Numer. Math.
DOI 10.1007/s00211-012-0473-9

Numerische Mathematik

A mixed finite element method for thin film epitaxy

Wenbin Chen · Yanqiu Wang

Received: 20 July 2011 / Revised: 25 April 2012

Abstract We present a mixed finite element method for the thin film epitaxy problem. Comparing to the primal formulation which requires C^2 elements in the discretization, the mixed formulation only needs to use C^1 elements, by introducing proper dual variables. The dual variable in our method is defined naturally from the nonlinear term in the equation, and its accurate approximation will be essential for understanding the long-time effect of the nonlinear term. For time-discretization, we use a backward-Euler semi-implicit scheme, which involves a convex–concave decomposition of the nonlinear term. The scheme is proved to be unconditionally stable and its convergence rate is analyzed.

Mathematics Subject Classification (2000) 35K55 · 35Q99 · 65M12 · 65M60

1 Introduction

Molecular beam epitaxy (MBE) [11,12] is a technology of depositing high-purity crystalline films with atomic thicknesses onto the surface of a base material. One distinguishing feature of MBE is the slow deposition rate of atoms or molecules, which allows the thin film on surface to grow epitaxially, or in other words, to grow as organized high-quality crystal. In this process, it is essential to have precise control on the surface morphology during epitaxial growth. This requires mathematical modeling, over multiple temporal and spatial scales, of particle adsorption, desorption, surface

W. Chen
School of Mathematical Sciences, Fudan University, Shanghai, China
e-mail: wbchen@fudan.edu.cn

Y. Wang (✉)
Department of Mathematics, Oklahoma State University, Stillwater, OK, USA
e-mail: yqwang@math.okstate.edu

Published online: 15 June 2012

diffusion, and step dynamics (the Ehrlich–Schwoebel barrier). Many different models, describing part or all of the above-mentioned physical phenomena, have been developed. Generally speaking, these models can be classified into three categories. The *atomistic models* [13,26,38] describe molecular dynamics using kinetic Monte Carlo methods. However, their applicability is limited due to high computational costs. The *continuum models* ([28,29,32,33,36,37,42,45] and references therein) are based on partial differential equations and the conservation of mass. They are able to capture large scale features of the crystal growth, and hence are interesting to physicists and mathematicians as well. There are also the *hybrid models* [9,22] that seek a compromise between atomistic and continuum models.

Here we consider a continuum model. Although the continuum model contains many inherent simplifications and heuristics, it can still provide a unique insight into the long-term evolution of the physical problem, especially into certain types of instabilities during the epitaxial growth. For decades, there have been large amount of research work on building the continuum model for thin film epitaxial growth. Most of them are conducted by physicists. Only in recent years, there has been an emerging trend of mathematician's involvement in this area. Most of these mathematical research has been focused on the existence, uniqueness and regularity of the solution to different types of governing evolution equations for MBE.

In 2002, Blömker and Gugg [6] have proved the global existence of the solution to a solid-on-solid model equation derived from [41]. The proof is based on Galerkin approximation and a priori estimates, using techniques similar to the proof of 2D Navier–Stokes equations. Later, Hoppe and Nash [24] proposed a combined spectral element/finite element approach for Blömker and Gugg's model. In 2003, King, Stein and Winkler [27] have studied the existence, uniqueness and regularity of the fourth-order governing equation proposed by Ortiz, Repetto and Si [36]. Their proof is based on an asymptotic analysis. However, the continuum model that probably has received most of the attention from mathematicians is a simplification with linearized surface diffusion [28,30–33]. In [28], Kohn and Yan proved there is an upper bound of the averaged coarsening rate for this model with finite Ehrlich–Schwoebel barrier and with slope selection. Another important work was done by Li and Liu [32], in which they have proved the well-posedness of the model for finite Ehrlich–Schwoebel barrier, with or without slope selection. The proof is based on a Galerkin spectral approximation and its a priori bounds. Numerical results using this spectral method are presented. Also given in [32] is the regularity of the global solution, which lays the foundation for further numerical study of the model problem. For the case of finite Ehrlich–Schwoebel barrier without slope selection, Li and Liu in another paper [33] has obtained two main theoretical results. First, by using the energy method and the convexity argument, they have derived the bounds for several important physical quantities including the interface width, average slope and average energy. These theoretical predictions agree with heuristic arguments. Second, by using the perturbation theory, they have shown that the system evolves in such a way that it always stays close to a sequence of periodic equilibria. In [30,31], Li has made further progress by generalizing the above results to the case of infinite Ehrlich–Schwoebel barriers and also for higher-order surface diffusion.

Numerical schemes for the simplified model problem proposed in [32] has been studied in [10,43–45], for the finite Ehrlich–Schwoebel barrier case, either with or without slope selection. All these previous research are based on the primal formulation. In [10], an energy-stable semi-implicit scheme, which has linear implicit parts, was developed for the without-slope-selection case only. In [43], an unconditionally stable semi-implicit scheme has been developed for both with and without slope selection cases. In [44], a fully implicit, stable scheme was analyzed for the without-slope-selection case only. And in [45], the authors studied the time-stability of the large time-stepping method. We mention that for the semi-implicit schemes presented in [10,43], a convex–concave decomposition is the key to the time-stability analysis. Indeed, we will also use this technique in our numerical method, and details shall be given later.

In this paper, we consider a mixed finite element method for the model problem, for the finite Ehrlich–Schwoebel barrier case and either with or without slope selection, as presented in [32]. Let Ω be a rectangular domain and $u(\boldsymbol{x}, t)$ be the height function of the thin film. The thin film epitaxy problem defined on $\Omega \times (0, T]$ can be written as

$$\begin{aligned}\partial_t u &= -\delta \Delta^2 u + \nabla \cdot \nabla_{\mathbb{F}} G(\nabla u) \\ &= \nabla \cdot [\nabla_{\mathbb{F}} G(\nabla u) - \delta \nabla \Delta u]\,,\end{aligned} \tag{1}$$

where δ is a positive constant, $\nabla_{\mathbb{F}}$ is the Fréchet gradient, and $G(\nabla u)$ is defined by

$$G(\nabla u) = \begin{cases} -\frac{1}{2}\ln(1 + |\nabla u|^2) & \text{without slope selection,} \\ \frac{1}{4}(|\nabla u|^2 - 1)^2 & \text{with slope selection.} \end{cases}$$

It is easy to check that their Fréchet gradients are, respectively

$$\nabla_{\mathbb{F}} G(\nabla u) = \begin{cases} -\frac{\nabla u}{1+|\nabla u|^2} & \text{without slope selection,} \\ (|\nabla u|^2 - 1)\nabla u & \text{with slope selection.} \end{cases}$$

Throughout the paper, we adopt the convention that a bold Latin or Greek character denotes a vector. Let $\boldsymbol{n}$ be the unit outward normal on $\partial\Omega$. To close the problem, we impose the following initial and boundary conditions. At $t = 0$, let $u = u_0$. Two different type of boundary conditions will be considered:

1. Dirichlet boundary condition. Let $u = \frac{\partial u}{\partial \boldsymbol{n}} = 0$ on $\partial\Omega$ for all time t.
2. Periodic boundary condition, where u is Ω-periodic for all time t. Since u is unique up to a constant, it is convenient to set it to be mean value zero.

Obviously, the initial condition u_0 should satisfy the same boundary condition for compatibility.

We adopt the usual notation $H^s(\Omega)$ for the Sobolev space with index s, equipped with the norm $\|\cdot\|_{H^s(\Omega)}$ and sometimes also the semi-norm $|\cdot|_{H^s(\Omega)}$. When $s = 0$, $H^0(\Omega)$ coincides with $L^2(\Omega)$, and for simplicity of the notation, we suppress the subscript in $\|\cdot\|_{L^2(\Omega)}$ and denote the norm by $\|\cdot\|$. Denote $(\cdot,\ \cdot)$ to be the L^2 inner-product on Ω. Define $L^p(0, T; H^s(\Omega))$, $1 \leq p \leq \infty$, to be the space of functions which are H^s in space and L^p in time. Finally, notice that all these notations can easily be extended to vector functions, by using product spaces. For convenience, when it is not ambiguous, some notations for product spaces will appear the same as those for a single space. For example, we write $\|\nabla u\|_{H^1(\Omega)}$ instead of $\|\nabla u\|_{(H^1(\Omega))^2}$, and $\|\nabla u\|_{L^\infty(0,T;H^1(\Omega))}$ instead of $\|\nabla u\|_{L^\infty(0,T;(H^1(\Omega))^2)}$.

For the periodic boundary problem, it has been proved [32] that for $u_0 \in H^s(\Omega), s \geq 2$,

$$\begin{aligned}
&\text{the initial-boundary value problem of (1) has a unique solution } u,\\
&u \in L^\infty(0,T;H^s(\Omega)) \cap L^2(0,T;H^{s+2}(\Omega)),\\
&\partial_t u \in L^2(0,T;H^{s-2}(\Omega)).
\end{aligned} \tag{2}$$

Such a result has not yet been proved for the Dirichlet boundary problem. However, the analysis in this paper will be based on the existence and regularity assumption (2), which is known to be true for at least the periodic boundary problem.

An important observation is that, the operator G can be decomposed into a convex (+) and a concave (−) part [43]:

$$G(\boldsymbol{w}) = G_+(\boldsymbol{w}) + G_-(\boldsymbol{w}),$$

such that the Fréchet Hessian $\nabla^2_{\mathbb{F}} G_+$ and $\nabla^2_{\mathbb{F}} G_-$ are positively and negatively semi-definite, respectively. Moreover, similar to [43], we assume both $\nabla^2_{\mathbb{F}} G_+(\boldsymbol{w})$ and $\nabla^2_{\mathbb{F}} G_-(\boldsymbol{w})$ have at most polynomial growth in $\boldsymbol{w}$, that is, there exists a positive integer m such that

$$|\nabla^2_{\mathbb{F}} G_+(\boldsymbol{w})| + |\nabla^2_{\mathbb{F}} G_-(\boldsymbol{w})| \leq C_G(1 + |\boldsymbol{w}|^m), \tag{3}$$

where C_G is a positive constant independent of $\boldsymbol{w}$, $|\boldsymbol{w}|$ is the vector 2-norm and $|\nabla^2_{\mathbb{F}} G_+(\boldsymbol{w})|$, $|\nabla^2_{\mathbb{F}} G_-(\boldsymbol{w})|$ are matrix 2-norms. An example of such a decomposition is to simply set [43]

$$G_-(\boldsymbol{w}) = -\frac{1}{2}|\boldsymbol{w}|^2, \quad G_+(\boldsymbol{w}) = G(\boldsymbol{w}) - G_-(\boldsymbol{w}). \tag{4}$$

Then $\nabla^2_{\mathbb{F}} G_+$ and $\nabla^2_{\mathbb{F}} G_-$ are positively and negatively semi-definite, respectively, and satisfy Inequality (3). Furthermore, it is easy to check that for both with and without slope selection, the decomposition defined in (4) satisfies

$$G_+(\boldsymbol{w}) \geq 0 \tag{5}$$

and

$$\begin{aligned}
&(\nabla_{\mathbb{F}} G(\boldsymbol{w}) - \nabla_{\mathbb{F}} G(\boldsymbol{\varphi}), \boldsymbol{w} - \boldsymbol{\varphi})\\
&\quad = (\nabla_{\mathbb{F}} G_+(\boldsymbol{w}) - \nabla_{\mathbb{F}} G_+(\boldsymbol{\varphi}), \boldsymbol{w} - \boldsymbol{\varphi}) + (\nabla_{\mathbb{F}} G_-(\boldsymbol{w}) - \nabla_{\mathbb{F}} G_-(\boldsymbol{\varphi}), \boldsymbol{w} - \boldsymbol{\varphi})\\
&\quad \geq (\nabla_{\mathbb{F}} G_-(\boldsymbol{w}) - \nabla_{\mathbb{F}} G_-(\boldsymbol{\varphi}), \boldsymbol{w} - \boldsymbol{\varphi})\\
&\quad = -\|\boldsymbol{w} - \boldsymbol{\varphi}\|^2.
\end{aligned} \tag{6}$$

The convex–concave decomposition defined above is essential in developing stable numerical schemes for problem (1). In the time discretization, the convex term will be approximated implicitly and the concave term explicitly. Such technique has been widely used in the time-discretization for Cahn–Hilliard equations [14,15,17–21]. For the thin film epitaxy problem, the use of convex–concave decomposition was first proposed in [43] for discretizing the primal formulation of problem (1). In this paper, we shall combine this decomposition with the mixed finite element method, and develop stable numerical schemes.

Although many ideas are borrowed from the previous research on primal finite element methods for the thin file epitaxy problem, we would like to point out that, the analysis of mixed finite element methods is quite different, due to its different finite element space settings. The time-stability and convergence analysis in this paper is relatively complicated. We are not sure whether an easier proof is possible or not, or whether a better convergence rate estimate can be achieved. The main contribution of this paper lies in that, it is the first in developing a mixed finite element method for thin film epitaxy model (1). New schemes, ideas and tools are introduced. Notice that the model problem (1) is essentially a fourth-order equation. A mixed formulation will break the fourth-order equation into more than one lower-order equations, hence avoiding the use of C^1 conforming or non-conforming finite elements in the numerical approximation. Also, as it will be explained in the next section, our mixed method involves a dual variable ∇u, which provides a natural and accurate approximation to the nonlinear term $G(\nabla u)$.

The paper is organized as follows. In Sect. 2, we introduce a mixed formulation for problem (1). Its finite element discretization, together with its time-stability, will be discussed in Sect. 3. Finally, in Sect. 4, we analyze the convergence rate of the discrete scheme.

2 The mixed formulation

In this section, we consider a mixed formulation for Eq. (1). Equation (1) is essentially a time-dependent fourth-order problem with a nonlinear second order term. Let us first recall the mixed formulation for the biharmonic problem

$$\Delta^2 u = f.$$

One popular method [39] is to define $w = \Delta u$. Then the biharmonic problem can be rewritten into

$$\begin{cases} w - \Delta u = 0, \\ \Delta w = f. \end{cases}$$

Another [23,25] is to define $\boldsymbol{w} = \nabla u$, $\boldsymbol{\lambda} = \Delta \boldsymbol{w}$ and it gives

$$\begin{cases} \boldsymbol{w} - \nabla u = 0, \\ \boldsymbol{\lambda} - \Delta \boldsymbol{w} = 0, \\ \nabla \cdot \boldsymbol{\lambda} = f, \end{cases}$$

where the last equation follows from $\nabla \cdot (\Delta \boldsymbol{w}) = \Delta(\nabla \cdot \boldsymbol{w}) = \Delta^2 u$. This mixed formulation is similar to the reduced integration method proposed in [25,35] for the biharmonic problem, which is also a popular numerical method for approximating the Reissner-Mindlin plate problems [1–4,7,8,16]. Indeed, we will use some existing theoretical results from these works. However, our analysis shall concentrate on the nonlinear well-posedness and the time-stability issue.

Since the nonlinear term in Eq. (1) depends solely on ∇u, it will be natural to use the second mixed formulation. Indeed, by defining $\boldsymbol{w} = \nabla u$ and $\boldsymbol{\lambda} = \delta \Delta \boldsymbol{w} - \nabla_{\mathbb{F}} G(\boldsymbol{w})$, Equation (1) can be rewritten into

$$\begin{cases} -\delta \Delta \boldsymbol{w} + \nabla_{\mathbb{F}} G(\boldsymbol{w}) + \boldsymbol{\lambda} = \boldsymbol{0}, \\ \partial_t u + \nabla \cdot \boldsymbol{\lambda} = 0, \\ \boldsymbol{w} - \nabla u = 0. \end{cases} \tag{7}$$

It is not hard to see that for the Dirichlet boundary problem, $\boldsymbol{w} = \boldsymbol{0}$ on the boundary, and for the periodic boundary problem, $\boldsymbol{w} = \nabla u$ is also periodic and each of its entries has mean value zero in Ω. Let $\dot{C}^{\infty}_{per}(\Omega)$ be the space of infinitely differentiable periodic functions with mean value zero in Ω. Define $H^1_{per}(\Omega)$ to be the closure of $\dot{C}^{\infty}_{per}(\Omega)$ in $H^1(\Omega)$. Denote spaces

$$S = \begin{cases} H^1_0(\Omega) & \text{for the Dirichlet boundary problem,} \\ H^1_{per}(\Omega) & \text{for the periodic boundary problem,} \end{cases}$$

and $\boldsymbol{Q} = (L^2(\Omega))^2$. We have the Poincaré inequality in S, that is, there exists a positive constant C such that

$$\|v\| \leq C\|\nabla v\| \quad \text{for all } v \in S.$$

By testing system (7) with $(v, \boldsymbol{\varphi}, \boldsymbol{\mu}) \in S \times S^2 \times \boldsymbol{Q}$, we end up with the following weak formulation: *find* $(u, \boldsymbol{w}, \boldsymbol{\lambda}) \in L^2(0, T; S) \times L^2(0, T; S^2) \times L^2(0, T; \boldsymbol{Q})$ *such that*

$$\begin{cases} \delta(\nabla \boldsymbol{w}, \nabla \boldsymbol{\varphi}) + (\nabla_{\mathbb{F}} G(\boldsymbol{w}), \boldsymbol{\varphi}) + (\boldsymbol{\lambda}, \boldsymbol{\varphi}) = 0 & \text{for all } \boldsymbol{\varphi} \in S^2, \\ (\partial_t u, v) - (\boldsymbol{\lambda}, \nabla v) = 0 & \text{for all } v \in S, \\ -(\boldsymbol{w} - \nabla u, \boldsymbol{\mu}) = 0, & \text{for all } \boldsymbol{\mu} \in \boldsymbol{Q}, \end{cases} \tag{8}$$

almost everywhere for $t \in (0, T]$. Notice that the weak solution should satisfy the initial condition

$$u|_{t=0} = u_0, \quad \boldsymbol{w}|_{t=0} = \nabla u_0, \quad \boldsymbol{\lambda}|_{t=0} = \delta \Delta(\nabla u_0) - \nabla_{\mathbb{F}} G(\nabla u_0).$$

Hence by the compatibility requirement, the entire mixed formulation is well-posed only when

$$u_0 \in H^2(\Omega) \quad \text{and} \quad \delta \Delta(\nabla u_0) - \nabla_{\mathbb{F}} G(\nabla u_0) \in L^2(\Omega). \tag{9}$$

This is mainly because of the introduction of the auxiliary variable $\boldsymbol{\lambda}$. For simplicity, we assume $u_0 \in H^3(\Omega)$ in this paper, which is enough to guarantee (9).

Theorem 2.1 *Given $u_0 \in H^3(\Omega)$, system (8) has a unique weak solution.*

Proof The existence of the solution follows from the existence and regularity assumptio n (2). By defining $\boldsymbol{w} = \nabla u$ and $\boldsymbol{\lambda} = \delta \Delta \boldsymbol{w} - \nabla_{\mathbb{F}} G(\boldsymbol{w})$, one immediately ends up with a weak solution for (8).

The uniqueness of the solution follows from a stability result: let $u_0^{(i)} \in H^3(\Omega)$, $i = 1, 2$, be two initial data, and $(u^{(i)}, \boldsymbol{w}^{(i)}, \boldsymbol{\lambda}^{(i)})$ be the corresponding weak solutions. Then

$$\|u^{(1)} - u^{(2)}\|_{L^\infty(0,T;L^2(\Omega))} + \|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\|_{L^2(0,T;H^1(\Omega))} \le C(\delta, T)\|u_0^{(1)} - u_0^{(2)}\|_{L^2(\Omega)},$$

where $C(\delta, T)$ is a positive constant. Next, we shall prove this stability result.

Denote $\tilde{u} = u^{(1)} - u^{(2)}$, $\tilde{\boldsymbol{w}} = \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}$ and $\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}$. Clearly,

$$\begin{cases} \delta(\nabla \tilde{\boldsymbol{w}}, \nabla \boldsymbol{\varphi}) + (\nabla_{\mathbb{F}} G(\boldsymbol{w}^{(1)}) - \nabla_{\mathbb{F}} G(\boldsymbol{w}^{(2)}), \boldsymbol{\varphi}) + (\tilde{\boldsymbol{\lambda}}, \boldsymbol{\varphi}) = 0, \\ (\partial_t \tilde{u}, v) - (\tilde{\boldsymbol{\lambda}}, \nabla v) = 0, \\ -(\tilde{\boldsymbol{w}} - \nabla \tilde{u}, \boldsymbol{\mu}) = 0. \end{cases}$$

By setting $\boldsymbol{\varphi} = \tilde{\boldsymbol{w}}$, $v = \tilde{u}$ and $\boldsymbol{\mu} = \tilde{\boldsymbol{\lambda}}$, and adding up all three equations, one gets

$$\frac{1}{2}\frac{d}{dt}\|\tilde{u}\|^2 + \delta\|\nabla \tilde{\boldsymbol{w}}\|^2 + (\nabla_{\mathbb{F}} G(\boldsymbol{w}^{(1)}) - \nabla_{\mathbb{F}} G(\boldsymbol{w}^{(2)}), \tilde{\boldsymbol{w}}) = 0.$$

By the lower bound (6), we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\tilde{u}\|^2 + \delta\|\nabla \tilde{\boldsymbol{w}}\|^2 &\le (\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{w}}) = (\nabla \tilde{u}, \tilde{\boldsymbol{w}}) \\ &= -(\tilde{u}, \nabla \cdot \tilde{\boldsymbol{w}}) \le \|\tilde{u}\| \, (2\|\nabla \tilde{\boldsymbol{w}}\|) \\ &\le \frac{2}{\delta}\|\tilde{u}\|^2 + \frac{\delta}{2}\|\nabla \tilde{\boldsymbol{w}}\|^2. \end{aligned}$$

The stability result then follows from the Gronwall's inequality. This completes the proof of the theorem. □

3 Finite element discretization

We use the rectangular finite element spaces defined in [23] to discretize the mixed problem (8). Given a quasi-uniform rectangular mesh $\mathcal{T}_h$ in Ω with characteristic mesh size h. Define $S_h \in S$ and $\boldsymbol{Q}_h \in \boldsymbol{Q}$ as follows:

$$\begin{aligned} S_h &= \{v \in S, v|_K \in Q_1(K) \text{ for all } K \in \mathcal{T}_h\}, \\ \boldsymbol{Q}_h &= \{\boldsymbol{\mu} \in \boldsymbol{Q}, \boldsymbol{\mu}|_K = \begin{pmatrix} a + by \\ c + dx \end{pmatrix} \text{ for all } K \in \mathcal{T}_h\}, \end{aligned}$$

where $Q_1(K)$ is the space of bilinear polynomials on K. It is clear that $\nabla S_h \subset \boldsymbol{Q}_h$. Let $I_h : S \cap H^2(\Omega) \to S_h$ be the nodal value interpolation and $\boldsymbol{P}_h : \boldsymbol{Q} \to \boldsymbol{Q}_h$ be the L^2 orthogonal projection. We have the following approximation properties [23]:

$$\begin{aligned} \|v - I_h v\| + h\|\nabla(v - I_h v)\| &\le Ch^2 |v|_{H^2(\Omega)} && \text{for all } v \in S \cap H^2(\Omega), \\ \|\boldsymbol{\mu} - \boldsymbol{P}_h \boldsymbol{\mu}\| &\le Ch|\boldsymbol{\mu}|_{H^1(\Omega)} && \text{for all } \boldsymbol{\mu} \in (H^1(\Omega))^2, \\ (\nabla(v - I_h v), \boldsymbol{\mu}_h) &\le Ch^2 |v|_{H^3(\Omega)} \|\boldsymbol{\mu}_h\| && \text{for all } v \in S \cap H^3(\Omega) \\ &&& \quad \text{and all } \boldsymbol{\mu}_h \in \boldsymbol{Q}_h, \end{aligned} \tag{10}$$

where $C > 0$ is a general constant independent of h.

Now we are able to introduce a fully-discrete scheme for the mixed problem (8). A convex-splitting semi-implicit scheme will be considered, whose main idea is to use implicit time discretization in G_+ and the fourth-order term, and to use explicit time discretization in G_-. Such an idea has been used in [14,15,17–21] for the Cahn–Hilliard flow, and in [43] for the primal formulation of the thin film epitaxy.

Denote $(u_h^n, \boldsymbol{w}_h^n, \boldsymbol{\lambda}_h^n) \in S_h \times S_h^2 \times \boldsymbol{Q}_h$ to be the approximation to the weak solution at time $t_n = n\Delta t$, the discrete problem for (8) can be written as:

$$\begin{cases} \delta(\nabla \boldsymbol{w}_h^{n+1}, \nabla \boldsymbol{\varphi}_h) + (\nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^{n+1}), \boldsymbol{\varphi}_h) + (\boldsymbol{\lambda}_h^{n+1}, \boldsymbol{\varphi}_h) & \\ \quad = -(\nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n), \boldsymbol{\varphi}_h) & \text{for all } \boldsymbol{\varphi}_h \in S_h^2, \\ \left(\frac{u_h^{n+1} - u_h^n}{\Delta t}, v_h\right) - (\boldsymbol{\lambda}_h^{n+1}, \nabla v_h) = 0 & \text{for all } v_h \in S_h, \\ \varepsilon(\boldsymbol{\lambda}_h^{n+1}, \boldsymbol{\mu}_h) - (\boldsymbol{w}_h^{n+1} - \nabla u_h^{n+1}, \boldsymbol{\mu}_h) = 0, & \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{Q}_h, \end{cases} \tag{11}$$

where $\varepsilon = O(h^2)$ is a penalty constant which is needed to ensure the solvability of the discrete problem [23,25]. Notice that given u_h^{n+1} and $\boldsymbol{w}_h^{n+1}$, $\boldsymbol{\lambda}_h^{n+1}$ is uniquely solvable from the third equation of (11). In other words, the third equation can be decoupled from the system.

Equation (11) is a stabilized formulation. In practice, any stabilized finite element spaces [23,25,35] for fourth-order elliptic equations can be used to discretize problem (8) and the discretization leads to system (11). There are also stable finite element spaces available for the mixed formulation of Reissner–Mindlin plate [1–4,7,8,16], which can be adopted for discretizing problem (8). However, here we prefer the stabilized finite elements, because system (11) can easily be reduced to a simple minimization problem, which will be discussed later. The drawback is that, the stabilization term limits the finite element approximation rate. In the future, researchers can work towards increasing the approximation rate or adopting stable finite elements.

Define functional

$$\mathcal{F}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda}) = \int_\Omega \Bigg(G_+(\boldsymbol{w}) + \frac{\delta}{2}|\nabla \boldsymbol{w}|^2 + \left[(\boldsymbol{w} - \nabla u)\cdot \boldsymbol{\lambda} - \frac{\varepsilon}{2}|\boldsymbol{\lambda}|^2\right] + \frac{1}{2\Delta t}|u|^2 + \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n)\cdot \boldsymbol{w} - \frac{1}{\Delta t} u_h^n u \Bigg)\, dx.$$

Then the Fréchet gradient $\nabla_{\mathbb{F}}\mathcal{F}^{n+1} = \mathbf{0}$, which can be written as

$$\begin{cases} \left[\mathcal{F}_u^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda})\right](v_h) = \frac{d}{dk}\mathcal{F}^{n+1}(u + kv_h, \boldsymbol{w}, \boldsymbol{\lambda})\big|_{k=0} = 0 & \text{for all } v_h \in S_h, \\ \left[\mathcal{F}_{\boldsymbol{w}}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda})\right](\boldsymbol{\varphi}_h) = \frac{d}{dk}\mathcal{F}^{n+1}(u, \boldsymbol{w} + k\boldsymbol{\varphi}_h, \boldsymbol{\lambda})\big|_{k=0} = 0 & \text{for all } \boldsymbol{\varphi}_h \in S_h^2, \\ [\mathcal{F}_{\boldsymbol{\lambda}}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda})](\boldsymbol{\mu}_h) = \frac{d}{dk}\mathcal{F}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda} + k\boldsymbol{\mu}_h)\big|_{k=0} = 0 & \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{Q}_h, \end{cases}$$

leads to exactly system (11) when being expanded. Indeed, given $(u_h^n, \boldsymbol{w}_h^n)$, the solution $(u_h^{n+1}, \boldsymbol{w}_h^{n+1}, \boldsymbol{\lambda}_h^{n+1})$ for system (11) can be characterized as the solution to the following saddle point problem

$$\min_{\substack{u \in S_h \\ \boldsymbol{w} \in S_h^2}} \max_{\boldsymbol{\lambda} \in \boldsymbol{Q}_h} \mathcal{F}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda}). \tag{12}$$

It is not hard to see that $\max_{\boldsymbol{\lambda}\in\boldsymbol{Q}_h} \mathcal{F}^{n+1}(u, \boldsymbol{w}, \boldsymbol{\lambda})$ is reached at $\boldsymbol{\lambda} = \boldsymbol{P}_h(\boldsymbol{w} - \nabla u)/\varepsilon$, hence the saddle problem (12) is also equivalent to the following minimization problem

$$\min_{\substack{u \in S_h \\ \boldsymbol{w} \in S_h^2}} F^{n+1}(u, \boldsymbol{w}) \tag{13}$$

where

$$F^{n+1}(u, \boldsymbol{w}) = \int_{\Omega} \Bigg(G_+(\boldsymbol{w}) + \frac{\delta}{2}|\nabla \boldsymbol{w}|^2 + \frac{1}{2\varepsilon}|\boldsymbol{P}_h(\boldsymbol{w} - \nabla u)|^2 + \frac{1}{2\Delta t}|u|^2 + \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n) \cdot \boldsymbol{w} - \frac{1}{\Delta t}u_h^n u \Bigg)\, dx.$$

Therefore, to prove the existence and uniqueness of the solution to problem (11), we only need to show that problem (13) has a unique solution. Indeed, we have the following theorem:

Theorem 3.1 *Given* $(u_h^n, \boldsymbol{w}_h^n)$*, the minimization problem* (13) *has a unique solution at* t_{n+1}.

Proof The minimization problem is an unconstrained convex optimization problem on finite dimensional spaces. According to the standard theory [5], we only need to prove the coercivity, which implies that $F^{n+1}(u, \boldsymbol{w})$ goes to infinity as $\|u\|_{H^1}$ or $\|\boldsymbol{w}\|_{H^1}$ goes to infinity, and the strict convexity of $F^{n+1}(u, \boldsymbol{w})$.

We first prove the coercivity of $F^{n+1}(u, \boldsymbol{w})$. Let c_1 be a positive constant such that

$$c_1\|\boldsymbol{w}\|^2 \le \frac{\delta}{2}\|\nabla \boldsymbol{w}\|^2 \quad \text{for all } \boldsymbol{w} \in S_h^2.$$

This is possible because of the Poincaré inequality. Then, by using the Schwarz inequality and the Young's inequality,

$$
\begin{aligned}
F^{n+1}(u, \boldsymbol{w}) \geq \int\limits_{\Omega} \Bigg(& G_{+}(\boldsymbol{w}) + \frac{\delta}{2}|\nabla \boldsymbol{w}|^2 + \frac{1}{2\varepsilon}|\boldsymbol{P}_h(\boldsymbol{w} - \nabla u)|^2 + \frac{1}{2\Delta t}|u|^2 \\
& - \frac{1}{2c_1}|\nabla_{\mathbb{F}} G_{-}(\boldsymbol{w}_h^n)|^2 - \frac{c_1}{2}|\boldsymbol{w}|^2 - \frac{1}{\Delta t}|u_h^n|^2 - \frac{1}{4\Delta t}|u|^2 \Bigg) dx \\
\geq \int\limits_{\Omega} & \left(\frac{\delta}{4}|\nabla \boldsymbol{w}|^2 + \frac{1}{2\varepsilon}|\boldsymbol{P}_h(\boldsymbol{w} - \nabla u)|^2 + \frac{1}{4\Delta t}|u|^2 - \beta \right) dx,
\end{aligned}
$$

where β is a constant depending only on $\boldsymbol{w}_h^n$ and u_h^n. Let $c_2 > 1$ be a constant satisfying

$$
\frac{c_2 - 1}{2\varepsilon}\|\boldsymbol{P}_h \boldsymbol{w}\|^2 \leq \frac{\delta}{8}\|\nabla \boldsymbol{w}\|^2 \quad \text{for all } \boldsymbol{w} \in S_h^2.
$$

Again, this is possible by the stability of the L^2 projection $\boldsymbol{P}_h$ and the Poincaré inequality. Clearly, c_2 is independent of the mesh size h. Then, since $\boldsymbol{P}_h \nabla u = \nabla u$ for all $u \in S_h$,

$$
\begin{aligned}
F^{n+1}(u, \boldsymbol{w}) \geq \int\limits_{\Omega} & \Bigg(\frac{\delta}{4}|\nabla \boldsymbol{w}|^2 + \frac{1}{2\varepsilon}(|\boldsymbol{P}_h \boldsymbol{w}|^2 - 2\boldsymbol{P}_h \boldsymbol{w} \cdot \nabla u + |\nabla u|^2) \\
& + \frac{1}{4\Delta t}|u|^2 - \beta \Bigg) dx, \\
\geq \int\limits_{\Omega} & \Bigg(\frac{\delta}{4}|\nabla \boldsymbol{w}|^2 + \frac{1}{2\varepsilon}\left((1 - c_2)|\boldsymbol{P}_h \boldsymbol{w}|^2 + \left(1 - \frac{1}{c_2}\right)|\nabla u|^2 \right) \\
& + \frac{1}{4\Delta t}|u|^2 - \beta \Bigg) dx, \\
\geq \int\limits_{\Omega} & \left(\frac{\delta}{8}|\nabla \boldsymbol{w}|^2 + \frac{c_2 - 1}{2\varepsilon c_2}|\nabla u|^2 + \frac{1}{4\Delta t}|u|^2 - \beta \right) dx.
\end{aligned}
$$

This completes the proof of coercivity.

Next, we prove that the functional $F^{n+1}(u, \boldsymbol{w})$ is strictly convex on $S_h \times S_h^2$, then the minimization problem (13) admits a unique solution. This can be done by showing that the Hessian of F^{n+1} is positively definite. Indeed, for any $(v, \boldsymbol{\varphi}) \in S_h \times S_h^2$,

$$
\begin{aligned}
\nabla_{\mathbb{F}} F^{n+1}(u, \boldsymbol{w}) \begin{pmatrix} v \\ \boldsymbol{\varphi} \end{pmatrix} &= \frac{d}{dk} F^{n+1}(u + kv, \boldsymbol{w} + k\boldsymbol{\varphi}) \Big|_{k=0} \\
= \int\limits_{\Omega} \Bigg[& \nabla_{\mathbb{F}} G_{+}(\boldsymbol{w}) \cdot \boldsymbol{\varphi} + \delta \nabla \boldsymbol{w} \cdot \nabla \boldsymbol{\varphi} + \frac{1}{\varepsilon}\boldsymbol{P}_h(\boldsymbol{w} - \nabla u) \cdot \boldsymbol{P}_h(\boldsymbol{\varphi} - \nabla v) \\
& + \frac{1}{\nabla t} u v + \nabla_{\mathbb{F}} G_{-}(\boldsymbol{w}_h^n) \cdot \boldsymbol{\varphi} - \frac{1}{\Delta t} u_h^n v \Bigg] dx.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\left[\nabla_{\mathbb{F}}^2 F^{n+1}(u,\boldsymbol{w})\begin{pmatrix} v\\ \boldsymbol{\varphi}\end{pmatrix}\right]\begin{pmatrix} v\\ \boldsymbol{\varphi}\end{pmatrix} = \frac{d}{dk}\left(\nabla_{\mathbb{F}} F^{n+1}(u+kv,\boldsymbol{w}+k\boldsymbol{\varphi})\begin{pmatrix} v\\ \boldsymbol{\varphi}\end{pmatrix}\right)\Bigg|_{k=0}\\
&\quad = \frac{d}{dk}\int\limits_{\Omega}\Bigg[\nabla_{\mathbb{F}} G_{+}(\boldsymbol{w}+k\boldsymbol{\varphi})\cdot\boldsymbol{\varphi} + \delta\nabla(\boldsymbol{w}+k\boldsymbol{\varphi})\cdot\nabla\boldsymbol{\varphi}\\
&\qquad\qquad + \frac{1}{\varepsilon}\boldsymbol{P}_h\left(\boldsymbol{w}-\nabla u + k(\boldsymbol{\varphi}-\nabla v)\right)\cdot\boldsymbol{P}_h(\boldsymbol{\varphi}-\nabla v)\\
&\qquad\qquad + \frac{1}{\nabla t}(u+kv)v + \nabla_{\mathbb{F}} G_{-}(\boldsymbol{w}_h^n)\cdot\boldsymbol{\varphi} - \frac{1}{\Delta t}u_h^n v\Bigg]\,dx\Bigg|_{k=0}\\
&\quad = \int\limits_{\Omega}\left[\boldsymbol{\varphi}\cdot\nabla_{\mathbb{F}}^2 G_{+}(\boldsymbol{w})\cdot\boldsymbol{\varphi} + \delta|\nabla\boldsymbol{\varphi}|^2 + \frac{1}{\varepsilon}|\boldsymbol{P}_h(\boldsymbol{\varphi}-\nabla v)|^2 + \frac{1}{\Delta t}v^2\right]dx\\
&\quad \geq \int\limits_{\Omega}\left[\delta|\nabla\boldsymbol{\varphi}|^2 + \frac{1}{\varepsilon}\left((1-c_3)|\boldsymbol{P}_h\boldsymbol{\varphi}|^2 + \left(1-\frac{1}{c_3}\right)|\nabla v|^2\right) + \frac{1}{\Delta t}v^2\right]dx\\
&\quad \geq \int\limits_{\Omega}\left[\frac{\delta}{2}|\nabla\boldsymbol{\varphi}|^2 + \frac{c_3-1}{\varepsilon c_3}|\nabla v|^2 + \frac{1}{\Delta t}v^2\right]dx\\
&\quad \geq \min\left(\frac{\delta}{2},\frac{c_3-1}{\varepsilon c_3}\right)\left(\|\nabla\boldsymbol{\varphi}\|^2 + \|\nabla v\|^2\right),
\end{aligned}
$$

where $c_3 > 1$ is a constant satisfying

$$
\frac{c_3-1}{\varepsilon}\|\boldsymbol{P}_h\boldsymbol{\varphi}\|^2 \leq \frac{\delta}{2}\|\nabla\boldsymbol{\varphi}\|^2 \quad \text{for all } \boldsymbol{\varphi}\in S_h^2.
$$

Finally, by applying the Poincaré inequality, we have shown that $F^{n+1}(u,\boldsymbol{w})$ is strictly convex. This completes the proof of the theorem. □

Next, we show that the numerical scheme is time-stable. Let $(u_h^n, \boldsymbol{w}_h^n, \boldsymbol{\lambda}_h^n)$ be the solution to problem (11). Define an "energy" functional

$$
\begin{aligned}
E^n &= \int\limits_{\Omega}\left(G(\boldsymbol{w}_h^n) + \frac{\delta}{2}|\nabla\boldsymbol{w}_h^n|^2 + \frac{1}{2\varepsilon}|\boldsymbol{P}_h(\boldsymbol{w}_h^n - \nabla u_h^n)|^2\right)dx\\
&= \int\limits_{\Omega}\left(G(\boldsymbol{w}_h^n) + \frac{\delta}{2}|\nabla\boldsymbol{w}_h^n|^2 + \frac{\varepsilon}{2}|\boldsymbol{\lambda}_h^n|^2\right)dx.
\end{aligned}
$$

Note that by definition, $G(\boldsymbol{w}_h^n)$ is always non-negative in the case with slope selection, but is negative in the case without slope selection. However, even when $G(\boldsymbol{w}_h^n)$ is negative, as long as E^n satisfies certain conditions, it can still be considered an "energy" functional and thus be used to prove the time-stability. We shall explain this in details below. For the case with slope selection, clearly,

$$E^n \geq \int\limits_{\Omega} \left(\frac{\delta}{2} |\nabla \boldsymbol{w}_h^n|^2 + \frac{\varepsilon}{2} |\boldsymbol{\lambda}_h^n|^2 \right) dx. \tag{14}$$

Now consider the case without slope selection. Using elementary Calculus, one can show that for any constant $c > 0$ and $x \geq 0$,

$$cx - \frac{1}{2}\ln(1+x) \geq \begin{cases} \frac{1}{2} - c + \frac{1}{2}\ln(2c) > \frac{1}{2}\ln(2c) & \text{for } 0 < c < \frac{1}{2} \\ 0 & \text{for } c \geq \frac{1}{2} \end{cases}.$$

Set $c = \delta/4$, then the "energy" functional E^n for the case without slope selection satisfies

$$\begin{aligned} E^n &\geq |\Omega| \min\left\{ \frac{1}{2}\ln\frac{\delta}{2}, 0 \right\} + \int\limits_{\Omega} \left(\frac{\delta}{4} |\nabla \boldsymbol{w}_h^n|^2 + \frac{\varepsilon}{2} |\boldsymbol{\lambda}_h^n|^2 \right) dx \\ &\geq \int\limits_{\Omega} \left(\frac{\delta}{4} |\nabla \boldsymbol{w}_h^n|^2 + \frac{\varepsilon}{2} |\boldsymbol{\lambda}_h^n|^2 \right) dx - C_\delta, \end{aligned} \tag{15}$$

where $|\Omega|$ is the measure of domain Ω, and C_δ is a non-negative constant that only depends on δ and Ω. We point out that such an observation has been used before in [43] to define a similar "energy" functional for thin film epitaxy in the primal formulation. Next, we prove that the "energy" functional E^n is non-increasing and thus the numerical scheme (11) is time-stable.

Theorem 3.2 *The energy functional E^n is non-increasing in time. Indeed,*

$$E^{n+1} \leq E^n - \frac{1}{2\Delta t} \|u_h^{n+1} - u_h^n\|^2. \tag{16}$$

Consequently,

$$\|u_h^n\|_{H^1(\Omega)}^2 + \|\boldsymbol{w}_h^n\|_{H^1(\Omega)}^2 + \varepsilon \|\boldsymbol{\lambda}_h^n\|^2 \leq C, \tag{17}$$

where C is a positive constant depending on Ω, δ and E^0, but not on h, n or Δt.

Proof Since $F^{n+1}(u_h^{n+1}, \boldsymbol{w}_h^{n+1}) \leq F^{n+1}(u_h^n, \boldsymbol{w}_h^n)$, we have

$$\begin{aligned} E^{n+1} &+ \int\limits_{\Omega} \left(-G_-(\boldsymbol{w}_h^{n+1}) + \frac{1}{2\Delta t} |u_h^{n+1}|^2 + \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n) \cdot \boldsymbol{w}_h^{n+1} - \frac{1}{\Delta t} u_h^n u_h^{n+1} \right) dx \\ &\leq E^n + \int\limits_{\Omega} \left(-G_-(\boldsymbol{w}_h^n) + \frac{1}{2\Delta t} |u_h^n|^2 + \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n) \cdot \boldsymbol{w}_h^n - \frac{1}{\Delta t} |u_h^n|^2 \right) dx. \end{aligned}$$

Then, Inequality (16) follows from

$$\begin{aligned}
&\int\limits_{\Omega} (G_-(\boldsymbol{w}_h^{n+1}) - G_-(\boldsymbol{w}_h^n) - \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n) \cdot (\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n))\, dx \\
&\quad = \int\limits_{\Omega} ((\nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n + s_1(\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n)) - \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n)) \cdot (\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n))\, dx \\
&\quad = \int\limits_{\Omega} s_1(\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n) \cdot \nabla_{\mathbb{F}}^2 G_-(\boldsymbol{w}_h^n + s_2(\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n)) \cdot (\boldsymbol{w}_h^{n+1} - \boldsymbol{w}_h^n)\, dx \\
&\quad \le 0,
\end{aligned}$$

where $0 \le s_2 \le s_1 \le 1$ are constants from the mean-value theorem.

For (17), by using (14) and (15), we only need to prove that $\|u_h^n\|_{H^1(\Omega)}$ is bounded. This is because

$$\|u_h^n\|_{H^1(\Omega)}^2 \le C\|\nabla u_h^n\|^2 = C\|\boldsymbol{P}_h \boldsymbol{w}_h^n - \varepsilon \boldsymbol{\lambda}_h^n\|^2 \le C\|\boldsymbol{w}_h^n\|^2 + C\varepsilon^2\|\boldsymbol{\lambda}_h^n\|^2.$$

As long as $\varepsilon \le O(1)$, Inequality (17) is true. □

4 Convergence

In this section, we analyze the convergence rate of the fully-discrete, mixed finite element approximation (11). Notice that the well-posedness and time-stability results in the previous section is proved for arbitrary convex–concave decomposition $G = G_+ + G_-$. However, for the convergence rate, so far we limit our analysis for the special decomposition defined in (4). Convergence rate analysis for general convex–concave decomposition can be non-trivial.

Let $(u, \boldsymbol{w}, \lambda)$ be the solution to (8) and $(u_h^n, \boldsymbol{w}_h^n, \boldsymbol{\lambda}_h^n)$ be the solution to (11). Denote $u^n = u(\cdot, t_n)$, $\boldsymbol{w}^n = \boldsymbol{w}(\cdot, t_n)$ and $\boldsymbol{\lambda}^n = \boldsymbol{\lambda}(\cdot, t_n)$. Define the error terms

$$\underline{u}^n = u^n - u_h^n, \quad \underline{\boldsymbol{w}}^n = \boldsymbol{w}^n - \boldsymbol{w}_h^n, \quad \underline{\boldsymbol{\lambda}}^n = \boldsymbol{\lambda}^n - \boldsymbol{\lambda}_h^n.$$

By subtracting (11) from (8), we have

$$\begin{cases}
\delta(\nabla \underline{\boldsymbol{w}}^{n+1}, \nabla \boldsymbol{\varphi}_h) + (\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^{n+1}), & \\
\quad \boldsymbol{\varphi}_h) + (\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{\varphi}_h) = -(\nabla_{\mathbb{F}} G_-(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^n), \boldsymbol{\varphi}_h) & \text{for all } \boldsymbol{\varphi}_h \in S_h^2, \\
(\partial_t u^{n+1} - \frac{u_h^{n+1} - u_h^n}{\Delta t}, v_h) - (\underline{\boldsymbol{\lambda}}^{n+1}, \nabla v_h) = 0 & \text{for all } v_h \in S_h, \\
-\varepsilon(\boldsymbol{\lambda}_h^{n+1}, \boldsymbol{\mu}_h) - (\underline{\boldsymbol{w}}^{n+1} - \nabla \underline{u}^{n+1}, \boldsymbol{\mu}_h) = 0, & \text{for all } \boldsymbol{\mu}_h \in \boldsymbol{Q}_h.
\end{cases} \tag{18}$$

We first introduce several technical lemmas. For simplicity, use C to denote a general positive constant that depends only on $\delta, C_G, m, \Omega, T$, $\|u\|_{L^\infty(0,T;H^3(\Omega))}$, $\|\boldsymbol{w}\|_{L^\infty(0,T;H^2(\Omega))}$, $\|\boldsymbol{\lambda}\|_{L^\infty(0,T;H^1(\Omega))}$ and E_0.

Lemma 4.1 $\|\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^n) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^n)\| \leq C\|\nabla \underline{\boldsymbol{w}}^n\|$.

Proof In two-dimension, $H^1(\Omega)$ is continuously embedded in the Hölder space $L^q(\Omega)$ for all $1 \leq q < \infty$. By Assumption (3), the Sobolev embedding theorem, Poincaré inequality, and Inequality (17),

$$\begin{aligned}
&\|\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^n) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^n)\| \\
&\quad \leq \left\| \max_{0\leq s\leq 1} |\nabla_{\mathbb{F}}^2 G_+(s\boldsymbol{w}^n + (1-s)\boldsymbol{w}_h^n)|\, |\underline{\boldsymbol{w}}^n| \right\| \\
&\quad \leq C \left\| (1 + |\boldsymbol{w}^n|^m + |\boldsymbol{w}_h^n|^m)|\underline{\boldsymbol{w}}^n| \right\| \\
&\quad \leq C(1 + \|\boldsymbol{w}^n\|_{L^{4m}(\Omega)}^m + \|\boldsymbol{w}_h^n\|_{L^{4m}(\Omega)}^m)\|\underline{\boldsymbol{w}}^n\|_{L^4(\Omega)} \\
&\quad \leq C(1 + \|\boldsymbol{w}^n\|_{H^1(\Omega)}^m + \|\boldsymbol{w}_h^n\|_{H^1(\Omega)}^m)\|\nabla \underline{\boldsymbol{w}}^n\| \\
&\quad \leq C(1 + \|\boldsymbol{w}\|_{L^\infty(0,T;H^1(\Omega))}^m)\|\nabla \underline{\boldsymbol{w}}^n\|.
\end{aligned}$$

Here $\boldsymbol{w}$ is the solution to the mixed problem (8). By the regularity assumption (2), $\|\boldsymbol{w}\|_{L^\infty(0,T;H^1(\Omega))}$ is bounded as long as $u_0 \in H^3(\Omega)$. This completes the proof of the lemma. □

Lemma 4.2 $\|\nabla \underline{u}^n\| \leq \|\underline{\boldsymbol{w}}^n\| + C(h + \sqrt{\varepsilon})$.

Proof Note that by definition, $\nabla u^n = \boldsymbol{w}^n$. By Eq. (11) and the fact that $\boldsymbol{\lambda}_h^n \in \boldsymbol{Q}_h$, $\nabla u_h^n \in \boldsymbol{Q}_h$, we have

$$\boldsymbol{0} = \boldsymbol{P}_h(\varepsilon\boldsymbol{\lambda}_h^n - \boldsymbol{w}_h^n + \nabla u_h^n) = \varepsilon\boldsymbol{\lambda}_h^n - \boldsymbol{P}_h\boldsymbol{w}_h^n + \nabla u_h^n.$$

Combining the above and using the triangle inequality, the property of the L^2 orthogonal projection $\boldsymbol{P}_h$, and Inequality (17),

$$\begin{aligned}
\|\nabla \underline{u}^n\| &= \|\nabla u^n - \nabla u_h^n\| = \|\boldsymbol{w}^n - (\boldsymbol{P}_h\boldsymbol{w}_h^n - \varepsilon\boldsymbol{\lambda}_h^n)\| \\
&\leq \|\underline{\boldsymbol{w}}^n\| + \|(I - \boldsymbol{P}_h)\boldsymbol{w}_h^n\| + \varepsilon\|\boldsymbol{\lambda}_h^n\| \leq \|\underline{\boldsymbol{w}}^n\| + C(h + \sqrt{\varepsilon}).
\end{aligned}$$

□

Lemma 4.3 *Let G_- be defined as in (4), then*

$$\begin{aligned}
&\frac{1}{4}(3\|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 - \|\nabla \underline{\boldsymbol{w}}^n\|^2) + \frac{\varepsilon}{2\delta}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2 \\
&\quad \leq C\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right) + C\Delta t \int_{t_n}^{t_{n+1}} \|\boldsymbol{w}_t(\cdot, s)\|^2\, ds + C\|\underline{u}^{n+1}\|^2 \\
&\qquad - \frac{2}{\delta}(\nabla \underline{u}^{n+1}, \boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}).
\end{aligned}$$

Proof For all $\boldsymbol{\psi}_h \in S_h^2$,

$$\begin{aligned}
&\|\nabla(\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h)\|^2 \\
&\quad = \|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + \|\nabla(\boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h)\|^2 + 2(\nabla \underline{\boldsymbol{w}}^{n+1}, \nabla(\boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h)).
\end{aligned}$$

By setting $\boldsymbol{\psi}_h$ to be the nodal value interpolation of $\boldsymbol{w}^{n+1}$ and the test function $\boldsymbol{\varphi}_h = \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h$ in (18), we have

$$\begin{aligned}
&\|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 \\
&\quad \le \|\nabla(\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h)\|^2 - 2(\nabla \underline{\boldsymbol{w}}^{n+1}, \nabla(\boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h)) \\
&\quad \le Ch^2 + \frac{2}{\delta}(\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^{n+1}), \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) \\
&\qquad + \frac{2}{\delta}(\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) + \frac{2}{\delta}(\nabla_{\mathbb{F}} G_-(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^{n}), \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) \\
&\quad \le Ch^2 + \frac{2}{\delta}(\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^{n+1}), \boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h) \\
&\qquad + \frac{2}{\delta}(\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) + \frac{2}{\delta}(\nabla_{\mathbb{F}} G_-(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_-(\boldsymbol{w}_h^{n}), \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) \\
&\quad = Ch^2 + I_1 + I_2 + I_3, \qquad (19)
\end{aligned}$$

where the second last step follows from the fact that G_+ is convex.

For I_1, by Lemma 4.1, we have

$$\begin{aligned}
I_1 &\le C\|\nabla_{\mathbb{F}} G_+(\boldsymbol{w}^{n+1}) - \nabla_{\mathbb{F}} G_+(\boldsymbol{w}_h^{n+1})\| \, \|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\| \\
&\le C\|\nabla \underline{\boldsymbol{w}}^{n+1}\| \, \|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\| \\
&\le Ch^2. \qquad (20)
\end{aligned}$$

Now we consider I_2. By the triangle inequality, Schwarz inequality, Poincaré inequality, and the Young's inequality, we have

$$\begin{aligned}
&(\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) = (\boldsymbol{\lambda}^{n+1} - \boldsymbol{P}_h\boldsymbol{\lambda}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) \\
&\qquad + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}) + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h) \\
&\quad \le \|(I - \boldsymbol{P}_h)\boldsymbol{\lambda}^{n+1}\|(\|\boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}\| + \|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\|) \\
&\qquad + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}) + \frac{\varepsilon}{4}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2 + \frac{1}{\varepsilon}\|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\|^2 \\
&\quad \le \frac{\delta}{16}\|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + \frac{1}{2}\|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\|^2 + C\|(I - \boldsymbol{P}_h)\boldsymbol{\lambda}^{n+1}\|^2 \\
&\qquad + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}) + \frac{\varepsilon}{4}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2 + \frac{1}{\varepsilon}\|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\|^2 \\
&\quad \le \frac{\delta}{16}\|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + C\left(h^2 + \frac{h^4}{\varepsilon}\right) + \frac{\varepsilon}{4}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2 \\
&\qquad + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}).
\end{aligned}$$

By setting $\boldsymbol{\mu}_h = \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}$ in (18), we have

$$\begin{aligned}
&(\boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{w}^{n+1}) \\
&\quad = \varepsilon(\boldsymbol{\lambda}_h^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}) - (\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}) \\
&\quad = -\varepsilon \|\boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}\|^2 + \varepsilon(\boldsymbol{P}_h \boldsymbol{\lambda}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}) - (\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}) \\
&\quad \le -\frac{\varepsilon}{2} \|\boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}\|^2 + \frac{\varepsilon}{2} \|\boldsymbol{P}_h \boldsymbol{\lambda}^{n+1}\|^2 - (\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}) \\
&\quad \le -\frac{\varepsilon}{2} \|\boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}\|^2 + C\varepsilon - (\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}).
\end{aligned}$$

Combining the above, we have

$$I_2 \le \frac{1}{8} \|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + C\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right) - \frac{\varepsilon}{2\delta} \|\boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}\|^2 - \frac{2}{\delta} (\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{\lambda}}^{n+1}). \tag{21}$$

Finally, we consider I_3. The analysis of I_3 depends on the definition of G_-. It is not trivial to get an upper bound of I_3 when G_- satisfying only (3), without making further assumptions. However, if G_- is defined as in (4), then $\nabla_{\mathbb{F}} G_-(\boldsymbol{w}) = -\boldsymbol{w}$ and the analysis is given as following:

$$\begin{aligned}
I_3 &= C(\boldsymbol{w}_h^n - \boldsymbol{w}^{n+1}, \boldsymbol{w}_h^{n+1} - \boldsymbol{\psi}_h) \\
&= C((\boldsymbol{w}^n - \boldsymbol{w}^{n+1}, \boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h) - (\boldsymbol{w}^n - \boldsymbol{w}^{n+1}, \underline{\boldsymbol{w}}^{n+1}) \\
&\qquad + (\underline{\boldsymbol{w}}^n, \underline{\boldsymbol{w}}^{n+1}) - (\underline{\boldsymbol{w}}^n, \boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h)) \\
&\le Ch^2 + C(\boldsymbol{w}^{n+1} - \boldsymbol{w}^n, \underline{\boldsymbol{w}}^{n+1}) + C\|\underline{\boldsymbol{w}}^n\| \, \|\boldsymbol{w}^{n+1} - \boldsymbol{\psi}_h\| \\
&\quad + C((I - \boldsymbol{P}_h)\underline{\boldsymbol{w}}^n, \underline{\boldsymbol{w}}^{n+1}) + C(\boldsymbol{P}_h \underline{\boldsymbol{w}}^n, \underline{\boldsymbol{w}}^{n+1}) \\
&\le Ch^2 + C(\boldsymbol{w}^{n+1} - \boldsymbol{w}^n, \underline{\boldsymbol{w}}^{n+1}) + \frac{1}{16} \|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + C(\boldsymbol{P}_h \underline{\boldsymbol{w}}^n, \underline{\boldsymbol{w}}^{n+1}).
\end{aligned}$$

Notice that by (18), the triangle inequality, Schwarz inequality, Poincaré inequality, Young's inequality, Theorem 3.2 and Lemma 4.2,

$$\begin{aligned}
&C(\boldsymbol{P}_h \underline{\boldsymbol{w}}^n, \underline{\boldsymbol{w}}^{n+1}) \\
&\quad = C((\nabla \underline{u}^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{w}}^n) - \varepsilon(\boldsymbol{\lambda}_h^{n+1}, \boldsymbol{P}_h \underline{\boldsymbol{w}}^n)) \\
&\quad = C((\nabla \underline{u}^{n+1}, \underline{\boldsymbol{w}}^n) - (\nabla \underline{u}^{n+1}, (I - \boldsymbol{P}_h)\underline{\boldsymbol{w}}^n) - \varepsilon(\boldsymbol{\lambda}_h^{n+1}, \underline{\boldsymbol{w}}^n)) \\
&\quad = C(-(\underline{u}^{n+1}, \nabla \cdot \underline{\boldsymbol{w}}^n) - (\nabla \underline{u}^{n+1}, (I - \boldsymbol{P}_h)\underline{\boldsymbol{w}}^n) - \varepsilon(\boldsymbol{\lambda}_h^{n+1}, \underline{\boldsymbol{w}}^n)) \\
&\quad \le \frac{1}{4} \|\nabla \underline{\boldsymbol{w}}^n\|^2 + C\|\underline{u}^{n+1}\|^2 + C\varepsilon^2 \|\boldsymbol{\lambda}_h^{n+1}\|^2 \\
&\qquad + C(\|\underline{\boldsymbol{w}}^{n+1}\| + C(h + \sqrt{\varepsilon}))\|(I - \boldsymbol{P}_h)\underline{\boldsymbol{w}}^n\| \\
&\quad \le \frac{1}{4} \|\nabla \underline{\boldsymbol{w}}^n\|^2 + C\|\underline{u}^{n+1}\|^2 + \frac{1}{32} \|\nabla \underline{\boldsymbol{w}}^{n+1}\|^2 + C(h^2 + \varepsilon).
\end{aligned}$$

Also,

$$\begin{aligned}C(\boldsymbol{w}^{n+1}-\boldsymbol{w}^{n},\underline{\boldsymbol{w}}^{n+1}) &= C\left(\int_{t_n}^{t_{n+1}}\boldsymbol{w}_t(\cdot,s)\,ds,\ \underline{\boldsymbol{w}}^{n+1}\right)\\ &\le C\int_{t_n}^{t_{n+1}}\|\boldsymbol{w}_t(\cdot,s)\|\,\|\underline{\boldsymbol{w}}^{n+1}\|\,ds\\ &\le \frac{1}{32}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 + C\Delta t\int_{t_n}^{t_{n+1}}\|\boldsymbol{w}_t(\cdot,s)\|^2\,ds.\end{aligned}$$

Combining the above,

$$\begin{aligned}I_3 \le C\Delta t\int_{t_n}^{t_{n+1}}\|\boldsymbol{w}_t(\cdot,s)\|^2\,ds + \frac{1}{4}\|\nabla\underline{\boldsymbol{w}}^{n}\|^2 + \frac{1}{8}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2\\ +C\|\underline{u}^{n+1}\|^2 + C(h^2+\varepsilon). \qquad (22)\end{aligned}$$

Combining (19), (20), (21), (22), we have proved the lemma. □

Finally, we are able to prove the main result of this section. The following discrete Gronwall's inequality will be needed [34]: let y^n, a^n, b^n, c^n , be non-negative sequences satisfying

$$y^n + \Delta t\sum_{i=1}^{n} a^i \le y^0 + \Delta t\sum_{i=1}^{n}(b^i y^i + c^i)$$

with $\Delta t b^i < 1$, then

$$y^n + \Delta t\sum_{i=1}^{n} a^i \le e^{C_b\Delta t\sum_{i=1}^{n} b^i}\left(\Delta t\sum_{i=1}^{n} c^i + y^0\right),$$

where $C_b = \max_{0\le i\le n}(1-\Delta t b^i)^{-1}$.

Theorem 4.4 *Let G_- be defined as in (4). Then there exists a constant C independent of h or ε such that for $O(h^2)\le\Delta t\le C$,*

$$\begin{aligned}\|\underline{u}^n\|^2 + \sum_{i=1}^{n}\Delta t\|\nabla\underline{\boldsymbol{w}}^i\|^2 + \sum_{i=1}^{n}\Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^i\|^2\\ \le Ce^{Ct_n}\Bigg(\|\underline{u}^0\|^2 + \Delta t\|\nabla\underline{\boldsymbol{w}}^0\|^2 + \Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^0\|^2 + t_n\left(h^2+\frac{h^4}{\varepsilon}+\varepsilon\right)\\ +\Delta t^2\int_0^{t_n}(\|u_{tt}(\cdot,s)\|^2 + \|\boldsymbol{w}_t(\cdot,s)\|^2)\,ds\Bigg). \qquad (23)\end{aligned}$$

Here again, all general constant C may depend on δ, C_G, m, Ω, $\|u\|_{L^\infty(0,T;H^3(\Omega))}$, $\|\boldsymbol{w}\|_{L^\infty(0,T;H^2(\Omega))}$, $\|\boldsymbol{\lambda}\|_{L^\infty(0,T;H^1(\Omega))}$ and E_0, but not on h, Δt or ε.

Proof By setting $v_h = I_h\underline{u}^{n+1}$ in (18), we have

$$\begin{aligned}\frac{1}{\Delta t}(\underline{u}^{n+1} - \underline{u}^n, I_h\underline{u}^{n+1}) &= (\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1})) - (\xi^{n+1}, I_h\underline{u}^{n+1})\\ &= (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1})) - (\xi^{n+1}, I_h\underline{u}^{n+1}),\end{aligned} \tag{24}$$

where the local truncation error ξ^{n+1} is

$$\xi^{n+1} = \partial_t u^{n+1} - \frac{u^{n+1} - u^n}{\Delta t} = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} (s - t_n)u_{tt}(\cdot, s)\,ds.$$

By the Schwarz inequality, Lemma 4.2, Poincaré inequality and Young's inequality, it is not hard to see that

$$\begin{aligned}&(\xi^{n+1}, I_h\underline{u}^{n+1})\\ &\le \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \|(s - t_n)u_{tt}(\cdot, s)\|\,\|I_h\underline{u}^{n+1}\|\,ds\\ &\le \frac{C}{\Delta t}\int_{t_n}^{t_{n+1}} \|(s - t_n)u_{tt}(\cdot, s)\|(\|\nabla\underline{u}^{n+1}\| + \|\underline{u}^{n+1} - I_h\underline{u}^{n+1}\|)\,ds\\ &\le \frac{C}{\Delta t}\int_{t_n}^{t_{n+1}} \|(s - t_n)u_{tt}(\cdot, s)\|(\|\nabla\underline{\boldsymbol{w}}^{n+1}\| + (h + \sqrt{\varepsilon}))\,ds\\ &\le \frac{\delta}{8}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 + C(h^2 + \varepsilon) + \frac{C}{\Delta t}\int_{t_n}^{t_{n+1}} \|(s - t_n)u_{tt}(\cdot, s)\|^2\,ds\\ &\le \frac{\delta}{8}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 + C(h^2 + \varepsilon) + C\Delta t\int_{t_n}^{t_{n+1}} \|u_{tt}(\cdot, s)\|^2\,ds.\end{aligned}$$

Then by using the Schwarz inequality and the Young's inequality,

$$\begin{aligned}&\frac{1}{2\Delta t}(\|\underline{u}^{n+1}\|^2 - \|\underline{u}^n\|^2 + \|\underline{u}^{n+1} - \underline{u}^n\|^2)\\ &\quad= \frac{1}{\Delta t}(\underline{u}^{n+1} - \underline{u}^n, (I - I_h)\underline{u}^{n+1}) + \frac{1}{\Delta t}(\underline{u}^{n+1} - \underline{u}^n, I_h\underline{u}^{n+1})\\ &\quad= \frac{1}{\Delta t}(\underline{u}^{n+1} - \underline{u}^n, (I - I_h)u^{n+1}) + \frac{1}{\Delta t}(\underline{u}^{n+1} - \underline{u}^n, I_h\underline{u}^{n+1})\end{aligned}$$

$$\le \frac{Ch^2}{\Delta t^2}(\|\underline{u}^n\|^2 + \|\underline{u}^{n+1}\|^2) + Ch^2 + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1}))$$

$$+\frac{\delta}{8}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 + C(h^2+\varepsilon) + C\Delta t\int_{t_n}^{t_{n+1}}\|u_{tt}(\cdot,s)\|^2\,ds. \tag{25}$$

Combine the above with Lemma 4.3, we have

$$\frac{1}{2\Delta t}(\|\underline{u}^{n+1}\|^2 - \|\underline{u}^n\|^2 + \|\underline{u}^{n+1} - \underline{u}^n\|^2)$$
$$+\frac{\delta}{8}(2\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 - \|\nabla\underline{\boldsymbol{w}}^n\|^2) + \frac{\varepsilon}{4}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2$$
$$\le \frac{Ch^2}{\Delta t^2}(\|\underline{u}^n\|^2 + \|\underline{u}^{n+1}\|^2) + C\Delta t\int_{t_n}^{t_{n+1}}\|u_{tt}(\cdot,s)\|^2\,ds$$
$$+C\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right) + C\Delta t\int_{t_n}^{t_{n+1}}\|\boldsymbol{w}_t(\cdot,s)\|^2\,ds + C\|\underline{u}^{n+1}\|^2$$
$$+(\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1}) - \nabla\underline{u}^{n+1})$$

Since by Inequality (10) and the Young's inequality

$$\begin{aligned}(\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1}) - \nabla\underline{u}^{n+1}) &= (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h u^{n+1} - u^{n+1}))\\ &\le Ch^2\|u^{n+1}\|_{H^3}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|,\\ &\le C\frac{h^4}{\varepsilon} + \frac{\varepsilon}{8}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2,\end{aligned}$$

then we can conclude that

$$\frac{1}{2\Delta t}(\|\underline{u}^{n+1}\|^2 - \|\underline{u}^n\|^2) + \frac{\delta}{8}(2\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 - \|\nabla\underline{\boldsymbol{w}}^n\|^2) + \frac{\varepsilon}{8}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}\|^2$$
$$\le C\left(1 + \frac{h^2}{\Delta t^2}\right)(\|\underline{u}^n\|^2 + \|\underline{u}^{n+1}\|^2) + C\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right)$$
$$+C\Delta t\int_{t_n}^{t_{n+1}}(\|u_{tt}(\cdot,s)\|^2 + \|\boldsymbol{w}_t(\cdot,s)\|^2)\,ds.$$

Finally, by taking summation of the above inequality with respect to n, and setting

$$y^0 = \frac{1}{2}\|\underline{u}^0\|^2 + \frac{\delta\Delta t}{4}\|\nabla\underline{\boldsymbol{w}}^0\|^2 + \frac{\varepsilon\Delta t}{8}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^0\|^2,$$

and for $n \geq 1$,

$$
\begin{aligned}
y^n &= \frac{1}{2}\|\underline{u}^n\|^2,\\
a^n &= \frac{\varepsilon}{8}\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^n\|^2 + \frac{\delta}{8}\|\nabla\underline{\boldsymbol{w}}^n\|^2,\\
b^n &= C\left(1 + \frac{h^2}{\Delta t^2}\right),\\
c^n &= C\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right) + C\Delta t\int_{t_n}^{t_{n+1}} (\|u_{tt}(\cdot,s)\|^2 + \|\boldsymbol{w}_t(\cdot,s)\|^2)\,ds,
\end{aligned}
$$

and using the Gronwall's inequality, we get Inequality (23). Notice that in order to guarantee $\Delta t b^n < 1$, we need $O(h^2) < \Delta t < C$. However, the upper bound of Δt does not depend on h. □

The constraint $O(h^2) \leq \Delta t$ is unusual, since most numerical schemes performs better when fixing h and decreasing Δt. Indeed, if we assume further regularity of the solution, then this condition can be dropped.

Theorem 4.5 *Let G_- be defined as in (4) and assume $u_t \in L^\infty(0,T;H^{1+s}(\Omega))$, where $s > 0$. Then there exists a constant C independent of h or ε such that for all $\Delta t \leq C$,*

$$
\begin{aligned}
&\|\underline{u}^n\|^2 + \sum_{i=1}^{n}\Delta t\|\nabla\underline{\boldsymbol{w}}^i\|^2 + \sum_{i=1}^{n}\Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^i\|^2\\
&\leq Ce^{Ct_n}\left(\|\underline{u}^0\|^2 + \Delta t\|\nabla\underline{\boldsymbol{w}}^0\|^2 + \Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^0\|^2 + t_n\left(h^2 + \frac{h^4}{\varepsilon} + \varepsilon\right)\right.\\
&\qquad\left. + \Delta t^2\int_0^{t_n}(\|I_h u_{tt}(\cdot,s)\|^2 + \|\boldsymbol{w}_t(\cdot,s)\|^2)\,ds\right).
\end{aligned}
$$

The general constant in the above inequality may depend on all parameters mentioned as in Theorem 4.4 plus $\|u_t\|_{L^\infty(0,T;H^{1+s}(\Omega))}$, but not on h, Δt, or ε.

Proof Similar to Eq. (24) in the beginning of the proof for Theorem 4.4, we have

$$
\begin{aligned}
&\frac{1}{\Delta t}(I_h\underline{u}^{n+1} - I_h\underline{u}^n, I_h\underline{u}^{n+1})\\
&= \frac{1}{\Delta t}(I_h(u^{n+1} - u^n), I_h\underline{u}^{n+1}) - \frac{1}{\Delta t}(u_h^{n+1} - u_h^n, \nabla(I_h\underline{u}^{n+1}))\\
&= (I_h(\partial_t u^{n+1} - \xi^{n+1}), I_h\underline{u}^{n+1}) - (\boldsymbol{\lambda}_h^{n+1}, \nabla(I_h\underline{u}^{n+1}))\\
&= (\boldsymbol{\lambda}^{n+1}, \nabla(I_h\underline{u}^{n+1})) - ((I - I_h)\partial_t u^{n+1}, I_h\underline{u}^{n+1}) - (I_h\xi^{n+1}, I_h\underline{u}^{n+1})\\
&\quad -(\boldsymbol{\lambda}_h^{n+1}, \nabla(I_h\underline{u}^{n+1}))
\end{aligned}
$$

$$= (\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1})) - (I_h\xi^{n+1}, I_h\underline{u}^{n+1}) - ((I-I_h)\partial_t u^{n+1}, I_h\underline{u}^{n+1})$$
$$= (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1})) - (I_h\xi^{n+1}, I_h\underline{u}^{n+1}) - ((I-I_h)\partial_t u^{n+1}, I_h\underline{u}^{n+1}).$$

Here $(I_h\xi^{n+1}, I_h\underline{u}^{n+1})$ can be bounded similarly as the term $(\xi^{n+1}, I_h\underline{u}^{n+1})$ in the proof of Theorem 4.4, with $\|u_{tt}(\cdot,s)\|$ being replaced by $\|I_h u_{tt}(\cdot,s)\|$. We also have an extra term, which can be bounded by

$$\begin{aligned}((I-I_h)\partial_t u^{n+1}, I_h\underline{u}^{n+1}) &\le Ch\|u_t\|_{L^\infty(0,T;H^{1+s}(\Omega))}\|I_h\underline{u}^{n+1}\| \\ &\le Ch^2 + \|I_h\underline{u}^{n+1}\|^2.\end{aligned}$$

Then, Eq. (25) in the proof of Theorem 4.4 can then be rewritten as

$$\begin{aligned}&\frac{1}{2\Delta}(\|I_h\underline{u}^{n+1}\|^2 - \|I_h\underline{u}^n\|^2 + \|I_h\underline{u}^{n+1} - I_h\underline{u}^n\|^2) \\ &\quad= \frac{1}{\Delta t}(I_h\underline{u}^{n+1} - I_h\underline{u}^n, I_h\underline{u}^{n+1}) \\ &\quad\le Ch^2 + \|I_h\underline{u}^{n+1}\|^2 + (\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^{n+1}, \nabla(I_h\underline{u}^{n+1})) \\ &\qquad+\frac{\delta}{8}\|\nabla\underline{\boldsymbol{w}}^{n+1}\|^2 + C(h^2+\varepsilon) + C\Delta t\int_{t_n}^{t_{n+1}}\|I_h u_{tt}(\cdot,s)\|^2\,ds.\end{aligned}$$

The rest of the proof is the same as the proof of Theorem 4.4, with the only differences that $\|\underline{u}^{n+1}\|^2$ and $\|\underline{u}^n\|^2$ are now substituted by $\|I_h\underline{u}^{n+1}\|^2$ and $\|I_h\underline{u}^n\|^2$ and we no longer have the term $1+\frac{h^2}{\Delta t^2}$. The Gronwall's inequality will now be applied with $y^n = \frac{1}{2}\|I_h\underline{u}^n\|^2$ and $b^n = C$. Since

$$\begin{aligned}\|\underline{u}^n\| &\le \|(I-I_h)\underline{u}^n\| + \|I_h\underline{u}^n\|^2 \le Ch^2 + \|I_h\underline{u}^n\|, \\ \|I_h\underline{u}^n\| &\le \|(I-I_h)\underline{u}^n\| + \|\underline{u}^n\|^2 \le Ch^2 + \|\underline{u}^n\|,\end{aligned}$$

we will be able to get the error estimation in this theorem. □

Remark 4.6 If

$$\|\underline{u}^0\|^2 + \Delta t\|\nabla\underline{\boldsymbol{w}}^0\|^2 + \Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^0\|^2 \le Ch^2, \quad \varepsilon = Ch^2,$$

and

$$\int_0^{t_n}(\|I_h u_{tt}(\cdot,s)\|^2 + \|\boldsymbol{w}_t(\cdot,s)\|^2)\,ds \le C,$$

then we have

$$\|\underline{u}^n\|^2 + \sum_{i=1}^n \Delta t\|\nabla\underline{\boldsymbol{w}}^i\|^2 + \sum_{i=1}^n \Delta t\varepsilon\|\boldsymbol{P}_h\underline{\boldsymbol{\lambda}}^i\|^2 \le C(h^2+\Delta t^2).$$

Furthermore, by Lemma 4.2 and the Poincaré inequality, it is easy to see that

$$\sum_{i=1}^{n} \Delta t \|\nabla \underline{u}^i\|^2 \le C(h^2 + \Delta t^2).$$

Remark 4.7 One may be able to slightly improve the result in Theorem 4.5, by defining I_h to be a Clément-type interpolation preserving homogeneous or periodic boundary conditions. Such an interpolation has been constructed in [40]. However, the advantage of doing so is not very obvious. For smooth solutions, Theorem 4.5 already gives the optimal convergence rate. For the future research, a more interesting direction would be to explore the role of parameters δ and ε.

Acknowledgments Chen was supported by the 111 project, Key Project National Science Foundation of China (91130004) and the Natural Science Foundation of China (11171077). He also thanks Jianguo Liu in Duke University and Xiaoming Wang in Florida State University for the fruitful discussions. Wang thanks the Key Laboratory of Mathematics for Nonlinear Sciences (EYH1140070), Fudan University, for the support during her visit. The authors are also grateful to the anonymous referees for their helpful comments and suggestions which greatly improved the quality of this paper.

References

1. Arnold, D.N., Falk, R.S.: A uniformly accurate finite element method for the Reissner–Mindlin plate. SIAM J. Numer. Anal. **26**, 1276–1290 (1989)
2. Bathe, K.J., Dvorkin, E.N.: A four-node plate bending element based on Mindlin–Reissner plate theory and a mixed interpolation. J. Numer. Methods Eng. **21**, 367–383 (1985)
3. Bathe, K.J., Brezzi, F.: On the convergence of a four-node plate bending element based on Mindlin–Reissner plate theory and a mixed interpolation. In: Whiteman, J.R. (ed.) MAFELAP V, pp. 491–503. Academic Press, London (1985)
4. Bathe, K.J., Brezzi, F.: A simplified analysis of two plate-bending elements-the MITC4 and MITC9 elements. In: Pande, G.N., Middleton, J. (eds.) MUNETA 87. Numerical Techniques for Engineering Analysis and Design, vol. 1 (1987)
5. Berkovitz, L.D.: Convexity and optimization in $\mathbb{R}^n$. Wiley, New York (2002)
6. Blomker, D., Gugg, C.: On the existence of solutions for amorphous molecular beam epitaxy. Nonlinear Anal. Real World Appl. **3**, 61–73 (2002)
7. Brezzi, F., Fortin, M.: Numerical approximation of Mindlin-Reissner plates. Math. Comp. **47**, 151–158 (1986)
8. Brezzi, F., Bathe, K.J., Fortin, M.: Mixed interpolated elements for Reissner–Mindlin plates. J. Numer. Methods Eng. **28**, 1787–1801 (1989)
9. Caflisch, R.E., Gyure, M.F., Merriman, B., Osher, S., Ratsch, C., Vvedensky, D.D.: Island dynamics and the level set method for epitaxial growth. Appl. Math. Lett. **12**, 13–22 (1999)
10. Chen, W., Conde, S., Wang, C., Wang, X., Wise, S.M.: A linear energy stable scheme for a thin film model without slope selection. J. Sci. Comput. **26**, 1–17 (2011)
11. Cho, A.: Film deposition by molecular beam techniques. J. Vac. Sci. Technol. **8**, S31–S38 (1971)
12. Cho, A., Arthur, J.: Molecular beam epitaxy. Prog. Solid State Chem. **10**, 157–192 (1975)
13. Clarke, S., Vvedensky, D.D.: Origin of reflection high-energy electron-diffraction intensity oscillations during molecular-beam epitaxy: a computational modeling approach. Phys. Rev. Lett. **58**, 2235–2238 (1987)
14. Copetti, M.I.M., Elliot, C.M.: Numerical Analysis of the Cahn–Hilliard equation with a logarithmic free energy. Numer. Math. **63**, 39–65 (1992)
15. Du, Q., Nicolaides, R.A.: Numerical analysis of a continuum model of phase transition. SIAM J. Numer. Anal. **28**, 1310–1322 (1991)
16. Duran, R., Liberman, E.: On mixed finite element methods for the Reissner–Mindlin plate model. Math. Comp. **58**, 561–573 (1992)

17. Elliot, C.M., French, D.A.: Numerical studies of the Cahn–Hilliard equation for phase separation. IMA J. Appl. Math. **38**, 97–128 (1987)
18. Elliot, C.M., French, D.A.: A nonconforming finite-element method for the two-dimensional Cahn–Hilliard equation. SIAM J. Numer. Anal. **26**, 884–903 (1989)
19. Elliot, C.M., French, D.A.: A second order splitting method for the Cahn–Hilliard equation. Numer. Math. **54**, 575–590 (1989)
20. Eyre, D.J.: Unconditionally gradient stable time marching the Cahn–Hilliard equation. In: Bullard, J.W., Kalia, R., Stoneham, M., , Chen, L.Q. (eds.) Computational and Mathematical Models of Microstructural Evolution, p. 1712. Materials Research Society, Warrendale (1998)
21. Feng, X., Prohl, A.: Error analysis of a mixed finite element method for the Cahn–Hilliard equation. Numer. Math. **99**, 47–84 (2004)
22. Gyure, M.F., Ratsch, C., Merriman, B., Caflisch, R.E., Osher, S.: Level-set methods for the simulation of epitaxial phenomena. Phys. Rev. E **58**, R6927–R6930 (1998)
23. Han, W., Cheng, X., Huang, H.: Some mixed finite element methods for biharmonic equation. J. Comp. Appl. Math. **126**, 91–109 (1999)
24. Hoppe, R.H., Nash, E.M.: A combined spectral element/finite element approach to the numerical solution of a nonlinear evolution equation describing amorphous surface growth of thin films. J. Numer. Math. **10**, 127–136 (2002)
25. Johnson, C., Pitkäranta, J.: Analysis of some mixed finite element methods related to reduced integration. Math. Comp. **38**, 375–400 (1982)
26. Kang, H.C., Weinberg, W.H.: Dynamic Monte Carlo with a proper energy barrier: surface diffusion and two-dimensional domain ordering. J. Chem. Phys. **90**, 2824–2830 (1989)
27. King, B.B., Stein, O., Winkler, M.: A fourth-order parabolic equation modeling epitaxial thin film growth. J. Math. Anal. Appl. **286**, 459–490 (2003)
28. Kohn, R.V., Yan, X.: Upper bounds on the coarsening rate for an epitaxial growth model. Commun. Pure Appl. Math. **56**, 1549–1564 (2003)
29. Krug, J.: Origins of scale invariance in growth processes. Adv. Phys. **46**, 139–282 (1997)
30. Li, B.: High-order surface relaxation versus the Ehrlich–Schwoebel effect. Nonlinearity **19**, 2581–2603 (2006)
31. Li, B.: Variational properties of unbounded order parameters. SIAM J. Math. Anal. **38**, 16–36 (2006)
32. Li, B., Liu, J.: Thin film epitaxy with or without slope selection. Eur. J. Appl. Math. **14**, 713–743 (2003)
33. Li, B., Liu, J.: Epitaxial growth without slope selection: energetics, coarsening, and dynamic scaling. J. Nonlinear Sci. **14**, 429–451 (2004)
34. Lu, X., Lin, P., Liu, J.: Analysis of a sequential regularization method for the unsteady Navier–Stokes equations. Math. Comp. **77**, 1467–1494 (2008)
35. Malkus, D.S., Hughes, T.J.R.: Mixed finite element methods-reduced and selective integration techniques: a unification of concepts. Comput. Methods Appl. Mech. Eng. **15**, 63–81 (1978)
36. Ortiz, M., Repetto, E., Si, H.: A continuum model of kinetic roughening and coarsening in thin films. J. Mech. Phys. Solids **47**, 697–730 (1999)
37. Rost, M.: Continuum models for surface growth. Int. Ser. Numer. Math. **149**, 195–208 (2005)
38. Schneider, M., Schuller, I.K., Rahman, A.: Epitaxial growth of silicon: a molecular-dynamics simulation. Phys. Rev. B **36**, 1340–1343 (1987)
39. Scholtz, R.: A mixed method for fourth-order problems using the linear finite elements. RAIRO Numer. Anal. **15**, 85–90 (1978)
40. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth function satisfying boundary conditions. Math. Comp. **54**, 483–493 (1990)
41. Siegert, M., Plischke, M.: Solid-on-solid models of molecular-beam epitaxy. Phys. Rev. E **50**, 917–931 (1994)
42. Villain, J.: Continuum models of crystal growth from atomistic beams with and without desorption. J. Phys. I **1**, 19–42 (1991)
43. Wang, C., Wang, X., Wise, S.: Unconditionally stable schemes for equations of thin film epitaxy. Discrete Contin. Dyn. Syst. **28**, 405–423 (2010)
44. Xia, X., Chen, W., Liu, J.: Convergence analysis of implicit full discretization for the epitaxial growth model of thin films. Numer. Math. J. Chin. Univ. **34**(1), 30–51 (2012). (in Chinese)
45. Xu, C., Tang, T.: Stability analysis of large time-stepping methods for epitaxial growth models. SIAM J. Numer. Anal. **44**, 1759–1779 (2006)