Interfaces and Free Boundaries **7** (2005), 1–28

Numerical analysis of the Cahn–Hilliard equation and approximation for the Hele–Shaw problem

XIAOBING FENG[†]
Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, U.S.A.

AND

ANDREAS PROHL[‡]
Department of Mathematics, ETHZ, CH-8092 Zurich, Switzerland

[Received 4 March 2003 and in revised form 5 March 2004]

This paper concerns numerical approximations for the Cahn–Hilliard equation $u_t + \Delta(\varepsilon\Delta u - \varepsilon^{-1}f(u)) = 0$ and its sharp interface limit as $\varepsilon \searrow 0$, known as the Hele–Shaw problem. The primary goal of this paper is to establish the convergence of the solution of the fully discrete mixed finite element scheme proposed in [29] to the solution of the Hele–Shaw (Mullins–Sekerka) problem, provided that the Hele–Shaw (Mullins–Sekerka) problem has a global (in time) classical solution. This is accomplished by establishing some improved a priori solution and error estimates, in particular, an $L^\infty(L^\infty)$ error estimate, and making full use of the convergence result of [2]. The cruxes of the analysis are to establish stability estimates for the discrete solutions, use a spectrum estimate result of Alikakos and Fusco [3] and Chen [15], and establish a discrete counterpart of it for a linearized Cahn–Hilliard operator to handle the nonlinear term.

Keywords: Cahn–Hilliard equation; Hele–Shaw (Mullins–Sekerka) problem; phase transition; biharmonic problem; fully discrete mixed finite element method; Ciarlet–Raviart element.

1. Introduction

In [29] we proposed and analyzed a semi-discrete (in time) and a fully discrete mixed finite element method for the Cahn–Hilliard equation:

$$u_t + \Delta\left(\varepsilon\Delta u - \frac{1}{\varepsilon}f(u)\right) = 0 \qquad \text{in } \Omega_T := \Omega \times J,\ J := (0,T), \tag{1.1}$$

$$\frac{\partial u}{\partial n} = \frac{\partial}{\partial n}\left(\varepsilon\Delta u - \frac{1}{\varepsilon}f(u)\right) = 0 \qquad \text{in } \partial\Omega_T := \partial\Omega \times J, \tag{1.2}$$

$$u = u_0^\varepsilon \qquad \text{in } \Omega \times \{0\}, \tag{1.3}$$

where $f(u) = F'(u)$ and F is a double well potential. Note that the super-index ε on the solution u^ε is suppressed for notational brevity. We established a priori solution estimates and optimal and quasi-optimal error estimates under *minimum regularity assumptions* on the domain $\Omega \subset \mathbb{R}^N$ ($N = 2, 3$) and the initial datum function u_0^ε. Special attention was given to the dependence of the error bounds on ε. It was shown that all the error bounds depend on $1/\varepsilon$ only in some low polynomial order for small ε.

[†]E-mail: xfeng@math.utk.edu
[‡]E-mail: apr@math.ethz.ch

In this paper, we are concerned with the second stage of the evolution of the concentration, that is, the motion of the interface. We focus on approximating the Hele–Shaw (Mullins–Sekerka) problem:

$$\Delta w = 0 \qquad \text{in } \Omega \setminus \Gamma_t,\ t \in [0, T], \tag{1.4}$$

$$\frac{\partial w}{\partial n} = 0 \qquad \text{on } \partial\Omega,\ t \in [0, T], \tag{1.5}$$

$$w = \sigma\kappa \qquad \text{on } \Gamma_t,\ t \in [0, T], \tag{1.6}$$

$$V = \frac{1}{2}\left[\frac{\partial w}{\partial n}\right]_{\Gamma_t} \qquad \text{on } \Gamma_t,\ t \in [0, T], \tag{1.7}$$

$$\Gamma_0 = \Gamma_{00} \qquad \text{when } t = 0 \tag{1.8}$$

via the Cahn–Hilliard equation as $\varepsilon \searrow 0$. Here

$$\sigma = \int_{-1}^{1} \sqrt{\frac{F(s)}{2}}\, \mathrm{d}s,$$

and κ and V are, respectively, the mean curvature and the normal velocity of the interface Γ_t, n is the unit outward normal to either $\partial\Omega$ or Γ_t,

$$\left[\frac{\partial w}{\partial n}\right]_{\Gamma_t} := \frac{\partial w^+}{\partial n} - \frac{\partial w^-}{\partial n},$$

and w^+ and w^- are respectively the restriction of w to Ω_t^+ and Ω_t^-, the exterior and interior of Γ_t in Ω (cf. [2, 29]). We remark that the orientation of Γ_t is chosen such that the unit outward normal n on Γ_t is pointing to Ω_t^+.

Numerical approximations for the Cahn–Hilliard equation with a *fixed* ε have been studied by several authors in the past fifteen years. Elliott and Zheng [24] analyzed a (continuous in time) semi-discrete conforming finite element discretization in one space dimension. Numerical experiments of the method in one space dimension were reported in [21]. Elliott and French [22] proposed a (continuous in time) semi-discrete nonconforming finite element method based on the Morley nonconforming finite element method [11, 17]. Optimal order error estimates were also established for the nonconforming method under the assumption that the solution is smooth. Elliott, French and Milner [23] proposed and analyzed a (continuous in time) semi-discrete splitting finite element method (mixed finite element method) which approximates simultaneously the concentration u and the chemical potential w. Optimal order error estimates were shown under the assumption that the finite element approximation u_h of the concentration u is bounded in L^∞. Later, Du and Nicolaides [19] analyzed a fully discrete splitting finite element method in one space dimension under weaker regularity assumptions on the solution u of the Cahn–Hilliard equation, and established optimal order error estimates by first proving the boundedness of u_h in L^∞. Copetti and Elliott [18] considered the Cahn–Hilliard equation with a nonsmooth logarithmic potential function. A fully discrete splitting finite element method was proposed and convergence of the method was also demonstrated. In one space dimension, French and Jensen [30] analyzed the long time behavior of the (continuous time) semi-discrete conforming hp-finite element approximations. Recently, extensive studies have been carried out by Barrett and Blowey and others on the finite element approximations of the Cahn–Hilliard system for multi-component alloys with constant or degenerate mobility; we refer to [5, 6, 7] and the references therein for detailed expositions. We like to emphasize that the results cited above were established for the Cahn–Hilliard equation with a *fixed* "interaction length" ε. No special effort and attention were given to address issues such as how the mesh sizes h and k depend on ε and how the error bounds depend on ε. In fact, since all error estimates were derived using a Gronwall inequality type argument at the end of the derivations, it is not hard to check that these error bounds contain a factor $\exp(T/\varepsilon)$, which clearly is not very useful when $\varepsilon \to 0$.

The main objective of this paper is to establish the convergence of the fully discrete mixed finite element method proposed in [29] to the solution of the Hele–Shaw problem (1.4)–(1.8) as $h, k, \varepsilon \to 0$, provided that the Hele–Shaw problem has a global (in time) classical solution. To our knowledge, such a numerical convergence result is not yet known in the literature for the Cahn–Hilliard equation. We also note that the convergence of the Cahn–Hilliard equation to the Hele–Shaw model was established in [2] under the same assumption.

To show convergence, we need to establish stronger error estimates, in particular, an $L^\infty(J; L^\infty)$ estimate. We are able to obtain the desired error estimates by first proving some improved a priori solution estimates, and then an improved discrete spectrum estimate under the assumption that the Hele–Shaw problem admits a global (in time) classical solution. As in [29], the cruxes of the analysis are to establish stability estimates for a discrete solution, use a spectrum estimate result of Alikakos and Fusco [3] and Chen [15], and establish a discrete counterpart of it for a linearized Cahn–Hilliard operator to handle the nonlinear term.

We also remark that parallel studies using a similar approach were also carried out by the authors in [28, 27] for the Allen–Cahn equation and the related curvature driven flows, and for the classical phase field model and the related Stefan problems, respectively. On the other hand, unlike the Allen–Cahn equation which is an L^2 gradient flow, the Cahn–Hilliard equation is an H^{-1} gradient flow; this makes the analysis for the Cahn–Hilliard equation much more delicate and complicated than that for the Allen–Cahn equation given in [28].

The paper is organized as follows: In Section 2, we shall derive some improved a priori estimates for the solution of (1.1)–(1.3) under the condition that the Hele–Shaw problem has a global (in time) classical solution. Special attention is given to dependence of the solution on ε in various norms. In Section 3, we analyze the fully discrete mixed finite element method proposed in [29] for the Cahn–Hilliard equation, which consists of the backward Euler discretization in time and the lowest order Ciarlet–Raviart mixed finite element (for the biharmonic operator) discretization in space. Optimal and quasi-optimal error estimates in stronger norms, including the $L^\infty(J; L^\infty)$ norm, are obtained for the fully discrete solution. It is shown that all the error bounds depend on $1/\varepsilon$ only in low polynomial orders for small ε. Finally, Section 4 is devoted to establishing the convergence of the fully discrete solution to the solution of the Hele–Shaw problem. Using the $L^\infty(J; L^\infty)$ error estimate and the convergence result of [2], we show that the fully discrete numerical solution converges to the solution (including the free boundary) of the Hele–Shaw problem, provided that the latter admits a global (in time) classical solution.

This paper is a condensed and revised version of [26], where one can find more details, and some additional results as well as helpful comments which could not be included here due to page limitation.

2. Energy estimates for the differential problem

In this section, we derive some energy estimates in various function spaces up to $L^\infty(J; H^4(\Omega)) \cap H^1(J; H^3(\Omega))$ in terms of negative powers of ε for the solution u of the Cahn–Hilliard problem

(1.1)–(1.3) for given $u_0^\varepsilon \in H^4(\Omega)$. The basic estimates are derived under general (minimum) regularities, while the improved estimates are established under the assumption that the Hele–Shaw problem admits a global (in time) solution. Throughout this paper, we assume that $\Omega \subset \mathbb{R}^N$ $(N = 2, 3)$ is a bounded domain with *smooth* boundary $\partial\Omega$. The standard notation is also adopted in this paper (cf. [29]), in particular, Δ^{-1} and $\Delta^{-1/2}$ stand for the inverse Laplacian and its gradient. For their detailed definitions, we refer to Section 2 of [29]. Again, C and $\widetilde{C}$ are used to denote generic positive constants which are independent of ε and the time and space mesh sizes k and h.

In this paper, we are mainly concerned with the second stage of the evolution of the concentration u, that is, the motion of the interface, and focus on approximating the Hele–Shaw problem via the Cahn–Hilliard equation discretized by a fully discrete mixed finite element method. For these purposes, we rewrite (1.1)–(1.3) as

$$u_t = \Delta w \qquad \text{in } \Omega_T, \tag{2.1}$$

$$w = \frac{1}{\varepsilon} f(u) - \varepsilon \Delta u \qquad \text{in } \Omega_T, \tag{2.2}$$

$$\frac{\partial u}{\partial \nu} = \frac{\partial w}{\partial \nu} = 0 \qquad \text{on } \partial\Omega \times (0, T), \tag{2.3}$$

$$u(x, 0) = u_0^\varepsilon(x) \qquad \forall x \in \Omega, \tag{2.4}$$

where w physically represents the chemical potential. We refer to [24, 10] and references therein for more discussions on well-posedness and regularities of the Cahn–Hilliard and the biharmonic problems. Unless stated otherwise, we define $w_0^\varepsilon(x) := w(x, 0)$ by setting $t = 0$ in (2.2).

As in [28, 29], we consider the following general double equal-well potential function F:

GENERAL ASSUMPTION 1 (GA$_1$) 1) $f = F'$ for $F \in C^3(\mathbb{R})$ such that $F(\pm 1) = 0$, and $F > 0$ elsewhere.

2) For some finite $p > 2$ and positive numbers $\tilde{c}_i > 0$, $i = 0, \ldots, 3$,

$$\tilde{c}_1 |a|^{p-2} - \tilde{c}_0 \leqslant f'(a) \leqslant \tilde{c}_2 |a|^{p-2} + \tilde{c}_3 \quad \forall a \in \mathbb{R}.$$

3) There exist $0 < \gamma_1 \leqslant 1$, $\gamma_2 > 0$, $\delta > 0$ and $C > 0$ such that

(i) $(f(a) - f(b), a - b) \geqslant \gamma_1 (f'(a)(a - b), a - b) - \gamma_2 |a - b|^{2+\delta} \quad \forall |a| \leqslant 2C,$

(ii) $a f''(a) \geqslant 0 \quad \forall |a| \geqslant C.$

REMARK We note that the above (GA$_1$) differs slightly from those of [29] in 2) and 3). It is trivial to check that (GA$_1$)$_2$ implies

$$-(f'(u)v, v) \leqslant \tilde{c}_0 \|v\|_{L^2}^2 \quad \forall v \in L^2(\Omega), \tag{2.5}$$

which will be utilized several times in the paper.

EXAMPLE The potential function $F(u) = \frac{1}{4}(u^2 - 1)^2$, and consequently, $f(u) = u^3 - u$, is often used in physical and geometrical applications [4, 12, 8, 2, 16]. For convenience, we verify (GA$_1$)$_1$–(GA$_1$)$_3$. First, (GA$_1$)$_1$ holds trivially. Since $f'(u) = 3u^2 - 1$, (GA$_1$)$_2$ holds with $\tilde{c}_1 = \tilde{c}_2 = 3$, $p = 4$ and $\tilde{c}_0 = \tilde{c}_3 = 1$. A direct calculation gives

$$f(a) - f(b) = (a - b)[f'(a) + (a - b)^2 - 3(a - b)a]. \tag{2.6}$$

Hence, (GA$_1$)$_3$ holds with $\gamma_1 = 1$, $\gamma_2 = 3$, $\delta = 1$ and any constant $C_0 \geqslant 0$. Also, (2.5) holds with $\tilde{c}_0 = 1$.

REMARK In this paper, we mainly consider the case $\gamma_1 = 1$; the analysis for this case is harder than that for the case $0 < \gamma_1 < 1$. We refer to [26] for the analysis of the latter case.

In the rest of this section, we shall establish some basic and improved a priori estimates for the solution of the Cahn–Hilliard equation under the assumption that the Hele–Shaw problem (1.4)–(1.8) has a global (in time) classical solution (cf. [29]). These improved a priori estimates are necessary for us to obtain error estimates in stronger norms in the next section.

LEMMA 2.1 Suppose that f satisfies (GA_1). Then the solution of (2.1)–(2.4) satisfies the following estimates:

$$\text{(i)}\quad \operatorname*{ess\,sup}_{[0,\infty)}\left\{\frac{\varepsilon}{2}\|\nabla u\|_{L^2}^2 + \frac{1}{\varepsilon}\|F(u)\|_{L^1}\right\} + \left\{\begin{matrix}\int_0^\infty \|u_t\|_{H^{-1}}^2\,\mathrm{d}s\\ \int_0^\infty \|\nabla w\|_{L^2}^2\,\mathrm{d}s\end{matrix}\right\} = \mathcal{J}_\varepsilon(u_0^\varepsilon),$$

$$\text{(ii)}\quad \operatorname*{ess\,sup}_{[0,\infty)} \|u\|_{L^p}^p \leqslant C(1 + \mathcal{J}_\varepsilon(u_0^\varepsilon)) \qquad (p \text{ as in } (\mathrm{GA}_1)_2),$$

$$\text{(iii)}\quad \operatorname*{ess\,sup}_{[0,\infty)} \||u| - 1\|_{L^2}^2 \leqslant C\varepsilon\mathcal{J}_\varepsilon(u_0^\varepsilon).$$

Proof. Assertion (i) follows from the basic energy law associated with the Cahn–Hilliard equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_\varepsilon(u(t)) = \begin{cases} -\|u_t(t)\|_{H^{-1}}^2, \\ -\|\nabla w(t)\|_{L^2}^2, \end{cases} \tag{2.7}$$

which is obtained from testing (1.1) with $-\Delta^{-1}u_t$, and integrating (2.7) in t from 0 to ∞. Here

$$\mathcal{J}_\varepsilon(u) := \int_\Omega \left[\frac{\varepsilon}{2}|\nabla u|^2 + \frac{1}{\varepsilon}F(u)\right]\mathrm{d}x \qquad \forall t \geqslant 0. \tag{2.8}$$

The conclusions of (ii) and (iii) follow from (i), the Mean Value Theorem ($F(u) - F(\pm 1) = (u \pm 1)F'(\xi^\pm)$), $(\mathrm{GA}_1)_1$ and $(\mathrm{GA}_1)_2$. □

The next lemma is a corollary of Theorems 2.1 and 2.3 of [2]. It shows the boundedness of the solution of the Cahn–Hilliard equation, provided that the Hele–Shaw problem (1.4)–(1.8) has a global (in time) classical solution. This boundedness result is the key for us to be able to establish improved a priori estimates for the solution of the Cahn–Hilliard equation. We remark that the estimates in [29] were obtained without assuming existence of a global (in time) classical solution for the Hele–Shaw problem, and hence, we were not able to show the boundedness of the solution of the Cahn–Hilliard equation there.

LEMMA 2.2 Suppose that f satisfies (GA_1), and the Hele–Shaw problem (1.4)–(1.8) has a global (in time) classical solution. Then there exists a family of smooth initial datum functions $\{u_0^\varepsilon\}_{0<\varepsilon\leqslant 1}$ and constants $\varepsilon_0 \in (0, 1]$ and $C_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the solution u of the Cahn–Hilliard equation (1.1)–(1.3) with the above initial data u_0^ε satisfies

$$\|u\|_{L^\infty(\Omega_T)} \leqslant \tfrac{3}{2}C_0. \tag{2.9}$$

Proof. A proof of the assertion is buried in the middle of the proof of Theorem 2.3 of [2]. In fact, the assertion of Theorem 2.3 of [2] was proved by establishing (2.9) first. Here we only sketch the main idea of the proof.

First, using a matched asymptotic expansion technique, a family of smooth approximate solutions $(u^\varepsilon_A, w^\varepsilon_A)$ to the solution (u, w) of (2.1)–(2.4) satisfying the assumption of Theorem 2.1 of [2] was constructed in Section 4 of [2]. One condition is $\|u^\varepsilon_A\|_{L^\infty(\Omega_T)} \leqslant C_0$ for some $C_0 > 0$. Second, it was proved in Theorem 2.1 of [2] that $(u^\varepsilon_A, w^\varepsilon_A)$ is very "close" to (u, w) in $L^p(\Omega_T)$ for some $p > 2$ (see (2.7) on p. 169 of [2]). Finally, (2.9) was proved using a regularization argument. The argument goes as follows in three steps: (i) f is modified into $\bar f$ such that $\bar f = f$ in $(-\frac{3}{2}C_0, \frac{3}{2}C_0)$ and $\bar f$ is linear for $|u| > 2C_0$; (ii) it was shown that the solution $\overline{u}$ of the Cahn–Hilliard equation with the new nonlinearity $\bar f$ satisfies the estimate (2.9) when $\varepsilon \in (0, \varepsilon_0)$ for some small $\varepsilon_0 \in (0, 1]$; (iii) it follows from the uniqueness of the solution of the Cahn–Hilliard equation that $u \equiv \overline{u}$. □

REMARK As in [2], the result of Lemma 2.2 is proved for a special family of initial data $\{u^\varepsilon_0(x)\}_{0<\varepsilon\leqslant 1}$. On the other hand, as explained in the introduction of [2], this is not a serious restriction on approximating the Hele–Shaw problem since (i) at the end of the first stage of the evolution of the concentration u has the required profile, and (ii) the solution of the Hele–Shaw problem (1.4)–(1.8) depends only on Γ_{00} and Ω.

The next lemma states a Poincaré–Friedrichs type inequality for any function w which has the form (2.2); it was proved in Lemma 3.4 of [16]. We note that u^ε in the lemma does not have to be the solution of the Cahn–Hilliard equation.

LEMMA 2.3 Suppose that u^ε satisfies

$$\frac{1}{|\Omega|}\int_\Omega u^\varepsilon(t)\,\mathrm{d}x = m_0 \in (-1, 1) \qquad \forall t \geqslant 0, \tag{2.10}$$

where m_0 is independent of ε. Let $\mathcal{J}_\varepsilon(u^\varepsilon)$ be defined by (2.8) and w^ε be defined by (2.2). Then there exist a (large) positive constant C and a (small) positive constant ε_0 such that for every $\varepsilon \in (0, \varepsilon_0]$,

$$\|w^\varepsilon(\cdot, t)\|_{L^2} \leqslant C(\mathcal{J}_\varepsilon(u^\varepsilon(\cdot, t)) + \|\nabla w^\varepsilon(\cdot, t)\|_{L^2}) \qquad \forall t \geqslant 0. \tag{2.11}$$

To derive a priori estimates in high norms we need to require that u^ε_0 satisfies the following conditions:

GENERAL ASSUMPTION 2 (GA_2) There exist positive ε-independent constants m_0 and σ_j for $j = 1, \dots, 4$ such that

$$1)\quad m_0 := \frac{1}{|\Omega|}\int_\Omega u^\varepsilon_0(x)\,\mathrm{d}x \in (-1, 1),$$

$$2)\quad \mathcal{J}_\varepsilon(u^\varepsilon_0) = \frac{\varepsilon}{2}\|\nabla u^\varepsilon_0\|^2_{L^2} + \frac{1}{\varepsilon}\|F(u^\varepsilon_0)\|_{L^1} \leqslant C\varepsilon^{-2\sigma_1},$$

$$3)\quad \|w^\varepsilon_0\|_{H^\ell} := \left\|\varepsilon\Delta u^\varepsilon_0 - \frac{1}{\varepsilon}f(u^\varepsilon_0)\right\|_{H^\ell} \leqslant C\varepsilon^{-\sigma_{2+\ell}}, \qquad \ell = 0, 1, 2.$$

LEMMA 2.4 Suppose f satisfies (GA_1), u^ε_0 satisfies (GA_2), and $\partial\Omega$ is of class $C^{3,1}$. Assume the solution u of (2.1)–(2.4) satisfies (2.9). Then (u, w) satisfies the following estimates:

$$\text{(i)}\quad \frac{1}{|\Omega|}\int_\Omega u(t)\,\mathrm{d}x = m_0 \in (-1, 1) \qquad \forall t \geqslant 0,$$

(ii) $\displaystyle\int_0^\infty \|\Delta u\|_{L^2}^2 \,\mathrm{d}s \leqslant C\varepsilon^{-(2\sigma_1+3)}$,

(iii) $\displaystyle\int_0^\infty \|\nabla\Delta u\|_{L^2}^2 \,\mathrm{d}s \leqslant C\varepsilon^{-(2\sigma_1+5)}$,

(iv) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)} \left\{ \begin{matrix} \|u_t\|_{H^{-1}}^2 \\ \|\nabla w\|_{L^2}^2 \end{matrix} \right\} + \varepsilon \int_0^\infty \|\nabla u_t\|_{L^2}^2 \,\mathrm{d}s \leqslant C\varepsilon^{-\max\{2\sigma_1+3.2\sigma_3\}}$,

(v) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)} \|\Delta u\|_{L^2} \leqslant C\varepsilon^{-\max\{\sigma_1+5/2,\sigma_3+1\}}$,

(vi) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)} \|\nabla\Delta u\|_{L^2} \leqslant C\varepsilon^{-\max\{\sigma_1+5/2,\sigma_3+1\}}$,

(vii) $\displaystyle\left\{ \begin{matrix} \int_0^\infty \|u_t\|_{L^2}^2 \,\mathrm{d}s \\ \int_0^\infty \|\Delta w\|_{L^2}^2 \,\mathrm{d}s \end{matrix} \right\} + \operatorname*{ess\,sup}_{[0,\infty)} \varepsilon\|\Delta u\|_{L^2}^2 \leqslant C\varepsilon^{-\max\{2\sigma_1+7/2,2\sigma_3+1/2,2\sigma_2+1\}}$,

(viii) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)}(\|u_t\|_{L^2}^2 + \|\Delta w\|_{L^2}^2) + \varepsilon \int_0^\infty \|\Delta u_t\|_{L^2}^2 \,\mathrm{d}s \leqslant C\varepsilon^{-\max\{2\sigma_1+13/2,2\sigma_3+7/2,2\sigma_2+4,2\sigma_4\}}$.

Moreover, in addition to (GA_2) suppose that there exists $\sigma_5 > 0$ such that

$$\lim_{s\searrow 0} \|\nabla u_t(s)\|_{L^2} \leqslant C\varepsilon^{-\sigma_5}. \tag{2.12}$$

Then the solution of (1.1)–(1.3) also satisfies the following estimates: for $N = 2, 3$,

(ix) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)} \|\nabla u_t\|_{L^2}^2 + \varepsilon \int_0^\infty \|\nabla\Delta u_t\|_{L^2}^2 \,\mathrm{d}s \leqslant C(\varepsilon^{-\max\{2\sigma_1+7,2\sigma_3+4\}}$

$\displaystyle\qquad + \varepsilon^{-\frac{2}{6-N}\max\{2\sigma_1+5,2\sigma_3+2\}-\max\{2\sigma_1+13/2,2\sigma_3+7/2,2\sigma_2+4\}} + \varepsilon^{-2\sigma_5}) \equiv C\rho_0(\varepsilon, N)$,

(x) $\displaystyle\int_0^\infty \|u_{tt}\|_{H^{-1}}^2 \,\mathrm{d}s \leqslant C\varepsilon\rho_0(\varepsilon, N) \equiv C\rho_1(\varepsilon, N)$,

(xi) $\displaystyle\operatorname*{ess\,sup}_{[0,\infty)} \|\Delta^2 u\|_{L^2} \leqslant C\varepsilon^{-\max\{\sigma_1+5,\sigma_3+7/2,\sigma_2+5/2,\sigma_4+1\}} \equiv C\rho_2(\varepsilon)$.

We refer the readers to [26] for the proof of the lemma.

REMARK From the construction of $(u_A^\varepsilon, w_A^\varepsilon)$ in Section 4 of [2] we know that $\{u_0^\varepsilon\}_{0<\varepsilon\leqslant 1}$ obtained in Lemma 2.2 satisfy (GA_2). In addition, the corresponding solution u satisfies (2.9) (see Lemma 2.2), provided that the Hele–Shaw problem (1.4)–(1.8) has a global (in time) classical solution.

We conclude this section by quoting the following result of [3, 15] on a lower bound estimate of the spectrum of the linearized Cahn–Hilliard operator

$$\mathcal{L}_{CH} := \Delta\left(\varepsilon\Delta - \frac{1}{\varepsilon} f'(u) I\right). \tag{2.13}$$

The estimate plays an important role in our error analysis to be given in Section 3.

LEMMA 2.5 Suppose the assumptions of Lemma 2.2 hold. Let $\gamma_1 = 1$ in $(\mathrm{GA}_1)_3$. Then there exist $0 < \varepsilon_0 \ll 1$ and another positive constant C_0 such that the principal eigenvalue of the linearized Cahn–Hilliard operator $\mathcal{L}_{\mathrm{CH}}$ satisfies

$$\lambda_{\mathrm{CH}} \equiv \inf_{0 \not\equiv \psi \in H^1(\Omega)} \frac{\varepsilon \|\nabla \psi\|_{L^2}^2 + \frac{1}{\varepsilon}(f'(u)\psi, \psi)}{\|\Delta^{-1/2}\psi\|_{L^2}^2} \geqslant -C_0, \tag{2.14}$$

or equivalently

$$\lambda_{\mathrm{CH}} \equiv \inf_{\substack{0 \not\equiv \psi \in H^1(\Omega) \\ \Delta w = \psi}} \frac{\varepsilon \|\nabla \psi\|_{L^2}^2 + \frac{1}{\varepsilon}(f'(u)\psi, \psi)}{\|\nabla w\|_{L^2}^2} \geqslant -C_0, \tag{2.15}$$

for all $\varepsilon \in (0, \varepsilon_0)$. Here u denotes the solution of the Cahn–Hilliard problem (1.1)–(1.3).

Proof. The estimate (2.14) was proved by Alikakos and Fusco [3] in the two-dimensional case and by X. Chen [15] for all dimensions, provided that the function u (which does not have to be the solution to the Cahn–Hilliard equation) has some special profile (cf. p. 638 of [3] and p. 1374 of [15]). It was shown in [2] that the solution to the Cahn–Hilliard problem (1.1)–(1.3) indeed has the required profile (cf. Theorems 4.12 and 2.1 of [2]) for sufficiently small ε. The conclusion of the lemma then follows from combining these two results. □

3. Error analysis for a fully discrete mixed finite element approximation

In this section we analyze the fully discrete mixed finite element method proposed in [29] for (2.1)–(2.4) under the condition that the Hele–Shaw problem has a global (in time) classical solution (cf. [2]). Under this assumption, we establish stronger error bounds than those of [29], which were shown under general (minimum) regularity assumptions, for the solution of the fully discrete mixed finite element method. In particular, we obtain an $L^\infty(J; L^\infty)$ error estimate, which is necessary for us to establish the convergence of the solution of the fully discrete mixed finite element scheme to the solution of the Hele–Shaw problem in the next section.

We recall that the weak formulation of (2.1)–(2.4) is defined as: Find $(u(t), w(t)) \in [H^1(\Omega)]^2$ such that for almost every $t \in (0, T)$,

$$(u_t, \eta) + (\nabla w, \nabla \eta) = 0 \qquad \forall \eta \in H^1(\Omega), \tag{3.1}$$

$$\varepsilon(\nabla u, \nabla v) + \frac{1}{\varepsilon}(f(u), v) = (w, v) \qquad \forall v \in H^1(\Omega), \tag{3.2}$$

$$u(x, 0) = u_0^\varepsilon(x) \qquad \forall x \in \Omega. \tag{3.3}$$

Note that $(u_t, 1) = 0$, that is, the mass $(u(t), 1) = (u_0^\varepsilon, 1)$ is conserved for all $t \geqslant 0$.

We also recall that the fully discrete mixed finite element discretization of (3.1)–(3.3) is defined as: Find $\{(U^m, W^m)\}_{m=1}^M \in [\mathcal{S}_h]^2$ such that

$$(d_t U^{m+1}, \eta_h) + (\nabla W^{m+1}, \nabla \eta_h) = 0 \qquad \forall \eta_h \in \mathcal{S}_h, \tag{3.4}$$

$$\varepsilon(\nabla U^{m+1}, \nabla v_h) + \frac{1}{\varepsilon}(f(U^{m+1}), v_h) = (W^{m+1}, v_h) \qquad \forall v_h \in \mathcal{S}_h, \tag{3.5}$$

with some suitable starting value U^0. Unless stated otherwise, we define W^0 by setting $m = -1$ in (3.5). Here $J_k := \{t_m\}_{m=0}^M$ is a uniform partition of $[0, T]$ of mesh size $k := T/M$ and $\mathcal{T}_h$ is a quasi-uniform "triangulation" of Ω. Also, $d_t U^{m+1} := (U^{m+1} - U^m)/k$ and $\mathcal{S}_h$ denotes the P_1 conforming finite element space defined by

$$\mathcal{S}_h := \{v_h \in C(\overline{\Omega});\, v_h|_K \in P_1(K),\ \forall K \in \mathcal{T}_h\}.$$

The mixed finite element space $\mathcal{S}_h \times \mathcal{S}_h$ is the lowest order element among a family of stable mixed finite spaces known as the Ciarlet–Raviart mixed finite elements for the biharmonic problem (cf. [17, 33]), which means that the following inf-sup condition holds:

$$\inf_{0\not\equiv\eta_h\in\mathcal{S}_h}\ \sup_{0\not\equiv\psi_h\in\mathcal{S}_h} \frac{(\nabla\psi_h, \nabla\eta_h)}{\|\psi_h\|_{H^1}\|\eta_h\|_{H^1}} \geqslant c_0 \tag{3.6}$$

for some $c_0 > 0$.

Also, we note that $(d_t U^{m+1}, 1) = 0$, which implies that $(U^{m+1}, 1) = (U^0, 1)$ for $m = 0, 1, \ldots, M-1$. Hence, the mass is also conserved by the fully discrete solution at each time step.

We define the $L^2(\Omega)$ projection $Q_h : L^2(\Omega) \to \mathcal{S}_h$ by

$$(Q_h v - v, \eta_h) = 0 \quad \forall\eta \in \mathcal{S}_h, \tag{3.7}$$

and the elliptic projection $P_h : H^1(\Omega) \to \mathcal{S}_h$ by

$$(\nabla[P_h v - v], \nabla\eta_h) = 0 \quad \forall\eta_h \in \mathcal{S}_h, \tag{3.8}$$

$$(P_h v - v, 1) = 0. \tag{3.9}$$

We refer to Section 4 of [28] for a list of approximation properties of Q_h and P_h. In what follows, we confine ourselves to meshes $\mathcal{T}_h$ that result in H^1 stability of Q_h (see [13] and reference therein for the details).

We also introduce the space notations

$$\mathring{\mathcal{S}}_h := \{v_h \in \mathcal{S}_h;\, (v_h, 1) = 0\}, \qquad L^2_0(\Omega) := \{v \in L^2(\Omega);\, (v, 1) = 0\},$$

and define the discrete inverse Laplace operator $-\Delta_h^{-1} : L^2_0(\Omega) \to \mathring{\mathcal{S}}_h$ such that

$$(\nabla(-\Delta_h^{-1} v), \nabla\eta_h) = (v, \eta_h) \quad \forall\eta_h \in \mathcal{S}_h. \tag{3.10}$$

To establish stability estimates for the solution of the fully discrete scheme (3.4)–(3.5) for general potential functions $F(u)$, we make the last structural assumption on $f(u)$.

GENERAL ASSUMPTION 3 (GA$_3$) *There exist $\alpha_0 \geqslant 0$, $0 < \gamma_3 < 1$, and $\tilde{c}_4 > 0$ such that f satisfies for any $0 < k \leqslant \varepsilon^{\alpha_0}$, any set of discrete (in time) functions $\{\phi^m\}_{m=0}^M \in H^1(\Omega)$, and all $\ell \leqslant M$,*

$$\gamma_3 k \sum_{m=1}^{\ell} (\|d_t\phi^m\|^2_{H^{-1}} + k\varepsilon\|\nabla d_t\phi^m\|^2_{L^2}) + \frac{k}{\varepsilon}\sum_{m=1}^{\ell}(f(\phi^m), d_t\phi^m) + \tilde{c}_4\mathcal{J}_\varepsilon(\phi^0) \geqslant \frac{\tilde{c}_4}{\varepsilon}\|F(\phi^\ell)\|_{L^1}. \tag{3.11}$$

We remark that the validity of (GA$_3$) was proved in [29, 28] for the case of the quartic potential $F(u) = \frac{1}{4}(u^2-1)^4$ with $\alpha_0 = 3$, $\gamma_3 = 1/4$ and $\tilde{c}_4 = 2$. With the help of (GA$_3$) we are able to show that the solution of (3.4)–(3.5) satisfies the following stability estimates.

LEMMA 3.1 The solution $\{(U^m, W^m)\}_{m=1}^M$ of (3.4)–(3.5) satisfies

(i) $\frac{1}{|\Omega|}\int_\Omega U^m\,\mathrm{d}x = \frac{1}{|\Omega|}\int_\Omega U^0\,\mathrm{d}x, \quad m = 1, \dots, M,$

(ii) $\|d_t U^m\|_{H^{-1}} \leqslant C\|\nabla W^m\|_{L^2}, \quad m = 1, \dots, M,$

(iii) $\max_{0\leqslant m\leqslant M}\left\{\varepsilon\|\nabla U^m\|_{L^2}^2 + \frac{1}{\varepsilon}\|F(U^m)\|_{L^1}\right\} + k\sum_{m=1}^M(\|\nabla W^m\|_{L^2}^2 + \varepsilon k\|\nabla d_t U^m\|_{L^2}^2) \leqslant C\mathcal{J}_\varepsilon(U^0),$

(iv) $k\sum_{m=1}^M \|d_t U^m\|_{H^{-1}}^2 \leqslant C\mathcal{J}_\varepsilon(U^0),$

(v) $\max_{0\leqslant m\leqslant M}\|U^m\|_{L^p}^p \leqslant C(1 + \mathcal{J}_\varepsilon(U^0)) \quad$ (p as in $(\mathrm{GA}_1)_2$),

(vi) $\max_{1\leqslant m\leqslant M}\|\nabla W^m\|_{L^2}^2 + k\sum_{m=1}^M(k\|d_t\nabla W^m\|_{L^2}^2 + \varepsilon\|\nabla d_t U^m\|_{L^2}^2) \leqslant C\varepsilon^{-3}\mathcal{J}_\varepsilon(U^0),$

(vii) $\max_{1\leqslant m\leqslant M}\|d_t U^m\|_{L^2}^2 + k\sum_{m=1}^M k\|d_t^2 U^m\|_{L^2}^2 \leqslant C[(k\varepsilon)^{-\frac{1}{2}} + h^{-2}]\varepsilon^{-3}\mathcal{J}_\varepsilon(U^0).$

Proof. Assertion (i) is an immediate consequence of setting $\eta_h = 1$ in (3.4).

For any $\phi \in H^1(\Omega)$, from (3.4), (3.7), and the stability of Q_h in $H^1(\Omega)$ (cf. [13] and references therein) we have

$$\begin{aligned}(d_t U^m, \phi) &= (d_t U^m, Q_h\phi) + (d_t U^m, \phi - Q_h\phi)\\ &= -(\nabla W^m, \nabla Q_h\phi) \leqslant C\|\nabla W^m\|_{L^2}\|\nabla\phi\|_{L^2}.\end{aligned} \tag{3.12}$$

Assertion (ii) then follows from

$$\|d_t U^m\|_{H^{-1}} = \sup_{0\not\equiv\phi\in H^1}\frac{(d_t U^m, \phi)}{\|\phi\|_{H^1}} \leqslant C\|\nabla W^m\|_{L^2}.$$

To show assertion (iii), setting $\eta_h = W^{m+1}$ in (3.4) and $v_h = d_t U^{m+1}$ in (3.5) and adding the resulting equations gives

$$\|\nabla W^{m+1}\|_{L^2}^2 + \frac{\varepsilon}{2}d_t\|\nabla U^{m+1}\|_{L^2}^2 + \frac{\varepsilon k}{2}\|d_t\nabla U^{m+1}\|_{L^2}^2 + \frac{1}{\varepsilon}(f(U^{m+1}), d_t U^{m+1}) = 0. \tag{3.13}$$

The statement then follows from (GA_3) and (ii) after multiplying (3.13) by k and summing over m from 0 to ℓ $(\leqslant M-1)$.

Assertions (iv) and (v) follow immediately from (ii), (iii), and the general assumption (GA_1) on F and f. To show (vi), choose $\eta_h = d_t W^{m+1}$ in (3.4) and $v_h = d_t U^{m+1}$ in (3.5) after applying the difference operator d_t to (3.5), and add the resulting equations to get

$$\frac{1}{2}d_t\|\nabla W^{m+1}\|_{L^2}^2 + \frac{k}{2}\|d_t\nabla W^{m+1}\|_{L^2}^2 + \varepsilon\|\nabla d_t U^{m+1}\|_{L^2}^2 + \frac{1}{\varepsilon}(d_t f(U^{m+1}), d_t U^{m+1}) = 0. \tag{3.14}$$

By the Mean Value Theorem and (2.5) we bound the last term on the left hand side of (3.14) as follows:

$$\frac{1}{\varepsilon}(d_t f(U^{m+1}), d_t U^{m+1}) = \frac{1}{\varepsilon}(f'(\xi), |d_t U^{m+1}|^2) \geqslant -\frac{\widetilde{c}_0}{\varepsilon}\|d_t U^{m+1}\|_{L^2}^2$$
$$\geqslant -\frac{\varepsilon}{2}\|\nabla d_t U^{m+1}\|_{L^2}^2 - \frac{C}{\varepsilon^3}\|d_t U^{m+1}\|_{H^{-1}}^2. \quad (3.15)$$

The assertion then follows from (3.14)–(3.15) and (iv) after multiplying (3.14) by k and summing over m from 0 to ℓ ($\leqslant M-1$).

Finally, to show (vii), we first apply the difference operator d_t to both sides of (3.4), then take $\eta_h = d_t U^{m+1}$ in the resulting equation to have

$$\frac{1}{2}d_t\|d_t U^{m+1}\|_{L^2}^2 + \frac{k}{2}\|d_t^2 U^{m+1}\|_{L^2}^2 = -(\nabla d_t W^{m+1}, \nabla d_t U^{m+1})$$
$$\leqslant \frac{1}{\sqrt{k\varepsilon}}(k\|\nabla d_t W^{m+1}\|_{L^2}^2 + \varepsilon\|\nabla d_t U^{m+1}\|_{L^2}^2). \quad (3.16)$$

Multiplying (3.16) by k and summing over m from 0 to ℓ ($\leqslant M-1$), the assertion then follows from (ii), (vi), and the inverse inequality (bounding the L^2 norm by the H^{-1} norm) to bound $\|d_t U^1\|_{L^2}^2$ on the right hand side. □

REMARK In view of Lemmas 2.4(i) and 3.1(i), in order for the scheme (3.4)–(3.5) to conserve the mass of the underlying physical problem, it is necessary to require $(U^0 - u_0^\varepsilon, 1) = 0$ for the starting value U^0. This condition will be assumed in the rest of this section.

As is shown in [29], in order to establish error bounds that depend on low order polynomials of $1/\varepsilon$, the crucial idea is to utilize the spectrum estimate result of Lemma 2.5 for the linearized Cahn–Hilliard operator. In the following we show that the spectrum estimate still holds if the function u, which is the solution of (1.1)–(1.3), is replaced by its elliptic projection $P_h u$ and the nonlinear term is scaled by a factor $1-\varepsilon$, provided that the mesh size h is small enough. As expected, this result plays a critical role in our error analysis for the fully discrete finite element discretization.

For u the solution of (1.1)–(1.3), let C_0 be as in (2.9), define

$$C_1 = \max_{|v|\leqslant 2C_0} |f''(v)|, \quad (3.17)$$

and let C_2 be the smallest positive ε-independent constant such that (cf. Chapter 7 of [11])

$$\|u - P_h u\|_{L^\infty(J;L^\infty)} \leqslant C_2 h^2|\ln h|\|u\|_{L^\infty(J;W^{2,\infty})} \leqslant C_2 h^2 |\ln h|\rho_3(\varepsilon, N) \quad (3.18)$$

for some (low order) polynomial function $\rho_3(\varepsilon, N)$ in $1/\varepsilon$. We remark that the existence of C_2 and $\rho_3(\varepsilon, N)$ follows easily from Lemma 2.4(xi) and the following Gagliardo–Nirenberg inequality [1]:

$$\|u\|_{W^{2,\infty}} \leqslant C(\|D^4 u\|_{L^2}^{4/(8-N)}\|u\|_{L^\infty}^{(4-N)/(8-N)} + \|u\|_{L^\infty}), \qquad N = 2, 3.$$

In fact, the above inequality, (2.9) and Lemma 2.4(xi) imply

$$\rho_3(\varepsilon, N) \leqslant \rho_2(\varepsilon)^{4/(8-N)}, \quad (3.19)$$

where $\rho_2(\varepsilon)$ is defined in Lemma 2.4(xi).

LEMMA 3.2 Let the assumptions of Lemma 2.5 hold, and ε_0 and C_0 be as there. Then for $\varepsilon \in (0, \varepsilon_0)$,

$$\lambda_{\mathrm{CH}}^h \equiv \inf_{\substack{0 \not\equiv \psi \in L_0^2(\Omega) \\ \Delta w = \psi,\, \partial w/\partial n = 0}} \frac{\varepsilon \|\nabla \psi\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon}(f'(P_h u)\psi, \psi)}{\|\nabla w\|_{L^2}^2} \geqslant -(1-\varepsilon^3)(C_0+1), \tag{3.20}$$

provided that h satisfies the constraint

$$h^2 |\ln h| \leqslant (C_1 C_2 \rho_3(\varepsilon, N))^{-1} \varepsilon^3. \tag{3.21}$$

Proof. From the definitions of C_1 and C_2, we immediately have

$$\|P_h u\|_{L^\infty(J;L^\infty)} \leqslant \|u\|_{L^\infty(J;L^\infty)} + \|P_h u - u\|_{L^\infty(J;L^\infty)} \leqslant \tfrac{4}{3}\|u\|_{L^\infty(J;L^\infty)} \leqslant 2C_0$$

if h satisfies (3.21). It then follows from the Mean Value Theorem that

$$\begin{aligned} \|f'(P_h u) - f'(u)\|_{L^\infty(J;L^\infty)} &\leqslant \max_{|\xi| \leqslant 2C_0} |f''(\xi)| \|P_h u - u\|_{L^\infty(J;L^\infty)} \\ &\leqslant C_1 C_2 h^2 |\ln h| \rho_3(\varepsilon, N) \leqslant \varepsilon^3. \end{aligned} \tag{3.22}$$

Using the inequality $a \geqslant b - |a-b|$ and (3.22) we get

$$f'(P_h u) \geqslant f'(u) - \|f'(P_h u) - f'(u)\|_{L^\infty(J;L^\infty)} \geqslant f'(u) - \varepsilon^3. \tag{3.23}$$

In addition, for any ψ as in (3.20) we have

$$\|\psi\|_{L^2}^2 = (\nabla \psi, \nabla w) \leqslant \frac{\varepsilon^2}{1-\varepsilon^3} \|\nabla \psi\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon^2} \|\nabla w\|_{L^2}^2. \tag{3.24}$$

Substituting (3.23)–(3.24) into the definition of λ_{CH}^h we get

$$\lambda_{\mathrm{CH}}^h \geqslant \inf_{\substack{0 \not\equiv \psi \in L_0^2(\Omega) \\ \Delta w = \psi,\, \partial w/\partial n = 0}} \frac{(1-\varepsilon^3)[\varepsilon \|\nabla \psi\|_{L^2}^2 + \frac{1}{\varepsilon}(f'(u)\psi, \psi)]}{\|\nabla w\|_{L^2}^2} - (1-\varepsilon^3)^2.$$

The proof is completed by applying Lemma 2.5. □

REMARK Under a slightly weaker mesh constraint than (3.21), a slightly weaker version of (3.20) was shown in Proposition 3.2 of [26].

The first main result of this section is stated in the following theorem.

THEOREM 3.1 Let $\{(U^m, W^m)\}_{m=1}^M$ solve (3.4)–(3.5) on a quasi-uniform space mesh $\mathcal{T}_h$ of size $O(h)$ and a quasi-uniform time mesh J_k of size $O(k)$. Suppose the assumptions of Lemma 2.4 and 3.1, Lemma 2.5 and 3.2 hold, in particular, α_0, σ_i, $\rho_i(\varepsilon, N)$ are as there. Let $0 < \delta < 16/(8-N)$ for $N = 2, 3$, and define, for any $\nu > 0$,

$$\mu := \mu(N, \delta, \nu) = \min\{\delta, \nu, 8 - N/8\}, \qquad \rho_4(\varepsilon) := \varepsilon^{\min\{2\sigma_1+21/2, 2\sigma_3+15/2, 2\sigma_2+8, 2\sigma_4+4\}}, \tag{3.25}$$
$$\rho_5(\varepsilon) := \varepsilon^{-\max\{2\sigma_1+9, 2\sigma_3+6, 2\sigma_2+4, 2\sigma_4+1\}},$$

$$\pi_1(k; \varepsilon, N, \delta, \sigma_i) := \rho_1(\varepsilon, N) + k^{\frac{16+(8-N)\delta}{16-(8-N)\delta}} \varepsilon^{-\frac{32(\sigma_1+3)+2\delta(8-N)(2\sigma_1-1)}{16-(8-N)\delta}} \rho_2(\varepsilon)^{\frac{4N\delta}{16-(8-N)\delta}}, \tag{3.26}$$

$$\begin{aligned} \pi_2(k, h; \varepsilon, N, \delta, \sigma_i, \nu) := {} & [h^{\frac{(8-N)\delta}{4} - 2\mu} \varepsilon^{-\frac{(2\sigma_1+1)[16+(8-N)\delta]}{16}} + h^{-2\mu} k^2 \rho_5(\varepsilon)] \rho_2(\varepsilon)^{\frac{\delta N}{8}} \\ & + [h^{2(\delta-\mu)}(\rho_2(\varepsilon)^{\frac{4(2+\delta)-2N}{8-N}} + \rho_4(\varepsilon)) + h^{2(\nu-\mu)} \varepsilon^{-2(\sigma_2+1)}], \end{aligned} \tag{3.27}$$

$$r(h, k; \varepsilon, N, \delta, \sigma_i, \nu) := k^2 \pi_1(k; \varepsilon, N, \delta, \sigma_i) + h^{2(2+\mu)} \pi_2(k, h; \varepsilon, N, \delta, \sigma_i, \nu). \tag{3.28}$$

Then, under the following mesh and starting value constraints:

1) $k \leqslant \varepsilon^{\alpha_0}$,
2) $h^2 |\ln h| \leqslant \varepsilon^3 \rho_2(\varepsilon)^{-\frac{4}{8-N}}$,
3) $r(h, k; \varepsilon, N, \delta, \sigma_i, \nu) \leqslant \varepsilon^{[2+\frac{48}{(8-N)\delta}]} \rho_2(\varepsilon)^{-\frac{2N}{8-N}}$,
4) $(U^0, 1) = (u_0^\varepsilon, 1)$,
5) $\|u_0^\varepsilon - U^0\|_{H^{-1}} \leqslant C h^{2+\nu} \|u_0^\varepsilon\|_{H^2}$,

the solution of (3.4)–(3.5) converges to the solution of (2.1)–(2.4) and satisfies

$$\text{(i)} \quad \max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{H^{-1}} + \Big(k \sum_{m=1}^{M} k \|d_t(u(t_m) - U^m)\|_{H^{-1}}^2\Big)^{1/2} \leqslant \tilde{C}[r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)]^{1/2},$$

$$\text{(ii)} \quad \Big(k \sum_{m=0}^{M} \|u(t_m) - U^m\|_{L^2}^2\Big)^{1/2} \leqslant \tilde{C}\{h^2 \varepsilon^{-(\sigma_1+1/2)} + \varepsilon^{-2}[r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)]^{1/2}\},$$

$$\text{(iii)} \quad \Big(k \sum_{m=0}^{M} \|\nabla(u(t_m) - U^m)\|_{L^2}^2\Big)^{1/2} \leqslant \tilde{C}\{h \varepsilon^{-(\sigma_1+1/2)} + \varepsilon^{-2}[r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)]^{1/2}\}$$

for some positive constant $\tilde{C} = \tilde{C}(u_0^\varepsilon; \gamma_2, C_0, T; \Omega)$. Here γ_2 and δ are defined in $(\mathrm{GA}_1)_3$.

Proof. The proof is divided into four steps. Step one deals with the consistency error due to the time discretization. Steps two and three use Lemma 3.2 and the stability estimates of Lemmas 2.4 and 3.1 to avoid an exponential blow-up in $1/\varepsilon$ of the error constants. In the final step, an inductive argument is used to handle the difficulty caused by the super-quadratic term in $(\mathrm{GA}_1)_3$.

STEP 1 Let $E^m := u(t_m) - U^m$ and $G^m := w(t_m) - W^m$. Subtracting (3.4)–(3.5) (after replacing $m+1$ by m) from (3.1)–(3.2) (after setting $t = t_m$), respectively, we get the following error equations at t_m:

$$(d_t E^m, \eta_h) + (\nabla G^m, \nabla \eta_h) = (\mathcal{R}(u_{tt}; m), \eta_h), \tag{3.29}$$

$$\varepsilon(\nabla E^m, \nabla v_h) + \frac{1}{\varepsilon}(f(u(t_m)) - f(U^m), v_h) = (G^m, v_h), \tag{3.30}$$

where

$$\mathcal{R}(u_{tt}; m) = -\frac{1}{k} \int_{t_{m-1}}^{t_m} (s - t_{m-1}) u_{tt}(s)\, \mathrm{d}s. \tag{3.31}$$

It is easy to check that

$$k \sum_{m=1}^{M} \|\mathcal{R}(u_{tt}; m)\|_{H^{-1}}^2 \leqslant \frac{1}{k} \sum_{m=1}^{M} \left[\int_{t_{m-1}}^{t_m} (s - t_{m-1})^2\, \mathrm{d}s\right] \left[\int_{t_{m-1}}^{t_m} \|u_{tt}\|_{H^{-1}}^2\, \mathrm{d}s\right] \leqslant C k^2 \rho_1(\varepsilon, N), \tag{3.32}$$

where $\rho_1(\varepsilon, N)$ is defined in Lemma 2.4(x).

STEP 2 Introduce the decompositions $E^m := \Theta^m + \Phi^m$ and $G^m := \Lambda^m + \Psi^m$, where

$$\Theta^m := u(t_m) - P_h u(t_m), \qquad \Phi^m := P_h u(t_m) - U^m,$$
$$\Lambda^m := w(t_m) - P_h w(t_m), \qquad \Psi^m := P_h w(t_m) - W^m.$$

Then from the definition of P_h in (3.8)–(3.9) we can rewrite (3.29)–(3.30) as follows:

$$(d_t\Phi^m, \eta_h) + (\nabla\Psi^m, \nabla\eta_h) = (\mathcal{R}(u_{tt}; m), \eta_h) - (d_t\Theta^m, \eta_h), \tag{3.33}$$

$$\begin{aligned}\varepsilon(\nabla\Phi^m, \nabla v_h) &+ \frac{1}{\varepsilon}(f(P_h u(t_m)) - f(U^m), v_h)\\ &= (\Psi^m, v_h) + (\Lambda^m, v_h) - \frac{1}{\varepsilon}(f(u(t_m)) - f(P_h u(t_m)), v_h).\end{aligned} \tag{3.34}$$

Since $E^m, \Phi^m \in L_0^2(\Omega)$ for $0 \leqslant m \leqslant M$, setting $\eta_h = -\Delta_h^{-1}\Phi^m$ in (3.33) and $v_h = \Phi^m$ in (3.34) and summing over m from 1 to ℓ ($\leqslant M$), after adding the equations we conclude

$$\begin{aligned}\frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^\ell\|_{L^2}^2 &+ k\sum_{m=1}^{\ell}\frac{k}{2}\|\nabla\Delta_h^{-1}d_t\Phi^m\|_{L^2}^2\\ &+ k\sum_{m=1}^{\ell}[\varepsilon\|\nabla\Phi^m\|_{L^2}^2 + \frac{1}{\varepsilon}(f(P_h u(t_m)) - f(U^m), \Phi^m)]\\ &= k\sum_{m=1}^{\ell}[(\mathcal{R}(u_{tt}; m), -\Delta_h^{-1}\Phi^m) - (d_t\Theta^m, -\Delta_h^{-1}\Phi^m) + (\Lambda^m, \Phi^m)]\\ &+ \frac{k}{\varepsilon}\sum_{m=1}^{\ell}(f(u(t_m)) - f(P_h u(t_m)), \Phi^m) + \frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^0\|_{L^2}^2.\end{aligned} \tag{3.35}$$

The first sum on the right hand side can be bounded as follows:

$$\begin{aligned}k\sum_{m=1}^{\ell}&[(\mathcal{R}(u_{tt}; m), -\Delta_h^{-1}\Phi^m) - (d_t\Theta^m, -\Delta_h^{-1}\Phi^m) + (\Lambda^m, \Phi^m)]\\ &\leqslant Ck\sum_{m=1}^{\ell}[\|\mathcal{R}(u_{tt}; m)\|_{H^{-1}}^2 + \|d_t\Theta^m\|_{H^{-1}}^2 + \varepsilon^{-4}\|\Lambda^m\|_{H^{-1}}^2]\\ &\quad + k\sum_{m=1}^{\ell}\{\|\nabla\Delta_h^{-1}\Phi^m\|_{L^2}^2 + \frac{\varepsilon^4}{2(1-\varepsilon^3)}\|\nabla\Phi^m\|_{L^2}^2\}\\ &\leqslant k\sum_{m=1}^{\ell}\{\|\nabla\Delta_h^{-1}\Phi^m\|_{L^2}^2 + \frac{\varepsilon^4}{2(1-\varepsilon^3)}\|\nabla\Phi^m\|_{L^2}^2\} + C[k^2\rho_1(\varepsilon, N) + h^6\rho_4(\varepsilon)],\end{aligned} \tag{3.36}$$

where $\rho_1(\varepsilon, N)$ is defined in Lemma 2.4(x), and $\rho_4(\varepsilon, N)$, which is defined in (3.25), comes from Lemma 2.4(viii). Here we have used the following approximation properties in H^{-1} of the elliptic projection P_h (cf. [20]):

$$\|u - P_h u\|_{H^{-1}} \leqslant Ch^3\|u\|_{H^2},$$
$$\|(u - P_h u)_t\|_{H^{-1}} \leqslant Ch^3\|u_t\|_{H^2},$$
$$\|w - P_h w\|_{H^{-1}} \leqslant Ch^3\|w\|_{H^2},$$

and $k\sum_{m=1}^{\ell}\|d_t u(t_m)\|_{H^2}^2 \leqslant \int_0^{t_\ell}\|u_t(s)\|_{H^2}^2\,\mathrm{d}s$, which follows from $kd_t u(t_m) = \int_{t_{m-1}}^{t_m} u_t(s)\,\mathrm{d}s$.

In view of (2.9) and the inequality at the beginning of the proof of Lemma 3.2, the second sum on the right hand side of (3.35) can be bounded by

$$\begin{aligned}\frac{k}{\varepsilon}\sum_{m=1}^{\ell}(f(u(t_m))-f(P_hu(t_m)),\Phi^m) &= \frac{k}{\varepsilon}\sum_{m=1}^{\ell}(f'(\xi)\Theta^m,\Phi^m)\\ &\leqslant k\sum_{m=1}^{\ell}\left[\frac{\varepsilon^4}{2(1-\varepsilon^3)}\|\nabla\Phi^m\|_{L^2}^2 + C\varepsilon^{-6}\|\Theta^m\|_{H^{-1}}^2\right]\\ &\leqslant k\sum_{m=1}^{\ell}\frac{\varepsilon^4}{2(1-\varepsilon^3)}\|\nabla\Phi^m\|_{L^2}^2\\ &\quad + Ch^6\varepsilon^{-\max\{2\sigma_1+11,2\sigma_3+8\}}. \end{aligned} \tag{3.37}$$

By $(GA_1)_3$ with $\gamma_1 = 1$, the last term on the left hand side of (3.35) is bounded from below by

$$k\sum_{m=1}^{\ell}\frac{1}{\varepsilon}(f(P_hu(t_m))-f(U^m),\Phi^m) \geqslant \frac{k}{\varepsilon}\sum_{m=1}^{\ell}[(f'(P_hu(t_m))\Phi^m,\Phi^m)-\gamma_2\|\Phi^m\|_{L^{2+\delta}}^{2+\delta}]. \tag{3.38}$$

Substituting (3.36)–(3.38) into (3.35) we arrive at

$$\begin{aligned}&\frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^\ell\|_{L^2}^2 + k\sum_{m=1}^{\ell}\frac{k}{2}\|\nabla\Delta_h^{-1}d_t\Phi^m\|_{L^2}^2 + k\sum_{m=1}^{\ell}2\varepsilon^2(f'(P_hu(t_m))\Phi^m,\Phi^m)\\ &\qquad + \frac{1-2\varepsilon^3}{1-\varepsilon^3}k\sum_{m=1}^{\ell}\left[\varepsilon\|\nabla\Phi^m\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon}(f'(P_hu(t_m))\Phi^m,\Phi^m)\right]\\ &\quad\leqslant C[k^2\rho_1(\varepsilon,N)+h^6\rho_4(\varepsilon)] + \frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^0\|_{L^2}^2\\ &\qquad + k\sum_{m=1}^{\ell}\|\nabla\Delta_h^{-1}\Phi^m\|_{L^2}^2 + \frac{\gamma_2 k}{\varepsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^{2+\delta}}^{2+\delta}. \end{aligned} \tag{3.39}$$

We could bound the last term on the left hand side from below using Lemma 3.2, however, this will consume all the contribution of $\varepsilon\|\nabla\Phi^m\|_{L^2}^2$ on the left hand side. On the other hand, in order to bound the super-quadratic term on the right hand side in Step 3 below, we do need a small help from this $\varepsilon\|\nabla\Phi^m\|_{L^2}^2$. For that reason, we are going to apply Lemma 3.2 with a scaling factor $1-\varepsilon^3$, that is, we first write

$$\begin{aligned}&\varepsilon\|\nabla\Phi^m\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon}(f'(P_hu(t_m))\Phi^m,\Phi^m)\\ &\qquad = (1-\varepsilon^3)\left[\varepsilon\|\nabla\Phi^m\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon}(f'(P_hu(t_m))\Phi^m,\Phi^m)\right]\\ &\qquad\quad + \varepsilon^3\left[\varepsilon\|\nabla\Phi^m\|_{L^2}^2 + \frac{1-\varepsilon^3}{\varepsilon}(f'(P_hu(t_m))\Phi^m,\Phi^m)\right]. \end{aligned}$$

From Lemma 3.2 we then bound the first term on the right hand side of the above equation as

$$(1-\varepsilon^3)\left[\varepsilon\|\nabla\Phi^m\|_{L^2}^2+\frac{1-\varepsilon^3}{\varepsilon}(f'(P_hu(t_m))\Phi^m,\Phi^m)\right]$$
$$\geqslant -(1-\varepsilon^3)^2(C_0+1)\|\nabla\Delta^{-1}\Phi^m\|_{L^2}^2\geqslant -(C_0+1)\|\nabla\Delta^{-1}\Phi^m\|_{L^2}^2. \quad (3.40)$$

We then keep $\varepsilon^4\|\nabla\Phi^m\|_{L^2}^2$ on the left hand side, and move the leftover term $\varepsilon^2(1-\varepsilon^3)\times(f'(P_hu(t_m))\Phi^m,\Phi^m)$ to the right side to bound it from above by

$$\varepsilon^2(1-\varepsilon^3)(f'(P_hu(t_m))\Phi^m,\Phi^m)\leqslant\varepsilon^2|(f'(P_hu(t_m))\Phi^m,\Phi^m)|$$
$$\leqslant\left[\frac{\varepsilon^4}{4}\|\nabla\Phi^m\|_{L^2}^2+\tilde{c}_0\|\nabla\Delta_h^{-1}\Phi^m\|_{L^2}^2\right]. \quad (3.41)$$

Combining (3.39)–(3.41) we finally get, for sufficiently small $\varepsilon>0$,

$$\frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^\ell\|_{L^2}^2+k\sum_{m=1}^{\ell}\left[\frac{k}{2}\|\nabla\Delta_h^{-1}d_t\Phi^m\|_{L^2}^2+\frac{\varepsilon^4}{4}\|\nabla\Phi^m\|_{L^2}^2\right]$$
$$\leqslant C[k^2\rho_1(\varepsilon,N)+h^6\rho_4(\varepsilon)]+\frac{1}{2}\|\nabla\Delta_h^{-1}\Phi^0\|_{L^2}^2$$
$$+(C_0+3\tilde{c}_0+3)k\sum_{m=1}^{\ell}\|\nabla\Delta_h^{-1}\Phi^m\|_{L^2}^2+\frac{\gamma_2k}{\varepsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^{2+\delta}}^{2+\delta}, \quad (3.42)$$

where we have used the fact that $0<\varepsilon<1$ and $\|\nabla\Delta^{-1}v_h\|_{L^2}=\|\nabla\Delta_h^{-1}v_h\|_{L^2}$ for any $v_h\in\mathring{\mathcal{S}}_h$.

STEP 3 It remains to bound the super-quadratic term at the end of (3.42). Since $\Phi^m=E^m-\Theta^m$, the triangle inequality implies

$$\|\Phi^m\|_{L^{2+\delta}}^{2+\delta}\leqslant C(\|E^m\|_{L^{2+\delta}}^{2+\delta}+\|\Theta^m\|_{L^{2+\delta}}^{2+\delta}). \quad (3.43)$$

To bound $\|E^m\|_{L^{2+\delta}}^{2+\delta}$ in (3.43), we first make a shift in the super-index to get

$$\|E^m\|_{L^{2+\delta}}^{2+\delta}\leqslant\sum_{K\in\mathcal{T}_h}[\|E^{m-1}\|_{L^{2+\delta}(K)}^{2+\delta}+k^{2+\delta}\|d_tE^m\|_{L^{2+\delta}(K)}^{2+\delta}]. \quad (3.44)$$

For each term in the first sum on the right hand side of (3.44), we use the Gagliardo–Nirenberg inequality [1] which interpolates $L^{2+\delta}(K)$ between $L^2(K)$ and $H^4(K)$,

$$\|E^{m-1}\|_{L^{2+\delta}(K)}^{2+\delta}\leqslant C[\|D^4E^{m-1}\|_{L^2(K)}^{\delta N/8}\|E^{m-1}\|_{L^2(K)}^{(16+(8-N)\delta)/8}+\|E^{m-1}\|_{L^2(K)}^{2+\delta}]$$
$$\leqslant C\|E^{m-1}\|_{L^2(K)}^{2+(8-N)\delta/8}[\|D^4u(t_{m-1})\|_{L^2(K)}^{\delta N/8}+\|E^{m-1}\|_{L^2(K)}^{\delta N/8}]$$
$$\leqslant C\|E^{m-1}\|_{L^2(K)}^{2+(8-N)\delta/8}\rho_2(\varepsilon)^{\delta N/8}. \quad (3.45)$$

Here we have used the estimates of Lemmas 2.4(xi), 2.1(ii) and 3.1(v) to obtain the last inequality.

Summing (3.45) over all $K\in\mathcal{T}_h$ and using the convexity of the function $g(s)=s^r$ for $r>1$ and $s\geqslant 0$ we have

$$\|E^{m-1}\|_{L^{2+\delta}}^{2+\delta}\leqslant C\rho_2(\varepsilon)^{\delta N/8}\|E^{m-1}\|_{L^2}^{2+(8-N)\delta/8}. \quad (3.46)$$

Similarly, we can bound the second sum on the right hand side of (3.44):

$$\begin{aligned} k^{2+\delta}\|d_t E^m\|^{2+\delta}_{L^{2+\delta}(K)} &\leqslant k^{2+\delta}[\|D^4 d_t E^m\|^{\delta N/8}_{L^2(K)}\|d_t E^m\|^{(16+(8-N)\delta)/8}_{L^2(K)} + \|d_t E^m\|^{2+\delta}_{L^2(K)}] \\ &\leqslant k^{2+\delta}\|d_t E^m\|^{2+(8-N)\delta/8}_{L^2(K)}[\|D^4 d_t u(t_m)\|^{\delta N/8}_{L^2(K)} + \|d_t E^m\|^{\delta N/8}_{L^2(K)}] \\ &\leqslant Ck^{2+\delta}\|d_t E^m\|^{2+(8-N)\delta/8}_{L^2(K)}[k^{-N\delta/8}\|D^4 u(t_m)\|^{\delta N/8}_{L^2(K)} + \|d_t E^m\|^{\delta N/8}_{L^2(K)}] \\ &\leqslant Ck^{2+(8-N)\delta/8}\|d_t E^m\|^{2+(8-N)\delta/8}_{L^2(K)}\rho_2(\varepsilon)^{\delta N/8}. \end{aligned} \tag{3.47}$$

Here we have used the estimates of Lemmas 2.4(viii),(xi) and 3.1(v) to control $\|d_t E^m\|_{L^2(K)} \leqslant 2k^{-1}\max\{\|E^m\|_{L^2(K)}, \|E^{m-1}\|_{L^2(K)}\}$ in the last inequality.

Summing both sides of (3.47) over all $K \in \mathcal{T}_h$ we get

$$k^{2+\delta}\|d_t E^m\|^{2+\delta}_{L^{2+\delta}} \leqslant Ck^{2+(8-N)\delta/8}\rho_2(\varepsilon)^{\delta N/8}\|d_t E^m\|^{2+(8-N)\delta/8}_{L^2}. \tag{3.48}$$

It now follows from the triangle inequality $\|E^m\|_{L^2} \leqslant \|\Theta^m\|_{L^2} + \|\Phi^m\|_{L^2}$, the inequalities (3.43), (3.44), (3.46) and (3.48) that

$$\begin{aligned} \|\Phi^m\|^{2+\delta}_{L^{2+\delta}} \leqslant C\rho_2(\varepsilon)^{\delta N/8}(&\|\Phi^{m-1}\|^{2+(8-N)\delta/8}_{L^2} + k^{2+(8-N)\delta/8}\|d_t\Phi^m\|^{2+(8-N)\delta/8}_{L^2} \\ &+ \|\Theta^{m-1}\|^{2+(8-N)\delta/8}_{L^2} + k^{2+(8-N)\delta/8}\|d_t\Theta^m\|^{2+(8-N)\delta/8}_{L^2}) + C\|\Theta^m\|^{2+\delta}_{L^{2+\delta}}. \end{aligned} \tag{3.49}$$

From [11, 17, 34] we know that

$$\|u - P_h u\|_{L^2} \leqslant Ch^2\|u\|_{H^2}, \tag{3.50}$$

$$\|(u - P_h u)_t\|_{L^2} \leqslant Ch^2\|u_t\|_{H^2}, \tag{3.51}$$

$$\|u - P_h u\|_{L^{2+\delta}} \leqslant Ch^2\|u\|_{W^{2,2+\delta}} \leqslant Ch^2\|\Delta^2 u\|^{\frac{4(2+\delta)-2N}{(8-N)(2+\delta)}}_{L^2}. \tag{3.52}$$

To control the two terms which involve Φ^{m-1} on the right hand side of (3.49), we only consider the case $\delta < 16/(8-N)$ because (i) it covers most useful ranges of δ, and (ii) the analysis for the case $\delta > 16/(8-N)$ is easier to carry out since $(16+(8-N)\delta)/8 > 4$ in this case.

From the definition of $-\Delta_h^{-1}$ in (3.10) and Young's inequality we have

$$\begin{aligned} \|\Phi^{m-1}\|^{(16+(8-N)\delta)/8}_{L^2} &\leqslant \|\nabla\Phi^{m-1}\|^{(16+(8-N)\delta)/16}_{L^2}\|\nabla\Delta_h^{-1}\Phi^{m-1}\|^{(16+(8-N)\delta)/16}_{L^2} \\ &\leqslant C[\varepsilon^5\rho_2(\varepsilon)^{-\delta N/8}]^{-\frac{16+(8-N)\delta}{16-(8-N)\delta}}\|\nabla\Delta_h^{-1}\Phi^{m-1}\|^{\frac{2[16+(8-N)\delta]}{16-(8-N)\delta}}_{L^2} + \frac{\varepsilon^5\rho_2(\varepsilon)^{-\delta N/8}}{4\gamma_2}\|\nabla\Phi^{m-1}\|^2_{L^2}. \end{aligned} \tag{3.53}$$

Similarly,

$$\begin{aligned} \|d_t\Phi^m\|^{(16+(8-N)\delta)/8}_{L^2} &\leqslant \|\nabla d_t\Phi^m\|^{(16+(8-N)\delta)/16}_{L^2}\|\nabla\Delta_h^{-1}d_t\Phi^m\|^{(16+(8-N)\delta)/16}_{L^2} \\ &\leqslant C[\varepsilon k^{-(8+(8-N)\delta)/8}\rho_2(\varepsilon)^{-\delta N/8}]^{-\frac{16+(8-N)\delta}{16-(8-N)\delta}}\|\nabla d_t\Phi^m\|^{\frac{2[16+(8-N)\delta]}{16-(8-N)\delta}}_{L^2} \\ &\quad + \frac{\varepsilon k^{-(8+(8-N)\delta)/8}\rho_2(\varepsilon)^{-\delta N/8}}{4\gamma_2}\|\nabla d_t\Delta_h^{-1}\Phi^m\|^2_{L^2}. \end{aligned}$$

Using

$$\|\nabla d_t \Phi^m\|_{L^2}^{\frac{2[16+(8-N)\delta]}{16-(8-N)\delta}} = \|\nabla d_t \Phi^m\|_{L^2}^2 \|\nabla d_t \Phi^m\|_{L^2}^{\frac{4(8-N)\delta}{16-(8-N)\delta}}$$
$$\leqslant k^{-\frac{4(8-N)\delta}{16-(8-N)\delta}} \varepsilon^{-(\sigma_1+1/2)\frac{4(8-N)\delta}{16-(8-N)\delta}} \|\nabla d_t \Phi^m\|_{L^2}^2,$$

we get

$$\|d_t \Phi^m\|_{L^2}^{(16+(8-N)\delta)/8} \leqslant (\varepsilon k^{-8+(8-N)\delta/8} \rho_2(\varepsilon)^{-\delta N/8})/4\gamma_2 \|\nabla d_t \Delta_h^{-1} \Phi^m\|_{L^2}^2$$
$$+ C\varepsilon^{-[1+\frac{4\delta(8-N)(\sigma_1+1)}{16-(8-N)\delta}]} \rho_2(\varepsilon)^{-\frac{\delta N[16+(8-N)\delta]}{8[16-(8-N)\delta]}} k^{1+\frac{(8-N)^2\delta^2}{8[16-(8-N)\delta]}} \|\nabla d_t \Phi^m\|_{L^2}^2. \tag{3.54}$$

Now, substituting (3.50)–(3.54) into (3.49), summing over m from 1 to ℓ ($\leqslant M$) after multiplying (3.49) by $\gamma_2 k/\varepsilon$ and using Lemmas 2.4(iv) and 3.1(vi) leads to the following estimate:

$$\frac{\gamma_2 k}{\varepsilon} \sum_{m=1}^{\ell} \|\Phi^m\|_{L^{2+\delta}}^{2+\delta} \leqslant k \sum_{m=1}^{\ell} \left[\frac{\varepsilon^4}{8} \|\nabla \Phi^m\|_{L^2}^2 + \frac{k}{8} \|\nabla \Delta_h^{-1} d_t \Phi^m\|_{L^2}^2 \right]$$
$$+ C\varepsilon^{-\frac{4[24+(8-N)\delta]}{16-(8-N)\delta}} \rho_2(\varepsilon)^{\frac{4N\delta}{16-(8-N)\delta}} k \sum_{m=1}^{\ell} \|\nabla \Delta_h^{-1} \Phi^{m-1}\|_{L^2}^{2+\frac{4(8-N)\delta}{16-(8-N)\delta}}$$
$$+ Ck^{3+\frac{2(8-N)\delta}{16-(8-N)\delta}} \varepsilon^{-[2+\frac{4\delta(8-N)(\sigma_1+1)}{16-(8-N)\delta}]} \rho_2(\varepsilon)^{\frac{4N\delta}{16-(8-N)\delta}} \varepsilon^{-2(\sigma_1+2)}$$
$$+ C\rho_2(\varepsilon)^{\delta N/8} [\varepsilon^{-\frac{(2\sigma_1+1)[16+(8-N)\delta]}{16}} h^{4+(8-N)\delta/4}$$
$$+ k^2 h^4 \varepsilon^{-\max\{2\sigma_1+9, 2\sigma_3+6, 2\sigma_2+4, 2\sigma_2+1\}}] + Ch^{2(2+\delta)} \rho_2(\varepsilon)^{\frac{4(2+\delta)-2N}{8-N}}. \tag{3.55}$$

Finally, substituting (3.55) into (3.42) we get

$$\frac{1}{2} \|\nabla \Delta_h^{-1} \Phi^\ell\|_{L^2}^2 + k \sum_{m=1}^{\ell} \left[\frac{k}{8} \|\nabla \Delta_h^{-1} d_t \Phi^m\|_{L^2}^2 + \frac{\varepsilon^4}{8} \|\nabla \Phi^m\|_{L^2}^2 \right]$$
$$\leqslant Cr(h, k; \varepsilon, N, \delta, \sigma_i, \nu) + (C_0 + 3\tilde{c}_0 + 3)k \sum_{m=1}^{\ell} \|\nabla \Delta_h^{-1} \Phi^m\|_{L^2}^2$$
$$+ \frac{1}{2} \|\nabla \Delta_h^{-1} \Phi^0\|_{L^2}^2 + Cs(\varepsilon, N, \delta) k \sum_{m=1}^{\ell-1} \|\nabla \Delta_h^{-1} \Phi^m\|_{L^2}^{2+\frac{4(8-N)\delta}{16-(8-N)\delta}}, \tag{3.56}$$

where $r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)$ is defined in (3.28) and

$$s(\varepsilon, N, \delta) = \varepsilon^{-\frac{4[24+(8-N)\delta]}{16-(8-N)\delta}} \rho_2(\varepsilon)^{\frac{4N\delta}{16-(8-N)\delta}}. \tag{3.57}$$

STEP 4 We now conclude the proof by the following induction argument. Suppose there exist two positive constants

$$c_1 = c_1(t_\ell, \Omega, u_0^\varepsilon, \sigma_i), \quad c_2 = c_2(t_\ell, \Omega, u_0^\varepsilon, \sigma_i; C_0),$$

independent of k and ε, such that

$$\max_{0 \leqslant m \leqslant \ell} \|\nabla \Delta_h^{-1} \Phi^m\|_{L^2}^2 + k \sum_{m=1}^{\ell} \left[\frac{k}{8} \|\nabla \Delta_h^{-1} d_t \Phi^m\|_{L^2}^2 + \frac{\varepsilon^4}{8} \|\nabla \Phi^m\|_{L^2}^2 \right]$$
$$\leqslant c_1 r(h, k; \varepsilon, N, \delta, \sigma_i, \nu) \exp(c_2 t_\ell). \tag{3.58}$$

In view of Lemma 2.4(xi) and (3.56), we can choose

$$c_1 = 2, \quad c_2 = 2(C_0 + 3\tilde{c}_0 + 3).$$

Since the exponent in the last term of (3.56) is greater than 2, we can recover (3.58) at the $(\ell+1)$th time step by using the discrete Gronwall inequality, provided that h, k satisfy

$$s(\varepsilon, N, \delta) \cdot r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)^{1+\frac{2(8-N)\delta}{16-(8-N)\delta}} \leqslant \frac{c_1}{2} r(h, k; \varepsilon, N, \delta, \sigma_i, \nu) \exp(c_2 t_{\ell+1}).$$

That is,

$$s(\varepsilon, N, \delta) \cdot r(h, k; \varepsilon, N, \delta, \sigma_i, \nu)^{\frac{2(8-N)\delta}{16-(8-N)\delta}} \leqslant C,$$

which gives the mesh condition 3) in the theorem. Hence, we have shown that

$$\max_{0\leqslant m\leqslant M} \|\nabla \Delta_h^{-1} \Phi^m\|_{L^2}^2 + k \sum_{m=1}^{\ell} [k \|\nabla \Delta_h^{-1} d_t \Phi^m\|_{L^2}^2 + \varepsilon^4 \|\nabla \Phi^m\|_{L^2}^2] \leqslant C r(h, k; \varepsilon, N, \delta, \sigma_i, \nu). \tag{3.59}$$

Finally, assertion (i) follows from (3.59) by applying the triangle inequality to $E^m = \Theta^m + \Phi^m$. Assertions (ii) and (iii) follow in the same way. Note that we need to apply the Poincaré inequality to Φ^m to show (ii), and since $\Phi^m \in \mathring{\mathcal{S}}_h$, we can bound $\|\Phi^m\|_{L^2}$ by $\|\nabla \Phi^m\|_{L^2}$. The proof is complete. □

REMARK (a) The $L^2(J; H^1)$ estimate is optimal with respect to h and k, and the $L^\infty(J; H^{-1})$ estimate is quasi-optimal.

(b) The proof clearly shows how the three mesh conditions arise. Condition 1) is for the stability of the time discretization (see (GA_3)), condition 2) is to ensure the discrete spectrum estimate (see Lemma 3.2), finally, condition 3) is caused by the super-quadratic nonlinearity of f (see $(GA_1)_3$). Also notice that only “smallness” of k and h with respect to ε is required but no restriction is imposed on the ratio between k and h in the $L^\infty(J; H^{-1})$ and $L^2(J; H^1)$ norm estimates.

(c) It is well known [34] that the finite element solutions of all linear and some nonlinear parabolic problems exhibit a superconvergence property (in h) when compared with the elliptic projections of the solutions of underlying problems. It is worth pointing out that this superconvergence also holds for the Cahn–Hilliard equation as shown by the inequality (3.59).

(d) Regarding the choices of the starting value U^0, clearly, both $U^0 = Q_h u_0^\varepsilon$ and $U^0 = P_h u_0^\varepsilon$ satisfy conditions 4) and 5) with $\nu = 1$ in view of (3.7) and (3.9). In fact, they also satisfy a stronger inequality (see (3.61) below). On the other hand, the L^2 projection $Q_h u_0^\varepsilon$ is cheaper to compute compared to the elliptic projection $P_h u_0^\varepsilon$. Note that condition 4) is necessary in order for the scheme (3.4)–(3.5) to conserve the mass.

(e) We remark that it is easy to check that for each fixed $\varepsilon > 0$, if $k \leqslant h$ then

$$r(h, k; \varepsilon, N, \delta, \sigma_i, \nu) \leqslant C(k^2 + h^{2(2+\mu)}). \tag{3.60}$$

In the next theorem we derive error estimates in stronger norms under a slightly stronger requirement on the starting value U^0, which nevertheless is satisfied by both the L^2 projection $U^0 = Q_h u_0^\varepsilon$ and the elliptic projection $U^0 = P_h u_0^\varepsilon$. In addition, a mild constraint on admissible choices of (k, h) is required to assure their validity.

THEOREM 3.2 In addition to the assumptions of Theorem 3.1, if

$$\|u_0^\varepsilon - U^0\|_{L^2} \leqslant Ch^2 \|u_0^\varepsilon\|_{H^2}, \tag{3.61}$$

then the solution of (3.4)–(3.5) also satisfies the following error estimates:

(i) $$\max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{L^2} + \Big(k \sum_{m=1}^M k\|d_t(u(t_m) - U^m)\|_{L^2}^2\Big)^{1/2} \leqslant \tilde{C}\{h^2\varepsilon^{-\max\{\sigma_1+5/2,\sigma_3+1/2\}} + k^{-1/4}\varepsilon^{-1}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\},$$

(ii) $$\max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{L^\infty} \leqslant \tilde{C}\{h^2|\ln h|\rho_2(\varepsilon)^{4/8-N} + h^{-N/2}k^{-1/4}\varepsilon^{-1}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\}.$$

Moreover, if $k = O(h^q)$ for some $2N/3 < q < (8-2N) + 4\mu$, then also

(iii) $$\max_{0\leqslant m\leqslant M} \|U^m\|_{L^\infty} \leqslant 3C_0,$$

(iv) $$\max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{L^2} + \Big(k \sum_{m=1}^M k\|d_t(u(t_m) - U^m)\|_{L^2}^2\Big)^{1/2} + \Big(\frac{k}{\varepsilon} \sum_{m=1}^M \|w(t_m) - W^m\|_{L^2}^2\Big)^{1/2} \leqslant \tilde{C}\{h^2\varepsilon^{-\max\{\sigma_1+7/2,\sigma_3+1/2\}} + \varepsilon^{-7/2}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\},$$

(v) $$\max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{L^\infty} \leqslant \tilde{C}\{h^2|\ln h|\rho_2(\varepsilon)^{4/(8-N)} + h^{(4-N)/2}\varepsilon^{-\max\{\sigma_1+7/2,\sigma_3+1/2\}} + h^{-N/2}\varepsilon^{-7/2}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\}$$

for some positive constant $\tilde{C} = \tilde{C}(u_0^\varepsilon; \gamma_2, C_0, T; \Omega)$.

Proof. Since

$$E^m = \Theta^m + \Phi^m, \qquad G^m = \Lambda^m + \Psi^m,$$

it suffices to show that assertions (i), (ii), (iv), (v) hold for Φ^m and Ψ^m without the first term on the right hand side of each inequality. Notice that Φ^m and Ψ^m satisfy (3.33)–(3.34).

Using the identity

$$(d_t\Phi^m, \Phi^m) = \frac{1}{2} d_t \|\Phi^m\|_{L^2}^2 + \frac{k}{2}\|d_t\Phi^m\|_{L^2}^2,$$

the definition of $-\Delta_h^{-1}$ in (3.10) and the estimate (3.59) we have

$$\begin{aligned}
\frac{1}{2}\|\Phi^\ell\|_{L^2}^2 + k\sum_{m=1}^{\ell}\frac{k}{2}\|d_t\Phi^m\|_{L^2}^2 &= k\sum_{m=1}^{\ell}(\nabla(-\Delta_h^{-1}d_t\Phi^m), \nabla\Phi^m) + \frac{1}{2}\|\Phi^0\|_{L^2}^2 \\
&\leqslant k^{-1/2}\varepsilon^{-2}\sum_{m=1}^{\ell}[k^2\|\nabla\Delta_h^{-1}d_t\Phi^m\|_{L^2}^2 + k\varepsilon^4\|\nabla\Phi^m\|_{L^2}^2] + \frac{1}{2}\|\Phi^0\|_{L^2}^2 \\
&\leqslant k^{-1/2}\varepsilon^{-2} r(h,k;\varepsilon,N,\delta,\sigma,\nu) + \frac{1}{2}\|\Phi^0\|_{L^2}^2.
\end{aligned} \tag{3.62}$$

Assertion (i) then follows from (3.62) and (3.61).

Assertion (ii) is an immediate consequence of (i), the inverse inequality bounding the L^∞ by the L^2-norm, and the L^∞ estimate of Θ^m (see Chapter 7 of [11]).

To show (iii), notice that under the mesh conditions of Theorem 3.1 and the assumption that $k = O(h^q)$ for some $2N/3 < q < (8-2N) + 4\mu$, we have, for sufficiently small ε,

$$\max_{0\leqslant m\leqslant M} \|u(t_m) - U^m\|_{L^\infty} \leqslant \tfrac{3}{2}C_0, \tag{3.63}$$

which together with (2.9) then implies

$$\max_{0\leqslant m\leqslant M} \|U^m\|_{L^\infty} \leqslant \max_{0\leqslant m\leqslant M} [\|u(t_m)\|_{L^\infty} + \|u(t_m) - U^m\|_{L^\infty}] \leqslant 3C_0. \tag{3.64}$$

Hence (iii) holds.

Now, taking $\eta_h = \Phi^m$ in (3.33) and $v_h = -\frac{1}{\varepsilon}\Psi^m$ in (3.34) and adding the resulting equations we get

$$\begin{aligned}\frac{1}{2}d_t\|\Phi^m\|_{L^2}^2 + \frac{k}{2}\|d_t\Phi^m\|_{L^2}^2 + \frac{1}{\varepsilon}\|\Psi^m\|_{L^2}^2 &= (\mathcal{R}(u_{tt};m), \Phi^m) - (d_t\Theta^m, \Phi^m) - \frac{1}{\varepsilon}(\Lambda^m, \Psi^m)\\ &\quad + \frac{1}{\varepsilon^2}(f(u(t_m)) - f(U^m), \Psi^m).\end{aligned} \tag{3.65}$$

The first three terms on the right hand side can be bounded as in (3.36), and the last term can be bounded as follows. By the Mean Value Theorem and Schwarz inequality we obtain

$$\frac{1}{\varepsilon^2}(f(u(t_m)) - f(U^m), \Psi^m) = \frac{1}{\varepsilon^2}(f'(\xi)E^m, \Psi^m) \leqslant \frac{1}{2\varepsilon}\|\Psi^m\|_{L^2}^2 + \frac{C}{\varepsilon^3}\|E^m\|_{L^2}^2. \tag{3.66}$$

Assertion (iv) follows from multiplying (3.65) by k, summing it over m from 1 to ℓ ($\leqslant M$) and using (3.66) and Theorem 3.1(ii).

Finally, (v) is a refinement of (ii), based on (iv) instead of (i). The proof is complete. □

REMARK (a) The estimate in (i) is optimal in h and suboptimal in k due to the factor $k^{-1/4}$ in the second term on the right hand side of the inequality. However, this estimate is important for establishing the $L^\infty(J; L^\infty)$ estimate in (ii), which then leads to the proof of the boundedness of U^m in (3.64), and the improved estimates (iv) and (v).

(b) Optimal estimates in stronger norms can also be obtained for both E^m and G^m under stronger regularity assumptions on the solution u (e.g. $u_{tt} \in L^2(J; L^2)$) of the Cahn–Hilliard equation and on the starting value U^0. These estimates include statements for E^m in $L^\infty(J; H^1)$ and $H^1(J; L^2)$, and in $L^\infty(J; L^2)$ and $L^2(J; H^1)$ for G^m. For more details in this direction, we refer to [23] (also see [19]), where a (continuous in time) semi-discrete splitting finite element method was analyzed for a *fixed* $\varepsilon > 0$ under the assumption that the semi-discrete finite element approximate solution for u is bounded in L^∞. Note that here we have indeed showed in (iii) that our fully discrete solution U^m is bounded in L^∞.

COROLLARY 3.1 Let the assumptions of Theorem 3.2 be valid, and W^0 be a value satisfying, for any $\beta > 1$,

$$\|P_h w(0) - W^0\|_{L^2} \leqslant Ch^\beta. \tag{3.67}$$

Then

$$\text{(i)} \quad \max_{0\leqslant m\leqslant M} \|w(t_m) - W^m\|_{L^2} + \Big(k\sum_{m=1}^{M} k\|d_t(w(t_m) - W^m)\|_{L^2}^2\Big)^{1/2}$$
$$\leqslant \tilde{C}\{h^2\rho_2(\varepsilon) + k^{-1/2}\{h^2\varepsilon^{-\max\{\sigma_1+3,\sigma_3\}} + \varepsilon^{-3}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\} + h^\beta\},$$

$$\text{(ii)} \quad \max_{0\leqslant m\leqslant M} \|w(t_m) - W^m\|_{L^\infty} \leqslant \tilde{C}\{h^{(4-N)/2}|\ln h|^{(3-N)/2}\rho_2(\varepsilon) + h^{-N/2}[k^{-1/2}\{h^2\varepsilon^{-\max\{\sigma_1+3,\sigma_3\}}$$
$$+ \varepsilon^{-3}[r(h,k;\varepsilon,N,\delta,\sigma_i,\nu)]^{1/2}\} + h^\beta]\}.$$

Proof. First, from [11, 34] we know that

$$\max_{0\leqslant m\leqslant M} \|\Lambda^m\|_{L^2} + \Big(k\sum_{m=1}^{M} k\|d_t\Lambda^m\|_{L^2}^2\Big)^{1/2} \leqslant \tilde{C}h^2\rho_2(\varepsilon). \tag{3.68}$$

Next, using the identity which immediately precedes (3.62) we get

$$\frac{1}{2}\|\Psi^\ell\|_{L^2}^2 + k\sum_{m=1}^{\ell}\frac{k}{2}\|d_t\Psi^m\|_{L^2}^2 = k\sum_{m=1}^{\ell}(d_t\Psi^m, \Psi^m) + \frac{1}{2}\|\Psi^0\|_{L^2}^2$$
$$\leqslant k\sum_{m=1}^{\ell}\Big[\frac{k}{4}\|d_t\Psi^m\|_{L^2}^2 + 4k^{-1}\|\Psi^m\|_{L^2}^2\Big] + \frac{1}{2}\|\Psi^0\|_{L^2}^2. \tag{3.69}$$

The first term on the right hand side can be absorbed by the second term on the left, and a desired bound for the second term on the right has been obtained in the proof of Theorem 3.2(iv). Hence, (i) follows by combining (3.68) and (3.69).

Assertion (ii) comes from applying the triangle inequality to $G^m = \Lambda^m + \Psi^m$, the estimate (cf. Section 4 of [28])

$$\|\Lambda^m\|_{L^\infty} \leqslant Ch^{(4-N)/2}|\ln h|^{(3-N)/2}\|w\|_{H^2} \leqslant Ch^{(4-N)/2}|\ln h|^{(3-N)/2}\rho_2(\varepsilon),$$

and the inverse inequality bounding $\|\Psi^m\|_{L^\infty}$ by $\|\Psi^m\|_{L^2}$. The proof is complete. □

REMARK (a) Clearly, the solution $\{(U^m, W^m)\}_{m=1}^M$ to (3.4)–(3.5) does not depend on W^0. However, estimates (i) and (ii), which will be needed for the convergence analysis in Section 4, do depend on the choice of W^0. Recall that W^m approximates $w = -\varepsilon\Delta u + (1/\varepsilon)f(u)$, hence, estimates (i) and (ii) bound the error $u(t_m) - U^m$ in higher norms, which in turn puts a constraint like (3.67) on the choice of W^0.

(b) w_0^ε is defined by setting $t = 0$ in (2.2). Clearly, both $Q_h w_0^\varepsilon$ and $P_h w_0^\varepsilon$ are valid candidates for W^0.

(c) Both estimates are not optimal due to the factor $k^{-1/2}$ in the second term on the right hand side of each inequality. It can be shown that the estimates will be improved to optimal order (first order in k and second order in h) under some stronger regularity assumptions and starting value constraint. See (b) of the remark after the proof of Theorem 3.2.

4. Approximation for the Hele–Shaw problem

The goal of this section is to establish the convergence of the solution $\{(U^m, W^m)\}_{m=0}^M$ of the fully discrete mixed finite element scheme (3.4)–(3.5) to the solution of the Hele–Shaw problem (1.4)–(1.8), provided that the Hele–Shaw problem has a global (in time) classical solution. It is shown that the fully discrete solution W^m, as $h, k \searrow 0$, converges to the solution w of the Hele–Shaw problem uniformly in $\overline{\Omega}_T$. In addition, the fully discrete solution U^m converges to ± 1 uniformly on every compact subset of the "outside" and "inside" of the free boundary Γ of the Hele–Shaw problem, respectively. Hence, the zero level set of U^m converges to the free boundary Γ. Our main ideas are to make full use of the convergence result that the Hele–Shaw problem is the distinguished limit, as $\varepsilon \searrow 0$, of the Cahn–Hilliard equation proved by Alikakos, Bates and Chen in [2], and to exploit the "closeness" between the solution u of the Cahn–Hilliard equation and its fully discrete approximation U^m, which is demonstrated by the error estimates in the previous section. We remark that as in [2], our numerical convergence result is also established under the assumption that the Hele–Shaw problem has a global (in time) classical solution. We refer to [2, 16] and references therein for further details on this assumption and related theoretical works on the Hele–Shaw problem.

Although it can be shown that the results of this section hold for a general potential $F(u)$ satisfying (GA_1), for the sake of clarity of the presentation, we only consider the quartic potential $F(u) = \frac{1}{4}(u^2 - 1)^2$ in this section. Let u^ε denote the solution of the Cahn–Hilliard problem (1.1)–(1.3). Note that we put back the super-index ε on the solution in this section. Let $(U_{\varepsilon,h,k}(x,t), W_{\varepsilon,h,k}(x,t))$ denote the piecewise linear interpolation (in time) of the fully discrete solution (U^m, W^m), that is,

$$U_{\varepsilon,h,k}(\cdot,t) := \frac{t - t_m}{k} U^{m+1}(\cdot) + \frac{t_{m+1} - t}{k} U^m(\cdot), \tag{4.1}$$

$$W_{\varepsilon,h,k}(\cdot,t) := \frac{t - t_m}{k} W^{m+1}(\cdot) + \frac{t_{m+1} - t}{k} W^m(\cdot), \tag{4.2}$$

for $t_m \leqslant t \leqslant t_{m+1}$ and $0 \leqslant m \leqslant M-1$. Note that W^0 is defined in Corollary 3.1, and $U_{\varepsilon,h,k}(x,t)$ and $W_{\varepsilon,h,k}(x,t)$ are continuous piecewise linear functions in space and time.

Let $\Gamma_{00} \subset \Omega$ be a smooth closed hypersurface and let $(w, \Gamma := \bigcup_{0 \leqslant t \leqslant T}(\Gamma_t \times \{t\}))$ be a smooth solution of the Hele–Shaw problem (1.4)–(1.8) starting from Γ_{00} such that $\Gamma \subset \Omega \times [0,T]$. Let $d(x,t)$ denote the *signed distance function* to Γ_t such that $d(x,t) < 0$ in $\mathcal{I}_t$, the *inside* of Γ_t, and $d(x,t) > 0$ in $\mathcal{O}_t := \Omega \setminus (\Gamma_t \cup \mathcal{I}_t)$, the *outside* of Γ_t. We also define the *inside* $\mathcal{I}$ and the *outside* $\mathcal{O}$ of Γ as follows:

$$\mathcal{I} := \{(x,t) \in \Omega \times [0,T];\, d(x,t) < 0\}, \qquad \mathcal{O} := \{(x,t) \in \Omega \times [0,T];\, d(x,t) > 0\}.$$

For the numerical solution $U_{\varepsilon,h,k}(x,t)$, we denote its zero level set at time t by $\Gamma_t^{\varepsilon,h,k}$, that is,

$$\Gamma_t^{\varepsilon,h,k} := \{x \in \Omega;\, U_{\varepsilon,h,k}(x,t) = 0\}. \tag{4.3}$$

Before we state our convergence theorem, Theorem 4.2, we need to recall the following convergence result (see Theorem 5.1 of [2]), which proved that the Hele–Shaw problem is the distinguished limit, as $\varepsilon \searrow 0$, of the Cahn–Hilliard equation.

THEOREM 4.1 Let Ω be a given smooth domain and Γ_{00} be a smooth closed hypersurface in Ω. Suppose that the Hele–Shaw problem (1.4)–(1.8) starting from Γ_{00} has a smooth solution $(w, \Gamma :=$

$\bigcup_{0\leqslant t\leqslant T}(\Gamma_t \times \{t\}))$ in the time interval $[0, T]$ such that $\Gamma_t \subset \Omega$ for all $t \in [0, T]$. Then there exists a family of smooth functions $\{u_0^\varepsilon(x)\}_{0<\varepsilon\leqslant 1}$ which are uniformly bounded in $\varepsilon \in (0, 1]$ and $(x, t) \in \overline{\Omega}_T$, such that if u^ε solves the Cahn–Hilliard equation (1.1)–(1.3), then

$$\text{(i)}\quad \lim_{\varepsilon\to 0} u^\varepsilon(x,t) = \begin{cases} 1 & \text{if } (x,t)\in\mathcal{O} \\ -1 & \text{if } (x,t)\in\mathcal{I} \end{cases} \quad \text{uniformly on compact subsets,}$$

$$\text{(ii)}\quad \lim_{\varepsilon\to 0} \left(\frac{1}{\varepsilon} f(u^\varepsilon) - \varepsilon\Delta u^\varepsilon\right)(x,t) = w(x,t) \quad \text{uniformly on } \overline{\Omega}_T.$$

We are now ready to state the following main theorem of this section.

THEOREM 4.2 Let Ω be a given smooth domain and Γ_{00} be a smooth closed hypersurface in Ω. Suppose that the Hele–Shaw problem (1.4)–(1.8) starting from Γ_{00} has a classical solution $(w, \Gamma := \bigcup_{0\leqslant t\leqslant T}(\Gamma_t \times \{t\}))$ in the time interval $[0, T]$ such that $\Gamma_t \subset \Omega$ for all $t \in [0, T]$. Let $\{u_0^\varepsilon(x)\}_{0<\varepsilon\leqslant 1}$ be the family of smooth uniformly bounded functions as in Theorem 5.1 of [2]. Let $(U_{\varepsilon,h,k}(x,t), W_{\varepsilon,h,k}(x,t))$ denote the piecewise linear interpolation (in time) of the fully discrete solution $\{(U^m, W^m)\}_{m=0}^M$. Also, let $\mathcal{I}$ and $\mathcal{O}$ stand for the "inside" and "outside" (in Ω_T) of Γ. Then, under the mesh and starting value constraints of Theorem 3.1, with $\nu = 1$ and $k = O(h^q)$ for some $2N/3 < q < (8-2N) + 4\mu$ we have

$$\text{(i)}\quad U_{\varepsilon,h,k}(x,t) \overset{\varepsilon\searrow 0}{\longrightarrow} 1 \qquad \text{uniformly on compact subsets of } \mathcal{O},$$

$$\text{(ii)}\quad U_{\varepsilon,h,k}(x,t) \overset{\varepsilon\searrow 0}{\longrightarrow} -1 \qquad \text{uniformly on compact subsets of } \mathcal{I}.$$

Moreover, when $N = 2$, let $k = O(h^q)$ for some $N < q < (4-N) + 2\mu$ and choose W^0 such that $\|w_0^\varepsilon - W^0\|_{L^2} \leqslant Ch^\beta$ for some $\beta > q/2$; then we also have

$$\text{(iii)}\quad W_{\varepsilon,h,k}(x,t) \overset{\varepsilon\searrow 0}{\longrightarrow} w(x,t) \qquad \text{uniformly on } \overline{\Omega}_T.$$

Proof. Let A be any compact subset of $\mathcal{O}$. For any $(x,t) \in A$, using the triangle inequality we have

$$\begin{aligned} |U_{\varepsilon,h,k}(x,t) - 1| &\leqslant |U_{\varepsilon,h,k}(x,t) - u^\varepsilon(x,t)| + |u^\varepsilon(x,t) - 1| \\ &\leqslant \|U_{\varepsilon,h,k} - u^\varepsilon\|_{L^\infty(\Omega_T)} + |u^\varepsilon(x,t) - 1|. \end{aligned} \tag{4.4}$$

Under the assumptions of Theorem 4.2, from Theorem 3.2(v) we know that there exists a constant $0 < \alpha < (4-N)/2$ such that

$$\|U_{\varepsilon,h,k} - u^\varepsilon\|_{L^\infty(\Omega_T)} \leqslant Ch^\alpha. \tag{4.5}$$

Here we have used the assumption $k = O(h^q)$ for some $2N/3 < q < (24-5N)/2$.

Clearly, the first term on the right hand side of (4.4) converges to zero uniformly on A (and on Ω) as $h \searrow 0$. From Theorem 4.1(i) we know that the second term on the right hand side of (4.4) also converges to zero uniformly on A. Note that $h \searrow 0$ as $\varepsilon \searrow 0$. Therefore,

$$U_{\varepsilon,h,k} \overset{\varepsilon\searrow 0}{\longrightarrow} 1 \quad \text{uniformly on } A.$$

This then completes the proof of (i).

The proof of (ii) is almost the same. The only change is to replace $\mathcal{O}$ by $\mathcal{I}$ and 1 by -1 in the above proof. So we omit it.

To show (iii), first we notice that

$$w^\varepsilon = \frac{1}{\varepsilon} f(u^\varepsilon) - \varepsilon \Delta u^\varepsilon$$

if the solution u^ε of the Cahn–Hilliard equation (1.1)–(1.3) belongs to $W^{1,\infty}(J; L^2) \cap L^\infty(J; H^4)$. Next, from Corollary 3.1(ii) we know that under the additional assumptions of Theorem 4.2 there exists a positive constant $0 < \zeta < (q-2)/2$ such that

$$\|W_{\varepsilon,h,k} - w^\varepsilon\|_{L^\infty(\Omega_T)} \leqslant C h^\zeta. \tag{4.6}$$

Here we have used the assumption $k = O(h^q)$ for some $2 < q < 7/2$.

By the triangle inequality for any $(x, t) \in \overline{\Omega}_T$ we have

$$\begin{aligned} |W_{\varepsilon,h,k}(x,t) - w| &\leqslant |W_{\varepsilon,h,k}(x,t) - w^\varepsilon(x,t)| + |w^\varepsilon(x,t) - w| \\ &\leqslant \|W_{\varepsilon,h,k} - w^\varepsilon\|_{L^\infty(\Omega_T)} + |w^\varepsilon(x,t) - w|. \end{aligned} \tag{4.7}$$

The first term on the right hand side of (4.7) clearly converges to zero uniformly as $h \searrow 0$, and so does the second term due to Theorem 4.1(ii). Hence,

$$W_{\varepsilon,h,k}(x,t) \overset{\varepsilon \searrow 0}{\longrightarrow} w(x,t) \quad \text{uniformly on } \overline{\Omega}_T.$$

The proof is complete. □

REMARK (a) The reason for us to only show assertion (iii) for $N = 2$ is that the current $L^\infty(J; L^\infty)$ estimate for $w^\varepsilon - W_{\varepsilon,h,k}$ in Corollary 3.1(ii) is not strong enough to give a positive power of h (note $k = O(h^q)$) in the error bound when $N = 3$. To circumvent the difficulty, we need a better $L^\infty(J; L^\infty)$ estimate for $w^\varepsilon - W_{\varepsilon,h,k}$ which is similar to the one for $u^\varepsilon - U_{\varepsilon,h,k}$ in Theorem 3.2(v). This can be done under the assumption that $u_{tt} \in L^2(J; L^2)$ (needed to derive a priori estimate in $1/\varepsilon$) and that the starting value U^0 satisfies the following stronger constraint:

$$\|\nabla(P_h u_0^\varepsilon - U^0)\|_{L^2} \leqslant C h^2. \tag{4.8}$$

See (b) of the remark after Theorem 3.2.

A corollary of Theorem 4.2 is the following convergence result of the zero level set $\Gamma_t^{\varepsilon,h,k}$ of $U_{\varepsilon,h,k}$ to the free boundary Γ_t.

THEOREM 4.3 Let $\Gamma_t^{\varepsilon,h,k} := \{x \in \Omega; U_{\varepsilon,h,k}(x,t) = 0\}$ denote the zero level set of $U_{\varepsilon,h,k}$. Then under the assumptions for Theorem 4.2(i),(ii), we have

$$\sup_{x \in \Gamma_t^{\varepsilon,h,k}} \operatorname{dist}(x, \Gamma_t) \overset{\varepsilon \searrow 0}{\longrightarrow} 0 \quad \text{uniformly on } [0, T].$$

Proof. For any $\eta \in (0, 1)$, define the (open) tubular neighborhood $\mathcal{N}_\eta$ of width 2η of Γ as

$$\mathcal{N}_\eta := \{(x,t) \in \Omega_T; d(x,t) < \eta\}. \tag{4.9}$$

Let A and B denote the complements of $\mathcal{N}_\eta$ in $\mathcal{O}$ and $\mathcal{I}$, respectively, that is,

$$A = \mathcal{O} \setminus \mathcal{N}_\eta, \qquad B = \mathcal{I} \setminus \mathcal{N}_\eta.$$

Note that A is a compact subset of $\mathcal{O}$ and B is a compact subset of $\mathcal{I}$. Hence, from Theorem 4.2(i),(ii) we know that there exists $\widehat{\varepsilon}_0 > 0$, which only depends on η, such that for all $\varepsilon \in (0, \widehat{\varepsilon}_0)$,

$$|U_{\varepsilon,h,k}(x,t) - 1| \leqslant \eta \quad \forall (x,t) \in A, \tag{4.10}$$

$$|U_{\varepsilon,h,k}(x,t) + 1| \leqslant \eta \quad \forall (x,t) \in B. \tag{4.11}$$

Now for any $t \in [0,T]$ and $x \in \Gamma_t^{\varepsilon,h,k}$, since $U_{\varepsilon,h,k}(x,t) = 0$, we have

$$|U_{\varepsilon,h,k}(x,t) - 1| = 1, \tag{4.12}$$

$$|U_{\varepsilon,h,k}(x,t) + 1| = 1. \tag{4.13}$$

Evidently, (4.10) and (4.12) imply that $(x,t) \notin A$, and (4.11) and (4.13) say that $(x,t) \notin B$. Hence (x,t) must reside in the tubular neighborhood $\mathcal{N}_\eta$. Since t is an arbitrary number in $[0,T]$ and x is an arbitrary point on $\Gamma_t^{\varepsilon,h,k}$, therefore, for all $\varepsilon \in (0, \widehat{\varepsilon}_0)$,

$$\sup_{x \in \Gamma_t^{\varepsilon,h,k}} \operatorname{dist}(x, \Gamma_t) \leqslant \eta \quad \text{uniformly on } [0,T]. \tag{4.14}$$

The proof is complete. □

We conclude this section and the paper with some discussions about the rate of convergence of $\Gamma_t^{\varepsilon,h,k}$ to Γ_t.

It is well known (see [4, 14, 25]) that the solution for the Allen–Cahn equation $u_t^\varepsilon = \Delta u^\varepsilon - (1/\varepsilon^2) f(u^\varepsilon)$ approaches ± 1 away from the interface exponentially fast. This property allows estimating the rate of convergence for the zero level set of the solution of the Allen–Cahn equation and its numerical approximations to the true interface (see [9, 28, 31, 32] and references therein).

Unlike the situation for the Allen–Cahn equation, the solution u^ε of the Cahn–Hilliard equation (1.1)–(1.3) *does not* approach ± 1 away from the interface exponentially fast, and the transition region from 1 to -1 could be “large” (see [2]). In fact, it was shown in Theorem 4.12 of [2] that this transition region is contained in a tubular neighborhood of width δ^* of Γ, where δ^* is a constant such that $\operatorname{dist}(\Gamma_t, \partial\Omega) > 2\delta^*$ for all $t \in [0,T]$. The combination of this result with the $L^\infty(J; L^\infty)$ estimate for $u^\varepsilon - U_{\varepsilon,h,k}$ immediately leads to the following theorem.

THEOREM 4.4 Let δ^* be a positive constant such that $\operatorname{dist}(\Gamma_t, \partial\Omega) > 2\delta^*$ for all $t \in [0,T]$. Then, under the assumptions for Theorem 4.2(i),(ii), there exists a (small) positive number $\widehat{\varepsilon} > 0$ such that

$$\sup_{x \in \Gamma_t^{\varepsilon,h,k}} \operatorname{dist}(x, \Gamma_t) \leqslant \delta^*/2 \quad \text{uniformly on } [0,T], \quad \forall \varepsilon \in (0, \widehat{\varepsilon}].$$

Proof. From Theorems 4.12 and 5.1 of [2] we know that there exists an $\widehat{\varepsilon}_1 > 0$ and a constant $C^* > 0$ such that for all $\varepsilon \in (0, \widehat{\varepsilon}_1)$,

$$\|u^\varepsilon - 1\|_{C^0(\mathcal{O} \setminus \mathcal{N}_{\delta^*/2})} \leqslant C^* \varepsilon, \tag{4.15}$$

$$\|u^\varepsilon + 1\|_{C^0(\mathcal{I} \setminus \mathcal{N}_{\delta^*/2})} \leqslant C^* \varepsilon. \tag{4.16}$$

Now for any $x \in \Gamma_t^{\varepsilon,h,k}$, since $U_{\varepsilon,h,k}(x,t) = 0$, from Theorem 3.2(v) we know that there exists an $\widehat{\varepsilon}_2 > 0$, independent of (x,t), such that

$$|u^\varepsilon \pm 1| \geqslant 1 - |u^\varepsilon - U_{\varepsilon,h,k}| \geqslant 2C^*\varepsilon \tag{4.17}$$

for all $\varepsilon \in (0, \widehat{\varepsilon}_2)$. Then (4.15)–(4.17) implies that (x,t) must be in the tubular neighborhood $\mathcal{N}_{\delta^*/2}$ of Γ for all $\varepsilon \in (0, \widehat{\varepsilon})$ with $\widehat{\varepsilon} = \min\{\widehat{\varepsilon}_1, \widehat{\varepsilon}_2\}$. The proof is complete. □

Acknowledgments

The first author would like to thank Nicholas Alikakos for introducing and explaining all the fascinating things related to the Allen–Cahn and Cahn–Hilliard equations to him. He would also like to thank Nicholas Alikakos and Xinfu Chen for answering his questions regarding the spectrum estimate in Lemma 2.5 and the L^∞ estimate in (2.9). The second author gratefully acknowledges financial support by the DFG. Both authors would also like to thank the referees for carefully reading the paper and for their helpful suggestions.

REFERENCES

1. ADAMS, R. A. *Sobolev Spaces*. Academic Press, New York, 1975. Zbl 0314.46030 MR 0450957
2. ALIKAKOS, N. D., BATES, P. W., & CHEN, X. Convergence of the Cahn–Hilliard equation to the Hele–Shaw model. *Arch. Rat. Mech. Anal.* **128** (1994), 165–205. Zbl 0828.35105 MR 1308851
3. ALIKAKOS, N. D. & FUSCO, G. The spectrum of the Cahn–Hilliard operator for generic interface in higher space dimensions. *Indiana Univ. Math. J.* **42** (1993), 637–674. Zbl 0798.35123 MR 1237062
4. ALLEN, S. & CAHN, J. W. A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metall.* **27** (1979), 1084–1095.
5. BARRETT, J. W. & BLOWEY, J. F. An error bound for the finite element approximation of a model for phase separation of a multi-component alloy. *IMA J. Numer. Anal.* **16** (1996), 257–287. Zbl 0849.65069 MR 1382718
6. BARRETT, J. W. & BLOWEY, J. F. Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy. *Numer. Math.* **77** (1997), 1–34. Zbl 0882.65129 MR 1464653
7. BARRETT, J. W., BLOWEY, J. F., & GARCKE, H. On fully practical finite element approximations of degenerate Cahn–Hilliard systems. *M2AN Math. Model. Numer. Anal.* **35** (2001), 713–748. Zbl 0987.35071 MR 1863277
8. BATES, P. W. & FIFE, P. C. The dynamics of nucleation for the Cahn–Hilliard equation. *SIAM J. Appl. Math.* **53** (1993), 990–1008. Zbl 0788.35061 MR 1232163
9. BELLETTINI, G. & PAOLINI, M. Quasi-optimal error estimates for the mean curvature flow with a forcing term. *Differential Integral Equations* **8** (1995), 735–752. Zbl 0820.49019 MR 1306590
10. BLUM, H. & RANNACHER, R. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Methods Appl. Sci.* **2** (1980), 556–581. Zbl 0445.35023 MR 0595625
11. BRENNER, S. C. & SCOTT, L. R. *The Mathematical Theory of Finite Element Methods*. Springer, New York (1994). Zbl 0804.65101 MR 1278258
12. CAHN, J. W. & HILLIARD, J. E. Free energy of a nonuniform system I. Interfacial free energy. *J. Chem. Phys.* **28** (1958), 258–267.
13. CARSTENSEN, C. Merging the Bramble–Pasciak–Steinbach and the Crouzeix–Thomée criterion for H^1-stability of the L^2-projection onto finite element spaces. *Math. Comp.* **71** (2002), 157–163. Zbl 0989.65123 MR 1862993

14. CHEN, X. Generation and propagation of interfaces for reaction-diffusion equations. *J. Differential Equations* **96** (1992), 116–141. Zbl 0765.35024 MR 1153311
15. CHEN, X. Spectrum for the Allen–Cahn, Cahn–Hilliard, and phase-field equations for generic interfaces. *Comm. Partial Differential Equations* **19** (1994), 1371–1395. Zbl 0811.35098 MR 1284813
16. CHEN, X. Global asymptotic limit of solutions of the Cahn–Hilliard equation. *J. Differential Geom.* **44** (1996), 262–311. Zbl 0874.35045 MR 1425577
17. CIARLET, P. G. *The Finite Element Method for Elliptic Problems*. Stud. Math. Appl. 4, North-Holland, Amsterdam (1978). Zbl 0383.65058 MR 0520174
18. COPETTI, M. I. M. & ELLIOTT, C. M. Numerical analysis of the Cahn–Hilliard equation with a logarithmic free energy. *Numer. Math.* **63** (1992), 39–65. Zbl 0762.65074 MR 1182511
19. DU, Q. & NICOLAIDES, R. A. Numerical analysis of a continuum model of phase transition. *SIAM J. Numer. Anal.* **28** (1991), 1310–1322. Zbl 0744.65089 MR 1119272
20. DUPONT, T. Some L^2 error estimates for parabolic Galerkin methods. *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (Baltimore, MD 1972), Academic Press, New York (1972), 491–504. Zbl 0279.65086 MR 0403255
21. ELLIOTT, C. M. & FRENCH, D. A. Numerical studies of the Cahn–Hilliard equation for phase separation. *IMA J. Appl. Math.* **38** (1987), 97–128. Zbl 0632.65113 MR 0983721
22. ELLIOTT, C. M. & FRENCH, D. A. A nonconforming finite-element method for the two-dimensional Cahn–Hilliard equation. *SIAM J. Numer. Anal.* **26** (1989), 884–903. Zbl 0686.65086 MR 1005515
23. ELLIOTT, C. M., FRENCH, D. A., & MILNER, F. A. A second order splitting method for the Cahn–Hilliard equation. *Numer. Math.* **54** (1989), 575–590. Zbl 0668.65097 MR 0978609
24. ELLIOTT, C. M. & SONGMU, Z. On the Cahn–Hilliard equation. *Arch. Rat. Mech. Anal.* **96** (1986), 339–357. Zbl 0624.35048 MR 0855754
25. EVANS, L. C., SONER, H. M., & SOUGANIDIS, P. E. Phase transitions and generalized motion by mean curvature. *Comm. Pure Appl. Math.* **45** (1992), 1097–1123. Zbl 0801.35045 MR 1177477
26. FENG, X. & PROHL, A. Numerical analysis of the Cahn–Hilliard equation and approximation for the Hele–Shaw problem, Part II: error analysis and convergence of the interface. IMA Technical Report #1799, Univ. Minnesota (2001).
27. FENG, X. & PROHL, A. Analysis of a fully discrete finite element method for the phase field model and approximation of its sharp interface limits. *Math. Comp.* **73** (2003), 541–567. Zbl pre02041046 MR 2028419
28. FENG, X. & PROHL, A. Numerical analysis of the Allen–Cahn equation and approximation for mean curvature flows. *Numer. Math.* **94** (2003), 33–65. Zbl 1029.65093 MR 1971212
29. FENG, X. & PROHL, A. Error analysis of a mixed finite element method for the Cahn–Hilliard equation. *Numer. Math.* **99** (2004), 47–84. MR 2101784
30. FRENCH, D. A. & JENSEN, S. Long-time behaviour of arbitrary order continuous time Galerkin schemes for some one-dimensional phase transition problems. *IMA J. Numer. Anal.* **14** (1994), 421–442. Zbl 0806.65132 MR 1283945
31. NOCHETTO, R. H., PAOLINI, M., & VERDI, C. Optimal interface error estimates for the mean curvature flow. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **21** (1994), 193–212. Zbl 0886.35079 MR 1288364
32. NOCHETTO, R. H. & VERDI, C. Convergence past singularities for a fully discrete approximation of curvature-driven interfaces. *SIAM J. Numer. Anal.* **34** (1997), 490–512. Zbl 0876.35053 MR 1442924
33. SCHOLZ, R. A mixed method for 4th order problems using linear finite elements. *RAIRO Anal. Numér.* **12** (1978), 85–90. Zbl 0382.65059 MR 0483557
34. WHEELER, M. F. A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM J. Numer. Anal.* **10** (1973), 723–759. Zbl 0258.35041 MR 0351124