Error Analysis with Polynomial Dependence on ε^{-1} in SAV Methods for the Cahn-Hilliard Equation

Ma, Shu; Qiu, Weifeng; Yang, Xiaofeng

Published in:
Journal of Scientific Computing

Published: 01/12/2024

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

Published version (DOI):
10.1007/s10915-024-02734-8

Publication details:
Ma, S., Qiu, W., & Yang, X. (2024). Error Analysis with Polynomial Dependence on ε^{-1} in SAV Methods for the Cahn-Hilliard Equation. *Journal of Scientific Computing*, *101*(3), Article 83. https://doi.org/10.1007/s10915-024-02734-8

Journal of Scientific Computing (2024) 101:83
https://doi.org/10.1007/s10915-024-02734-8

Error Analysis with Polynomial Dependence on ε^{-1} in SAV Methods for the Cahn-Hilliard Equation

Shu Ma[1] · **Weifeng Qiu**[1] · **Xiaofeng Yang**[2]

Received: 16 March 2024 / Revised: 16 October 2024 / Accepted: 30 October 2024 /
Published online: 14 November 2024

Abstract
The optimal error estimate that depends only on the polynomial degree of ε^{-1} is established for the temporal semi-discrete scheme of the Cahn-Hilliard equation based on the scalar auxiliary variable (SAV) formulation. The key to our analysis is converting the structure of the SAV time-stepping scheme back to a form compatible with the original format of the Cahn-Hilliard equation, which makes it feasible to use spectral estimates to handle the nonlinear term. Based on the transformation of the SAV numerical scheme, the optimal error estimate for the temporal semi-discrete scheme which depends only on the low polynomial order of ε^{-1} instead of the exponential order, is derived by using mathematical induction, spectral arguments, and the superconvergence properties of some nonlinear terms. Numerical examples are provided to illustrate the discrete energy decay property and validate our theoretical convergence analysis.

Keywords Cahn-Hilliard equation · SAV formulation · Energy decay · Spectral estimates · Polynomial order · Error estimates

1 Introduction

In this paper we consider the initial boundary value problem for the Cahn-Hilliard (CH) phase field equation

$$\partial_t u = \Delta\big(-\varepsilon\Delta u + \frac{1}{\varepsilon} f(u)\big) \qquad \text{in } \Omega \times (0, T], \tag{1.1a}$$

$$\partial_{\boldsymbol{n}} u = \partial_{\boldsymbol{n}}\big(-\varepsilon\Delta u + \frac{1}{\varepsilon} f(u)\big) = 0 \qquad \text{on } \partial\Omega \times (0, T], \tag{1.1b}$$

✉ Weifeng Qiu
weifeqiu@cityu.edu.hk

Shu Ma
shuma2@cityu.edu.hk

Xiaofeng Yang
xfyang@math.sc.edu

[1] Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong, People's Republic of China

[2] Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

$$u(\cdot, 0) = u_0, \qquad \text{in } \Omega, \tag{1.1c}$$

where $\Omega \subset R^d$, $d = 2, 3$ is a bounded domain, $\boldsymbol{n}$ is the outward normal, ε is a small parameter, and f is the derivative of a non-negative potential function F with two local minima, i.e., $f = F'$. For instance, the Ginzburg-Landau energy function

$$F(v) = \frac{1}{4}(v^2 - 1)^2 \quad \text{and} \quad f(v) = v^3 - v.$$

In view of its wide application as a phase field model [10, 39], many numerical methods and analyses have been developed to approximate the Cahn-Hilliard equation (1.1). However, most of these methods have been developed for the Cahn-Hilliard equation with fixed $\varepsilon > 0$. As pointed out in [4, 15, 17, 18], error estimates using the direct Gronwall inequality argument yield a constant factor $e^{T/\varepsilon}$, resulting in exponential growth of the error as $\varepsilon \to 0$. Such an estimate is clearly not useful for very small value of ε, especially in solving the problem of whether the computed numerical interface converges to the original sharp interface of the Hele-Shaw problem when $\varepsilon \to 0$ (see [14, 28] for details). To overcome this difficulty, Feng and Prohl [25] first established a priori error estimates with polynomial dependence on ε^{-1} for the time-discrete format of the Cahn-Hilliard equation. Subsequently, Feng and Wu obtained a posteriori error estimates with polynomial-order of ε^{-1} in [26] for the same time-discrete methods. The main idea of the polynomial-order error estimates of ε^{-1} is to use the spectral estimates given by Alikakos and Fusco [3] and Chen [13] for the linearized Cahn-Hilliard operator to handle the nonlinear term in the error analysis. After this, spectral estimates have been frequently used to eliminate the exponential dependence on ε^{-1} in the error analyses of other numerical methods for the Cahn-Hilliard equations (see [6, 19, 21, 29, 34] and references therein) and the related phase field equations, including the Allen-Cahn equations in [1, 6–8, 20, 23, 24, 27], the Ginzburg-Landau equations in [5], and the phase field models with nonlinear constitutive laws in [16].

On the other hand, it is well known that the Cahn-Hilliard equation (1.1) is the H^{-1}-gradient flow of the energy functional

$$E[u] := \int_\Omega \left(\frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon}F(u)\right) \mathrm{d}x. \tag{1.2}$$

As a result, the solution of the Cahn-Hilliard equation has decaying energy. Indeed, testing (1.1) by $-\varepsilon\Delta u + \frac{1}{\varepsilon}f(u)$ yields

$$E[u(\cdot, t_2)] - E[u(\cdot, t_1)] = -\int_{t_1}^{t_2} \|\partial_t u\|_{H^{-1}}^2 \,\mathrm{d}t \le 0, \tag{1.3}$$

where $0 \le t_1 \le t_2 \le T$ and $\|\partial_t u\|_{H^{-1}} = \|\nabla w\|$ with $w = -\varepsilon\Delta u + \frac{1}{\varepsilon}f(u)$.

Accordingly, great efforts have been devoted to the construction of efficient and accurate numerical methods that preserve the energy decay properties at the discrete level. In particular, for widely used linear time-stepping methods with energy decay, including stabilized semi-implicit schemes [11, 12, 33], invariant energy quadratization (IEQ) methods [35–37], and the scalar auxiliary variable (SAV) approach in [30, 31] and [2], optimal error estimates of these schemes have been well established for the Cahn-Hilliard equations with fixed ε^{-1}. However, the main challenge with error analysis in these methods is how to establish error bounds that depend only on the polynomial order of ε^{-1} rather than the exponential order for small $\varepsilon \to 0$. This difficulty arises because, unlike in Feng's previous work [19, 21, 25, 26, 29], the numerical methods based on the IEQ/SAV formulation break the standard structure of the nonlinear term of the Cahn-Hilliard equation, which is crucial to the utilization of the

spectral arguments. As a result, spectral estimates become ineffective in estimating errors for the IEQ/SAV approach. To the best of our knowledge, optimal error estimates that depend only on the polynomial order of ε^{-1} for the IEQ/SAV methods to the Cahn-Hilliard equation remain open.

The objective of this paper is to establish error bounds that depend only on low polynomial order of ε^{-1} for semi-discrete methods based on the SAV formulation of the Cahn-Hilliard equation. The SAV reformulation of the Cahn-Hilliard equation was introduced in [31, 32] as an enhanced version of the invariant energy quadratization (IEQ) approach [35–38], for developing energy-decay methods at the discrete level. By reconstructing the system based on the SAV reformulation, a linear and easy-to-implement time semi-discrete scheme is obtained. The SAV formulation introduces new difficulties to the error analysis for the Cahn-Hilliard equation due to the presence of a new scalar r in the nonlinear part (see equation (2.5b)), which alters the structure of the original Cahn-Hilliard equation and makes the spectral argument used in [19, 21, 25, 26, 29] not directly applicable. To overcome this problem, we transform the structure of the SAV scheme using the new scalar variable into a form that is compatible with the original format of the Cahn-Hilliard equation (1.1). This transformation enable us to use spectral estimates in our error analysis. However, this transformation will accordingly introduces a strong perturbation term (see equation (4.11)) that needs to be delicately controlled. To address this issue, an inductive argument is used to deal with the difficulties that arise in the structural transformation. Specifically, we need to establish error bounds of $\|\nabla e^{i-1} - \nabla e^{i-2}\|$ and $\|e^{i-1} - e^{i-2}\|_{H^{-1}}$ for $i \leq n$. By integrating them into the perturbation terms, and utilizing the super-convergence characteristics of resulting nonlinear terms, we can complete the mathematical induction method. Therefore, by reconverting the scheme format using the new variable back to a form compatible with the original format of the Cahn-Hilliard equation, and using mathematical induction, the superconvergence characteristics of some nonlinear terms, as well as the spectrum argument, we derive optimal error estimates that depend only on polynomial degree of ε^{-1}. To the best of our knowledge, our numerical analysis has the following properties that were not possessed in the existing literature:

(1) We have addressed the challenge of spectral estimates being ineffective in the SAV approach for the Cahn-Hilliard equation;

(2) This is the first error estimate with polynomial dependence on ε^{-1} for the IEQ/SAV-type schemes of the Cahn-Hilliard equation (1.1).

The techniques developed here can also be applied for other IEQ/SAV-type schemes of phase-field models.

The rest of this paper is organized as follows. In Sect. 2, we present the SAV reformulation of the Cahn-Hilliard equation and introduce an equivalent transformation of the SAV time-stepping scheme. In Sect. 3, we show the properties of energy decay and derive the consistency estimates for the proposed method. In Sect. 4, we present an error estimate of the semi-discrete SAV scheme to derive a convergence rate that does not depend on ε^{-1} exponentially. The spectrum estimate plays a crucial role in the proof. Finally, in Sect. 5, we present a few numerical experiments to validate the theoretical results.

2 Formulation of the Semi-Discrete SAV Scheme

In this section, we construct a backward Euler implicit-explicit type temporal semi-discrete numerical scheme based on the SAV reformulation of the CH equation (1.1), and also present an equivalent formulation of the SAV scheme.

2.1 Function Spaces

Let $W^{s,p}(\Omega)$ denote the usual Sobolev spaces, and $H^s(\Omega)$ denote the Hilbert spaces $W^{s,2}(\Omega)$ with norm $\|\cdot\|_{H^s}$. Let $\|\cdot\|$ and $(\cdot,\cdot)$ represent the L^2 norm and L^2 inner product, respectively. In addition, define for $p \geq 0$

$$H^{-p}(\Omega) := (H^p(\Omega))^*, \quad H_0^{-p}(\Omega) := \{u \in H^{-p}(\Omega) \mid \langle u, 1\rangle_p = 0\}, \tag{2.1}$$

where $\langle\cdot,\cdot\rangle_p$ stands for the dual product between $H^p(\Omega)$ and $H^{-p}(\Omega)$. We denote $L_0^2(\Omega) := H_0^0(\Omega)$. For $v \in L_0^2(\Omega)$, let $-\Delta^{-1}v := v_1 \in H^1(\Omega) \cap L_0^2(\Omega)$, where v_1 is the solution to

$$-\Delta v_1 = v \text{ in } \Omega, \qquad \partial_{\boldsymbol{n}} v_1 = 0 \text{ on } \partial\Omega, \tag{2.2}$$

and $\|v\|_{-1} := \sqrt{(v, -\Delta^{-1}v)}$.

For $v \in L_0^2(\Omega) \cap H^1(\Omega)$, we have the following inequality

$$\|v\|^2 = (\nabla v, \nabla(-\Delta)^{-1}v) \leq \|\nabla v\|\|v\|_{-1}. \tag{2.3}$$

We denote by C generic constant and C_i, $\tilde{C}_i$, $\tilde{C}$, κ_i' and κ_i specific constants, which are independent of τ, h and ε, but may possibly depend on the domain Ω, T and the constants of Sobolev inequalities. We use notation $\lesssim$ in the sense that $f \lesssim g$ means that $f \leq Cg$ with positive constant C independent of τ, h and ε.

2.2 The SAV Reformulation

The SAV formulation of the CH equation (cf. [30, 31]) introduces a scalar auxiliary variable

$$r = \sqrt{\int_\Omega F(u)\mathrm{d}x + c_0} \quad \text{with} \quad g(u) = \frac{f(u)}{\sqrt{\int_\Omega F(u)\mathrm{d}x + c_0}}, \tag{2.4}$$

with a positive c_0 (which guarantees that the function r has a positive lower bound), and reformulate (1.1) as

$$\partial_t u = \Delta w \qquad \text{in } \Omega \times (0,T], \tag{2.5a}$$

$$w = -\varepsilon\Delta u + \frac{1}{\varepsilon} r g(u) \qquad \text{in } \Omega \times (0,T], \tag{2.5b}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{1}{2}\big(g(u), \partial_t u\big) \qquad \text{in } \Omega \times (0,T], \tag{2.5c}$$

$$\partial_{\boldsymbol{n}} u = \partial_{\boldsymbol{n}} w = 0 \qquad \text{on } \partial\Omega \times (0,T], \tag{2.5d}$$

$$u(\cdot,0) = u_0 \qquad \text{in } \Omega, \tag{2.5e}$$

$$r(0) = \sqrt{\int_\Omega F(u_0)\mathrm{d}x + c_0}. \tag{2.5f}$$

We define an energy functional with respect to u and r:

$$E(u,r) = \frac{\varepsilon}{2}\|\nabla u\|^2 + \frac{1}{\varepsilon}r^2, \tag{2.6}$$

and taking the L^2 inner product of the first equation (2.5a) with w, of the second equation (2.5b) with $\partial_t u$, and of the third equation (2.5c) with $\frac{1}{\varepsilon}2r$, performing integration by parts and summing up the two obtained equations, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}E(u,r) = -\|\nabla w\|^2 = -\|\partial_t u\|_{-1}^2 \le 0. \tag{2.7}$$

2.3 The Equivalent Formulation of the SAV Scheme

Let $\{t_n\}_{n=0}^{N+1}$ be a uniform partition of $[0, T]$ with the time step size τ, where N is a positive integer and hence $\tau = \frac{T}{N+1}$. We consider the following temporal semi-discrete SAV scheme for solving the system (2.5):

$$\begin{cases} \dfrac{u^{n+1}-u^n}{\tau} = \Delta w^{n+1}, \\ w^{n+1} = -\varepsilon\Delta u^{n+1} + \dfrac{1}{\varepsilon}r^{n+1}g(u^n), \\ r^{n+1} - r^n = \dfrac{1}{2}\big(g(u^n), u^{n+1}-u^n\big), \\ \partial_{\boldsymbol{n}} u^{n+1}|_{\partial\Omega} = \partial_{\boldsymbol{n}} w^{n+1}|_{\partial\Omega} = 0, \end{cases} \tag{2.8}$$

with $u^0 = u_0$ and $r^0 = r(0)$ for $n = 0, 1, \ldots, N$.

By taking the L^2 inner product of the first equation in (2.8) with $v = 1$, we have the following conservation property, which is important to the error estimates.

Lemma 2.1 *The numerical solution of* (2.8) *satisfies*

$$\frac{1}{|\Omega|}\int_\Omega u^n(x)\,\mathrm{d}x = \frac{1}{|\Omega|}\int_\Omega u^0(x)\,\mathrm{d}x, \qquad n = 1, \ldots, N, \tag{2.9}$$

and the error function $e^n := u(t_n) - u^n$ *satisfies*

$$\int_\Omega e^n\,\mathrm{d}x = 0, \qquad n = 1, \ldots, N. \tag{2.10}$$

Due to (2.9–2.10), both $u^{n+1} - u^n$ and e^n belong to $L_0^2(\Omega)$, allowing us to define their $\|\cdot\|_{-1}$ norm. Let $\hat{w}^{n+1} = w^{n+1} - \int_\Omega w^{n+1}\mathrm{d}x$ such that $\hat{w}^{n+1} \in L_0^2(\Omega)$. The first equation in (2.8) satisfies

$$\frac{u^{n+1}-u^n}{\tau} = \Delta w^{n+1} = \Delta\hat{w}^{n+1}. \tag{2.11}$$

Since both sides of the equation belong to $L_0^2(\Omega)$, we can apply Δ^{-1} to get

$$\Delta^{-1}\frac{u^{n+1}-u^n}{\tau} = \hat{w}^{n+1},$$

which combines with

$$(\hat{w}^{n+1}, v) = (w^{n+1}, v) \qquad \forall\ v \in L_0^2(\Omega), \tag{2.12}$$

gives

$$\Big(\Delta^{-1}\frac{u^{n+1}-u^n}{\tau}, v\Big) = (w^{n+1}, v) \qquad \forall\ v \in L_0^2(\Omega). \tag{2.13}$$

Thus, by testing the second equation in (2.8) with $v \in L_0^2(\Omega)$ and using (2.13), the semi-discrete SAV scheme (2.8) can be written as

$$\begin{cases} \Big(\Delta^{-1}\dfrac{u^{n+1}-u^n}{\tau}, v\Big) = \varepsilon\big(\nabla u^{n+1}, \nabla v\big) + \dfrac{1}{\varepsilon} r^{n+1}\big(g(u^n), v\big) \qquad \forall\ v \in L_0^2(\Omega), \\ r^{n+1} - r^n = \dfrac{1}{2}\big(g(u^n), u^{n+1} - u^n\big), \qquad n = 0, 1, \ldots, N. \end{cases} \tag{2.14}$$

In order to avoid the exponentially dependence of the error bound on $\frac{1}{\varepsilon}$ induced by using the Gronwall inequality, we need to use a spectral estimate of the linearized Cahn-Hilliard operator, which is given in [3, 13, 25] and will be described in Sect. 4.

However, compared with the previous work [25], the SAV method (2.8) alters the structure of the CH equation such that the spectral argument can not be applied directly. To achieve the ideal error bound, we need to transform the structure into a form compatible with the CH equation (1.1) so that the spectral estimate of the linearized Cahn-Hilliard operator can be used. To this end, we define $A(v) = \sqrt{\int_\Omega F(v)\,\mathrm{d}x + c_0}$, the Gateaux derivatives of $A(v)$ can be defined as follows:

$$DA(v, w) := \frac{1}{2}\Big(\frac{f(v)}{\sqrt{\int_\Omega F(v)\,\mathrm{d}x + c_0}}, w\Big) = \frac{1}{2}\big(g(v), w\big), \tag{2.15}$$

$$D^2A(v, w) := \frac{1}{2}\frac{\int_\Omega f'(v)w^2\,\mathrm{d}x}{\sqrt{\int_\Omega F(v)\,\mathrm{d}x + c_0}} - \frac{1}{4}\frac{\big(\int_\Omega f(v)w\,\mathrm{d}x\big)^2}{\Big(\int_\Omega F(v)\,\mathrm{d}x + c_0\Big)^{\frac{3}{2}}}. \tag{2.16}$$

By Taylor expansion, we derive

$$A(u^i) = A(u^{i-1}) + \frac{1}{2}\big(g(u^{i-1}), u^i - u^{i-1}\big) + \frac{1}{2}D^2A(\xi_i\,; u^i - u^{i-1}), \tag{2.17}$$

where $\xi_i = \theta u^i + (1-\theta)u^{i-1}$ with $\theta \in (0, 1)$. Thus we get

$$\frac{1}{2}\big(g(u^{i-1}), u^i - u^{i-1}\big) = A(u^i) - A(u^{i-1}) - \frac{1}{2}D^2A(\xi_i\,; u^i - u^{i-1}),$$

which together with the second equation in (2.14) implies

$$r^i - r^{i-1} = A(u^i) - A(u^{i-1}) - \frac{1}{2}D^2A(\xi_i\,; u^i - u^{i-1}).$$

After summing up the above equation from $i = 1$ to n, we derive

$$r^n - r^0 = A(u^n) - A(u^0) - \frac{1}{2}\sum_{i=1}^{n} D^2A(\xi_i\,; u^i - u^{i-1}).$$

Since $r^0 = A(u^0)$, we obtain

$$r^n = A(u^n) - \frac{1}{2}\sum_{i=1}^{n} D^2A(\xi_i\,; u^i - u^{i-1}). \tag{2.18}$$

Then the SAV scheme (2.14) can be written as

$$\begin{aligned}\Big(\Delta^{-1}\frac{u^{n+1}-u^n}{\tau},v\Big) &= \varepsilon\big(\nabla u^{n+1},\nabla v\big)+\frac{1}{\varepsilon}r^n\big(g(u^n),v\big)+\frac{1}{\varepsilon}\big(g(u^n),v\big)(r^{n+1}-r^n)\\ &= \varepsilon\big(\nabla u^{n+1},\nabla v\big)+\frac{1}{\varepsilon}r^n\big(g(u^n),v\big)+\frac{1}{2\varepsilon}\big(g(u^n),v\big)\big(g(u^n),u^{n+1}-u^n\big),\end{aligned}\tag{2.19}$$

for all $v\in L_0^2(\Omega)$, which together with (2.18) gives

$$\begin{aligned}\Big(\Delta^{-1}\frac{u^{n+1}-u^n}{\tau},v\Big) =& \varepsilon\big(\nabla u^{n+1},\nabla v\big)+\frac{1}{\varepsilon}\big(f(u^n),v\big)\\ &-\frac{1}{2\varepsilon}\big(g(u^n),v\big)\sum_{i=1}^{n}D^2A(\xi_i\,;u^i-u^{i-1})\\ &+\frac{1}{2\varepsilon}\big(g(u^n),v\big)\big(g(u^n),u^{n+1}-u^n\big)\qquad\forall\ v\in L_0^2(\Omega).\end{aligned}\tag{2.20}$$

The above equation (2.20) provides an equivalent formulation of the semi-discrete SAV scheme (2.8), which will be frequently used in the subsequent error analysis.

3 Energy Decay and Consistency Analysis

In this section, we present several inequalities related to the proposed numerical method.

3.1 Assumption and Regularity

Before presenting the detailed numerical analysis, we first make some assumptions. The quartic growth of the Ginzburg-Landau energy function $F(u)=\frac{1}{4}(u^2-1)^2$ at infinity poses various technical difficulties for the analysis and approximation of CH equations. Although the CH equation does not satisfy the maximum principle, if the maximum norm of the initial condition u^0 is bounded, it has been shown in [9] that the maximum norm of the solution of the CH equation for the truncation potential $F(u)$ with quadratic growth rate at infinity is bounded. Therefore, it has been a common practice (cf. [33]) to consider the CH equations with a truncated $F(u)$.

Assumption 3.1 We assume that the potential function $F(u)$ whose derivative $f(u)=F'(u)$ satisfies the following condition:

(1) $F\in C^4(\mathbb{R})$, $F(\pm1)=0$, and $F>0$ elsewhere.
(2) $f(\pm1)=0$, $f'(\pm1)>0$, and there exists a non-negative constant L such that

$$\max_{v\in\mathbb{R}}|f(v)|\le L,\quad \max_{v\in\mathbb{R}}|f'(v)|\le L\quad\text{and}\quad\max_{v\in\mathbb{R}}|f''(v)|\le L.\tag{3.1}$$

In order to trace the dependence of the solution on the small parameter $\varepsilon>0$, we assume that the solution of (1.1) satisfies the following conditions:

Assumption 3.2 Suppose there exist positive ε-independent constants m_0 and ρ_j for $j=1,2,3$ such that the solution of (1.1) satisfies

$$\frac{1}{|\Omega|}\int_\Omega u\,\mathrm{d}x=m_0\in(-1,1),\tag{3.2a}$$

$$\operatorname*{ess\,sup}_{t\in[0,\infty]}\left\{\frac{\varepsilon}{2}\|\nabla u\|^2+\frac{1}{\varepsilon}\int_\Omega F(u)\,\mathrm{d}x\right\}+\int_0^\infty \|u_t\|_{-1}^2\,\mathrm{d}s\lesssim\varepsilon^{-\rho_1},\tag{3.2b}$$

$$\int_0^\infty \|u_t\|^2\,\mathrm{d}s\lesssim\varepsilon^{-\rho_2},\tag{3.2c}$$

$$\int_0^\infty \|\Delta^{-1}u_{tt}\|_{-1}^2\,\mathrm{d}t\lesssim\varepsilon^{-\rho_3}.\tag{3.2d}$$

Remark 3.1 (a) Note that the conditions (1) and (2) are satisfied by restricting the growth of $F(v)$ for $|v|\geq M$. More precisely, for a given $M\geq 1$, we can replace $F(v)=\frac{1}{4}(v^2-1)^2$ by a cut-off function $\hat{F}(v)\in C^4(\mathbb{R})$ as follows:

$$\hat{F}(v)=\begin{cases}((2M)^2-1)2M(v-2M)+\frac{1}{4}((2M)^2-1)^2 & \text{for } v>2M,\\ \Phi_+(v) & \text{for } v\in(M,2M],\\ \frac{1}{4}(v^2-1)^2 & \text{for } v\in[-M,M],\\ \Phi_-(v) & \text{for } v\in[-2M,-M),\\ -((2M)^2-1)2M(v+2M)+\frac{1}{4}((2M)^2-1)^2 & \text{for } v<-2M,\end{cases}\tag{3.3}$$

where $\Phi_+(v)$ and $\Phi_-(v)>0$ elsewhere between $M<|v|<2M$ and satisfy the required conditions at $|v|=M$ and $|v|=2M$, respectively. Then we replace $f(v)=(v^2-1)v$ by $\hat{F}'(v)$ which is

$$\hat{f}(v)=\hat{F}'(v)=\begin{cases}((2M)^2-1)2M & \text{for } v>2M,\\ \Phi_+'(v) & \text{for } v\in(M,2M],\\ (v^2-1)v & \text{for } v\in[-M,M],\\ \Phi_-'(v) & \text{for } v\in[-2M,-M),\\ -((2M)^2-1)2M & \text{for } v<-2M.\end{cases}\tag{3.4}$$

In simplicity, we still denote the modified function $\hat{F}$ by F. It is then obvious that there exists L such that (3.1) are satisfied with f replaced by $\hat{f}$.

(b) The transformed SAV scheme (2.20) introduced a complicated term. With the condition (2) in the Assumption 3.1, we can get

$$D^2A(v;w)\lesssim\left(\|f'(v)\|_{L^\infty}+\|f(v)\|^2\right)\|w\|^2\lesssim\|w\|^2,\tag{3.5}$$

which will be frequently used to control the difficult term in the error analysis.

(c) Assumption 3.2 can be achieved in many cases. For example, suppose that f satisfies Assumption 3.1, $\partial\Omega$ is of class $C^{2,1}$, $u^0\in H^3(\Omega)$, and there exist positive ε-independent constants σ_j for $j=1,2,3$ such that

$$E[u^0]=\frac{\varepsilon}{2}\|\nabla u^0\|^2+\frac{1}{\varepsilon}\int_\Omega F(u^0)\,\mathrm{d}x\lesssim\varepsilon^{-2\sigma_1},\tag{3.6a}$$

$$\|-\varepsilon\Delta u^0+\frac{1}{\varepsilon}f(u^0)\|\lesssim\varepsilon^{-2\sigma_2},\tag{3.6b}$$

$$\|-\varepsilon\Delta u^0+\frac{1}{\varepsilon}f(u^0)\|_{H^1}\lesssim\varepsilon^{-2\sigma_3}.\tag{3.6c}$$

Then the estimates (3.2a)–(3.2d) can be derived by standard test function techniques and satisfy:

$$\rho_1 = 2\sigma_1, \quad \rho_2 = \max\{2\sigma_1 + 2, 2\sigma_2 - 1\} \quad \text{and} \quad \rho_3 = \max\{2\sigma_1 + 4, 2\sigma_2 + 1, 2\sigma_3 - 1\}.$$

We refer to [22, 25] for their detailed proof.

3.2 Energy Decay Structure

In this subsection, we prove the following energy decay property of the numerical solution, which comprise of the first theorem of this paper.

Theorem 3.1 (energy decay) *The scheme* (2.8) *is unconditionally energy stable in the sense that*

$$E(u^{n+1}, r^{n+1}) - E(u^n, r^n) \le -\frac{1}{\tau}\|u^{n+1} - u^n\|_{-1}^2 \le 0 \quad \textit{for } n \ge 1. \tag{3.7}$$

Proof Taking the inner product of the first equation in (2.8) with $-\Delta^{-1}(u^{n+1} - u^n)$, and of the second equation with $u^{n+1} - u^n$, using (2.12) and multiplying the third equation in (2.8) by $\frac{2}{\varepsilon}r^{n+1}$, we derive that

$$\begin{aligned}\frac{1}{\tau}\|u^{n+1} - u^n\|_{-1}^2 + \frac{\varepsilon}{2}\Big(\|\nabla u^{n+1}\|^2 - \|\nabla u^n\|^2 + \|\nabla u^{n+1} - \nabla u^n\|^2\Big)\\ + \frac{1}{\varepsilon}r^{n+1}\big(g(u^n), u^{n+1} - u^n\big) = 0,\end{aligned} \tag{3.8}$$

$$\frac{1}{\varepsilon}\Big((r^{n+1})^2 - (r^n)^2 + (r^{n+1} - r^n)^2\Big) = \frac{1}{\varepsilon}r^{n+1}\big(g(u^n), u^{n+1} - u^n\big). \tag{3.9}$$

Taking the summation of the above equations, we get

$$\begin{aligned}\frac{1}{\tau}\|u^{n+1} - u^n\|_{-1}^2 + \frac{\varepsilon}{2}\Big(\|\nabla u^{n+1}\|^2 - \|\nabla u^n\|^2 + \|\nabla u^{n+1} - \nabla u^n\|^2\Big)\\ + \frac{1}{\varepsilon}\Big((r^{n+1})^2 - (r^n)^2 + (r^{n+1} - r^n)^2\Big) = 0.\end{aligned} \tag{3.10}$$

which gives (3.7). □

Remark 3.2 After summing up (3.10) from $n = 0$ to N, we get

$$\begin{aligned}\frac{\varepsilon}{2}\|\nabla u^{N+1}\|^2 + \frac{1}{\varepsilon}(r^{N+1})^2 + \frac{1}{\tau}\sum_{n=0}^{N}\|u^{n+1} - u^n\|_{-1}^2 + \frac{\varepsilon}{2}\sum_{n=0}^{N}\|\nabla u^{n+1} - \nabla u^n\|^2\\ + \frac{1}{\varepsilon}\sum_{n=0}^{N}(r^{n+1} - r^n)^2 = \frac{\varepsilon}{2}\|\nabla u^0\|^2 + \frac{1}{\varepsilon}(r^0)^2 \lesssim \varepsilon^{-\rho_1}.\end{aligned} \tag{3.11}$$

It follows from $u^{n+1} - u^n \in L_0^2(\Omega)$ and (2.3) that

$$\begin{aligned}\sum_{n=0}^{N}\|u^{n+1} - u^n\|^2 &\le \sum_{n=0}^{N}\|u^{n+1} - u^n\|_{-1}\|\nabla u^{n+1} - \nabla u^n\|\\ &\le \Big(\sum_{n=0}^{N}\|u^{n+1} - u^n\|_{-1}^2\Big)^{\frac{1}{2}}\Big(\sum_{n=0}^{N}\|\nabla u^{n+1} - \nabla u^n\|^2\Big)^{\frac{1}{2}}\\ &\lesssim \varepsilon^{-(\rho_1+\frac{1}{2})}\tau^{\frac{1}{2}}.\end{aligned} \tag{3.12}$$

3.3 Consistency

It follows from (1.1b) that $\int_\Omega \Delta u \mathrm{d}x = 0$ and $\int_\Omega \partial_t u \mathrm{d}x = 0$. Since $\partial_t u \in L_0^2(\Omega)$, we have $u(t_{n+1}) - u(t_n) \in L_0^2(\Omega)$. For any $f \in L^2(\Omega)$, let $\hat{f} = f - \int_\Omega f \mathrm{d}x$ such that $\hat{f} \in L_0^2(\Omega)$. Following the derivation of (2.11), the CH equation (1.1) can be written as

$$\Delta^{-1}\partial_t u = -\varepsilon \Delta u + \frac{1}{\varepsilon}\hat{f}(u).$$

Similar to (2.13), we use the relation

$$(\hat{f}(u), v) = (f(u), v) \qquad \forall\ v \in L_0^2(\Omega),$$

to get

$$\Big(\Delta^{-1}\partial_t u(t_{n+1}), v\Big) = \varepsilon\big(\nabla u(t_{n+1}), \nabla v\big) + \frac{1}{\varepsilon}\big(f(u(t_{n+1})), v\big) \qquad \forall\ v \in L_0^2(\Omega).$$

To derive the error estimates of the equivalent transformation (2.20) of the semi-discrete SAV scheme (2.8), we reformulate the CH equation (1.1) as the truncated form

$$\begin{aligned}&\Big(\Delta^{-1}\frac{u(t_{n+1}) - u(t_n)}{\tau}, v\Big)\\ &= \varepsilon\big(\nabla u(t_{n+1}), \nabla v\big) + \frac{1}{\varepsilon}\big(f(u(t_n)), v\big) + \big(\mathcal{R}^{n+1}, v\big) \quad \forall\ v \in L_0^2(\Omega),\end{aligned} \tag{3.13}$$

where the truncation error $\mathcal{R}^{n+1}$ is given by

$$\mathcal{R}^{n+1} := \Big[\Delta^{-1}\frac{u(t_{n+1}) - u(t_n)}{\tau} - \Delta^{-1}\partial_t u(t_{n+1})\Big] + \frac{1}{\varepsilon}\Big[f(u(t_{n+1})) - f(u(t_n))\Big]. \tag{3.14}$$

Lemma 3.1 (consistency estimate) *Suppose that assumptions 3.1 and 3.2 hold, then we have the following consistency estimate:*

$$\tau \sum_{n=0}^{N} \|\mathcal{R}^{n+1}\|_{H^{-1}}^2 \le C\varepsilon^{-\max\{\rho_2+2,\,\rho_3\}}\tau^2. \tag{3.15}$$

Proof For any $\varphi \in L_0^2(\Omega)$, there holds $\varphi_1 = -\Delta^{-1}\varphi \in L_0^2(\Omega) \cap H^1(\Omega)$ with $\partial_{\boldsymbol{n}}\varphi_1 = 0$ on $\partial\Omega$ and $\|\varphi\|_{-1} = \|\nabla\varphi_1\|$, which gives

$$\begin{aligned}\|\varphi\|_{H^{-1}} &= \sup_{v\in H^1(\Omega)} \frac{(\varphi, v)}{\|v\|_{H^1}} = \sup_{v\in H^1(\Omega)} \frac{\big(-\Delta(-\Delta^{-1})\varphi, v\big)}{\|v\|_{H^1}} = \sup_{v\in H^1(\Omega)} \frac{\big(\nabla\varphi_1, \nabla v\big)}{\|v\|_{H^1}}\\ &\le \|\nabla\varphi_1\| = \|\varphi\|_{-1}.\end{aligned} \tag{3.16}$$

Since $\partial_t u \in L_0^2(\Omega)$, it follows that $u_{tt} \in L_0^2(\Omega)$. By the definition of $\Delta^{-1} : L_0^2(\Omega) \to H^1(\Omega) \cap L_0^2(\Omega)$, we have $\Delta^{-1}u_{tt} \in L_0^2(\Omega)$. By performing standard calculations, we obtain

$$\begin{aligned}\Big\|\Delta^{-1}\frac{u(t_{n+1}) - u(t_n)}{\tau} - \Delta^{-1}\partial_t u(t_{n+1})\Big\|_{H^{-1}}^2 &= \Big\|\frac{1}{\tau}\int_{t_n}^{t_{n+1}} (s - t_n)\Delta^{-1}u_{tt}\,\mathrm{d}s\Big\|_{H^{-1}}^2\\ &\lesssim \tau\int_{t_n}^{t_{n+1}} \|\Delta^{-1}u_{tt}\|_{H^{-1}}^2\,\mathrm{d}s\\ &\le \tau\int_{t_n}^{t_{n+1}} \|\Delta^{-1}u_{tt}\|_{-1}^2\,\mathrm{d}s,\end{aligned} \tag{3.17}$$

and

$$\begin{aligned}\left\|\frac{1}{\varepsilon}f(u(t_{n+1}))-\frac{1}{\varepsilon}f(u(t_n))\right\|_{H^{-1}}^2 &= \varepsilon^{-2}\sup_{v\in H^1(\Omega)}\frac{\left(f(u(t_{n+1}))-f(u(t_n)),v\right)^2}{\|v\|_{H^1}^2}\\ &\le \varepsilon^{-2}\sup_{v\in H^1(\Omega)}\frac{\|f'(\xi^n)\|_{L^3}^2\|u(t_{n+1})-u(t_n)\|^2\|v\|_{L^6}^2}{\|v\|_{H^1}^2}\\ &\lesssim \varepsilon^{-2}\|u(t_{n+1})-u(t_n)\|^2\\ &\lesssim \varepsilon^{-2}\tau\int_{t_n}^{t_{n+1}}\|u_t\|^2\,\mathrm{d}s, \end{aligned}\tag{3.18}$$

where ξ^n is between $u(t_n)$ and $u(t_{n+1})$. Thus, we have

$$\begin{aligned}\tau\sum_{n=0}^{N}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2 &\lesssim \tau^2\int_0^T\|\Delta^{-1}u_{tt}\|_{-1}^2\,\mathrm{d}s+\varepsilon^{-2}\tau^2\int_0^T\|u_t\|^2\,\mathrm{d}s\\ &\lesssim \varepsilon^{-\rho_3}\tau^2+\varepsilon^{-(\rho_2+2)}\tau^2. \end{aligned}\tag{3.19}$$

The proof is completed. □

4 Error Estimates

In this section, we will derive the error bound of the semi-discrete scheme (2.8), in which the focus is to obtain the polynomial type dependence of the error bound on ε^{-1}. If we use the usual error estimate of the SAV numerical scheme (2.8), the error growth depends on ε^{-1} exponentially. To avoid the exponential dependence on ε^{-1} induced by using the Gronwall inequality, we need to use a spectral estimate of the linearized Cahn-Hilliard operator, which is given in [3, 13, 25].

Lemma 4.1 (spectral estimate) *Suppose that Assumption 3.1 holds. Then there exist* $0 < \varepsilon_0 << 1$ *and a positive constant* λ_0 *such that the principle eigenvalue of the linearized Cahn-Hilliard operator*

$$\mathcal{L}_{CH} := \Delta(\varepsilon\Delta-\frac{1}{\varepsilon}f'(u)I)$$

satisfies for all $t\in[0,T]$

$$\lambda_{CH}=\inf_{\substack{0\neq v\in H^1(\Omega)\\ \Delta w=v}}\frac{\varepsilon\|\nabla v\|^2+\frac{1}{\varepsilon}\left(f'(u(\cdot,t)v,v)\right)}{\|\nabla w\|^2}\ge-\lambda_0,\tag{4.1}$$

for $\varepsilon\in(0,\varepsilon_0)$*, where I denotes the identity operator and u is the solution of the Cahn-Hilliard problem (1.1).*

We will now prove the following error estimates for the semi-discrete numerical scheme, which is the main result of this paper.

Theorem 4.1 (error estimate) *We assume that assumptions 3.1 and 3.2 hold and that*

$$\tau\le\tilde{C}\varepsilon^{\beta_0}\quad with\quad \beta_0=\frac{4\alpha_0+32+4d}{4-d},\tag{4.2}$$

then the discrete solution given by (2.8) satisfies the following error estimate for $e^n = u(t_n) - u^n$*:*

$$\max_{1\le n\le m}\|e^n\|_{-1}^2+\frac{1}{2}\sum_{n=1}^{m}\|e^n-e^{n-1}\|_{-1}^2+\frac{1}{2}\varepsilon^4\sum_{n=1}^{m}\tau\|\nabla e^n\|^2+\max_{1\le n\le m}\tau\varepsilon^4\|\nabla e^m\|^2 + \frac{1}{2}\varepsilon^4\sum_{n=1}^{m}\tau\|\nabla e^n-\nabla e^{n-1}\|^2\le\kappa_0\varepsilon^{-\alpha_0}\tau^2. \tag{4.3}$$

where $\alpha_0 := \max\{\rho_1+3,\ 2\rho_2+4,\ \rho_2+6,\ \rho_3+4\}$ *and the constant* κ_0 *is independent of* τ *when* τ *is sufficiently small. The specific values of* $\tilde{C}$ *and* κ_0 *will be given in the proof.*

Proof We use the mathematical induction as follows. The proof is split into four steps. The first step gives the error estimate for the first step $t = t^1$. Steps two and three use the spectral estimate (4.1) to avoid exponential blow-up in ε^{-1} of the error constants. In the last step, an inductive argument is used to conclude the proof.

Step 1: Estimation of $\|e^1\|_{-1}^2+\tau\varepsilon\|\nabla e^1\|^2$. For $n=0$ in (2.14), we have

$$\begin{cases}\Big(\Delta^{-1}\dfrac{u^1-u^0}{\tau},v\Big)=\varepsilon\big(\nabla u^1,\nabla v\big)+\dfrac{1}{\varepsilon}r^1\big(g(u^0),v\big) & \forall\ v\in L_0^2(\Omega),\\ r^1=r^0+\dfrac{1}{2}\big(g(u^0),u^1-u^0\big),\end{cases} \tag{4.4}$$

After plugging the second equation into the first equation, we get

$$\Big(\Delta^{-1}\frac{u^1-u^0}{\tau},v\Big)=\varepsilon\big(\nabla u^1,\nabla v\big)+\frac{1}{\varepsilon}\big(f(u^0),v\big)+\frac{1}{2\varepsilon}\big(g(u^0),v\big)\big(g(u^0),u^1-u^0\big), \tag{4.5}$$

for all $v\in L_0^2(\Omega)$. Subtracting (4.5) from (3.13), we get the corresponding error equation

$$\Big(\Delta^{-1}\frac{e^1}{\tau},v\Big)=\varepsilon\big(\nabla e^1,\nabla v\big)+\big(\mathcal{R}^1,v\big)-\frac{1}{2\varepsilon}\big(g(u^0),v\big)\big(g(u^0),u^1-u^0\big)\qquad\forall\ v\in L_0^2(\Omega). \tag{4.6}$$

Testing the above equation with $v=e^1\in L_0^2(\Omega)$, we have

$$\|e^1\|_{-1}^2+\tau\varepsilon\|\nabla e^1\|^2=-\tau(\mathcal{R}^1,e^1)+\frac{\tau}{2\varepsilon}\big(g(u^0),u^1-u^0\big)\big(g(u^0),e^1\big). \tag{4.7}$$

Using Poincaré's inequality for $e^1\in L_0^2(\Omega)$ and Lemma 3.1, we have

$$\begin{aligned}\tau(\mathcal{R}^1,e^1)&\le\frac{\varepsilon\tau}{4}\|\nabla e^1\|^2+C\varepsilon^{-1}\tau\|\mathcal{R}^1\|_{H^{-1}}^2\\&\le\frac{\varepsilon\tau}{4}\|\nabla e^1\|^2+C\varepsilon^{-\max\{\rho_2+3,\,\rho_3+1\}}\tau^2.\end{aligned} \tag{4.8}$$

From (3.12), we have

$$\begin{aligned}\frac{\tau}{2\varepsilon}\big(g(u^0),u^1-u^0\big)\big(g(u^0),e^1\big)&\le C\tau\varepsilon^{-1}\|g(u^0)\|^2\|u^1-u^0\|\|e^1\|\\&\le C\tau\varepsilon^{-1}\|u^1-u^0\|\|e^1\|_{-1}^{\frac{1}{2}}\|\nabla e^1\|^{\frac{1}{2}}\\&\le\frac{1}{2}\tau^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\|e^1\|_{-1}\|\nabla e^1\|+C\tau^{\frac{3}{2}}\varepsilon^{-\frac{5}{2}}\|u^1-u^0\|^2\\&\le\frac{1}{4}\|e^1\|_{-1}^2+\frac{\varepsilon\tau}{4}\|\nabla e^1\|^2+C\varepsilon^{-(\rho_1+3)}\tau^2\end{aligned} \tag{4.9}$$

Thus, there exists a positive constant κ_0' such that

$$\|e^1\|_{-1}^2 + \tau\varepsilon\|\nabla e^1\|^2 \le \kappa_0'\varepsilon^{-\max\{\rho_1+3,\,\rho_2+3,\,\rho_3+1\}}\tau^2. \tag{4.10}$$

We use the mathematical induction as follows. For $m = 1$, (4.3) holds from (4.10). We suppose (4.3) holds for $m = 1, 2, \ldots, N$, and show that (4.3) is also valid for $m = N + 1$.

Subtracting (2.20) from (3.13), by denoting $e^n = u(t_n) - u^n$, we get the error equation

$$\begin{aligned}\Big(\Delta^{-1}\frac{e^{n+1}-e^n}{\tau}, v\Big) =& \varepsilon\big(\nabla e^{n+1}, \nabla v\big) + \frac{1}{\varepsilon}\Big(f(u(t_n)) - f(u^n), v\Big)\\ &+ \frac{1}{2\varepsilon}\big(g(u^n), v\big)\sum_{i=1}^{n} D^2A(\xi_i\,; u^i - u^{i-1})\\ &- \frac{1}{2\varepsilon}\big(g(u^n), v\big)\big(g(u^n), u^{n+1}-u^n\big) + \big(\mathcal{R}^{n+1}, v\big)\quad \forall\ v \in L_0^2(\Omega).\end{aligned} \tag{4.11}$$

Step 2: Estimation of $\|\nabla e^{N+1}\|^2 + \sum_{n=0}^{N}\|\nabla e^{n+1} - \nabla e^n\|^2$. By testing (4.11) with $v = e^{n+1} - e^n \in L_0^2(\Omega)$, we have

$$\begin{aligned}&\frac{1}{2}\|\nabla e^{n+1}\|^2 - \frac{1}{2}\|\nabla e^n\|^2 + \frac{1}{2}\|\nabla e^{n+1} - \nabla e^n\|^2 + \frac{\tau}{\varepsilon}\Big\|\frac{e^{n+1}-e^n}{\tau}\Big\|_{-1}^2\\ &\quad= -\frac{1}{\varepsilon^2}\Big(f(u(t_n)) - f(u^n), e^{n+1}-e^n\Big)\\ &\qquad- \frac{1}{2\varepsilon^2}\big(g(u^n), e^{n+1}-e^n\big)\sum_{i=1}^{n} D^2A(\xi_i\,; u^i - u^{i-1})\\ &\qquad+ \frac{1}{2\varepsilon^2}\big(g(u^n), e^{n+1}-e^n\big)\big(g(u^n), u^{n+1}-u^n\big) - \frac{1}{\varepsilon}\big(\mathcal{R}^{n+1}, e^{n+1}-e^n\big)\\ &\quad=: K_1 + K_2 + K_3 + K_4.\end{aligned} \tag{4.12}$$

It follows from (2.3) that

$$\begin{aligned}K_1 &\le \varepsilon^{-2}\|f'(w^n)\|_{L^\infty}\|e^n\|\|e^{n+1}-e^n\|\\ &\le C\varepsilon^{-2}\|\nabla e^n\|^{\frac{1}{2}}\|e^n\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|^{\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\\ &\le C\frac{1}{\gamma_0}\varepsilon^{-2}\tau^{\frac{1}{2}}\|\nabla e^n\|\|e^n\|_{-1} + \gamma_0\varepsilon^{-2}\tau^{-\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|\|e^{n+1}-e^n\|_{-1}\\ &\le C\tau\|\nabla e^n\|^2 + C\varepsilon^{-4}\|e^n\|_{-1}^2 + \frac{1}{2}\gamma_0\tau^{-1}\varepsilon^{-4}\|e^{n+1}-e^n\|_{-1}^2 + \frac{1}{2}\gamma_0\|\nabla e^{n+1}-\nabla e^n\|^2\end{aligned} \tag{4.13}$$

where $w_n = \theta u^n + (1-\theta)u(t_n)$ with $\theta \in (0, 1)$, and γ_0 is a sufficiently small constant such that the two terms involving $\|e^{n+1} - e^n\|_{-1}^2$ and $\|\nabla e^{n+1} - \nabla e^n\|^2$ can be absorbed by the left hand side in the Step 4.

By using $D^2A(v; w) \lesssim \big(\|f'(v)\|_{L^\infty} + \|f(v)\|^2\big)\|w\|^2$, the term K_2 can be bounded as

$$K_2 \lesssim \frac{1}{2\varepsilon^2}\big|\big(g(u^n), e^{n+1}-e^n\big)\big|\sum_{i=1}^{n}\|u^i - u^{i-1}\|^2$$

$$\begin{aligned}
&\lesssim \frac{1}{\varepsilon^2}\big|\big(g(u^n), e^{n+1}-e^n\big)\big| \sum_{i=1}^{n} \|e^i - e^{i-1}\|^2 \\
&\quad + \frac{1}{\varepsilon^2}\big|\big(g(u^n), e^{n+1}-e^n\big)\big| \sum_{i=1}^{n} \|u(t_i) - u(t_{i-1})\|^2 \\
&=: K_{21} + K_{22}. \qquad (4.14)
\end{aligned}$$

From the assumptions of induction, we have

$$\begin{aligned}
\big(g(u^n), e^{n+1}-e^n\big) &= \frac{1}{\sqrt{\int_\Omega F(u^n)\,\mathrm{d}x + c_0}}\big(f(u^n), e^{n+1}-e^n\big) \\
&\le C\|f(u^n)\|\|e^{n+1}-e^n\| \\
&\le C\|\nabla e^{n+1} - \nabla e^n\|^{\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}, \qquad (4.15)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n}\|e^i - e^{i-1}\|^2 &\lesssim \Big(\sum_{i=1}^{n}\|e^i - e^{i-1}\|_{-1}^2\Big)^{\frac{1}{2}}\Big(\sum_{i=1}^{n}\|\nabla e^i - \nabla e^{i-1}\|^2\Big)^{\frac{1}{2}} \\
&\lesssim \kappa_0\varepsilon^{-(\alpha_0+2)}\tau^{\frac{3}{2}}, \qquad (4.16)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n}\|u(t_i) - u(t_{i-1})\|^2 &\lesssim \tau\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\|u_t\|^2\,\mathrm{d}s \lesssim \tau\int_0^{t_n}\|u_t\|^2\,\mathrm{d}s \\
&\lesssim \varepsilon^{-\rho_2}\tau. \qquad (4.17)
\end{aligned}$$

Using the estimates above, we have

$$\begin{aligned}
K_{21} &\le C\varepsilon^{-2}\|\nabla e^{n+1} - \nabla e^n\|^{\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\kappa_0\varepsilon^{-(\alpha_0+2)}\tau^{\frac{3}{2}} \\
&\le \gamma_0\varepsilon^{-2}\tau^{-\frac{1}{2}}\|\nabla e^{n+1} - \nabla e^n\|\|e^{n+1}-e^n\|_{-1} + C\frac{1}{\gamma_0}\kappa_0^2\varepsilon^{-(2\alpha_0+6)}\tau^{\frac{7}{2}} \\
&\le \frac{1}{2}\gamma_0\tau^{-1}\varepsilon^{-4}\|e^{n+1}-e^n\|_{-1}^2 + \frac{1}{2}\gamma_0\|\nabla e^{n+1} - \nabla e^n\|^2 + C\kappa_0^2\varepsilon^{-(2\alpha_0+6)}\tau^{\frac{7}{2}}. \qquad (4.18)
\end{aligned}$$

and

$$\begin{aligned}
K_{22} &\le C\varepsilon^{-2}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1} - \nabla e^n\|^{\frac{1}{2}}\varepsilon^{-\rho_2}\tau \\
&\le \gamma_0\varepsilon^{-2}\tau^{-\frac{1}{2}}\|\nabla e^{n+1} - \nabla e^n\|\|e^{n+1}-e^n\|_{-1} + C\frac{1}{\gamma_0}\varepsilon^{-(2\rho_2+2)}\tau^{\frac{5}{2}} \\
&\le \frac{1}{2}\gamma_0\tau^{-1}\varepsilon^{-4}\|e^{n+1}-e^n\|_{-1}^2 + \frac{1}{2}\gamma_0\|\nabla e^{n+1} - \nabla e^n\|^2 + C\varepsilon^{-(2\rho_2+2)}\tau^{\frac{5}{2}}. \qquad (4.19)
\end{aligned}$$

In addition, we divided the term K_3 into two parts as

$$\begin{aligned}
K_3 &= \frac{1}{2\varepsilon^2}\big(g(u^n), e^{n+1}-e^n\big)\big(g(u^n), u^{n+1}-u^n\big) \\
&= -\frac{1}{2\varepsilon^2}\big(g(u^n), e^{n+1}-e^n\big)\big(g(u^n), e^{n+1}-e^n\big) \\
&\quad + \frac{1}{2\varepsilon^2}\big(g(u^n), e^{n+1}-e^n\big)\big(g(u^n), u(t_{n+1}) - u(t_n)\big) \\
&=: K_{31} + K_{32}, \qquad (4.20)
\end{aligned}$$

where

$$K_{31} \le C\varepsilon^{-2}\|g(u^n)\|^2\|e^{n+1}-e^n\|_{-1}\|\nabla e^{n+1}-\nabla e^n\| \le C\frac{1}{\gamma_0}\varepsilon^{-4}\|e^{n+1}-e^n\|_{-1}^2+\frac{1}{2}\gamma_0\|\nabla e^{n+1}-\nabla e^n\|^2, \tag{4.21}$$

and

$$\begin{aligned} K_{32} &\le C\varepsilon^{-2}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|^{\frac{1}{2}}\|u(t_{n+1})-u(t_n)\| \\ &\le \gamma_0\varepsilon^{-2}\tau^{-\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}\|\nabla e^{n+1}-\nabla e^n\|+C\gamma_0^{-1}\varepsilon^{-2}\tau^{\frac{3}{2}}\int_{t_n}^{t_{n+1}}\|u_t\|^2\,\mathrm{d}s \\ &\le \frac{1}{2}\gamma_0\tau^{-1}\varepsilon^{-4}\|e^{n+1}-e^n\|_{-1}^2+\frac{1}{2}\gamma_0\|\nabla e^{n+1}-\nabla e^n\|^2+C\varepsilon^{-2}\tau^{\frac{3}{2}}\int_{t_n}^{t_{n+1}}\|u_t\|^2\,\mathrm{d}s, \end{aligned} \tag{4.22}$$

By using Poincaré's inequality for $e^{n+1}-e^n \in L_0^2(\Omega)$, the last term on the right hand side of (4.12) can be bounded by

$$K_4=\frac{1}{\varepsilon}\big(\mathcal{R}^{n+1},e^{n+1}-e^n\big)\le\frac{1}{2}\gamma_0\|\nabla e^{n+1}-\nabla e^n\|^2+C\frac{1}{\gamma_0}\varepsilon^{-2}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2. \tag{4.23}$$

Combining these estimates (4.13)–(4.23) together with (4.12), and taking the summation for $n=0$ to N, we have

$$\begin{aligned} \|\nabla e^{N+1}\|^2+(1-6\gamma_0)\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2 &\le C\tau\sum_{n=0}^{N}\|\nabla e^n\|^2+C\varepsilon^{-4}\sum_{n=0}^{N}\|e^n\|_{-1}^2 \\ &+(C+4\gamma_0\tau^{-1})\varepsilon^{-4}\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2+C\varepsilon^{-2}\sum_{n=0}^{N}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2 \\ &+C\kappa_0^2\varepsilon^{-(2\alpha_0+6)}\tau^{\frac{5}{2}}+C\varepsilon^{-(2\rho_2+2)}\tau^{\frac{3}{2}}+C\varepsilon^{-2}\tau^{\frac{3}{2}}\int_0^T\|u_t\|^2\,\mathrm{d}s. \end{aligned} \tag{4.24}$$

Step 3: Estimation of $\|e^{N+1}\|_{-1}^2+\sum_{n=0}^{N}\|e^n-e^{n-1}\|_{-1}^2$. By testing (4.11) with $v=e^{n+1}\in L_0^2(\Omega)$, we get

$$\begin{aligned} &\frac{1}{2\tau}\Big(\|e^{n+1}\|_{-1}^2-\|e^n\|_{-1}^2+\|e^{n+1}-e^n\|_{-1}^2\Big)+\varepsilon\|\nabla e^{n+1}\|^2+\frac{1}{\varepsilon}\Big(f(u(t_n))-f(u^n),e^{n+1}\Big) \\ &\quad+\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\sum_{i=1}^{n}D^2A(\xi_i;u^i-u^{i-1}) \\ &\quad-\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\big(g(u^n),u^{n+1}-u^n\big)+\big(\mathcal{R}^{n+1},e^{n+1}\big)=0. \end{aligned} \tag{4.25}$$

We denote $w^n=\theta u^n+(1-\theta)u(t_n)$ with $\theta\in(0,1)$, and using the Taylor expansion to get

$$\frac{1}{\varepsilon}\Big(f(u(t_n))-f(u^n),e^{n+1}\Big)=\frac{1}{\varepsilon}\Big(f'(u(t_n))e^n-\frac{1}{2}f''(w^n)(e^n)^2,e^{n+1}\Big)$$
$$=\frac{1}{\varepsilon}\Big(f'(u(t_n))e^{n+1},e^{n+1}\Big)-\frac{1}{\varepsilon}\Big(f'(u(t_n))(e^{n+1}-e^n),e^{n+1}\Big)$$
$$-\frac{1}{2\varepsilon}\Big(f''(w^n)(e^n)^2,e^{n+1}\Big).$$

Then, (4.25) becomes

$$\frac{1}{2\tau}\Big(\|e^{n+1}\|_{-1}^2-\|e^n\|_{-1}^2+\|e^{n+1}-e^n\|_{-1}^2\Big)+\varepsilon\|\nabla e^{n+1}\|^2+\frac{1}{\varepsilon}\Big(f'(u(t_n))e^{n+1},e^{n+1}\Big)$$
$$=\frac{1}{\varepsilon}\Big(f'(u(t_n))(e^{n+1}-e^n),e^{n+1}\Big)+\frac{1}{2\varepsilon}\Big(f''(w^n)(e^n)^2,e^{n+1}\Big)$$
$$-\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\sum_{i=1}^{n}D^2A(\xi_i;u^i-u^{i-1})$$
$$+\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\big(g(u^n),u^{n+1}-u^n\big)-\big(\mathcal{R}^{n+1},e^{n+1}\big)$$
$$=:T_1+T_2+T_3+T_4+T_5. \tag{4.26}$$

Using Lemma 4.1, there holds

$$\varepsilon\|\nabla e^{n+1}\|^2+\frac{1}{\varepsilon}\Big(f'(u(t_n))e^{n+1},e^{n+1}\Big)\geq-\lambda_0\|e^{n+1}\|_{-1}^2. \tag{4.27}$$

If the entire $\varepsilon\|\nabla e^{n+1}\|^2$ term is used to control the the term $\varepsilon^{-1}\big(f'(u(t_n))e^{n+1},e^{n+1}\big)$, we will not be able to control the $\|\nabla e^{n+1}\|^2$ terms in T_j, $j=1,\dots,5$. So we apply (4.27) with a scaling factor $(1-\eta)$ close to but smaller than 1, to get

$$-(1-\eta)\frac{1}{\varepsilon}\Big(f'(u(t_n))e^{n+1},e^{n+1}\Big)\leq(1-\eta)\lambda_0\|e^{n+1}\|_{-1}^2+(1-\eta)\varepsilon\|\nabla e^{n+1}\|^2. \tag{4.28}$$

On the other hand,

$$-\frac{\eta}{\varepsilon}\Big(f'(u(t_n))e^{n+1},e^{n+1}\Big)\leq\frac{C\eta}{\varepsilon}\|e^{n+1}\|^2\leq\frac{C\eta}{\varepsilon^2\eta_1}\|e^{n+1}\|_{-1}^2+\frac{\eta\eta_1}{4}\|\nabla e^{n+1}\|^2. \tag{4.29}$$

The first term T_1 on the right-hand side of (4.26) can be bounded by

$$\begin{aligned}T_1&\leq\varepsilon^{-1}\|f'(u(t_n))\|_{L^\infty}\|e^{n+1}-e^n\|\|e^{n+1}\|\\&\leq C\varepsilon^{-1}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|^{\frac{1}{2}}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}\|^{\frac{1}{2}}\\&\leq\tau^{-\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}\|e^{n+1}\|_{-1}+C\varepsilon^{-2}\tau^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|\|\nabla e^{n+1}\|\\&\leq\frac{1}{2}\gamma_0\tau^{-1}\|e^{n+1}-e^n\|_{-1}^2+C\frac{1}{\gamma_0}\|e^{n+1}\|_{-1}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2+T_1^*.\end{aligned} \tag{4.30}$$

where $T_1^*:=C\varepsilon^{-(\eta_2+4)}\tau\|\nabla e^{n+1}-\nabla e^n\|^2$ and γ_0 is sufficiently small. To control the last term T_1^* on the right-hand side of (4.30), we assume that $\tau\leq\tilde{C}_1\varepsilon^{\eta_2+8}$ to get

$$\begin{aligned}\tau\sum_{n=0}^{N}T_1^* &:= C\varepsilon^{-(\eta_2+4)}\tau^2\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2\\ &= C\varepsilon^{-(\eta_2+8)}\tau\Big(\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2\Big)\\ &\le \frac{1}{16}\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2.\end{aligned}$$

Using the Sobolev interpolation inequality, we have for $v\in L_0^2(\Omega)\cap H^1(\Omega)$

$$\|v\|_{L^4}\lesssim\|\nabla v\|^{\frac{d}{4}}\|v\|^{1-\frac{d}{4}}\lesssim\|\nabla v\|^{\frac{d}{4}}\|v\|_{-1}^{\frac{1}{2}-\frac{d}{8}}\|\nabla v\|^{\frac{1}{2}-\frac{d}{8}}=\|\nabla v\|^{\frac{1}{2}+\frac{d}{8}}\|v\|_{-1}^{\frac{1}{2}-\frac{d}{8}},\tag{4.31}$$

which together with the assumptions of induction yields

$$\begin{aligned}T_2&\lesssim\varepsilon^{-1}\|e^n\|_{L^4}^2\|e^{n+1}\|\\ &\le C\varepsilon^{-1}\|\nabla e^n\|^{1+\frac{d}{4}}\|e^n\|_{-1}^{1-\frac{d}{4}}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}\|^{\frac{1}{2}}\\ &\le C\kappa_0\varepsilon^{-(\alpha_0+3+\frac{d}{2})}\tau^{\frac{3}{2}-\frac{d}{8}}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}\|^{\frac{1}{2}}\\ &\le C\varepsilon^{\frac{\eta_2}{2}}\|\nabla e^{n+1}\|\|e^{n+1}\|_{-1}+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+d+\frac{\eta_2}{2})}\tau^{3-\frac{d}{4}}\\ &\le C\|e^{n+1}\|_{-1}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+d+\frac{\eta_2}{2})}\tau^{3-\frac{d}{4}}.\end{aligned}\tag{4.32}$$

It follows from (4.16)–(4.17) that

$$\begin{aligned}T_3&\lesssim\frac{1}{2\varepsilon}\big|\big(g(u^n),e^{n+1}\big)\big|\sum_{i=1}^{n}\|u^i-u^{i-1}\|^2\\ &\le\frac{1}{\varepsilon}\big|\big(g(u^n),e^{n+1}\big)\big|\sum_{i=1}^{n}\|e^i-e^{i-1}\|^2+\frac{1}{\varepsilon}\big|\big(g(u^n),e^{n+1}\big)\big|\sum_{i=1}^{n}\|u(t_i)-u(t_{i-1})\|^2\\ &\le C\varepsilon^{-1}\|g(u^n)\|\|e^{n+1}\|\Big(\kappa_0\varepsilon^{-(\alpha_0+2)}\tau^{\frac{3}{2}}+\varepsilon^{-\rho_2}\tau\Big)\\ &\le C\varepsilon^{-1}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}\|^{\frac{1}{2}}\Big(\kappa_0\varepsilon^{-(\alpha_0+2)}\tau^{\frac{3}{2}}+\varepsilon^{-\rho_2}\tau\Big)\\ &\le\varepsilon^{\frac{\eta_2}{2}}\|e^{n+1}\|_{-1}\|\nabla e^{n+1}\|+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+\frac{\eta_2}{2})}\tau^3+C\varepsilon^{-(2\rho_2+2+\frac{\eta_2}{2})}\tau^2\\ &\le C\|e^{n+1}\|_{-1}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+\frac{\eta_2}{2})}\tau^3+C\varepsilon^{-(2\rho_2+2+\frac{\eta_2}{2})}\tau^2.\end{aligned}\tag{4.33}$$

In order to estimate the term T_4, we divided it into two parts as

$$\begin{aligned}T_4=&-\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\big(g(u^n),e^{n+1}-e^n\big)\\ &+\frac{1}{2\varepsilon}\big(g(u^n),e^{n+1}\big)\big(g(u^n),u(t_{n+1})-u(t_n)\big)\\ =:&\,T_{41}+T_{42}.\end{aligned}\tag{4.34}$$

Similarly to the estimate of T_1, the term T_{41} can be bounded by

$$\begin{aligned} T_{41} &\le C\varepsilon^{-1}\|g(u^n)\|^2\|e^{n+1}\|\|e^{n+1}-e^n\| \\ &\le C\varepsilon^{-1}\|e^{n+1}-e^n\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|^{\frac{1}{2}}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|\nabla e^{n+1}\|^{\frac{1}{2}} \\ &\le \tau^{-\frac{1}{2}}\|e^{n+1}-e^n\|_{-1}\|e^{n+1}\|_{-1}+C\varepsilon^{-2}\tau^{\frac{1}{2}}\|\nabla e^{n+1}-\nabla e^n\|\|\nabla e^{n+1}\| \\ &\le \frac{1}{2}\gamma_0\tau^{-1}\|e^{n+1}-e^n\|_{-1}^2+C\frac{1}{\gamma_0}\|e^{n+1}\|_{-1}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2+T_1^*, \end{aligned} \tag{4.35}$$

which together with the condition $\tau \le \tilde{C}_1\varepsilon^{\eta_2+8}$ yields

$$\tau\sum_{n=0}^{N}T_1^* = C\varepsilon^{-(\eta_2+4)}\tau^2\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2 \le \frac{1}{16}\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2.$$

In addition, we obtain

$$\begin{aligned} T_{42} &\le C\varepsilon^{-1}\|\nabla e^{n+1}\|^{\frac{1}{2}}\|e^{n+1}\|_{-1}^{\frac{1}{2}}\|u(t_{n+1})-u(t_n)\| \\ &\le \varepsilon^{\frac{\eta_2}{2}}\|\nabla e^{n+1}\|\|e^{n+1}\|_{-1}+C\varepsilon^{-(2+\frac{\eta_2}{2})}\tau\int_{t_n}^{t_{n+1}}\|u_t\|^2\,\mathrm{d}s \\ &\le C\|e^{n+1}\|_{-1}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2+C\varepsilon^{-(2+\frac{\eta_2}{2})}\tau\int_{t_n}^{t_{n+1}}\|u_t\|^2\,\mathrm{d}s. \end{aligned} \tag{4.36}$$

For T_5, using Cauchy-Schwartz inequality and Poincaré's inequality for $e^{n+1}\in L_0^2(\Omega)$, we have

$$T_5 \le \|\mathcal{R}^{n+1}\|_{H^{-1}}\|e^{n+1}\|_{H^1} \le 8\varepsilon^{-\eta_2}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2+\frac{1}{16}\varepsilon^{\eta_2}\|\nabla e^{n+1}\|^2. \tag{4.37}$$

Combining the estimates (4.30)–(4.37) together with (4.26), taking the summation for $n=0$ to N, and assuming that $\tau \le \tilde{C}_1\varepsilon^{\eta_2+8}$, we have

$$\begin{aligned} &\frac{1}{2}\|e^{N+1}\|_{-1}^2+\frac{1}{2}(1-2\gamma_0)\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2+\varepsilon\tau\sum_{n=0}^{N}\|\nabla e^{n+1}\|^2 \\ &\quad\le\Big[C+(1-\eta)\lambda_0+\frac{C\eta}{\varepsilon^2\eta_1}\Big]\tau\sum_{n=0}^{N}\|e^{n+1}\|_{-1}^2+8\varepsilon^{-\eta_2}\tau\sum_{n=0}^{N}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2 \\ &\qquad+\Big[(1-\eta)\varepsilon+\frac{\eta\eta_1}{4}+\frac{\varepsilon^{\eta_2}}{2}\Big]\tau\sum_{n=0}^{N}\|\nabla e^{n+1}\|^2+\frac{1}{8}\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2 \\ &\qquad+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+d+\frac{\eta_2}{2})}\tau^{3-\frac{d}{4}}+C\kappa_0^2\varepsilon^{-(2\alpha_0+6+\frac{\eta_2}{2})}\tau^3 \\ &\qquad+C\varepsilon^{-(2\rho_2+2+\frac{\eta_2}{2})}\tau^2+C\varepsilon^{-(2+\frac{\eta_2}{2})}\tau^2\int_0^T\|u_t\|^2\,\mathrm{d}s. \end{aligned} \tag{4.38}$$

By taking $\eta=\varepsilon^3$, $\eta_1=\varepsilon$ and $\eta_2=4$, we have

$$(1-\eta)\varepsilon+\frac{\eta\eta_1}{4}+\frac{\varepsilon^{\eta_2}}{2} = \varepsilon-\frac{1}{4}\varepsilon^4, \tag{4.39}$$

which together with (3.2c) gives

$$\begin{aligned}
&\|e^{N+1}\|_{-1}^2+(1-2\gamma_0)\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2+\frac{1}{2}\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}\|^2\\
&\quad\le C\tau\sum_{n=0}^{N}\|e^{n+1}\|_{-1}^2+\frac{1}{4}\tau\varepsilon^4\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2+C\varepsilon^{-4}\tau\sum_{n=0}^{N}\|\mathcal{R}^{n+1}\|_{H^{-1}}^2\\
&\qquad+C\kappa_0^2\varepsilon^{-(2\alpha_0+8+d)}\tau^{3-\frac{d}{4}}+C\kappa_0^2\varepsilon^{-(2\alpha_0+8)}\tau^3\\
&\qquad+C\varepsilon^{-(2\rho_2+4)}\tau^2+C\varepsilon^{-(\rho_2+4)}\tau^2.
\end{aligned}\tag{4.40}$$

Step 4: Completion of the proof. We now conclude the proof by the following induction argument which is based on the results from Step 1 to Step 3. By multiplying $\tau\varepsilon^4$ on both sides of (4.24), combining the estimate (4.40), and together with Lemma 3.1, we obtain

$$\begin{aligned}
&\|e^{N+1}\|_{-1}^2+\tau\varepsilon^4\|\nabla e^{N+1}\|^2+(1-C\tau-6\gamma_0)\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2\\
&\qquad+(\frac{3}{4}-6\gamma_0)\tau\varepsilon^4\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2+\frac{1}{2}\varepsilon^4\tau\sum_{n=0}^{N}\|\nabla e^{n+1}\|^2\\
&\quad\le C_0\tau\sum_{n=0}^{N}\Big(\|e^n\|_{-1}^2+\tau\varepsilon^4\|\nabla e^n\|^2\Big)+C_1\kappa_0^2\varepsilon^{-(2\alpha_0+8+d)}\tau^{3-\frac{d}{4}}\\
&\qquad+C_2\kappa_0^2\varepsilon^{-(2\alpha_0+8)}\tau^3+C_3\kappa_0^2\varepsilon^{-(2\alpha_0+2)}\tau^{\frac{7}{2}}+C_4\varepsilon^{-\max\{\rho_1+3,\,2\rho_2+4,\,\rho_2+6,\,\rho_3+4\}}\tau^2.
\end{aligned}\tag{4.41}$$

in which the term $\kappa_0'\varepsilon^{-\max\{\rho_1+3,\,\rho_2+3,\,\rho_3+1\}}\tau^2$ is absorbed in $C_4\varepsilon^{-\max\{\rho_1+3,\,2\rho_2+4,\,\rho_2+6,\,\rho_3+4\}}\tau^2$.

Suppose that for sufficiently small constant γ_0 satisfying $\frac{3}{4}-6\gamma_0\ge\frac{1}{2}$ and sufficiently small τ satisfying

$$\tau\le\Big(\frac{C_4}{C_1}\kappa_0^{-2}\Big)^{\frac{4}{4-d}}\varepsilon^{\frac{4\alpha_0+32+4d}{4-d}},\quad \tau\le\Big(\frac{C_4}{C_2}\kappa_0^{-2}\Big)\varepsilon^{\alpha_0+8},\quad \tau\le\Big(\frac{C_4}{C_3}\kappa_0^{-2}\Big)^{\frac{2}{3}}\varepsilon^{\frac{(2\alpha_0+4)}{3}},$$

then, by denoting $\alpha_0:=\max\{\rho_1+3,\ 2\rho_2+4,\ \rho_2+6,\ \rho_3+4\}$, we derive

$$\begin{aligned}
&\|e^{N+1}\|_{-1}^2+\tau\varepsilon^4\|\nabla e^{N+1}\|^2+\frac{1}{2}\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2+\frac{1}{2}\varepsilon^4\sum_{n=0}^{N}\tau\|\nabla e^{n+1}\|^2\\
&\qquad+\frac{1}{2}\tau\varepsilon^4\sum_{n=0}^{N}\|\nabla e^{n+1}-\nabla e^n\|^2\le C_0\tau\sum_{n=0}^{N}\Big(\|e^n\|_{-1}^2+\tau\varepsilon^4\|\nabla e^n\|^2\Big)+4C_4\varepsilon^{-\alpha_0}\tau^2.
\end{aligned}\tag{4.42}$$

We denote $\kappa_0:=4C_4e^{(C_0T)}$ and use the Gronwall's inequality to get

$$\begin{aligned}
&\|e^{N+1}\|_{-1}^2+\frac{1}{2}\sum_{n=0}^{N}\|e^{n+1}-e^n\|_{-1}^2+\frac{1}{2}\varepsilon^4\sum_{n=0}^{N}\tau\|\nabla e^{n+1}\|^2+\tau\varepsilon^4\|\nabla e^{N+1}\|^2\\
&\qquad+\frac{1}{2}\tau\varepsilon^4\sum_{n=0}^{m}\|\nabla e^{n+1}-\nabla e^n\|^2\le 4C_4e^{(C_0T)}\varepsilon^{-\alpha_0}\tau^2=\kappa_0\varepsilon^{-\alpha_0}\tau^2,
\end{aligned}\tag{4.43}$$

where κ_0 is independent of τ when τ is sufficiently small. The induction is completed.

In the above proof, we have used these conditions:

$$\tau \le \tilde{C}_1 \varepsilon^{12}, \quad \tau \le \tilde{C}_2 \varepsilon^{\frac{4\alpha_0+32+4d}{4-d}}, \quad \tau \le \tilde{C}_3 \varepsilon^{\alpha_0+8}, \quad \tau \le \tilde{C}_4 \varepsilon^{\frac{(2\alpha_0+4)}{3}}, \tag{4.44}$$

where $\tilde{C}_2 = \left(\frac{C_4}{C_1}\kappa_0^{-2}\right)^{\frac{4}{4-d}}$, $\tilde{C}_3 = \left(\frac{C_4}{C_2}\kappa_0^{-2}\right)$, and $\tilde{C}_4 = \left(\frac{C_4}{C_3}\kappa_0^{-2}\right)^{\frac{2}{3}}$. By denoting

$$\beta_0 = \frac{4\alpha_0 + 32 + 4d}{4-d} \quad \text{and} \quad \tilde{C} = \min\{\tilde{C}_1, \tilde{C}_2, \tilde{C}_3, \tilde{C}_4\}, \tag{4.45}$$

we specify the final condition on τ, that is, $\tau \le \tilde{C}\varepsilon^{\beta_0}$. □

5 Numerical Experiments

In this section, we present a two-dimensional numerical test to validate the theoretical results on the energy decay properties proved in Theorem 3.1, as well as the convergence rates of the proposed method given in Theorem 4.1. All the computations are performed using the software package NGSolve (https://ngsolve.org).

We solve the Cahn-Hilliard equation (1.1) on the two-dimensional square $\Omega = [0, 1] \times [0, 1]$ under Neumann boundary conditions by using the proposed scheme (2.8) with the following initial condition

$$\begin{aligned} u_0(x, y) = &\tanh\left(((x-0.65)^2 + (y-0.5)^2 - 0.1^2)/\varepsilon\right) \\ &\times \tanh\left(((x-0.35)^2 + (y-0.5)^2 - 0.125^2)/\varepsilon\right), \end{aligned} \tag{5.1}$$

where $\tanh(x) := (e^x - e^{-x})/(e^x + e^{-x})$. This type of initial condition is also adopted in [21, 26], where the set of the zero-level of the initial function $u_0(x, y)$ encloses two circles of radius 0.1 and 0.125, respectively.

To obtain a C^4 potential function $F(v)$ that satisfies the assumption 3.1, we modify the common double-well potential $F(v) = \frac{1}{4}(v^2-1)^2$ by setting $M = 2$ in (3.3) to get a cut-off function $\hat{F}(v) \in C^4(\mathbb{R})$. Correspondingly, the ninth-order polynomials $\Phi_+(v)$ and $\Phi_-(v)$ in (3.3) are determined with the following conditions

$$\begin{cases} \Phi_+^{(i)}(M) = F^{(i)}(M) \quad \text{and} \quad \Phi_-^{(i)}(-M) = F^{(i)}(-M) \qquad \text{for } i = 0, 1, 2, 3, 4, \\ \Phi_+(2M) = \Phi_-(-2M) = \dfrac{1}{4}((2M)^2 - 1)^2, \\ \Phi_+^{(1)}(2M) = \Phi_-^{(1)}(-2M) = (2M)^3 - 2M, \\ \Phi_+^{(i)}(2M) = \Phi_-^{(i)}(-2M) = 0 \qquad \text{for } i = 2, 3, 4. \end{cases} \tag{5.2}$$

Note that the truncation point $M = 2$ used here are for convenience only. For simplicity, we still denote the modified function $\hat{F}(u)$ by $F(u)$.

The spatial discretization is done by using the Galerkin finite element method. Let S_h denotes the P_s conforming finite element space defined by

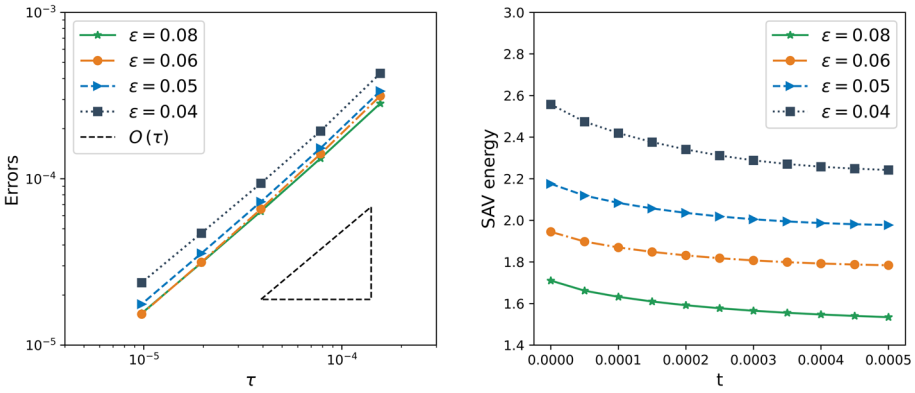

Fig. 1 (left) Time discretization errors; (right) Evolution of the SAV energy

$$S_h := \{v_h \in C(\bar{\Omega});\ \ v_h|_K \in P_s(K),\ \ \forall\, K \in \mathcal{T}_h\},$$

where $\mathcal{T}_h$ is a quasi-uniform triangulation of Ω. We introduce space notation $S_h^\circ := \{v_h \in S_h;\ (v_h, 1) = 0\}$, and define the discrete inverse Laplace operator $-\Delta_h^{-1} : L_0^2(\Omega) \to S_h^\circ$ such that

$$\left(\nabla(-\Delta_h^{-1})v, \nabla\eta_h\right) = (v, \eta_h) \qquad \forall\ \eta_h \in S_h. \tag{5.3}$$

Since the exact solution of the considered problem is not known, we compute the orders of convergence by the formula

$$\text{order of convergence} = \log\left(\frac{\|u_N^{(\tau)} - u_N^{(\tau/2)}\|_{h,-1}}{\|u_N^{(\tau/2)} - u_N^{(\tau/4)}\|_{h,-1}}\right)/\log(2)$$

based on the finest three meshes, where $u_N^{(\tau)}$ denotes the numerical solution at $t_N = T$ computed by using a stepsize τ, and $\|v\|_{h,-1} := \sqrt{(v, -\Delta_h^{-1}v)}$ for $v \in L_0^2(\Omega)$.

The time discretization errors in $\|\cdot\|_{h,-1}$-norm are presented in Fig. 1 (left) for four different $\varepsilon = 0.08, 0.06, 0.05, 0.04$ at $T = 0.005$, where we have used finite elements of degree $s = 3$ with a sufficiently spatial mesh $h = 1/64$ so that the error from spatial discretization is negligibly small in observing the temporal convergence rates. From Fig. 1 (left), we see that the error of time discretization is $O(\tau)$, which is consistent with the theoretical results proved in Theorem 4.1. In addition, Fig. 1 (right) shows the evolution in time of the discrete SAV energy for four different ε, which should be decreasing according to Theorem 3.1. This graph clearly confirms this decay property. Therefore, the numerical experiments are in accordance with our theoretical results.

Figure 2 shows snapshots of the numerical interface for four different $\varepsilon = 0.08, 0.06, 0.05, 0.04$ at six fixed time points. They clearly indicate that at each time point, as ε tends to zero, the numerical interface converges to the sharp interface of the Hele-haw flow, which is consistent with the phenomenon stated in [21, 26]. It also shows that for larger ε, the numerical interface evolves faster in time.

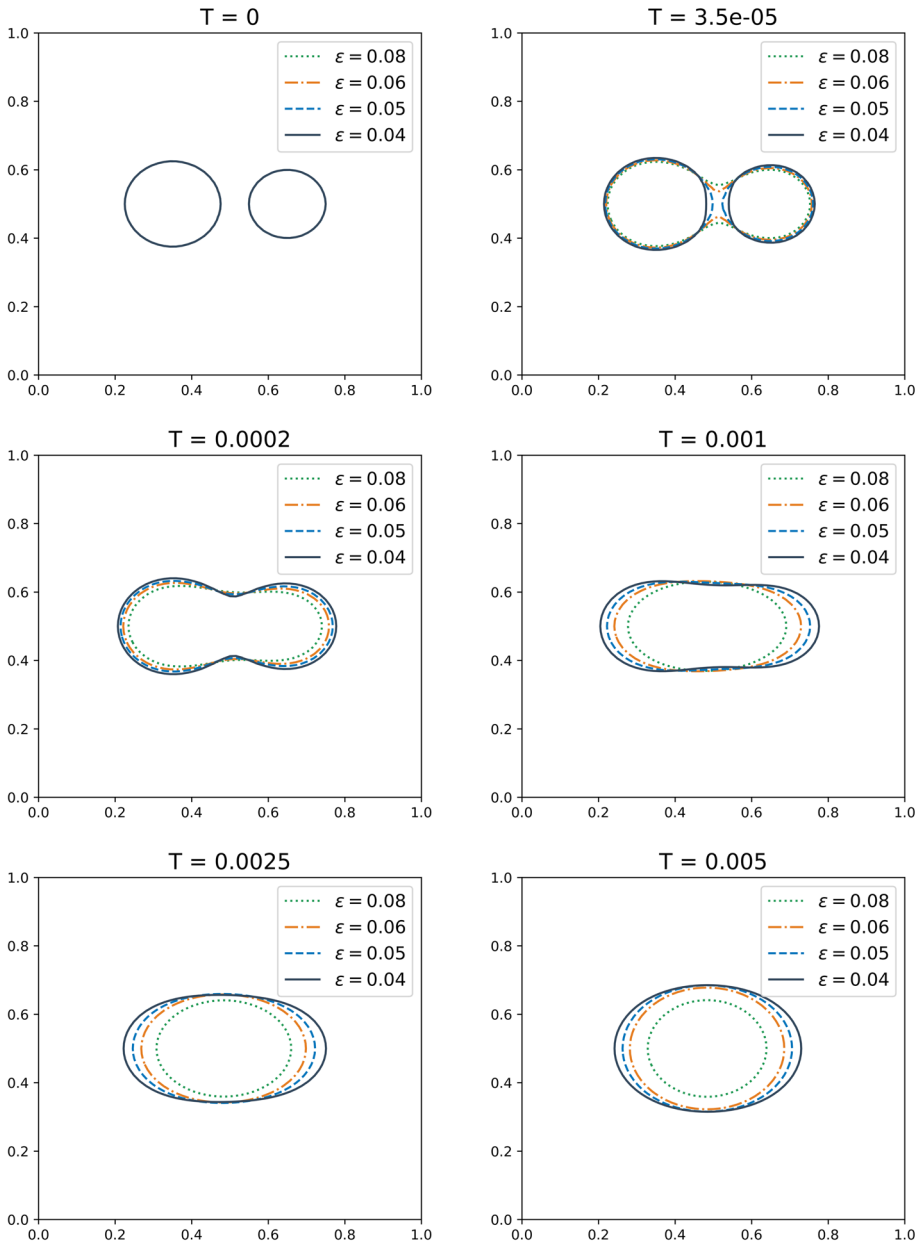

Fig. 2 Snapshots of the zero-level sets of the numerical solutions

Author Contributions Shu Ma, Weifeng Qiu and Xiaofeng Yang have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, method, analysis and writing. All authors certify that this material or similar material has not been and will not be submitted to or published in any other publication.

Funding Open access publishing enabled by City University of Hong Kong Library's agreement with Springer Nature The work of Shu Ma and Weifeng Qiu was partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China. (Project No. CityU 11300621).

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of interest No conflict of interest exists.

References

1. Akrivis, G., Li, B.: Error estimates for fully discrete bdf finite element approximations of the Allen-Cahn equation. IMA J. Numer. Anal. **42**(1), 363–391 (2022)
2. Akrivis, G., Li, B., Li, D.: Energy-decaying extrapolated RK-SAV methods for the Allen-Cahn and Cahn-Hilliard equations. SIAM J. Sci. Comput. **41**(6), A3703–A3727 (2019)
3. Alikakos, N.D., Fusco, G.: The spectrum of the Cahn-Hilliard operator for generic interface in higher space dimensions. Indiana Univ. Math. J. **42**(2), 637–674 (1993)
4. Barrett, J.W., Blowey, J.F.: An error bound for the finite element approximation of the Cahn-Hilliard equation with logarithmic free energy. Numer. Math. **72**(1), 1–20 (1995)
5. Bartels, S.: Robust a priori error analysis for the approximation of degree-one Ginzburg-Landau vortices. ESAIM: Math. Model. Numer. Anal. **39**(5), 863–882 (2005)
6. Bartels, S., Müller, R.: Error control for the approximation of Allen-Cahn and Cahn-Hilliard equations with a logarithmic potential. Numer. Math. **119**(3), 409–435 (2011)
7. Bartels, S., Müller, R.: Quasi-optimal and robust a posteriori error estimates in $L^\infty(L^2)$ for the approximation of Allen-Cahn equations past singularities. Math. Comput. **80**(274), 761–780 (2011)
8. Bartels, S., Müller, R., Ortner, C.: Robust a priori and a posteriori error analysis for the approximation of Allen-Cahn and Ginzburg-Landau equations past topological changes. SIAM J. Numer. Anal. **49**(1), 110–134 (2011)
9. Caffarelli, L.A., Muler, N.E.: An L^∞ bound for solutions of the Cahn-Hilliard equation. Arch. Ration. Mech. Anal. **133**(2), 129–144 (1995)
10. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. interfacial free energy. J. Chem. Phys. **28**(2), 258–267 (1958)
11. Cai, Y., Choi, H., Shen, J.: Error estimates for time discretizations of Cahn-Hilliard and Allen-Cahn phase-field models for two-phase incompressible flows. Numer. Math. **137**(2), 417–449 (2017)
12. Cai, Y., Shen, J.: Error estimates for a fully discretized scheme to a Cahn-Hilliard phase-field model for two-phase incompressible flows. Math. Comput. **87**(313), 2057–2090 (2018)
13. Chen, X.: Spectrum for the Allen-Chan, Chan-Hillard, and phase-field equations for generic interfaces. Commun. Partial Differ. Equ. **19**(7–8), 1371–1395 (1994)
14. Chen, X.: Global asymptotic limit of solutions of the Cahn-Hilliard equation. J. Differ. Geom. **44**(2), 262–311 (1996)

15. Du, Q., Nicolaides, R.A.: Numerical analysis of a continuum model of phase transition. SIAM J. Numer. Anal. **28**(5), 1310–1322 (1991)
16. Eck, C., Jadamba, B., Knabner, P.: Error estimates for a finite element discretization of a phase field model for mixtures. SIAM J. Numer. Anal. **47**(6), 4429–4445 (2010)
17. Elliott, C.M., French, D.A.: A nonconforming finite-element method for the two-dimensional Cahn-Hilliard equation. SIAM J. Numer. Anal. **26**(4), 884–903 (1989)
18. Elliott, C.M., French, D.A., Milner, F.: A second order splitting method for the Cahn-Hilliard equation. Numer. Math. **54**(5), 575–590 (1989)
19. Feng, X., Karakashian, O.: Fully discrete dynamic mesh discontinuous Galerkin methods for the Cahn-Hilliard equation of phase transition. Math. Comput. **76**(259), 1093–1117 (2007)
20. Feng, X., Li, Y.: Analysis of symmetric interior penalty discontinuous Galerkin methods for the Allen-Cahn equation and the mean curvature flow. IMA J. Numer. Anal. **35**(4), 1622–1651 (2015)
21. Feng, X., Li, Y., Xing, Y.: Analysis of mixed interior penalty discontinuous Galerkin methods for the Cahn-Hilliard equation and the hele-shaw flow. SIAM J. Numer. Anal. **54**(2), 825–847 (2016)
22. Feng, X., Prohl, A.: Numerical analysis of the Cahn-Hilliard equation and approximation for the Hele-Shaw problem, Part I: error analysis under minimum regularities. IMA Technical Report, (2001)
23. Feng, X., Prohl, A.: Numerical analysis of the Allen-Cahn equation and approximation for mean curvature flows. Numer. Math. **94**(1), 33–65 (2003)
24. Feng, X., Prohl, A.: Analysis of a fully discrete finite element method for the phase field model and approximation of its sharp interface limits. Math. Comput. **73**(246), 541–567 (2004)
25. Feng, X., Prohl, A.: Error analysis of a mixed finite element method for the Cahn-Hilliard equation. Numer. Math. **99**(1), 47–84 (2004)
26. Feng, X., Wu, H.: A posteriori error estimates for finite element approximations of the Cahn-Hilliard equation and the hele-shaw flow. Journal of Computational Mathematics, pages 767–796, (2008)
27. Feng, X., Wu, H.-J.: A posteriori error estimates and an adaptive finite element method for the Allen-Cahn equation and the mean curvature flow. J. Sci. Comput. **24**(2), 121–146 (2005)
28. Pego, R.L.: Front migration in the nonlinear Cahn-Hilliard equation. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 422(1863):261–278, (1989)
29. Prohl, A., Feng, X.H.: Numerical analysis of the Cahn-Hilliard equation and approximation for the hele-shaw problem. Interfaces Free Bound. **7**(1), 1–28 (2005)
30. Shen, J., Xu, J.: Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows. SIAM J. Numer. Anal. **56**(5), 2895–2912 (2018)
31. Shen, J., Xu, J., Yang, J.: The scalar auxiliary variable (SAV) approach for gradient flows. J. Comput. Phys. **353**, 407–416 (2018)
32. Shen, J., Xu, J., Yang, J.: A new class of efficient and robust energy stable schemes for gradient flows. SIAM Rev. **61**(3), 474–506 (2019)
33. Shen, J., Yang, X.: Numerical approximations of Allen-Cahn and Cahn-Hilliard equations. Discret. & Contin. Dyn. Syst. **28**(4), 1669 (2010)
34. Wang, L., Yu, H.: On efficient second order stabilized semi-implicit schemes for the Cahn-Hilliard phase-field equation. J. Sci. Comput. **77**(2), 1185–1209 (2018)
35. Yang, X.: Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends. J. Comput. Phys. **327**, 294–316 (2016)
36. Yang, X., Ju, L.: Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model. Comput. Methods Appl. Mech. Eng. **315**, 691–712 (2017)
37. Yang, X., Ju, L.: Linear and unconditionally energy stable schemes for the binary fluid-surfactant phase field model. Comput. Methods Appl. Mech. Eng. **318**, 1005–1029 (2017)
38. Yang, X., Zhao, J., Wang, Q., Shen, J.: Numerical approximations for a three-component Cahn-Hilliard phase-field model based on the invariant energy quadratization method. Math. Models Methods Appl. Sci. **27**(11), 1993–2030 (2017)
39. Yue, P., Feng, J.J., Liu, C., Shen, J.: A diffuse-interface method for simulating two-phase flows of complex fluids. J. Fluid Mech. **515**, 293–317 (2004)