Numerical methods for the mass-conserved Ohta-Kawasaki equation

Juan Zhang · Shifeng Li · Kai Jiang

Abstract In this paper, we propose a numerical method to solve the mass-conserved Ohta-Kawasaki equation with finite element discretization. An unconditional stable convex splitting scheme is applied to time approximation. The Newton method and its variant are used to address the implicitly nonlinear term. We rigorously analyze the convergence of the Newton iteration methods. Theoretical results demonstrate that two Newton iteration methods have the same convergence rate, and the Newton method has a smaller convergent factor than the variant one. To reduce the condition number of discretized linear system, we design two efficient block preconditioners and analyze their spectral distribution. Finally, we offer numerical examples to support the theoretical analysis and indicate the efficiency of the proposed numerical methods for the mass-conserved Ohta-Kawasaki equation.

Keywords Mass-conserved Ohta-Kawasaki equation · Convex splitting scheme · Finite element discretization · Newton method · Block preconditioners

1 Introduction

Diblock copolymers are macromolecules composed of two incompatible blocks linked together by covalent bonds. The incompatibility between the two blocks drives the system to phase separation, while the chemical bonding of two blocks prevents the macroscopic phase separation. These competition factors lead diblock copolymers to self-assembling into a rich class of complex nanoscale structures [12]. Modeling and numerical simulation are effective means to investigate phase behaviors of block copolymers, such as the self-consistent field theory, coarse-grained density functional theory [3–5,6]. Among these theories, Ohta
and Kawasaki [7] presented an effective free energy functional to study diblock copolymers, which can be rescaled as

$$E(u) = \int_{\Omega} \left( \frac{\varepsilon^2}{2} |\nabla u|^2 + \frac{1}{4}(1-u^2)^2 + \frac{\sigma}{2} |(-\Delta)^{-\frac{1}{2}}(u-m)|^2 \right) dx.$$  \hspace{1cm} (1)

$u(x)$ is the order parameter which measures the order of diblock copolymer system. $m = \frac{1}{|\Omega|} \int_{\Omega} u(x) dx$ denotes the average mass of the melt on the domain $\Omega$. The parameters $\varepsilon \ll 1$ and $\sigma$ measure the interfacial thickness in the region of pure phases and the non-local interaction potential, respectively. In the energy functional (1), the first term penalizes the jump in the solution, the second term favors $u = \pm 1$, and the last term penalizes variation from the mean by a long-range interaction. More physical background about the Ohta-Kawasaki free energy functional can refer to [7], and corresponding mathematical theories can be found in the literature [8] and references therein.

Using the Ohta-Kawasaki free energy functional, a mass-conserved dynamic equation can be given as

$$u_t = \Delta \mu.$$  \hspace{1cm} (2)

$\mu$ is the chemical potential, i.e., the variation derivative of $E$ with respect to $u$

$$\mu = \delta E/\delta u = -\varepsilon^2 \Delta u - u(1-u^2) - \sigma \Delta^{-1}(u-m).$$  \hspace{1cm} (3)

By introducing a new variable

$$w = -\varepsilon^2 \Delta u - u(1-u^2),$$

the forth-order dynamic equation (2) can be split into two second-order equations on $\Omega \times [0,T]$

$$u_t - \Delta w + \sigma (u-m) = 0,$$  \hspace{1cm} (4a)

$$w + \varepsilon^2 \Delta u - u(u^2 - 1) = 0.$$  \hspace{1cm} (4b)

It is easy to verify that the energy functional (1) is nonincreasing in time along the solution trajectories of (4) with homogeneous Neumann boundary condition

$$\nabla u \cdot n = 0 \text{ and } \nabla w \cdot n = 0 \text{ on } [0,T]$$  \hspace{1cm} (5)

and a given initial value $u(x,0) = u_0(x), x \in \Omega$.

From the numerical computation viewpoint, it is necessary to construct an efficient numerical method to solve the gradient flow (4). For time discretization direction, in recent literature, many energy stable approaches have been proposed, e.g., convex-splitting schemes [9], stabilized factor methods [10], auxiliary variable approaches [11], exponential time differencing schemes [12]. Among these time discretization approaches, the convex splitting scheme (CSS in short) splits the non-convex nonlinear term into two convex parts and explicitlyimplicitly treats them, which does not have time step restriction in theory. Besides the time discretization, we adopt the widely used finite element method (FEM in short) to discretize the spatial variable in (4).

Solving the discretization system faces twofold difficulties. One is the implicitly nonlinear term that will take the most time-consuming part in each time update. The common approaches to update the nonlinear term include the Picard method [13,14], the Newton
method [15][16], the nonlinear multigrid method [17][18], the preconditioned steepest descent algorithm [19][20], etc. The Picard method and the Newton method have been applied to the Ohta-Kawasaki equation [15][21]. However, the convergence rate of the Picard method is usually unsatisfactory [22]. An alternative selection is to turn to the Newton method. Therefore, we choose the Newton method to update the nonlinearity term of Ohta-Kawasaki equation (4). More significantly, we present rigorous theoretical analysis.

The other difficulty is that we need to solve the ill-conditioned linear algebra system at each time step, leading to high computational cost. Due to this, we begin to explore feasible calculation schemes and present corresponding theoretical analysis. As we know, the previous methods, such as Krylov subspace methods for solving the linear system, may not be convergent or converge slowly without appropriate preconditioners. Thus, recent researchers have paid attention to the construction of preconditioners for the Ohta-Kawasaki equation [23][24]. Our other goal is to design efficient preconditioners using Schur complement approximation and a modified Hermitian and skew-Hermitian splitting (MHSS in short), respectively.

The main contribution of our paper is summarized as follows:

1. We present an unconditionally stable energy method to mass-conserved Ohta-Kawasaki equation based on the convex-splitting scheme;
2. We apply and theoretically analyze two Newton methods to update implicitly nonlinear terms;
3. We propose two block preconditioners to solve the ill-conditioned system.

The rest of the paper is organized as follows. In Section 2, the Ohta-Kawasaki dynamic equation is discretized via the CSS and the FEM. Section 3 presents the convergent analysis for two Newton methods, and compares their convergent speed. Section 4 gives two block triangular preconditioners, and analyzes corresponding spectral distributions. Section 5 offers numerical examples to verify theoretical results and demonstrate the efficiency of our proposed method. Some concluding remarks are drawn in Section 6.

Throughout the paper, the set of $n \times n$ complex and real matrices are denoted by $\mathbb{C}^{n \times n}$ and $\mathbb{R}^{n \times n}$. If $X, Y \in \mathbb{R}^{n \times n}$, let $X^{-1}, X^T, ||X||_2$ represent the inverse, conjugate transpose and the spectral norm of $X$, respectively. The expression $X \succ 0$ ($X \succeq 0$) means that $X$ is symmetric (semi-) positive definite. $X \succ Y$ ($X \succeq Y$) represents that $X - Y$ is symmetric (semi-) positive definite. The identity matrix is expressed by $I$. Let $W^{h,p}(\Omega)$ be a standard Sobolev space equipped with the norm $|| \cdot ||_{d_h,p}$ and the semi-norm $| \cdot |_{d_h,p}$,

$$
||v||_{d_h,p} = \begin{cases} 
\left\{ \sum_{|\alpha| \leq d_h} ||D^\alpha v||_{L^p(\Omega)} \right\}^{\frac{1}{p}}, & 1 \leq p < \infty, \\
\max_{|\alpha| \leq d_h} ||D^\alpha v||_{L^\infty(\Omega)}, & p = \infty,
\end{cases}
$$

$$
|v|_{d_h,p} = \begin{cases} 
\left\{ \sum_{|\alpha| = d_h} ||D^\alpha v||_{L^p(\Omega)} \right\}^{\frac{1}{p}}, & 1 \leq p < \infty, \\
\max_{|\alpha| = d_h} ||D^\alpha v||_{L^\infty(\Omega)}, & p = \infty.
\end{cases}
$$

The notation $|| \cdot ||$ stands for the $L^2(\Omega)$-norm.
2 Numerical discretization

Before we go further, it is necessary to give the weak form of (4). Using $L^2(\Omega)$-inner product and test function $v \in H^1(\Omega)$, we can have the variational formulation of equation (4). Denote the bulk energy density as

$$\Phi(u) = \frac{1}{4}(1 - u^2)^2. \quad (6)$$

We seek to find $u(\cdot, t) \in H^1(\Omega)$ and $w(\cdot, t) \in H^1(\Omega)$ such that

$$\begin{align*}
(u, v) + (\nabla w, \nabla v) + \sigma(u - m, v) &= 0, \\
(w, v) - \varepsilon^2(\nabla u, \nabla v) - (\Phi'(u), v) &= 0,
\end{align*} \quad (7)$$

where

$$\Phi'(u) = -u(1 - u^2).$$

In the following subsections, we will discretize (7) through CSS and FEM.

2.1 CSS discretization

The CSS, originally proposed by Eyre \cite{9}, splits the non-convex bulk energy density $\Phi(u)$ into two convex functions:

$$\Phi(u) = \Phi_+(u) - \Phi_-(u),$$

with

$$\Phi_+(u) = \frac{1}{4}(u^4 + 1), \quad \Phi_-(u) = \frac{1}{2}u^2.$$  

Using the above notations, (7) becomes

$$\begin{align*}
(u, v) + (\nabla w, \nabla v) + \sigma(u - m, v) &= 0, \\
(w, v) - \varepsilon^2(\nabla u, \nabla v) - (\Phi'_+(u) - \Phi'_-(u), v) &= 0,
\end{align*} \quad (8)$$

with

$$\Phi'_+(u) = u^3, \quad \Phi'_-(u) = u.$$

The CSS treats $\Phi'_+$ implicitly, and $\Phi'_-$ explicitly. In particular, let $\Delta t = T/N, N \geq 1$, be the uniform time step size. $u^n$ and $w^n$ represent the approximation of $u(\cdot, t_n)$ and $w(\cdot, t_n)$ in $H^1(\Omega)$, where $t_n = n\Delta t, n = 0, 1, \cdots, N$. The continuous system (7) can be discretized as

$$\begin{align*}
\left(\frac{u^n - u^{n-1}}{\Delta t}, v\right) + (\nabla w^n, \nabla v) + \sigma(u^n - m, v) &= 0, \\
(w^n, v) - \varepsilon^2(\nabla u^n, \nabla v) - (\Phi'_+(u^n) - \Phi'_-(u^{n-1}), v) &= 0,
\end{align*} \quad (9)$$

where the test function $v \in H^1(\Omega)$. The CSS \cite{9} has the energy dissipation law as shown in Theorem 1.

**Theorem 1** The scheme (9) is unconditionally energy stable. Moreover, we have

$$E(u^n) \leq E(u^{n-1}) - \Delta t\|\nabla \mu\|^2 - \frac{\varepsilon^2}{2}\|\nabla (u^n - u^{n-1})\|^2. \quad (10)$$

**Proof** See Appendix A.
2.2 Finite element discretization

Next, we discretize the semi-discrete system (9) in space by the FEM. Denote $X = H^1(\Omega) = W^{1,2}(\Omega)$. $h$ is a real positive number and $\tau_h = \{K_\Omega : \cup_{K_\Omega} \subset \Omega, \Omega = \Omega\}$ is a quasi-uniform regular partition of $\Omega$ with diameters bounded by $h$. For a given $\tau_h, V_h \subset X$ is the FEM space defined by

$$V_h = \{v_h \in C(\Omega) \cap X; v_h|_{K_\Omega} \in P_1(K_\Omega), \forall K_\Omega \in \tau_h\},$$

where $P_1(K_\Omega)$ is the polynomial space of degree, not greater than 1. The full-discrete system for (7) is obtained by seeking for the test function $v_h \in V_h$ such that

$$\begin{align*}
&\frac{d}{dt} (u_h^n - u_h^{n-1}) + (\nabla u_h^n, \nabla v_h) + \sigma (u_h^n - m, v_h) = 0, \quad (11a) \\
&(w_h^n, v_h) - \epsilon^2 (\nabla u_h^n, \nabla v_h) - (\Phi'_h (u_h^n) - \Phi'_h (u_h^{n-1}), v_h) = 0. \quad (11b)
\end{align*}$$

The sequence $u_h^n$ generated by the finite element approximation (11) is bounded uniformly in $h$, as shown in Theorem 2.

**Theorem 2** The sequence $u_h^n$ defined by the finite element approximation is bounded, i.e.,

$$\|u_h^n\|_{1,\infty} \leq C(\epsilon, \sigma, u_0, m, c, c_e, T, |\Omega|), \quad n = 1, 2, \ldots, N,$$

where $c$ and $c_e$ depend only on the space dimension $d$ and $\Omega$.

**Proof** See Appendix B.

3 Convergent analysis for Newton iteration methods

In this section, we apply two Newton methods to treat the nonlinear term. The first Newton method of linearizing $(u_h^n)^3$ at $u_h^{n,k}$ in [15] and [21] is

$$(u_h^n)^3 \approx (u_h^{n,k})^3 + 3(u_h^{n,k})^2(u_h^{n,k} - u_h^{n,k-1}).$$

The second variant Newton (V-N) method of linearizing $(u_h^n)^3$ in [24] and [25] is

$$(u_h^n)^3 \approx (u_h^{n,k-1})^3 u_h^{n,k},$$

where $u_h^{n,0} = u_h^{n-1}$.

To the best of our knowledge, there has been no convergence analysis for the two Newton methods in related literature. One of our main work is providing the convergent analysis for two Newton methods and comparing their convergent speed.

Next, we will present a unified framework to show the convergence analysis of two nonlinear Newton methods. We first establish the theoretical framework to the Newton method, then apply it to the V-N method. Using the Newton approximation (13), the discrete scheme (9) can be rewritten as

$$\begin{align*}
&\frac{d}{dt} (u_h^n - u_h^{n-1}) + \Delta t (\nabla u_h^n, \nabla v_h) - \Delta t \sigma (m, v_h) + (u_h^{n-1}, v_h), \quad (15a) \\
&(w_h^n, v_h) - \epsilon^2 (\nabla u_h^n, \nabla v_h) + (-3(u_h^{n,k-1})^2(u_h^{n,k}), v_h) + (u_h^{n-1}, v_h) = -2((u_h^{n,k-1})^3, v_h). \quad (15b)
\end{align*}$$
3.1 Bilinear forms

From (11), we have
\[ ε^2(\nabla u_h^n, \nabla v_h) + ((u_h^n)^2, v_h) + (1 + Δtσ)(u_h^n, v_h) - (w_h^n, v_h) + Δt(\nabla w_h^n, \nabla v_h) = (f, v_h), \] (16)
where
\[ f = 2u_h^{n-1} + Δtσm. \] (17)

To simplify the following analysis, we introduce the bilinear forms. For \( u, w \in X \), denote
\[ Ψ(u, w; v) = ε^2(\nabla u, \nabla v) + (1 + Δtσ)(u, v) + (u^3, v) - (f, v) + Δt(\nabla w, \nabla v) - (w, v), \]
where
\[ Ψ_1(u, v) = (F_1(., u, \nabla u), \nabla v) + (g_1(., u, \nabla u), v), \]
\[ Ψ_2(w, v) = (F_2(., w, \nabla w), \nabla v) + (g_2(., w, \nabla w), v), \]
and
\[ F_1(x, u, \nabla u) = ε^2 \nabla u, \ g_1(x, u, \nabla u) = (1 + Δtσ)u + u^3 - f, \]
\[ F_2(x, w, \nabla w) = Δt \nabla w, \ g_2(x, w, \nabla w) = -w. \]

Then (16) is equivalent to
\[ Ψ(u_h^n, w_h^n; v_h) = 0. \] (18)

Evidently, \( Ψ(u, w; v) \) is a linear combination of bilinear forms \( Ψ_1 \) and \( Ψ_2 \). For \( i = 1, 2 \), note that
\[ F_i(x, y, z): Ω × \mathbb{R}^1 × \mathbb{R}^2 \rightarrow \mathbb{R}^2 \] and \( g_i(x, y, z): Ω × \mathbb{R}^1 × \mathbb{R}^2 \rightarrow \mathbb{R}^1 \)
are smooth functions. Thus for any \( u, w \in X \), we have
\[ F_{1,z}(x, u, \nabla u) = ε^2 I_2, \ F_{1,y}(x, u, \nabla u) = 0_2, \ g_{1,z}(x, u, \nabla u) = 1 + Δtσ + 3u^2, \]
\[ F_{2,z}(x, w, \nabla w) = Δt I_2, \ F_{2,y}(x, w, \nabla w) = 0_2, \ g_{2,z}(x, w, \nabla w) = 0_2, \ g_{2,y}(x, w, \nabla w) = -1, \]
where \( I_2 \) and \( 0_2 \) represent 2-order identity matrix and 2-order zero matrix, respectively.

Similar to (20), for (18), we introduce the bilinear form
\[ Ψ'(u, w; v_1, v_2, v) = Ψ_1'(u; v_1, v) + Ψ_2'(w; v_2, v), \] (19)
where
\[ Ψ_1'(u; v_1, v) = (F_{1,z} \nabla v_1 + F_{1,y} v_1, \nabla v) + (g_{1,z} \nabla v_1 + g_{1,y} v_1, v) \]
\[ = (ε^2 \nabla v_1, \nabla v) + (1 + Δtσ)(v_1, v) + 3(u^2 v_1, v), \]
\[ Ψ_2'(w; v_2, v) = (F_{2,z} \nabla v_2 + F_{2,y} v_2, \nabla v) + (g_{2,z} \nabla v_2 + g_{2,y} v_2, v) \]
\[ = Δt(\nabla v_2, \nabla v) - (v_2, v). \]

Obviously,
\[ Ψ_2(w, v) = Ψ_2'(w; v). \] (20)
3.2 Some useful lemmas

The following lemmas are helpful to prove the convergence result.

**Lemma 1** \([27]\) For any \(u_1, u_2, w_1, v \in X\), denote
\[
\Psi(u_1, w_1; v) = \Psi(u_2, w_2; v) + \Psi'(u_2, w_2; u_1 - u_2, w_1 - w_2, v) + R(u_2, w_2, u_1, w_1, v).
\]  \hspace{1cm} (21)

For any given \(k_0 > 0\), if
\[
||u_1||_{1,∞} ≤ k_0, \quad ||u_2||_{1,∞} ≤ k_0,
\]  \hspace{1cm} (22)
then the remainder \(R\) satisfies
\[
|R(u_2, w_2, u_1, w_1, v)| = |R_1(u_1, u_2, v)| ≤ C(k_0)||e||_{H^2, p}||v||_{0, q},
\]
where \(C \geq \max_{x \in \Omega, |y| ≤ k_0} |g(x, y)|\) and \(e = u_1 - u_2\).

**Lemma 2** \([26]\) If \(h\) is sufficiently small and \(u \in X\) is an isolated solution of \([4]\), then
\[
||\Psi'(u; v, \phi)|| ≤ c||v||_1||\phi||_1, \quad \forall v \in W^{1,p}(\Omega), \quad \phi \in W^{1,q}(\Omega).
\]

**Lemma 3** \([28, 29]\) For a given \(z \in \Omega\), we can find \(\delta_h(z) \in \mathcal{V}_h\) such that
\[
\Psi_1'(u; \delta_h(z), \delta_h(z)) = \partial z \forall u \in \mathcal{V}_h,
\]
where
\[
\sup_{z \in \Omega} ||\delta_h(z)||_{1,1} ≤ c|\log h|.
\]

3.3 Convergent analysis for Newton method

In this subsection, we present the error estimate for the Newton method. Applying the Newton method to \([18]\), we have
\[
\Psi_1'(u_h^{n, k-1}; u_h^{n, k}, v_h) + \Psi_2'(w_h^{n, k-1}; w_h^{n, k}, v_h) = \Psi_1'(u_h^{n, k-1}; u_h^{n, k-1}, v_h) + \Psi_2'(w_h^{n, k-1}; w_h^{n, k-1}, v_h)
\]
\[
- \Psi_1'(u_h^{n, k-1}, v_h) - \Psi_2'(w_h^{n, k-1}, v_h), \quad \forall v_h \in \mathcal{V}_h.
\]  \hspace{1cm} (23)

The error estimate for the Newton method is illustrated by Theorem [3].

**Theorem 3** Assume that \(h\) is sufficiently small, \(u \in X\) is an isolated solution of \([4]\) and \(0 < \eta < 1\). Take \(u_h^{n, 0} = u_h^{n-1}\) such that
\[
||u_h^{n, 0} - u_h^{n-1}||_{1,∞} ≤ \eta.
\]  \hspace{1cm} (24)

For constants \(c_3\) and \(2 \leq p ≤ \infty\), if the convergence factor \(\rho\) satisfies
\[
\rho = \frac{c_3 C(k_0)|\log h|}{1 - c_3 [(C(k_0) + 1)(c_2 + \eta) + c||\log h||]} < 1,
\]  \hspace{1cm} (25)
then
\[
||u_h^n - u_h^{n, m}||_{1, p} ≤ c\rho^n.
\]  \hspace{1cm} (26)
Proof Using the definition (21) in Lemma 1 we get

$$
\Psi_1(u_h^{n,k-1}, v_h) + \Psi_2(w_h^{n,k-1}, v_h) + \Psi'_1(u_h^{n,k-1}, u_h^n - u_h^{n,k-1}, v_h) \\
+ \Psi'_2(u_h^{n,k-1}, w_h^n - w_h^{n,k-1}, v_h) + R_1(u_h^{n,k-1}, u_h^n, v_h) = 0. \tag{27}
$$

Combining (23) with (27), it is obvious that

$$
\Psi_1(u_h^{n,k-1}, v_h) + \Psi_2(w_h^{n,k}, v_h) + \Psi'_1(u_h^{n,k-1}, u_h^n - u_h^{n,k-1}, v_h) = 0, \tag{28}
$$

$$
\Psi_1(u_h^{n,k-1}, v_h) + \Psi_2(w_h^n, v_h) + \Psi'_1(u_h^{n,k-1}, u_h^n - u_h^{n,k-1}, v_h) + R_1(u_h^{n,k-1}, u_h^n, v_h) = 0. \tag{29}
$$

Subtracting (29) from (28), we have the following error equation

$$
\Psi_2(w_h^n - w_h^{n,k}, v_h) + \Psi'_1(u_h^{n,k-1}; u_h^n - u_h^{n,k}, v_h) + R_1(u_h^{n,k-1}, u_h^n, v_h) = 0. \tag{30}
$$

1) Firstly, using mathematical induction method for $\tilde{m}$, we prove the following estimate

$$
||u_h^n - u_h^{n,\tilde{m}}||_{1,\infty} \leq \rho^{2^{n-1}} \eta^{2^n}. \tag{31}
$$

i) When $\tilde{m} = 0$, (31) is evidently true according to the initial condition (24).

ii) Assume that (31) is true for $\tilde{m} = k - 1$.

iii) It is only required to prove (31) for $\tilde{m} = k$. From (25), we know that

$$
||u||_{1,\infty} \leq C(\varepsilon, \sigma, u_0, c_p, c_s, T, m, |\Omega|)
$$

is true. Combining with Theorem 2 it follows that

$$
||u - u_h^n||_{1,\infty} \leq C(\varepsilon, \sigma, u_0, c_p, c_s, T, m, |\Omega|) \leq c_2.
$$

Thus, according to the induction assumption of $\tilde{m} = k - 1$, we have

$$
||u - u_h^{n,k-1}||_{1,\infty} \leq ||u - u_h^n||_{1,\infty} + ||u_h^n - u_h^{n,k-1}||_{1,\infty} \\
\leq c_2 + (\rho \eta)^{2^{n-1}} \eta \leq c_2 + \eta.
$$

Then

$$
||u_h^{n,k-1}||_{1,\infty} \leq C + c_2 + \eta \leq \tilde{k}_0. \tag{32}
$$

According to the definition of $\Psi'_1(\cdot, \cdot, \cdot)$, we obtain

$$
|\Psi'_1(u_h^{n,k-1} - u_h^n, v_h) - \Psi'_1(u_h^{n,k-1}, u_h^n - u_h^{n,k}, v_h)| \leq \mathcal{F}(u, u_h^{n,k-1}) ||u_h^{n,k-1}||_{1,\infty} ||v_h||_{1,1}, \tag{33}
$$

where

$$
\mathcal{F}(u, u_h^{n,k-1}) = ||F_{1,z}(x, u, \nabla u) - F_{1,z}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1})||_{0,\infty} \\
+ ||F_{1,y}(x, u, \nabla u) - F_{1,y}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1})||_{0,\infty} \\
+ ||g_{1,z}(x, u, \nabla u) - g_{1,z}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1})||_{0,\infty} \\
+ ||g_{1,y}(x, u, \nabla u) - g_{1,y}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1})||_{0,\infty} \\
= ||g_{1,y}(x, u, \nabla u) - g_{1,y}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1})||_{0,\infty}. \tag{34}
$$
By setting
\[ \Phi(t) = g_{1,y}(x, u + t(u_h^{n,k-1} - u), \nabla u_1 + t\nabla(u_h^{n,k-1} - u)), \]
for \( u, u_h^{n,k} \in X \cap W^{1,\infty}(\Omega) \) and using the definition of \( g_{1,y} \), we have
\[
|g_{1,y}(x, u_h^{n,k-1}, \nabla u_h^{n,k-1}) - g_{1,y}(x, u, \nabla u)| = \left| \int_0^1 \Phi'(t) \, dt \right|
\leq \max_{x \in \Omega, |t| \leq h_0, z \leq k} |(g_{1,y})||u - u_h^{n,k-1}|_{1,\infty}. \tag{35}
\]
Combining (34) with (35), it follows that
\[
\mathcal{F}(u, u_h^{n,k-1}) \leq C(k_0)||u - u_h^{n,k-1}||_{1,\infty}. \tag{36}
\]
Substituting (36) into (33) yields
\[
|\Psi'_1(u; u_h^{n,k} - u_h^n, \nu_h) - \Psi'_1(u_h^{n,k-1}; u_h^{n,k} - u_h^n, \nu_h)|
\leq C(k_0)||u - u_h^{n,k-1}||_{1,\infty}||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1}
\leq (c_2 + \eta) C(k_0)||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1}. \tag{37}
\]
Moreover, using Lemma 2 and (32), it is evident that
\[
|\Psi'_1(u_h^{n,k-1}; u_h^n - u_h^n, \nu_h)| \leq (c_2 + \eta)||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1} + |\Psi'_1(u_h^{n,k} - u_h^n, \nu_h)|
\leq (c_2 + \eta)||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1} + c||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1}
= (c_2 + \eta) + c||u_h^n - u_h^{n,k}||_{1,\infty}||v_h||_{1,1}. \tag{38}
\]
Now for any \( z \in \Omega \), let \( \hat{g}_h \) be the discrete Green function. Note that
\[
\partial (w_h^{n,k} - w_h^n)(z) = \Psi'_2(w; w_h^{n,k} - w_h^n, \hat{g}_h)
= \Psi'_2(w; w_h^{n,k} - w_h^n, \hat{g}_h) + \Psi_2(w_h^{n,k} - w_h^n, \hat{g}_h)
+ \Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) + R_1(u_h^{n,k-1}; u_h^n, \hat{g}_h)
= \Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) + R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h).
\]
Applying Lemma 3 to (39), we have
\[
|\partial (u_h^{n,k} - u_h^n)(z)| = |\Psi'_2(u; u_h^{n,k} - u_h^n, \hat{g}_h) - \Psi_2(w_h^{n,k} - w_h^n, \hat{g}_h)|
- |\Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) - R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h)|
= |\Psi'_2(u; u_h^{n,k} - u_h^n, \hat{g}_h) - \partial (w_h^{n,k} - w_h^n)(z)|
- |\Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) - R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h)|. \tag{39}
\]
Therefore, combining (37), (38) with (39), by Lemma 1 we obtain
\[
|\partial (u_h^{n,k} - u_h^n)(z)| = |\Psi'_2(u; u_h^{n,k} - u_h^n, \hat{g}_h) - 2\Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) - 2R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h)|
\leq |\Psi'_2(u; u_h^{n,k} - u_h^n, \hat{g}_h) - \Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h) - 2R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h)|
+ |\Psi'_2(u_h^{n,k-1}; u_h^n - u_h^{n,k}, \hat{g}_h)| + |R_1(u_h^{n,k-1}, u_h^n, \hat{g}_h)|
\leq |\mathcal{F}(u, u_h^{n,k-1}) + (c_2 + \eta) + c||\log h||u_h^n - u_h^{n,k}||_{1,\infty} + C(k_0)||\log h||||u_h^n - u_h^{n,k-1}||_{1,\infty}. \tag{40}
\]
Applying Lemma 3 to (40) yields
\[ \|u_h^{n,k} - u_h^n\|_{1,\infty} \leq c_3 \|C(k_0) + 1\| \|C(k_0) + 1\| \log h \|\|u_h^{n,k} - u_h^n\|_{1,\infty} \]
\[ + c_5 \log h \|\|u_h^{n,k} - u_h^{n,k-1}\|_{1,\infty} \].

Thus, it can be seen that
\[ \|u_h^{n,k} - u_h^n\|_{1,\infty} \leq \frac{c_3 \log h}{1 - c_3 \|C(k_0) + 1\|} \|\|u_h^{n,k} - u_h^{n,k-1}\|_{1,\infty} \]
\[ = \rho \|\|u_h^{n,k} - u_h^{n,k-1}\|_{1,\infty} \]. \quad (41)

Using induction assumption, it is evident that
\[ \|u_h^{n,k} - u_h^{n,k-1}\|_{1,\infty} \leq \rho^{2^k-1} \eta^{2^k-1}. \quad (42) \]

Combining (41) with (42) yields
\[ \|u_h^{n,k} - u_h^n\|_{1,\infty} \leq \rho^{2^k-1} \eta^{2^k}. \]

Then (41) is true for any \( k \).

2) Secondly, note that
\[ \|u_h^{n,0} - u_h^{n,0}\|_{1,\infty} \leq \|\|u_h^n - u_h^{n,0}\|_{1,\infty} \]. \quad (43)

In terms of the condition (25), substituting (43) into (41) yields
\[ \|u_h^n - u_h^{n,0}\|_{1,\infty} \leq c\rho^k, \]
i.e., (26). The proof is completed.

3.4 Convergent analysis for the V-N method

Next, we analyze the convergence for the V-N method. Furthermore, we compare the convergent factor of two Newton methods.

Applying (14) to (9) can lead to the V-N scheme
\[ (1 + \sigma A)u_h^{n,k} + \Delta t(\nabla u_h^{n,k}, \nabla v_h) = \Delta t(\nabla m, v_h) + (u_h^{n-1} - v_h), \quad (44a) \]
\[ (u_h^{n,k} - v_h) - \epsilon^2(\nabla u_h^{n,k}, \nabla v_h) + (-u_h^{n,k-1})^2 (u_h^{n,k} - v_h) = 0. \quad (44b) \]

Subtracting (44b) from (44a) yields
\[ \epsilon^2(\nabla u_h^{n,k}, \nabla v_h) + ((u_h^{n,k-1})^2 (u_h^{n,k} - v_h) + (1 + \Delta t)(u_h^{n,k} - v_h) - \Delta t(\nabla u_h^{n,k}, \nabla v_h) = (f, v_h), \quad (45) \]

where \( f \) is defined by (17). Then we can rewrite (45) as
\[ \Psi_1(u_h^{n,k-1}, v_h) + \Psi_2(w_h^{n,k}, v_h) + \Psi_1'(u_h^{n,k-1}, u_h^{n,k} - u_h^{n,k-1}, v_h) - 2((u_h^{n,k-1})^2 (u_h^{n,k} - u_h^{n,k-1}), v_h) = 0. \quad (46) \]

For (46), similarly to the proof of Theorem 3, we obtain the convergent conclusion for the V-N method.
Theorem 4 Assume that $h$ is sufficiently small, $u \in X$ is an isolated solution of (4) and $0 \leq \eta < 1$. Take $u_h^{n,0} = u_h^{n-1}$ such that (4) is satisfied. For constants $c_3, c_4$ and $2 \leq p \leq \infty$, if the convergence factor $\hat{\rho}$ satisfies

$$\hat{\rho} = \frac{c_3C(k_0)[\log h]}{1 - c_3[(C(k_0) + 1)(c_2 + \eta) + c + c_4][\log h]} < 1,$$

then

$$||u_h^n - u_h^{n-1}||_{1,p} \leq c\hat{\rho}^3n.$$

Remark 1 Obviously, $\rho < \hat{\rho}$. This shows that the convergent speed of the Newton method is faster than that of the V-N method. It will be verified by numerical experiments in Section 5 as well.

4 Block preconditioners

In practical implementation, we impose a uniform discretization on the spatial domain $\mathcal{T}_h$ for each $t_n = n\Delta t$ ($0 \leq n \leq N$). We assume that the dimension of the finite element subspace $V_h$ is $\kappa$, and use the set of piecewise linear $\phi_i$ as the basis functions which are defined in the usual way. Then $V_h$ can be spanned in terms of these basis functions as

$$V_h = \text{span} \{ \phi_i \}_{i=0}^{\kappa-1}.$$

Let $\{\phi_i\}_{i=0}^{\kappa-1}$ be a basis of $V_h$, then $u_h^n$ and $w_h^n$ can be expressed as

$$u_h^n = \sum_{i=0}^{\kappa-1} U_i^n \phi_i(x), \quad w_h^n = \sum_{i=0}^{\kappa-1} W_i^n \phi_i(x).$$

When taking $v_h = \phi_j$, the discretized system (11) can be written as

$$\begin{align*}
(1 + \sigma\Delta t) \sum_{i=0}^{\kappa-1} U_i^n (\phi_i, \phi_j) + \Delta t \sum_{i=0}^{\kappa-1} W_i^n (\nabla \phi_i, \nabla \phi_j) &= \sum_{i=0}^{\kappa-1} U_i^{n-1} (\phi_i, \phi_j) + \sigma\Delta t (m, \phi_j), \quad (47a) \\
\sum_{i=0}^{\kappa-1} W_i^n (\phi_i, \phi_j) &= \epsilon^2 \sum_{i=0}^{\kappa-1} U_i^n (\nabla \phi_i, \nabla \phi_j) + \left( \sum_{i=0}^{\kappa-1} U_i^n \phi_i \right)^3 - \sum_{i=0}^{\kappa-1} U_i^{n-1} \phi_i \phi_j. \quad (47b)
\end{align*}$$

We use the Newton method or V-N method to approximate the implicit nonlinear term in (47b). For simplicity, we show the implementation scheme of the Newton method for each $j (0 \leq j \leq \kappa - 1)$ below:

$$\begin{align*}
\left( \left( \sum_{i=0}^{\kappa-1} U_i^n \phi_i \right)^3, \phi_j \right) &\approx \left( 3 \sum_{i=0}^{\kappa-1} U_i^{n,k} \phi_i \left( \sum_{i=0}^{\kappa-1} U_i^{n,k-1} \phi_i \right)^2, \phi_j \right) - \left( 2 \sum_{i=0}^{\kappa-1} U_i^{n,k-1} \phi_i \right)^3, \phi_j \right) \\
&= \sum_{i=0}^{\kappa-1} U_i^{n,k} \left( 3 \int_{\Omega} \left( \sum_{i=0}^{\kappa-1} U_i^{n,k-1} \phi_i \right)^2 \phi_i \phi_j d\Omega \right) - 2 \int_{\Omega} U_i^{n,k-1} \phi_i \phi_j d\Omega. \quad (48)
\end{align*}$$

Define the mass and stiffness matrices as

$$M = (m_{ij}), \quad m_{ij} = (\phi_i, \phi_j) = \int_{\Omega} \phi_i \phi_j d\Omega,$$

$$S = (s_{ij}), \quad s_{ij} = (\nabla \phi_i, \nabla \phi_j) = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j d\Omega.$$
Note that $M > 0$ and $S \succeq 0$. For $0 \leq j \leq \kappa - 1$, let

$$L^{(n,k)} = (l_{ij}^{(n,k)}), l_{ij}^{(n,k)} = 3 \int_\Omega \left( \sum_{l=0}^{\kappa-1} u_j^{n,k-1} \phi_l \right)^2 \phi_j \, dx,$$

$$U_j^{(n-1)} = (U_0^{(n-1)}, U_1^{(n-1)}, \ldots, U_{\kappa-1}^{(n-1)})^T,$$

$$F^{(n)} = (F_0^{(n)}, \ldots, F_{\kappa-1}^{(n)})^T, \quad F_j^{(n)} = (MU_j^{n-1}) + \sigma \Delta t \mu_j,$$

$$E_j^{(n,k)} = (E_0^{(n,k)}, \ldots, E_{\kappa-1}^{(n,k)})^T, \quad E_j^{(n,k)} = -2 \int_\Omega u_j^{n,k-1} \phi_j \, dx.$$

$L^{(n,k)} \succeq 0$ depends on the previous iteration solution at each time step. The vectors $F^{(n)}$ and $E^{(n)}$ are obtained from the previous step. Using the approximation of (48), the scheme for (47) is summarized as follows:

$$d \begin{pmatrix} U_j^{(n,k)} \\ W_j^{(n,k)} \end{pmatrix} = \begin{pmatrix} \int_\Omega (1 + \sigma \Delta t) \mu_j \Delta t S & \int_\Omega U_j^{(n,k)} \\ -\varepsilon^2 S - L^{(n,k)} & M \end{pmatrix} \begin{pmatrix} U_j^{(n,k)} \\ W_j^{(n,k)} \end{pmatrix} = \begin{pmatrix} F_j^{(n)} \\ E_j^{(n,k)} \end{pmatrix}, \quad (49)$$

where

$$U_j^{n+1} = \lim_{k \to \infty} U_j^{(n,k)}, \quad W_j^{n+1} = \lim_{k \to \infty} W_j^{(n,k)}.$$ 

The starting conditions are $U_j^{(n,0)} = U_j^{(n-1)}$ and $W_j^{(n,0)} = W_j^{(n-1)}$.

Subsequently, it is required to solve the discretized linear problem (49) using linear solvers at each time step. Due to a fast increase of memory requirement and bad scaling properties for massively parallel problems, the direct solvers, like UMFPACK [30], MUMPS [31], or SuperLU-DIST [32], may be difficult to solve (49) efficiently. Hence we use iteration methods to address these problems. The linear system (49) becomes more ill-conditioned as the mesh is refined. For instance, in one dimension, Figure 1 and Table 1 give the spectral distribution and corresponding condition number for different subdivisions. The ill-conditioned system reduces the performance of linear solvers and impedes the convergence of nonlinear solvers, which is the difficulty in numerical computation.

To improve the condition number of the linear system and accelerate the convergence of nonlinear iteration, a natural selection is using preconditioning approach.

**Table 1** The condition number of the linear system (49) for one dimensional dynamic equation with different subdivision $\kappa$ when $\varepsilon = 0.1$, $\Delta t = 0.01$, $\sigma = 100$, $n = 1$ and $\kappa - 1$.

| $\kappa$ | 100 | 500 | 1000 | 2000 |
|----------|-----|-----|------|------|
| $\text{cond}(\mathcal{M})$ | 0.407e+03 | 1.009e+04 | 4.037e+04 | 1.624e+05 |

4.1 Schur complement preconditioner

From the above discussion, it is necessary to find appropriate preconditioners for solving the discretized system (49). In [21] and [24], two preconditioners have been proposed for Ohta-Kawasaki equation when using the backward Euler scheme. In this subsection, we will present Schur complement preconditioners for CSS similar to the construction technique in [24]. Moreover, we will analyze the spectral distribution of the preconditioning system.
Denote $\zeta = \Delta t / (1 + \sigma \Delta t)$. Using Schur complement method and ignoring the influence of $L^{(n,k)}$, the Schur block preconditioner (Schur in short) for (49) is
\[
\mathcal{P}_{\text{Schur}} = \begin{pmatrix}
M & \zeta S \\
-\varepsilon^2 S M + 2\varepsilon \sqrt{\zeta} S 
\end{pmatrix}.
\]

The spectral distribution for the preconditioning system is shown as follows.

**Theorem 5** The eigenvalues of $\mathcal{P}^{-1}_{\text{Schur}} \hat{A}$ are real and satisfy
\[
\lambda(\mathcal{P}^{-1}_{\text{Schur}} \hat{A}) \in \left(1/2, 1 + \frac{1}{4\varepsilon} \sqrt{\zeta} \lambda_+\right),
\]
with $\lambda_+ = \lambda_{\text{max}}(M^{-1}L^{(n,k)})$ and
\[
\hat{A} = \begin{pmatrix}
M \\
-\varepsilon^2 S - L^{(n,k)} M
\end{pmatrix}.
\]

**Proof** Notice that
\[
\hat{K} = (M + \varepsilon \sqrt{\zeta} S)M^{-1}(M + \varepsilon \sqrt{\zeta} S)
\]
is a good approximation of the Schur complement:
\[
K = M + \varepsilon^2 \zeta S M^{-1} S + \zeta L^{(n,k)} M^{-1} S.
\]

It is obvious that
\[
\mathcal{P}^{-1}_{\text{Schur}} \hat{A} = \begin{pmatrix}
I & M^{-1} B_1 \\
0 & I
\end{pmatrix}^{-1} \begin{pmatrix}
I & 0 \\
K^{-1} L^{(n,k)} & K^{-1} K
\end{pmatrix} \begin{pmatrix}
I & M^{-1} B_1 \\
0 & I
\end{pmatrix}.
\]

(50)
Then, it is required to prove

\[ \lambda(\tilde{K}^{-1}K) \in \left( \frac{1}{2}, 1 + \frac{1}{4\epsilon} \sqrt{\zeta} \lambda_+ \right) \]

and the corresponding eigenvalues are real.

Suppose \( v \) is the corresponding eigenvector of the eigenvalue \( \lambda(\tilde{K}^{-1}K) \), thus

\[ K_v = \lambda \tilde{K} v. \] (51)

i) If \( v \in \text{null}(S) \), by (51), then \( Mv = \lambda Mv \), which implies that \( \lambda = 1 \).

ii) If \( v \notin \text{null}(S) \), using (51), we have

\[ SM^{-1}K_v = \lambda SM^{-1}\tilde{K} v. \]

It follows that

\[ v^* F v = \lambda v^* G v, \] (52)

where

\[ F = SM^{-1}K = S + \epsilon^2 \zeta SM^{-1}S + \zeta SM^{-1}L(\alpha, \epsilon)M^{-1}S, \]

\[ G = SM^{-1}\tilde{K} = S + \epsilon^2 \zeta SM^{-1}S + 2\epsilon \sqrt{\zeta} SM^{-1}S. \]

As \( F \) and \( G \) are real symmetric, then \( v^* F v \) and \( v^* G v \) are real (with \( v^* G v \) also positive). The eigenvalues of \( \lambda(\tilde{K}^{-1}K) \) are real. Using (52), we get

\[ \lambda = \frac{v^* F v}{v^* G v} = \frac{v^* S v + \epsilon^2 \zeta v^* SM^{-1}S v + \zeta v^* SM^{-1}L(\alpha, \epsilon)M^{-1}S v}{v^* S v + \epsilon^2 \zeta v^* SM^{-1}S v + 2\epsilon \sqrt{\zeta} v^* SM^{-1}S v} = 1 - \frac{2\epsilon \sqrt{\zeta} v^* SM^{-1}S v}{v^* S v + 2\epsilon \sqrt{\zeta} v^* SM^{-1}S v + \epsilon^2 \zeta v^* SM^{-1}S v} \]

\[ + \frac{\zeta v^* SM^{-1}L(\alpha, \epsilon)M^{-1}S v}{v^* S v + 2\epsilon \sqrt{\zeta} v^* SM^{-1}S v + \epsilon^2 \zeta v^* SM^{-1}S v} \]

\[ = 1 - \mathcal{A}_1 + \mathcal{A}_2. \] (53)

It is evident that \( \mathcal{A}_1 > 0 \). Writing

\[ a = \frac{1}{2} S^1 v, \quad b = \epsilon \sqrt{\zeta} S^\frac{1}{2} M^{-1} S v, \]

combining with the inequality \( a^* a + b^* b \geq a^* b + b^* a \), it follows that

\[ \mathcal{A}_1 = (a^* b + b^* a)/(a^* a + b^* b + a^* b + b^* a) \leq \frac{1}{2}. \]

Therefore, we have

\[ \mathcal{A}_1 \in (0, \frac{1}{2}]. \] (54)

Noting

\[ \mathcal{A}_2 = \frac{\zeta}{\epsilon \sqrt{\zeta}} v^* SM^{-1}L(\alpha, \epsilon)M^{-1}S v \]

\[ = \zeta \mathcal{A}_{21} \mathcal{A}_{22}, \]

where

\[ \mathcal{A}_{21} = \frac{v^* S v + 2\epsilon \sqrt{\zeta} v^* SM^{-1}S v + \epsilon^2 \zeta v^* SM^{-1}S v}{v^* S v + 2\epsilon \sqrt{\zeta} v^* SM^{-1}S v + \epsilon^2 \zeta v^* SM^{-1}S v}. \]
then
\[ R_{22} = \frac{1}{2\varepsilon \sqrt{\zeta}} R_1, \]
which shows that
\[ R_{22} \in \left(0, \frac{1}{4\varepsilon \sqrt{\zeta}}\right]. \]
Writing
\[ \bar{v} = M^{-\frac{1}{2}} Sv, \]
we immediately have
\[ R_{21} = \bar{v}^* M^{-\frac{1}{2}} L^{(n,k)} M^{-\frac{1}{2}} \bar{v} \in \left[ \lambda_{\min}(M^{-1} L^{(n,k)}), \lambda_{\max}(M^{-1} L^{(n,k)}) \right], \]
where \( \lambda_{\min}(M^{-1} L^{(n,k)}) > 0 \). Combining with the derived bounds, it is obvious that
\[ R_2 \in \left(0, \frac{1}{4\varepsilon \sqrt{\zeta}} \lambda_+ \right]. \quad (55) \]
Substituting (54) and (55) into (53), it follows that
\[ \lambda(\tilde{K}^{-1}K) \in \left(\frac{1}{2}, 1 + \frac{1}{4\varepsilon \sqrt{\zeta}} \lambda_+ \right). \]
Combining this with (50), we complete the proof.

4.2 MHSS preconditioner

In this section, using the structure of the linear system (49), we propose a block preconditioner based on the MHSS and analyze the eigenvalue distribution of the preconditioning system.

The linear system (49) can be rearranged equivalently to the two-by-two block system:
\[
\begin{pmatrix}
\varepsilon^2 S + L^{(n,k)} & -M \\
M & \zeta S
\end{pmatrix}
\begin{pmatrix}
U^{(n,k)} \\
W^{(n,k)}
\end{pmatrix} =
\begin{pmatrix}
-E^{(n,k)} \\
\frac{1}{1+\sigma} F^{(n)}
\end{pmatrix} = b.
\] (56)

Let \( A = \varepsilon^2 S + L^{(n,k)} \), \( B = \zeta S \), then
\[ \bar{A} = \begin{pmatrix} A & -M \\ M & B \end{pmatrix}. \]

Denote \( \bar{\kappa} = \kappa - 1 \), thus \( A \in \mathbb{R}^{\bar{\kappa} \times \bar{\kappa}} \) and \( A \succ 0, B \in \mathbb{R}^{\bar{\kappa} \times \bar{\kappa}} \) and \( B \succeq 0 \). The rearranged linear system (56) is a generalized saddle point problem, which widely exists in scientific computing and numerical algebra.

In recent years, many works have been devoted to developing efficient preconditioners for the generalized saddle point problem, such as block diagonal and triangular preconditioners (33–34), matrix splitting preconditioners (35–37). Among the preconditioners, the HSS preconditioner is an efficient method to solve the generalized saddle point problem, originally developed by Bai et al. (36). However, directly applying the HSS method to (56) will result in inefficiency. As an improvement, we present an MHSS block preconditioner for the linear system (56).
Using the similar constructing technique as \cite{37}, for any constant \(\alpha > 0\), the MHSS block preconditioner \(P_{\text{MHSS}}\) is defined by:

\[
P_{\text{MHSS}} = \begin{pmatrix} \frac{1}{\alpha} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \alpha I - M \\ M & B \end{pmatrix}.
\] (57)

Let

\[
\mathcal{R} = P_{\text{MHSS}} - I = \begin{pmatrix} 0 & M - \frac{1}{\alpha} AM \\ 0 & 0 \end{pmatrix}.
\]

then

\[
\mathcal{I} = P_{\text{MHSS}} - \mathcal{R}.
\]

Decomposing \(P_{\text{MHSS}}\) as

\[
P_{\text{MHSS}} = \frac{1}{\alpha} \begin{pmatrix} A & 0 \\ 0 & \alpha I \end{pmatrix} \begin{pmatrix} I & 0 \\ \frac{1}{\alpha} M & I + \frac{1}{\alpha^2} \end{pmatrix} \begin{pmatrix} \alpha I - M \\ M I - \frac{1}{\alpha} \end{pmatrix} \begin{pmatrix} I - \frac{1}{\alpha} M \\ 0 \\ I \end{pmatrix},
\]

then we obtain

\[
\mathcal{T}(\alpha) = P^{-1}_{\text{MHSS}} \mathcal{R} = \begin{pmatrix} 0 & A^{-1}M - \frac{1}{\alpha} M + \frac{1}{\alpha^2} M(B + \frac{1}{\alpha} M^2)^{-1}(\frac{1}{\alpha} M^2 - MA^{-1}M) \\ 0 & B + \frac{1}{\alpha^2} M + \frac{1}{\alpha} M(B + \frac{1}{\alpha} M^2)^{-1}(\frac{1}{\alpha} M^2 - MA^{-1}M) \end{pmatrix}.
\] (58)

Writing

\[
D = B + \frac{1}{\alpha^2} M^2,
\]

\[
F = \frac{1}{\alpha} M^2 - MA^{-1}M,
\]

thus (58) turns to

\[
\mathcal{T}(\alpha) = \begin{pmatrix} 0 & A^{-1}M - \frac{1}{\alpha} M + \frac{1}{\alpha^2} MD^{-1}F \\ 0 & D^{-1}F \end{pmatrix}.
\] (60)

Therefore, we obtain the preconditioned system

\[
P^{-1}_{\text{MHSS}} \mathcal{I} = P^{-1}_{\text{MHSS}}(P_{\text{MHSS}} - \mathcal{R}) = I - P^{-1}_{\text{MHSS}} \mathcal{R}
\]

\[
= I - \mathcal{T}(\alpha)
\]

\[
= \begin{pmatrix} I - A^{-1}M + \frac{1}{\alpha} M - \frac{1}{\alpha^2} MD^{-1}F \\ 0 \\ I - D^{-1}F \end{pmatrix},
\] (61)

where \(D\) and \(F\) are defined by (59).

Now, we analyze the spectral distribution for the preconditioned matrix \(P^{-1}_{\text{MHSS}} \mathcal{I}\).

**Theorem 6** Let \(A \in \mathbb{R}^{\tilde{p} \times \tilde{p}}\), \(M \in \mathbb{R}^{\tilde{p} \times \tilde{p}}\) and \(M > 0\), \(B \in \mathbb{R}^{\tilde{p} \times \tilde{p}}\) and \(B \succeq 0\), \(\alpha\) be a positive constant. Then the preconditioned matrix \(P^{-1}_{\text{MHSS}} \mathcal{I}\) has at least \(\tilde{p}\) eigenvalues \(1\). The remaining eigenvalues \(\lambda\) are real and located in the positive interval

\[
\frac{\alpha c^2}{\theta_1(c^2 + \alpha \sigma_1)} \leq \lambda \leq \frac{\alpha (\sigma_1 \theta_1 + c_1^2)}{\theta_\sigma c^2},
\] (62)

where \(\sigma_\beta, c_\beta, \theta_\beta\) and \(\sigma_1, c_1, \theta_1\) are the minimum and maximum eigenvalues of \(B, M, A\), respectively. Furthermore, if \(A = \alpha I\), then all the eigenvalues of \(P^{-1}_{\text{MHSS}} \mathcal{I}\) are \(1\).
Proof From (61), it is evident that $\mathcal{P}^{-1}_{\text{HS}}$ has at least $\tilde{\rho}$ eigenvalues 1. The remaining are the eigenvalues of $I - D^{-1}F$. Let

$$I - D^{-1}F = D^{-1}(D - F) = (B + \frac{1}{\alpha} M^2)^{-1}(B + MA^{-1}M),$$

$\lambda$ and $u$ be the eigenvalue and corresponding eigenvector of $I - D^{-1}F$, thus

$$(B + MA^{-1}M)u = \lambda (B + \frac{1}{\alpha} M^2)u. \quad (63)$$

Premultiplying both sides of (63) by $u^*/(u^*u)$ gets

$$u^*(B + MA^{-1}M)u = \lambda u^*(B + \frac{1}{\alpha} M^2)u. \quad (64)$$

Define

$$d = \frac{u^*MA^{-1}M}{u^*u}, \quad \gamma = \frac{u^*Bu}{u^*u}, \quad \text{and} \quad \tau = \frac{u^*M^2u}{u^*u}, \quad (65)$$

then

$$\lambda = \frac{\alpha(\gamma + d)}{\alpha \gamma + \tau}. \quad (66)$$

Note that $A \succ 0$, $M \succ 0$ and $B \succeq 0$. Using (64), then

$$0 < \frac{\sigma_1}{\theta_1} \leq d \leq \frac{\sigma_1}{\theta_u}, \quad 0 < c_1^2 \leq \tau \leq c_1^2, \quad \text{and} \quad 0 \leq \gamma \leq \sigma_1. \quad (66)$$

Combining (65) with (66), we obtain (62).

Remark 2 Noting that the positive interval presented in Theorem 6 is not very tight. Therefore, choosing appropriate $\lambda$ could obtain a better spectral distribution. This is illustrated by a simple numerical experiment in Figure 2.

In the following, we present an approach to get $\alpha$ such that $\rho(\mathcal{F}(\alpha))$ as small as possible. However, it is difficult to compute $\rho(\mathcal{F}(\alpha))$ exactly. Instead, we give an upper bound of $\rho(\mathcal{F}(\alpha))$.

Theorem 7 Let $A \succ 0$, $M \succ 0 \in \mathbb{R}^{\tilde{\rho} \times \tilde{\rho}}$ and, $B \succeq 0 \in \mathbb{R}^{\tilde{\rho} \times \tilde{\rho}}$. Assume that $\mathcal{F}(\alpha)$ is defined in (60). Then, we get

$$\rho(\mathcal{F}(\alpha)) \leq \max_{1 \leq i \leq \tilde{\rho}} \left| 1 - \frac{\alpha}{\lambda_i(A)} \right| = \tilde{\sigma}(\alpha).$$

Proof By the definition of $\mathcal{F}(\alpha)$ in (60), it follows that

$$\rho(\mathcal{F}(\alpha)) = \rho(D^{-1}F),$$

where $D$ and $F$ are defined in (59). Note that

$$D^{-1}F = M^{-1}D^{-1}FM,$$
Fig. 2 The spectral distribution of the preconditioned linear system with $\tilde{M} = 1000$, $\epsilon = 0.001$, $dt = 0.001$, $\sigma = 10$, $n = 1$, $k = 1$, $\alpha = 0.001$.

where

$$\tilde{D} = \frac{1}{\alpha} I + M^{-1}BM^{-1}, \quad \tilde{F} = \frac{1}{\alpha} I - A^{-1}.$$  

Thus

$$\rho(\tilde{T}(\alpha)) = \rho(D^{-1}F) = \rho(\tilde{D}^{-1}\tilde{F}) \leq \|\tilde{D}^{-1}\|_2 \|\tilde{F}\|_2.$$  

Since $A > 0$, $M > 0$ and $B \geq 0$, we have

$$\|\tilde{D}^{-1}\|_2 = \left\| \left( \frac{1}{\alpha} I + M^{-1}BM^{-1} \right)^{-1} \right\|_2 = \max_{1 \leq i \leq \tilde{p}} \left| \frac{1}{\alpha} - \frac{1}{\tilde{\lambda}_i(M^{-1}BM^{-1})} \right| = \alpha,$$

and

$$\|\tilde{F}\|_2 = \left\| \frac{1}{\alpha} I - A^{-1} \right\|_2 = \max_{1 \leq i \leq \tilde{p}} \left| \frac{1}{\alpha} - \tilde{\lambda}_i(A^{-1}) \right|.$$

Hence,

$$\rho(\tilde{T}(\alpha)) = \rho(D^{-1}F) \leq \alpha \max_{1 \leq i \leq \tilde{p}} \left| \frac{1}{\alpha} - \tilde{\lambda}_i(A^{-1}) \right| = \max_{1 \leq i \leq \tilde{p}} \left| 1 - \frac{\alpha}{\tilde{\lambda}_i(A)} \right|.$$  

Remark 3 By Theorem 7 we can choose $\alpha$ as the optimal $\alpha_*$ to minimize $\tilde{\sigma}(\alpha)$. Therefore,

$$\alpha = \alpha_* = \arg \min_{\alpha} \tilde{\sigma}(\alpha) = \frac{2\tilde{\lambda}_1(A)\tilde{\lambda}_\beta(A)}{\tilde{\lambda}_1(A) + \tilde{\lambda}_\beta(A)}.$$  

Since $\alpha_*$ is related to the eigenvalues of $A$, it is expensive to compute the optimal parameter $\alpha_*$ for large scale matrix $A$. Therefore, in practical implementation, we need to make an approximation for $\alpha_*$, refer to [33]. The practical choice strategy for parameter $\alpha$ is

$$\alpha = \frac{\text{trace}(A)}{\rho} \quad \text{or} \quad \alpha = \frac{\text{trace}(M^4)}{\text{trace}(M^4A^{-1})}.$$
5 Numerical results

In this section, we offer several examples to support the theoretical analysis. The spatial domain $\Omega_h$ is uniformly discreted for $t_n = n\Delta t$ ($0 \leq n \leq N$). The piecewise linear functions are used as the basis functions. GMRES is adopted to solve the preconditioned linear system (49) in each iteration. The elapsed CPU time per nonlinear iteration process in seconds is denoted by ‘CPU’ and the average of GMRES iteration per Newton step by $IT_{GM}$. Let $k_n$ be the iteration steps of the $n$-th nonlinear iteration, and $IT_{tol}$ be the total iteration steps after $n$-th nonlinear iteration, i.e., $IT_{tol} = \sum_{i=1}^{n} k_i$. The nonlinear iteration error is defined as $\varepsilon = ||u^{(n,k)} - u^{(n,k-1)}||_2$. The stop criteria $\varepsilon$ is chosen as $10^{-6}$.

First, we demonstrate the performance of two Newton iteration methods. When the total number of time steps is $N = 100$, Tab. 2 compares two Newton methods against CPU and $IT_{tol}$ in one and two dimensional space with different degree of freedom (DOF). Obviously, the Newton method requires less CPU and $IT_{tol}$ than those of the V-N method. Furthermore, Fig. 3 gives the concrete iteration process when $\Omega = [0, 1]$, DOF = $10^4$, $\sigma = 100$, $\Delta t = \varepsilon^2$ and $\varepsilon = 0.01$. Note that the CPU time required for each nonlinear iteration step of two Newton methods is almost the same. Fig. 3(a) represents that the Newton method requires fewer nonlinear iteration steps than that of the V-N method for each time step $n$, as well as CPU time. When $n = 23$, Fig. 3(b) illustrates that the error $\varepsilon$ of the Newton method decreases faster than that of the V-N method, as the theory in Sec. 3 predicted. Similar phenomena also appear for other cases in Tab. 2.

Second, we demonstrate the efficiency of proposed preconditioners when solving linear system (49). As an one-dimensional example, Table 3 presents the performance of solving non-preconditioned and preconditioned systems in 500 time steps ($\Delta t = \varepsilon^2$) with different spatial DOF for domains $\Omega = [0, 1]$. From these results, it can be seen that two preconditioners accelerate computation, almost three times faster than solving non-preconditioned system in terms of CPU time.

Last, we observe the long-time behavior of the proposed method in the domain $\Omega = [0, \pi]^2$ with 1000 $\times$ 1000 triangular grids. Using random initial values, Figs. 4 and 5 show the coarsening dynamic process with different parameters. The final converged morphologies are lamellar and cylindrical phases, respectively. As shown in these figures, our method keeps energy dissipation as time evolves.

6 Conclusion

This paper presents a systematic numerical method to solve the mass-conserved Ohta-Kawasaki equation. An unconditionally energy stable scheme, the CSS, is used to discretize the time
variable, while the FEM to the spatial variable. Two Newton methods are applied to update the implicit nonlinear terms. To reduce the condition number of the discretized linear system, we propose Schur complement preconditioner and MHSS block preconditioner. The convergent rate of the two Newton iteration method has been proven to be the same, while the Newton method has a smaller convergent factor than the variant one. The spectral distribution of two block preconditioning linear systems has been analyzed. Numerical investigations provided sufficient support for the theoretical analysis and demonstrated the efficiency of the proposed numerical methods.

References

1. F. S. Bates, M. A. Hillmyer, T. P. Lodge, C. M. Bates, K. T. Delaney, G. H. Fredrickson, Multiblock polymers: Panacea or Pandora’s box? Science, 336, 434-440 (2012).
2. K. Jiang, C. Wang, Y. Q. Huang, P. W. Zhang, Discovery of new metastable patterns in diblock copolymers, Commun. Comput. Phys., 14, 443-460 (2013).
Numerical methods for the mass-conserved Ohta-Kawasaki equation

Fig. 4 The dynamic process in domain $[0, \pi]^2$ when $\varepsilon = 0.08$, $\sigma = 10$, $m = 0$ and $\Delta t = 0.01$ with a random initial value. (B1), (B2), (B3) and (B4) correspond to $t = 0, 2, 4, 13.43$.

Fig. 5 The dynamic process in domain $[0, \pi]^2$ when $\varepsilon = 0.08$, $\sigma = 10$, $m = 0.4$ and $\Delta t = 0.0064$ with a random initial value. (T1), (T2), (T3) and (T4) correspond to $t = 0, 0.448, 1.344, 2.224$.

3. L. Leibler, Theory of microphase separation in block copolymer, Macromolecules, 13, 1602-1617 (1980).
4. R. Guo, Y. Xu, Local discontinuous Galerkin method and high order semi-implicit scheme for the phase field crystal equation, SIAM J. Sci. Comput., 38(1), A105–A127 (2016).
5. R. Guo, Y. Xu, A high order adaptive time-stepping strategy and local discontinuous Galerkin method for the modified phase field crystal equation, Commun. Comput. Phys., 24, 123-151 (2018).
6. G. H. Fredrickson, The equilibrium theory of inhomogeneous polymers, Oxford University Press: New York (2006).
7. T. Ohta, K. Kawasaki, Equilibrium morphology of block copolymer melts, Macromolecules, 19, 2621-2632 (1986).
8. R. Choksi, X. Ren, On the derivation of a density functional theory for microphase separation of diblock copolymers, J. Stat. Phys, 113, 151-176 (2002).
9. D. Eyre, Unconditionally gradient stable time marching the Cahn-Hilliard equation, Computational and Mathematical Models of Microstructural Evolution, 529, 39-46 (1998).
10. C. Xu, T. Tang, Stability analysis of large time-stepping methods for epitaxial growth models, SIAM J. Numer. Anal., 44, 1759-1779 (2006).
11. J. Shen, J. Xu, J. Yang, The scalar auxiliary variable (SAV) approach for gradient flows, J. Comput. Phys., 353, 407-416 (2018).
12. Q. Du, W.X. Zhu, Stability analysis and application of the exponential time differencing schemes, J. Comput. Math., 22, 200-209 (2004).
13. W. Hackbusch, On the fast solution of nonlinear elliptic equations, Numer. Math., 32, 83-95 (1979).
14. W. T. Li, C. F. Wei, An efficient finite element procedure for analyzing three-phase porous media based on the relaxed Picard method, Int. J. Numer. Methods Eng., 101, 825-846 (2015).
15. K. Junseok, S. Jaemin, An unconditionally gradient stable numerical method for the Ohta-Kawasaki model, J. Korean Math. Soc., 54, 145-158 (2017).
16. X. X. Dai, X. L. Cheng, A two-grid method based on Newton iteration for the Navier-Stokes equations, J. Comput. Appl. Math., 220, 566-573 (2008).
17. S. M. Wise, Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn-Hilliard-Hele-Shaw system of equations, J. Sci. Comput., 44, 38-68 (2010).
18. Z. Hu, S. M. Wise, C. Wang, J. S. Lowengrub, Stable and Efficient Finite-difference Nonlinear-multigrid Schemes for the Phase-field Crystal Equation, J. Comput. Phys., 228, 5323-5339 (2009).
19. X. Xu, Y. X. Zhao, Energy stable semi-implicit schemes for Allen-Cahn-Ohta-Kawasaki model in Binary System, J. Sci. Comput., 80, 1656-1680 (2019).
20. W. Feng, Z. Guan, J. Lowengrub, C. Wang, S. Wise, Y. Chen, A uniquely solvable, energy stable numerical scheme for the functionalized Cahn-Hilliard equation and its convergence analysis, J. Sci. Comput., 76, 1938-1967 (2018).
21. P. E. Farrell, J. W. Pearson, A preconditioner for the Ohta-Kawasaki equation, SIAM J. Matrix Anal. Appl., 38, 217-225 (2017).
22. J. C. Xu, Y. K. Li, S. N. Wu, On the stability and accuracy of partially and fully implicit schemes for phase field modeling, Comput. Method Appl. M., 345, 829-853 (2019).
23. J. Bosch, M. Stoll, P. Benner, Fast solution of Cahn-Hilliard variational inequalities using implicit time discretization and finite elements, J. Comput. Phys., 262, 38-57 (2014).
24. R. X. Li, Z. Z. Liang, G. F. Zhang, L. D. Liao, A note on preconditioner for the Ohta-Kawasaki equation, Appl. Math. Lett., 85, 132-138 (2018).
25. P. Quentin, Numerical approximation of the Ohta-Kawasaki functional, Master’s thesis, University of Oxford, Oxford, UK (2012).
26. Y. N. He, Y. Zhang, H. Xu, Two-level Newton’s method for nonlinear elliptic PDEs, J. Sci. Comput., 57, 124-145 (2013).
27. J. C. Xu, Two-grid discretization techniques for linear and nonlinear PDEs, SIAM J. Numer. Anal., 33, 1759-1777 (1996).
28. R. Rannacher, R. Scott, Some optimal error estimates for piecewise linear finite element approximation, Math. Comp., 54, 437-445 (1982).
29. A. Schatz, An observation concerning Ritz-Galerkin methods with indefinite bilinear forms, Math. Comp., 28, 959-962 (1974).
30. T. A. Davis, Algorithm 832: UMFPACK V4.3-an unsymmetric-pattern multifrontal method, ACM Trans. Math. Software, 30, 196-199 (2004).
31. P. R. Amestoy, I. S. Duff, J. Y. L. Excellent, J. Koster, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM J. Matrix Anal. Appl., 23, 15-41 (2001).
32. X. S. Li, J. W. Demmel, SuperLU DIST: A scalable distributed-memory sparse direct solver for unsymmetric linear systems, ACM Trans. Math. Software, 29, 110-140 (2003).
33. O. Axelsson, Unified analysis of preconditioning methods for saddle point matrices, Numer. Linear Algebra Appl., 22, 233-253 (2015).
34. M. Q. Jiang, Y. Cao, L. Q. Yao, On parameterized block triangular preconditioners for generalized saddle point problems, Appl. Math. Comput., 216, 1777-1789 (2010).
35. J. Zhang, C. Gu, A variant of the deteriorated PSS preconditioner for nonsymmetric saddle point problems, BIT Numer. Math., 56, 587-604 (2016).
36. Z. Z. Bai, G. H. Golub, M. K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, SIAM J. Matrix Anal. Appl., 24, 603-626 (2003).
37. L. D. Liao, G. F. Zhang, A generalized variant of simplified HSS preconditioner for generalized saddle point problems, Appl. Math. Comput., 346, 790-799 (2019).
38. Y. M. Huang, A practical formula for computing optimal parameters in the HSS iteration methods, J. Comput. Appl. Math., 255, 142-149 (2014).
39. K. Jiang, S. F. Li, J. Zhang, A modified Hermitian and skew-Hermitian preconditioner for the Ohta-Kawasaki equation, arXiv:1910.09297.
40. E. Suli, Numerical solution of ordinary differential equations, Mathematical Institute, University of Oxford (2010).

Appendix A. Proof of Theorem 1

Proof Using the definition of $E(u)$ in (1), then

$$E(u^n) - E(u^{n-1}) = \frac{1}{2} \varepsilon^2 (\|\nabla u^n\|^2 - \|\nabla u^{n-1}\|^2) + \Phi(u^n) - \Phi(u^{n-1}),$$

$$+ \frac{1}{2} \sigma \left(\|(-\Delta)^{-1}(u^n - m)\|^2 - \|(-\Delta)^{-1}(u^{n-1} - m)\|^2\right). (67)$$
Using Taylor expansion, we have
\[
\Phi_+(u^n) - \Phi_+(u^n) = \Phi'_+(u^n)(u^{n-1} - u^n) + \frac{\Phi''_+(u^n)}{2} (u^{n-1} - u^n)^2,
\]
\[
\Phi_-(u^n) - \Phi_-(u^{n-1}) = \Phi'_-(u^{n-1})(u^n - u^{n-1}) + \frac{\Phi''_-(u^{n-1})}{2} (u^n - u^{n-1})^2.
\]

Thus,
\[
\Phi(u^n) - \Phi(u^{n-1}) = - \left( \Phi_+(u^{n-1}) - \Phi_+(u^n) \right) - \left( \Phi_-(u^n) - \Phi_-(u^{n-1}) \right)
\]
\[
= (\Phi'_+(u^n) - \Phi'_+(u^{n-1}))(u^n - u^{n-1}) - \frac{1}{2}(\Phi''_+(u^{n-1}) + \Phi''_-(u^{n-1}))(u^n - u^{n-1})^2.
\]

(68)

Combing the identity
\[
a^2 - b^2 = 2a(a - b)
\]

with (68), (67) is equivalent to
\[
E(u^n) - E(u^{n-1})
\]
\[
= e^2(\nabla u^n, \nabla (u^n - u^{n-1})) - \frac{1}{2} e^2 \| \nabla (u^n - u^{n-1}) \|^2
\]
\[
+ (\Phi'_+(u^n) - \Phi'_+(u^{n-1}), u^n - u^{n-1}) - \frac{1}{2}(\Phi''_+(u^{n-1}) + \Phi''_-(u^{n-1}))(u^n - u^{n-1})^2
\]
\[
+ \sigma((\Delta)^{-\frac{1}{2}}(u^n - m), (\Delta)^{-\frac{1}{2}}(u^n - u^{n-1})) - \frac{1}{2}\sigma((\Delta)^{-\frac{1}{2}}(u^n - u^{n-1}))
\]
\[
= -e^2 \Delta u^n, u^n - u^{n-1}) - \frac{1}{2} e^2 \| \nabla (u^n - u^{n-1}) \|^2
\]
\[
+ \Phi'_+(u^n) - \Phi'_+(u^{n-1}), u^n - u^{n-1}) - \frac{1}{2}(\Phi''_+(u^{n-1}) + \Phi''_-(u^{n-1}))(u^n - u^{n-1})^2
\]
\[
+ \sigma((\Delta)^{-\frac{1}{2}}(u^n - m), u^n - u^{n-1}) - \frac{1}{2}\sigma((\Delta)^{-\frac{1}{2}}(u^n - u^{n-1}))
\]
\[
= -e^2 \Delta u^n + \Phi'_+(u^n) - \Phi'_+(u^{n-1}) + \sigma((\Delta)^{-1}(u^n - m), u^n - u^{n-1})
\]
\[
- \frac{1}{2} e^2 \| \nabla (u^n - u^{n-1}) \|^2 - \frac{1}{2}(\Phi''_+(u^{n-1}) + \Phi''_-(u^{n-1}))(u^n - u^{n-1})^2
\]
\[
- \frac{1}{2}\sigma((\Delta)^{-\frac{1}{2}}(u^n - u^{n-1}))
\]

(69)

Taking the test function \( v = u^n - u^{n-1} \) in (70), then
\[
\left( \frac{u^n - u^{n-1}}{\Delta t}, u^n - u^{n-1} \right) + \left( \nabla w^n, \nabla (u^n - u^{n-1}) \right) + \sigma(u^n - m, u^n - u^{n-1}) = 0,
\]
\[
(w^n, u^n - u^{n-1}) = e^2 (\nabla u^n, \nabla (u^n - u^{n-1})) - (\Phi'_+(u^n) - \Phi'_+(u^{n-1}), u^n - u^{n-1}) = 0.
\]

(70a)

(70b)

Putting
\[
\mu^n = w^n + \sigma((\Delta)^{-1}(u^n - m))
\]
into \((70a)\), rearranging \((70b)\), thus
\[
\begin{align*}
\frac{u^n - u^{n-1}}{\Delta t}, u^n - u^{n-1} &= (\Delta \mu^n, u^n - u^{n-1}), \\
(w^n, u^n - u^{n-1}) &= -e^2(\Delta u^n, u^n - u^{n-1}) + (\Phi'_+(u^n) - \Phi'_-(u^n-1), u^n - u^{n-1}).
\end{align*}
\] (71a) (71b)

Note that
\[\Phi_+(u) \geq 0, \quad \Phi_-(u) \geq 0.\]

Substituting \((71a)\) and \((71b)\) into \((69)\) yields
\[
\begin{align*}
E(u^n) - E(u^{n-1}) &= (\mu^n, u^n - u^{n-1}) - \frac{1}{2}e^2\|\nabla(u^n - u^{n-1})\|^2 \\
&\quad - \frac{1}{2}(\Phi'_+(\xi^{n-1}) + \Phi''_-(\eta^{n-1}), (u^n - u^{n-1})^2) - \frac{1}{2}\sigma\|(-\Delta)^{-\frac{1}{2}}(u^n - u^{n-1})\|^2.
\end{align*}
\]

By the definition of \(\mu^n\) and let the test function \(v = \mu^n\) in \((\theta)\), we can obtain
\[
(\mu^n, u^n - u^{n-1}) = \Delta t(\mu^n, \Delta \mu^n),
\]
then
\[
\begin{align*}
E(u^n) - E(u^{n-1}) &= \Delta t(\mu^n, \Delta \mu^n) - \frac{1}{2}e^2\|\nabla(u^n - u^{n-1})\|^2 \\
&\quad - \frac{1}{2}(\Phi'_+(\xi^{n-1}) + \Phi''_-(\eta^{n-1}), (u^n - u^{n-1})^2) - \frac{1}{2}\sigma\|(-\Delta)^{-\frac{1}{2}}(u^n - u^{n-1})\|^2 \leq 0,
\end{align*}
\]
which implies \((10)\). Thus we complete the proof of Theorem \(1\).

Appendix B. Proof of Theorem \(2\).

Lemma 4 \([40]\) (Poincaré inequality) Let \(\Omega\) be a bounded, connected, open subset of \(\mathbb{R}^d\) with boundary \(\partial \Omega\) of \(\mathbb{C}^1\). Then there exists a constant \(c\), depending only on \(d\) and \(\Omega\) such that
\[
\left\|v - \int_{\Omega} v\,dx\right\| \leq c\|\nabla v\|.
\]

Definition 1 \([25]\) For \(u \in H^1(\Omega)\), we define \(\Delta_h u \in V_h\) such that
\[
(-\Delta_h u, v_h) = (\nabla u, \nabla v_h), \quad \forall v_h \in V_h,
\]
where the subscript ‘\(h\)’ means that \(\Delta_h u\) depends on the discretization mesh.

Proof of Theorem \(2\)
Proof From Poincaré inequality (Lemma 4), we have
\[
\|u_h^n - m\| = \left\| u_h^n - \int_{\Omega} u_h^n \, dx \right\| \leq c \| \nabla u_h^n \|.
\]
Then
\[
\|u_h^n - m|\Omega \| \leq \|u_h^n - m\| \leq c \| \nabla u_h^n \|. \tag{72}
\]
Set \( v_h = u_h^n \) in (11a), then
\[
\frac{(u_h^n - u_h^{n-1})}{\Delta t} + (\nabla w_h^n, \nabla u_h^n) + \sigma(u_h^n - m, w_h^n) = 0. \tag{73}
\]
Taking \( v_h = (u_h^n - u_h^{n-1})/\Delta t \) in (11b) yields
\[
(w_h^n, \frac{u_h^n - u_h^{n-1}}{\Delta t}) = \varepsilon^2 (\nabla u_h^n, \frac{u_h^n - u_h^{n-1}}{\Delta t}) + (\Phi'_+(u_h^n) - \Phi'_-(u_h^{n-1}), \frac{u_h^n - u_h^{n-1}}{\Delta t}). \tag{74}
\]
Subtracting (73) from (74), using the symmetry in the inner product, thus
\[
e^2 \left( \nabla u_h^n, \frac{\nabla u_h^n - \nabla u_h^{n-1}}{\Delta t} \right) + \left( \Phi'_+(u_h^n) - \Phi'_-(u_h^{n-1}), \frac{u_h^n - u_h^{n-1}}{\Delta t} \right) + \| \nabla w_h^n \|^2 + \sigma(u_h^n - m, w_h^n) = 0. \tag{75}
\]
Note that
\[
\nabla u_h^n = \frac{\nabla u_h^n - \nabla u_h^{n-1}}{2} + \frac{\nabla u_h^n + \nabla u_h^{n-1}}{2}. \tag{76}
\]
Then (75) is equivalent to
\[
\frac{\varepsilon^2}{2\Delta t} \left\{ (\nabla (u_h^n - u_h^{n-1}), \nabla (u_h^n - u_h^{n-1})) + (\nabla u_h^n + \nabla u_h^{n-1}, \nabla u_h^n - \nabla u_h^{n-1}) \right\}
+ \left( \Phi'_+(u_h^n) - \Phi'_-(u_h^{n-1}), \frac{u_h^n - u_h^{n-1}}{\Delta t} \right) + \| \nabla w_h^n \|^2 + \sigma(u_h^n - m, w_h^n) = 0. \tag{76}
\]
The sequence \( u_h^n \) keeps mass conservation in (25), i.e.,
\[
(u_h^n - m, 1) = 0 \text{ for } n = 0, 1, \cdots, N. \tag{77}
\]
Multiplying (77) by \( \int_{\Omega} w_h^n \, dx \) gets
\[
(u_h^n - m, \int_{\Omega} w_h^n \, dx) = 0.
\]
Hence,
\[
(u_h^n - m, w_h^n) = (u_h^n - m, w_h^n - \int_{\Omega} w_h^n \, dx). \tag{78}
\]
Combining (76) with (78), we have
\[
\frac{\varepsilon^2}{2\Delta t} \left\{ \| \nabla (u_h^n - u_h^{n-1}) \|^2 + \| \nabla u_h^n \|^2 - \| \nabla u_h^{n-1} \|^2 \right\}
+ \| \nabla w_h^n \|^2 + \sigma(u_h^n - m, w_h^n - \int_{\Omega} w_h^n \, dx) = 0. \tag{79}
\]
Using Taylor expansion for $\Phi(t)$ at $a$, then

$$
\Phi'(a)(a - b) = \Phi(a) - \Phi(b) + \frac{1}{2} \Phi''(\eta)(b - a)^2, \quad \eta \in (a, b).
$$

Notice that

$$
\Phi'_+(u) = a^3, \quad \Phi'_-(u) = u, \quad \Phi''_+(u) = 3a^2 > -1, \quad \Phi''_-(u) = 1.
$$

Hence,

$$
(\Phi_+(a^n_k), 1) - (\Phi_+(a^{n-1}_k), 1) - \frac{1}{2} \|a^n_k - a^{n-1}_k\|^2 < (\Phi'_+(a^n_k), a^n_k - a^{n-1}_k), \quad (80a)
$$

$$
(\Phi_-(a^{n-1}_k), 1) - (\Phi_-(a^n_k), 1) + \frac{1}{2} \|a^n_k - a^{n-1}_k\|^2 = (\Phi'_-(a^{n-1}_k), a^n_k - a^{n-1}_k). \quad (80b)
$$

Substituting $(80a)$ and $(80b)$ in $(79)$, multiplying by $2\Delta t$, then

$$
e^2 \|\nabla u^n_k\|^2 + e^2 \|\nabla (a^n_k - a^{n-1}_k)\|^2 + 2(\Phi_+(a^n_k), 1) + 2(\Phi_-(a^{n-1}_k), 1) + 2\Delta t \|\nabla w^n_k\|^2
$$

$$
< e^2 \|\nabla u^n_k\|^2 + 2(\Phi_+(a^n_k), 1) + 2(\Phi_-(a^{n-1}_k), 1) + 2\Delta t \|\nabla w^n_k\|^2
$$

$$
< e^2 \|\nabla u^n_k\|^2 + 2(\Phi_+(a^n_k), 1) + 2(\Phi_-(a^{n-1}_k), 1) + 2\Delta t \|\nabla w^n_k\|^2.
$$

Applying the inequality $2ab \leq a^2 + b^2$ to the last term on the right-hand side of $(81)$, subtracting $\Delta t \|\nabla w^n_k\|^2$ from both sides of $(81)$, then

$$
e^2 \sum_{n=1}^{k} \|\nabla u^n_k\|^2 + e^2 \sum_{n=1}^{k} \|\nabla (a^n_k - a^{n-1}_k)\|^2 + 2 \sum_{n=1}^{k} (\Phi_+(a^n_k), 1) + 2 \sum_{n=1}^{k} (\Phi_-(a^{n-1}_k), 1) + \Delta t \sum_{n=1}^{k} \|\nabla w^n_k\|^2
$$

$$
< e^2 \sum_{n=1}^{k} \|\nabla u^n_k\|^2 + 2 \sum_{n=1}^{k} (\Phi_+(a^n_k), 1) + 2 \sum_{n=1}^{k} (\Phi_-(a^{n-1}_k), 1) + \sigma^2 c^4 \Delta t \sum_{n=1}^{k} \|\nabla u^n_k\|^2.
$$

Summing $(82)$ over $n = 1, 2, \cdots, k \ (k \leq N)$ yields

$$
e^2 \sum_{n=1}^{k} \|\nabla u^n_k\|^2 + e^2 \sum_{n=1}^{k} \|\nabla (a^n_k - a^{n-1}_k)\|^2 + 2 \sum_{n=1}^{k} (\Phi_+(a^n_k), 1) + 2 \sum_{n=1}^{k} (\Phi_-(a^{n-1}_k), 1) + \Delta t \sum_{n=1}^{k} \|\nabla w^n_k\|^2
$$

$$
< e^2 \sum_{n=1}^{k} \|\nabla u^n_k\|^2 + 2 \sum_{n=1}^{k} (\Phi_+(a^n_k), 1) + 2 \sum_{n=1}^{k} (\Phi_-(a^{n-1}_k), 1) + \sigma^2 c^4 \Delta t \sum_{n=1}^{k} \|\nabla u^n_k\|^2.
$$

Canceling the same terms on each side of $(83)$, then

$$
e^2 \|\nabla u^n_k\|^2 + e^2 \sum_{n=1}^{k} \|\nabla (a^n_k - a^{n-1}_k)\|^2 + 2(\Phi_+(a^n_k), a^n_k - a^{n-1}_k) + \Delta t \sum_{n=1}^{k} \|\nabla w^n_k\|^2
$$

$$
< e^2 \|\nabla u^n_k\|^2 + 2(\Phi_+(a^n_k), a^n_k - a^{n-1}_k) + \sigma^2 c^4 \Delta t \sum_{n=1}^{k} \|\nabla u^n_k\|^2.
$$

For $k = 1, 2, \cdots, N$, note that

$$
\Phi(u^n_k) = \Phi_+(a^n_k) - \Phi_-(a^n_k) \geq 0.
$$
Hence, (84) reduces to
\[ \varepsilon^2 \| \nabla u_h^n \|^2 < \varepsilon^2 \| \nabla u_h^0 \|^2 + 2(\Phi_+(u_h^0) - \Phi_-(u_h^0), 1) + \sigma^2 \varepsilon^4 \Delta t \sum_{n=1}^k \| \nabla u_h^n \|^2. \]

Thus,
\[ (\varepsilon^2 - \sigma^2 \varepsilon^4 \Delta t) \| \nabla u_h^n \|^2 < \varepsilon^2 \| \nabla u_h^0 \|^2 + 2(\Phi_+(u_h^0) - \Phi_-(u_h^0), 1) + \sigma^2 \varepsilon^4 \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2. \quad (85) \]

If we suppose
\[ \frac{\varepsilon^2}{2} \leq \varepsilon^2 - \sigma^2 \varepsilon^4 \Delta t, \]
by (85), we get
\[ \frac{\varepsilon^2}{2} \| \nabla u_h^n \|^2 < \varepsilon^2 \| \nabla u_h^0 \|^2 + 2(\Phi_+(u_h^0) - \Phi_-(u_h^0), 1) + \sigma^2 \varepsilon^4 \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2. \]

It is equivalent to
\[ \| \nabla u_h^n \|^2 < 2 \| \nabla u_h^0 \|^2 + 4 \frac{\varepsilon^2}{\varepsilon^2} (\Phi(u_h^0), 1) + \frac{2 \sigma^2 \varepsilon^4}{\varepsilon^2} \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2 \]
\[ = \mathcal{K}^* + \mathcal{L}^* \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2, \quad (86) \]
where
\[ \mathcal{K}^* = 2 \| \nabla u_h^0 \|^2 + 4 \frac{\varepsilon^2}{\varepsilon^2} (\Phi(u_h^0), 1) \geq 0, \quad \mathcal{L}^* = \frac{2 \sigma^2 \varepsilon^4}{\varepsilon^2} \geq 0. \]

Denote
\[ \kappa_k = \mathcal{K}^* + \mathcal{L}^* \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2. \]

Using (86), then
\[ \hat{\kappa}_{k+1} - \hat{\kappa}_k = \mathcal{L}^* \Delta t \| \nabla u_h^k \|^2 \]
\[ < \mathcal{L}^* \Delta t (\mathcal{K}^* + \mathcal{L}^* \Delta t \sum_{n=1}^{k-1} \| \nabla u_h^n \|^2) \]
\[ = \mathcal{L}^* \Delta t \mathcal{K}_k. \]

Hence, we have
\[ \hat{\kappa}_{k+1} < (1 + \mathcal{L}^* \Delta t) \hat{\kappa}_k < \cdots < (1 + \mathcal{L}^* \Delta t)^k \hat{\kappa}_1 \leq e^{\mathcal{L}^* T} \mathcal{K}^*. \]

Thus, for \( \hat{k} = 1, 2, \cdots, N \), combining with (86), we get
\[ \| \nabla u_h^{\hat{k}+1} \|^2 < \hat{\kappa}_{\hat{k}+1} < e^{\mathcal{L}^* T} \mathcal{K}^*. \]
Moreover, using (86), when $k = 1$, then

$$\|\nabla u_h^k\|^2 < \mathcal{K}^* \leq e^{\mathcal{K}^* T} \mathcal{K}^*.$$ 

Therefore, for $n = 1, 2, \ldots, N$,

$$\|\nabla u_h^n\|^2 < C_2(e, \sigma, u_0, c, T).$$ 

(87)

Substituting (57) into (2) yields

$$\|u_h^n\| \leq C(e, \sigma, u_0, m, c, T, |\Omega|), \quad n = 1, 2, \ldots, N.$$ 

(88)

Since $\Phi'_1(u_h^n) = (u_h^n)^3$ for $n = 1, 2, \ldots, N$ and by Sobolev’s inequality (see Theorem 3 in [25]), it follows

$$\|\Phi'_1\| = \|(u_h^n)^3\| = \|u_h^n\|^3_{L^2(\Omega)} \leq c_3\|u_h^n\|^3_{H^1(\Omega)}.$$ 

Now, we put $v_h = \Delta_h u_h^n$ in (11b) and have

$$(w_h^n, \Delta_h u_h^n) = e^2(\nabla u_h^n, \nabla \Delta_h u_h^n) + (\Phi'_1(u_h^n) - \Phi'_1(u_h^{n-1}), \Delta_h u_h^n)$$

$$= -e^2(\Delta_h u_h^n, \Delta_h u_h^n) + (\Phi'_1(u_h^n) - \Phi'_1(u_h^{n-1}), \Delta_h u_h^n),$$

thus

$$e^2\|\Delta_h u_h^n\|^2 = -(w_h^n, \Delta_h u_h^n) + (\Phi'_1(u_h^n) - \Phi'_1(u_h^{n-1}), \Delta_h u_h^n)$$

$$\leq \|w_h^n\|\|\Delta_h u_h^n\| + \|\Phi'_1(u_h^n)\|\|\Delta_h u_h^n\| + \|\Phi'_1(u_h^{n-1})\|\|\Delta_h u_h^n\|,$$

i.e.,

$$e^2\|\Delta_h u_h^n\|^2 \leq \|w_h^n\| + \|\Phi'_1(u_h^n)\| + \|\Phi'_1(u_h^{n-1})\|$$

$$\leq \|w_h^n\| + c_3\|u_h^n\|^3_{H^1(\Omega)} + \|u_h^{n-1}\|.$$ 

(89)

Moreover, we take $v_h = w_h^n$ in (11b) and obtain

$$\|w_h^n\|^2 = e^2(\nabla u_h^n, \nabla w_h^n) + (\Phi'_1(u_h^n) - \Phi'_1(u_h^{n-1}), w_h^n)$$

$$\leq e^2\|\nabla u_h^n\|\|\nabla w_h^n\| + (\|\Phi'_1(u_h^n)\| + \|\Phi'_1(u_h^{n-1})\|)\|w_h^n\|,$$

then

$$\|w_h^n\| \leq e^2\|\nabla u_h^n\| + \|\Phi'_1(u_h^n)\| + \|\Phi'_1(u_h^{n-1})\|.$$ 

(90)

Hence, combining (90) with (89), we have

$$e^2\|\Delta_h u_h^n\|^2 \leq e^2\|\nabla u_h^n\|^2 + 2c_3\|u_h^n\|^3_{H^1(\Omega)} + 2\|u_h^{n-1}\|$$

$$\leq e^2\|\nabla u_h^n\|^2 + 2c_3(\|\nabla u_h^n\|^2 + \|u_h^n\|^3 + 2\|u_h^{n-1}\|).$$

From (87) and (88), we obtain

$$\|\Delta_h u_h^n\| \leq C(e, \sigma, u_0, m, c, c_5, T, |\Omega|), \quad n = 1, 2, \ldots, N.$$

Assume that the finite element triangulation is quasi-uniform, it follows

$$\|u_h^n - \int_{\Omega} u_h^n d\Omega\|_{L^\infty(\Omega)} \leq C\|u_h^n\|^{1-\theta}\|\Delta_h u_h^n\|^\theta.$$ 

Thus we deduce that

$$\|u_h^n\|_{1,\infty} \leq C(e, \sigma, u_0, m, c, c_5, T, |\Omega|), \quad n = 1, 2, \ldots, N,$$

Hence we complete the proof of Theorem 2.