Abstract. In this paper we present PDE and finite element analyses for a system of partial differential equations (PDEs) consisting of the Darcy equation and the Cahn-Hilliard equation, which arises as a diffuse interface model for the two phase Hele-Shaw flow. In the model the two sets of equations are coupled through an extra phase induced force term in the Darcy equations and a fluid induced transport term in the Cahn-Hilliard equation. We propose a fully discrete implicit finite element method for approximating the PDE system, which consists of the implicit Euler method combined with a convex splitting energy strategy for the temporal discretization, the standard finite element discretization for the pressure and a split (or mixed) finite element discretization for the fourth order Cahn-Hilliard equation. It is shown that the proposed numerical method satisfies a mass conservation law in addition to a discrete energy law that mimics the basic energy law for the Darcy-Cahn-Hilliard phase field model and holds uniformly in the phase field parameter $\varepsilon$. With help of the discrete energy law, we first prove that the fully discrete finite method is unconditionally energy stable and uniquely solvable at each time step. We then show that, using the compactness method, the finite element solution has an accumulation point that is a weak solution of the PDE system. As a result, the convergence result also provides a constructive proof of the existence of global-in-time weak solutions to the Darcy-Cahn-Hilliard phase field model in both two and three dimensions. Finally, we propose a nonlinear multigrid iterative algorithm to solve the finite element equations at each time step. Numerical experiments based on the overall solution method of combining the proposed finite element discretization and a nonlinear multigrid solver are presented to validate the theoretical results and to show the effectiveness of the proposed fully discrete finite element method for approximating the Darcy-Cahn-Hilliard phase field model.

Key words. Two phase Hele-Shaw flow, diffuse interface model, Darcy law, Cahn-Hilliard equation, energy splitting, finite element method, nonlinear multigrid.

AMS subject classifications. 65M60, 35K55, 76D05

1. Introduction. Hele-Shaw flow refers to the motion of (one or more) viscous fluids between two flat parallel plates separated by an infinitesimally small gap. Such a physical setup is often called a Hele-Shaw cell and was originally designed by Hele-Shaw to study two dimensional potential flows [17]. Various fluid mechanics problems can be approximated by Hele-Shaw flows and thus the research of those flows is of great practical importance. In addition, the relative simplicity of the governing equations of these flows makes Hele-Shaw flows ideal test cases in which rigorous mathematical theory and efficient numerical methods can be developed for studying interfacial dynamics — such as the formation of singularities and topological changes — in immiscible fluids (cf. [20, 21, 24] and the references therein).

The governing equation of Hele-Shaw flow is identical to that of the inviscid potential flow and to the flow of fluids through a porous medium, because the gap-averaged velocity of the fluid is given by Darcy’s law. Specifically, the two phase
Hele-Shaw flow takes the form (cf. [20] and the references therein):

\begin{align}
\mathbf{u} &= -\frac{1}{12\eta}(\nabla p - \rho g), \quad \text{div } \mathbf{u} = 0 \quad \text{in } \Omega_T \setminus \Gamma_t, \\
[p] &= \gamma \kappa, \quad [\mathbf{u} \cdot \mathbf{n}] = 0 \quad \text{on } \Gamma_t,
\end{align}

with a given set of initial and boundary conditions. Here \(\Omega_T = \Omega \times (0, T)\), where \(\Omega \subset \mathbb{R}^2\) is a bounded domain. \(\Gamma_t\) denotes the interface between the fluids at the time \(t\) with the normal \(\mathbf{n}\). \(\mathbf{u}\) is the fluid velocity and \(p\) stands for the pressure of the fluids. The symbol \([p]\) stands for the jump of \(p\) across the interface \(\Gamma_t\). \(\eta\) is the viscosity, which may have different (positive constant) values on both sides of \(\Gamma_t\). \(g\) is the gravitational force per unit mass; and \(\rho\) is the mass density of the fluid, which again can take different (positive constant) values on both sides of the interface. Equation (1.1)a is Darcy’s law [4], and (1.1)b implies that the fluids are incompressible. Equations (1.2)a and (1.2)b are the boundary conditions at the fluid-fluid interface, which represent the mathematical descriptions of the balance of the surface tension forces and the balance of mass, respectively. Equation (1.2)a is called the Laplace-Young condition, where \(\gamma\) is the dimensionless surface tension coefficient and \(\kappa\) is the (mean) curvature of the interface \(\Gamma_t\). Notice that the tangential component of the velocity \(\mathbf{u}\) may experience a jump across the interface [20].

Computationally, the above moving interface problem is difficult to solve directly due to the existence of the surface tension on the interface. In addition, during the evolution the fluid interface may experience topological changes such as self-intersection, pinch-off, splitting, and fattening. When that happens, the classical solution of the moving interface problem ceases to exist. In such cases it is very delicate matter to develop a proper notion of generalized solutions, and it becomes even more challenging to compute the generalized solutions when they can be defined.

To overcome the difficulties, an alternative approach for solving moving interface problems is the diffuse interface theory, which was originally developed as methodology for modeling and approximating solid-liquid phase transitions in which the effects of surface tension and non-equilibrium thermodynamic behavior may be important at the interface. In the theory, the interface is represented as a thin layer of finite thickness, as opposed to a sharp interface. Such an idea dates to Poisson, Gibbs, Rayleigh, van der Waals, and Korteweg (see [23, 20] and the references therein). The approach then uses an auxiliary function (called the phase field function) to indicate the “phase”. The phase field function, denoted by \(\varphi\) below, assumes distinct values in the bulk phases away from the interfacial region, through which the phase function varies smoothly. The interface itself can be associated with an intermediate contour/level set of the phase function (cf. [2, 3, 8, 22, 23] and the references therein). Generally speaking, the diffuse interface models are expected to converge to some corresponding sharp interface models as the width of the interfacial layer tends to zero.

The diffuse interface model for Hele-Shaw flows to be studied in this paper is given as follows:

\begin{align}
\mathbf{u} &= -\nabla p - \gamma \varphi \nabla \mu \quad \text{in } \Omega_T, \\
\text{div } \mathbf{u} &= 0 \quad \text{in } \Omega_T, \\
\varphi_t + \mathbf{u} \cdot \nabla \varphi - \varepsilon \Delta \mu &= 0 \quad \text{in } \Omega_T, \\
\mu &= -\varepsilon \Delta \varphi + \frac{1}{\varepsilon} f(\varphi) \quad \text{in } \Omega_T,
\end{align}
where \( f(\phi) = F'(\phi) \) and \( F(\phi) = \frac{1}{4}(\phi^2 - 1)^2 \) is the so-called double-well (potential) energy density, and \( 0 < \varepsilon < 1 \) is a fixed constant. To close the system, we impose the following initial and boundary conditions

\[
\begin{align*}
\frac{\partial p}{\partial n} &= \frac{\partial \mu}{\partial n} = \frac{\partial \phi}{\partial n} = 0 & \text{on } \partial \Omega_T := \partial \Omega \times (0, T], \\
\phi(\cdot, 0) &= \phi_0(\cdot) & \text{in } \Omega.
\end{align*}
\]

Note that we have suppressed the superscript \( \varepsilon \) in \((u^\varepsilon, p^\varepsilon, \phi^\varepsilon)\) for the sake of notational simplicity. Although \( \Omega \) is a two-dimensional domain in the original Hele-Shaw problem, in this paper we consider \( \Omega \subset \mathbb{R}^d \) (\( d = 2, 3 \)) because the three-dimensional problem also has a mathematical interest and arises from biological applications [28]. Here the vector \( u(x, t) \in \mathbb{R}^d \) and the scalar \( p(x, t) \in \mathbb{R} \) denote the velocity and the pressure of the fluid mixture at the space-time point \((x, t)\), respectively. The variables \( \phi(x, t), \mu(x, t) \in \mathbb{R} \) are the phase field function and the chemical potential, respectively. \( \phi \) assumes distinct values — namely, \( \pm 1 \) based on our choice of \( F(\phi) \) — in the bulk phases away from a thin layer of width \( O(\varepsilon) \). This thin layer is called the diffuse interface region. It is natural to define the zero level curve of \( \phi \), \( \Gamma_\varepsilon(t) = \{ x \in \mathbb{R}^d : \phi(x, t) = 0 \} \), as the \( d - 1 \) dimensional interface. Eq. (1.3) with \( \gamma = 0 \) is the Darcy equation [4]. (1.5) and (1.6) without the convection term \( u \cdot \nabla \phi \) is the Cahn-Hilliard equation [8, 11, 23]. Note that if \( \gamma = 0 \), the velocity vanishes, and the Cahn-Hilliard equation results.

The system (1.3)–(1.8) is a special case of the BHSCH (Boussinesq-Hele-Shaw-Cahn-Hilliard) model proposed by Lee, Lowengrub, and Goodman in [20]. They showed, using formal asymptotics, that solutions of the BHSCH system converge to those of the Hele-Shaw model (1.1)–(1.2) as their interfacial parameter converges to zero. We note that the pressure \( p \) in (1.3) has a different scaling from that in the BHSCH model in [20]. To obtain a similarly scaled pressure, one can simply introduce a redefined pressure in our model via \( \tilde{p} = p + \gamma \phi \mu \). We shall refer (1.3)–(1.8) as the DCH (Darcy-Cahn-Hilliard) model/system herein.

Define the Cahn-Hilliard energy functional

\[
\mathcal{J}_\varepsilon(\phi) := \int_\Omega \left[ \frac{\varepsilon}{2} |\nabla \phi|^2 + \frac{1}{\varepsilon} F(\phi) \right] dx.
\]

Like many diffuse interface models (cf. [2] [12] [22] [20] [23]), the DCH system is also a dissipative system as it satisfies the following energy dissipation law (see Sec. 2 for the details):

\[
\frac{d\mathcal{J}_\varepsilon(\phi)}{dt} + \varepsilon \| \nabla \mu \|^2_{L^2} + \frac{1}{\gamma} \| u \|^2_{L^2} = 0.
\]

As expected, the above energy law plays a vital role in the analysis of the DCH system and in the design and analysis of numerical methods for the system (see Secs. 2, 4 for the details).

This paper consists of four additional sections. Section 2 is devoted to the PDE analysis of the initial-boundary value problem (1.3)–(1.8). Weak solutions are defined and the uniqueness and regularities of weak solutions are established. Section 3 contains the formulation of our fully discrete implicit finite element method for problem (1.3)–(1.8). It is shown that the proposed numerical method satisfies a mass conservation law in addition to a discrete energy law that mimics the basic energy...
law for the Darcy-Cahn-Hilliard phase field model and holds uniformly in the phase field parameter $\varepsilon$. With help of the discrete energy law, it also proved that the fully discrete finite method is unconditionally energy stable and uniquely solvable at each time step. Section 4 presents a convergence analysis for the proposed fully discrete finite element method. Using the compactness method it is shown that the finite element solution has an accumulation point that is a weak solution of problem (1.3)–(1.8). As a byproduct, this convergence result also provides a constructive proof of the existence global-in-time weak solutions to the PDE system (1.3)–(1.8) in both two and three dimensions. Finally, in Sec. 5 we provide some results of numerical experiments validating our theoretical results and showing the effectiveness of the proposed fully discrete finite element method. To solve the nonlinear finite element equations at each time step, we propose a nonlinear multigrid iterative method to do the job. The details of the nonlinear multigrid solver and some other algorithmic and implementation issues are described in Appendix A.

2. PDE analysis. The standard space notations are used in this paper, we refer to [1, 9] for their exact definitions. In particular, $B^*$ denotes the dual space of a Banach space $B$, and $B^d$ denotes the vector Banach space $B^d$, where $d$ is the dimension space. Here we shall assume $d = 2$ or 3. The symbol $(\cdot, \cdot)$ is used to denote the standard $L^2(\Omega)$ inner product, $(\cdot, \cdot)$ stands for the dual product between $H^1(\Omega)$ and $(H^1(\Omega))^*$. $L_0^2(\Omega)$ denotes the subspace of $L^2(\Omega)$ whose functions have zero mean. Throughout the paper, unless stated otherwise, $c$ and $C$ will be used to denote generic positive constants which are independent of $\mu$, $\phi$, $u$, and $\epsilon$. If, for example, there is a dependence on $\epsilon$, we shall explicitly write $C = C(\epsilon)$. As indicated earlier, we shall assume that $0 < \epsilon < 1$.

In the next section we shall construct a finite element method which directly approximates variables $p$, $\mu$, and $\phi$, but not $u$, which will be computed as an auxiliary variable as needed. Specifically, we shall approximate the pressure equation by a standard finite element method and the phase equation by a mixed finite element method. We remark that it is also a viable strategy that approximates both the pressure equation and the phase equation by mixed finite element methods, which we shall study in a separate work.

To get the governing equations without using $u$, substituting the expression of $u$ in (1.3) into (1.4) and (1.5) we get

\begin{align}
\text{(2.1)} & \quad \text{div} \ (\nabla p + \gamma \varphi \nabla \mu) = 0 \quad \text{in } \Omega_T, \\
\text{(2.2)} & \quad \varphi_t - \varepsilon \Delta \mu - \text{div} \ (\varphi [\nabla p + \gamma \varphi \nabla \mu]) = 0 \quad \text{in } \Omega_T.
\end{align}

Then the PDE system to be studied and approximated in this paper consists of equations (2.1), (2.2), and (1.6), which are complimented with the boundary and initial conditions (1.7)–(1.8).

Motivated by the energy law (1.10), we define the following weak formulation and solutions to the initial-boundary value problem.

**Definition 2.1.** Let $\varphi_0^\epsilon \in H^1(\Omega)$. A triple $(p, \mu, \varphi)$ is called a weak solution of
Problem (2.1), (2.2), and (1.6)–(1.8) if it satisfies

\begin{align}
(2.3) & 
   p \in L^\frac{d}{d-2}((0,T); H^1(\Omega) \cap L^2(\Omega)), \\
(2.4) & 
   \mu \in L^2((0,T); H^1(\Omega)), \\
(2.5) & 
   \nabla p + \gamma \varphi \nabla \mu \in L^2((0,T); L^2(\Omega)), \\
(2.6) & 
   \varphi \in L^\infty((0,T); H^1(\Omega) \cap L^4((0,T); L^\infty(\Omega))), \\
(2.7) & 
   \varphi_t \in L^\frac{d}{d-2}((0,T); (H^1(\Omega))^*),
\end{align}

and there hold for almost all \( t \in (0,T) \)

\begin{align}
(2.8) & 
   (\nabla p + \gamma \varphi \nabla \mu, \nabla q) = 0 \quad \forall q \in H^1(\Omega), \\
(2.9) & 
   \langle \varphi_t, \nu \rangle + \varepsilon (\nabla \mu, \nabla \nu) + (\varphi [\nabla p + \gamma \varphi \nabla \mu], \nabla \nu) = 0 \quad \forall \nu \in H^1(\Omega), \\
(2.10) & 
   (\mu, \psi) - \varepsilon (\nabla \varphi, \nabla \psi) - \frac{1}{\varepsilon} (f(\varphi), \psi) = 0 \quad \forall \psi \in H^1(\Omega),
\end{align}

with the initial condition \( \varphi(0) = \varphi_0^\varepsilon \).

**Remark 2.1.** The reason for not breaking the sum \( \nabla p + \gamma \varphi \nabla \mu \) is that it has better regularity/integrability than each of its two terms. By the Aubin-Lions lemma (cf. [25]), the regularity on \( \varphi \) ensures that \( \varphi \in C^0([0,T]; L^2(\Omega)) \). Hence, the initial condition \( \varphi(0) = \varphi_0^\varepsilon \) makes sense.

**Remark 2.2.** The regularities imposed on the solution \( (p, \mu, \varphi) \) in the definition are not the “minimum” required to make all terms in (2.8)–(2.10) be well defined. These regularities are imposed because they are suggested by the energy law (1.10). Moreover, the product space of the spaces used in the definition for \( (p, \mu, \varphi) \) is indeed the energy space associated with the DCH system.

As we mentioned in Sec. 1, a key feature of the Darcy-Cahn-Hilliard (DCH) system is that it is a dissipative system in the sense that it satisfies an energy law, namely, (1.10). Below we demonstrate that this is indeed the case for weak solutions of the DCH system.

**Lemma 2.2.** Suppose that \( \varphi_0^\varepsilon \in H^1(\Omega) \) and \( \Omega \subset \mathbb{R}^d \) (\( d = 2,3 \)) is a Lipschitz domain, let \( (p, \mu, \varphi) \) be a weak solution defined by (2.8)–(2.10). In addition, suppose that the initial value \( \varphi_0^\varepsilon \) satisfies \( J_\varepsilon(\varphi_0^\varepsilon) \leq C_0 \) for some \( \varepsilon \)-independent constant \( C_0 \), i.e., the initial energy is uniformly bounded in \( \varepsilon \). Set \( u := - (\nabla p + \gamma \varphi \nabla \mu) \). Then, for almost all \( t \in (0,T) \),

\begin{align}
(2.11) & 
   \int_\Omega \varphi(x,t) \, dx = \int_\Omega \varphi_0^\varepsilon(x) \, dx,
\end{align}
and for all \( t \in (0, T) \) and some \( \varepsilon \)-independent constant \( C = C(E(0)) > 0 \)

\[
E(t) + \int_0^t \left[ \varepsilon \| \nabla \mu(s) \|_{L^2}^2 + \frac{1}{\gamma} \| u(s) \|_{L^2}^2 \right] ds = E(0) < \infty, \tag{2.12}
\]

\[
\max_{0 \leq s \leq t} \| \varphi(s) \|_{H^1}^2 \leq \frac{C}{\varepsilon}, \tag{2.13}
\]

\[
\int_0^t \| \mu(s) \|_{H^1}^2 \ ds \leq \frac{(T + 1)C}{\varepsilon^5}, \tag{2.14}
\]

\[
\int_0^t \| \mu(s) - \varepsilon^{-1} f(\varphi(s)) \|_{L^2}^2 \ ds \leq \frac{(T + 1)C}{\varepsilon^5}, \tag{2.15}
\]

\[
\int_0^t \| \varphi_t(s) + u(s) \cdot \nabla \varphi(s) \|_{(H^1)^*}^2 \ ds \leq C\varepsilon, \tag{2.16}
\]

\[
\int_0^t \| \varphi_t(s) \|_{L^{2+3}}^2 \ ds \leq \frac{C}{\varepsilon}, \tag{2.17}
\]

\[
\max_{0 \leq s \leq t} \| \varphi(s) \| - 1 \|_{L^2}^2 \leq C\varepsilon, \tag{2.18}
\]

where \( E(t) := \mathcal{J}_\varepsilon(\varphi(t)) \) and \( \mathcal{J}_\varepsilon(\cdot) \) is defined in \( (1.9) \).

Proof. (2.11) follows trivially from setting \( \nu = 1 \) in \( (2.9) \). To prove (2.12), we consider two cases separately. First, suppose that \( \varphi_t \in L^2 ((0, T); H^1(\Omega)) \). Then setting \( q = \frac{p}{\gamma} \) in \( (2.8) \), \( \nu = \mu \) in \( (2.9) \) and \( \psi = -\varphi_t \) in \( (2.10) \) (note that by the assumption, \( -\varphi_t \) is a valid test function), and adding the resulting equations we get

\[
\frac{d}{dt} \left[ \varepsilon \| \nabla \varphi \|_{L^2}^2 + \frac{1}{\varepsilon} \| F(\varphi, 1) \|_{L^2}^2 + \frac{1}{\gamma} \| \nabla p + \gamma \varphi \nabla \varphi \|_{L^2}^2 \right] = 0.
\]

Integrating over the interval \( (0, t) \) yields (2.12).

For the general case \( \varphi_t \in L^2 ((0, T); (H^1(\Omega))^*) \), and we note that \( \psi = -\varphi_t \) is not a valid test function in (2.10). However, this technical difficulty can be overcome by using a Steklov average technique. For \( t \in (0, T) \), let \( \delta > 0 \) be a small number. Define the Steklov average \( \varphi^\delta \) of \( \varphi \) by (cf. \[19, Ch. 2\])

\[
\varphi_t^\delta(\cdot, t) := S^\delta_+(\varphi)(\cdot, t) := \frac{1}{\delta} \int_{t}^{t+\delta} \varphi(\cdot, s) \ ds \quad \forall t \in (0, T),
\]

Trivially, for small enough \( \delta \),

\[
\varphi_t^\delta(\cdot, t) := \left( \varphi_t^\delta(\cdot, t) \right)_t = \frac{\varphi(\cdot, t + \delta) - \varphi(\cdot, t)}{\delta}.
\]

Hence, \( \varphi_t^\delta(\cdot, t) \in H^1(\Omega) \) for almost every \( t \in (0, T - \delta) \). It is well known that (cf. \[19, Ch. 2\])

\[
S^\delta_+(\varphi) = (S^\delta_+(\varphi))_t = \varphi_t^\delta.
\]

Note that the derivative on the left-hand side of the above identity is understood as a distributional derivative, while the derivative on the right-hand side is understood in the classical sense.
Now applying $S^\delta_+$ to both sides of (2.8)–(2.10) after replacing $t$ by $s$ yields

\begin{equation}
(\nabla p^\delta + \gamma (\varphi \nabla \mu)^\delta, \nabla q) = 0 \quad \forall q \in H^1(\Omega),
\end{equation}

(2.20)

\begin{equation}
(\varphi^\delta, \nu) + \varepsilon (\nabla \mu^\delta, \nabla \nu) + ((\varphi |\nabla p + \gamma \varphi \nabla \mu)^\delta, \nabla \nu) = 0 \quad \forall \nu \in H^1(\Omega),
\end{equation}

(2.21)

\begin{equation}
(\mu^\delta, \psi) - \varepsilon (\nabla \varphi^\delta, \nabla \psi) - \frac{1}{\varepsilon} ((f(\varphi))^\delta, \psi) = 0 \quad \forall \psi \in H^1(\Omega).
\end{equation}

(2.22)

Setting $\psi = -\varphi^\delta$ in (2.22), $\nu = \mu^\delta$ in (2.21) and $q = \frac{\varepsilon}{\gamma}$ in (2.20) and adding the resulting equations we get

\begin{equation}
\frac{d}{dt} \mathcal{J}_\varepsilon(\varphi^\delta) + \varepsilon J(\nabla \mu^\delta)^2 + \frac{1}{\gamma} \frac{1}{\varepsilon} \left| \nabla p^\delta + \gamma (\varphi \nabla \mu)^\delta \right|_{L^2}^2 = R^\delta(t),
\end{equation}

(2.23)

where

\[ R^\delta(t) := \frac{1}{\varepsilon} (f(\varphi^\delta) - (f(\varphi))^\delta, \varphi^\delta) + (\nabla p^\delta + \gamma (\varphi \nabla \mu)^\delta, \nabla \mu^\delta - (\varphi \nabla \mu)^\delta) + (\varphi |\nabla p^\delta + \gamma (\varphi \nabla \mu)^\delta| - (\varphi |\nabla p + \gamma \varphi \nabla \mu|^\delta, \nabla \mu^\delta). \]

Integrating (2.23) in $t$ gives

\[
\mathcal{J}_\varepsilon(\varphi^\delta(s)) + \int_0^s \left( \varepsilon \frac{1}{\gamma} \left| \nabla p^\delta(t) + \gamma (\varphi \nabla \mu)^\delta(t) \right|_{L^2}^2 \right) dt = \mathcal{J}_\varepsilon(\varphi^\delta(0)) + \int_0^s R^\delta(t) dt \quad \forall s \in (0, T).
\]

Note that, for each fixed $\varepsilon > 0$, $f(\varphi) \in L^2((0, T); H^1(\Omega))$, since $\varphi \in L^4((0, T); L^\infty(\Omega))$, and $f(\cdot)$ is continuous. Sending $\delta \to 0^+$ and using properties of the Steklov average $S^\delta_+$ (cf. [19, Ch. 2]) we get

\[
\lim_{\delta \to 0^+} \int_0^s R^\delta(t) dt = 0,
\]

\[
\mathcal{J}_\varepsilon(\varphi^\delta(s)) = \int_0^s \left( \varepsilon \frac{1}{\gamma} \left| \nabla p(t) + \gamma (\varphi(t) \nabla \mu(t)) \right|_{L^2}^2 \right) dt = \mathcal{J}_\varepsilon(\varphi(0)).
\]

Hence, we recover (2.12). Observe that from (2.12) we can conclude that $E(t)$ is an absolutely continuous function of time.

Using (2.12) and the estimate

\begin{equation}
(F(\varphi), 1) \geq \frac{1}{2} \| \varphi \|_{L^2}^2 - \frac{3}{4} |\Omega|
\end{equation}

(2.24)

we discover that $\| \varphi \|_{L^2}^2 \leq C$, for all time and independent of $\varepsilon$, and inequality (2.13) follows. Inequalities (2.14) and (2.15) follow straightforwardly from (2.10), (2.13), and the Sobolev embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$ for $d = 2, 3$. Inequality (2.16) is an immediate consequence of (2.9) and the fact that $\varepsilon \| \nabla \mu \|_{L^2(L^2)} < C < \infty$ from (2.12). To show (2.17), by (2.9), the Sobolev embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$ for $d = 2, 3$, and (2.13), we get for any $\nu \in W^{1,3}(\Omega)$

\[
\langle \varphi \mu, \nu \rangle = -\varepsilon (\nabla \mu, \nabla \nu) + (\varphi \mu, \nabla \nu) \leq \varepsilon \| \nabla \mu \|_{L^2} \| \nabla \nu \|_{L^2} + \| \varphi \|_{L^5} \| \mu \|_{L^2} \| \nabla \nu \|_{L^3}
\]

\[
\leq C \left( \varepsilon \| \nabla \mu \|_{L^2} + \| \varphi \|_{H^1} \| \mu \|_{L^2} \| \nabla \nu \|_{L^3} \right)
\]

\[
\leq C \left( \varepsilon \| \nabla \mu \|_{L^2} + \frac{1}{\sqrt{\varepsilon}} \| \mu \|_{L^2} \right) \| \nabla \nu \|_{L^3}.
\]
The above inequality and (2.12) infer (2.17).

Finally, (2.18) is an immediate consequence of (2.12) and the inequality \((\varphi^2 - 1)^2 \geq (|\varphi| - 1)^2\). The proof is complete. \(\square\)

The next lemma shows that weak solutions have some additional regularities.

**Lemma 2.3.** Suppose that \(\varphi_0^2 \in H^1(\Omega)\) and \(\Omega \subset \mathbb{R}^d\) \((d = 2, 3)\) is a Lipschitz domain. Let \((p, \mu, \varphi)\) be a weak solution defined by (2.8)–(2.10). Then \(\varphi \in L^2((0, T); H^2(\Omega))\). Moreover, if \(\Omega\) is a convex polygonal or polyhedral domain, then \(\varphi \in L^2((0, T); H^3(\Omega))\).

**Proof.** We begin by rewriting (2.10) as

\[
(2.25) \\
\varepsilon(\nabla \varphi, \nabla \psi) = (\mu - \varepsilon^{-1} f(\varphi), \psi) \quad \forall \psi \in H^1(\Omega).
\]

Hence, \(\varphi\) is a weak solution to a Poisson equation with homogeneous Neumann boundary conditions and the right-hand side “source” function \(g := \mu - \varepsilon^{-1} f(\varphi)\). Since \(\varphi \in L^1((0, T); L^\infty(\Omega))\), then \(g \in L^2((0, T); H^{-1}(\Omega))\). By elliptic regularity theory (cf. [10] Ch. 7) we conclude that \(\varphi \in L^2((0, T); H^2(\Omega))\) and \(\frac{\partial \varphi}{\partial n} = 0\) on \(\partial \Omega\) in the distributional sense.

Introduce the function space

\[
U := \left\{ \theta \in H^2(\Omega); \frac{\partial \theta}{\partial n} = 0 \text{ on } \partial \Omega \right\}.
\]

For any \(\theta \in U \cap C^\infty(\Omega)\), setting \(\psi = \Delta \theta\) in (2.25), integrating by parts and using a density argument we get

\[
\varepsilon(\Delta \varphi, \Delta \theta) = (\nabla [\mu - \varepsilon^{-1} f(\varphi)], \nabla \theta) \quad \forall \theta \in U,
\]

which implies that \(\varphi \in H^2(\Omega)\) with \((\varphi, 1) = \text{const}\) is the unique weak solution to the following biharmonic problem:

\[
\Delta^2 \varphi = -\Delta g \quad \text{in } \Omega, \\
\frac{\partial \varphi}{\partial n} = \frac{\partial \Delta \varphi}{\partial n} = 0 \quad \text{on } \partial \Omega.
\]

Since \(-\Delta g \in L^2((0, T); H^{-1}(\Omega))\), it follows from a well-known regularity result of [5] that \(\varphi \in L^2((0, T); H^3(\Omega))\) when \(\Omega\) is a convex polygonal domain. The proof is complete. \(\square\)

We conclude this section by establishing the following uniqueness theorem for weak solutions of problem (2.1), (2.2), and (1.6)–(1.8) defined in Definition 2.1.

**Theorem 2.4.** Suppose that \(\varphi_0^2 \in H^1(\Omega)\) with \(J_\varepsilon(\varphi_0^2) \leq C_0\) for some \(\varepsilon\)-independent constant \(C_0\) and \(\Omega \subset \mathbb{R}^d\) \((d = 2, 3)\) is a Lipschitz domain. We say that a weak solution \((p, \mu, \varphi)\) belongs to the function space \(\mathcal{F}\) if it satisfies the additional regularity conditions \(\nabla p + \gamma \varphi \mu \nabla \mu \in L^{\frac{12}{13}}((0, T); L^2(\Omega)), \mu \in L^{\frac{12}{11}}((0, T); H^1(\Omega)), \) and \(\varphi \in L^2((0, T); (H^3(\Omega))^*)\). Then weak solutions of (2.8)–(2.10) in the function class \(\mathcal{F}\) are unique.

**Proof.** Since the proof is long, we divide it into six steps.

**Step 1:** Suppose \((p_1, \mu_1, \varphi_1), i = 1, 2, \) are two weak solutions, and define \(u_i := -\nabla p_i - \gamma \varphi_i \nabla \mu_i, i = 1, 2\). Let \(p = p_1 + p_2, u = u_1 - u_2, \mu = \mu_1 - \mu_2, \) and \(\varphi = \varphi_1 - \varphi_2\). Subtracting the corresponding equations of (2.8)–(2.10) satisfied by \((p_1, \mu_1, \varphi_1)\) and
(p_2, \mu_2, \varphi_2) 
we get the following “error” equations:
\begin{align}
(2.26) \quad & (\mathbf{u}, \nabla q) = 0 \quad \forall q \in H^1(\Omega), \\
(2.27) \quad & \langle \varphi_t, \nu \rangle + \varepsilon (\nabla \mu, \nabla \nu) - (\varphi_1 \mathbf{u} + \varphi_2 \nabla \nu) = 0 \quad \forall \nu \in H^1(\Omega), \\
(2.28) \quad & (\mu, \psi) - \varepsilon (\nabla \varphi, \nabla \psi) - \frac{1}{\varepsilon} (f(\varphi_1) - f(\varphi_2), \psi) = 0 \quad \forall \psi \in H^1(\Omega).
\end{align}

We will frequently use the fact that \ref{2.28},\ref{2.27},\ref{2.26}.

Using Schwarz inequality and the following Gagliardo-Nirenberg inequality (cf. [1][II])
\[ ||\varphi||_{L^\infty} \leq C ||\Delta \varphi||_{L^2}^{\frac{d}{4}} + C ||\varphi||_{L^2} \quad (d = 2, 3) \]
in \ref{2.29} we get
\[ \frac{d}{dt} ||\varphi||_{L^2}^2 + 2 ||\mu||_{L^2}^2 \leq 2 ||\mathbf{u}||_{L^2} ||\nabla \varphi||_{L^2} ||\varphi||_{L^\infty} + \frac{2}{\varepsilon} ||g(\varphi_1, \varphi_2)||_{L^\infty} ||\varphi||_{L^2} ||\mu||_{L^2} \\
\leq \frac{\varepsilon}{4\gamma} ||\mathbf{u}||_{L^2}^2 + \frac{C}{\varepsilon} ||\nabla \varphi||_{L^2} ||\Delta \varphi||_{L^2} ||\varphi||_{L^\infty} + \frac{2}{\varepsilon} ||g(\varphi_1, \varphi_2)||_{L^\infty} ||\varphi||_{L^2} ||\mu||_{L^2} \\
\leq \frac{\varepsilon}{4\gamma} ||\mathbf{u}||_{L^2}^2 + \frac{\varepsilon^2}{16} ||\Delta \varphi||_{L^2}^2 + C(\varepsilon) ||\nabla \varphi||_{L^2}^2 ||\varphi||_{L^2}^2 + ||\mu||_{L^2}^2 \\
+ \frac{1}{\varepsilon^2} ||g(\varphi_1, \varphi_2)||_{L^\infty}^2 ||\varphi||_{L^2}^2.
\]

Hence, it follows from \ref{2.12} that
\[ \frac{d}{dt} ||\varphi||_{L^2}^2 + ||\mu||_{L^2}^2 \leq \frac{\varepsilon}{4\gamma} ||\mathbf{u}||_{L^2}^2 + \frac{\varepsilon^2}{16} ||\Delta \varphi||_{L^2}^2 \\
+ C(\varepsilon) \left( 1 + ||g(\varphi_1, \varphi_2)||_{L^\infty}^2 \right) ||\varphi||_{L^2}^2.
\]

**Step 2:** Setting \( \psi = \Delta \varphi \) in \ref{2.28} gives
\[ \varepsilon ||\Delta \varphi||_{L^2}^2 = - (\mu, \Delta \varphi) + \frac{1}{\varepsilon} (g(\varphi_1, \varphi_2) \varphi, \Delta \varphi) \\
\leq \frac{\varepsilon}{2} ||\Delta \varphi||_{L^2}^2 + \frac{1}{\varepsilon} ||\mu||_{L^2}^2 + \frac{1}{\varepsilon} ||g(\varphi_1, \varphi_2)||_{L^\infty}^2 ||\varphi||_{L^2}^2.
\]

Hence
\[ \frac{\varepsilon^2}{4} ||\Delta \varphi||_{L^2}^2 \leq \frac{1}{2} ||\mu||_{L^2}^2 + C(\varepsilon) ||g(\varphi_1, \varphi_2)||_{L^\infty}^2 ||\varphi||_{L^2}^2.
\]

**Step 3:** First, on noting that
\[ \mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2 = - \nabla p - \gamma (\varphi_1 \nabla \mu_1 - \varphi_2 \nabla \mu_2) = - \nabla p - \gamma \varphi_1 \nabla \mu - \gamma \varphi \nabla \mu_2,
\]
and using (2.26) with \( q = p \) we obtain

\[
(2.32) \quad \frac{1}{\gamma} \| u \|^2_{L^2} = -\frac{1}{\gamma} (u, \nabla p) - (u, \varphi_1 \nabla \mu) - (u, \varphi \nabla \mu_2) + \langle \varphi_1 u, \nabla \mu \rangle - (\varphi u, \nabla \mu_2).
\]

Second, setting \( \nu = \mu \) in (2.27) yields

\[
(2.33) \quad \langle \varphi_t, \mu \rangle + \varepsilon \| \nabla \mu \|^2_{L^2} = (\varphi_1 u + \varphi u_2, \nabla \mu).
\]

Third, applying the Steklov average operator \( S^\delta_+ \) to (2.28) (we use the same notion as in the proof of Lemma 2.2) yields

\[
(\mu^\delta, \psi) - \varepsilon (\nabla \varphi^\delta, \nabla \psi) - \frac{1}{\varepsilon} ((g(\varphi_1, \varphi_2)\varphi)^\delta, \psi) = 0 \quad \forall \psi \in H^1(\Omega).
\]

Where \( g(\varphi_1, \varphi_2) = \varphi_1^2 + \varphi_1 \varphi_2 + \varphi_2^2 - 1 \). Setting \( \psi = -\varphi^\delta \) in the above equation gives

\[
-(\mu^\delta, \varphi_1^\delta) + \varepsilon \frac{d}{dt} \| \nabla \varphi^\delta \|^2_{L^2} + \frac{1}{\varepsilon} ((g(\varphi_1, \varphi_2)\varphi)^\delta, \varphi^\delta) = 0.
\]

Taking the limit \( \delta \to 0^+ \) and using the properties of Steklov average operator (cf. [19]) we get

\[
(2.34) \quad -\langle \varphi_1, \mu \rangle + \frac{\varepsilon}{2} \frac{d}{dt} \| \nabla \varphi \|^2_{L^2} + \frac{1}{\varepsilon} \langle \varphi_1, g(\varphi_1, \varphi_2) \varphi \rangle = 0.
\]

Finally, adding (2.32), (2.33), and (2.34), using the fact that \( (u_2, \nabla (\varphi \mu)) = (u, \nabla (\varphi \mu_2)) = 0 \) and Young’s and Gagliardo-Nirenberg inequalities (cf. [11]) we get

\[
\frac{1}{\gamma} \| u \|^2_{L^2} + \varepsilon \| \nabla \mu \|^2_{L^2} + \frac{\varepsilon}{2} \frac{d}{dt} \| \nabla \varphi \|^2_{L^2} + \frac{1}{\varepsilon} \langle \varphi_t, g(\varphi_1, \varphi_2) \varphi \rangle
\]

\[
= (\varphi u_2, \nabla \mu) - (\varphi u, \nabla \mu_2) = -(u_2 \cdot \nabla \varphi, \mu) + (u \cdot \nabla \varphi, \mu_2)
\]

\[
\leq \| u_2 \|_{L^2} \| \nabla \varphi \|_{L^3} \| \mu \|_{L^6} + \| u \|_{L^2} \| \nabla \varphi \|_{L^3} \| \mu_2 \|_{L^6}
\]

\[
\leq C \left( \| u_2 \|_{L^2} \| \mu \|_{H^1} + \| u \|_{H^1} \| \mu_2 \|_{L^6} \right) \left( \| \Delta \varphi \|_{L^2}^2 + \| \nabla \varphi \|_{L^2}^2 \right)
\]

\[
\leq \frac{1}{4\gamma} \| u \|^2_{L^2} + \frac{\varepsilon}{2} \| \mu \|^2_{H^1} + \frac{\varepsilon}{16} \| \Delta \varphi \|^2_{L^2}
\]

\[
+ C(\varepsilon) \left( \| u_2 \|_{L^2}^2 \| \mu \|^2_{H^1} + \| \mu_2 \|_{H^1}^2 \right) \| \nabla \varphi \|^2_{L^2}.
\]

Hence

\[
(2.35) \quad \frac{3}{4\gamma} \| u \|^2_{L^2} + \frac{\varepsilon}{2} \| \nabla \mu \|^2_{L^2} + \frac{\varepsilon}{2} \frac{d}{dt} \| \nabla \varphi \|^2_{L^2} + \frac{1}{\varepsilon} \langle \varphi_t, g(\varphi_1, \varphi_2) \varphi \rangle
\]

\[
\leq \frac{\varepsilon}{16} \| \Delta \varphi \|^2_{L^2} + \| \mu \|^2_{H^1} + C(\varepsilon) \left( \| u_2 \|_{L^2}^2 + \| \mu_2 \|_{H^1}^2 \right) \| \nabla \varphi \|^2_{L^2}.
\]

**Step 4:** To control the last term on the left-hand side of (2.35), we rewrite

\[
g(\varphi_1, \varphi_2) = \varphi_1^2 + \varphi_1 \varphi_2 + \varphi_2^2 - 1 = (\varphi_1 - \varphi_2)^2 + 3\varphi_1 \varphi_2 - 1 = (\varphi^2 - 1) + 3\varphi_1 \varphi_2.
\]
Thus

\[ \langle \varphi_t, g(\varphi_1, \varphi_2) \varphi \rangle = \frac{1}{2} \langle (\varphi^2)_t, \varphi^2 \rangle + \frac{3}{2} \langle \varphi^2, \varphi_1 \varphi_2 \rangle \]

\[ = \frac{d}{dt} \left[ \frac{1}{4} \| \varphi \|_{L^4}^4 - \frac{1}{2} \| \varphi \|_{L^2}^2 + \frac{3}{2} \| (\varphi^2, \varphi_1 \varphi_2) \| \right] - \frac{3}{2} \langle (\varphi_1 \varphi_2)_t, \varphi^2 \rangle \]

\[ = \frac{d}{dt} \left[ \frac{1}{4} \| \varphi \|_{L^4}^4 - \frac{1}{2} \| \varphi \|_{L^2}^2 + \frac{3}{2} \| (\varphi^2, \varphi_1 \varphi_2) \| \right] - \frac{3}{2} \left[ \langle \varphi_1 t, \varphi_2 \varphi^2 \rangle + \langle \varphi_2 t, \varphi_1 \varphi^2 \rangle \right]. \]

To bound the last term on the right-hand side of (2.36), we introduce the inverse Laplace operator $\Delta^{-1} : (H^1(\Omega))^* \to H^1(\Omega)$. For any $w \in (H^1(\Omega))^*$ with $\langle w, 1 \rangle = 0$, let $\Delta^{-1} w \in H^1(\Omega)$ be the unique solution of the following problem:

\[ \langle \nabla (\Delta^{-1} w), \nabla \eta \rangle = - \langle w, \eta \rangle \quad \forall \eta \in H^1(\Omega), \]

\[ (\Delta^{-1} w, 1) = 0. \]

It is straightforward to show that, for all $w \in (H^1(\Omega))^*$ with $\langle w, 1 \rangle = 0$,

\[ \| w \|_{(H^1)^*} = \| \nabla (\Delta^{-1} w) \|_{L^2}. \]

Then for $j, k = 1, 2$ and $j \neq k$ using Sobolev inequality (cf. [1]) we have

\[ \langle \varphi_{jt}, \varphi_k \varphi^2 \rangle = - \langle \nabla (\Delta^{-1} \varphi_{jt}), \varphi^2 \nabla \varphi_k + 2 \varphi_k \varphi \nabla \varphi \rangle \]

\[ \leq \| \nabla (\Delta^{-1} \varphi_{jt}) \|_{L^2} \| \varphi^2 \nabla \varphi_k + 2 \varphi_k \varphi \nabla \varphi \|_{L^2} \]

\[ \leq \| \varphi_{jt} \|_{(H^1)^*} \left( \| \nabla \varphi_k \|_{L^6} \| \varphi \|_{L^6}^2 + 2 \| \varphi_k \|_{L^6} \| \varphi \|_{L^6} \| \nabla \varphi \|_{L^6} \right) \]

\[ \leq C \| \varphi_{jt} \|_{(H^1)^*} \left( \| \Delta \varphi_k \|_{L^2} \| \nabla \varphi \|_{L^6}^2 + \| \varphi_k \|_{H^1} \| \nabla \varphi \|_{L^2} \| \Delta \varphi \|_{L^2} \right) \]

\[ \leq \frac{\varepsilon^2}{48} \| \Delta \varphi \|_{L^2}^2 + C(\varepsilon) \left( \| \varphi_{jt} \|_{(H^1)^*} \| \Delta \varphi_k \|_{L^2} + \| \varphi_{jt} \|_{(H^1)^*}^2, \| \varphi_k \|_{H^1}^2 \right) \| \nabla \varphi \|_{L^2}^2. \]

**Step 5:** Adding (2.30), (2.31) and $\varepsilon$ times of (2.33), and utilizing (2.36) and (2.40) we get

\[ \frac{d}{dt} \left[ \frac{1}{2} \| \varphi \|_{L^2}^2 + \frac{\varepsilon^2}{2} \| \nabla \varphi \|_{L^2}^2 + \frac{1}{4} \| \varphi \|_{L^4}^4 + \frac{3}{2} \langle \varphi^2, \varphi_1 \varphi_2 \rangle \right] \]

\[ + \frac{1}{2} \left( 1 - \varepsilon^2 \right) \| \mu \|_{L^2}^2 + \frac{\varepsilon^2}{2} \| \mu \|_{L^4}^4 + \frac{\varepsilon^2}{2} \| \nabla \mu \|_{L^2}^2 + \frac{\varepsilon^2}{16} \| \Delta \varphi \|_{L^2}^2 \]

\[ \leq C(\varepsilon) \left( 1 + \| g(\varphi_1, \varphi_2) \|_{L^\infty} \right) \| \varphi \|_{L^2}^2 \]

\[ + C(\varepsilon) \left[ \left( \| u_2 \|_{L^2}^2 + \| u_2 \|_{H^1}^2 \right) + \sum_{j,k=1}^2 \left( \| \varphi_{jt} \|_{(H^1)^*} \| \Delta \varphi_k \|_{L^2} \right) \| \nabla \varphi \|_{L^2}^2 \]

\[ + \| \varphi_{jt} \|_{(H^1)^*}^2, \| \varphi_k \|_{H^1}^2 \left) \right\} \| \nabla \varphi \|_{L^2}^2 \]

\[ \leq a(t) \left( \| \varphi \|_{L^2}^2 + \| \nabla \varphi \|_{L^2}^2 \right), \]
where

\[ a(t) := C(\epsilon) \left[ 1 + \|g(\varphi_1, \varphi_2)\|_{L^\infty}^2 + \|u_2\|_{L^2}^2 + \|\mu_2\|_{H^1}^2 + \sum_{j \neq k} \left\| \varphi_{jk}(H^{1/2}) \right\|_{L^2} + \|\Delta \varphi_k\|_{L^2} + \|\varphi_{jk}(H^{1/2}) \right\| \right]. \]

Integrating (2.41) in \( t \) over the interval \((0, t)\) we get

\[ \|\varphi(t)\|_{L^2}^2 + \epsilon^2 \|\nabla \varphi(t)\|_{L^2}^2 + \|\varphi(t)\|_{L^2}^4 + 6 (\varphi^2(t), \varphi_1(t)\varphi_2(t)) \leq \int_0^t a(s) \left( \|\varphi(s)\|_{L^2}^2 + \|\nabla \varphi(s)\|_{L^2}^2 \right) \, dt. \]

**Step 6:** Define

\[ \tau := \max \left\{ t \in [0, T] ; \ (\varphi^2(s), \varphi_1(s)\varphi_2(s)) \geq 0 \ \forall s \in [0, t] \right\}. \]

We now show that \( \tau > 0 \). Since \( \varphi_1(0) = \varphi_2(0) = \varphi_0^\epsilon \), by continuity there exists \( t_1 > 0 \) such that for \( j = 1, 2 \)

\[ \left( |\varphi_0^\epsilon|^2 - \frac{1}{12}, \eta \right) \leq (\varphi_1(t)\varphi_2(t), \eta) \leq \left( |\varphi_0^\epsilon|^2 + \frac{1}{12}, \eta \right) \ \forall t \in [0, t_1], \ \forall 0 \leq \eta \in L^\infty(\Omega). \]

Consequently,

\[ (\varphi^2(t), \varphi_1(t)\varphi_2(s)) \geq \|\varphi_0^\epsilon\|_{L^2}^2 - \frac{1}{12} \|\varphi\|_{L^2}^2 \ \forall t \in [0, t_1]. \]

Substituting (2.43) into (2.42) yields

\[ \frac{1}{2} \|\varphi(t)\|_{L^2}^2 + \epsilon^2 \|\nabla \varphi(t)\|_{L^2}^2 + \|\varphi(t)\|_{L^2}^4 + 6 \|\varphi_0^\epsilon\|_{L^2}^2 \leq \int_0^t a(s) \left( \|\varphi(s)\|_{L^2}^2 + \|\nabla \varphi(s)\|_{L^2}^2 \right) \, ds \ \forall t \in [0, t_1]. \]

By Gronwall’s inequality we get

\[ \frac{1}{2} \|\varphi(t)\|_{L^2}^2 + \epsilon^2 \|\nabla \varphi(t)\|_{L^2}^2 \leq \exp \left\{ \int_0^T a(s) \, ds \right\} \]

Here we have used the fact that \( \int_0^T a(s) \, ds < \infty \). Thus, \( \varphi(t) = 0 \) for \( t \in [0, t_1] \). Therefore, \( \tau \geq t_1 > 0 \). In fact, the above proof also shows that \( \varphi(t) = 0 \) for \( t \in [0, \tau] \).

Suppose that \( \tau < T \), by the definition of \( \tau \) we have \( \varphi(\tau) = 0 \), that is, \( \varphi_1(\tau) = \varphi_2(\tau) \). Repeating the above Gronwall’s inequality argument with \( \tau \) in place of \( t = 0 \), we conclude that there exists \( t_2 > \tau \) such that \( \varphi(t) = 0 \) for \( t \in [0, t_2] \). Hence, \((\varphi^2(s), \varphi_1(s)\varphi_2(s)) = 0 \ \forall s \in [0, t_2] \). By the definition of \( \tau \) we must have \( \tau \geq t_2 \). So we get a contradiction. Therefore, \( \tau = T \) and \( \varphi(t) = 0 \), i.e., \( \varphi_1(t) = \varphi_2(t) \), for \( t \in [0, T] \). The proof is complete. \( \square \)
3. Fully discrete finite element method.

3.1. Formulation of the finite element method. For simplicity we assume that $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) is a polygonal or polyhedral domain. Let $J = \{ t_m \}_{m=0}^M$ be a uniform partition of $[0, T]$ of mesh size $\tau := \frac{T}{M}$, and $d_t v^m := (v^m - v^{m-1})/\tau$. (This is for simplicity; we could also use a quasi-uniform partition of the time interface.) Let $T_h$ be a quasi-uniform “triangulation” of the domain $\Omega$ of mesh size $h \in (0, 1)$ and $\bar{\Omega} = \bigcup_{K \in T_h} K$ ($K \in T_h$ are tetrahedrons in the case $d = 3$). For a nonnegative integer $r$, let $P_r(K)$ denote the space of polynomials of degree less than or equal to $r$ on $K$, and define

$$S^r_h = \{ v_h \in C^0(\bar{\Omega}) \mid v_h|_K \in P_r(K) \forall K \in T_h \}.$$ 

For fixed positive integers $r$ and $\ell$, we introduce the finite element spaces $V_h = S^r_h$ and $\hat{W}_h = S^\ell_h$. Define $\hat{V}_h := V_h \cap L^2_0(\Omega)$ and similarly for $\hat{W}_h$.

We now are ready to introduce our fully discrete finite element method for problem \((2.1), (2.2), (1.6)-(1.8)\) based on the variational formulation \((2.8)-(2.10)\). Find \(\{ (p^m_h, \mu^m_h, \varphi^m_h) \}_{m=1}^M \subset \hat{W}_h \times \hat{V}_h \times \hat{V}_h\) such that

\[
\begin{align*}
(3.1) & \quad (\nabla p^m_h + \gamma \varphi^{m-1}_h \nabla \mu^m_h, \nabla q_h) = 0 \quad \forall q_h \in W_h, \\
(3.2) & \quad (d_t \varphi^m_h, v_h) + \varepsilon (\nabla \mu^m_h, \nabla v_h) \\
& \quad \quad + (\varphi^{m-1}_h [\nabla p^m_h + \gamma \varphi^{m-1}_h \nabla \mu^m_h], \nabla v_h) = 0 \quad \forall v_h \in V_h, \\
(3.3) & \quad (\mu^m_h, \psi_h) - \varepsilon (\nabla \varphi^m_h, \nabla \psi_h) - \frac{1}{\varepsilon} (f^m_h, \psi_h) = 0 \quad \forall \psi_h \in V_h, \\
(3.4) & \quad \varphi^0_h = \varphi_{0h},
\end{align*}
\]

where $\varphi_{0h} \in V_h$, to be specified in the next section, is an approximation of $\varphi^0_0$, and

\[
(3.5) \quad f^m_h := (\varphi^m_h)^3 - \varphi^{m-1}_h.
\]

We will prove that $(\hat{W}_h, \hat{V}_h, \hat{V}_h)$ is a stable triple for our mixed finite element approximation. The techniques we use are based on energy estimates and convexity analysis, rather than an inf-sup-type condition, which would be used for the analysis of linear biharmonic-type equations \([9, 10, 14, 26]\).

3.2. Well-posedness of the finite element method. The goal of this subsection is to show that the fully discrete finite element scheme \((3.1)-(3.5)\) is uniquely solvable and energy stable for all $h, \tau, \varepsilon > 0$. To prove unconditional unique solvability, we shall show that at each time step the scheme can be reformulated as a minimization problem for a strictly convex and coercive functional. We begin by defining an inner product on the subspace $\hat{V}_h$.

**Lemma 3.1.** Define the bilinear form $a : \hat{V}_h \times \hat{V}_h \to \mathbb{R}$ via

\[
(3.6) \quad a(\mu, \nu) := \tau \left( \mathcal{M}(\varepsilon, \varphi^{m-1}_h) \nabla \mu + \varphi^{m-1}_h \nabla p(\mu), \nabla \nu \right)_{L^2} ,
\]

where $\mathcal{M}(\varepsilon, \varphi) = \varepsilon + \gamma \varphi^2$ and $p(\mu) \in \hat{W}_h$ solves

\[
(3.7) \quad (\nabla p(\mu), \nabla q)_{L^2} = -\gamma \left( \varphi^{m-1}_h \nabla \mu, \nabla q \right)_{L^2} \quad \forall q \in \hat{W}_h .
\]

Then $a(\cdot, \cdot)$ is an inner product on $\hat{V}_h$.

The proof is omitted for brevity. Note that in the next few calculations, the pressure, $p$, will be regarded as an auxiliary variable that can be calculated when the
chemical potential, \( \mu \), is known. Owing to the last result, we can define an invertible linear operator \( \mathcal{L} : \hat{V}_h \to \hat{V}_h \) via the following problem: given \( \zeta \in \hat{V}_h \), find \( \mu \in \hat{V}_h \) such that

\[
(3.8) \quad a(\mu, \nu) = - (\zeta, \nu)_{L^2} \quad \forall \nu \in \hat{V}_h.
\]

This clearly has a unique solution because \( a(\cdot, \cdot) \) is an inner product on \( \hat{V}_h \). We write \( \mathcal{L}(\mu) = -\zeta \), or, equivalently, \( \mu = -\mathcal{L}^{-1}(\zeta) \).

We now wish to define a negative norm, i.e., a discrete analogue to the \( H^{-1} \) norm. Again we omit the details for brevity.

**Lemma 3.2.** Let \( \zeta, \xi \in \hat{V}_h \) and suppose \( \mu_\zeta, \mu_\xi \in \hat{V}_h \) are the unique weak solutions to \( \mathcal{L}(\mu_\zeta) = -\zeta \) and \( \mathcal{L}(\mu_\xi) = -\xi \). Define

\[
(3.9) \quad (\zeta, \xi)_{L^{-1}} := a(\mu_\zeta, \mu_\xi) = - (\zeta, \mu_\xi)_{L^2} = -(\mu_\zeta, \xi)_{L^2}.
\]

\( (\cdot, \cdot)_{L^{-1}} \) defines an inner product on \( \hat{V}_h \), and the induced norm is

\[
(3.10) \quad \|\zeta\|_{L^{-1}} = \sqrt{(\zeta, \zeta)_{L^{-1}}}.
\]

Using this last norm we can define a variational problem closely related to our fully discrete scheme.

**Lemma 3.3.** Set \( K_1 := (\varphi_h^{m-1}, 1)_{L^2} \), and define \( \varphi^*_m := \varphi_h^{m-1} - K_1 \in \hat{V}_h \). For all \( \varphi \in \hat{V}_h \), define the nonlinear functional

\[
(3.11) \quad G(\varphi) := \frac{1}{2} \|\varphi - \varphi^*_m\|^2_{L^{-1}} + \frac{1}{4\varepsilon} \|\varphi + K_1\|^4_{L^2} + \frac{\varepsilon}{2} \|\nabla \varphi\|_{L^2}^2 - \frac{1}{\varepsilon} (\varphi_h^{m-1}, \varphi)_{L^2}.
\]

\( G \) is strictly convex and coercive on the linear subspace \( \hat{V}_h \). Consequently, \( G \) has a unique minimizer, call it \( \varphi^*_m \in \hat{V}_h \). Moreover, \( \varphi^*_m \in \hat{V}_h \) is the unique minimizer of \( G \) if and only if it is the unique solution to

\[
(3.12) \quad \frac{1}{\varepsilon} \left( (\varphi^*_m + K_1)^3 - (\varphi^*_m - K_1)^3, \psi \right)_{L^2} + \varepsilon \left( \nabla \varphi^*_m, \nabla \psi \right)_{L^2} - (\mu^*_m, \psi)_{L^2} = \frac{1}{\varepsilon} (\varphi_h^{m-1}, \psi)_{L^2}
\]

for all \( \psi \in \hat{V}_h \), where \( \mu^*_m \in \hat{V}_h \) is the unique solution to

\[
(3.13) \quad a(\mu^*_m, \nu) = - (\varphi^*_m - \varphi^*_m - K_1, \nu)_{L^2} \quad \forall \nu \in \hat{V}_h.
\]

**Proof.** In detail, the first variation, i.e. the gradient, of the first term of \( G \) is

\[
\frac{d}{ds} \left[ \frac{1}{2} \|\varphi + s\psi - \varphi^*_m\|^2_{L^{-1}} \right]_{s=0} = \left[ (\varphi + s\psi - \varphi^*_m, \psi)_{L^{-1}} \right]_{s=0}
\]

\[
= (\varphi - \varphi^*_m, \psi)_{L^{-1}}
\]

\[
= - (\mu, \psi)_{L^2},
\]

where, owing to the definition of the inner product \( (\cdot, \cdot)_{L^{-1}} \), \( \mu \in \hat{V}_h \) is the unique solution to

\[
(3.14) \quad a(\mu, \nu) = - (\varphi - \varphi^*_m, \nu)_{L^2} \quad \forall \nu \in \hat{V}_h.
\]
The second variation is
\[
\frac{d^2}{ds^2} \left[ \frac{1}{2} \| \varphi + s \psi - \varphi^{m-1}_h \|_{L^2}^2 \right]_{s=0} = (\psi, \psi)_{L^2-1} > 0 \quad \forall \psi \neq 0,
\]
which establishes the strict convexity of the term. The strict convexity of \( G \) follows because each of the other terms is at least convex. The coercivity of \( G \) follows from an estimate of the form
\[
(3.15) \quad G(\varphi) \geq C_1(\varepsilon) \| \varphi \|_{L^2}^2 - C_2(\varepsilon) \quad \forall \varphi \in \tilde{V}_h,
\]
where \( 0 < C_1(\varepsilon), C_2(\varepsilon) < \infty \) are constants. By the standard theory of convex optimization, \( G \) has a unique (bounded) minimizer in \( \tilde{V}_h \), call it \( \varphi^*_m \). Moreover, \( \varphi^*_m \) is the unique minimizer of \( G \) if and only if it is the unique solution to \( A_{\varphi}(\psi) = 0 \), for all \( \psi \in \tilde{V}_h \), where
\[
A_{\varphi}(\psi) := \frac{d}{ds} G(\varphi + s \psi) \bigg|_{s=0} - \frac{1}{\varepsilon} \left[ ((\varphi + K_1)^3, \psi) + \varepsilon (\nabla \varphi, \nabla \psi)_{L^2} \right. \\
+ (\varphi - \varphi^{m-1}_*, \psi)_{L^2-1} - \frac{1}{\varepsilon} (\varphi^{m-1}_h, \psi)_{L^2}.
(3.16)
\]
The rest of the details follow from the definition of the inner product \((\cdot, \cdot)_{L^2-1}\). \( \square \)

Finally, we are in the position to prove the unconditional unique solvability of our scheme.

**Theorem 3.4.** The scheme \( (3.1) - (3.3) \) is uniquely solvable for any mesh parameters \( \tau \) and \( h \) and for any phase parameter \( \varepsilon \).

**Proof.** First it is clear that a necessary condition for solvability of \( (3.2) \) is that
\[
(3.17) \quad (\varphi^*_m, 1)_{L^2} = (\varphi^{m-1}_h, 1)_{L^2} =: K_1,
\]
as can be found by taking \( v_h \equiv 1 \) in \( (3.2) \). Now, let \( (\mu^*_h, \varphi^*_m) \in \tilde{V}_h \times \tilde{V}_h \) be a solution of \( (3.12) - (3.13) \). Define \( \varphi^*_m := \varphi^m + K_1/|\Omega| \). Set
\[
K_2 := \frac{1}{\varepsilon} \left( (\varphi^*_m)^3, 1 \right)_{L^2} - K_1\right).
(3.18)
\]
and define \( \mu^*_h := \mu^*_m + K_2/|\Omega| \). Then it is straightforward to show that \( (\mu^*_h, \varphi^*_m) \in \tilde{V}_h \times \tilde{V}_h \) is a solution to \( (3.2) - (3.3) \). In fact, there is a one-to-one correspondence of the respective solution sets. Namely, if \( (\mu^*_h, \varphi^*_m) \in \tilde{V}_h \times \tilde{V}_h \) is a solution to \( (3.1) - (3.3) \), then \( (\mu^*_m - K_2/|\Omega|, \varphi^*_m - K_1/|\Omega|) \in \tilde{V}_h \times \tilde{V}_h \) is a solution to \( (3.12) - (3.13) \). But \( (3.12) - (3.13) \) admits only a unique solution, which proves that \( (3.1) - (3.3) \) is uniquely solvable. \( \square \)

We now establish a discrete energy law for the numerical scheme that mimics the continuous version \( (2.12) \).

**Lemma 3.5.** Let \( (p^*_h, \mu^*_h, \varphi^*_m) \) denote the unique solution of the scheme \( (3.2) - (3.3) \) and define \( \mathbf{u}^*_h := -\nabla p^*_h - \gamma \varphi^{m-1}_h \nabla \mu^*_h \), then there holds
\[
E_h^* + \tau \sum_{m=1}^l \left\{ \varepsilon \| \nabla \mu^*_h \|_{L^2}^2 + \frac{1}{\gamma} \| \mathbf{u}^*_h \|_{L^2}^2 + \tau \frac{2}{\varepsilon \gamma} \left[ \| \nabla \varphi^*_m \|_{L^2}^2 + \| d_1 \varphi^*_m \|_{L^2}^2 + 2 \| d_1 \varphi^*_m \|_{L^2}^2 \right] \right\} = E_h^0
(3.19)
\]
for all \(0 \leq \ell \leq M\). Here \(E_h^m := \mathcal{J}_\ell(\varphi_h^m)\) and \(\mathcal{J}_\ell(\cdot)\) is defined in \((1.9)\).

Proof. The desired estimate \((3.19)\) follows from setting \(q_h = p_h^m\) in \((3.1)\), \(v_h = \mu_h^m\) in \((3.2)\), \(\psi_h = -d\varphi_h^m\) in \((3.3)\), adding the resulting equations, using the identities

\[
\begin{align*}
(\nabla \varphi_h^m, d_\ell \nabla \varphi_h^m) &= \frac{1}{2} \left[ d_\ell \| \nabla \varphi_h^m \|_{L^2}^2 + \tau \| d_\ell \nabla \varphi_h^m \|_{L^2}^2 \right], \\
(f_h^m, d_\ell \varphi_h^m) &= \frac{1}{4} d_\ell \left( \| \varphi_h^m \|^2 - 1 \right) + \frac{\tau}{4} \left[ \| d_\ell \varphi_h^m \|_{L^2}^2 + 2 \| \varphi_h^m d_\ell \varphi_h^m \|_{L^2}^2 + 2 \| d_\ell \varphi_h^m \|_{L^2}^2 \right],
\end{align*}
\]

and applying the operator \(\tau \sum_{\ell=1}^\ell\) to the combined equation. \(\square\)

The discrete energy law immediately implies the following uniform (in \(\epsilon, h, \tau\)) \(a priori\) estimates for \((p_h^m, \mu_h^m, \varphi_h^m)\).

Lemma 3.6. Let \((p_h^m, \mu_h^m, \varphi_h^m)\) be the unique solution of \((3.1)-(3.5)\) and define \(u_h^m := -\nabla p_h^m - \gamma \varphi_h^{m-1} \nabla \mu_h^m\). Suppose that \(E_h^0 < \infty\). Then, for all \(m \geq 1\),

\[
(3.20) \quad \int_\Omega \varphi_h^m \, dx = \int_\Omega \varphi_h^0 \, dx,
\]

and, in addition, there hold the following estimates:

\[
\begin{align*}
(3.21) \quad & \max_{0 \leq m \leq M} \left[ \epsilon \| \nabla \varphi_h^m \|_{L^2}^2 + \frac{1}{\epsilon} \langle F(\varphi_h^m), 1 \rangle \right] \leq C, \\
(3.22) \quad & \max_{0 \leq m \leq M} \| \varphi_h^m \|_{\mathcal{H}}^2 \leq \frac{C}{\epsilon}, \\
(3.23) \quad & \tau \sum_{m=1}^M \left[ \epsilon \| \nabla \mu_h^m \|_{L^2}^2 + \frac{1}{\gamma} \| u_h^m \|_{L^2}^2 \right] \leq C, \\
(3.24) \quad & \sum_{m=1}^M \left[ \epsilon \| \nabla \varphi_h^m - \nabla \varphi_h^{m-1} \|_{L^2}^2 + \frac{1}{\epsilon} \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \right] \leq C, \\
(3.25) \quad & \sum_{m=1}^M \left[ \| \varphi_h^m (\varphi_h^m - \varphi_h^{m-1}) \|_{L^2}^2 + \| (\varphi_h^m)^2 - (\varphi_h^{m-1})^2 \|_{L^2}^2 \right] \leq C \epsilon, \\
(3.26) \quad & \tau \sum_{m=1}^M \| \nabla p_h^m \|_{L^2}^2 \leq \frac{C}{\epsilon^2}, \\
(3.27) \quad & \tau \sum_{m=1}^M \| d_\ell \varphi_h^m \|_{W^{1,3}}^2 \leq \frac{C}{\epsilon},
\end{align*}
\]

for some \(\epsilon, h,\) and \(\tau\)-independent constant \(C = C(E_h^0) > 0\).

Proof. \((3.20)\) follows immediately from setting \(\nu_h = 1\) in \((3.2)\), and \((3.21)-(3.25)\) are the immediate corollaries of the discrete energy law \((3.19)\). \((3.26)\) follows from the identity \(\nabla p_h^m = u_h^m - \gamma \varphi_h^{m-1} \nabla \mu_h^m\), the Sobolev embedding \(H^1(\Omega) \hookrightarrow L^6(\Omega)\) for \(d = 2, 3\), Young’s inequality and estimates \((3.22)\) and \((3.23)\).

Let \(Q_h\) denote the standard \(L^2\) projection operator into \(V_h\) (cf. \([7,9]\)), and for any \(\nu \in W^{1,3}(\Omega)\), set \(\nu_h = Q_h \nu\) in \((3.2)\). To prove \((3.27)\), we use the Schwarz inequality.
and the Sobolev embedding $H^1(\Omega) \hookrightarrow L^6(\Omega)$ (for $d = 2, 3$) to we get

$$(d_t \varphi_h^m, \nu) = (d_t \varphi_h^m, \mathcal{Q}_h \nu)$$

$$= -\varepsilon (\nabla \mu_h^m, \nabla \mathcal{Q}_h \nu) + (\varphi_h^{m-1}, \nabla \mathcal{Q}_h \nu)$$

$$\leq \varepsilon \| \nabla \mu_h^m \|_{L^2} \| \nabla \mathcal{Q}_h \nu \|_{L^2} + \| \varphi_h^{m-1} \|_{L^6} \| \mathcal{Q}_h \nu \|_{L^3}$$

$$\leq C \varepsilon \| \nabla \mu_h^m \|_{L^2} + C \| \mu_h^{m-1} \|_{H^1} \| \mathcal{Q}_h \nu \|_{L^3}$$

$$\leq C \varepsilon \| \nabla \mu_h^m \|_{L^2} + \frac{1}{\sqrt{\varepsilon}} \| \mathcal{Q}_h \nu \|_{L^2} \| \nabla \mathcal{Q}_h \nu \|_{L^3},$$

where we have used the $W^{1,3}$ stability of the $L^2$ projection $\mathcal{Q}_h$ (cf. [7, 9]) to get the last inequality. Property (3.27) now follows immediately from the above inequality and estimates (3.21) and (3.23). The proof is complete. \( \square \)

**Remark 3.1.** Property (3.20) says that the proposed numerical method enjoys the same mass conservation law as the phase field model (cf. (2.11)). This property will be validated numerically in Sec. 5.

4. Convergence analysis. The goal of this section is to prove that the fully discrete finite element solution has a unique accumulation point (in some function space) and this accumulation point is necessarily a weak solution to problem (2.8)–(2.10). A byproduct of this convergence result is to provide a constructive proof of the existence of weak solutions to problem (2.8)–(2.10).

First, we derive some additional estimates for the finite element solution. To the end, we introduce the discrete Laplacian $\Delta_h : V_h \to V_h$ which is defined as follows: for any $v_h \in V_h$, $\Delta_h v_h \in V_h$ denotes the unique solution to the problem

$$(\Delta_h v_h, w_h) = - (\nabla v_h, \nabla w_h) \quad \forall w_h \in V_h.$$  

In particular, setting $w_h = \Delta_h v_h$ in (4.1), we obtain

$$\| \Delta_h v_h \|_{L^2}^2 = - (\nabla v_h, \nabla \Delta_h v_h).$$

**Lemma 4.1.** Let $(p_h^m, \mu_h^m, \varphi_h^m)$ and $\mathcal{Q}_h^m$ be same as in Lemma 3.6. Then, under the assumption $E_h^0 < \infty$, there hold the following additional estimates: for $d = 2, 3$,

$$\tau \sum_{m=1}^M \| \Delta_h \varphi_h^m \|_{L^2}^2 \leq C_0 (T + 1) C(\varepsilon),$$

$$\tau \sum_{m=1}^M \| \varphi_h^m \|_{L^6}^{\frac{4(d-5)}{d}} \leq C_0 (T + 1) C(\varepsilon),$$

$$\tau \sum_{m=1}^M \| \nabla \varphi_h^m \|_{L^4} \leq C_0 (T + 1) C(\varepsilon),$$

$$\tau \sum_{m=1}^M \| \nabla p_h^m \|_{L^2}^{\frac{4(d-5)}{d}} \leq C_0 (T + 1) C(\varepsilon),$$

$$\tau \sum_{m=1}^M \| d_t \varphi_h^m \|_{H^1}^{\frac{4(d-5)}{d}} \leq C_0 (T + 1) C(\varepsilon),$$

for some $\varepsilon, h$, and $\tau$-independent constant $C_0 = C_0(E_h^0) > 0$ and some $h$ and $\tau$-independent constant $C(\varepsilon) > 0$ that grows like $\varepsilon^{-r}$, for some $r \in \mathbb{Z}$, as $\varepsilon \to 0$. 

We now appeal to the following discrete Gagliardo-Nirenberg inequality (cf. [18] and [11]),

\[ \varepsilon \| \Delta_h \varphi_h^m \|_{L^2}^2 = -\varepsilon (\nabla \varphi_h^m, \nabla \Delta_h \varphi_h^m) \]

\[ = - (\mu_h^m, \Delta_h \varphi_h^m) + \frac{1}{\varepsilon} (f_h^m, \Delta_h \varphi_h^m) \]

\[ \leq (\nabla \mu_h^m, \nabla \varphi_h^m) + \frac{1}{\varepsilon} \left( \frac{\varepsilon^2}{2} \| \Delta_h \varphi_h^m \|_{L^2}^2 + \frac{1}{2\varepsilon^2} \| f_h^m \|_{L^2}^2 \right) \]

\[ \leq \frac{1}{2} \| \nabla \mu_h^m \|_{L^2}^2 + \frac{1}{2} \| \nabla \varphi_h^m \|_{L^2}^2 + \frac{\varepsilon}{2} \| \Delta_h \varphi_h^m \|_{L^2}^2 + \frac{1}{2\varepsilon^2} \| f_h^m \|_{L^2}^2 . \]

Hence,

\[ (4.7) \quad \varepsilon \| \Delta_h \varphi_h^m \|_{L^2}^2 \leq \| \nabla \mu_h^m \|_{L^2}^2 + \| \nabla \varphi_h^m \|_{L^2}^2 + \frac{1}{\varepsilon^2} \| f_h^m \|_{L^2}^2 . \]

To bound \( \| f_h^m \|_{L^2}^2 \), we write

\[ f_h^m := (\varphi_h^m)^3 - \varphi_h^{m-1} = \varphi_h^m ((\varphi_h^m)^2 - 1) + \varphi_h^m - \varphi_h^{m-1}. \]

Then by (3.21) we have

\[ \| f_h^m \|_{L^2}^2 \leq 2 \| \varphi_h^m \|_{L^\infty}^2 \| (\varphi_h^m)^2 - 1 \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \]

\[ = 8 \| \varphi_h^m \|_{L^\infty}^2 (F(\varphi_h^m), 1) + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \]

\[ \leq C \varepsilon \| \varphi_h^m \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 . \]

We now appeal to the following discrete Gagliardo-Nirenberg inequality (cf. [18] and [11]):

\[ (4.8) \quad \| \varphi_h^m \|_{L^\infty} \leq C \| \Delta_h \varphi_h^m \|_{L^2}^{\frac{d-4}{d}} \| \varphi_h^m \|_{L^2}^{\frac{3(4-d)}{2d}} + C \| \varphi_h^m \|_{L^6} \quad (d = 2, 3) \]

and get

\[ (4.9) \quad \| f_h^m \|_{L^2}^2 \leq C \varepsilon \| \Delta_h \varphi_h^m \|_{L^2}^{\frac{d}{d-4}} \| \varphi_h^m \|_{L^2}^{\frac{3(4-d)}{2d}} + C \varepsilon \| \varphi_h^m \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \]

\[ \leq \varepsilon \left( \frac{\varepsilon^3}{2} \| \Delta_h \varphi_h^m \|_{L^2}^2 + C \varepsilon \| \varphi_h^m \|_{L^2}^2 \right) + C \varepsilon \| \varphi_h^m \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \]

\[ \leq \frac{\varepsilon^4}{2} \| \Delta_h \varphi_h^m \|_{L^2}^2 + C \left( \varepsilon \frac{4-2d}{4-2d} + \varepsilon \right) \| \varphi_h^m \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 \]

\[ \leq \frac{\varepsilon^4}{2} \| \Delta_h \varphi_h^m \|_{L^2}^2 + C \left( \varepsilon \frac{4-2d}{4-2d} + \varepsilon \right) \| \varphi_h^m \|_{L^2}^2 + \| \varphi_h^m - \varphi_h^{m-1} \|_{L^2}^2 , \]

where we used the Sobolev embedding \( H^1(\Omega) \hookrightarrow L^6(\Omega) \) for \( d = 2, 3 \) in the last step.

Then (4.2) follows from applying the operator \( \tau \sum_{m=1}^M \) to (4.7) and using (4.9), (3.22), and (3.23). (4.3) is an immediate consequence of (4.8) and (4.2).

To prove (4.4), we recall another discrete Gagliardo-Nirenberg inequality (cf. [18] and [11]):

\[ (4.10) \quad \| \nabla \nu_h \|_{L^2} \leq C \| \nabla \nu_h \|_{L^2}^\frac{d-2}{2} \| \Delta_h \nu_h \|_{L^2}^{\frac{d}{2}} + C \| \nabla \nu_h \|_{L^2} \quad \forall \nu_h \in V_h, \quad d = 2, 3. \]
It follows from the above inequality and estimates \((3.21)\) and \((4.2)\) that
\[
\tau \sum_{m=1}^{M} \| \nabla \phi_h^m \|_{L^2}^2 \leq C \tau \sum_{m=1}^{M} \| \nabla \phi_h^m \|_{L^2}^{2(\delta-d)} \| \Delta_h \phi_h^m \|_{L^2}^2 + C \sum_{m=1}^{M} \| \nabla \phi_h^m \|_{L^2}^{\frac{8}{\epsilon}}
\]
which proves \((4.4)\).

Inequality \((4.5)\) follows from the estimate
\[
\| \nabla p_h^m \|_{L^2} \leq C \| u_h^m \|_{L^2}^{2(\delta-d)} + C \| \nu^{m-1} \|_{L^\infty} \| \nabla \mu_h^m \|_{L^2}
\]
and estimates \((3.23)\) and \((4.3)\).

Now, let \(Q_h\) denote the standard \(L^2\) projection operator into \(V_h\) (cf. \(11\)). For any \(\nu \in H^1(\Omega)\), setting \(\nu_h = Q_h \nu\) in \((3.2)\), we get
\[
(d_t \phi_h^m, \nu) = (d_t \phi_h^m, Q_h \nu)
\]
\[
= -\varepsilon (\nabla \mu_h^m, \nabla Q_h \nu) + (\phi_h^{m-1} u_h^m, \nabla Q_h \nu)
\]
\[
\leq \left[ \varepsilon \| \nabla \mu_h^m \|_{L^2} + \| \phi_h^{m-1} \|_{L^\infty} \| u_h^m \|_{L^2} \right] \| \nabla Q_h \nu \|_{L^2}
\]
\[
\leq C \left[ \varepsilon \| \nabla \mu_h^m \|_{L^2} + \| \phi_h^{m-1} \|_{L^\infty} \| u_h^m \|_{L^2} \right] \| \nabla \nu \|_{L^2},
\]
where we have used the \(H^1\) stability of the \(L^2\) projection \(Q_h\) (cf. \(11\)) to get the last inequality.\(\Box\) now follows immediately from the above inequality and estimates \((4.3)\) and \((3.23)\). The proof is complete. \(\Box\)

Next, let \(\phi_{h,t}(x,t)\) denote the piecewise linear interpolant (in \(t\)) of the fully discrete solution \(\{\phi_h^m\}\), that is,
\[
\phi_{h,t}(\cdot,t) := \frac{t-t_{m-1}}{\tau} \phi_h^m(\cdot) + \frac{t_{m-1} - t}{\tau} \phi_h^{m-1}(\cdot) \quad \forall t \in [t_{m-1}, t_m],
\]
for \(1 \leq m \leq M\). Let \(\overline{p}_{h,t}(x,t), \overline{u}_{h,t}(x,t), \overline{\phi}_{h,t}(x,t), \overline{\varphi}_{h,t}(x,t), \) and \(\overline{\varphi}_{h,t}(x,t)\), denote the piecewise constant extensions of \(\{p_h^m\}\), \(\{u_h^m\}\), \(\{\mu_h^m\}\), and \(\{\phi_h^m\}\), respectively, defined as follows
\[
\overline{p}_{h,t}(\cdot,t) := p_h^m \quad \forall t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M,
\]
\[
\overline{u}_{h,t}(\cdot,t) := u_h^m \quad \forall t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M,
\]
\[
\overline{\mu}_{h,t}(\cdot,t) := \mu_h^m \quad \forall t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M,
\]
\[
\overline{\phi}_{h,t}(\cdot,t) := \phi_h^m \quad \forall t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M,
\]
\[
\overline{\varphi}_{h,t}(\cdot,t) := \phi_h^{m-1} \quad \forall t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M.
\]

We remark that \(\phi_{h,t}(x,t)\) is a continuous piecewise polynomial function in space and time, \(\overline{p}_{h,t}(x,t), \overline{u}_{h,t}(x,t), \overline{\phi}_{h,t}(x,t), \) and \(\overline{\varphi}_{h,t}(x,t)\) are right continuous at the nodes \(\{t_m\}\), and \(\overline{\varphi}_{h,t}(x,t)\) is left continuous at the nodes \(\{t_m\}\).
The main result of this section is the following convergence theorem.

**Theorem 4.2.** Let Ω ⊂ \(\mathbb{R}^d\) (\(d = 2, 3\)) be a polygonal or polyhedral domain. For each fixed \(\varepsilon > 0\), suppose that \(\mathcal{J}_\varepsilon(\varphi_{0h}) \leq C_0 < \infty\), where \(C_0\) is independent of \(h\), and

\[
\lim_{h \to 0} \|\varphi_{0h} - \varphi_0\|_{L^2} = 0.
\]

Then the sequence \(\{\varphi_{h,\tau}, \overline{u}_{h,\tau}, \overline{p}_{h,\tau}, \overline{\varphi}_{h,\tau}\}\) has an accumulation point \((p^\varepsilon, u^\varepsilon, \mu^\varepsilon, \varphi^\varepsilon)\) with \(u^\varepsilon = -\nabla p^\varepsilon - \gamma \varphi^\varepsilon \nabla \mu^\varepsilon\), and \((p^\varepsilon, \mu^\varepsilon, \varphi^\varepsilon)\) is a weak solution to problem \((2.8) + (2.10)\).

**Proof.** We divide the proof into two steps.

**Step 1: Extracting convergent subsequences.** The estimates of Lemmas 3.6 and 4.1 immediately give the following (uniform in \(h\) and \(\tau\)) estimates:

\[
\begin{align*}
\|\nabla \varphi_{h,\tau}\|_{L^\infty(L^2)} + \|\varphi_{h,\tau}^2 - 1\|_{L^\infty(L^2)} &\leq C; \\
\|\overline{u}_{h,\tau}\|_{L^2(L^2)} + \|\nabla \overline{p}_{h,\tau}\|_{L^2(L^2)} &\leq C; \\
\|\nabla \overline{p}_{h,\tau}\|_{L^2(L^2)} + \|\nabla \overline{p}_{h,\tau}\|_{L^\infty(L^2)} &\leq C; \\
\|\overline{\varphi}_{h,\tau}\|_{L^\infty(L^\infty)} &\leq C; \\
\|\nabla \overline{\varphi}_{h,\tau}\|_{L^\infty(L^2)} &\leq C; \\
\|\overline{\varphi}_{h,\tau}\|_{L^2((W^{1,3}_t)^\ast)} + \|\{(\overline{\varphi}_{h,\tau})_t\|_{L^2((H^1_t)^\ast)} &\leq C,
\end{align*}
\]

where \(\beta := \frac{4(d-6)}{d} \geq 4\) and \(\sigma := \frac{4(d-6)}{2d} > \frac{4}{3}\). Note, we have suppressed the dependences of the constants on \(T\) and \(\varepsilon\) above. \(\{\nabla \overline{p}_{h,\tau}\}\) is uniformly (with respect to \(h\) and \(\tau\)) integrable in \(L^\sigma((0, T); L^2(\Omega))\) and \(\{(\overline{\varphi}_{h,\tau})_t\}\) is uniformly integrable in \(L^\sigma((0, T); (H^1_t)^\ast)\).

Then there exists a convergent subsequence of \(\{(\overline{p}_{h,\tau}, \overline{u}_{h,\tau}, \overline{p}_{h,\tau}, \overline{\varphi}_{h,\tau})\}\) (still denoted by the same symbols) and a quadruple \((p^\varepsilon, u^\varepsilon, \mu^\varepsilon, \varphi^\varepsilon)\) such that

\[
\begin{align*}
p^\varepsilon &\in L^\sigma((0, T); H^1(\Omega) \cap L^3_0(\Omega)), \\
\varphi^\varepsilon &\in L^\infty((0, T); H^1(\Omega)) \cap L^{4d}((0, T); L^\infty(\Omega)), \\
\phi^\varepsilon &\in L^\sigma((0, T); (H^1(\Omega))^\ast) \cap L^2((0, T); (W^{1,3}(\Omega))^\ast), \\
\mu^\varepsilon &\in L^2((0, T); H^1(\Omega)),
\end{align*}
\]

and

\[
\begin{align*}
\overline{p}_{h,\tau} &\overset{h,\tau \to 0}{\rightharpoonup} p^\varepsilon \text{ weakly in } L^\sigma((0, T); H^1(\Omega) \cap L^3_0(\Omega)), \\
\overline{u}_{h,\tau} &\overset{h,\tau \to 0}{\rightharpoonup} u^\varepsilon \text{ weakly in } L^2((0, T); L^2(\Omega)), \\
\overline{p}_{h,\tau} &\overset{h,\tau \to 0}{\rightharpoonup} \mu^\varepsilon \text{ weakly in } L^2((0, T); H^1(\Omega)), \\
\overline{\varphi}_{h,\tau} &\overset{h,\tau \to 0}{\rightharpoonup} \varphi^\varepsilon \text{ weakly* in } L^\infty((0, T); H^1(\Omega)) \cap L^3((0, T); L^\infty(\Omega)) \text{ strongly in } L^2((0, T); L^2(\Omega)),
\end{align*}
\]

weakly in \(H^1((0, T); (W^{1,3}(\Omega))^\ast) \cap W^{1,\sigma}((0, T); (H^1(\Omega))^\ast)\), weakly in \(L^\frac{3}{2}((0, T); W^{1,4}(\Omega))\).

We have used Aubin-Lions lemma (cf. [25]) to conclude (4.27).
From (3.24) we also have
\[
\|\phi_{h,\tau} - \bar{\phi}_{h,\tau}\|_{L^2(\Omega)}^2 = \sum_{m=1}^{M} \|\phi_{h}^m - \phi_{h}^{m-1}\|_{H^1}^2 + \int_{t_{m-1}}^{t_m} \left( \frac{t - t_{m-1}}{\tau} \right)^2 dt \\
= \frac{\tau}{3} \sum_{m=1}^{M} \|\phi_{h}^m - \phi_{h}^{m-1}\|_{H^1}^2 \rightarrow 0.
\]
Hence, \(\{\phi_{h,\tau}\}, \{\overline{\phi}_{h,\tau}\}\), and \(\{\bar{\phi}_{h,\tau}\}\) converge to the same limit as \(h, \tau \to 0\).

**Step 2: Passing to the limit.** We now want to pass to the limit in (3.1)–(3.4) and to show that \((p^\varepsilon, mu^\varepsilon, \varphi^\varepsilon)\) is a weak solution to problem (2.8)–(2.10) with the initial data \(\varphi^\varepsilon(0) = \varphi_0^\varepsilon\). To this end, we rewrite (3.1)–(3.4) as
\[
(\mathbf{u}_{h,\tau}, \nabla q_h) = 0 \quad \forall q_h \in W_h, \\
(\varphi_{h,\tau} + \varepsilon(\nabla \varphi_{h,\tau}, \nabla \nu_h) - \bar{\varphi}_{h,\tau}, \nabla \nu_h) = 0 \quad \forall \nu_h \in V_h, \\
(\mathbf{p}_{h,\tau} + \varepsilon(\nabla \mathbf{p}_{h,\tau}, \nabla \psi_h) - \frac{1}{\varepsilon} (f_{h,\tau}, \psi_h) = 0 \quad \forall \psi_h \in V_h,
\]
where \(f_{\varepsilon, h, \tau}\) denotes the right continuous constant extension of \(\{f_h^m\}\).

For any \(\eta \in C^0([0,T]),\) multiplying (4.28)–(4.30) by \(\eta,\) respectively, and integrating the resulting equations in \(t\) from 0 to \(T\) we get
\[
\int_0^T (\mathbf{u}_{h,\tau}, \nabla q_h) \eta(t) dt = 0 \quad \forall q_h \in W_h, \\
\int_0^T (\varphi_{h,\tau} + \varepsilon(\nabla \varphi_{h,\tau}, \nabla \nu_h) - \bar{\varphi}_{h,\tau}, \nabla \nu_h) \eta(t) dt = 0 \quad \forall \nu_h \in V_h, \\
\int_0^T (\mathbf{p}_{h,\tau} + \varepsilon(\nabla \mathbf{p}_{h,\tau}, \nabla \psi_h) - \frac{1}{\varepsilon} (f_{h,\tau}, \psi_h) \eta(t) dt = 0 \quad \forall \psi_h \in V_h.
\]
For any \((q, \nu, \psi) \in [H^1(\Omega) \cap C^1(\Omega)]^3,\) let \((q_h, \nu_h, \psi_h) \in W_h \times V_h \times V_h\) be the standard finite element (nodal) interpolations of \((q, \nu, \psi)\) in (4.31)–(4.33). Since
\[
q_h \rightharpoonup q, \quad \nu_h \rightharpoonup \nu, \quad \psi_h \rightharpoonup \psi \quad \text{strongly in } H^1(\Omega),
\]
sending \(h, \tau \to 0\) in (4.31) and (4.33) and using (4.25)–(4.26) we get \(\varphi^\varepsilon(0) = \varphi_0^\varepsilon\) and
\[
\int_0^T (\mathbf{u}^\varepsilon, \nabla q) \eta(t) dt = 0 \quad \forall q \in H^1(\Omega), \\
\int_0^T (\varphi^\varepsilon + \varepsilon(\nabla \mu^\varepsilon, \nabla \nu) + (\varphi^\varepsilon \mathbf{u}^\varepsilon, \nabla \nu) \eta(t) dt = 0 \quad \forall \nu \in H^1(\Omega), \\
\int_0^T (\mu^\varepsilon, \psi) - \varepsilon(\nabla \varphi^\varepsilon, \nabla \psi) - \frac{1}{\varepsilon} (f(\varphi^\varepsilon), \psi) \eta(t) dt = 0 \quad \forall \psi \in H^1(\Omega).
\]
Moreover, from the identity \(\mathbf{u}_h^m = -\nabla p_h^m - \gamma \varphi_h^m \nabla \mu_h^m\) we have
\[
\int_0^T (\mathbf{u}_{h,\tau}, \nabla q) \eta(t) dt = -\int_0^T (\nabla \overline{p}_{h,\tau} + \gamma \varphi_{h,\tau} \nabla \mathbf{p}_{h,\tau}, \nabla q) \eta(t) dt, \\
\int_0^T (\nabla \mathbf{p}_{h,\tau}, \nabla \nu) \eta(t) dt = -\int_0^T (\nabla \overline{p}_{h,\tau} + \gamma \varphi_{h,\tau} \nabla \mathbf{p}_{h,\tau}, \nabla \nu) \eta(t) dt.
\]
Sending \( h, \tau \to 0 \) and using (4.24)–(4.26) yields

\[
\int_0^T (u^\epsilon, \nabla q) \eta(t) \, dt = - \int_0^T (\nabla p^\epsilon + \gamma \varphi^\epsilon \nabla \mu^\epsilon, \nabla q) \eta(t) \, dt,
\]

(4.37)

\[
\int_0^T (\varphi^\epsilon u^\epsilon, \nabla \nu) \eta(t) \, dt = - \int_0^T (\varphi^\epsilon \nabla p^\epsilon + \gamma \varphi^\epsilon \nabla \mu^\epsilon, \nabla \nu) \eta(t) \, dt.
\]

(4.38)

Combining (4.34)–(4.36) and (4.37)–(4.38) we obtain (2.8)–(2.10), since \( C^0[0,T] \) is dense in \( L^2(0,T) \). Hence, \((p^\epsilon, \mu^\epsilon, \varphi^\epsilon)\) is a weak solution to (2.8)–(2.10). The proof is complete. \( \Box \)

**Corollary 4.3.** The whole sequence \( \{(p_h, \mu_h, \varphi_h)\} \) converges if weak solutions to problem (2.8)–(2.10) are unique.

**Proof.** We have shown in the above proof that \( \{(p_h, \mu_h, \varphi_h)\} \) has a convergent subsequence and its limit \( (p^\epsilon, \mu^\epsilon, \varphi^\epsilon) \) is a weak solution of (2.8)–(2.10). Moreover, the proof also implies that the limit of every convergent subsequence of \( \{(p_h, \mu_h, \varphi_h)\} \) is necessarily a weak solution of (2.8)–(2.10). Hence, by the uniqueness assumption of weak solutions, the whole sequence \( \{(p_h, \mu_h, \varphi_h)\} \) must converge to the unique weak solution. \( \Box \)

**Remark 4.1.** In Theorem 4.2 and Corollary 4.3, \( \Omega \) is assumed to be a polygonal or polyhedral domain. This assumption is imposed only to avoid the technicalities for defining our finite element method (3.1)–(3.5). It is not used or needed in the proofs of the theorem and the corollary. By using the standard numerical integration technique or the approximated boundary technique (i.e., to approximate a bounded Lipschitz domain by a sequence of polygonal or polyhedral domains) (cf. [9]), it can be proved that the modified finite element methods would also possess all the properties proved in Lemmas 3.1, 3.2, 3.3, 3.5, 3.6, and Theorem 3.4 as well as Lemma 4.1. As a result, the conclusions of Theorem 4.2 and Corollary 4.3 still hold when \( \Omega \) is a bounded Lipschitz domain.

From Theorem 4.2 and Theorem 2.4 we immediately have

**Theorem 4.4.** There exists a weak solution to problem (2.8)–(2.10) and weak solutions are unique in the function class \( \mathcal{F} \).

We conclude this section with a remark on the error estimates for the solution of the fully discrete scheme (3.1)–(3.4). Using the standard (perturbation) technique as presented in [13], it is not hard to prove that the scheme converges optimally in the energy norm. However, the error constant would contain a factor of \( \exp(\varepsilon^{-2}) \). Such an error bound is clearly not very useful for small \( \varepsilon \). A better error bound would only depend on \( \varepsilon^{-1} \) in some low polynomial order (cf. [13, 15]). Deriving such a polynomial order error bound is an on-going project and the result will be reported in a forthcoming paper.

5. Numerical experiments. In this section we provide some numerical experiments to gauge the accuracy and reliability of the fully discrete finite element method developed in the previous sections. For the experiments we take \( V_h = W_h = S^1_h \) for simplicity. We use a square domain \( \Omega = (0,1)^2 \) and take \( T_h \) to be a regular triangulation of \( \Omega \) consisting of right isosceles triangles, as depicted in Fig. A.1. We use a nonlinear multigrid method, which is detailed in Appendix A, to solve the scheme (3.1)–(3.4) at each time step. We perform a battery of three tests on the scheme. First, we measure numerical convergence of the scheme in the presence of added, artificial source terms. Second, we measure the numerical convergence of the scheme without source terms using a Cauchy-convergence method. Third, we conduct a test
corresponding continuous problem is precisely
\[ \begin{align*} L^2 \text{ convergence test. The final time is } T = 1.0, \text{ and the refinement path is taken to be } \tau = 25.6h^2. \text{ The other parameters are } \varepsilon = \gamma = 1.0; \Omega = (0, 1)^2. \text{ The global error at } T \text{ is expected to be } O(\tau) + O(h^2) = O(h^2), \text{ and this is confirmed.} \end{align*} \]

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\( h \) & \( |||\varphi|||_{L^2} \) & \text{rate} & \( |||\mu|||_{L^2} \) & \text{rate} & \( |||\rho|||_{L^2} \) & \text{rate} \\
\hline
\( \sqrt{2}/16 \) & \( 8.683 \times 10^{-3} \) & – & \( 1.088 \times 10^{-2} \) & – & \( 1.270 \times 10^{-2} \) & – \\
\( \sqrt{2}/32 \) & \( 1.850 \times 10^{-3} \) & 2.23 & \( 2.701 \times 10^{-3} \) & 2.01 & \( 2.479 \times 10^{-3} \) & 2.35 \\
\( \sqrt{2}/64 \) & \( 4.568 \times 10^{-4} \) & 2.01 & \( 6.759 \times 10^{-4} \) & 2.00 & \( 5.759 \times 10^{-4} \) & 2.11 \\
\( \sqrt{2}/128 \) & \( 1.141 \times 10^{-4} \) & 2.00 & \( 1.691 \times 10^{-4} \) & 2.00 & \( 1.413 \times 10^{-4} \) & 2.03 \\
\( \sqrt{2}/256 \) & \( 2.852 \times 10^{-5} \) & 2.00 & \( 4.227 \times 10^{-5} \) & 2.00 & \( 3.515 \times 10^{-5} \) & 2.00 \\
\hline
\end{tabular}
\caption{Table 5.1}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\( h \) & \( |||\varphi|||_{H^1} \) & \text{rate} & \( |||\mu|||_{H^1} \) & \text{rate} & \( |||\rho|||_{H^1} \) & \text{rate} \\
\hline
\( \sqrt{2}/16 \) & \( 2.886 \times 10^{-1} \) & – & \( 2.907 \times 10^{-1} \) & – & \( 2.943 \times 10^{-1} \) & – \\
\( \sqrt{2}/32 \) & \( 1.455 \times 10^{-1} \) & 0.99 & \( 1.462 \times 10^{-1} \) & 0.99 & \( 1.466 \times 10^{-1} \) & 1.01 \\
\( \sqrt{2}/64 \) & \( 7.290 \times 10^{-2} \) & 1.00 & \( 7.320 \times 10^{-2} \) & 1.00 & \( 7.313 \times 10^{-2} \) & 1.00 \\
\( \sqrt{2}/128 \) & \( 3.647 \times 10^{-2} \) & 1.00 & \( 3.660 \times 10^{-2} \) & 1.00 & \( 3.653 \times 10^{-2} \) & 1.00 \\
\( \sqrt{2}/256 \) & \( 1.824 \times 10^{-2} \) & 1.00 & \( 1.839 \times 10^{-2} \) & 1.00 & \( 1.826 \times 10^{-2} \) & 1.00 \\
\hline
\end{tabular}
\caption{Table 5.2}
\end{table}

of spinodal decomposition using varying values of the excess surface tension \( \gamma \), and demonstrate the discrete energy dissipation and mass conservation properties of the scheme.

For the convergence of the problem with source terms, we solve a problem of the following form: find \((p_h^m, \mu_h^m, \varphi_h^m) \in V_h \times V_h \times V_h\), such that
\begin{align*}
\text{(5.1)} & \quad (\nabla p_h^m + \gamma \varphi_h^{m-1} \nabla \mu_h^m, \nabla q_h) = (s_1(x, t_m), q_h) \quad \forall q_h \in V_h, \\
\text{(5.2)} & \quad (d \varphi_h^m, \nu_h) + \varepsilon (\nabla \mu_h^m, \nabla \nu_h) \\
& \quad + (\varphi_h^{m-1} [\nabla p_h^m + \gamma \varphi_h^{m-1} \nabla \mu_h^m], \nabla \nu_h) = (s_2(x, t_m), \nu_h) \quad \forall \nu_h \in V_h, \\
\text{(5.3)} & \quad (\mu_h^m, \psi_h) - \varepsilon (\nabla \varphi_h^m, \nabla \psi_h) - \frac{1}{\varepsilon} (f_h^m, \psi_h) = (s_3(x, t_m), \psi_h) \quad \forall \psi_h \in V_h, \\
\text{(5.4)} & \quad \varphi_h^0 = \varphi_{0h},
\end{align*}
for \( m = 1, \ldots, M \), where the source terms are chosen so that the solution of the corresponding continuous problem is precisely
\begin{align*}
\text{(5.5)} & \quad p(x, y, t) = \mu(x, y, t) = \varphi(x, y, t) = \cos(\pi t) \cdot g(x) \cdot g(y),
\end{align*}
with \( g(\xi) = 16\xi^2(\xi - 1)^2 \). The initial data are precisely given by \( \varphi_{0h} = I_h(\varphi(\cdot, 0)) \), where \( I_h : H^2(\Omega) \to V_h \) is the standard nodal interpolation operator. All integrations are done exactly using the appropriate Gauss-quadrature rules. This is of course made possible since we are using polynomials in space. The exact values of all of the other parameters used in the test are given in the captions of Tables 5.1 and 5.2. The results of an \( L^2 \) error analysis using a quadratic refinement path are found in Tab. 5.1 and confirm the expected optimal second-order convergence rate in this case. The results of an \( H^1 \) error analysis using a linear refinement path are found in Tab. 5.2 and confirm the expected optimal first-order convergence rate for this case. Notice that
use the same randomized initial data for the three simulations represented in Fig. 5.1, \( \tau \)

where we solve the scheme (5.1)–(5.3) with \( s \)

the approximations \( \delta \)

Tab. 5.4 and confirm first-order convergence in this case. \( L^2 \) Cauchy convergence test. The final time is \( T = 4.0 \times 10^{-2} \), and the refinement path is taken to be \( \tau = 1.024h^2 \). The other parameters are \( \varepsilon = 6.25 \times 10^{-2} \); \( \gamma = 1.25 \times 10^{-1} \); \( \Omega = (0,1)^2 \). The Cauchy difference is defined via \( \delta_\varphi := \varphi_{h_f} - \varphi_{h_i} \), where the approximations are evaluated at time \( t = T \), and analogously for \( \delta_\mu \), and \( \delta_\psi \). The norm of the Cauchy difference at \( T \) is expected to be \( O(\tau) + O(h^2) = O(h^2) \).

\[
\begin{array}{ccccccccc}
 h_c & h_f & \|\delta_\varphi\|_{L^2} & \text{rate} & \|\delta_\mu\|_{L^2} & \text{rate} & \|\delta_\psi\|_{L^2} & \text{rate} \\
\sqrt{2}/16 & \sqrt{2}/32 & 5.514 \times 10^{-2} & - & 2.890 \times 10^{-1} & - & 3.099 \times 10^{-2} & - \\
\sqrt{2}/32 & \sqrt{2}/64 & 2.165 \times 10^{-2} & 1.35 & 1.292 \times 10^{-1} & 1.23 & 1.148 \times 10^{-2} & 1.43 \\
\sqrt{2}/64 & \sqrt{2}/128 & 6.284 \times 10^{-3} & 1.78 & 3.588 \times 10^{-2} & 1.78 & 3.250 \times 10^{-3} & 1.82 \\
\sqrt{2}/128 & \sqrt{2}/256 & 1.636 \times 10^{-3} & 1.94 & 9.327 \times 10^{-3} & 1.94 & 8.420 \times 10^{-4} & 1.95 \\
\sqrt{2}/256 & \sqrt{2}/512 & 4.132 \times 10^{-4} & 1.99 & 2.355 \times 10^{-3} & 1.99 & 2.128 \times 10^{-4} & 1.98 \\
\end{array}
\]

Tab. 5.3

\[
\begin{array}{ccccccccc}
 h_c & h_f & \|\delta_\varphi\|_{H^1} & \text{rate} & \|\delta_\mu\|_{H^1} & \text{rate} & \|\delta_\psi\|_{H^1} & \text{rate} \\
\sqrt{2}/16 & \sqrt{2}/32 & 8.569 \times 10^{-1} & - & 1.301 \times 10^{0} & - & 8.371 \times 10^{-2} & - \\
\sqrt{2}/32 & \sqrt{2}/64 & 4.160 \times 10^{-1} & 1.04 & 6.295 \times 10^{-1} & 1.04 & 3.715 \times 10^{-1} & 1.17 \\
\sqrt{2}/64 & \sqrt{2}/128 & 2.061 \times 10^{-1} & 1.01 & 3.111 \times 10^{-1} & 1.02 & 1.779 \times 10^{-2} & 1.06 \\
\sqrt{2}/128 & \sqrt{2}/256 & 1.029 \times 10^{-1} & 1.00 & 1.554 \times 10^{-1} & 1.00 & 8.834 \times 10^{-3} & 1.01 \\
\sqrt{2}/256 & \sqrt{2}/512 & 5.146 \times 10^{-2} & 1.00 & 7.777 \times 10^{-2} & 1.00 & 4.422 \times 10^{-3} & 1.00 \\
\end{array}
\]

Tab. 5.4

\( H^1 \) Cauchy convergence test. The final time is \( T = 4.0 \times 10^{-2} \), and the refinement path is taken to be \( \tau = 2.0 \times 10^{-1} h \). The other parameters are \( \varepsilon = 6.25 \times 10^{-2} \); \( \gamma = 1.25 \times 10^{-1} \); \( \Omega = (0,1)^2 \). The norm of the Cauchy difference at \( T \) is expected to be \( O(\tau) + O(h) = O(h) \).

the approximations \( p_{h}^{0}, \varphi_{h}^{m}, \mu_{h}^{m} \) all appear to converge to the same optimal rates, in both cases.

We now give the results of a test without any artificial sources. In other words, we solve the scheme (5.1)–(5.3) with \( s_i = 0 \), \( i = 1, 2, 3 \). The initial data are taken to be

\[
\varphi_{0h} = I_h \left( \frac{[1.0 - \cos(4.0 \pi x)] \cdot [1.0 - \cos(2.0 \pi y)]}{2} - 1.0 \right),
\]

and the parameters are given in the captions of Tabs. 5.3 and 5.4. Note that in this case we are not in possession of the exact solutions. To circumvent this, we measure the difference of the computed solutions at successive resolutions. Specifically, we compute the rate at which the Cauchy difference \( \delta_\varphi := \psi_{h_f}^{M_f} - \psi_{h_c}^{M_c} \) converges to zero, where \( h_f = 2h_c \); \( \tau_f = 2^p \tau_c \) \( (p = 1 \text{ for a linear refinement path and } p = 2 \text{ for a quadratic refinement path}) \), and \( \tau_f M_f = \tau_c M_c = T \). A quadratic refinement path, \( \text{i.e.}, \tau = Ch^2 \), is taken when measurements are made in the \( L^2 \) norm, and a linear refinement path, \( \text{i.e.}, \tau = Ch \), when measurements are made in the \( H^1 \) norm. The results of an \( L^2 \) Cauchy error analysis are found in Tab. 5.3 and confirm second-order convergence in this case. The results of an \( H^1 \) Cauchy error analysis are found in Tab. 5.4 and confirm first-order convergence in this case.

Our final test is a simulation of spinodal decomposition with different values of \( \gamma \). Specifically, we solve the scheme (5.1)–(5.3) with \( s_i = 0 \), \( i = 1, 2, 3 \), and with three values of \( \gamma \); namely, \( \gamma = 0 \), which yields the familiar Cahn-Hilliard model; \( \gamma = 0.01 \); and \( \gamma = 0.04 \). Furthermore, we take \( \varepsilon = 0.01, h = \sqrt{\frac{2}{35}} \), \( \tau = 1 \times 10^{-3} \), and \( T = 0.1 \). We use the same randomized initial data for the three simulations represented in Fig. 5.1.
Fig. 5.1. Spinodal decomposition for three values of $\gamma$. The domain is $\Omega = (0, 1) \times (0, 1)$ and $\varepsilon = 0.01$. The initial data are exactly the same for the three simulations. The time step size is $\tau = 1.0 \times 10^{-3}$, and $h = \sqrt[3]{2} / 256$. We use a uniform mesh, as in Fig. A.1. The corresponding energy plots are shown in Fig. 5.2. The average value of $\varphi$ for all three simulations is approximately $-0.1$. For $\gamma = 0.01$, the mass variation over the simulated time is only $1 \times 10^{-12}$. The max and min values of $\varphi$ are very near the values $+1$ and $-1$, respectively.

where the average value of $\varphi$ is approximately $-0.1$. As expected, the mixture phase separates into domains wherein $\varphi \approx -1$ and $\varphi \approx +1$. Afterwards the system coarsens, as larger phase regions grow at the expense of smaller ones. The energy for the three simulations is displayed in Fig. 5.2. A general trend emerges, where, at least at early times, the energy decreases faster and the coarsening process is appears to be accelerated as the excess surface tension $\gamma$ increases. This behavior is expected and was observed in similar finite difference calculations undertaken in [28].

Note that we have proved that (at the theoretical level) the energy is non-increasing at each time step. This is observed in our computations. In addition to this, the mass, i.e., $\int_{\Omega} \varphi h \, dx$, at the theoretical level is expected to be unchanging from one time step to the next. On the practical level, we observe very little mass variation. For example, for the $\gamma = 0.01$ case depicted in Fig. 5.1, where initially $\int_{\Omega} \varphi_0 h \, dx \approx -0.1$, we observe mass variation of only $1 \times 10^{-12}$ over the whole of the
Fig. 5.2. Energy plots for the spinodal decomposition simulations depicted in Fig. 5.1. The parameters for the simulations are given in the caption of Fig. 5.1. The energy is observed to decrease at each time step. The general trend, at least at early times, is that the energy decreases faster with increasing values of the excess surface tension $\gamma$.

Simulation time. Note that our multigrid iteration stopping tolerance is of the same order, namely, $\text{tol} = 1 \times 10^{-12}$ in (A.35).

Acknowledgments. The authors would like to thank Professor Xiaoming Wang of Florida State University for his helpful discussions and for bringing the reference [27] to their attention.

Appendix A. Nonlinear Multigrid Solver. In this appendix we give the full details of the nonlinear multigrid solver that is used to march the scheme in time. Suppose $\Omega \subseteq \mathbb{R}^2$ is polygonal, and assume that $\mathcal{T}_\ell$, $\ell = 0, 1, \ldots, L$, is a hierarchy of nested triangulations of $\Omega$ as suggested in Fig. 5.1. In particular, $\mathcal{T}_\ell$ is obtained by subdividing the triangles of $\mathcal{T}_{\ell-1}$ into 4 congruent sub-triangles. Note that $h_{\ell-1} = 2h_\ell$, $\ell = 1, \ldots, L$, and that $\{\mathcal{T}_\ell\}$ is a quasi-uniform family. For simplicity, we shall use $P_1$ finite element spaces and use the same space for the pressure as is used for the other variables. We define

$$V_\ell = \{v \in C^0(\Omega) \mid v|_K \in P_1(K) \; \forall K \in \mathcal{T}_\ell\},$$

for $\ell = 0, \ldots, L$ and observe the nested space chain $V_0 \subset V_1 \subset V_3 \subset \cdots \subset V_L$. Because of this nestedness, there is a natural injection operation $I_{\ell-1,\ell} : V_{\ell-1} \hookrightarrow V_\ell$ defined by $I_{\ell-1,\ell}(v) = v$, for all $v \in V_{\ell-1}$, $\ell = 1, \ldots, L$. Now, let $B_\ell = \{u_{\ell,i}(x)\}_{i=1}^{N_\ell}$ be the nodal basis for $V_\ell$, $\ell = 0, 1, \ldots, L$. In other words, $u_{\ell,j}(x)_{i} = \delta_{i,j}$, where $\{x_{\ell,i}\}_{i=1}^{N_\ell}$ are the nodes of $\mathcal{T}_\ell$. We have level-wise representations of the unknowns of the form

$$\varphi_\ell(x) = \sum_{i=1}^{N_\ell} \varphi_{\ell,i}u_{\ell,i}(x) \iff \varphi_\ell = (\varphi_{\ell,1}, \varphi_{\ell,2}, \ldots, \varphi_{\ell,N_\ell})^T,$$

and similarly for $\mu_\ell(x)$ and $p_\ell(x)$. Define the prolongation matrix via $P_{\ell-1,\ell} = I_{\ell-1,\ell}$, where $I_{\ell-1,\ell}$ is the $N_\ell \times N_{\ell-1}$ matrix representation of the injection operator $I_{\ell-1,\ell}$ with respect to the bases $B_{\ell-1}$ and $B_\ell$. There are two restriction operations — i.e., operations transferring information from the finer space $V_\ell$ to the coarser space $V_{\ell-1}$...
that we shall use. The first is called the canonical restriction and, in matrix form, is the \( N_{\ell-1} \times N_{\ell} \) matrix defined via \( R_{\ell,\ell-1} := I_{\ell}^T I_{\ell-1,\ell} \) [6] [7]. The second is defined via

\[
R_{\ell,\ell-1}(v) = \sum_{i=1}^{N_{\ell-1}} v(x_{\ell-1,i}) u_{\ell-1,i}(x) \quad \forall v \in V_\ell,
\]

where the points \( x_{\ell-1,i} \) are the nodes of the mesh \( T_{\ell-1} \). Note that \( \{x_{\ell-1,i}\}_{i=1}^{N_{\ell-1}} \subset \{x_{\ell,i}\}_{i=1}^{N_{\ell}} \) by construction. By \( \hat{R}_{\ell,\ell-1} \) we denote the matrix representation of \( \hat{R}_{\ell,\ell-1} \) with respect to the bases \( B_{\ell} \) and \( B_{\ell-1} \).

In the present framework, our nonlinear finite element scheme is defined on the finest level, \( \ell = L \), as follows: find the triple \((p_L, \mu_L, \varphi_L) \in V_L \times V_L \times V_L\) such that

\[
\begin{align*}
(\nabla p_L + \gamma \varphi_L^{m-1} \nabla \mu_L, \nabla q_L) &= 0 \quad \forall q_L \in V_L, \\
(\varphi_L, \nu_L) + \tau \varepsilon (\nabla \mu_L, \nabla \nu_L) &+ \nu_L \left( \varphi_L^{m-1} + \nu_L \right) = (\varphi_L^{m-1}, \nu_L) \quad \forall \nu_L \in V_L, \\
(\mu_L, \psi_L) - \varepsilon (\nabla \varphi_L, \nabla \psi_L) - \frac{1}{\varepsilon} \left( (\varphi_L)^2 \varphi_L, \psi_L \right) &= -\frac{1}{\varepsilon} \left( \varphi_L^{m-1}, \psi_L \right) \quad \forall \psi_L \in V_L,
\end{align*}
\]

where \( \varphi_L^{m-1} \in V_L \) is given. We have dropped the superscript \( m \) (the time step index) on the unknowns for simplicity. Theorem 3.4 guarantees that this problem always has a unique solution. The nonlinear system \((A.3) - (A.5)\) may be written as

\[
\begin{align*}
A_L p_L + \gamma C_L \mu_L &= 0, \\
M_L \varphi_L + \tau (\varepsilon A_L + \gamma B_L) \mu_L + \tau C_L p_L &= M_L \varphi_L^{m-1}, \\
\varepsilon A_L \varphi_L + \frac{1}{\varepsilon} Q_L (\varphi_L) \varphi_L - M_L \mu_L &= \frac{1}{\varepsilon} M_L \varphi_L^{m-1},
\end{align*}
\]

where \( A_L, B_L, C_L, M_L, \) and \( Q_L (\varphi_L) \) are \( N_L \times N_L \) matrices whose components are

\[
\begin{align*}
[A_L]_{i,j} &:= (\nabla u_{L,j}, \nabla u_{L,i}), & [B_L]_{i,j} &:= \left( (\varphi_L^{m-1})^2 \nabla u_{L,j}, \nabla u_{L,i} \right), \\
[C_L]_{i,j} &:= (\varphi_L^{m-1} \nabla u_{L,j}, \nabla u_{L,i}), & [M_L]_{i,j} &:= (u_{L,j}, u_{L,i}), \\
[Q_L (\varphi_L)]_{i,j} &:= \left( (\varphi_L)^2 u_{L,j}, u_{L,i} \right).
\end{align*}
\]

We solve \((A.6) - (A.8)\) using a nonlinear multigrid method [6] Ch. 5, §6. This
requires that we split the equations into source (s) and operator (N) terms:

\[(A.12) \quad s_L^{(1)} := 0, \quad N_L^{(1)} (\phi_L) := A_L p_L + \gamma C_L \mu_L,\]
\[(A.13) \quad s_L^{(2)} := M_L \phi_L^{-1}, \quad N_L^{(2)} (\phi_L) := M_L \phi_L + \tau (\varepsilon A_L + \gamma B_L) \mu_L + \tau C_L p_L,\]
\[(A.14) \quad s_L^{(3)} := \frac{1}{\varepsilon} M_L \phi_L^{-1}, \quad N_L^{(3)} (\phi_L) := \varepsilon A_L \phi_L + \frac{1}{\varepsilon} Q_L (\phi_L) \phi_L - M_L \mu_L,\]

where \(\phi_L := [p_L, \mu_L, \varphi_L]\) is the \(N_L \times 3\) array of unknowns. We must also define a “consistent” version of the nonlinear operator on all of the coarser levels. There are a number of ways to proceed in this task [6]; we choose the following path. Suppose that \(\ell \in \{0, 1, \ldots, L - 1\}\) is given. We restrict the known solution from the previous time step to the coarser levels via

\[(A.15) \quad \varphi_{\ell}^{m-1} := \prod_{j=\ell+1}^{L} \hat{R}_{j,j-1} \varphi_{\ell,j-1}^{m-1} = \left(\varphi_{\ell,1}^{m-1}, \varphi_{\ell,2}^{m-1}, \ldots, \varphi_{\ell,N_{\ell}}^{m-1}\right)^T,\]
\[\iff \quad \varphi_{\ell}^{m-1}(x) = \sum_{i=1}^{N_{\ell}} \varphi_{\ell,i}^{m-1} u_{\ell,i}(x).\]

Now, given any \(\psi_\ell \in V_\ell\) with the representation

\[(A.16) \quad \psi_\ell = (\psi_{\ell,1}, \psi_{\ell,2}, \ldots, \psi_{\ell,N_{\ell}})^T \iff \psi_\ell(x) = \sum_{i=1}^{N_{\ell}} \psi_{\ell,i} u_{\ell,i}(x),\]

we define

\[(A.17) \quad [A_\ell]_{i,j} := (\nabla u_{\ell,j}, \nabla u_{\ell,i}), \quad [B_\ell]_{i,j} := \left(\left(\varphi_{\ell}^{m-1}\right)^2 \nabla u_{\ell,j}, \nabla u_{\ell,i}\right),\]
\[(A.18) \quad [C_\ell]_{i,j} := (\varphi_{\ell}^{m-1} \nabla u_{\ell,j}, \nabla u_{\ell,i}), \quad [M_\ell]_{i,j} := (u_{\ell,j}, u_{\ell,i}),\]
\[(A.19) \quad [Q_\ell (\psi_\ell)]_{i,j} := \left(\left(\psi_\ell\right)^2 u_{\ell,j}, u_{\ell,i}\right).\]

Observe that

\[(A.20) \quad A_\ell = R_{\ell+1,\ell} A_{\ell+1} P_{\ell,\ell+1}, \quad M_\ell = R_{\ell+1,\ell} M_{\ell+1} P_{\ell,\ell+1},\]

which is standard in the finite element setting [6, 7] and is the reason for the term “canonical” describing \(R_{\ell+1,\ell}\). On the other hand,

\[(A.21) \quad B_\ell \approx R_{\ell+1,\ell} B_{\ell+1} P_{\ell,\ell+1}, \quad C_\ell \approx R_{\ell+1,\ell} C_{\ell+1} P_{\ell,\ell+1}.\]

(Note that we could have recursively defined \(B_\ell = R_{\ell+1,\ell} B_{\ell+1} P_{\ell,\ell+1}\), and similarly for \(C_\ell\). But it turns out that this is an unnecessary complication from the point of view of the convergence of the algorithm.) Finally, we have

\[(A.22) \quad N_\ell^{(1)} (\xi_\ell) := A_\ell q_\ell + \gamma C_\ell \nu_\ell,\]
\[(A.23) \quad N_\ell^{(2)} (\xi_\ell) := M_\ell \psi_\ell + \tau (\varepsilon A_\ell + \gamma B_\ell) \nu_\ell + \tau C_\ell q_\ell,\]
\[(A.24) \quad N_\ell^{(3)} (\xi_\ell) := \varepsilon A_\ell \varphi_\ell + \frac{1}{\varepsilon} Q_\ell (\psi_\ell) \psi_\ell - M_\ell \nu_\ell,\]

where \(\xi_\ell := [q_\ell, \nu_\ell, \psi_\ell]\) is any given \(N_\ell \times 3\) array of unknowns.
We are now in a position to define the recursive nonlinear multigrid V-Cycle operator \([6, Ch. 5, §6]\), which is the heart of our solver. In the following the superscript \(k\) is the V-Cycle loop index (not the time step index). Let \(\phi_{\ell}^{k-1} := [p_{\ell}^{k-1}, \mu_{\ell}^{k-1}, \varphi_{\ell}^{k-1}]\) denote the current, level-\(\ell\) multigrid iterate. For any \(N_{\ell} \times 3\) array of unknowns \(\xi_{\ell}\), define \(N_{\ell}(\xi_{\ell}) := [N_{\ell}^{(1)}(\xi_{\ell}), N_{\ell}^{(2)}(\xi_{\ell}), N_{\ell}^{(3)}(\xi_{\ell})]\), and \(s_{\ell} := [s_{\ell}^{(1)}, s_{\ell}^{(2)}, s_{\ell}^{(3)}]\). Note that these last two objects are \(N_{\ell} \times 3\) arrays by design. We define the action of the recursive nonlinear multigrid V-Cycle operator

\[
(\text{A.25}) \quad \phi_{\ell}^{k} = \text{NMGM}(\ell, \phi_{\ell}^{k-1}, N_{\ell}, s_{\ell}, \lambda)
\]

in the following 3 steps:

1. Pre-smoothing:
   - Given \(\phi_{\ell}^{k-1}\), compute a smoothed level-\(\ell\) approximation \(\bar{\phi}_{\ell}\):
     \[
     (\text{A.26}) \quad \bar{\phi}_{\ell} = S^{\lambda}(\phi_{\ell}^{k-1}, N_{\ell}, s_{\ell}),
     \]
     where \(S\) is a smoothing (or relaxation) operator, and \(\lambda > 0\) is the number of smoothing sweeps.
   - Compute coarse-grid correction:
     - Compute coarse-level initial iterate:
       \[
       (\text{A.27}) \quad \bar{\phi}_{\ell-1} = R_{\ell,\ell-1} \bar{\phi}_{\ell}.
       \]
     - Compute the coarse-level right-hand side:
       \[
       (\text{A.28}) \quad s_{\ell-1} = R_{\ell,\ell-1} (s_{\ell} - N_{\ell}(\bar{\phi}_{\ell})) + N_{\ell-1}(\bar{\phi}_{\ell-1}).
       \]
     - Compute an approximate solution \(\hat{\xi}_{\ell-1}\) of the following coarse grid equation:
       \[
       (\text{A.29}) \quad N_{\ell-1}(\xi_{\ell-1}) = s_{\ell-1}.
       \]
       Note that this equation is uniquely solvable by Theorem 3.4.
         - If \(\ell = 1\) employ \(\lambda\) smoothing steps:
           \[
           (\text{A.30}) \quad \hat{\xi}_{0} = S^{\lambda}(\bar{\phi}_{0}, N_{0}, s_{0}).
           \]
         - If \(\ell > 1\) get an approximate solution to Eq. \((\text{A.29})\) using \(\bar{\phi}_{\ell-1}\) as initial guess:
           \[
           (\text{A.31}) \quad \hat{\xi}_{\ell-1} = \text{NMGM}(\ell - 1, \bar{\phi}_{\ell-1}, N_{\ell-1}, s_{\ell-1}, \lambda).
           \]
     - Compute the coarse-grid correction:
       \[
       (\text{A.32}) \quad \hat{\phi}_{\ell-1} = \hat{\xi}_{\ell-1} - \bar{\phi}_{\ell-1}.
       \]
     - Compute the coarse-grid-corrected approximation at level \(k\):
       \[
       (\text{A.33}) \quad \hat{\phi}_{\ell} = P_{\ell-1,\ell} \hat{\phi}_{\ell-1} + \bar{\phi}_{\ell}.
       \]
2. Post-smoothing:
• Finally, compute $\phi_k^\ell$ by applying $\lambda$ smoothing steps:

$$\phi_k^\ell = S^\lambda \left( \hat{\phi}_\ell, N_\ell, s_\ell \right).$$

When

$$\sqrt{\frac{1}{3N_L} \sum_{j=1}^{3} \sum_{i=1}^{N_L} \left( \left[ s_{L}^{(j)} - N_{L}^{(j)} \left( \phi_L^k \right) \right] \right)^2} < \text{tol}$$

we stop iterating and set $\phi_L^k \rightarrow \phi_L = [p_L, \mu_L, \varphi_L]$, the fine-level solution. For smoothing, we use a nonlinear block Gauss-Seidel method, like that discussed in [28] for a similar finite-difference nonlinear multigrid method. The exact details are omitted for brevity, but the principal idea is that the nodal values $(p_{L,i}, \mu_{L,i}, \varphi_{L,i})$ are always obtained simultaneously in the smoothing operation. We use $\lambda = 2$ or 3 in the smoothing step.

REFERENCES

[1] R. A. Adams, Sobolev Spaces, Academic press, New York, 1975.
[2] D. M. Anderson and G. B. McFadden, Diffuse-interface method in fluid mechanics, Annual Review of Fluid Mech, vol. 30 (1998), pp. 139–165.
[3] N. D. Alikakos, P. W. Bates, and X. Chen, Convergence of the Cahn-Hilliard equation to the Hele-Shaw model, Arch. Rational Mech. Anal., 128 (1994), pp. 165–205.
[4] J. Bear, Dynamics of Fluids in Porous Media, Dover Publications, Inc., New York, 1972.
[5] H. Blum and R. Rannacher, On the boundary value problem of biharmonic operator on domains with angular corners, Math. Methods Appl. Sci., 2 (1980), pp. 556–581.
[6] D. Braess, Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics, third edition, Cambridge, 2007.
[7] S. C. Brenner and L. R. Scott, The Mathematical Theory of Finite Element Methods, third edition, Springer, 2008.
[8] J. W. Cahn and J. E. Hilliard, Free energy of a nonuniform system I. Interfacial free energy, J. Chem. Phys., 28 (1958), pp. 258–267.
[9] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[10] C. M. Elliott, D. A. French, and F. A. Milner, A second order splitting method for the Cahn-Hilliard equation, Numer. Math., 54 (1989), pp. 575–590.
[11] C. M. Elliott and Z. Songmu, On the Cahn-Hilliard equation, Arch. Rational Mech. Anal., 96 (1986), pp. 339–357.
[12] X. Feng, Fully discrete finite element approximations of the Navier-Stokes-Cahn-Hilliard diffuse interface model for two-phase fluids, SIAM J. Numer. Anal., 40 (2002), pp. 1019–1072.
[13] X. Feng, Y. He, and C. Liu, Analysis of finite element approximations of a phase field model for two-phase fluids, Math. Comp., 76 (2007), pp. 539–571.
[14] X. Feng and A. Prohl, Error analysis of a mixed finite element method for the Cahn-Hilliard equation, Numer. Math, 99 (2004), pp. 47–84.
[15] X. Feng and A. Prohl, Numerical analysis of the Cahn-Hilliard equation and approximation for the Hele-Shaw problem, Interfaces and Free Boundaries, 7 (2005), pp. 1–28.
[16] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Second Edition, Springer, New York, 2000.
[17] H. S. Hele-Shaw, The flow of water, Nature, 58 (1898), pp. 34–35.
[18] J. G. Heywood and R. Rannacher, Finite element approximation of the non-stationary Navier-Stokes problem I. Regularity of solutions and second-order error estimates for spatial discretization, SIAM J. Numer. Anal., 19 (1982), pp. 275–311.
[19] O. A. Ladyženskaja, V. A. Solonnikov and N. N. Ural’tseva, Linear and quasilinear equations of parabolic type, Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1967.
[20] H.-G. Lee, J. Lowengrub and J. Goodman, Modeling pinch-off and reconstruction in a Hele-Shaw cell. I. The models and their calibration, Phys Fluids, 14 (2002), pp. 492–513.
[21] H.-G. Lee, J. Lowengrub and J. Goodman, *Modeling pinch-off and reconstruction in a Hele-Shaw cell. I. The analysis and simulation in the nonlinear regime*, Phys Fluids, 14 (2002), pp. 514–545.

[22] J. Lowengrub and I. Truskinovsky, *Cahn-Hilliard fluids and topological transitions*, Proc. R. Soc. London A, 454 (1998), pp. 2617–2654.

[23] G. B. McFadden, *Phase field models of solidification*, Contemporary Mathematics, 295 (2002), pp. 107–145.

[24] Q. Nie and F. Tian, *Singularities in Hele-Shaw Flows*, SIAM J. on Appl. Math., 58 (1998), pp. 34-54.

[25] J. Simon, *Compact sets in the space $L^p(0,T;B)$*, Ann. Mat. Pura Appl., (1986), pp. 65–96.

[26] R. Scholz, *A mixed method for 4th order problems using linear finite elements*, RAIRO Anal. Numér., 12 (1978), pp. 85–90.

[27] X. Wang and Z. Zhang, *Well-posedness of the Hele-Shaw-Cahn-Hilliard system*, Ann. de l'Inst. Henri Poincare (C) Nonli. Anal., DOI:10.1016/j.anihpc.2012.06.003.

[28] S.M. Wise, *Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn-Hilliard-Hele-Shaw system of equations*, J. Sci. Comput. 44 (2010) 38-68.