# FULLY DISCRETE DYNAMIC MESH DISCONTINUOUS GALERKIN METHODS FOR THE CAHN-HILLIARD EQUATION OF PHASE TRANSITION 

XIAOBING FENG AND OHANNES A. KARAKASHIAN


#### Abstract

Fully discrete discontinuous Galerkin methods with variable meshes in time are developed for the fourth order Cahn-Hilliard equation arising from phase transition in materials science. The methods are formulated and analyzed in both two and three dimensions, and are proved to give optimal order error bounds. This coupled with the flexibility of the methods demonstrates that the proposed discontinuous Galerkin methods indeed provide an efficient and viable alternative to the mixed finite element methods and nonconforming (plate) finite element methods for solving fourth order partial differential equations.


## 1. Introduction

This paper develops and analyzes discontinuous Galerkin (DG) methods for the the following fourth order parabolic problem:

$$
\begin{align*}
u_{t}+\Delta\left(\varepsilon \Delta u-\frac{1}{\varepsilon} f(u)\right) & =0 & & \text { in } \Omega_{T}:=\Omega \times(0, T),  \tag{1.1}\\
\partial_{n} u=\partial_{n} \Delta u & =0 & & \text { on } \partial \Omega_{T}:=\partial \Omega \times(0, T),  \tag{1.2}\\
u & =u_{0} & & \text { on } \Omega \times\{0\}, \tag{1.3}
\end{align*}
$$

where $f(s)=F^{\prime}(s)$ and $F(s)=\frac{1}{4}\left(s^{2}-1\right)^{2}$ and $\Omega \subset \mathbf{R}^{d}, d=2,3$, is a bounded domain. $\partial_{n}$ denotes the normal derivative operator on $\partial \Omega$.

Equation (1.1) is the well-known Cahn-Hilliard equation; it was originally introduced by Cahn and Hilliard [10 to describe the complicated phase separation and coarsening phenomena in a melted alloy that is quenched to a temperature at which only two different concentration phases can stably exist. The Cahn-Hilliard equation has been widely accepted as a good (conservative) model to describe the phase separation and coarsening phenomena in a melted alloy. We note that equation (1.1) differs from the original Cahn-Hilliard equation (see [10) in the scaling of the time so that $t$ here, called the fast time, represents $\frac{t}{\varepsilon}$ in the original formulation. The function $u$ represents the concentration of one of the two metallic components of the alloy. The parameter $\varepsilon$ is an "interaction length", which is small compared to the characteristic dimensions on the laboratory scale. For the physical background,

[^0]derivation, and discussion of the Cahn-Hilliard equation and related equations, we refer to [2, 10, 19, 22, 32] and the references therein.

Another motivation for developing efficient numerical methods for the CahnHilliard equation is its applications far beyond its original role in phase transition. The Cahn-Hilliard equation is indeed a fundamental equation and an essential building block in the phase field theory for moving interface problems (cf. 32). It is often combined with other fundamental equations of mathematical physics such as the Navier-Stokes equation (cf. [20, 28, 31] and the references therein) to be used as diffuse interface models for describing various interface dynamics, such as flow of two-phase fluids, from various applications. In addition, the Cahn-Hilliard equation (1.1) has also been extensively studied in the past due to its connection to the free boundary problem, known as the Hele-Shaw problem and the MullinsSekerka problem, when $\varepsilon \rightarrow 0^{+}$; see [2, 23] and the references therein for a detailed exposition.

Numerical approximations of the Cahn-Hilliard equation (mainly in one and two dimensions) have been studied by several authors in the past twenty years. Elliott and Zheng [19] analyzed a (continuous in time) semi-discrete conforming finite element discretization in one space dimension. Elliott and French 17 proposed a (continuous in time) semi-discrete nonconforming finite element method based on the Morley nonconforming finite element method [11]. Optimal order error estimates were also established for the nonconforming method. Elliott, French and Milner [18] 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 $w:=-\varepsilon \Delta u+\varepsilon^{-1} f(u)$. Optimal order error estimates were shown under the assumption that the finite element approximation $u_{h}$ of the concentration $u$ is bounded in $L^{\infty}$. Later, Du and Nicolaides [16] 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 $L^{\infty}$. In one space dimension, French and Jensen 25 ] analyzed the long time behavior of the (continuous time) semi-discrete conforming $h p$-finite element approximations. Recently, Feng and Prohl [22, 23] carried out a rigorous study of the convergence of mixed finite element approximations of the Cahn-Hilliard equation to the Hele-Shaw problem as mesh sizes, and the parameter $\varepsilon$ all tend to zero by first establishing polynomial order a priori error estimates in $\frac{1}{\varepsilon}$. Feng and Wu [24] obtained polynomial order a posteriori error estimates in $\frac{1}{\varepsilon}$ for both mixed finite element and $C^{1}$ finite element approximations of the Cahn-Hilliard equation.

Three main difficulties arise in finite element approximations of the Cahn-Hilliard equation. First, no practical 3-d conforming or nonconforming finite elements are known in the literature for fourth order equations. All best known plate elements for the biharmonic operator were designed in 2-d (cf. Chapter 6 of [11), and their generalizations to the 3 -d case is not available and seems not easy, either. Hence, the mixed finite element method becomes the only viable method to solve the Cahn-Hilliard equation in $3-$ d. Second, since the Cahn-Hilliard equation is a singularly perturbed equation, to control the parameter $\varepsilon$, which is either small or approaching zero, both error analysis and numerical simulations are big challenges, and using adaptive meshes becomes necessary, in particular, for 3-d simulations.

Third, no matter which discretization method is used to approximate the CahnHilliard equation, the resulting linear or nonlinear algebraic systems are large and strongly ill-conditioned and hence are difficult to solve either directly or iteratively. Consequently, the development of fast-converging numerical algorithms (including efficient preconditioners) for solving the algebraic systems is crucial to the overall adaptive solution procedure for the Cahn-Hilliard equation and its sharp interface limit, the Hele-Shaw problem.

In this paper we mainly focus on addressing the first difficulty and only slightly touch the second one. Specifically, we shall develop and analyze a family of fully discrete discontinuous Galerkin (DG) methods with dynamic meshes for the CahnHilliard equation. As is now well known, DG methods have several advantages over other types of finite element methods. For example, the trial and test spaces are very easy to construct; they can naturally handle inhomogeneous boundary conditions and curved boundaries; they also allow the use of highly nonuniform and unstructured meshes, and have built-in parallelism which permits coarse-grain parallelization. In addition, the fact that the mass matrices are block diagonal is an attractive feature in the context of time-dependent problems, especially if explicit time discretizations are used. We refer to [3, 4, 5, 6, 7, 12, 13, 15, 27, 34, 36, and the references therein for a detailed account on DG methods.

In addition to the advantages listed above, for fourth order equations, such as the Cahn-Hilliard equation, DG methods have another big advantage over other finite element methods in view of their simplicity and dimension-independence in construction (cf. [6, 21]). As we shall see later in this paper, the formulations of our DG methods for the Cahn-Hilliard problem (1.1)-(1.3) are exactly the same for $d=2$ and $d=3$. In fact, the formulations are exactly the same for all $d \geq 1$. We emphasize that due to the existence of the small scale $\varepsilon$ the use of dynamic/adaptive meshes, which allow local refinement and coarsening, is necessary for simulating the Cahn-Hilliard problem, in particular, in the 3-d case. Although it will not be addressed in this paper, we note that at each time step the parallel Schwarz domain decomposition preconditioners developed by the authors in 21 have an immediate application for solving algebraic systems resulting from the discontinuous Galerkin discretizations of the Cahn-Hilliard problem proposed in this paper.

The paper is organized as follows. In Section 2, notations of this paper are introduced, and the trace inequalities and approximation properties of the interpolation operator are recalled. In Section 3 we consider the (stationary) biharmonic problem with the boundary conditions (1.2). We first present the DG method of Baker 6]. The results of this section serve as building blocks for us to construct and analyze our DG methods for the Cahn-Hilliard problem in the next section. In addition, we present a variant of Baker's method and a Gårding type weak coercivity result for the bilinear forms on the energy space. These results are of independent interest. In Section 4, we propose a family of fully discrete dynamic mesh DG methods for the problem (1.1)-(1.3), whereby the time step size as well as the (spatial) meshes, and consequently the discontinuous finite element spaces, may change at every time level. The time stepping is done using the implicit Euler method, but the schemes as well as the error estimates may be extended, with minor modifications, to second order methods such as the midpoint rule. To alleviate the problem of solving the nonlinear systems of algebraic equations resulting from the $\Delta f(u)$ term, we propose an implicit-explicit variant whereby at each time $t_{m}$, the implicit fully discrete
approximation $U^{m}$ is replaced by a suitable projection $P^{m} U^{m-1}$ in the nonlinear term. As a consequence, only linear systems need be solved at every time step. Given that the convergence rates for this method are the same as those of the fully implicit method, this variant is of great practical interest. We also point out that in contrast to schemes using static spatial meshes, if the mesh is changed at certain times, then the error bounds will contain terms involving jumps of the Riesz projection of the solution $u$. In addition, the total number of such changes also enters into the error bounds. We also give a result that replaces the Riesz projection by the Lagrange interpolant of $u$. This improves on the previous result in the sense that unlike the Riesz projection operator, the interpolant is a local operator. In Section [5] we give a summary and point out a few possible extentions of the work presented in this paper.

## 2. Notation and preliminaries

Throughout this paper, we adopt the standard norm and inner product notation on the $L^{p}$ spaces and the Sobolev spaces $H^{m}$ (cf. [1]). In particular, for a regular domain $D,\|\cdot\|_{D}$ and $(\cdot, \cdot)_{D}$ will denote the norm and inner product on $L^{2}(D)$ (we shall use $(\cdot, \cdot):=(\cdot, \cdot)_{\Omega},\|\cdot\|:=\|\cdot\|_{\Omega}$, and $\|\cdot\|_{m, D}$ will denote the norm on $H^{m}(D)$. Also, $|\cdot|_{m, D}$ will denote the seminorm of derivatives of order $m$. We shall also use $|\cdot|_{\partial D}$ and $\langle\cdot, \cdot\rangle_{\partial D}$ to denote the norm and inner product respectively on $L^{2}(\partial D)$.

Let $\mathcal{T}_{h}=\{K\}$ be a family of star-like partitions (triangulations) of the domain $\Omega$ parametrized by $0<h<1$. We assume that $\mathcal{T}_{h}$ satisfies the following assumptions:
(i) The elements (cells) of $\mathcal{T}_{h}$ satisfy the minimal angle condition.
(ii) $\mathcal{T}_{h}$ is locally quasi-uniform, that is, if two elements $K$ and $K^{\prime}$ are adjacent (i.e., the $(n-1)$-dimensional measure of $\partial K \cap \partial K^{\prime}$ is positive), then $h_{K} \approx$ $h_{K^{\prime}}$, where $h_{K}, h_{K^{\prime}}$ denote the diameters of $K$ and $K^{\prime}$ respectively.
The weak formulations as well as the approximations themselves will involve functions that are discontinuous across interelement boundaries. This motivates the use of so-called "broken" spaces

$$
H^{m}\left(\mathcal{T}_{h}\right)=\Pi_{K \in \mathcal{T}_{h}} H^{m}(K) .
$$

In particular, the "energy space" for fourth order problems will be $E_{h}:=H^{4}\left(\mathcal{T}_{h}\right)$. Note that members of these spaces are not functions in the proper sense since they can be multivalued on the interelement boundaries; so care must be applied in interpreting traces and other related quantities.

Another consequence of the discontinuous nature of the functions is that the edges/faces of the partition $\mathcal{T}_{h}$ play a prominent role in the formulation of the methods as well as their analysis. So we define

$$
\begin{aligned}
\mathcal{E}^{I} & :=\text { set of all interior edges/faces of } \mathcal{T}_{h}, \\
\mathcal{E}^{B} & :=\text { set of all boundary edges/faces of } \mathcal{T}_{h}, \\
\mathcal{E} & :=\mathcal{E}^{I} \cup \mathcal{E}^{B}=\text { set of all edges/faces of } \mathcal{T}_{h} .
\end{aligned}
$$

For $e \in \mathcal{E}^{I}$, we have $e=\partial K^{+} \cap \partial K^{-}$for some $K^{+}, K^{-} \in \mathcal{T}_{h}$. For $v \in E_{h}$ we define the jump $[v]$ of $v$ on $e$ as $\left.[v]\right|_{e}=\left.v^{+}\right|_{e}-\left.v^{-}\right|_{e}$, where $v^{+}$and $v^{-}$denote the restrictions of $v$ to $K^{+}$and $K^{-}$respectively. For $e \in \mathcal{E}^{B}$, we set $\left.[v]\right|_{e}=\left.v\right|_{e}$. For $e \in \mathcal{E}, h_{e}$ will denote the length of $e$ for $n=2$, or the diameter of $e$ for $n \geq 3$. It follows from the local quasiuniformity assumption that $h_{e} \approx h_{K^{+}} \approx h_{K^{-}}$. This fact is used repeatedly in this paper. For $e \in \mathcal{E}^{I}$ we define the average of $v$ on $e$
to be $\left.\{v\}\right|_{e}:=\frac{1}{2}\left(\left.v^{+}\right|_{e}+\left.v^{-}\right|_{e}\right)$. If $e \in \mathcal{E}^{B}$, set $\left.\{v\}\right|_{e}=\left.v\right|_{e}$. In [6] $\{v\}$ was set to $\left.v^{+}\right|_{e}$ for $e \in \mathcal{E}$. We remark that the results of this paper cover both of the above conventions. We also let $\partial_{n}$ denote the normal derivative operator in the direction outward from $K^{+}$.

The following trace inequality is well known
Lemma 2.1. There exists a positive constant $C$, which is independent of $h$, such that for any $K \in \mathcal{T}_{h}$, if $\phi \in H^{1}(K)$, then

$$
\begin{equation*}
|\phi|_{\partial K}^{2} \leq C\left(h_{K}^{-1}\|\phi\|_{K}^{2}+h_{K}\|\nabla \phi\|_{K}^{2}\right), \tag{2.1}
\end{equation*}
$$

where $h_{K}$ is the diameter of $K$.
For any $K \in \mathcal{T}_{h}$ and integer $r \geq 0$, let $P_{r-1}(K)$ denote the set of all polynomials of degree less than or equal to $r-1$ on $K$ (we let $P_{-1}=\{0\}$ ). The (discontinuous) finite element space $V_{h}$ is defined by

$$
V_{h}:=\prod_{K \in \mathcal{T}_{h}} P_{r-1}(K)
$$

Clearly, $V_{h} \subset E_{h} \subset L^{2}(\Omega)$. But $V_{h} \not \subset H^{2}(\Omega)$. In fact, $V_{h} \not \subset H^{1}(\Omega)$.
We shall make frequent use of so-called inverse inequalities that hold on spaces of polynomial functions.
Lemma 2.2. There exists a constant $c$ depending only on the minimum angle of $K$ and $r$ such that

$$
\begin{equation*}
\|\chi\|_{j, K} \leq c h_{K}^{-j}\|\chi\|_{K} \quad \forall \chi \in P_{r-1}(K), j=1, \ldots, r-1 . \tag{2.2}
\end{equation*}
$$

An immediate consequence of the trace and the inverse inequalities for polynomials are the following trace inequalities (cf. [3, 21]). For $e=\mathcal{E}^{I}$ and $v \in V_{h}$ there hold

$$
\begin{align*}
|\{v\}|_{e}^{2} & \leq C h_{e}^{-1}\left(\|v\|_{K^{+}}^{2}+\|v\|_{K^{-}}^{2}\right)  \tag{2.3}\\
\left|\left\{\partial_{n} v\right\}\right|_{e}^{2} & \leq C h_{e}^{-3}\left(\|v\|_{K^{+}}^{2}+\|v\|_{K^{-}}^{2}\right) \tag{2.4}
\end{align*}
$$

For $e \in \mathcal{E}^{B}$, the above inequalities hold without $K^{-}$.
The spaces $V_{h}$ have good approximation properties due to the fact that the approximations can be localized to individual elements. From a result of ScottDupont (cf. [9] and also [7]) we have
Lemma 2.3. For $K \in \mathcal{T}_{h}$ let $\phi \in H^{s}(K), s \geq 0$. Then for each $r$ with $0 \leq r \leq s$, there exists $\chi \in P_{r-1}(K)$ such that

$$
\begin{equation*}
|\phi-\chi|_{j, K} \leq C h_{K}^{r-j}|\phi|_{r, K}, \quad 0 \leq j \leq r \tag{2.5}
\end{equation*}
$$

where $C$ is independent of $h_{K}, \phi, r$.
We will also make use of the usual nodal based Lagrangian interpolation operators $I_{K}: C(\bar{K}) \rightarrow P_{r-1}(K)$. The approximation properties of this operator are well known and can be found in [11.
Lemma 2.4. For $K \in \mathcal{T}_{h}$, let $\phi \in H^{s}(K) \cap C(\bar{K})$, with $2 \leq r \leq s$. Then

$$
\begin{equation*}
\left|\phi-I_{K} \phi\right|_{j, K} \leq C h_{K}^{r-j}|\phi|_{r, K}, \quad 0 \leq j \leq r \tag{2.6}
\end{equation*}
$$

Furthermore, if $\phi \in W^{2, \infty}$, then

$$
\begin{equation*}
\left|\phi-I_{K} \phi\right|_{L^{\infty}(K)} \leq c h_{K}^{2}|\phi|_{W^{2, \infty}(K)} . \tag{2.7}
\end{equation*}
$$

## 3. DG methods for the biharmonic equation

In this section, we shall consider discontinuous Galerkin approximations of the biharmonic problem

$$
\begin{array}{ll}
\Delta^{2} u=g & \text { in } \Omega \\
\partial_{n} u=\partial_{n} \Delta u=0 & \text { on } \partial \Omega \tag{3.2}
\end{array}
$$

The discontinuous Galerkin method considered in this paper for discretizing problem (1.1)-(1.3) is related to one proposed in (6). We emphasize that the discontinuous Galerkin method and the results of this paper are valid for both $d=2$ and $d=3$.

The results of this section, which are of independent interest and appear in the name of elliptic projections in the next section, will serve as a basis for us to analyze our fully discrete DG methods for the Cahn-Hilliard problem (1.1)- (1.3).

First, we recall the family of DG methods developed by Baker in 6 (also see [21), and then quote the main properties of the DG methods. Second, we shall propose a variant of Baker's methods and briefly analyze this new family of DG methods.

In [6] G. Baker constructed and analyzed DG formulations for (3.1) with homogeneous Dirichlet conditions $u=\partial_{n} u=0$ on $\partial \Omega$ using as template the formula

$$
\begin{equation*}
\int_{K} \Delta^{2} u v d x=\int_{K} \Delta u \Delta v d x+\int_{\partial K}\left(\partial_{n_{K}} \Delta u\right) v d \sigma-\int_{\partial K} \Delta u\left(\partial_{n_{K}} v\right) d \sigma \tag{3.3}
\end{equation*}
$$

Since the boundary conditions considered in this paper are different from those in [6], our bilinear form was modified to

$$
\begin{gather*}
\widehat{b}_{h}(u, v)=\sum_{K \in \mathcal{T}_{h}}(\Delta u, \Delta v)_{K}+\sum_{e \in \mathcal{E}^{I}}\left(\left\langle\left\{\partial_{n} \Delta u\right\},[v]\right\rangle_{e}+\left\langle\left\{\partial_{n} \Delta v\right\},[u]\right\rangle_{e}\right. \\
\left.+\gamma h_{e}^{-3}\langle[u],[v]\rangle_{e}\right)-\sum_{e \in \mathcal{E}}\left(\left\langle\{\Delta u\},\left[\partial_{n} v\right]\right\rangle_{e}+\left\langle\{\Delta v\},\left[\partial_{n} u\right]\right\rangle_{e}\right.  \tag{3.4}\\
\left.-\gamma h_{e}^{-1}\left\langle\left[\partial_{n} u\right],\left[\partial_{n} v\right]\right\rangle_{e}\right)
\end{gather*}
$$

Here $(\cdot, \cdot)_{K}$ denotes the $L^{2}$ integral over $K ;\langle\cdot, \cdot\rangle_{e}$ stands for the $L^{2}$ integral over the edge $e ; \gamma$ is a positive constant independent of $h$, and the terms including $\gamma$ are the so-called penalty terms. The difference between our formulation and Baker's is that here the first sum on edges does not include the boundary edges for the simple reason that boundary values of the solution $u$ are not known on $\partial \Omega$. The absence of the boundary edges has the interesting consequence that our bilinear form is no longer coercive on the finite element spaces. This is just as it should be since the solution to (3.1), (3.2) is unique up to an additive constant.

With the bilinear form $\widehat{b}_{h}(\cdot, \cdot)$ we naturally associate the following seminorm on the space $E_{h}$ :

$$
\begin{gather*}
\|v\|_{2, h, \widehat{b}}=\left(\sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2}+\sum_{e \in \mathcal{E}^{I}}\left(h_{e}^{3}\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2}+h_{e}^{-3}|[v]|_{e}^{2}\right)\right.  \tag{3.5}\\
\left.+\sum_{e \in \mathcal{E}}\left(h_{e}|\{\Delta v\}|_{e}^{2}+h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}\right)\right)^{1 / 2}
\end{gather*}
$$

Before stating our next result, we recall the definition of a quotient space $X / R$ of a Banach $\left(X,\|\cdot\|_{X}\right) . X / R$ is the quotient space of equivalence classes of functions in $X$ that differ by constants. $X / R$ is a Banach space with norm $\inf _{c \in R}\|x-c\|_{X}$.

Lemma 3.1. (i) $\|\cdot\|_{2, h, \widehat{b}}$ is a norm on the quotient space $E_{h} / R$.
(ii)

$$
\begin{equation*}
\left|\widehat{b}_{h}(u, v)\right| \leq(1+\gamma)\|u\|_{2, h, \widehat{b}}\|v\|_{2, h, \widehat{b}} \quad \forall u, v \in E_{h} \tag{3.6}
\end{equation*}
$$

(iii) There exist positive constants $\gamma_{0}$ and $c_{0}$ such that for $\gamma \geq \gamma_{0}$

$$
\begin{equation*}
\widehat{b}_{h}(v, v) \geq c_{0}\|v\|_{2, h, \widehat{b}}^{2} \quad \forall v \in V_{h} \tag{3.7}
\end{equation*}
$$

Proof. The proof of (ii) is a simple application of the Cauchy-Schwarz inequality. The proof of (iii) follows the outline of a similar proof in [6]. Essential use is made of the trace and inverse inequalities. As for $(i)$, it suffices to show that $\|v\|_{2, h, \widehat{b}}$ can vanish only if $v$ is constant on $\Omega$. So, integrating by parts, we easily get the well-known identity

$$
\begin{equation*}
\sum_{K \in \mathcal{T}_{h}}\|\nabla v\|_{K}^{2}=-\sum_{K \in \mathcal{T}_{h}}(\Delta v, v)_{K}+\sum_{e \in \mathcal{E}}\left(\left\langle\left\{\partial_{n} v\right\},[v]\right\rangle_{e}+\left\langle\{v\},\left[\partial_{n} v\right]\right\rangle_{e}\right) \tag{3.8}
\end{equation*}
$$

It is easy to see that the vanishing of $\|v\|_{2, h, \widehat{b}}$ implies the vanishing of the right hand side above. Thus $v$ must be piecewise constant on $\mathcal{T}_{h}$. On the other hand, since the jumps of $v$ on the interior edges are also zero, it follows that $v$ is constant.

The weak formulation of (3.1)-(3.2) can be phrased as seeking $u \in E_{h}$ such that

$$
\begin{equation*}
\widehat{b}_{h}(u, v)=(g, v) \quad \forall v \in E_{h} \tag{3.9}
\end{equation*}
$$

This formulation is indeed consistent with the boundary value problem (3.1)-(3.2)
Based on the weak formulation (3.9), we define the DG formulation as follows: find $u_{h} \in V_{h}$ such that

$$
\begin{equation*}
\widehat{b}_{h}\left(u_{h}, v_{h}\right)=\left(g, v_{h}\right) \quad \forall v_{h} \in V_{h} \tag{3.10}
\end{equation*}
$$

3.1. A variant of Baker's DG methods. In contrast to DG methods for secondorder elliptic problems where $\|\nabla v\|$ is a seminorm on $H^{1}$, the term $\|\Delta v\|$ is not a seminorm on $H^{2}$. This creates additional technical problems not all of which are resolved. Interestingly, the following integration by parts formula

$$
\int_{K} \Delta u \Delta v d x=\int_{K} D^{2} u \cdot D^{2} v d x-\int_{\partial K}\left(\partial_{n_{K}} \nabla u\right) \cdot \nabla v d \sigma+\int_{\partial K} \Delta u\left(\partial_{n_{K}} v\right) d \sigma
$$

where $D^{2} v$ denotes the Hessian of $v$, can be used to rewrite (3.3) as

$$
\begin{align*}
\int_{K} \Delta^{2} u v d x= & \int_{K} D^{2} u \cdot D^{2} v d x+\int_{\partial K}\left(\partial_{n_{K}} \Delta u\right) v d \sigma  \tag{3.11}\\
& -\int_{\partial K}\left(\partial_{n_{K}} \nabla u\right) \cdot \nabla v d \sigma
\end{align*}
$$

Based on (3.11) we propose the following variant of Baker's DG method:

$$
\begin{equation*}
\widetilde{b}_{h}\left(u_{h}, v_{h}\right)=\left(g, v_{h}\right) \quad \forall v_{h} \in V_{h} \tag{3.12}
\end{equation*}
$$

where

$$
\begin{align*}
& \widetilde{b}_{h}(u, v)= \sum_{K \in \mathcal{T}_{h}}\left(D^{2} u, D^{2} v\right)_{K}+\sum_{e \in \mathcal{E}^{I}}\left(\left\langle\left\{\partial_{n} \Delta u\right\},[v]\right\rangle_{e}+\left\langle\left\{\partial_{n} \Delta v\right\},[u]\right\rangle_{e}\right. \\
&\left.+\gamma h_{e}^{-3}\langle[u],[v]\rangle_{e}\right)-\sum_{e \in \mathcal{E}}\left(\left\langle\left\{\partial_{n} \nabla u\right\},[\nabla v]\right\rangle_{e}+\left\langle\left\{\partial_{n} \nabla v\right\},[\nabla u]\right\rangle_{e}\right.  \tag{3.13}\\
&\left.+\gamma h_{e}^{-1}\langle[\nabla u],[\nabla v]\rangle_{e}\right)
\end{align*}
$$

With the new bilinear form $\widetilde{b}_{h}(\cdot, \cdot)$ we associate the following mesh-dependent seminorm on the energy space $E_{h}$ :

$$
\begin{align*}
\|v\|_{2, h, \tilde{b}}= & \left(\sum_{K \in \mathcal{T}_{h}}\left\|D^{2} v\right\|_{K}^{2}+\sum_{e \in \mathcal{E}^{I}}\left(h_{e}^{3}\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2}+h_{e}^{-3}|[v]|_{e}^{2}\right)\right. \\
& \left.+\sum_{e \in \mathcal{E}}\left(h_{e}|\{\Delta v\}|_{e}^{2}+h_{e}\left|\left\{\partial_{n} \nabla v\right\}\right|_{e}^{2}+h_{e}^{-1}|[\nabla v]|_{e}^{2}\right)\right)^{\frac{1}{2}} \tag{3.14}
\end{align*}
$$

It is not hard to check that (3.9) can be reformulated as seeking $u \in E_{h}$ satisfying

$$
\begin{equation*}
\widetilde{b}_{h}(u, v)=(g, v) \quad \forall v \in E_{h} \tag{3.15}
\end{equation*}
$$

which is also consistent with the boundary value problem (3.1)-(3.2).
The results of Lemma 3.1 apply to the bilinear form $\widetilde{b}(\cdot, \cdot)$ and the associated seminorm $\|\cdot\|_{2, h, \widetilde{b}}$ as well. Therefore, we introduce the corresponding DG approximation $u_{h} \in V_{h}$ by

$$
\begin{equation*}
\widetilde{b}_{h}\left(u_{h}, v_{h}\right)=\left(g, v_{h}\right) \quad \forall v_{h} \in V_{h} \tag{3.16}
\end{equation*}
$$

Given the common properties of the two methods introduced in this section and the fact that the error estimates turn out to be similar, we will henceforth, unless explicitly indicated otherwise, use the common symbol $b(\cdot, \cdot)$ to stand for either of the two bilinear forms (3.4) and (3.13) and $\|\cdot\|_{2, h}$ to denote either of the two seminorms $\|\cdot\|_{2, h, \widehat{b}},\|\cdot\|_{2, h, \widetilde{b}}$ respectively.

We introduce the negative norm $\|\psi\|_{-2, h}:=\sup \frac{(\psi, v)}{\|v\|_{2, h}}$, the supremum being taken over all nonconstant $v$ in $H^{4}\left(\mathcal{T}_{h}\right)$, and the space $H^{-2}\left(\mathcal{T}_{h}\right)$ of all measurable functions with finite $\|\cdot\|_{-2, h}$ norm.
3.2. A priori error estimates. In [6] Baker obtained optimal a priori error estimates for his method in the energy norm as well as negative norms under the assumption that $u \in H^{s}(\Omega), s \geq 4$ and $r \geq 4$. Estimates for the case $r=3$ can also be obtained except that the rate for the $L^{2}$-norm of the error is suboptimal. We have obtained similar results for both of our formulations for the BVP (3.1), (3.2). Theorem 3.1 summarizes these results. The proof follows the same lines as those found in [6] and is omitted. The basic approach is classical: First, estimates in the energy norm $\|\cdot\|_{2, h}$ are obtained. Then, Nitsche's duality argument can be employed to derive $L^{2}$ as well as negative norm estimates.

Since our estimates encompass variable meshes in space, we opt to cast the energy estimates in terms of local quantities. Indeed, it follows easily from the approximation properties (2.5) that for $v \in H^{s}\left(\mathcal{T}_{h}\right), s \geq 4$,

$$
\|\phi-\chi\|_{2, h} \leq \begin{cases}C\left|h^{r-2} \phi\right|_{H^{r}\left(\mathcal{T}_{h}\right)}, & 4 \leq r \leq s  \tag{3.17}\\ C|h \phi|_{H^{3}\left(\mathcal{T}_{h}\right)}+\left|h^{2} \phi\right|_{H^{4}\left(\mathcal{T}_{h}\right)}, & r=3\end{cases}
$$

where $\left|h^{j} v\right|_{H^{\ell}\left(\mathcal{T}_{h}\right)}:=\left(\sum_{K \in \mathcal{T}_{h}} h_{K}^{2 j}|v|_{\ell, K}^{2}\right)^{1 / 2}$.
Theorem 3.1. Assume that the solution $u$ of the $B V P$ (3.1)-(3.2) is in $H^{3}(\Omega) \cap$ $H^{s}\left(\mathcal{T}_{h}\right), s \geq 4$, and let $u_{h} \in V_{h}$ be given by (3.10) or (3.12). Then,
(i) For $4 \leq r \leq s$, there holds

$$
\begin{equation*}
\left\|u-u_{h}\right\|_{2, h} \leq c\left|h^{r-2} u\right|_{H^{r}\left(\mathcal{T}_{h}\right)} . \tag{3.18}
\end{equation*}
$$

If in addition $\int_{\Omega} u_{h} d x=\int_{\Omega} u d x$, then

$$
\left\|u-u_{h}\right\| \leq c h^{2}\left|h^{r-2} u\right|_{H^{r}\left(\mathcal{T}_{h}\right)} .
$$

For $K \in \mathcal{T}_{h}$ and multi-index $\alpha, 1 \leq|\alpha| \leq r$, we have

$$
\begin{equation*}
\left\|D^{\alpha}\left(u-u_{h}\right)\right\|_{K} \leq c h_{K}^{r-|\alpha|}|u|_{H^{r}(K)}+c h_{K}^{-|\alpha|}\left\|u-u_{h}\right\|_{K} \tag{3.20}
\end{equation*}
$$

(ii) Similarly, for $r=3$, there holds

$$
\begin{equation*}
\left\|u-u_{h}\right\|_{2, h} \leq c|h u|_{H^{3}\left(\mathcal{T}_{h}\right)}+c\left|h^{2} u\right|_{H^{4}\left(\mathcal{T}_{h}\right)} . \tag{3.21}
\end{equation*}
$$

If in addition $\int_{\Omega} u_{h} d x=\int_{\Omega} u d x$, then

$$
\begin{equation*}
\left\|u-u_{h}\right\| \leq \operatorname{ch}\left(|h u|_{H^{3}\left(\mathcal{T}_{h}\right)}+\left|h^{2} u\right|_{H^{4}\left(\mathcal{T}_{h}\right)}\right) \tag{3.22}
\end{equation*}
$$

For $K \in \mathcal{T}_{h}$ and multi-index $\alpha, 1 \leq|\alpha| \leq 3$, we have

$$
\left\|D^{\alpha}\left(u-u_{h}\right)\right\|_{K} \leq c h_{K}^{3-|\alpha|}|u|_{H^{r}(K)}+c h_{K}^{-|\alpha|}\left\|u-u_{h}\right\|_{K}
$$

Remark 3.1. We first note the sub-optimality of the $L^{2}$ estimate (3.22). Also, the requirement of $\int_{\Omega} u_{h} d x=\int_{\Omega} u d x$ in the estimates (3.19) and (3.22) is due to the fact that these are obtained via Nitsche's trick; this involves using the error $u-u_{h}$ as the forcing term $g$ in the BVP (3.1), (3.2). This in turn requires the compatibility condition $\int_{\Omega}\left(u-u_{h}\right) d x=0$ which can be imposed since $u$ and $u_{h}$ are determined modulo additive constants anyway.

We next revisit the issue of coercivity of the bilinear form $b_{h}(\cdot, \cdot)$. Lemma 3.1(iii) establishes coercivity on $V_{h}$ under the condition that the penalty parameter $\gamma$ is larger than a threshold value $\gamma_{0}$ which depends on the minimum angle of the elements and quadratically on $r$ through the trace inequalities and inverse estimates used in "hiding" the terms $\left\{\partial_{n} \Delta v\right\}$ and other similar terms. The dependence of $\gamma_{0}$ on $r$ indicates that $b_{h}(\cdot, \cdot)$ cannot be coercive on the energy space $E_{h}$, since the latter can be thought of as being the limiting case of $V_{h}$ as $r \rightarrow \infty$. We next establish a weaker result than full coercivity on $E_{h}$ which is reminiscent of Gårding's inequality and which proves useful.

Theorem 3.2. Let $3 \leq r \leq s$ with $s \geq 4$. Then there exist constants $c_{0}, c_{1}, \gamma_{0}$ with $\gamma_{0}$ depending on $r$ and the minimum angles of the elements such that for $\gamma \geq \gamma_{0}$ there holds

$$
\begin{equation*}
b_{h}(v, v) \geq c_{0}\|v\|_{2, h}^{2}-c_{1}\left|h^{r} v\right|_{H^{r}\left(\mathcal{T}_{h}\right)}^{2} \quad \forall v \in H^{s}\left(\mathcal{T}_{h}\right) \tag{3.24}
\end{equation*}
$$

Proof. We will consider only the case of $\widehat{b}(\cdot, \cdot)$, that of $\widetilde{b}(\cdot, \cdot)$ being entirely similar. We have

$$
\begin{aligned}
& \widehat{b}(v, v)=\sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2}+\sum_{e \in \mathcal{E}^{I}}\left(2\left\langle\left\{\partial_{n} \Delta v\right\},[v]\right\rangle_{e}+\gamma h_{e}^{-3}|[v]|_{e}^{2}\right) \\
&+\sum_{e \in \mathcal{E}}\left(-2\left\langle\{\Delta v\},\left[\partial_{n} v\right]\right\rangle_{e}+\gamma h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}\right)
\end{aligned}
$$

For a constant $c_{0}$ to be chosen later, we have from Cauchy-Schwarz and the arithmetic geometric mean inequality

$$
\begin{align*}
& \widehat{b}(v, v)-c_{0}\|v\|_{2, h, \widehat{b}}^{2}=\left(1-c_{0}\right) \sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2} \\
& \quad+\sum_{e \in \mathcal{E}^{I}}\left(2\left\langle\left\{\partial_{n} \Delta v\right\},[v]\right\rangle_{e}-c_{0} h_{e}^{3}\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2}+\left(\gamma-c_{0}\right) h_{e}^{-3}|[v]|_{e}^{2}\right) \\
& \quad+\sum_{e \in \mathcal{E}}\left(-2\left\langle\{\Delta v\},\left[\partial_{n} v\right]\right\rangle_{e}-c_{0} h_{e}|\{\Delta v\}|_{e}^{2}+\left(\gamma-c_{0}\right) h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}\right) \\
& \geq  \tag{3.25}\\
& \quad\left(1-c_{0}\right) \sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2}+\left(\gamma-c_{0}-\frac{1}{\epsilon}\right) \sum_{e \in \mathcal{E}^{I}} h_{e}^{-3}|[v]|_{e}^{2} \\
& \quad+\left(\gamma-c_{0}-\frac{1}{\epsilon}\right) \sum_{e \in \mathcal{E}} h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}-\left(c_{0}+\epsilon\right) \sum_{e \in \mathcal{E}^{I}} h_{e}^{3}\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2} \\
& \quad-\left(c_{0}+\epsilon\right) \sum_{e \in \mathcal{E}} h_{e}|\{\Delta v\}|_{e}^{2} \quad \text { for any } \epsilon>0 .
\end{align*}
$$

Now using the trace inequality, for any $\chi \in V_{h}$ we have

$$
\begin{aligned}
\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2} \leq & 2\left|\left\{\partial_{n} \Delta(v-\chi)\right\}\right|_{e}^{2}+2\left|\left\{\partial_{n} \Delta \chi\right\}\right|_{e}^{2} \\
\leq & c \sum_{K=K^{+}, K^{-}} \\
& \left.\quad+h_{K}^{-3}\|\Delta \chi\|_{K}^{2}\right) \\
\leq & c \sum_{K=K^{+}, K^{-}} \\
& \left(h_{K}^{-1}|\Delta(v-\chi)|_{1, K}^{2}+h_{K}|\Delta(v-\chi)|_{2, K}^{2}\right. \\
& \left.\quad+h_{K}^{-3}\|\Delta v\|_{K}^{2}+h_{K}^{-3}\|\Delta(v-\chi)\|_{K}^{2}\right),
\end{aligned}
$$

where we have also used the inverse inequality on the $\Delta \chi$ terms. We now choose $\chi$ so that by the approximation properties (2.5)

$$
\begin{equation*}
h_{e}^{3}\left|\left\{\partial_{n} \Delta v\right\}\right|_{e}^{2} \leq c \sum_{K=K^{+}, K^{-}}\left(h_{K}^{2 r}|v|_{r, K}^{2}+\|\Delta v\|_{K}^{2}\right) \tag{3.26}
\end{equation*}
$$

In an entirely similar way we obtain

$$
\begin{equation*}
h_{e}|\{\Delta v\}|_{e}^{2} \leq c \sum_{K=K^{+}, K^{-}}\left(h_{K}^{2 r}|v|_{r, K}^{2}+\|\Delta v\|_{K}^{2}\right) . \tag{3.27}
\end{equation*}
$$

Now using (3.26) and (3.27) in (3.25), we obtain

$$
\begin{aligned}
& \widehat{b}(v, v)-c_{0}\|v\|_{2, h, \widehat{b}}^{2} \\
& \geq\left(1-c_{0}-c\left(c_{0}+\epsilon\right)\right) \sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2}+\left(\gamma-c_{0}-\frac{1}{\epsilon}\right) \sum_{e \in \mathcal{E}^{I}} h_{e}^{-3}|[v]|_{e}^{2} \\
&+\left(\gamma-c_{0}-\frac{1}{\epsilon}\right) \sum_{e \in \mathcal{E}} h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}-c\left(c_{0}+\epsilon\right)\left|h^{r} u\right|_{H^{r}\left(\mathcal{T}_{h}\right)}^{2} .
\end{aligned}
$$

We now choose $c_{0}=\epsilon$ small enough so that $1-c_{0}-c\left(c_{0}+\epsilon\right) \geq 0$. We then choose $\gamma_{0}$ such that $\gamma_{0}-c_{0}-\frac{1}{\epsilon} \geq 0$. This implies (3.24) with $c_{1}=c\left(c_{0}+\epsilon\right)$.

The next lemma, which is Lemma 3.1 of [21], establishes an interpolation result which bounds the piecewise $H^{1}$-seminorm in terms of the $L^{2}$-norm and the $\|\cdot\|_{2, h^{-}}$ norm for totally discontinuous functions.

Lemma 3.2. There exists a constant $C>0$, which is independent of $h$, such that for any $v \in V_{h}$ and any $\delta>0$

$$
\begin{align*}
\sum_{K \in \mathcal{T}_{h}}\|\nabla v\|_{K}^{2} & \leq C\left(\delta^{-1}\|v\|^{2}+\delta|v|_{2, h}^{2}\right)+\sum_{e \in \mathcal{E}^{B}}\left\langle\partial_{n} v, v\right\rangle_{e}  \tag{3.28}\\
\sum_{K \in \mathcal{T}_{h}}\|\nabla v\|_{K}^{2} & \leq C\left(\delta^{-1}\|v\|^{2}+\delta\|v\|_{2, h}^{2}\right) \tag{3.29}
\end{align*}
$$

Here

$$
|v|_{2, h}^{2}=\sum_{K \in \mathcal{T}_{h}}\|\Delta v\|_{K}^{2}+\sum_{e \in \mathcal{E}^{I}}\left(h_{e}^{-3}|[v]|_{e}^{2}+h_{e}^{-1}\left|\left[\partial_{n} v\right]\right|_{e}^{2}\right) .
$$

## 4. Fully discrete dynamic mesh DG methods for the Cahn-Hilliard equation

4.1. Formulation of fully discrete dynamic mesh DG methods. Let $J_{m}:=$ $\left(t_{m-1}, t_{m}\right], m=1, \ldots, M$ be a partition of $[0, T]$ and $\tau_{m}:=t_{m}-t_{m-1}$. For each $J_{m}, m=1, \ldots, M$, let $\mathcal{T}_{h}^{m}$ be a partition of $\Omega$ as defined in Section 2 and let $V_{h}^{m}$ denote the finite element space associated with the partition $\mathcal{T}_{h}^{m}$. We set $V_{h}^{0}=V_{h}^{1}$.

At certain times $t_{m}$ the spatial mesh may be changed via a process of refinement and coarsening based on information supplied by an a posteriori error estimator. Both the algorithmic implementation and the error estimation were seen to benefit from the imposition of the following mild conditions designed to govern the process $\mathcal{T}_{h}{ }^{m-1} \rightarrow \mathcal{T}_{h}{ }^{m}:$
(M1) A cell (the father) in $\mathcal{T}_{h}^{m-1}$ marked for refinement is cut into a number of cells (the sons). In 2-d, a triangle may be subdivided into four similar triangles.
(M2) A cell in $\mathcal{T}_{h}^{m-1}$ marked for coarsening is removed from the mesh only if the remaining sons of its father are all marked for coarsening. Then all sons are removed from the mesh.
Supposing that a new mesh $\mathcal{T}_{h}^{m}$ has been obtained from $\mathcal{T}_{h}^{m-1}$ by the process of refinement/coarsening described above, we shall need an operator that serves as a natural injection operator from spaces defined on $\mathcal{T}_{h}^{m-1}$ to those defined on $\mathcal{T}_{h}^{m}$. We define $P^{m}: L^{2}\left(\mathcal{T}_{h}^{m-1}\right) \rightarrow L^{2}\left(\mathcal{T}_{h}^{m}\right)$ as follows: Let $v \in L^{2}\left(\mathcal{T}_{h}^{m-1}\right)$ and let $K \in \mathcal{T}_{h}^{m}$. Then the restriction $\left.P^{m} v\right|_{K}$ of $P^{m} v$ to $K$ is given by:
(1) If $K$ also belongs to $\mathcal{T}_{h}^{m-1}$ or if $K$ is the son of an element in $\mathcal{T}_{h}^{m-1}$, then $\left.P^{m} v\right|_{K}=\left.v\right|_{K}$.
(2) If $K$ is obtained by the merger of its sons that belonged to $\mathcal{T}_{h}^{m-1}$, then $\left.P^{m} v\right|_{K}$ is the $L^{2}$ projection of $\left.v\right|_{K}$ into $P_{r-1}(K)$.
We first note that $P^{m}$ is defined as a local operator, and we let $P_{K}^{m}$ denote the restriction of $P^{m}$ to $K$. Moreover, $P^{m}$ is the identity operator except on the part of $\mathcal{T}_{h}^{m}$ which has been obtained from $\mathcal{T}_{h}^{m-1}$ by coarsening.

The operator $P^{m}$ has good approximation properties. Indeed, it follows from Lemma 2.3 and the polynomial inverse inequalities that

$$
\left|\phi-P_{K}^{m} \phi\right|_{j, K} \leq c h_{K}^{r-j}|\phi|_{r, K} \quad \forall \phi \in H^{s}\left(\mathcal{T}_{h}^{m}\right), s \geq 0, \quad 0 \leq j \leq r \leq s
$$

Then, our fully discrete dynamic mesh DG methods for (1.1)-(1.3) are defined as follows:

Fully implicit scheme. For each $m=1, \ldots, M$, given $U^{m-1} \in V_{h}^{m-1}$, find $U^{m} \in$ $V_{h}^{m}$ such that

$$
\begin{equation*}
\left(d_{t} U^{m}, v_{h}\right)+\varepsilon b_{h}^{m}\left(U^{m}, v_{h}\right)+\frac{1}{\varepsilon} c_{h}^{m}\left(U^{m}, v_{h}\right)=0 \quad \forall v_{h} \in V_{h}^{m} \tag{4.1}
\end{equation*}
$$

with some starting value $U^{0} \in V_{h}^{0}$. Here $d_{t} U^{m}:=\left(U^{m}-U^{m-1}\right) / \tau_{m}$, and $b_{h}^{m}(\cdot, \cdot)$ represents the mesh-dependent bilinear form $\widehat{b}_{h}(\cdot, \cdot)$ or $\widetilde{b}_{h}(\cdot, \cdot)$ defined on the partition $\mathcal{T}_{h}^{m}$, and the mesh-dependent nonlinear-linear form $c_{h}^{m}(\cdot, \cdot)$ is defined by

$$
\begin{align*}
c_{h}^{m}(u, v)=-\sum_{K \in \mathcal{T}_{h}^{m}}(f(u), \Delta v)_{K} & +\sum_{e \in \mathcal{E}_{m}}\left\langle f(\{u\}),\left[\partial_{n} v\right]\right\rangle_{e}  \tag{4.2}\\
& -\sum_{e \in \mathcal{E}_{m}^{I}}\left\langle f^{\prime}(\{u\})\left\{\partial_{n} u\right\},[v]\right\rangle_{e}
\end{align*}
$$

where $\mathcal{E}_{m}^{I}, \mathcal{E}_{m}$ denote respectively the interior and total edge sets of the partition $\mathcal{T}_{h}{ }^{m}$.

## Implicit-Explicit scheme.

$$
\begin{equation*}
\left(d_{t} U^{m}, v_{h}\right)+\varepsilon b_{h}^{m}\left(U^{m}, v_{h}\right)+\frac{1}{\varepsilon} c_{h}^{m}\left(P^{m} U^{m-1}, v_{h}\right)=0 \quad \forall v_{h} \in V_{h}^{m} \tag{4.3}
\end{equation*}
$$

Remark 4.1. (a) The existence of solutions of the fully implicit scheme can be established by using the standard fixed point argument in finite dimensional spaces under the condition that $\tau_{m}$ is sufficiently small. As for the implicit-explicit scheme, existence follows without such conditions given that $U^{m}$ is the solution of a linear system with a positive definite matrix.
(b) We emphasize that the above fully discrete DG methods are consistent methods since $\left(\Delta^{2} u, v\right)=b_{h}^{m}(u, v)$ for any $v \in V_{h}^{m}$ and $u \in H^{4}(\Omega)$ satisfying the boundary conditions in (1.2), and $-(\Delta f(u), v)=c_{h}^{m}(u, v)$ for any $v \in V_{h}^{m}$ and $u \in H^{2}(\Omega)$ satisfying the homogeneous Neumann boundary condition.
(c) Since $V_{h}^{m} \subset L^{2}(\Omega)$ for all $m \geq 0$, the first term on the left-hand side of (4.1) and (4.3) is well-defined although $U^{m-1}$ and $U^{m}$ reside in different finite element spaces.
(d) In defining the nonlinear-linear form $c_{h}^{m}(\cdot, \cdot)$, we have performed integration by parts twice after multiplying the nonlinear term $-\Delta f(u)$ by a test function $v$. It turns out that this step plays an important role in simplifying the error estimates
in the next subsection. We note that this is not possible with conforming methods without generating jump terms on the edges.
(e) Since both $b_{h}^{m}(\cdot, v)$ and $c_{h}^{m}(\cdot, v)$ are zero when $v$ is constant, it follows that both of our schemes (4.1) and (4.3) conserve mass exactly, i.e. $\int_{\Omega} U^{m} d x=\int_{\Omega} U^{0} d x$ for all $m=1, \ldots, M$.
4.2. Convergence analysis. In this subsection we will derive optimal error estimates for the fully discrete dynamic mesh DG methods (4.1) and (4.3). The strategy will be to compare the fully discrete solution $U^{m}$ to the "elliptic projection" $u_{h}^{m}$ of $u\left(t_{m}\right)$ which we define as follows: For $t \in\left(t_{m-1}, t_{m}\right], m=1, \ldots, M$, $u_{h}(t)$ is the solution of the stationary problem

$$
\begin{equation*}
b_{h}^{m}\left(u_{h}(t), v\right)=b_{h}^{m}(u(t), v) \quad \forall v \in V_{h}^{m} . \tag{4.4}
\end{equation*}
$$

It is clear that for any fixed $t \in[0, T], u_{h}(t)$ satisfies the a priori estimates of Theorem 3.1. We also introduce the function $u_{h}^{m+}=\lim _{t \rightarrow t_{m}^{+}} u_{h}(t)$. If $V_{h}^{m} \neq V_{h}^{m+1}$, then $u_{h}^{m+} \neq u_{h}^{m}$. In that case, the difference $\left[u_{h}^{m}\right]:=u_{h}^{m+}-u_{h}^{m}$ is called the jump of the elliptic projection at time $t_{m}$ and will appear in an important way in the error estimates.

Since the nonlinear function $f$ is not globally Lipschitz, we will use a well-known technique which consists in analyzing instead a modified scheme where $f$ is replaced by a smooth and globally Lipschitz function $f_{L}$ that agrees with $f$ on a sufficiently large interval $[-L, L]$. After obtaining the error estimates, we then show that the (modified) fully discrete solutions are such that they also satisfy the original schemes. Let $L=2 \max _{0 \leq t \leq T}\|u(t)\|_{\infty(\Omega)}$. Then, there exists a function $f_{L}$ such that $f_{L}, f_{L}^{\prime}, f_{L}^{\prime \prime}$ are continuous and uniformly bounded on $R$ and

$$
f_{L}(x)=f(x), x \in[-L, L], \quad \text { and } \quad\left|f_{L}(x)-f_{L}(y)\right| \leq C_{L}|x-y| \forall x, y \in R
$$

With the function $f_{L}$ at hand we prove a result which will allow the estimation of the nonlinear term $c_{h}^{m}$.

Proposition 4.1. Let $c_{h}^{m}$ be defined in terms of $f_{L}$. Let $\phi, \psi, \xi \in E_{h}^{m}$ where also $\nabla \phi \in L^{\infty}(\Omega)$. Then for any $\delta>0$ we have

$$
\begin{gather*}
\left|c_{h}^{m}(\phi, \xi)-c_{h}^{m}(\psi, \xi)\right| \leq \frac{C(L)}{\delta} \sum_{K \in \mathcal{T}_{h}^{m}}\left(\|\phi-\psi\|_{K}^{2}+h_{K}^{2}\|\nabla(\phi-\psi)\|_{K}^{2}\right. \\
\left.+h_{K}^{4}|\phi-\psi|_{2, K}^{2}\right)+C(L) \delta|\xi|_{2, h}^{2} \tag{4.5}
\end{gather*}
$$

where $|\cdot|_{2, h}$ is defined in (3.29).
Proof.

$$
\begin{aligned}
c_{h}^{m}(\phi, \xi)-c_{h}^{m}(\psi, \xi)=- & \sum_{K \in \mathcal{T}_{h}^{m}}\left(f_{L}(\phi, \xi)-f_{L}(\psi, \xi)\right)_{K} \\
& +\sum_{e \in \mathcal{E}_{m}}\left\langle f_{L}(\{\phi\})-f_{L}(\{\psi\}),\left[\partial_{n} \xi\right]\right\rangle_{e} \\
& +\sum_{e \in \mathcal{E}_{m}^{I}}\left\langle f_{L}^{\prime}(\{\phi\})\left\{\partial_{n} \phi\right\}-f_{L}^{\prime}(\{\psi\})\left\{\partial_{n} \psi\right\},[\xi]\right\rangle_{e} .
\end{aligned}
$$

Since $f_{L}^{\prime}$ is bounded, for any $\delta>0$ we have

$$
\begin{align*}
\mid\left(f_{L}(\phi)\right. & \left.-f_{L}(\psi), \Delta \xi\right)_{K}\left|=\left|\int_{0}^{1}\left(f_{L}^{\prime}(\psi+s(\phi-\psi))(\phi-\psi) d s, \Delta \xi\right)_{K}\right|\right. \\
& \leq c(L)\|\phi-\psi\|_{K}\|\Delta \xi\|_{K} \leq c(L)\left\{\frac{1}{\delta}\|\phi-\psi\|_{K}^{2}+\delta\|\Delta \xi\|_{K}^{2}\right\} \tag{4.7}
\end{align*}
$$

Also, with $\tilde{K}=K^{+} \cup K^{-}$and using the trace inequality (2.1), we obtain

$$
\begin{align*}
\mid\left\langle f_{L}(\{\phi\})-\right. & \left.f_{L}(\{\psi\}),\left[\partial_{n} \xi\right]\right\rangle_{e} \mid \\
& =\left|\int_{0}^{1}\left\langle f_{L}^{\prime}(\{\psi\}+s(\{\phi-\psi\}))(\{\phi-\psi\}) d s,\left[\partial_{n} \xi\right]\right\rangle_{e}\right| \\
& \leq c(L)|\{\phi-\psi\}|_{e}\left|\left[\partial_{n} \xi\right]\right|_{e}  \tag{4.8}\\
& \leq c(L)\left\{\frac{1}{\delta}\left(\|\phi-\psi\|_{\tilde{K}}^{2}+h_{K}^{2}\|\nabla(\phi-\psi)\|_{\tilde{K}}^{2}\right)+\delta h_{e}^{-1}\left|\left[\partial_{n} \xi\right]\right|_{e}^{2}\right\}
\end{align*}
$$

where $\|\nabla(\phi-\psi)\|_{\tilde{K}}^{2}$ is shorthand for $\sum_{K=K^{+}, K^{-}}\|\nabla(\phi-\psi)\|_{K}^{2}$. Finally, since $f_{L}^{\prime}$ and $f_{L}^{\prime \prime}$ are bounded

$$
\begin{gather*}
\left|\left\langle f_{L}^{\prime}(\{\phi\})\{\phi\}-f_{L}^{\prime}(\{\psi\})\{\psi\},[\xi]\right\rangle_{e}\right| \leq\left|\int_{0}^{1}\left\langle f_{L}^{\prime}(\{\psi\})\left\{\partial_{n}(\phi-\psi)\right\},[\xi]\right\rangle_{e}\right| \\
+\left|\int_{0}^{1}\left\langle f_{L}^{\prime \prime}(\{\psi\}+s(\{\phi-\psi\}))\left\{\partial_{n} \phi\right\}(\{\phi-\psi\}) d s,[\xi]\right\rangle_{e}\right| \\
\leq c(L)\left\{\left|\left\{\partial_{n}(\phi-\psi)\right\}\right|_{e}+\|\nabla \phi\|_{L^{\infty}(\tilde{K})}|\{\phi-\psi\}|_{e}\right\}|[\xi]|_{e}  \tag{4.9}\\
\leq \frac{c(L)}{\delta}\left\{h_{K}^{2}\|\nabla(\phi-\psi)\|_{\tilde{K}}^{2}+h_{K}^{4}|\phi-\psi|_{2, \tilde{K}}^{2}\right. \\
\left.+h_{K}^{2}\|\nabla \phi\|_{L^{\infty}(\tilde{K})}^{2}\left(\|\phi-\psi\|_{\tilde{K}}^{2}+h_{K}^{2}\|\nabla(\phi-\psi)\|_{\tilde{K}}^{2}\right)\right\} \\
+C(L) \delta h_{e}^{-3}|[\xi]|_{e}^{2} .
\end{gather*}
$$

Combining (4.7), (4.8) and (4.9) we obtain (4.5).
As further preparation for the first main result of this paper, we define the quantities

$$
\begin{aligned}
\tau & :=\max _{1 \leq m \leq M} \tau_{m}, \\
h & :=\max _{1 \leq m \leq M} h_{m}, \quad h_{m}=\max _{K \in \mathcal{T}_{h}^{m}} h_{K}, \quad h_{\min }=\min _{1 \leq m \leq M} \min _{K \in \mathcal{T}_{h}^{m}} h_{K}, \\
\tilde{r} & =r \quad \text { if } r \geq 4 \text { and } \tilde{r}=2 \quad \text { if } r=3,
\end{aligned}
$$

and assume that the solution $u$ of (1.1)-(1.3) satisfies the following regularity assumptions:

$$
\begin{aligned}
& u \in C((0, T) \times \Omega) \\
& u, u_{t} \in L^{2}\left((0, T) ; H^{3}(\Omega) \cap H^{s}\left(\mathcal{T}_{h}^{m}\right)\right), s \geq 4, m=1, \ldots, M \\
& u_{t t} \in L^{2}\left((0, T) ; H^{-2}\left(\mathcal{T}_{h}^{m}\right)\right), m=1, \ldots, M
\end{aligned}
$$

Theorem 4.1. Suppose that the solution $u$ of (1.1)-(1.3) satisfies the regularity assumptions above and let the fully discrete approximations $\left\{U^{m}\right\}_{m=1}^{M}$ be given by (4.1) or (4.3) with $f$ replaced by $f_{L}$ and $U^{0}$ chosen so that $\left\|u_{0}-U^{0}\right\|=O\left(h_{1}^{\tilde{r}}\right)$ and
$\left(U^{0}, 1\right)=\left(u_{0}, 1\right)$. Then for $\tau$ and $h$ sufficiently small the following estimates hold for the error $e^{m}:=u\left(t_{m}\right)-U^{m}$ :

$$
\begin{align*}
& \max _{1 \leq m \leq M}\left\|e^{m}\right\| \leq c \mathrm{e}^{\mathrm{CT}}\left(\tau+\mathrm{h}^{\tilde{\mathrm{r}}}+\mathcal{N}_{\mathrm{c}} \max _{1 \leq \mathrm{m} \leq \mathrm{M}-1}\left\|\left[\mathrm{u}_{\mathrm{h}}^{\mathrm{m}}\right]\right\|\right)  \tag{4.10}\\
& \left(\sum_{m=1}^{M} \tau_{m}\left\|e^{m}\right\|_{2, h}^{2}\right)^{\frac{1}{2}} \leq c \mathrm{e}^{\mathrm{CT}}\left(\tau+\mathrm{h}^{\mathrm{r}-2}+\mathcal{N}_{\mathrm{c}} \max _{1 \leq \mathrm{m} \leq \mathrm{M}-1}\left\|\left[\mathrm{u}_{\mathrm{h}}^{\mathrm{m}}\right]\right\|\right)  \tag{4.11}\\
& \left(\sum_{m=1}^{M} \tau_{m} \sum_{K \in \mathcal{T}_{h}^{m}}\left\|\nabla e^{m}\right\|_{K}^{2}\right)^{\frac{1}{2}} \leq c \mathrm{e}^{\mathrm{CT}}\left(\tau+\mathrm{h}^{\tilde{\mathrm{r}}-1}+\mathcal{N}_{\mathrm{c}} \max _{1 \leq \mathrm{m} \leq \mathrm{M}-1}\left\|\left[\mathrm{u}_{\mathrm{h}}^{\mathrm{m}}\right]\right\|\right) \tag{4.12}
\end{align*}
$$

where the constant $C$ is proportional to $C(L)^{2} / \varepsilon$ and $\mathcal{N}_{c}$ denotes the total number of times where $V_{h}^{m} \neq V_{h}^{m-1}, m=1,2, \cdots, M$. Furthermore, there exists a constant $c_{0}$ such that if also $h_{\min }^{-\frac{d}{2}}\left(\tau+h^{\tilde{r}}+\mathcal{N}_{c} \max _{1 \leq m \leq M-1}\left\|\left[u_{h}^{m}\right]\right\|\right) \leq c_{0}$, then the estimates (4.10) -(4.12) also hold for the (unmodified) schemes (4.1) and (4.3).

Proof. Since the proof is long, we divide it into several steps.
Step 1: Derivation of the error equation.
Let $u^{m}=u\left(t_{m}\right), u_{h}^{m}=u_{h}\left(t_{m}\right), m=1, \ldots, M\left(u^{0}=u_{0}\right)$, where $u$ is the solution of the differential problem (1.1)-(1.3) and $u_{h}$ denotes its elliptic projection defined by (4.4). We see that

$$
\begin{equation*}
\left(d_{t} u^{m}, v_{h}\right)+\varepsilon b_{h}^{m}\left(u^{m}, v_{h}\right)+\frac{1}{\varepsilon} c_{h}^{m}\left(u^{m}, v_{h}\right)=\left(\mathcal{R}_{m}, v_{h}\right) \quad \forall v_{h} \in V_{h}^{m} \tag{4.13}
\end{equation*}
$$

where $\mathcal{R}_{m}=-\tau_{m}^{-1} \int_{t_{m-1}}^{t_{m}}\left(t-t_{m-1}\right) u_{t t}(t) d t$.
Subtracting (4.1) or (4.3) from (4.13) yields the following error equation: For any $v_{h} \in V_{h}^{m}$,

$$
\begin{equation*}
\left(d_{t} e^{m}, v_{h}\right)+\varepsilon b_{h}^{m}\left(e^{m}, v_{h}\right)+\frac{1}{\varepsilon} c_{h}^{m}\left(u^{m}, v_{h}\right)-\frac{1}{\varepsilon} c_{h}^{m}\left(\psi^{m}, v_{h}\right)=\left(\mathcal{R}_{m}, v_{h}\right) \tag{4.14}
\end{equation*}
$$

where $\psi^{m}=U^{m}$ for (4.1) and $\psi^{m}=P^{m} U^{m-1}$ for (4.3). Now introduce the decomposition

$$
e^{m}:=\eta^{m}+\xi_{h}^{m}, \quad \text { where } \eta^{m}:=u^{m}-u_{h}^{m}, \quad \xi_{h}^{m}:=u_{h}^{m}-U^{m} .
$$

Then from the definition of the elliptic projection $u_{h}^{m}$, we can rewrite (4.14) as

$$
\begin{align*}
\left(d_{t} \xi_{h}^{m}, v_{h}\right)+\varepsilon b_{h}^{m}\left(\xi_{h}^{m}, v_{h}\right)+\frac{1}{\varepsilon} c_{h}^{m}\left(u^{m}, v_{h}\right) & -\frac{1}{\varepsilon} c_{h}^{m}\left(\psi^{m}, v_{h}\right)  \tag{4.15}\\
& =\left(\mathcal{R}_{m}, v_{h}\right)-\left(d_{t} \eta^{m}, v_{h}\right)
\end{align*}
$$

Setting $v_{h}=\xi_{h}^{m}$ in (4.14) and using the coercivity of the bilinear form $b_{h}^{m}(\cdot, \cdot)$ yield

$$
\begin{array}{r}
\frac{1}{2} d_{t}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\tau_{m}}{2}\left\|d_{t} \xi_{h}^{m}\right\|^{2}+c_{b} \varepsilon\left\|\xi_{h}^{m}\right\|_{2, h}^{2} \leq\left(\mathcal{R}_{m}, \xi_{h}^{m}\right)-\left(d_{t} \eta^{m}, \xi_{h}^{m}\right) \\
+\frac{1}{\varepsilon}\left(\left(c_{h}^{m}\left(\psi^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)\right)\right. \tag{4.16}
\end{array}
$$

Step 2: Estimation of the $\left(\mathcal{R}_{m}, \xi_{h}^{m}\right)$ and $\left(d_{t} \eta^{m}, \xi_{h}^{m}\right)$ terms.
Without loss of generality, we may assume that $\xi_{h}^{m}$ is nonconstant over $\Omega$. Thus,

$$
\begin{gather*}
\left(\mathcal{R}_{m}, \xi_{h}^{m}\right) \leq\left\|\mathcal{R}_{m}\right\|_{-2, h}\left\|\xi_{h}^{m}\right\|_{2, h} \leq \frac{c_{b} \varepsilon}{8}\left\|\xi_{h}^{m}\right\|_{2, h}^{2}+\frac{2}{\varepsilon c_{b}}\left\|\mathcal{R}_{m}\right\|_{-2, h}^{2} \\
\leq \frac{c_{b} \varepsilon}{8}\left\|\xi_{h}^{m}\right\|_{2, h}^{2}+\frac{2 \tau_{m}}{\varepsilon c_{b}} A_{0}^{m} \tag{4.17}
\end{gather*}
$$

where

$$
\begin{equation*}
A_{0}^{m}=\left\|u_{t t}\right\|_{\left(L^{2}\left(J_{m} ; H^{-2}\left(\mathcal{T}_{h}^{m}\right)\right)\right.}^{2}=\int_{t_{m-1}}^{t_{m}}\left\|u_{t t}(t)\right\|_{-2, h}^{2} d t \tag{4.18}
\end{equation*}
$$

Let $d_{t}^{+} \eta^{m}=\tau_{m}^{-1}\left(\eta^{m}-\eta^{(m-1)+}\right)$. We have

$$
\begin{align*}
& \left|\left(d_{t} \eta^{m}, \xi_{h}^{m}\right)\right|=\left(d_{t}^{+} \eta^{m}, \xi_{h}^{m}\right)+\tau_{m}^{-1}\left(\left[\eta^{m-1}\right], \xi_{h}^{m}\right) \\
& \quad \leq \frac{1}{2}\left\|d_{t}^{+} \eta^{m}\right\|^{2}+\frac{1}{2}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\beta_{m} \tau_{m}^{-1}}{2}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\beta_{m}^{-1} \tau_{m}^{-1}}{2}\left\|\left[\eta^{m-1}\right]\right\|^{2} \tag{4.19}
\end{align*}
$$

where $\beta_{m}=0$ if $V_{h}^{m-1}=V_{h}^{m}$ in which case $\left[\eta^{m-1}\right]=0$ and will be chosen appropriately if $V_{h}^{m-1} \neq V_{h}^{m}$.

Now $d_{t}^{+} \eta^{m}=\tau_{m}^{-1} \int_{t_{m-1}}^{t_{m}} \eta_{t}(t) d t$, so using the a priori estimate (3.19) or (3.22), we obtain

$$
\begin{equation*}
\left\|d_{t}^{+} \eta^{m}\right\|^{2} \leq \tau_{m}^{-2}\left(\int_{t_{m-1}}^{t_{m}}\left\|\eta_{t}(t)\right\| d t\right)^{2} \leq \tau_{m}^{-1} \int_{t_{m-1}}^{t_{m}}\left\|\eta_{t}(t)\right\|^{2} d t \leq \tau_{m}^{-1} A_{1}^{m} \tag{4.20}
\end{equation*}
$$

where

$$
A_{1}^{m}=\left\{\begin{array}{l}
\left(h^{2}\left|h^{r-2} u_{t}\right|_{L^{2}\left(J_{m} ; H^{r}\left(\mathcal{T}_{h}^{m}\right)\right)}\right)^{2}, \quad r \geq 4,  \tag{4.21}\\
\left(h\left(\left|h u_{t}\right|_{L^{2}\left(J_{m} ; H^{3}\left(\mathcal{T}_{h}\right)\right)}+\left|h^{2} u_{t}\right|_{L^{2}\left(J_{m} ; H^{4}\left(\mathcal{T}_{h}\right)\right)}\right)\right)^{2}, \quad r=3
\end{array}\right.
$$

Step 3: Estimation of the nonlinear terms.
For the fully implicit scheme, we use Proposition 4.1 with $\phi=u^{m}, \psi=U^{m}$ and $\xi=\xi_{h}^{m}$. Writing $u^{m}-U^{m}=\eta^{m}+\xi_{h}^{m}$, from (4.5) it follows that

$$
\begin{gather*}
\left|c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(U^{m}, \xi_{h}^{m}\right)\right| \leq \frac{C}{\delta}\left\{\left\|\eta^{m}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla \eta^{m}\right\|_{K}^{2}+h_{K}^{4}\left|\eta^{m}\right|_{2, K}^{2}\right. \\
\left.+\left\|\xi_{h}^{m}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla \xi_{h}^{m}\right\|_{K}^{2}+h_{K}^{4}\left|\xi_{h}^{m}\right|_{2, K}^{2}\right\}+C \delta\left|\xi_{h}^{m}\right|_{2, h}^{2} \tag{4.22}
\end{gather*}
$$

Now using the approximation properties of the elliptic projection for the terms in $\eta^{m}$, specifically (3.19) and (3.20) for $r \geq 4$ or (3.22) and (3.23) for $r=3$, we obtain

$$
\begin{equation*}
\sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|\eta^{m}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla \eta^{m}\right\|_{K}^{2}+h_{K}^{4}\left|\eta^{m}\right|_{2, K}^{2}\right) \leq c A_{2}^{m} \tag{4.23}
\end{equation*}
$$

where

$$
A_{2}^{m}=\left\{\begin{array}{l}
\left(h^{2}\left|h^{r-2} u^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}\right)^{2}, \quad r \geq 4,  \tag{4.24}\\
\left(h\left(\left|h u^{m}\right|_{H^{3}\left(\mathcal{T}_{h}\right)}+\left|h^{2} u^{m}\right|_{H^{4}\left(\mathcal{T}_{h}\right)}\right)\right)^{2}, \quad r=3 .
\end{array}\right.
$$

For the $\xi_{h}^{m}$ terms in (4.22) we use the inverse inequalities to get

$$
\begin{equation*}
\sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|\xi_{h}^{m}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla \xi_{h}^{m}\right\|_{K}^{2}+h_{K}^{4}\left|\xi_{h}^{m}\right|_{2, K}^{2}\right) \leq c\left\|\xi_{h}^{m}\right\|^{2} \tag{4.25}
\end{equation*}
$$

Using (4.23) and (4.25) in (4.22), we obtain

$$
\begin{equation*}
\left|c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(U^{m}, \xi_{h}^{m}\right)\right| \leq \frac{C(L)}{\delta}\left(\left\|\xi_{h}^{m}\right\|^{2}+A_{2}^{m}\right)+C(L) \delta\left|\xi_{h}^{m}\right|_{2, h}^{2} \tag{4.26}
\end{equation*}
$$

For the implicit-explicit method we use Proposition 4.1 with $\phi=u^{m}$ and $\psi=$ $P^{m} U^{m-1}$. Then,

$$
\phi-\psi=u^{m}-u^{m-1}+u^{m-1}-P^{m} u^{m-1}+P^{m} \eta^{m-1}+P^{m} \xi_{h}^{m-1}
$$

Now $u^{m}-u^{m-1}=\int_{t_{m-1}}^{t_{m}} u_{t}(s) d s$, thus,

$$
\begin{equation*}
\sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|u-u^{m-1}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla\left(u-u^{m-1}\right)\right\|_{K}^{2}+h_{K}^{4}\left|u-u^{m-1}\right|_{2, K}^{2}\right) \leq \tau_{m} A_{3}^{m} \tag{4.27}
\end{equation*}
$$

where

$$
\begin{equation*}
A_{3}^{m}=\int_{t_{m-1}}^{t_{m}}\left(\left\|u_{t}\right\|^{2}+\left|h u_{t}\right|_{H^{1}\left(\mathcal{T}_{h}^{m}\right)}^{2}+\left|h^{2} u_{t}\right|_{H^{2}\left(\mathcal{T}_{h}^{m}\right)}^{2}\right) \tag{4.28}
\end{equation*}
$$

Also note that $u^{m-1}-P^{m} u^{m-1}=\left.\sum_{K \in \mathcal{T}_{h, C}^{m}}\left(u^{m-1}-P_{K}^{m} u^{m-1}\right)\right|_{K}$ where $\mathcal{T}_{h, C}^{m}$ is the set of elements in $\mathcal{T}_{h}^{m}$ that were obtained from $\mathcal{T}_{h}^{m-1}$ through coarsening. Hence, using the approximation properties of the operator $P^{m}$, we obtain

$$
\begin{align*}
\sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|u^{m-1}-P^{m} u^{m-1}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla\left(u^{m-1}-P^{m} u^{m-1}\right)\right\|_{K}^{2}\right. \\
\left.+h_{K}^{4}\left|u^{m-1}-P^{m} u^{m-1}\right|_{2, K}^{2}\right) \leq c A_{4}^{m-1} \tag{4.29}
\end{align*}
$$

where

$$
\begin{equation*}
A_{4}^{m-1}=\left|h^{r} u^{m-1}\right|_{H^{r}\left(\mathcal{T}_{h, C}^{m}\right)}^{2} \tag{4.30}
\end{equation*}
$$

Now using the polynomial inverse inequalities and the stability of the operator $P^{m}$ in the $L^{2}$ norm, we obtain

$$
\begin{align*}
\sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|P^{m} \eta^{m-1}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla P^{m} \eta^{m-1}\right\|_{K}^{2}+h_{K}^{4}\left|P^{m} \eta^{m-1}\right|_{2, K}^{2}\right) \\
\leq c \sum_{K \in \mathcal{T}_{h}^{m}}\left\|P^{m} \eta^{m-1}\right\|_{K}^{2} \leq c \sum_{K \in \mathcal{T}_{h}^{m}}\left\|\eta^{m-1}\right\|_{K}^{2} \leq c A_{5}^{m-1} \tag{4.31}
\end{align*}
$$

where

$$
A_{5}^{m-1}= \begin{cases}\left(h^{2}\left|h^{r-2} u^{m-1}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}\right)^{2}, & r \geq 4  \tag{4.32}\\ \left(h\left(\left|h u^{m-1}\right|_{H^{3}\left(\mathcal{T}_{h}\right)}+\left|h^{2} u^{m-1}\right|_{H^{4}\left(\mathcal{T}_{h}\right)}\right)\right)^{2}, & r=3\end{cases}
$$

In the same way we obtain

$$
\begin{align*}
& \sum_{K \in \mathcal{T}_{h}^{m}}\left(\left\|P^{m} \xi_{h}^{m-1}\right\|_{K}^{2}+h_{K}^{2}\left\|\nabla P^{m} \xi_{h}^{m-1}\right\|_{K}^{2}+h_{K}^{4}\left|P^{m} \xi_{h}^{m-1}\right|_{2, K}^{2}\right) \\
& \leq c \sum_{K \in \mathcal{T}_{h}^{m}}\left\|P^{m} \xi_{h}^{m-1}\right\|_{K}^{2} \leq c \sum_{K \in \mathcal{T}_{h}^{m}}\left\|\xi_{h}^{m-1}\right\|_{K}^{2} \tag{4.33}
\end{align*}
$$

So gathering (4.27)-(4.33), for the implicit-explicit scheme we obtain

$$
\begin{align*}
& \left|c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(P^{m} U^{m-1}, \xi_{h}^{m}\right)\right|  \tag{4.34}\\
& \quad \leq \frac{C(L)}{\delta}\left(\tau_{m} A_{3}^{m}+A_{4}^{m-1}+A_{5}^{m-1}+\left\|\xi_{h}^{m-1}\right\|^{2}\right)+C(L) \delta\left|\xi_{h}^{m}\right|_{2, h}^{2}
\end{align*}
$$

Step 4: Stability and convergence.
From (4.16) and using (4.17), (4.19), (4.20) and (4.26), for the fully implicit scheme we have

$$
\begin{align*}
\frac{1}{2} d_{t}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\tau_{m}}{2} \| & d_{t} \xi_{h}^{m}\left\|^{2}+c_{b} \varepsilon\right\| \xi_{h}^{m}\left\|_{2, h}^{2} \leq \frac{c_{b} \varepsilon}{4}\right\| \xi_{h}^{m} \|_{2, h}^{2}+\frac{\tau_{m}}{c_{b} \varepsilon} A_{0}^{m} \\
& +\tau_{m}^{-1} A_{1}^{m}+\frac{1}{2}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\beta_{m} \tau_{m}^{-1}}{2}\left\|\xi_{h}^{m}\right\|^{2}+\frac{\beta_{m}^{-1} \tau_{m}^{-1}}{2}\left\|\left[\eta^{m-1}\right]\right\|^{2}  \tag{4.35}\\
& +\frac{C(L)}{\delta}\left(\left\|\xi_{h}^{m}\right\|^{2}+A_{2}^{m}\right)+C(L) \delta\left|\xi_{h}^{m}\right|_{2, h}
\end{align*}
$$

We now choose $\delta=\frac{c_{b} \varepsilon}{4 C(L)}$ and multiply (4.35) by $2 \tau_{m}$ to get

$$
\begin{equation*}
\left(1-\alpha_{m}\right)\left\|\xi_{h}^{m}\right\|^{2}+\Gamma_{1}^{m} \leq\left\|\xi_{h}^{m-1}\right\|^{2}+\Gamma_{2}^{m}, \quad m=1, \ldots, M \tag{4.36}
\end{equation*}
$$

where

$$
\begin{aligned}
\alpha_{m} & =\tau_{m}+8 \frac{C(L)^{2}}{c_{b} \varepsilon} \tau_{m}+\beta_{m} \\
\Gamma_{1}^{m} & =\left\|\xi_{h}^{m}-\xi_{h}^{m-1}\right\|^{2}+\tau_{m} c_{b} \varepsilon\left\|\xi_{h}^{m}\right\|_{2, h}^{2} \\
\Gamma_{2}^{m} & =\frac{2}{c_{b} \varepsilon} \tau_{m}^{2} A_{0}^{m}+2 A_{1}^{m}+\beta_{m}^{-1}\left\|\left[\eta^{m-1}\right]\right\|^{2}+8 \frac{C(L)^{2}}{c_{b} \varepsilon} \tau_{m} A_{2}^{m}
\end{aligned}
$$

We assume that $\alpha_{m}<1$. This is achieved if $0 \leq \beta_{j} \leq 1 / 2$ (see below) and if $\tau_{m}$ is sufficiently small. We then replace $m$ by $\ell$ in (4.36), multiply (4.36) by $\prod_{j=1}^{\ell-1}\left(1-\alpha_{j}\right)$, sum over $\ell$ from 1 to $m$, and finally multiply the sum by $\prod_{j=1}^{m-1}\left(1-\alpha_{j}\right)^{-1}$ to obtain

$$
\begin{align*}
\left\|\xi_{h}^{m}\right\|^{2}+\sum_{\ell=1}^{m} \prod_{j=\ell}^{m}\left(1-\alpha_{j}\right)^{-1} \Gamma_{1}^{\ell} \leq & \prod_{j=1}^{m}\left(1-\alpha_{j}\right)^{-1}\left\|\xi_{h}^{0}\right\|^{2}  \tag{4.37}\\
& +\sum_{\ell=1}^{m} \prod_{j=\ell}^{m}\left(1-\alpha_{j}\right)^{-1} \Gamma_{2}^{\ell}
\end{align*}
$$

Next, we bound the quantities $\prod_{j=\ell}^{m}\left(1-\alpha_{j}\right)^{-1}$. First we have $\left(1-\alpha_{j}\right)^{-1} \leq \mathrm{e}^{2 \alpha_{\mathrm{j}}}$, the inequality holding for $0 \leq \alpha_{j}<1 / 2$ which we assume. Also, to simplify the notation we let $\kappa:=1+8 \frac{C(L)^{2}}{c_{b} \varepsilon}$. Let $\mathcal{N}_{c}$ denote the number of times where $\mathcal{T}_{h}^{m-1} \neq$ $\mathcal{T}_{h}^{m}, m=1, \ldots, M$. If $\mathcal{N}_{c}=0$, we let $\beta_{j}=0$. Otherwise we let $\beta_{j}=\min \left\{\frac{1}{2}, \frac{1}{\mathcal{N}_{c}}\right\}$. Hence,

$$
\begin{equation*}
\prod_{j=\ell}^{m}\left(1-\alpha_{j}\right)^{-1} \leq \prod_{\substack{j=\ell \\ \beta_{j}<\kappa \tau_{j}}}^{m} \mathrm{e}^{4 \kappa \tau_{\mathrm{j}}} \prod_{\substack{\mathrm{j}=\ell \\ \beta_{\mathrm{j}} \geq \kappa \tau_{\mathrm{j}}}}^{\mathrm{m}} \mathrm{e}^{4 / \mathcal{N}_{\mathrm{c}}} \leq \mathrm{e}^{4 \kappa\left(\mathrm{t}_{\mathrm{m}}-\mathrm{t}_{\ell-1}\right)} \mathrm{e}^{4} \tag{4.38}
\end{equation*}
$$

Here we have used the fact that the number of times when $\beta_{j} \neq 0$ is bounded by $\mathcal{N}_{c}$. Similarly, using the inequality $\left(1-\alpha_{j}\right)^{-1} \geq e^{\alpha_{j}}$, which holds for $0 \leq \alpha_{j}<1$, we obtain the lower bound

$$
\begin{equation*}
\prod_{j=\ell}^{m}\left(1-\alpha_{j}\right)^{-1} \geq \mathrm{e}^{\kappa\left(\mathrm{t}_{\mathrm{m}}-\mathrm{t}_{\ell-1}\right)} \tag{4.39}
\end{equation*}
$$

Thus, using (4.38) and (4.39) in (4.37), we obtain

$$
\begin{equation*}
\left\|\xi_{h}^{m}\right\|^{2}+\sum_{\ell=1}^{m} \mathrm{e}^{\kappa\left(\mathrm{t}_{\mathrm{m}}-\mathrm{t}_{\ell-1}\right)} \Gamma_{1}^{\ell} \leq \mathrm{e}^{4+4 \kappa \mathrm{t}_{\mathrm{m}}}\left\|\xi_{\mathrm{h}}^{0}\right\|^{2}+\sum_{\ell=1}^{\mathrm{m}} \mathrm{e}^{4+4 \kappa\left(\mathrm{t}_{\mathrm{m}}-\mathrm{t}_{\ell-1}\right)} \Gamma_{2}^{\ell} \tag{4.40}
\end{equation*}
$$

Now recalling the definition of $\Gamma_{2}^{\ell}$, for $m=1, \ldots, M$, we obtain

$$
\begin{align*}
& \left\|\xi_{h}^{m}\right\|^{2}+c_{b} \varepsilon \sum_{\ell=1}^{m} \tau_{\ell} e^{\kappa\left(t_{m}-t_{\ell}\right)}\left\|\xi_{h}^{\ell}\right\|_{2, h}^{2} \\
& \quad \leq c \mathrm{e}^{\mathrm{c}_{1} \mathrm{t}_{\mathrm{m}}}\left(\left\|\xi_{\mathrm{h}}^{0}\right\|^{2}+\tau^{2}+\mathrm{h}^{2 \tilde{\mathrm{r}}}+\left(\mathcal{N}_{\mathrm{c}}\right)^{2} \max _{1 \leq \ell \leq \mathrm{m}}\left\|\left[\mathrm{u}_{\mathrm{h}}^{\ell-1}\right]\right\|^{2}\right) \tag{4.41}
\end{align*}
$$

Here the constant $c_{1}$ is proportional to $\frac{C(L)^{2}}{c_{b} \varepsilon}$. The implicit-explicit scheme is treated in a similar manner: using (4.34) instead of (4.26) the analog of (4.36) is

$$
\begin{equation*}
\left(1-\tau_{m}-\beta_{m}\right)\left\|\xi_{h}^{m}\right\|^{2}+\Gamma_{1}^{m} \leq\left(1+8 \frac{C(L)^{2}}{c_{b} \varepsilon} \tau_{m}\right)\left\|\xi_{h}^{m-1}\right\|^{2}+\Gamma_{3}^{m} \tag{4.42}
\end{equation*}
$$

with
$\Gamma_{3}^{m}=\frac{2}{c_{b} \varepsilon} \tau_{m}^{2} A_{0}^{m}+2 A_{1}^{m}+\beta_{m}^{-1}\left\|\left[\eta^{m-1}\right]\right\|^{2}+8 \frac{C(L)^{2}}{c_{b} \varepsilon}\left(\tau_{m}^{2} A_{3}^{m}+\tau_{m} A_{4}^{m-1}+\tau_{m} A_{5}^{m-1}\right)$,
leading again to the estimate (4.41).
Given the choice of $U^{0}$, it follows from the triangle inequality and the approximation properties of the elliptic projection $u_{h}^{m}$ (Theorem 3.1) that $\left\|\xi_{h}^{0}\right\|=O\left(h^{\tilde{r}}\right)$. Thus (4.10) and (4.11) follow readily from (4.41), Theorem 3.1 and the triangle inequality. As to (4.12), from (4.41) and the interpolation inequality (3.29) $(\delta=1)$ it follows that

$$
\begin{aligned}
\sum_{m=1}^{M} \tau_{m} \sum_{K \in \mathcal{T}_{h}^{m}}\left\|\nabla \xi_{h}^{m}\right\|_{K}^{2} & \leq c T \max _{1 \leq m \leq M}\left\|\xi_{h}^{m}\right\|^{2}+c \sum_{m=1}^{M} \tau_{m}\left\|\xi_{h}^{m}\right\|_{2, h}^{2} \\
& \leq c \mathrm{e}^{\mathrm{c}_{1} \mathrm{t}_{\mathrm{m}}}\left(\tau^{2}+\mathrm{h}^{2 \tilde{\mathrm{r}}}+\left(\mathcal{N}_{\mathrm{c}}\right)^{2} \max _{1 \leq \ell \leq \mathrm{m}}\left\|\left[\mathrm{u}_{\mathrm{h}}^{\ell-1}\right]\right\|^{2}\right)
\end{aligned}
$$

Thus, (4.12) follows from this, (3.20) or (3.23) and the triangle inequality.
Step 5: Back to $f$. We have proved that the estimates (4.10)-(4.12) hold for the solution of the modified schemes (4.1) and (4.3) obtained by replacing the function $f$ by the Lipschitz function $f_{L}$. As the final step of the proof, we will show that under the additional CFL condition $h_{\min }^{-\frac{d}{2}}\left(\tau+h^{\tilde{r}}+\mathcal{N}_{c} \max _{1 \leq m \leq M-1}\left\|\left[u_{h}^{m}\right]\right\|\right) \leq c_{0}$ the same estimates also hold for the solutions of the schemes themselves. This is a simple consequence of the bound $\max _{0 \leq m \leq M}\left\|U^{m}\right\|_{L^{\infty}(\Omega)} \leq L$. which is established next. Indeed, let $I_{h}^{m}$ denote the interpolant on $\mathcal{T}_{h}^{m}$ induced by the local interpolant operators $I_{K}$ in Lemma 2.4. Then, using the well-known $L^{\infty}-L^{2}$ inverse inequality $\|\chi\|_{L^{\infty}(\Omega)} \leq c h^{-\frac{d}{2}}\|\chi\|$ together with the estimate (4.10) and (2.7), for each $m, 1 \leq$ $m \leq M$, we get

$$
\begin{gathered}
\left\|U^{m}\right\|_{L^{\infty}(\Omega)} \leq\left\|U^{m}-I_{h}^{m} u\left(t_{m}\right)\right\|_{L^{\infty}(\Omega)}+\left\|I_{h}^{m} u\left(t_{m}\right)-u\left(t_{m}\right)\right\|_{L^{\infty}(\Omega)} \\
\quad+\left\|u\left(t_{m}\right)\right\|_{L^{\infty}(\Omega)} \\
\leq \operatorname{ch}_{\min }^{-\frac{1}{2}}\left(\tau+h^{\tilde{r}}+\mathcal{N}_{c} \max _{1 \leq m \leq M-1}\left\|\left[u_{h}^{m}\right]\right\|\right) \\
\quad+\operatorname{ch}_{m}^{2}\left|u\left(t_{m}\right)\right|_{W^{2, \infty}(\Omega)}+\left\|u\left(t_{m}\right)\right\|_{L^{\infty}(\Omega)} .
\end{gathered}
$$

Thus the required $L^{\infty}$ bound follows from the stated conditions. The same bound on $U^{0}$ (needed only for the implicit-explicit scheme) is obtained in a similar manner. This concludes the proof.

Remark 4.2. (a) If $V_{h}^{m}=V_{h}, m=1, \ldots, M$, then $\mathcal{N}_{c}=0$ and the error estimates obtained are optimal in $\tau$ and $h$. In addition, $\left\|\xi_{h}^{m}\right\|_{2, h}$ enjoys superconvergence in $h$.
(b) If $V_{h}^{m} \neq V_{h}^{(m+1)}$, then one still has

$$
\left\|\left[u_{h}^{m}\right]\right\|=\left\|\left(u\left(t_{m}\right)-u_{h}^{m}\right)-\left(u\left(t_{m}\right)-u_{h}^{(m+1)+}\right)\right\|=O\left(h^{r}\right)
$$

i.e. the jump terms are optimal in $h$. However, one cannot say that the estimates of Theorem 4.1 are optimal unless $\mathcal{N}_{c} \leq c$ independently of $\tau$. This is not an unreasonable assumption in many practical situations.
(c) If $V_{h}^{m}$ and $V_{h}^{m+1}$ differ only slightly, e.g. if $\mathcal{T}_{h}^{m}$ is modified in a small region only, then it is possible to obtain an improved estimate for the jump term. For instance, the bound $\left\|\left[u_{h}^{m}\right]\right\|=O\left(h^{r+\frac{d}{2}}\right)$ was obtained in 30 for a standard Galerkin formulation of the nonlinear Schrödinger equation.
(d) As expected, the constants in (4.10)-(4.12) depend exponentially on $\varepsilon$, more precisely $c=O\left(e^{\frac{1}{\varepsilon}}\right)$. By using nonstandard error analysis techniques, this dependence was reduced to a low order polynomial in [22, [23] for mixed finite element approximations. We will present such a result and analysis for the DG method in a forthcoming work.

To further elaborate on the issues raised in Remark 4.2, we observe that the inability to estimate precisely $\left\|\left[u_{h}^{m}\right]\right\|$ in terms of the local changes in the mesh stems from the global nature of the elliptic projection operator. Since the Lagrange interpolation operator $I_{h}^{m}$ is local to each cell, we consider modifying the proof of Theorem 4.1 by comparing $U^{m}$ to $I_{h}^{m} u^{m}$ instead. This technique goes back to Douglas and Dupont [14 and was largely forgotten after the introduction by M. Wheeler [35] of the elliptic projection as a tool in the error estimation of evolution equations.

Theorem 4.2. Let $I_{h}^{m}$ denote the standard nodal interpolation operator from $C(\bar{\Omega})$ to $V_{h}^{m}$. Then under the assumptions of Theorem 4.1 the errors $e^{m}=u\left(t_{m}\right)-U^{m}$ satisfy

$$
\begin{align*}
\max _{1 \leq m \leq M}\left\|e^{m}\right\| & +\left(\sum_{m=1}^{M} \tau_{m}\left\|e^{m}\right\|_{2, h}^{2}\right)^{\frac{1}{2}}  \tag{4.43}\\
& \leq C\left(\tau+h^{r-2}+\mathcal{N}_{c} \max _{1 \leq m \leq M-1}\left\|\left[I_{h}^{m} u\left(t_{m}\right)\right]\right\|\right)
\end{align*}
$$

Proof. Since the proof is similar to that of Theorem4.1, we only highlight the main points. First, we introduce the new error decomposition

$$
e^{m}=\eta^{m}+\xi_{h}^{m}, \quad \text { where now } \eta^{m}:=u^{m}-I_{h}^{m} u^{m}, \xi_{h}^{m}:=I_{h}^{m} u^{m}-U^{m}
$$

Setting $v_{h}=\xi_{h}^{m}=e^{m}-\eta^{m}$ in the error equation (4.14) yields

$$
\begin{align*}
\left(d_{t} e^{m}, e^{m}\right)+\varepsilon b_{h}^{m}\left(e^{m}, e^{m}\right)= & \left(\mathcal{R}_{m}, e^{m}-\eta^{m}\right)+\left(d_{t} e^{m}, \eta^{m}\right)+\varepsilon b_{h}^{m}\left(e^{m}, \eta^{m}\right) \\
& +\frac{1}{\varepsilon}\left(c_{h}^{m}\left(\psi^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)\right) \tag{4.44}
\end{align*}
$$

with $\psi$ as in Theorem 4.1. We now bound each term (except the second term on the right-hand side) in (4.44) as follows:

$$
\begin{align*}
& \left(d_{t} e^{m}, e^{m}\right)=\frac{1}{2} d_{t}\left\|e^{m}\right\|^{2}+\frac{\tau_{m}}{2}\left\|d_{t} e^{m}\right\|^{2}  \tag{4.45}\\
& \varepsilon b_{h}^{m}\left(e^{m}, e^{m}\right) \geq \varepsilon c_{0}\left\|e^{m}\right\|_{2, h}^{2}-\varepsilon c_{1}\left|h^{r} e^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}  \tag{4.46}\\
& \left|\left(\mathcal{R}_{m}, e^{m}-\eta^{m}\right)\right| \leq \frac{\varepsilon c_{0}}{8}\left\|e^{m}\right\|_{2, h}^{2}+c\left\|\eta^{m}\right\|_{2, h}^{2}+\frac{c}{\varepsilon}\left\|\mathcal{R}_{m}\right\|_{-2, h}^{2}  \tag{4.47}\\
& \left|\varepsilon b_{h}^{m}\left(e^{m}, \eta^{m}\right)\right| \leq \frac{\varepsilon c_{0}}{8}\left\|e^{m}\right\|_{2, h}^{2}+c \varepsilon\left\|\eta^{m}\right\|_{2, h}^{2} \tag{4.48}
\end{align*}
$$

(4.45) is obvious. (4.47) is obtained as in Theorem 4.1; (4.48) expresses the continuity of the form $b_{h}$ while (4.46) follows from weak coercivity estimate (3.24). In this context it is important to notice that $\left|h^{r} e^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}=\left|h^{r} u^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}$.

We next bound the nonlinear terms. As in Theorem 4.1, it follows from Proposition 4.1 that for the fully implicit scheme we have

$$
\begin{align*}
\left|c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(U^{m}, \xi_{h}^{m}\right)\right| \leq & \frac{C(L)}{\delta}\left(\left\|\xi_{h}^{m}\right\|^{2}+A_{2}^{m}\right)+C(L) \delta\left|\xi_{h}^{m}\right|_{2, h}^{2} \\
\leq & \frac{C(L)}{\delta}\left(\left\|e^{m}\right\|^{2}+\left\|\eta^{m}\right\|^{2}+A_{2}^{m}\right)  \tag{4.49}\\
& +C(L) \delta\left(\left|e^{m}\right|_{2, h}^{2}+\left|\eta^{m}\right|_{2, h}^{2}\right)
\end{align*}
$$

where the term $A_{2}^{m}$ now satisfies (compare with (4.24))

$$
\begin{equation*}
A_{2}^{m} \leq c\left|h^{r} u^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}, \quad r \geq 3 \tag{4.50}
\end{equation*}
$$

For the implicit-explicit scheme we have

$$
\begin{align*}
& \left|c_{h}^{m}\left(u^{m}, \xi_{h}^{m}\right)-c_{h}^{m}\left(P^{m} U^{m-1}, \xi_{h}^{m}\right)\right| \\
& \quad \leq \frac{C(L)}{\delta}\left(\tau_{m} A_{3}^{m}+A_{4}^{m-1}+A_{5}^{m-1}+\left\|\xi_{h}^{m-1}\right\|^{2}\right)+C(L) \delta\left|\xi_{h}^{m}\right|_{2, h}^{2} \\
& \quad \leq \frac{C(L)}{\delta}\left(\tau_{m} A_{3}^{m}+A_{4}^{m-1}+A_{5}^{m-1}+\left\|e^{m-1}\right\|^{2}+\left\|\eta^{m-1}\right\|^{2}\right)  \tag{4.51}\\
& \quad+C(L) \delta\left(\left|e^{m}\right|_{2, h}^{2}+\left|\eta^{m}\right|_{2, h}^{2}\right)
\end{align*}
$$

where $A_{3}^{m}, A_{4}^{m-1}$ are exactly as in Theorem 4.1 and satisfy (4.28) and (4.30) respectively, while $A_{h}^{m-1}$ satisfies

$$
\begin{equation*}
A_{5}^{m-1} \leq\left|h^{r} u^{m-1}\right|_{H^{r}\left(\mathcal{T}_{h}^{m-1}\right)} \tag{4.52}
\end{equation*}
$$

Now multiplying the error equation (4.44) by $2 \tau_{m}$, from (4.45)- (4.48) and (4.49) with $\delta=\frac{\varepsilon c_{0}}{4 C(L)}$, for the fully implicit scheme we obtain

$$
\begin{equation*}
\left\|e^{m}\right\|^{2}+\Gamma_{1}^{m} \leq\left\|e^{m-1}\right\|^{2}+\tau_{m} \alpha\left\|e^{m}\right\|^{2}+\Gamma_{2}^{m}+2 \tau_{m}\left(d_{t} e^{m}, \eta^{m}\right) \tag{4.53}
\end{equation*}
$$

where

$$
\begin{aligned}
\alpha & =8 \frac{C(L)^{2}}{c_{b} \varepsilon} \\
\Gamma_{1}^{m} & =\left\|e^{m}-e^{m-1}\right\|^{2}+\tau_{m} c_{0} \varepsilon\left\|e^{m}\right\|_{2, h}^{2}, \\
\Gamma_{2}^{m} & =c \tau_{m} \frac{C(L)^{2}}{c_{0} \varepsilon}\left\|\eta^{m}\right\|_{2, h}^{2}+c \tau_{m} \varepsilon\left|h^{r} u^{m}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}^{2}+\frac{c}{\varepsilon} \tau_{m}^{2} A_{0}^{m}+c \tau_{m} \frac{C(L)^{2}}{c_{0} \varepsilon} A_{2}^{m},
\end{aligned}
$$

with $A_{0}^{m}$ given by (4.18). Now replacing $m$ in (4.53) by $\ell$ and summing over it from 1 to $m=1, \ldots, M$ gives

$$
\begin{equation*}
\left\|e^{m}\right\|^{2}+\sum_{\ell=1}^{m} \Gamma_{1}^{\ell} \leq\left\|e^{0}\right\|^{2}+\alpha \sum_{\ell=1}^{m} \tau_{\ell}\left\|e^{\ell}\right\|^{2}+\sum_{\ell=1}^{m} \Gamma_{2}^{\ell}+2 \sum_{\ell=1}^{m} \tau_{\ell}\left(d_{t} e^{\ell}, \eta^{\ell}\right) \tag{4.54}
\end{equation*}
$$

We next consider the last term in (4.54). To the end, we recall the following discrete integration by parts formula:

$$
\sum_{\ell=1}^{m} \tau_{\ell}\left(d_{t} e^{\ell}, \eta^{\ell}\right)=-\sum_{\ell=1}^{m} \tau_{\ell}\left(e^{\ell-1}, d_{t} \eta^{\ell}\right)+\left(e^{m}, \eta^{m}\right)-\left(e^{0}, \eta^{0}\right)
$$

As seen before, $\tau_{l} d_{t} \eta^{\ell}=\eta^{\ell}-\eta^{(\ell-1)+}+\left[\eta^{\ell-1}\right]$, thus,

$$
\begin{array}{r}
\left|\tau_{\ell}\left(e^{\ell-1}, d_{t} \eta^{\ell}\right)\right| \leq \frac{1}{2}\left(\tau_{\ell}\left\|e^{\ell-1}\right\|^{2}+\tau_{\ell}^{-1}\left\|\eta^{\ell}-\eta^{(\ell-1)+}\right\|^{2}\right.  \tag{4.55}\\
\left.+\beta_{\ell}\left\|e^{\ell-1}\right\|^{2}+\beta_{\ell}^{-1}\left\|\left[\eta^{\ell-1}\right]\right\|^{2}\right)
\end{array}
$$

Now $\eta^{\ell}-\eta^{(\ell-1)+}=\int_{t_{\ell-1}}^{t_{\ell}} \eta_{t} d t$, thus,

$$
\begin{equation*}
\left\|\eta^{\ell}-\eta^{(\ell-1)+}\right\|^{2} \leq \tau_{\ell} \int_{t_{\ell-1}}^{t_{\ell}}\left|h^{r} u_{t}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}^{2} d t \tag{4.56}
\end{equation*}
$$

Furthermore,

$$
\begin{equation*}
\left|\left(e^{m}, \eta^{m}\right)+\left(e^{0}, \eta^{0}\right)\right| \leq \frac{1}{4}\left\|e^{m}\right\|^{2}+\left\|\eta^{m}\right\|^{2}+\frac{1}{2}\left\|e^{0}\right\|^{2}+\frac{1}{2}\left\|\eta^{0}\right\|^{2} \tag{4.57}
\end{equation*}
$$

Using these estimates in (4.54), we obtain

$$
\begin{align*}
& \left(1-2 \alpha \tau_{m}\right)\left\|e^{m}\right\|^{2}+\sum_{\ell=1}^{m} \Gamma_{1}^{\ell} \leq c\left(\left\|e^{0}\right\|^{2}+\left\|\eta^{0}\right\|^{2}+\left\|\eta^{m}\right\|^{2}\right)+\sum_{\ell=1}^{m-1} \gamma_{\ell}\left\|e^{\ell}\right\|^{2} \\
&  \tag{4.58}\\
& \quad+\sum_{\ell=1}^{m}\left(\Gamma_{2}^{\ell}+\beta_{m}^{-1}\left\|\left[\eta^{\ell-1}\right]\right\|^{2}+\tau_{\ell} \int_{t_{\ell-1}}^{t_{\ell}}\left|h^{r} u_{t}\right|_{H^{r}\left(\mathcal{T}_{h}^{m}\right)}^{2} d t\right)
\end{align*}
$$

where $\gamma_{\ell}=(2 \alpha+1) \tau_{\ell}+\beta_{\ell}$. Note that $\left\|e^{0}\right\|^{2},\left\|\eta^{0}\right\|^{2},\left\|\eta^{m}\right\|^{2}$ are $O\left(h^{r}\right)$ terms while $\sum_{\ell=1}^{m} \Gamma_{2}^{\ell}=O\left(\tau^{2}+h^{r-2}\right)$. Thus, a discrete version of Gronwall's inequality readily implies (4.43). In particular, the choice of the $\beta_{m}$ terms and their contribution to the error are as in Theorem 4.1. Furthermore, the implicit-explicit scheme is handled in the same, and so is the the transition from $f_{L}$ to $f$. The details are omitted. The proof is now complete.

## 5. Concluding remarks

In this paper we have presented fully discrete approximations for the CahnHilliard equation arising from phase transitions in materials science. The discontinuous Galerkin method is used in space, while the time stepping is effected by the backward Euler method. The schemes allow changing the spatial mesh as well as the time step size at every time level. One of the methods analyzed is of an implicit-explicit type, whereby the nonlinear term is rendered explicit via a simple extrapolation from the previous level. While yielding the same convergence rates as the fully implicit version, also analyzed in this paper, this method offers the advantage of solving only one linear system at a given time level. Furthermore, the coefficient matrix remains the same unless the spatial mesh is changed. Energy as
well as $L^{\infty}\left(L^{2}\right)$ a priori estimates are shown. These estimates are optimal provided that the number of times the spatial mesh is changed is bounded independently of the total number of time steps taken.

The extension of the analysis to higher order methods in time using multistep or implicit Runge-Kutta methods should be rather straightforward. There are other more interesting directions in which this work might be further pursued. Indeed, the constant appearing in the error estimates depends exponentially on the final time and also on $\epsilon^{-1}$. This is unfortunately, with very few exceptions, the state of the art in the error analysis of nonlinear evolution equations. In [23], this dependence was shown to be of a polynomial, rather than exponential, type for a mixed finite element approximation using continuous elements. Such a result in this context would also be quite significant since there is great interest in the sharp interface limit (as $\epsilon \rightarrow 0$ ) of the Cahn-Hilliard equation which is known as the Hele-Shaw problem. Another worthwhile endeavor would be one that would fully take advantage of the flexibility offered by the discontinuous Galerkin method, to obtain a posteriori error estimates and to construct adaptive algorithms based on them.

## Acknowledgment

The authors would like to thank the anonymous referee for carefully reading the paper and for his valuable suggestions.

## References

[1] R. Adams Sobolev Spaces Academic Press, New York, 1975. MR0450957(56:9247)
[2] 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. MR1308851 (97b:35174)
[3] D. Arnold, An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Anal., (19), 1982, pp. 742-760. MR664882 (83f:65173)
[4] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini, Discontinuous Galerkin methods for elliptic problems. Proceedings of the International Symposium on the discontinuous Galerkin method (B. Cockburn, G.E. Karniadakis, C.-W. Shu Eds.), Springer Lecture Notes in Computational Science and Engineering, vol. 11, 2000, pp. 89-101.
[5] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal., (39), 2001, pp. 1749-1779. MR 1885715 (2002k:65183)
[6] G. A. Baker, Finite element methods for elliptic equations using nonconforming elements. Math. Comp., (31), 1977, pp. 44-59. MR0431742 (55:4737)
[7] G. A. Baker, W. N. Jureidini and O. A. Karakashian, Piecewise solenoidal vector fields and the Stokes problems. SIAM J. Numer. Anal., (27), 1990, pp. 1466-1485. MR1080332 (91m:65246)
[8] J. H. Bramble, T. Dupont, and V. Thomée, Projection methods for Dirichlet's problem in approximating polygonal domains with boundary-value corrections, Math. Comp., (26), 1972, pp.869-879. MR0343657(49:8397)
[9] S. Brenner and R. Scott, The Mathematical Theory of Finite Element Methods, SpringerVerlag, New York, 1994. MR1278258(95f:65001)
[10] 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.
[11] P. G. Ciarlet, The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam, 1978. MR0520174 (58:25001)
[12] B. Cockburn, G. E. Karniadakis, C.-W. Shu, Discontinuous Galerkin Methods, Theory, Computation, and Applications, Springer Lecture Notes in Computational Science and Engineering, vol. 11, Springer-Verlag, 2000. MR1842160 (2002b:65004)
[13] B Cockburn and C. -W. Shu, The local discontinuous Galerkin method for convectiondiffusion systems, SIAM J. Numer. Anal., (35), 1998, pp. 2440-2463. MR1655854(99j:65163)
[14] J. Douglas, Jr. and T. Dupont, Galerkin methods for parabolic problems, SIAM J. Numer. Anal., (7), 1970, pp. 575-626. MR 0277126 (43:2863)
[15] J. Douglas, Jr. and T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, Lecture Notes In Physics 58, Springer Verlag, Berlin, 1976. MR0440955 (55:13823)
[16] Q. Du and R. A. Nicolaides, Numerical analysis of a continuum model of phase transition, SIAM J. Numer. Anal., (28), 1991, pp. 1310-1322. MR1119272 (92h:65166)
[17] C. M. Elliott and D. A. French, A nonconforming finite-element method for the twodimensional Cahn-Hilliard equation, SIAM J. Numer. Anal., (26), 1989, pp. 884-903. MR1005515 (90k:65163)
[18] C. M. Elliott, D. A. French, and F. A. Miller, A second order splitting method for the CahnHilliard equation, Numer. Math. (54), 1989, pp. 575-590. MR978609 (90k:65173)
[19] C. M. Elliott and S. Zheng, On the Cahn-Hilliard equation, Arch. Rational Mech. Anal., (96), 1986, pp. 339-357. MR855754(87k:80007)
[20] X. Feng, Fully discrete finite element approximations of the Navier-Stokes-Cahn-Hilliard diffuse interface model for two-phase fluid flows, SIAM J. Numer. Anal. (44), 2006, pp. 1049-1072. MR2231855
[21] X. Feng and O. A. Karakashian, Two-level nonoverlapping additive Schwarz methods for a discontinuous Galerkin approximation of the biharmonic problem, J. Scient. Comput. (22), 2005, pp. 299-324. MR2142199 (2006b:65156)
[22] 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. MR2101784 (2005m:65191)
[23] 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. MR 2126141 (2005j:35098)
[24] X. Feng and H. Wu, A posteriori error estimates and an adaptive finite element algorithm for the Cahn-Hilliard equation and the Hele-Shaw flow, M2AN Math. Model. Numer. Anal. (submitted).
[25] D. A. French and S. Jensen, Long-time behaviour of arbitrary order continuous time Galerkin schemes for some one-dimensional phase transition problems, IMA J. Numer. Anal., (14), 1994, pp. 421-442. MR 1283945 (95d:65083)
[26] P. Grisvard, Singularities in Boundary Value Problems, Research Notes in Applied Mathematics, Vol. 22, MASSON and Springer-Verlag, 1992. MR1173209 (93h:35004)
[27] P. Houston, E. Süli and C. Schwab, Discontinuous hp-finite element methods for advectiondiffusion problems, SIAM J. Numer. Anal. (39), 2002, pp. 2133-2163. MR1897953 (2003d:65108)
[28] D. Jacqmin, Calculation of two-phase Navier-Stokes flows using phase-field modeling, J. Comp. Phys., (115), 1999, pp. 96-127. MR1716497 (2000g:76088)
[29] O. A. Karakashian and W. N. Jureidini, A nonconforming finite element method for the stationary Navier-Stokes Equations. SIAM J. Numer. Anal., (35), 1998, pp. 93-120. MR1618436 (99d:65320)
[30] O. A. Karakashian, C. Makridakis, Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equations, Math. Comp. (74), 2005, pp. 85-102 MR2085403 (2005g:65147)
[31] C. Liu and J. Shen, A phase field model for the mixture of two incompressible fluids and its approximation by a Fourier-spectral method, Physica D, (179), 2003, pp. 211-228. MR 1984386 (2004j:35233)
[32] G. B. McFadden, Phase field models of solidification, Contemporary Mathematics, (295), 2002, pp. 107-145. MR 1940624 (2003k:80004)
[33] I. Mozolevski and E. Süli, A priori error analysis for the hp-version of the discontinuous Galerkin finite element method for the biharmonic equation, Comput. Meth. Appl. Math. (3), 2003, pp. 596-607. MR2048235 (2005c:65106)
[34] J.T. Oden and C.E. Baumann, A conservative DGM for convection-diffusion and NavierStokes problems, Proceedings of the International Symposium on the discontinuous Galerkin method (B. Cockburn, G. E. Karniadakis, C.-W. Shu Eds.), Springer Lecture Notes in Computational Science and Engineering, vol. 11, 2000, pp. 179-196. MR 1842173 (2002d:65128)
[35] M. F. Wheeler, A priori $L^{2}$ error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal., (10), 1973, pp. 723-759. MR0351124|(50:3613)
[36] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM J. Numer. Anal., (15), 1978, pp. 152-161 MR0471383 (57:11117)

Department of Mathematics, The University of Tennessee, Knoxville, Tennessee 37996

E-mail address: xfeng@math.utk.edu
Department of Mathematics, The University of Tennessee, Knoxville, Tennessee 37996

E-mail address: ohannes@math.utk.edu


[^0]:    Received by the editor February 1, 2006 and, in revised form, August 31, 2006.
    2000 Mathematics Subject Classification. Primary 65M15, 65M60, 74N20.
    Key words and phrases. Biharmonic equation, Cahn-Hilliard equation, discontinuous Galerkin methods, dynamic meshes, error estimates.

    The work of the first author was partially supported by the NSF grant DMS-0410266.
    The work of the second author was partially supported by the NSF grant DMS-0411448.

