorem 3.1.

We recall first some preliminary results which play a critical role in error analysis.

Lemma 4.1. *[27] For $1 \le k \le 5$, there exist $0 \le \tau_k < 1$, a positive definite symmetric matrix $G = (g_{ij}) \in \mathcal{R}^{k,k}$ and real numbers $a_0, \ldots, a_k$ such that*

$$\begin{aligned}\left(\alpha_k\phi^{n+1} - A_k\left(\phi^n\right), \phi^{n+1} - \tau_k\phi^n\right) &= \sum_{i,j=1}^{k} g_{ij}\left(\phi^{n+1+i-k}, \phi^{n+1+j-k}\right) \\ &\quad - \sum_{i,j=1}^{k} g_{ij}\left(\phi^{n+i-k}, \phi^{n+j-k}\right) + \left\|\sum_{i=0}^{k} a_i\phi^{n+1+i-k}\right\|^2,\end{aligned}$$

where the smallest possible values of τ_k are

$$\tau_1 = \tau_2 = 0, \quad \tau_3 = 0.0836, \quad \tau_4 = 0.2878, \quad \tau_5 = 0.8160, \tag{4.1}$$

and α_k, A_k are constant and operator related to BDFk IMEX schemes as described in the last section.

The following regularity results for (1.2) are given in [30,35].

Theorem 4.1. *Assume $\phi^0 \in H^2(\Omega)$ and the following holds*

$$\left|F''(x)\right| < C\left(|x|^p + 1\right), \quad p > 0 \text{ arbitrary if } d = 1, 2; \quad 0 < p < 4 \text{ if } d = 3. \tag{4.2}$$

Then for $\mathcal{G} = I$, problem (1.2) has a unique solution for any $T > 0$ in the space

$$C\left([0,T];H^2(\Omega)\right)\cap L^2\left(0,T;H^3(\Omega)\right). \tag{4.3}$$

Furthermore, assume

$$\left|F'''(x)\right| < C\left(|x|^{p'}+1\right), \quad p'>0 \text{ arbitrary if } d=1,2; \quad 0<p'<3 \text{ if } d=3. \tag{4.4}$$

Then for $\mathcal{G} = -\Delta$, there exists a unique solution for any $T > 0$ such that

$$\phi\in C\left([0,T];H^2(\Omega)\right)\cap L^2\left(0,T;H^4(\Omega)\right). \tag{4.5}$$

We also recall the following useful results to deal with the nonlinear term [30].

Theorem 4.2. *Assume that $\|\phi\|_{H^1} \le M$.*

- *Assume that (4.2) holds. Then for any $\phi \in H^3$, there exists $0 \le \sigma < 1$ and a constant $C(M)$ such that the following inequality holds*

$$\|\nabla F'(\phi)\|^2 \le C(M)\left(1+\|\nabla\Delta\phi\|^{2\sigma}\right). \tag{4.6}$$

- *Assume that (4.2) and (4.4) hold. Then, for any $\phi \in H^4$, there exists $0 \le \sigma < 1$ and a constant $C(M)$ such that the following inequality holds*

$$\|\Delta F'(\phi)\|^2 \le C(M)\left(1+\|\Delta^2\phi\|^{2\sigma}\right). \tag{4.7}$$

We consider first the Allen-Cahn type equation

$$\frac{\partial\phi}{\partial t}-\Delta\phi+\lambda\phi+F'(\phi)=0, \quad (\boldsymbol{x},t)\in\Omega\times(0,T], \tag{4.8}$$

which is (1.2) with $\mathcal{L} = -\Delta + \lambda I$ and $\mathcal{G} = I$. It can also be written in the form of (2.17) with $\mathcal{A} = -\Delta + \lambda I$ and $g(\phi) = F'(\phi)$. The corresponding (2.18) is with $E_{tot}(\phi) = \int_\Omega \frac{1}{2}|\nabla\phi|^2 + \lambda|\phi|^2 + F(\phi)\mathrm{d}\boldsymbol{x}$ and $\mathcal{K}(\phi) = \|-\Delta\phi + \lambda\phi + F'(\phi)\|^2$.

In the following, we follow a similar procedure as in [19] to carry out a unified error analysis for the R-GSAV/BDFk ($1 \le k \le 5$) defined by (3.1)-(3.5).

Theorem 4.3. *Given initial conditions $\bar{\phi}^i = \phi^i = \phi(t^i)$, $R^i = E(\phi^i)$, $i = 0, 1, ..., k-1$. Let $\bar{\phi}^{n+1}$ and ϕ^{n+1} be computed with the R-GSAV/BDFk ($1 \le k \le 5$) scheme (3.1)-(3.5) for (4.8), except that when $k = 1$, we set*

$$\eta_1^{n+1}=1-\left(1-\xi_1^{n+1}\right)^3, \tag{4.9}$$

and for Case 4 of $k = 1$, we set

$$\zeta_0^{n+1}=1-\frac{\delta t^2\tilde{R}^{n+1}\mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})(E(\phi^{n+1})-\tilde{R}^{n+1})}. \tag{4.10}$$

We assume (4.2) holds and

$$\phi^0\in H^3, \quad \frac{\partial^j\phi}{\partial t^j}\in L^2\left(0,T;H^1\right), 1\le j\le k+1. \tag{4.11}$$

Then for $\delta t < \min\{\frac{1}{1+2C_0^{k+2}}, \frac{1-\tau_k}{3k}\}$, we have

$$\left\|\bar{\phi}^{n+1}-\phi(\cdot,t^{n+1})\right\|_{H^2}, \quad \left\|\phi^{n+1}-\phi(\cdot,t^{n+1})\right\|_{H^2}\le C\delta t^k, \quad \forall n+1\le T/\delta t,$$

where τ_k is given in (4.1), and the constants C_0, C are independent of δt.

Remark 4.1. The choices (4.9) when $k = 1$ and (4.10) in Case 4 of $k = 1$ are made for purely technical reasons in the proof. In this case, we can choose $\gamma^{n+1} = \frac{\tilde{R}^{n+1}\|h(\bar{\phi}^{n+1})\|^2}{E(\bar{\phi}^{n+1})\|h(\phi^{n+1})\|^2}(1-\delta t)$, which satisfies energy stability (3.9) if $\delta t \le 1$. And it is clear that the original choices of η_1^{n+1} and ζ_0^{n+1} in Theorem 3.1 still lead to first-order accuracy which is confirmed by our numerical experiments.

Proof. We denote $t^n = n\delta t, \bar{e}^n = \bar{\phi}^n - \phi(\cdot, t^n), e^n = \phi^n - \phi(\cdot, t^n), \tilde{s}^n = \tilde{R}^n - R(t^n), s^n = R^n - R(t^n)$.

We will prove by induction

$$\left|1 - \xi_k^q\right| \le C_0 \delta t, \quad \forall q \le T/\delta t, \tag{4.12}$$

where C_0 is dependent on T, Ω and the exact solution ϕ but is independent of δt. Under the assumption, (4.12) holds for $q = 0$. Assuming

$$\left|1 - \xi_k^q\right| \le C_0 \delta t, \quad \forall q \le m, \tag{4.13}$$

we need to prove

$$\left|1 - \xi_k^{m+1}\right| \le C_0 \delta t. \tag{4.14}$$

We first consider the cases $k = 2, 3, 4, 5$, and point out the slightly different proof process for the case $k = 1$ later. The **Step 1** and **Step 2** below are essentially the same as in [19]. So we only list the results that will be used here and refer to [19] for more details.

Step 1: H^2 bound for ϕ^n and $\bar{\phi}^n$ for all $n \le m$. The first step of the scheme (3.1)-(3.2) for (4.8) is

$$\frac{\alpha_k \bar{\phi}^{n+1} - A_k\left(\phi^n\right)}{\delta t} = \Delta \bar{\phi}^{n+1} - \lambda \bar{\phi}^{n+1} + F'\left[B_k\left(\bar{\phi}^n\right)\right]. \tag{4.15}$$

Multiplying the above by η_k^{n+1}, we obtain

$$\frac{\alpha_k \phi^{n+1} - \eta_k^{n+1} A_k\left(\phi^n\right)}{\delta t} = \Delta \phi^{n+1} - \lambda \phi^{n+1} + \eta_k^{n+1} F'\left[B_k\left(\bar{\phi}^n\right)\right]. \tag{4.16}$$

Under the assumption (4.13), it can be shown that for δt small enough such that

$$\delta t \le \min\left\{\frac{1}{2C_0^{k+1}}, 1\right\}, \tag{4.17}$$

we have

$$1 - \frac{\delta t^k}{2} \le \left|\eta_k^q\right| \le 1 + \frac{\delta t^k}{2}, \quad \left|1 - \eta_k^q\right| \le \frac{\delta t^k}{2}, \quad \forall q \le m. \tag{4.18}$$

Taking the inner product of (4.16) with $\Delta^2 \phi^q - \tau_k \Delta^2 \phi^{q-1}$ and using Theorem 3.1, Lemma 4.1 and the property of symmetric positive definite matrix $G = (g_{ij})$, we can obtain

$$\|\phi^n\|_{H^2} \le C_1, \quad \forall \delta t < 1, n \le m. \tag{4.19}$$

We derive from the above and (4.18) that

$$\left\|\bar{\phi}^n\right\|_{H^2} \le 2C_1, \quad \forall \delta t < 1, n \le m. \tag{4.20}$$

Step 2: estimate for $\|\bar{e}^{n+1}\|_{H^2}$ for all $n \le m$. Subtracting (4.15) from (4.8), and using (4.19) and (4.20), we can derive that for $\delta t < \frac{1}{C_2}$, we have

$$\left\|\bar{e}^{n+1}\right\|_{H^2} \le \sqrt{C_2\left(1 + C_0^{2k+2}\right)}\delta t^k, \quad \forall 0 \le n \le m, \tag{4.21}$$

$$\left\|\bar{\phi}^{n+1}\right\|_{H^2} \le \bar{C}, \quad \forall 0 \le n \le m, \tag{4.22}$$

and

$$\left\|F'\left(\bar{\phi}^{n+1}\right)\right\|, \left\|F''\left(\bar{\phi}^{n+1}\right)\right\| \le \bar{C} \quad \forall 0 \le n \le m. \tag{4.23}$$

Step 3: estimate for $1 - \xi_k^{m+1}$. By direct calculation,

$$R_{tt} = \int_\Omega \left(|\nabla \phi_t|^2 + \nabla \phi \cdot \nabla \phi_{tt} + \lambda \phi_t^2 + \lambda \phi \phi_{tt} + F''(\phi)\phi_t^2 + F'(\phi)\phi_{tt}\right) \mathrm{d}\boldsymbol{x}. \tag{4.24}$$

It follows from (3.2) that the equation for the errors can be written as

$$\tilde{s}^{n+1}-s^{n}=\delta t\left(\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right)+T_{1}^{n}, \tag{4.25}$$

where $h(\phi)=\mu=-\Delta\phi+\lambda\phi+F'(\phi)$, and

$$T_{1}^{n}=R\left(t^{n}\right)-R\left(t^{n+1}\right)+\delta t R_{t}\left(t^{n+1}\right)=\int_{t^{n}}^{t^{n+1}}\left(s-t^{n}\right)R_{tt}(s)\mathrm{d}s. \tag{4.26}$$

Since $R^{n}=\zeta_{0}^{n}\tilde{R}^{n}+\left(1-\zeta_{0}^{n}\right)E\left(\phi^{n}\right)$ and $R\left(t^{n}\right)=\zeta_{0}^{n}R\left(t^{n}\right)+\left(1-\zeta_{0}^{n}\right)E\left(\phi\left(t^{n}\right)\right)$, we obtain

$$s^{n}=\zeta_{0}^{n}\tilde{s}^{n}+\left(1-\zeta_{0}^{n}\right)\left[E\left(\phi^{n}\right)-E\left(\phi\left(t^{n}\right)\right)\right]. \tag{4.27}$$

Plugging (4.27) into (4.25), we obtain

$$\tilde{s}^{n+1}-\zeta_{0}^{n}\tilde{s}^{n}-\left(1-\zeta_{0}^{n}\right)\left[E\left(\phi^{n}\right)-E\left(\phi\left(t^{n}\right)\right)\right]=\delta t\left(\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right)+T_{1}^{n}, \tag{4.28}$$

by using the triangle inequality principle, we derive

$$\begin{aligned}&\left|\tilde{s}^{n+1}\right|-\zeta_{0}^{n}\left|\tilde{s}^{n}\right|\leq\left|\tilde{s}^{n+1}-\zeta_{0}^{n}\tilde{s}^{n}\right|\\&\leq\delta t\left|\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right|+\left(1-\zeta_{0}^{n}\right)\left|E\left(\phi^{n}\right)-E\left(\phi\left(t^{n}\right)\right)\right|+\left|T_{1}^{n}\right|.\end{aligned} \tag{4.29}$$

Taking the sum of (4.29) for n from 0 to m, and noting that $\tilde{s}^{0}=0$, we have

$$\begin{aligned}&\left|\tilde{s}^{m+1}\right|+\sum_{n=1}^{m}\left(1-\zeta_{0}^{n}\right)\left|\tilde{s}^{n}\right|\leq\delta t\sum_{n=0}^{m}\left|\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right|\\&+\sum_{n=0}^{m}\left(1-\zeta_{0}^{n}\right)\left|E\left(\phi^{n}\right)-E\left(\phi\left(t^{n}\right)\right)\right|+\sum_{n=0}^{m}\left|T_{1}^{n}\right|.\end{aligned} \tag{4.30}$$

Similarly to the analysis of GSAV approach in [19], we can bound the right hand terms of (4.30) as follows. First, thanks to (4.24), we have

$$\left|T_{1}^{n}\right|\leq C\delta t\int_{t^{n}}^{t^{n+1}}\left|R_{tt}\right|ds\leq C\delta t\int_{t^{n}}^{t^{n+1}}\left(\left\|\phi_{t}(s)\right\|_{H^{1}}^{2}+\left\|\phi_{tt}(s)\right\|_{H^{1}}\right)\mathrm{d}s. \tag{4.31}$$

Next, by (3.2) and (3.9), we have $\tilde{R}^{n+1}<C$ and

$$\begin{aligned}&\left|\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\tfrac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right|\\&\leq\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}\left|1-\tfrac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\right|+\tfrac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\left|\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}-\left\|h\left(\bar{\phi}^{n+1}\right)\right\|^{2}\right|\\&:=P_{1}^{n}+P_{2}^{n}.\end{aligned} \tag{4.32}$$

By Theorem 4.1, we have $\left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^{2}<C$, and by $E(u)>\underline{C}>0$, we find

$$\begin{aligned}P_{1}^{n}&\leq C\left|1-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\right|\\&\leq C\left|\frac{R\left(t^{n+1}\right)}{E\left[\phi\left(t^{n+1}\right)\right]}-\frac{\tilde{R}^{n+1}}{E\left[\phi\left(t^{n+1}\right)\right]}\right|+C\left|\frac{\tilde{R}^{n+1}}{E\left[\phi\left(t^{n+1}\right)\right]}-\frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)}\right|\\&\leq C\left(\left|E\left[\phi\left(t^{n+1}\right)\right]-E\left(\bar{\phi}^{n+1}\right)\right|+\left|\tilde{s}^{n+1}\right|\right),\end{aligned} \tag{4.33}$$

and

$$\left|E\left[\phi\left(t^{n+1}\right)\right]-E\left(\bar{\phi}^{n+1}\right)\right|\leq C\bar{C}\left(\left\|\nabla\bar{e}^{n+1}\right\|+\left\|\bar{e}^{n+1}\right\|\right). \tag{4.34}$$

On the other hand,

$$P_2^n \le C\bar{C}\left(\left\|\Delta\bar{e}^{n+1}\right\| + \left\|\bar{e}^{n+1}\right\|\right), \tag{4.35}$$

and

$$\left|E\left(\phi^n\right) - E\left(\phi\left(t^n\right)\right)\right| \le \left|E\left(\phi^n\right) - E\left(\bar{\phi}^n\right)\right| + \left|E\left(\bar{\phi}^n\right) - E\left(\phi\left(t^n\right)\right)\right|. \tag{4.36}$$

Thanks to (4.13), Theorem 3.1 and (4.20), we have

$$\begin{aligned}
\left|E\left(\phi^n\right) - E\left(\bar{\phi}^n\right)\right| &\le \frac{1}{2}\left(\left\|\nabla\phi^n\right\| + \left\|\nabla\bar{\phi}^n\right\|\right)\left\|\nabla\phi^n - \nabla\bar{\phi}^n\right\| \\
&\quad + \frac{\lambda}{2}\left(\left\|\phi^n\right\| + \left\|\bar{\phi}^n\right\|\right)\left\|\phi^n - \bar{\phi}^n\right\| + \int_\Omega F\left(\phi^n\right) - F\left(\bar{\phi}^n\right)\mathrm{d}\boldsymbol{x} \\
&\le C\bar{C}\left(M_k + C_1\right)\left(\left\|\nabla\phi^n - \nabla\bar{\phi}^n\right\| + \left\|\phi^n - \bar{\phi}^n\right\|\right) \\
&\le C\bar{C}\left(M_k + C_1\right)\left|1 - \eta_k^n\right|\left\|\bar{\phi}^n\right\|_{H^1} \\
&\le C\bar{C}\left(M_k + C_1\right)C_0^{k+1}\delta t^{k+1},
\end{aligned} \tag{4.37}$$

and

$$\left|E\left(\bar{\phi}^n\right) - E\left(\phi\left(t^n\right)\right)\right| \le C\bar{C}\left(\left\|\nabla\bar{e}^n\right\| + \left\|\bar{e}^n\right\|\right) \le C\bar{C}\delta t^k. \tag{4.38}$$

Now, combining (4.21), (4.30)-(4.38), we obtain

$$\begin{aligned}
\left|\tilde{s}^{m+1}\right| &\le C\delta t\sum_{n=0}^{m}\left|\tilde{s}^{n+1}\right| + C\bar{C}\delta t\sum_{n=0}^{m}\left\|\bar{e}^{n+1}\right\|_{H^2} + C\bar{C}\left(M_k + C_1\right)C_0^{k+1}\delta t^k \\
&\quad + C\bar{C}\delta t^{k-1} + C\delta t\int_0^T\left(\|\phi_t(s)\|_{H^1}^2 + \|\phi_{tt}(s)\|_{H^1}\right)\mathrm{d}s \\
&\le C\delta t\sum_{n=0}^{m}\left|\tilde{s}^{n+1}\right| + C\bar{C}\left(\sqrt{C_2\left(1 + C_0^{2k+2}\right)} + \left(M_k + C_1\right)C_0^{k+1}\right)\delta t^k + C\delta t.
\end{aligned} \tag{4.39}$$

Applying the discrete Gronwall lemma to the above inequality with $\delta t < \frac{1}{2C}$, we derive

$$\begin{aligned}
\left|\tilde{s}^{m+1}\right| &\le C\exp\left(\left(1 - C\delta t\right)^{-1}\right)\delta t\left(\bar{C}\left(\sqrt{C_2\left(1 + C_0^{2k+2}\right)} + \left(M_k + C_1\right)C_0^{k+1}\right)\delta t^{k-1} + 1\right) \\
&\le C_3\delta t\left(\bar{C}\left(\sqrt{C_2\left(1 + C_0^{2k+2}\right)} + \left(M_k + C_1\right)C_0^{k+1}\right)\delta t^{k-1} + 1\right),
\end{aligned} \tag{4.40}$$

where C_3 is independent of C_0 and δt, can be defined as

$$C_3 := C\exp(2), \tag{4.41}$$

then $\delta t < \frac{1}{2C}$ can be guaranteed by

$$\delta t < \frac{1}{C_3}. \tag{4.42}$$

Hence, using (4.33), (4.34), (4.40) and (4.21), we have

$$\begin{aligned}
\left|1 - \xi_k^{m+1}\right| &\le C\left(\left|E\left[\phi\left(t^{m+1}\right)\right] - E\left(\bar{\phi}^{m+1}\right)\right| + \left|\tilde{s}^{m+1}\right|\right) \\
&\le C\left(\bar{C}\left\|\bar{e}^{m+1}\right\|_{H^1} + \left|\tilde{s}^{m+1}\right|\right) \\
&\le C\delta t\left(\bar{C}\sqrt{C_2\left(1 + C_0^{2k+2}\right)}\delta t^{k-1} + C_3\left(\bar{C}\left(\sqrt{C_2\left(1 + C_0^{2k+2}\right)} + \left(M_k + C_1\right)C_0^{k+1}\right)\delta t^{k-1} + 1\right)\right) \\
&\le C_4\delta t\left(\sqrt{1 + C_0^{2k+2}}\delta t^{k-1} + 1\right),
\end{aligned}$$

where C_4 is independent of C_0 and δt.

For the cases $k \geq 2$, we can define C_0 exactly the same as [19] with the condition $\delta t \leq \frac{1}{1+C_0^{k+1}}$ to obtain $\left|1-\xi_k^{m+1}\right| < C_0\delta t$.

The case $k=1$ needs special attention, and we need to modify the choices of η_1^{n+1} and ζ_0^{n+1} as specified in (4.9) and (4.10) for purely technical reasons. We can repeat the same process as the **Step 1** and **Step 2** in cases $k \geq 2$ and arrive at a similar result. In **Step 1**, we can get H^2 bound for ϕ^n and $\bar{\phi}^n$ for all $n \leq m$. In **Step 2**, we can obtain estimate for $\|\bar{e}^{n+1}\|_{H^2}$ for all $n \leq m$:

$$\|\bar{e}^{n+1}\|_{H^2} \leq \sqrt{C_2\left(1+C_0^6\delta t^2\right)}\delta t, \quad \forall 0 \leq n \leq m. \tag{4.43}$$

In **Step 3**, we need to discuss it case by case according to the different values of ζ_0^m.

(i) For Cases 1-3, we have $\zeta_0^m = 0$, then let $n=m$, we derive from (4.29) that

$$\left|\tilde{s}^{m+1}\right| \leq \delta t\left|\left\|h\left[\phi\left(t^{m+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{m+1}}{E\left(\bar{\phi}^{m+1}\right)}\left\|h\left(\bar{\phi}^{m+1}\right)\right\|^2\right| + \left|E\left(\phi^m\right)-E\left(\phi\left(t^m\right)\right)\right| + \left|T_1^m\right|. \tag{4.44}$$

Combining (4.43), (4.31)-(4.36), (4.38) and

$$\left|E\left(\phi^n\right)-E\left(\bar{\phi}^n\right)\right| \leq C\bar{C}\left(M_1+C_1\right)C_0^3\delta t^3, \tag{4.45}$$

we obtain

$$\begin{aligned}
\left|\tilde{s}^{m+1}\right| &\leq \delta t\sum_{n=0}^{m}\left|\left\|h\left[\phi\left(t^{m+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{m+1}}{E\left(\bar{\phi}^{m+1}\right)}\left\|h\left(\bar{\phi}^{m+1}\right)\right\|^2\right| + \left|E\left(\phi^m\right)-E\left(\phi\left(t^m\right)\right)\right| + \sum_{n=0}^{m}\left|T_1^n\right| \\
&\leq C\delta t\sum_{n=0}^{m}\left|\tilde{s}^{n+1}\right| + C\bar{C}\delta t\sum_{n=0}^{m}\left\|\bar{e}^{m+1}\right\|_{H^2} + C\bar{C}\left(M_1+C_1\right)C_0^3\delta t^3 + C\bar{C}\delta t \\
&\leq C\delta t\sum_{n=0}^{m}\left|\tilde{s}^{n+1}\right| + C\bar{C}\left(\sqrt{C_2\left(1+C_0^6\delta t^2\right)} + \left(M_k+C_1\right)C_0^3\delta t\right)\delta t + C\delta t.
\end{aligned} \tag{4.46}$$

Applying the discrete Gronwall lemma to the above inequality with $\delta t < \frac{1}{2C}$, we derive

$$\begin{aligned}
\left|\tilde{s}^{m+1}\right| &\leq C\exp\left(\left(1-C\delta t\right)^{-1}\right)\delta t\left(\bar{C}\left(\sqrt{C_2\left(1+C_0^6\delta t^2\right)} + \left(M_1+C_1\right)C_0^3\delta t\right)+1\right) \\
&\leq C_3\delta t\left(\bar{C}\left(\sqrt{C_2\left(1+C_0^6\delta t^2\right)} + \left(M_1+C_1\right)C_0^3\delta t\right)+1\right),
\end{aligned} \tag{4.47}$$

where C_3 is independent of C_0 and δt, can be defined as

$$C_3 := C\exp(2), \tag{4.48}$$

then $\delta t < \frac{1}{2C}$ can be guaranteed by

$$\delta t < \frac{1}{C_3}. \tag{4.49}$$

Similarly, we have

$$\begin{aligned}
\left|1-\xi_1^{m+1}\right| &\leq C\left(\left|E\left[\phi\left(t^{m+1}\right)\right]-E\left(\bar{\phi}^{m+1}\right)\right| + \left|\tilde{s}^{m+1}\right|\right) \\
&\leq C\left(\bar{C}\left\|\bar{e}^{m+1}\right\|_{H^1} + \left|\tilde{s}^{m+1}\right|\right) \\
&\leq C\delta t\left(\bar{C}\sqrt{C_2\left(1+C_0^6\delta t^2\right)} + C_3\left(\bar{C}\left(\sqrt{C_2\left(1+C_0^6\delta t^2\right)} + \left(M_1+C_1\right)C_0^3\delta t\right)+1\right)\right) \\
&\leq C_4\delta t\left(\sqrt{1+C_0^6\delta t^2}+1\right),
\end{aligned}$$

where C_4 is independent of C_0 and δt.

(ii) For Case 4, we choose $\zeta_0^m = 1 - \frac{\delta t^2\tilde{R}^m\|h(\bar{\phi}^m)\|^2}{E(\bar{\phi}^m)\left(E(\phi^m)-\tilde{R}^m\right)}$ as specified in (4.10), then it follows from (4.25) and (4.27) that

$$\tilde{s}^{m+1}-\tilde{s}^m = \delta t\left(\left\|h\left[\phi\left(t^{m+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{m+1}}{E\left(\bar{\phi}^{m+1}\right)}\left\|h\left(\bar{\phi}^{m+1}\right)\right\|^2\right) + T_1^m + \delta t^2\frac{\tilde{R}^m\|h(\bar{\phi}^m)\|^2}{E(\bar{\phi}^m)}. \tag{4.50}$$

Taking the sum of above from p $(0 \leq p \leq m)$ to m with $\zeta_{p-1} = 0$, and using (4.44), we obtain

$$\begin{aligned}
\left|\tilde{s}^{m+1}\right| &\leq \left|\tilde{s}^{p}\right| + \delta t \sum_{n=p}^{m} \left| \left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)} \left\|h\left(\bar{\phi}^{n+1}\right)\right\|^2 \right| + \sum_{n=p}^{m} \left|T_1^n\right| + \sum_{n=p}^{m} \delta t^2 \frac{\tilde{R}^n \|h(\bar{\phi}^n)\|^2}{E(\bar{\phi}^n)} \\
&\leq \delta t \sum_{n=p-1}^{m} \left| \left\|h\left[\phi\left(t^{n+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{n+1}}{E\left(\bar{\phi}^{n+1}\right)} \left\|h\left(\bar{\phi}^{n+1}\right)\right\|^2 \right| + \sum_{n=p-1}^{m} \left|T_1^n\right| + \sum_{n=p-1}^{m} \delta t^2 \frac{\tilde{R}^n \|h(\bar{\phi}^n)\|^2}{E(\bar{\phi}^n)} \\
&+ \left|E\left(\phi^{p-1}\right) - E\left(\phi\left(t^p - 1\right)\right)\right|.
\end{aligned} \tag{4.51}$$

Furthermore, we have

$$\left|\tilde{s}^{m+1}\right| \leq \delta t \sum_{n=0}^{m} \left| \left\|h\left[\phi\left(t^{m+1}\right)\right]\right\|^2 - \frac{\tilde{R}^{m+1}}{E\left(\bar{\phi}^{m+1}\right)} \left\|h\left(\bar{\phi}^{m+1}\right)\right\|^2 \right| + \sum_{n=0}^{m} \left|T_1^n\right| + C\delta t + \left|E\left(\phi^{p-1}\right) - E\left(\phi\left(t^p - 1\right)\right)\right|. \tag{4.52}$$

Then we can repeat the same process as (i) and derive that

$$\left|1 - \xi_1^{m+1}\right| \leq C_4 \delta t \left(\sqrt{1 + C_0^6 \delta t^2} + 1\right).$$

We can also define C_0 exactly the same as [19] with the condition $\delta t \leq \frac{1}{C_0^3}$ to obtain $\left|1 - \xi_1^{m+1}\right| < C_0 \delta t$.

Finally, we can show that

$$\begin{aligned}
\left\|e^{m+1}\right\|_{H^2}^2 &\leq 2\left\|\bar{e}^{m+1}\right\|_{H^2}^2 + 2\left\|\phi^{m+1} - \bar{\phi}^{m+1}\right\|_{H^2}^2 \\
&\leq 2\left\|\bar{e}^{m+1}\right\|_{H^2}^2 + 2\left|\eta_k^{m+1} - 1\right|^2 \left\|\bar{\phi}^{m+1}\right\|_{H^2}^2.
\end{aligned}$$

Then for $k \geq 2$, we have

$$\left\|e^{m+1}\right\|_{H^2}^2 \leq 2C_2\left(1 + C_0^{2(k+1)}\right)\delta t^{2k} + 2\bar{C}^2 C_0^{2(k+1)} \delta t^{2(k+1)},$$

and for $k = 1$, we have

$$\left\|e^{m+1}\right\|_{H^2}^2 \leq 2C_2\left(1 + C_0^6 \delta t^2\right)\delta t^2 + 2\bar{C}^2 C_0^6 \delta t^6,$$

provided that $\delta t < \min\left\{\frac{1}{1+2C_0^{k+2}}, \frac{1-\tau_k}{3k}\right\}$. The proof is complete. □

Similar results can also be established for the Cahn-Hilliard type equation

$$\frac{\partial \phi}{\partial t} = \Delta(-\Delta\phi + \lambda\phi + F'(\phi)), \quad (\boldsymbol{x}, t) \in \Omega \times (0, T], \tag{4.53}$$

which is (1.2) with $\mathcal{L} = -\Delta + \lambda I$ and with $\mathcal{G} = -\Delta$, with initial condition $\phi(x, 0) = \phi^0(x)$ and periodic or Neumann boundary condition. It can also be written in the form of (2.17) with $\mathcal{A} = -\Delta(-\Delta + \lambda I)$ and $g(\phi) = -\Delta(F'(\phi))$. The corresponding (2.18) is with $E_{tot}(\phi) = \int_\Omega \frac{1}{2}|\nabla\phi|^2 + \lambda|\phi|^2 + F(\phi)\mathrm{d}\boldsymbol{x}$ and $\mathcal{K}(\phi) = \|\nabla(-\Delta\phi + \lambda\phi + F'(\phi))\|^2$.

Theorem 4.4. *Given initial condition $\bar{\phi}^i = \phi^i = \phi(t^i)$, $R^i = E(\phi^i)$, $i = 0, 1, ..., k-1$. Let $\bar{\phi}^{n+1}$ and ϕ^{n+1} be computed with the R-GSAV/BDFk ($1 \leq k \leq 5$) scheme (3.1)-(3.5) for (4.53), except that when $k = 1$, we set*

$$\eta_1^{n+1} = 1 - \left(1 - \xi_1^{n+1}\right)^3, \tag{4.54}$$

and for Case 4 of $k = 1$, we set

$$\zeta_0^{n+1} = 1 - \frac{\delta t^2 \tilde{R}^{n+1} \mathcal{K}(\bar{\phi}^{n+1})}{E(\bar{\phi}^{n+1})(E(\phi^{n+1}) - \tilde{R}^{n+1})}. \tag{4.55}$$

We assume (4.2) and (4.4) holds and

$$\phi \in C\left([0, T]; H^3\right), \quad \frac{\partial^j \phi}{\partial t^j} \in L^2\left(0, T; H^2\right), 1 \leq j \leq k, \quad \frac{\partial^{k+1}\phi}{\partial t^{k+1}} \in L^2\left(0, T; H^1\right). \tag{4.56}$$

Then for $\delta t < \min\{\frac{1}{1+4C_0^{k+2}}, \frac{1-\tau_k}{3k}\}$, we have

$$\left\|\bar{\phi}^{n+1}-\phi(\cdot,t^{n+1})\right\|_{H^2}, \quad \left\|\phi^{n+1}-\phi(\cdot,t^{n+1})\right\|_{H^2} \le C\delta t^k, \quad \forall n+1 \le T/\delta t,$$

where τ_k is given in (4.1), and the constants C_0, C are independent of δt.

The above results can be established by combining the proofs of the above theorem and Theorem in [19], we leave the detail to the interested readers.

5. Extension to the multiple SAV approach

In some cases, the nonlinear part of the free energy may contain disparate terms such that schemes with a single SAV may require excessively small time steps to obtain correct simulations [7]. It is shown in [7] that the multiple SAV (MSAV) approach can overcome this difficulty.

In this section, we demonstrate how to construct relaxed MGSAV schemes for gradient flow. Without loss of generality, we consider the following gradient flow with two disparate nonlinear terms (extension to more than two disparate nonlinear terms is straightforward):

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = -\mathcal{G}\mu, \\ \mu = \mathcal{L}\phi + F_1'(\phi) + F_2'(\phi), \end{cases} \tag{5.1}$$

where $\mathcal{L}$ is a linear self-adjoint elliptic operator, $F_1(\phi), F_2(\phi)$ are nonlinear potential function, $\mathcal{G}$ is a positive definite linear operator. The system (5.1) satisfies an energy dissipation law as follows

$$\frac{\mathrm{d}E_{tot}(\phi)}{\mathrm{d}t} = -\left(\mathcal{G}\mu, \mu\right), \tag{5.2}$$

where

$$E_{tot}(\phi) = \frac{1}{2}(\mathcal{L}\phi, \phi) + \int_\Omega F_1(\phi)\mathrm{d}\boldsymbol{x} + \int_\Omega F_2(\phi)\mathrm{d}\boldsymbol{x} \tag{5.3}$$

is a free energy with lower bound $-C_0$. Setting $E(\phi) = E_{tot}(\phi) + C_0 = E_1(\phi) + E_2(\phi)$ with $E_1(\phi) = \frac{1}{2}(\mathcal{L}\phi, \phi) + \int_\Omega F_1(\phi)\mathrm{d}\boldsymbol{x} + C_1 > 0$, $E_2(\phi) = \int_\Omega F_2(\phi)\mathrm{d}\boldsymbol{x} + C_2 > 0$ and introducing two SAVs $R_1(t) = E_1(\phi)$, $R_2(t) = E_2(\phi)$, we can rewrite the equation (5.1) as

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = -\mathcal{G}\mu, \\ \mu = \dfrac{\delta E}{\delta \phi} = \mathcal{L}\phi + F_1'(\phi) + F_2'(\phi), \\ \dfrac{\mathrm{d}R_1(t)}{\mathrm{d}t} = -\dfrac{R_1(t)+R_2(t)}{E_1(\phi)+E_2(\phi)}\left(\mathcal{G}\dfrac{\delta E_1}{\delta\phi}, \mu\right), \\ \dfrac{\mathrm{d}R_2(t)}{\mathrm{d}t} = -\dfrac{R_1(t)+R_2(t)}{E_1(\phi)+E_2(\phi)}\left(\mathcal{G}\dfrac{\delta E_2}{\delta\phi}, \mu\right). \end{cases} \tag{5.4}$$

Note that the above MSAV formulation is different from that used in [7]. Then we construct the relaxed MGSAV BDFk schemes as follows:

Given $\phi^{n-k}, ..., \phi^n, R_1^{n-k}, ..., R_1^n, R_2^{n-k}, ..., R_2^n$, we compute $\phi^{n+1}, R_1^{n+1}, R_2^{n+1}$ via the following two steps:

Step 1:

- solve ϕ_1^{n+1} and ϕ_2^{n+1} from

$$\frac{\alpha_k\phi_1^{n+1} - \frac{1}{2}A_k(\phi^n)}{\delta t} = -\mathcal{G}\left(\mathcal{L}\phi_1^{n+1} + F_1'(B_k(\phi^n))\right), \tag{5.5}$$

$$\frac{\alpha_k\phi_2^{n+1} - \frac{1}{2}A_k(\phi^n)}{\delta t} = -\mathcal{G}\left(\mathcal{L}\phi_2^{n+1} + F_2'(B_k(\phi^n))\right); \tag{5.6}$$

and set

$$\bar{\phi}^{n+1} = \phi_1^{n+1} + \phi_2^{n+1}, \tag{5.7}$$

$$\bar{\mu}^{n+1} = \mathcal{L}\bar{\phi}^{n+1} + F_1'(\bar{\phi}^{n+1}) + F_2'(\bar{\phi}^{n+1}); \tag{5.8}$$

- solve $\tilde{R}_1^{n+1}$ and $\tilde{R}_2^{n+1}$ from

$$\frac{\tilde{R}_1^{n+1}-R_1^n}{\delta t}=-\frac{\tilde{R}_1^{n+1}+\tilde{R}_2^{n+1}}{E_1(\bar{\phi}^{n+1})+E_2(\bar{\phi}^{n+1})}\left(\mathcal{G}\frac{\delta E_1}{\delta\phi}\left(\bar{\phi}^{n+1}\right),\bar{\mu}^{n+1}\right), \tag{5.9}$$

$$\frac{\tilde{R}_2^{n+1}-R_2^n}{\delta t}=-\frac{\tilde{R}_1^{n+1}+\tilde{R}_2^{n+1}}{E_1(\bar{\phi}^{n+1})+E_2(\bar{\phi}^{n+1})}\left(\mathcal{G}\frac{\delta E_2}{\delta\phi}\left(\bar{\phi}^{n+1}\right),\bar{\mu}^{n+1}\right); \tag{5.10}$$

- set

$$\xi_{k,1}^{n+1}=\frac{\tilde{R}_1^{n+1}}{E_1(\bar{\phi}^{n+1})},\quad \xi_{k,2}^{n+1}=\frac{\tilde{R}_2^{n+1}}{E_2(\bar{\phi}^{n+1})}, \tag{5.11}$$

$$\eta_{k,1}^{n+1}=1-(1-\xi_{k,1}^{n+1})^{k+1},\quad \eta_{k,2}^{n+1}=1-(1-\xi_{k,2}^{n+1})^{k+1}, \tag{5.12}$$

$$\phi^{n+1}=\eta_{k,1}^{n+1}\phi_1^{n+1}+\eta_{k,2}^{n+1}\phi_2^{n+1}, \tag{5.13}$$

$$\mu^{n+1}=\mathcal{L}\phi^{n+1}+F_1'(\phi^{n+1})+F_2'(\phi^{n+1}); \tag{5.14}$$

Step 2: Update R_1^{n+1}, R_2^{n+1} by

$$R_1^{n+1}=\zeta_0^{n+1}\tilde{R}_1^{n+1}+(1-\zeta_0^{n+1})E_1(\phi^{n+1}),\quad R_2^{n+1}=\zeta_0^{n+1}\tilde{R}_2^{n+1}+(1-\zeta_0^{n+1})E_2(\phi^{n+1}),\quad \zeta_0^{n+1}\in\mathcal{V}. \tag{5.15}$$

Here $\mathcal{V}$ is a set defined by

$$\mathcal{V}=\Bigg\{\zeta\in[0,1]\ s.t.\ \frac{\left(R_1^{n+1}+R_2^{n+1}\right)-\left(\tilde{R}_1^{n+1}+\tilde{R}_2^{n+1}\right)}{\delta t}= -\gamma^{n+1}\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)+\frac{\tilde{R}_1^{n+1}+\tilde{R}_2^{n+1}}{E_1(\bar{\phi}^{n+1})+E_2(\bar{\phi}^{n+1})}\left(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1}\right),\quad \gamma^{n+1}\geq 0\Bigg\}, \tag{5.16}$$

with $\gamma^{n+1}\geq 0$ to be determined so that $\mathcal{V}$ is not empty.

Setting $\tilde{R}^{n+1}=\tilde{R}_1^{n+1}+\tilde{R}_2^{n+1}$, $R^{n+1}=R_1^{n+1}+R_2^{n+1}$, $E(\bar{\phi}^{n+1})=E_1(\bar{\phi}^{n+1})+E_2(\bar{\phi}^{n+1})$ and plugging (5.15) into the equality of (5.16), we find that if we choose ζ_0^{n+1} and γ^{n+1} such that the following condition is satisfied:

$$(\tilde{R}^{n+1}-E(\phi^{n+1}))\zeta_0^{n+1}=\tilde{R}^{n+1}-E(\phi^{n+1})-\delta t\gamma^{n+1}\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\left(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1}\right), \tag{5.17}$$

then $\zeta_0^{n+1}\in\mathcal{V}$. Following the same arguments as in the proof of Theorem 3.1, we can prove the following results for the schemes (5.5)-(5.15).

Theorem 5.1. *We choose ζ_0^{n+1} in (5.15) and γ^{n+1} in (5.16) as follows:*

1. *If $\tilde{R}^{n+1}=E(\phi^{n+1})$, we set $\zeta_0^{n+1}=0$ and $\gamma^{n+1}=\frac{\tilde{R}^{n+1}(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1})}{E(\bar{\phi}^{n+1})(\mathcal{G}\mu^{n+1},\mu^{n+1})}$.*
2. *If $\tilde{R}^{n+1}>E(\phi^{n+1})$, we set $\zeta_0^{n+1}=0$ and*

$$\gamma^{n+1}=\frac{\tilde{R}^{n+1}-E(\phi^{n+1})}{\delta t\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)}+\frac{\tilde{R}^{n+1}\left(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1}\right)}{E(\bar{\phi}^{n+1})\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)}. \tag{5.18}$$

3. *If $\tilde{R}^{n+1}<E(\phi^{n+1})$ and $\tilde{R}^{n+1}-E(\phi^{n+1})+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\left(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1}\right)\geq 0$, we set $\zeta_0^{n+1}=0$ and γ^{n+1} the same as (5.18).*
4. *If $\tilde{R}^{n+1}<E(\phi^{n+1})$ and $\tilde{R}^{n+1}-E(\phi^{n+1})+\delta t\frac{\tilde{R}^{n+1}}{E(\bar{\phi}^{n+1})}\left(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1}\right)<0$, we set $\zeta_0^{n+1}=1-\frac{\delta t\tilde{R}^{n+1}(\mathcal{G}\bar{\mu}^{n+1},\bar{\mu}^{n+1})}{E(\bar{\phi}^{n+1})\left(E(\phi^{n+1})-\tilde{R}^{n+1}\right)}$ and $\gamma^{n+1}=0$.*

Then, (5.17) is satisfied in all cases above and $\zeta_0^{n+1}\in\mathcal{V}$. Furthermore, given $R^n\geq 0$, we have $R^{n+1}\geq 0$, and the scheme (5.5)-(5.15) with the above choice of ζ_0^{n+1} and γ^{n+1} is unconditionally energy stable in the sense that

$$R^{n+1}-R^n=-\delta t\gamma^{n+1}\left(\mathcal{G}\mu^{n+1},\mu^{n+1}\right)\leq 0. \tag{5.19}$$

Furthermore, we have

$$R^{n+1}\leq E(\phi^{n+1})\quad \forall n\geq 0. \tag{5.20}$$

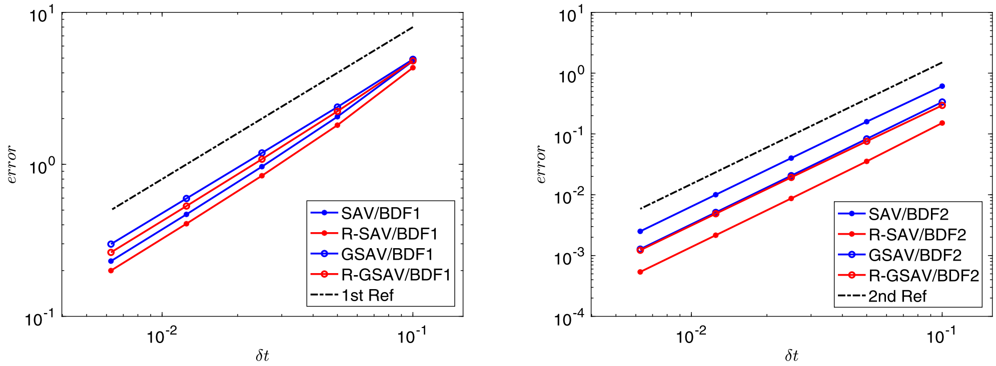


Fig. 1. Example 1A. Convergence rates for Allen-Cahn equation using various schemes. Left: first-order; Right: Second-order. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

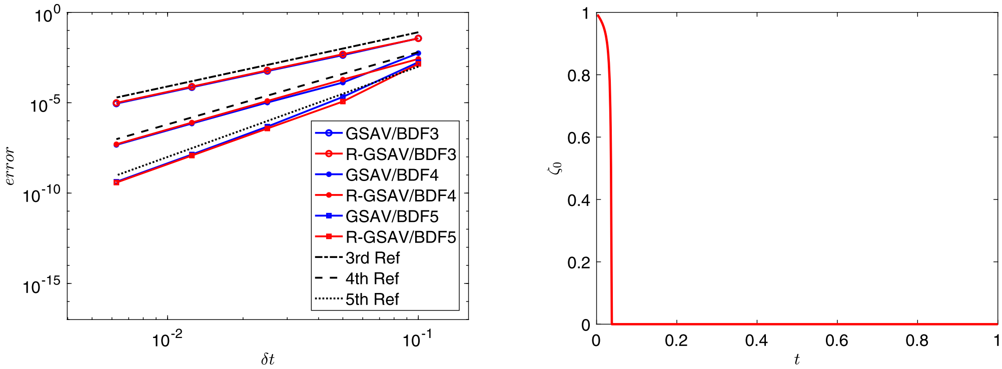


Fig. 2. Example 1A. Left: GSAV/BDFk and R-GSAV/BDFk ($k = 3, 4, 5$) schemes; Right: evolution of relaxation ζ_0^{n+1} using R-GSAV/BDF2 scheme with $\delta t = 1e-3$.

6. Numerical results and discussions

We present in this section some numerical results to validate the efficiency and accuracy of the R-GSAV approach, and provide detailed comparisons between the original SAV, R-SAV, GSAV and R-GSAV approaches.

Unless specified otherwise, we consider examples with periodic boundary condition and use the Fourier spectral method for spatial discretization. The default value of the parameter γ is set to 0.95 for the R-SAV approaches.

Example 1. The Allen-Cahn equation

$$\frac{\partial \phi}{\partial t} = \alpha \Delta \phi - \left(1 - \phi^2\right)\phi. \tag{6.1}$$

Case A. We add an external force f to (6.1) so that its exact solution is

$$\phi(x, y, t) = \exp(\sin(\pi x)\sin(\pi y))\sin(t). \tag{6.2}$$

We set $\Omega = [0, 2] \times [0, 2]$, $\alpha = 0.01^2$, and use 64^2 Fourier modes for space discretization so that the spatial discretization error is negligible when compared with the time discretization error.

In Fig. 1, we plot the convergence rate of the H^2 error at $T = 1$ by using various first- and second-order schemes. We observe that (i) the expected convergence rates are obtained for all cases; (ii) the errors of R-SAV (resp. R-GSAV) schemes are significantly smaller than that of SAV (resp. GSAV) schemes; (iii) the R-SAV approach is the most accurate but it requires solving two linear systems. In the left of Fig. 2, we plot the convergence rate of the H^2 error at $T = 1$ by using GSAV/BDFk and R-GSAV/BDFk ($k = 3, 4, 5$) schemes, and observe that all schemes achieve their desired order of accuracy, but the improvements by R-SAV and R-GSAV over SAV and GSAV for higher-order schemes are not as significant as for lower-order schemes. In the right of Fig. 2, we present evolution of relaxation parameter ζ_0^{n+1} using R-GSAV/BDF2 scheme with $\delta t = 1e-3$, and observe that, except at an initial time interval, ζ_0^{n+1} takes the value zero.

Table 1
Example 1B. A comparison of L^2-error by SAV/BDF2, R-SAV/BDF2, GSAV/BDF2 and R-GSAV/BDF2 schemes for Allen-Cahn equation at $T = 200$ with different time step.

	SAV	R-SAV	GSAV	R-GSAV
1E-1	4.30E-04	2.72E-04	1.27E-03	2.65E-04
1E-2	4.53E-05	2.93E-06	1.47E-04	2.90E-06
1E-3	6.26E-07	2.96E-08	2.38E-06	2.94E-08

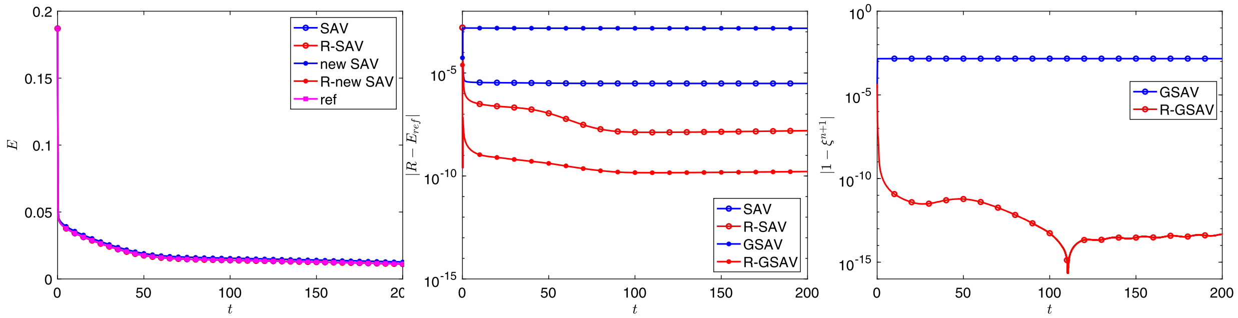


Fig. 3. Example 1B. Allen-Cahn equation: a comparison of energy (left) and energy error (middle) of SAV/BDF2, R-SAV/BDF2, GSAV/BDF2 and R-GSAV/BDF2 schemes; and a comparison of error of ξ^{n+1} of GSAV/BDF2 and R-GSAV/BDF2 schemes (right).

Case B. We set $\Omega = [0, L_x] \times [0, L_y]$ with $L_x = L_y = 1$, and choose the initial condition as

$$\begin{aligned} &\phi(x, y) = \tanh \frac{1.5 + 1.2\cos(6\theta) - 2\pi r}{\sqrt{2\alpha}}, \\ &\theta = \arctan \frac{y - 0.5L_y}{x - 0.5L_x}, \quad r = \sqrt{\left(x - \frac{L_x}{2}\right)^2 + \left(y - \frac{L_y}{2}\right)^2}, \end{aligned} \tag{6.3}$$

where (θ, r) are the polar coordinates of (x, y). The other parameters are $\alpha = 0.01^2, m_0 = 0.1$ and 128^2 Fourier modes. We use the results of the semi-implicit/BDF2 scheme with $\delta t = 1e-5$ as the reference solution. The L^2-norm error of four schemes at $T = 200$ with different time steps are shown in Table 1. We observe that R-GSAV (resp. R-SAV) schemes can significantly reduce the error of the solution compared with GSAV (resp. SAV) schemes, and the effect of R-GSAV scheme on improving accuracy is more obvious. In Fig. 3, we present a comparison of energy (left) and energy error (middle) of schemes with $\delta t = 1e-3$. Fig. 3 (right) shows the evolution of error of ξ^{n+1}, which indicates that the R-GSAV scheme can improve the accuracy of ξ^{n+1}.

Example 2. The Cahn-Hilliard equation

$$\frac{\partial \phi}{\partial t} = -m_0 \Delta \left(\alpha \Delta \phi - \left(1 - \phi^2\right)\phi\right). \tag{6.4}$$

Case A. We set the exact solution to be (6.2), and set $\alpha = 0.04, m_0 = 0.005$. Convergence rates of different schemes are presented in Fig. 4. The results are similar to those for the Allen-Cahn equation.

Case B. We set the initial condition as in (6.3), and set $m_0 = 0.1, \alpha = 0.01^2$. The other parameters are chosen to be the same as in Case B of Example 1. Numerical solutions at $T = 0.1$ using the GSAV/BDF2 and R-GSAV/BDF2 schemes with $\delta t = 1e-3$ are plotted in Fig. 5 along with the reference solution obtained by semi-Implicit/BDF2 scheme with time step $\delta t = 1e-5$. We observe that with $\delta t = 1e-3$, the solution by the GSAV scheme is totally wrong while the solution by the R-GSAV scheme is indistinguishable with the reference solution. We also observe that for this example $\zeta_0^{n+1} = 0$ at all times.

Example 3. In order to show that the R-GSAV approach can be used to simulate more complex nonlinear phenomena, we consider, as an example, the phase-field crystal model

$$\begin{cases} \frac{\partial \phi}{\partial t} = M\Delta\mu, \quad \boldsymbol{x} \in \Omega, t > 0, \\ \mu = (\Delta + \beta)^2\phi + \phi^3 - \epsilon\phi, \quad \boldsymbol{x} \in \Omega, t > 0, \\ \phi(\boldsymbol{x}, 0) = \phi_0(\boldsymbol{x}), \end{cases} \tag{6.5}$$

which is a gradient flow based on the following total free energy

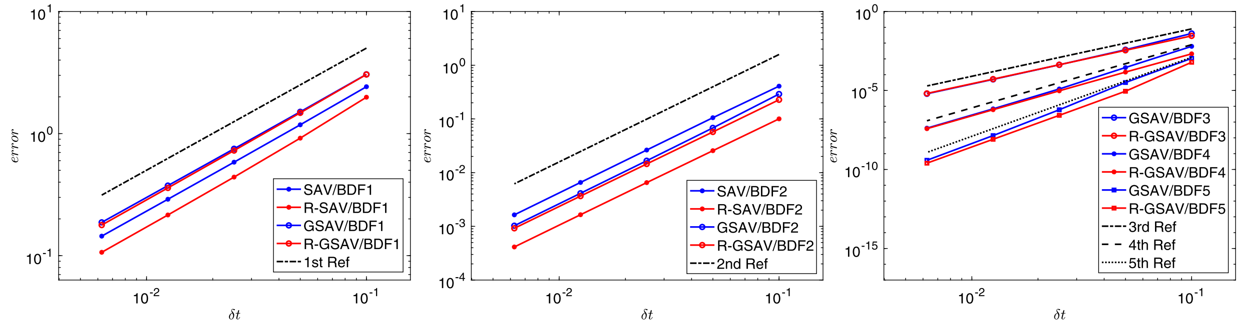


Fig. 4. Example 2A. Convergence test for Cahn-Hilliard equation using SAV/BDFk, R-SAV/BDFk ($k = 1, 2$), GSAV/BDFk and R-GSAV/BDFk ($k = 1, 2, 3, 4, 5$) schemes.

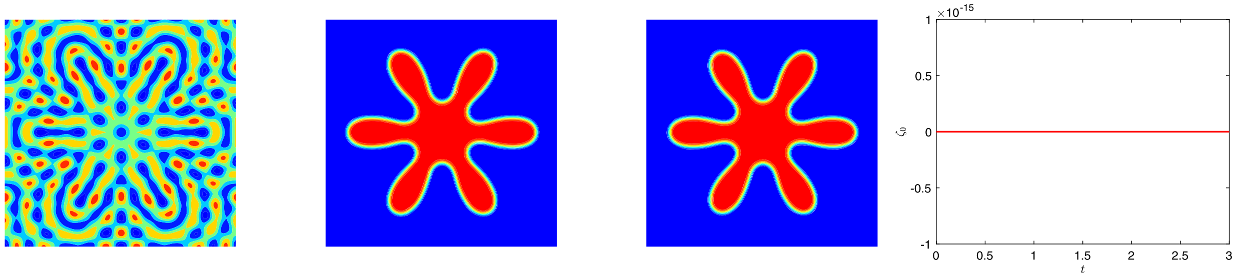


Fig. 5. Example 2B. Profiles of ϕ at $T = 0.1$. Dynamics driven by the Cahn-Hilliard equation using GSAV/BDF2 scheme (first), R-GSAV/BDF2 scheme (second) and reference solution (third), and the evolution of ζ_0^{n+1} (fourth).

$$E(\phi) = \int_\Omega \left(\frac{1}{2}\phi(\Delta + \beta)^2\phi + \frac{1}{4}\phi^4 - \frac{\epsilon}{2}\phi^2 \right) d\boldsymbol{x}, \tag{6.6}$$

where $M > 0$ is the mobility coefficient. In the following simulations, we choose $M = 1, \beta = 1$.

Case A. Crystal growth in a super-cooled liquid in $2D$. We set the initial condition to be

$$\phi(x_l, y_l, 0) = \bar{\phi} + C_1 \left(\cos\left(\frac{C_2}{\sqrt{3}} y_l \right) \cos(C_2 x_l) - 0.5 \cos\left(\frac{2C_2}{\sqrt{3}} y_l \right) \right), \quad l = 1, 2, 3, \tag{6.7}$$

where x_l and y_l define a local system of Cartesian coordinates that is oriented with the crystallite lattice, and the constant parameters $\bar{\phi} = 0.285, C_1 = 0.446, C_2 = 0.66$. Then, three crystallites in three small square patches with each length of 40 which located at $(350, 400)$, $(200, 200)$, and $(600, 300)$ respectively, are defined perfectly. In order to generate crystallites with different orientations, we use the following affine transformation to produce rotation

$$x_l(x, y) = x\sin(\theta) + y\cos(\theta), \quad y_l(x, y) = -x\cos(\theta) + y\sin(\theta), \tag{6.8}$$

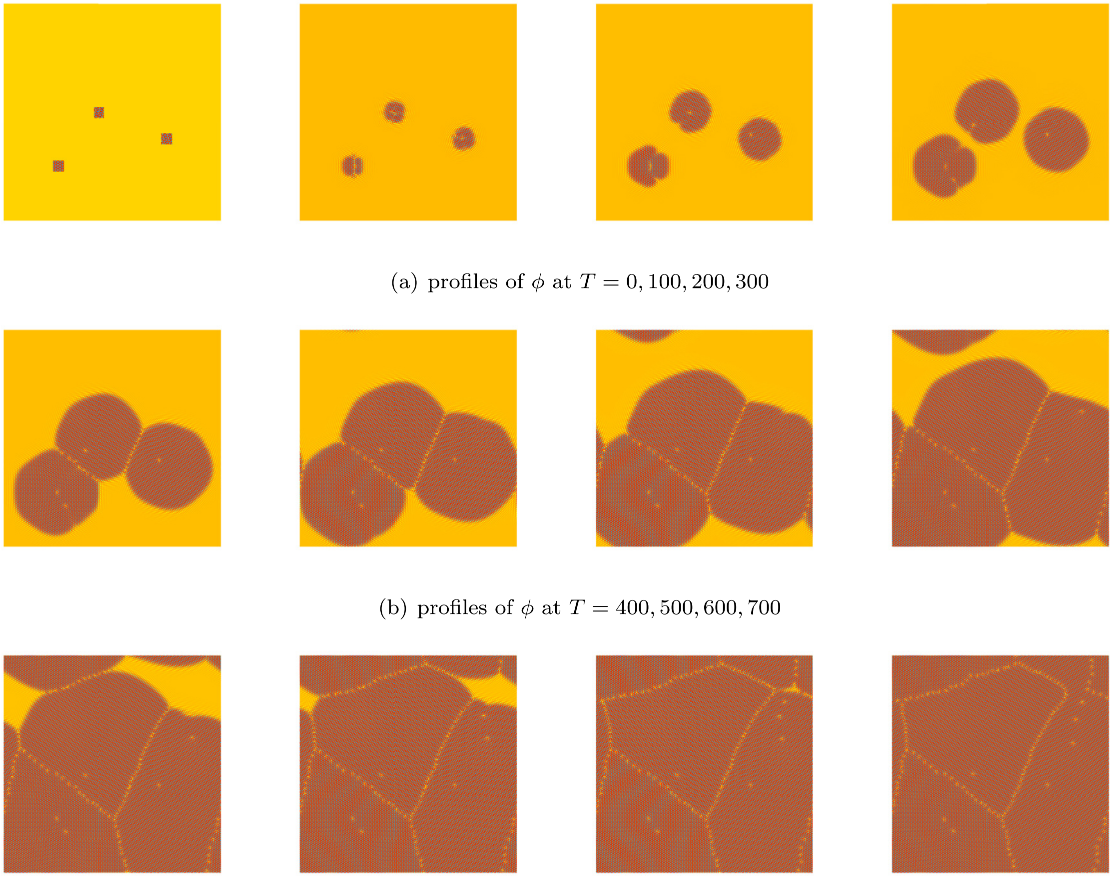
where angles are chosen as $\theta = -\frac{\pi}{4}, 0, \frac{\pi}{4}$ respectively. We choose 1024^2 Fourier modes to discretize the space and use relatively small the time step $\delta t = 0.02$ for better accuracy. And we take the other parameters $\epsilon = 0.25, T = 2000$. Fig. 6 shows crystal growth in a super-cooled liquid driven by the PFC equation using the R-GSAV/BDF2 scheme. It also demonstrates that the different alignment of the crystallites causes defects and dislocations. These results are consistent with those in [24,38]. For this example, the relaxation parameter ζ_0^{n+1} is also zero at all times.

Case B. Phase transition behaviors in $3D$. We choose the initial data $\phi(x, y, t = 0) = \bar{\phi} + 0.01 rand$ and computational domains $[0, 50]^3$. Other parameters are chosen as $\epsilon = 0.56, \delta t = 0.02, T = 3000$ and 64^3 Fourier modes. Fig. 7 shows the steady state microstructure of the phase transition behavior for $\bar{\phi} = -0.20, -0.35$ and -0.43, respectively. These results are also consistent with those in [22].

Example 4. In this example, we use the phase-field vesicle membrane (PFVM) model [7,8] as an example to demonstrate how to construct relaxed MSAV schemes.

Since the vesicle membrane is area and volume preserving, we consider the following penalized free energy

$$E_{tot}(\phi) = E_b(\phi) + \frac{1}{2\sigma_1}(A(\phi) - \alpha)^2 + \frac{1}{2\sigma_2}(B(\phi) - \beta)^2, \tag{6.9}$$



(a) profiles of ϕ at $T = 0, 100, 200, 300$

(b) profiles of ϕ at $T = 400, 500, 600, 700$

(c) profiles of ϕ at $T = 800, 900, 1000, 2000$

Fig. 6. Example 3A. The dynamic evolution of crystal growth in a supercooled liquid driven by the PFC equation using R-GSAV/BDF2 scheme. Snapshots of the numerical solution ϕ at $T = 0$, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, respectively.

where σ_1 and σ_2 are two small parameters, and α, β represent the initial volume and surface area, the bending energy $E_b(\phi)$, volume $A(\phi)$ and surface area $B(\phi)$ of the vesicle are defined by

$$E_b(\phi) = \frac{\epsilon}{2}\int_\Omega \left(-\Delta\phi + \frac{1}{\epsilon^2}G(\phi)\right)^2 \mathrm{d}\boldsymbol{x} = \frac{\epsilon}{2}\int_\Omega w^2 \mathrm{d}\boldsymbol{x}, \tag{6.10}$$

$$A(\phi) = \int_\Omega (\phi + 1)\mathrm{d}\boldsymbol{x} \quad \text{and} \quad B(\phi) = \int_\Omega \left(\frac{\epsilon}{2}|\nabla\phi|^2 + \frac{1}{\epsilon}F(\phi)\right)\mathrm{d}\boldsymbol{x}, \tag{6.11}$$

where

$$w := -\Delta\phi + \frac{1}{\epsilon^2}G(\phi), \quad G(\phi) := F'(\phi), \quad F(\phi) = \frac{1}{4}\left(\phi^2 - 1\right)^2.$$

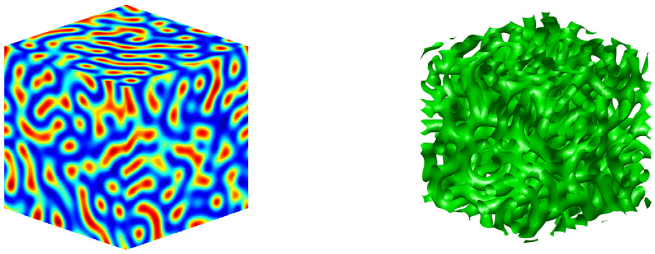
Then, the L^2 gradient flow associated with the above free energy is

$$\begin{cases} \phi_t = -M\mu, \\ \mu = -\epsilon\Delta w + \frac{1}{\epsilon}G'(\phi)w + \frac{1}{\sigma_1}(A(\phi) - \alpha) + \frac{1}{\sigma_2}(B(\phi) - \beta)\left(-\epsilon\Delta\phi + \frac{1}{\epsilon}F'(\phi)\right), \\ w = -\Delta\phi + \frac{1}{\epsilon^2}G(\phi), \end{cases} \tag{6.12}$$

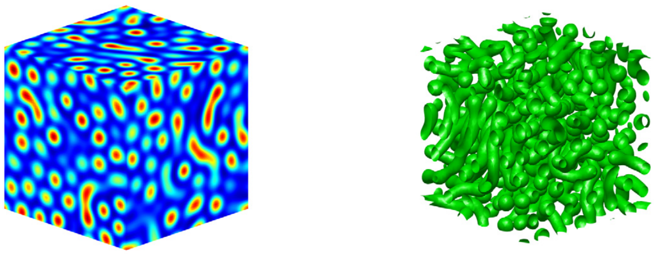
with the boundary conditions being

$$\text{(i) periodic or (ii)} \ \ \partial_{\mathbf{n}}\phi|_{\partial\Omega} = \partial_{\mathbf{n}}\Delta\phi|_{\partial\Omega} = 0, \tag{6.13}$$

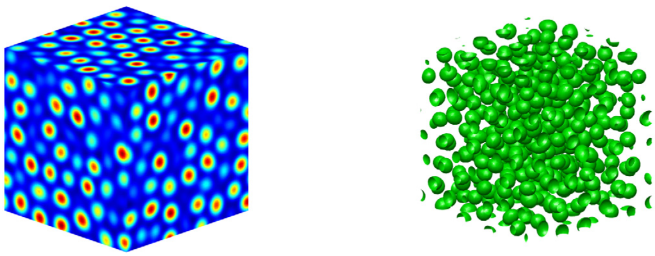
and M is the mobility constant. Then, one can easily see that the system (6.12) admits the following energy law



(a) $\bar{\phi} = -0.2$



(b) $\bar{\phi} = -0.35$



(c) $\bar{\phi} = -0.43$

Fig. 7. Example 3B. Evolution of ϕ driven by the PFC equation using R-GSAV/BDF2 scheme. Snapshots of density field ϕ (left) and isosurface plots of $\phi = 0$ (right) at $T = 3000$.

$$\frac{\mathrm{d}}{\mathrm{d}t} E_{tot}(\phi) = -M\|\mu\|^2. \tag{6.14}$$

We observe that (6.12) contains small parameters ϵ, σ_1, σ_2, but σ_1 is only associated with the linear non-local term $A(\phi)$ so it can be treated implicitly. Hence, we need to introduce two SAVs to deal with the nonlinear terms associated with ϵ and σ_2 separately. More precisely, we set

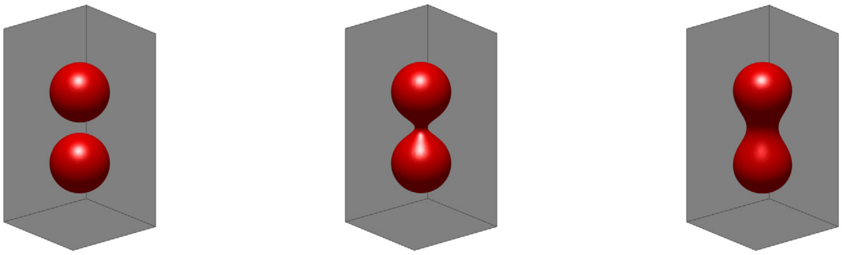
$$E_1(\phi) = E_b(\phi) + \frac{1}{2\sigma_1}(A(\phi) - \alpha)^2, \quad E_2(\phi) := \frac{1}{2\sigma_2}(B(\phi) - \beta)^2. \tag{6.15}$$

Then, we can apply the R-MGSAV scheme (5.5)-(5.15) directly.

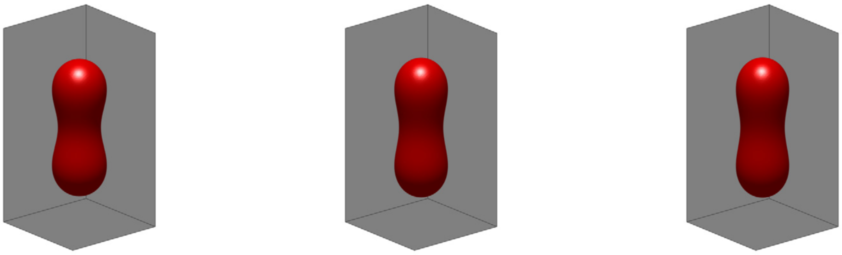
We consider the phase-field vesicle membrane model (6.12) in $\Omega = (-\pi, \pi)^3$ with $\epsilon = \frac{6\pi}{128}$, $M = 1$ and $\sigma_1 = \sigma_2 = 0.01$. We use the R-MSAV/BDF2 scheme with $\delta t = 1e - 4$ and 128^3 Fourier modes.

Case A. We first simulate the evolution of two close-by spherical vesicles by consider the following initial condition for ϕ to describe two close-by spherical vesicles in 3D

$$\begin{aligned}\phi(x, y, z, 0) = \tanh\left(\frac{0.28\pi - \sqrt{x^2 + y^2 + (z - 0.35\pi)^2}}{\sqrt{2}\epsilon}\right) \\ + \tanh\left(\frac{0.28\pi - \sqrt{x^2 + y^2 + (z + 0.35\pi)^2}}{\sqrt{2}\epsilon}\right) + 1.\end{aligned} \tag{6.16}$$

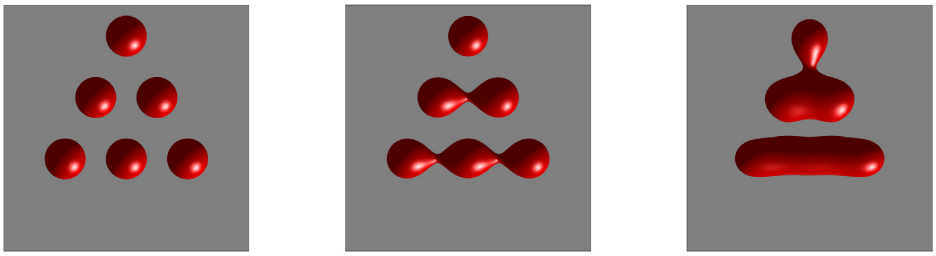


(a) profiles of $\phi = 0$ at $T = 0, 0.02, 0.1$

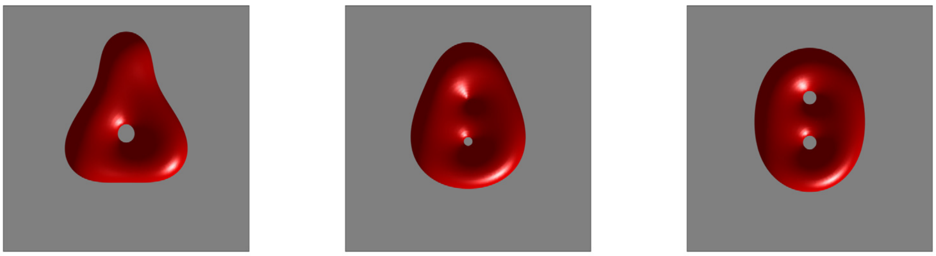


(b) profiles of $\phi = 0$ at $T = 0.5, 1, 2$

Fig. 8. Example 4A. The evolution of two close-by spherical vesicles: Snapshots of iso-surfaces of $\phi = 0$ driven by the PFVM equation at $T = 0, 0.02, 0.1, 0.5, 1, 2$.



(a) profiles of $\phi = 0$ at $T = 0, 0.02, 0.1$



(b) profiles of $\phi = 0$ at $T = 0.5, 1, 2$

Fig. 9. Example 4B. The evolution of six close-by spherical vesicles. Snapshots of iso-surfaces of $\phi = 0$ driven by the PFVM equation at $T = 0, 0.02, 0.1, 0.5, 1, 2$.

We depict the evolution process in Fig. 8. We observe that two spheres connect within a small time interval, then merge into a capsule shape which is a steady state. The results are consistent with those presented in [7].

Case B. Then we consider six closeby spheres as initial condition given by

$$\phi(x,y,z,0)=\sum_{i=1}^{6}\tanh\left(\frac{r_i-\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}}{\sqrt{2}\epsilon}\right)+5, \tag{6.17}$$

where $r_i=\frac{\pi}{6}, z_i=0$ for $i=1,2,\dots,6$, $(x_1,x_2,x_3,x_4,x_5,x_6)=\left(-\frac{\pi}{4},\frac{\pi}{4},0,\frac{\pi}{2},-\frac{\pi}{2},0\right)$, and $(y_1,y_2,y_3,y_4,y_5,y_6)=(-\frac{\pi}{4},-\frac{\pi}{4},\frac{\pi}{4},\frac{\pi}{4},\frac{\pi}{4},-\frac{3\pi}{4})$.

In Fig. 9, we plot snapshots of iso-surfaces $\phi=0$ at $T=0,0.02,0.1,0.5,2$ by using the R-MGSAV/BDF2 scheme. It shows that the initially separated spheres connect with each other gradually and finally merge into a big vesicle with two small holes in upper and lower parts respectively. The results are also consistent with those presented in [8].

Finally, we note that for the simulations in all examples except in the accuracy test, the relaxation parameter ζ_0^{n+1} is zero at all times. This indicates that, at least for these simulations, the modified energy is in fact equal to the original energy, which means that the R-GSAV schemes are effectively energy stable with the original energy.

In summary, the R-GSAV approach fixes a flaw in the GSAV approach and leads to more robust and accurate numerical schemes while keeping the simplicity, efficiency and generality of the GSAV approach.

CRediT authorship contribution statement

Yanrong Zhang and Jie Shen contributed to the conceptualization, methodology and writing. Yanrong Zhang carried out the numerical simulations.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Georgios Akrivis, Buyang Li, Dongfang Li, Energy-decaying extrapolated RK-SAV methods for the Allen-Cahn and Cahn-Hilliard equations, SIAM J. Sci. Comput. 41 (6) (2019) A3703–A3727.
[2] Georgios Akrivis, Dongfang Li, Structure-preserving Gauss methods for the nonlinear Schrödinger equation, Calcolo 58 (2) (2021) 1–25.
[3] Xavier Antoine, Jie Shen, Qinglin Tang, Scalar Auxiliary Variable/Lagrange multiplier based pseudospectral schemes for the dynamics of nonlinear Schrödinger/Gross-Pitaevskii equations, J. Comput. Phys. 437 (2021) 110328.
[4] Arvind Baskaran, John S. Lowengrub, Cheng Wang, Steven M. Wise, Convergence analysis of a second order convex splitting scheme for the modified phase field crystal equation, SIAM J. Numer. Anal. 51 (5) (2013) 2851–2873.
[5] Anass Bouchriti, Morgan Pierre, Nour Eddine Alaa, Remarks on the asymptotic behavior of scalar auxiliary variable (SAV) schemes for gradient-like flows, J. Appl. Anal. Comput. 10 (5) (2020) 2198–2219.
[6] Qing Cheng, Chun Liu, Jie Shen, A new Lagrange multiplier approach for gradient flows, Comput. Methods Appl. Mech. Eng. 367 (2020) 113070.
[7] Qing Cheng, Jie Shen, Multiple scalar auxiliary variable (MSAV) approach and its application to the phase-field vesicle membrane model, SIAM J. Sci. Comput. 40 (6) (2018) A3982–A4006.
[8] Qing Cheng, Jie Shen, Global constraints preserving scalar auxiliary variable schemes for gradient flows, SIAM J. Sci. Comput. 42 (4) (2020) A2489–A2513.
[9] Qiang Du, Xiaobing Feng, The phase field method for geometric moving interfaces and their numerical approximations, Handb. Numer. Anal. 21 (2020) 425–508.
[10] Qiang Du, Lili Ju, Xiao Li, Zhonghua Qiao, Maximum principle preserving exponential time differencing schemes for the nonlocal Allen–Cahn equation, SIAM J. Numer. Anal. 57 (2) (2019) 875–898.
[11] Qiang Du, Lili Ju, Xiao Li, Zhonghua Qiao, Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes, SIAM Rev. 63 (2) (2021) 317–359.
[12] Charles M. Elliott, A.M. Stuart, The global dynamics of discrete semilinear parabolic equations, SIAM J. Numer. Anal. 30 (6) (1993) 1622–1663.
[13] David J. Eyre, Unconditionally gradient stable time marching the Cahn-Hilliard equation, in: Computational and Mathematical Models of Microstructural Evolution, San Francisco, CA, 1998, in: Materials Research Society Symposium Proceedings, vol. 529, MRS, Warrendale, PA, 1998, pp. 39–46.
[14] Xiaobing Feng, Buyang Li, Shu Ma, High-order mass- and energy-conserving SAV-Gauss collocation finite element methods for the nonlinear Schrödinger equation, SIAM J. Numer. Anal. 59 (3) (2021) 1566–1591.
[15] Yuezheng Gong, Qi Hong, Qi Wang, Supplementary variable method for thermodynamically consistent partial differential equations, Comput. Methods Appl. Mech. Eng. 381 (2021) 113746.
[16] Yuezheng Gong, Jia Zhao, Qi Wang, Arbitrarily high-order unconditionally energy stable schemes for thermodynamically consistent gradient flow models, SIAM J. Sci. Comput. 42 (1) (2020) B135–B156.
[17] Dianming Hou, Chuanju Xu, Robust and stable schemes for time fractional molecular beam epitaxial growth model using SAV approach, J. Comput. Phys. 445 (2021) 110628.
[18] Dianming Hou, Chuanju Xu, A second order energy dissipative scheme for time fractional L^2 gradient flows using SAV approach, J. Sci. Comput. 90 (1) (2022) 1–22.
[19] Fukeng Huang, Jie Shen, A new class of implicit-explicit BDFk SAV schemes for general dissipative systems and their error analysis, Comput. Methods Appl. Mech. Eng. 392 (2022) 114718.
[20] Fukeng Huang, Jie Shen, Zhiguo Yang, A highly efficient and accurate new scalar auxiliary variable approach for gradient flows, SIAM J. Sci. Comput. 42 (4) (2020) A2514–A2536.
[21] Maosheng Jiang, Zengyan Zhang, Jia Zhao, Improving the accuracy and consistency of the scalar auxiliary variable (SAV) method with relaxation, J. Comput. Phys. (2022) 110954.

[22] Qi Li, Liquan Mei, Xiaofeng Yang, Yibao Li, Efficient numerical schemes with unconditional energy stabilities for the modified phase field crystal equation, Adv. Comput. Math. 45 (3) (2019) 1551–1580.
[23] Xiaoli Li, Jie Shen, On a SAV-MAC scheme for the Cahn-Hilliard-Navier-Stokes phase-field model and its error analysis for the corresponding Cahn-Hilliard-Stokes case, Math. Models Methods Appl. Sci. 30 (12) (2020) 2263–2297.
[24] Xiaoli Li, Jie Shen, Stability and error estimates of the SAV Fourier-spectral method for the phase field crystal equation, Adv. Comput. Math. 46 (2020) 1–20.
[25] Xiaoli Li, Jie Shen, Zhengguang Liu, New SAV-pressure correction methods for the Navier-Stokes equations: stability and error analysis, Math. Comput. 91 (333) (2022) 141–167.
[26] Lianlei Lin, Zhiguo Yang, Suchuan Dong, Numerical approximation of incompressible Navier-Stokes equations based on an auxiliary energy variable, J. Comput. Phys. 388 (2019) 1–22.
[27] Olavi Nevanlinna, F. Odeh, Multiplier techniques for linear multistep methods, Numer. Funct. Anal. Optim. 3 (4) (1981) 377–423.
[28] Yanxia Qian, Zhiguo Yang, Fei Wang, Suchuan Dong, gPAV-based unconditionally energy-stable schemes for the Cahn-Hilliard equation: stability and error analysis, Comput. Methods Appl. Mech. Eng. 372 (2020) 113444.
[29] Jie Shen, Cheng Wang, Xiaoming Wang, Steven M. Wise, Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: application to thin film epitaxy, SIAM J. Numer. Anal. 50 (1) (2012) 105–125.
[30] Jie Shen, Jie Xu, Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows, SIAM J. Numer. Anal. 56 (5) (2018) 2895–2912.
[31] Jie Shen, Jie Xu, Jiang Yang, The scalar auxiliary variable (SAV) approach for gradient flows, J. Comput. Phys. 353 (2018) 407–416.
[32] Jie Shen, Jie Xu, Jiang Yang, A new class of efficient and robust energy stable schemes for gradient flows, SIAM Rev. 61 (3) (2019) 474–506.
[33] Jie Shen, Xiaofeng Yang, Numerical approximations of Allen-Cahn and Cahn-Hilliard equations, Discrete Contin. Dyn. Syst. 28 (4) (2010) 1669.
[34] Jie Shen, Xiaofeng Yang, The IEQ and SAV approaches and their extensions for a class of highly nonlinear gradient flow systems, Contemp. Math. 754 (2020) 217–245.
[35] Roger Temam, Infinite-Dimensional Dynamical Systems in Mechanics and Physics, second edition, Applied Mathematical Sciences., vol. 68, Springer-Verlag, New York, 1997.
[36] Xiaoqiang Wang, Lili Ju, Qiang Du, Efficient and stable exponential time differencing Runge–Kutta methods for phase field elastic bending energy models, J. Comput. Phys. 316 (2016) 21–38.
[37] Xiaofeng Yang, Linear first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends, J. Comput. Phys. 327 (2016) 294–316.
[38] Xiaofeng Yang, Daozhi Han, Linearly first-and second-order, unconditionally energy stable schemes for the phase field crystal model, J. Comput. Phys. 330 (2017) 1116–1134.
[39] Xiaofeng Yang, Lili Ju, Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model, Comput. Methods Appl. Mech. Eng. 315 (2017) 691–712.
[40] Jun Zhang, Xiaofeng Yang, Efficient and accurate numerical scheme for a magnetic-coupled phase-field-crystal model for ferromagnetic solid materials, Comput. Methods Appl. Mech. Eng. 371 (2020) 113310.
[41] Jun Zhang, Xiaofeng Yang, A fully decoupled, linear and unconditionally energy stable numerical scheme for a melt-convective phase-field dendritic solidification model, Comput. Methods Appl. Mech. Eng. 363 (2020) 112779.
[42] Jingzhi Zhu, Long-Qing Chen, Jie Shen, Veena Tikare, Coarsening kinetics from a variable-mobility Cahn-Hilliard equation: application of a semi-implicit Fourier spectral method, Phys. Rev. E 60 (4) (1999) 3564.
[43] Qingqu Zhuang, Jie Shen, Efficient SAV approach for imaginary time gradient flows with applications to one-and multi-component Bose-Einstein condensates, J. Comput. Phys. 396 (2019) 72–88.